

Corso di Analisi

seconda parte

introduzione all'analisi matematica su \mathbb{R}^n

versione preliminare, Aprile 2023

Indice

1	Topologia euclidea e continuità	7
1	Struttura euclidea	8
1.1	Prodotto scalare e norma euclidea	8
1.2	Norme; norme equivalenti	11
2	Topologia standard di \mathbb{R}^n	12
3	Successioni e completezza	14
4	Continuità e limiti	15
5	Compattezza	18
6	Connessione	22
7	Spazi metrici e Lemma delle contrazioni	25
	Complementi	28
	Successioni di funzioni	28
	Equicontinuità e Teorema di Ascoli–Arzelà	32
	Integrazione di funzioni vettoriali di variabile reale	33
	Disuguaglianze di Hölder e Minkowski. Norme p su \mathbb{R}^n	34
	Norme ed esponenziale di matrici	36
	Spazi di Banach e spazi di Hilbert	39
	La proprietà di Heine–Borel	42
	Esercizi	43
2	Differenziabilità	53
1	Derivate e differenziali	53
2	Derivate successive	59
2.1	Derivate seconde e matrice hessiana	59
2.2	Punti critici	61
2.3	Funzioni C^k	66
3	Differenziale di funzioni vettoriali	67
3.1	Differenziali e matrici jacobiane	67
3.2	Differenziali di funzioni composte. Regola della catena	69
3.3	Funzioni $C^k(E, \mathbb{R}^m)$	70
	Complementi	71
	Teorema di Eulero sulle funzioni omogenee	71
	Derivazione ‘sotto segno di integrale’	72
	Formula di Taylor in \mathbb{R}^n	73
	Esercizi	79
3	Funzioni inverse e funzioni implicite	85
1	Teorema della Funzione Inversa	85
2	Teorema delle Funzioni Implicite	90
3	Problemi vincolati	92
	Complementi	95
	Teorema quantitativo delle Funzioni Implicite	95

4	Teoria classica dell'integrazione	99
1	Integrale di Riemann e misura di Peano–Jordan	99
1.1	Rettangoli, insieme elementari, funzioni a scalini	99
1.2	Integrale di Riemann	102
1.3	Misura di Peano–Jordan	112
1.4	Teorema del cambio di variabili	112
2	I teoremi classici del calcolo integrale	119
2.1	Curve e superfici	119
2.2	Integrali su curve e superfici	119
2.3	1–forme differenziali. Lavoro e circuitazione	119
2.4	Lemma di Poincarè. Campi conservativi e irrotazionali	121
2.5	Il teorema della divergenza	124
2.6	I teoremi classici di Gauss, Green e Stokes	124
3	Integrale di Lebesgue	126
3.1	Insiemi trascurabili e Teorema di Vitali–Lebesgue	126
3.2	Funzioni \mathcal{L}^1	129
3.3	Teoremi sull'integrazione di limiti	135
3.4	Lo spazio di Banach L^1	137
3.5	Integrali iterati e Teorema di Fubini	139
3.6	Funzioni ed insiemi misurabili. Misura di Lebesgue	141
3.7	Integrazione su insiemi misurabili. Insiemi normali	145
	Esercizi	146
	Complementi	150
	Prodotto vettoriale in \mathbb{R}^3	150
	Rette e piani tangenti	151
	Volumi, aree e curvatura	151
	Dimostrazione del Teorema della divergenza in \mathbb{R}^n	151
	Esercizi	156
A	Serie	159
A.1	Serie di funzioni	159
A.2	Serie di potenze	163
A.3	Serie di Fourier	165
	Complementi	176
	Funzioni reali–analitiche	176
	Trasformata di Fourier	179
	Esercizi	185
B	Equazioni differenziali ordinarie	187
B.1	Introduzione	187
B.1.1	Equazioni scalari	187
B.1.2	Equazioni vettoriali (sistemi di equazioni)	189
B.1.3	Tre esempi risolvibili	190
B.1.4	Osservazioni	193
B.2	Teorema di esistenza e unicità locale	195
B.2.1	Definizioni ed enunciato del Teorema	195
B.2.2	Lemma di Picard	197
B.2.3	Lemmi di Gronwall	198
B.2.4	Unicità e dipendenza continua dai dati iniziali	199
B.3	Soluzioni massimali e soluzioni globali	201
B.3.1	Definizioni	202
B.3.2	Criteri di estensione e di globalità	203
B.3.3	Il caso autonomo: flussi e sistemi dinamici	206
B.3.4	Analisi qualitativa di equazioni scalari del prim'ordine	208
B.4	Sistemi lineari	213

B.4.1	Spazio delle soluzioni	213
B.4.2	Il caso non autonomo	214
	Complementi	214
	Altre classi di equazioni differenziali ordinarie risolubili	214
	Sistemi meccanici unidimensionali. Spazio delle fasi	218
	Dipendenza C^k da dati iniziali e parametri	221
	Stabilità (cenni)	221
	Equazioni delle onde e del calore (alle derivate parziali)	222
	Esercizi	223
C	Varietà immerse in \mathbb{R}^n	225
C.1	Definizioni	225
C.2	Misure e integrali su k -varietà	229
C.2.1	Definizioni e osservazioni	229
C.2.2	Il Teorema della divergenza in \mathbb{R}^n	231
C.3	k -forme differenziali	232
C.4	Derivata esterna di forme differenziali	236
C.5	Forme chiuse. Forme esatte. Lemma di Poincaré	239
C.6	Varietà con bordo	241
C.7	Integrazione di forme differenziali su varietà	243
C.8	Formula di Stokes	244
	Complementi	246
	Retta tangente a una curva in \mathbb{R}^n	246
	Spazio tangente ad una varietà	248
	Volumi di parallelepipedi k -dimensionali in \mathbb{R}^n ($k \leq n$)	249
	Esercizi	249
D	Algebra lineare	255
D.1	Spazi vettoriali	255
D.2	Applicazioni lineari	261
D.3	Matrici	263
D.4	Determinanti	265
D.5	Matrici e applicazioni lineari	273
D.6	Sottospazi invarianti e matrici a blocchi	278
D.7	Diagonalizzazione	279
D.8	Matrici definite positive/negative	282
D.9	Forma normale (o canonica) di Jordan	283

Capitolo 1

Topologia euclidea e continuità

In questo capitolo, tramite la nozione di norma euclidea – che generalizza a \mathbb{R}^n la nozione di modulo – e la relativa nozione di distanza, si estende a \mathbb{R}^n gran parte delle proprietà topologiche di \mathbb{R} . Infatti, molte dimostrazioni in \mathbb{R}^n , per $n \geq 2$, avendo dato le giuste definizioni, sono sostanzialmente uguali a quelle unidimensionali¹.

\mathbb{R}^n come spazio vettoriale

Dato² $n \in \mathbb{N}$, \mathbb{R}^n denota il prodotto cartesiano $\mathbb{R} \times \cdots \times \mathbb{R}$ (n -volte), ossia, l'insieme delle n -uple ordinate $x = (x_1, \dots, x_n)$ con $x_i \in \mathbb{R}$, per ogni $1 \leq i \leq n$. Di norma, se $x := (x_1, \dots, x_n) \in \mathbb{R}^n$, denotiamo con x_i la sua i -ma componente e chiameremo i -esima proiezione l'applicazione (lineare) che associa a x la sua i -esima componente:

$$x \in \mathbb{R}^n \mapsto \pi_i(x) := x_i \in \mathbb{R}. \quad (1.1)$$

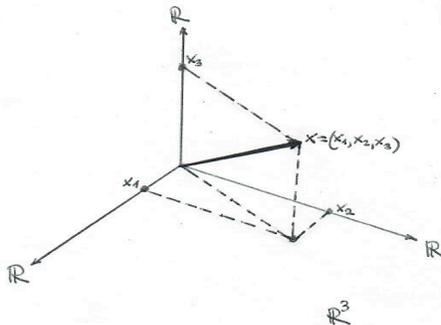


Figura 1.1: Un vettore $x \in \mathbb{R}^3$

\mathbb{R}^n è uno spazio vettoriale sul campo³ \mathbb{R} , ossia, è definita la somma di due elementi (o ‘punti’ o ‘vettori’) x e y di \mathbb{R}^n ed il prodotto di un vettore x di \mathbb{R}^n con uno ‘scalare’ $a \in \mathbb{R}$:

$$(S) \quad x + y = (x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n);$$

$$(P) \quad ax = (ax_1, \dots, ax_n);$$

¹Per la teoria unidimensionale (ossia, per l’analisi su \mathbb{R}), rimandiamo alla *prima parte* del Corso di Analisi (*una introduzione rigorosa all’analisi matematica su \mathbb{R}*), a cui ci riferiremo con [C2019], cfr. <https://www.mheducation.it/corso-di-analisi-prima-parte-9788838695438-italy>.

²Come in [C2019], \mathbb{N} denota l’insieme dei numeri naturali $\mathbb{N} = \{1, 2, 3, \dots\}$, mentre $\mathbb{N}_0 := \{0, 1, 2, \dots\}$ denota l’insieme dei numeri interi non negativi (cfr., anche <https://mathworld.wolfram.com/NaturalNumber.html>).

³Per richiami di algebra lineare si veda l’Appendice D.

le proprietà (S) e (P) sono tali che: la somma (S) è commutativa e associativa; l'elemento neutro è $0 := (0, \dots, 0)$; per ogni vettore $x = (x_1, \dots, x_n)$ esiste l'opposto $-x := (-x_1, \dots, -x_n)$ tale che $x - x := x + (-x) = 0$; inoltre si hanno le seguenti relazioni tra (S) e (P):

$$a(bx) = (ab)x, \quad 1x = x, \quad a(x + y) = ax + ay, \quad (a + b)x = ax + bx, \quad \forall a, b \in \mathbb{R}, \quad \forall x, y \in \mathbb{R}^n.$$

Da (S) e (P) segue immediatamente che ogni vettore $x = (x_1, \dots, x_n) \in \mathbb{R}^n$ si esprime (in modo unico) come combinazione lineare

$$x = \sum_{i=1}^n x_i e^{(i)}, \quad \text{dove :} \quad \begin{cases} e^{(1)} := (1, 0, 0, \dots, 0), \\ e^{(2)} := (0, 1, 0, \dots, 0), \\ \vdots \\ e^{(n)} := (0, 0, \dots, 0, 1), \end{cases} \quad (1.2)$$

ossia, i vettori $\{e^{(j)}\} := \{e^{(j)} \mid 1 \leq j \leq n\}$, le cui componenti sono date da⁴

$$(e^{(j)})_i := \pi_i(e^{(j)}) = \delta_{ij} := \begin{cases} 1 & \text{se } i = j, \\ 0 & \text{se } i \neq j, \end{cases} \quad (1.3)$$

formano una base⁵, detta **base standard** di \mathbb{R}^n ; i vettori $e^{(i)}$ sono detti **versori** (standard).

Osservazione 1.1 (i) La notazione adottata qui per indicare i vettori di \mathbb{R}^n è standard; ciononostante in molti testi introduttivi di matematica o fisica si usano notazioni che enfatizzano la differenza dal caso unidimensionale; per esempio, vettori in \mathbb{R}^n vengono, a volte, indicati con simboli quali \vec{x} o \mathbf{x} . La notazione adottata in questo testo è più 'leggera' ma richiede, soprattutto all'inizio, una certa attenzione. Per esempio, descrivendo l'elemento neutro $0 \in \mathbb{R}^n$ rispetto alla somma (S), abbiamo usato lo stesso simbolo usato per un oggetto, in generale, diverso, ossia, l'elemento neutro $0 \in \mathbb{R}$ del campo degli scalari \mathbb{R} . Il contesto dovrebbe chiarire questi 'abusi di notazione', che, però, richiedono – come si diceva – una certa attenzione iniziale.

(ii) Nel caso $n = 2$ o $n = 3$, spesso, adatteremo la notazione 'classica' (x, y) per elementi di \mathbb{R}^2 e (x, y, z) per elementi di \mathbb{R}^3 al posto di (x_1, x_2) e (x_1, x_2, x_3) .

(iii) Vi è, però, un altro aspetto più profondo da discutere in relazione alla definizione di \mathbb{R}^n . Come sottolineato varie volte in [C2019], l'uso dei 'puntini' nelle definizioni matematiche, anche se 'efficace' e d'uso assai comune, non è rigoroso ma piuttosto si basa sull'intuizione. Una definizione di \mathbb{R}^n più precisa è la seguente. Sia $n \in \mathbb{N}$ e $\mathcal{F}_n := \{i \in \mathbb{N} \mid i \leq n\}$, allora $\mathbb{R}^n := \{x \mid x : i \in \mathcal{F}_n \mapsto x_i \in \mathbb{R}\}$, ossia \mathbb{R}^n è l'insieme delle funzioni con dominio \mathcal{F}_n e codominio \mathbb{R} e x_i non è altro che il valore della funzione x nel 'punto' $i \in \mathcal{F}_n$. Detto questo una volta per tutte, useremo sistematicamente la 'notazione coi puntini'.

1 Struttura euclidea

1.1 Prodotto scalare e norma euclidea

La 'struttura euclidea' di \mathbb{R}^n si fonda sulle seguenti definizioni:

Definizione 1.2 *Dati $x = (x_1, \dots, x_n)$ e $y = (y_1, \dots, y_n)$ in \mathbb{R}^n definiamo:*

(i) **il prodotto scalare** (o 'prodotto scalare euclideo') di x e y :

$$x \cdot y := \sum_{i=1}^n x_i y_i; \quad (1.4)$$

⁴ δ_{ij} viene normalmente chiamato 'delta di Kronecker'; cfr. https://en.wikipedia.org/wiki/Kronecker_delta.

⁵Cfr. Appendice D.

(ii) la **norma euclidea** di x :

$$|x| := \sqrt{x \cdot x} = \sqrt{\sum_{i=1}^n x_i^2}. \quad (1.5)$$

(iii) la **distanza euclidea** di x e y :

$$d(x, y) := |x - y|. \quad (1.6)$$

Facciamo alcune osservazioni iniziali.

Osservazione 1.3 (i) Nel caso $n = 1$, la norma di x coincide con il modulo di x e per questo motivo usiamo lo stesso simbolo⁶.

(ii) Come è immediato verificare, il *prodotto scalare* è *commutativo* e *lineare*, cioè:

$$\begin{aligned} x \cdot y &= y \cdot x, \quad \forall x, y \in \mathbb{R}^n, \\ (ax) \cdot y &= a(x \cdot y), \quad \forall a \in \mathbb{R}, \forall x, y \in \mathbb{R}^n, \\ x \cdot (y + z) &= (x \cdot y) + (x \cdot z), \quad \forall x, y, z \in \mathbb{R}^n. \end{aligned}$$

(iii) La base standard di \mathbb{R}^n è una base **ortonormale** nel senso che

$$e^{(i)} \cdot e^{(j)} = \delta_{ij}. \quad (1.7)$$

Da questa relazione segue, in particolare, che la i -esima componente di un vettore $x = (x_1, \dots, x_n)$ è dato da $x \cdot e^{(i)}$:

$$x \cdot e^{(i)} = (x_1, \dots, x_n) \cdot e^{(i)} = \left(\sum_{j=1}^n x_j e^{(j)} \right) \cdot e^{(i)} \stackrel{(ii)}{=} \sum_{j=1}^n x_j e^{(j)} \cdot e^{(i)} \stackrel{(1.7)}{=} \sum_{j=1}^n x_j \delta_{ij} = x_i. \quad (1.8)$$

(iv) La parola ‘ortonormale’ nasce dalla fusione di ‘ortogonale’ e di ‘normale’. Dati due vettori $x, y \in \mathbb{R}^n$, si dice che x è **ortogonale** (o ‘perpendicolare’) a y , e si scrive $x \perp y$, se $x \cdot y = 0$. La nozione di ortogonalità è, come noto, una nozione fondamentale per investigare aspetti geometrici.

Il termine, ‘normale’ fa riferimenti al fatto che i vettori $e^{(j)}$ sono vettori di norma 1.

(v) Due proprietà del modulo che si generalizzano immediatamente a \mathbb{R}^n sono:

$$|ax| = |a| |x|, \quad \forall a \in \mathbb{R}, x \in \mathbb{R}^n; \quad (\text{positiva omogeneità di } |\cdot|) \quad (1.9)$$

$$|x| = 0 \iff x = 0, \quad (x \in \mathbb{R}^n). \quad (\text{non degenerazione di } |\cdot|) \quad (1.10)$$

(vi) Da (ii) segue che, per ogni $x, y \in \mathbb{R}^n$ si ha

$$|x + y|^2 = (x + y) \cdot (x + y) = x \cdot x + y \cdot y + 2x \cdot y = |x|^2 + |y|^2 + 2x \cdot y, \quad (1.11)$$

che generalizza a \mathbb{R}^n la ‘formula del quadrato del binomio’.

In \mathbb{R} , il modulo del prodotto di due numeri coincide col prodotto dei moduli di tali numeri. Questa proprietà *non* si generalizza a \mathbb{R}^n , dove vale, invece, la seguente (fondamentale) *disuguaglianza*:

Proposizione 1.4 (Disuguaglianza di Cauchy)

$$|x \cdot y| \leq |x| |y|, \quad \forall x, y \in \mathbb{R}^n. \quad (1.12)$$

⁶In altri testi si preferisce usare, per la norma euclidea in \mathbb{R}^n , il simbolo $\|\cdot\|$ oppure $\|\cdot\|_2$.

Dimostrazione Se $x = 0$ la (1.12) vale col segno di uguale. Se $x \neq 0$, consideriamo la parabola

$$t \in \mathbb{R} \mapsto p(t) := |tx + y|^2 \stackrel{(1.11)}{=} t^2(x \cdot x) + 2t(x \cdot y) + (y \cdot y).$$

Poiché $p(t) \geq 0$ per ogni t , la parabola p deve avere discriminante non positivo, il che è equivalente a (1.12). ■

Osservazione 1.5 Si noti che la disuguaglianza di Cauchy è *equivalente* a

$$\sum_{i=1}^n |x_i| |y_i| \leq |x| |y|, \quad \forall x, y \in \mathbb{R}^n; \quad (1.13)$$

infatti, da (1.13) segue immediatamente (1.12), essendo $|\sum x_i y_i| \leq \sum |x_i| |y_i|$; viceversa, se vale (1.12), essa vale anche applicata ai vettori $(|x_1|, \dots, |x_n|)$ e $(|y_1|, \dots, |y_n|)$, e dunque, vale anche (1.13).

La disuguaglianza di Cauchy permette di estendere a \mathbb{R}^n anche la fondamentale ‘disuguaglianza triangolare’⁷ di \mathbb{R} :

Proposizione 1.6 (Disuguaglianza triangolare in \mathbb{R}^n)

$$|x + y| \leq |x| + |y|, \quad \forall x, y \in \mathbb{R}^n. \quad (1.14)$$

Dimostrazione Elevando al quadrato, ed usando la (1.11) per $|x + y|^2$, si vede immediatamente che la relazione (1.14) è equivalente a $x \cdot y \leq |x| |y|$, che segue da (1.12). ■

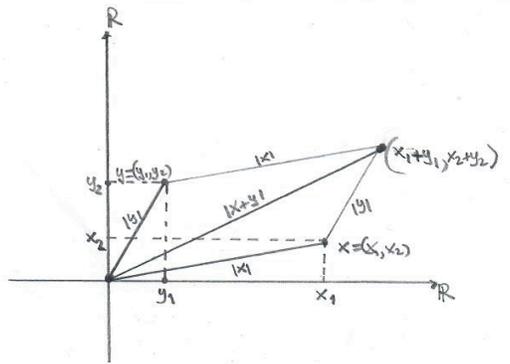


Figura 1.2: Disuguaglianza triangolare in \mathbb{R}^2

Osservazione 1.7 (i) Il segno di uguaglianza in (1.12) vale se e solo se $y = 0$ oppure se $y \neq 0$ e $x = ty$ con⁸ $t \in \mathbb{R}$ (x e y sono ‘co-lineari’).

(ii) La disuguaglianza triangolare (1.14) è equivalente alla disuguaglianza⁹:

$$d(u, v) = |u - v| \geq ||u| - |v||, \quad \forall u, v \in \mathbb{R}^n. \quad (1.15)$$

Infatti, da (1.14) segue che

$$\begin{cases} |u| = |u + v - v| = |(u - v) + v| \leq |u - v| + |v| \\ |v| = |v + u - u| = |(v - u) + u| \leq |u - v| + |u| \end{cases} \implies ||u| - |v|| \leq |u - v|.$$

Viceversa, (1.15) implica $|u| \leq |u - v| + |v|$ e ponendo $u := x + y$ e $v := y$ si ottiene (1.14).

⁷Per inciso, guardando la figura in \mathbb{R}^2 si capirà finalmente il nome dato a tale disuguaglianza.

⁸Es 1.1.

⁹Si faccia attenzione al ‘diverso’ significato del simbolo ‘|’ nel membro di destra, essendo $|u| - |v| \in \mathbb{R}$ e $u, v \in \mathbb{R}^n$.

Riassumendo, la norma euclidea $|\cdot|$ è una funzione su \mathbb{R}^n a valori in $[0, +\infty)$ positivamente omogenea, non degenera e che soddisfa la disuguaglianza triangolare (cfr., rispettivamente, (1.9), (1.10) e (1.14)).

1.2 Norme; norme equivalenti

La coppia $(\mathbb{R}^n, |\cdot|)$ è un esempio di ‘spazio normato’ nel senso della seguente definizione generale:

Definizione 1.8 *Uno spazio normato è una coppia $(X, \|\cdot\|)$ dove $X \neq \emptyset$ è uno spazio vettoriale sul campo¹⁰ K e $\|\cdot\|$ è una **norma** su X , ossia, funzione da X in $[0, +\infty)$ che verifica:*

- (i) $\|ax\| = |a| \|x\|, \forall a \in K, \forall x \in X$; (positiva omogeneità)
- (ii) $\|x\| = 0 \implies x = 0$; (non degenerazione)
- (iii) $\|x + y\| \leq \|x\| + \|y\|, \forall x, y \in X$. (disuguaglianza triangolare)

Osservazione 1.9 (i) Si noti che da (i) segue che $\|0\| = 0$ e dunque $\|x\| = 0 \iff x = 0$.

(ii) $(\mathbb{R}^n, |\cdot|)$ è uno spazio normato, ma come ci si aspetta, la norma euclidea non è l’unica norma possibile. Per esempio, useremo spesso anche le seguenti due norme, dette, rispettivamente, **norma uniforme** (anche: ‘norma del sup’, ‘norma infinito’ o ‘norma del massimo’) e la **norma 1**:

$$|x|_\infty := \max\{|x_1|, \dots, |x_n|\}, \quad |x|_1 := \sum_{i=1}^n |x_i|, \quad (x \in \mathbb{R}^n). \quad (1.16)$$

La verifica che queste due funzioni siano effettivamente delle norme nel senso della Defini-

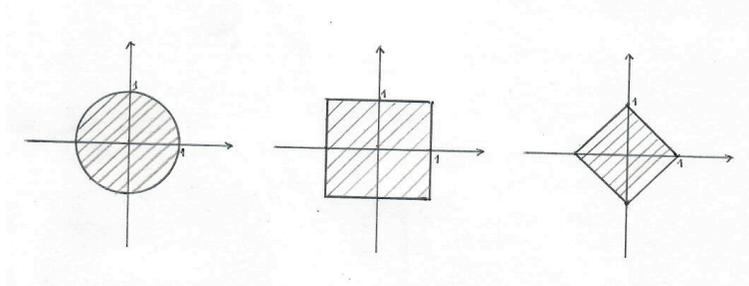


Figura 1.3: ‘Sfere unitarie’ nelle norme $|\cdot|, |\cdot|_\infty$ e $|\cdot|_1$ su \mathbb{R}^2

zione 1.8 è immediata¹¹. La scelta di usare di default la norma euclidea è basato sulla sua relazione ‘naturale’ con il prodotto scalare.

(iii) Dal punto di vista degli spazi normati, la norma euclidea, la norma uniforme e la norma 1 su \mathbb{R}^n sono equivalenti, nel senso della seguente definizione:

Definizione 1.10 *Sia X uno spazio vettoriale. Due norme $\|\cdot\|_a$ e $\|\cdot\|_b$ su X si dicono equivalenti se esistono due costanti positive c_1 e c_2 tali che*

$$\|x\|_a \leq c_1 \|x\|_b \leq c_2 \|x\|_a, \quad \forall x \in X. \quad (1.17)$$

Infatti, su \mathbb{R}^n , valgono le seguenti relazioni¹²:

$$|x|_\infty \leq |x| \leq |x|_1 \leq \sqrt{n}|x| \leq n|x|_\infty, \quad \forall x \in \mathbb{R}^n. \quad (1.18)$$

¹⁰Normalmente, in questo testo $K = \mathbb{R}$ o, più raramente, $K = \mathbb{C}$.

¹¹Es. 1.2.

¹²La prima disuguaglianza è ovvia; $|x| \leq |x|_1$ si verifica elevando al quadrato; $|x|_1 \leq \sqrt{n}|x|$ deriva dalla disuguaglianza di Cauchy (1.12) con $y_i = 1$ per ogni i e, infine, $|x| \leq \sqrt{n}|x|_\infty$ è ovvia.

In realtà, dimostremo tra breve¹³ che:

In \mathbb{R}^n tutte le norme sono equivalenti.

Questa affermazione, in generale, non è vera su spazi normati infinito dimensionali¹⁴.

2 Topologia standard di \mathbb{R}^n

Per studiare proprietà fini degli insiemi di \mathbb{R}^n è indispensabile – così come nel caso unidimensionale – introdurre elementi di topologia¹⁵.

Definizione 1.11 (i) Dato $r > 0$ e $x \in \mathbb{R}^n$, definiamo **sfera aperta** (o, semplicemente, ‘sfera’) di raggio r e centro x , l’insieme

$$B_r(x) := \{y \in \mathbb{R}^n \mid d(y, x) = |y - x| < r\}. \quad (1.19)$$

(ii) Un insieme $A \subseteq \mathbb{R}^n$ si dice **aperto** (‘nella topologia standard’) se esiste $r > 0$ tale che $B_r(x) \subseteq A$. Un insieme $C \subseteq \mathbb{R}^n$ si dice **chiuso** se il suo complementare $C^c := \mathbb{R}^n \setminus C$ è aperto.

(iii) La famiglia \mathcal{A} di tutti gli insiemi aperti definiti in (ii) prende il nome di **topologia standard** di \mathbb{R}^n .

È immediato verificare¹⁶ che la topologia standard è ‘una topologia’ nel senso della seguente definizione generale:

Definizione 1.12 Una **topologia** su di un insieme arbitrario X è, per definizione, una famiglia \mathcal{F} di sottoinsiemi di X – detti ‘insiemi aperti’ – che soddisfa:

- (i) X e \emptyset appartengono a \mathcal{F} ;
- (ii) un’unione arbitraria di elementi di \mathcal{F} è un elemento di \mathcal{F} ;
- (iii) un’intersezione finita di elementi di \mathcal{F} è un elemento di \mathcal{F} .

I complementari degli insiemi aperti si **chiamano insiemi chiusi**.

Uno **spazio topologico** è una coppia (X, \mathcal{F}) con X insieme arbitrario e \mathcal{F} una topologia su X .

Raccogliamo nella seguente definizione la terminologia topologica essenziale:

Definizione 1.13 Sia $E \subseteq \mathbb{R}^n$.

La **chiusura** di E , denotata \bar{E} , è il più piccolo chiuso contenente E (ossia è l’intersezione di tutti i chiusi contenenti E).

L’**interno** di E , denotato $\overset{\circ}{E}$, è il più grande aperto contenuto in E (ossia è l’unione di tutti gli aperti contenuti in¹⁷ E).

La **frontiera** (insiemistica) di E , denotata ∂E , è l’insieme $\partial E := \bar{E} \setminus \overset{\circ}{E}$.

E è un **intorno** di x se¹⁸ $x \in \overset{\circ}{E}$.

Un punto x si dice **interno** ad E se $x \in \overset{\circ}{E}$.

x è un **punto d’accumulazione** (o ‘punto limite’) per E se $B_r(x) \cap (E \setminus \{x\}) \neq \emptyset, \forall r > 0$.

x è un **punto isolato** di E se esiste $r > 0$ tale che $B_r(x) \cap E = \{x\}$.

E è **limitato** se $\exists r > 0$ tale che $E \subseteq B_r(0)$.

Infine, diremo che una proprietà $\mathcal{P}(x)$ è vera **vicino a** \bar{x} se esiste un intorno U di \bar{x} tale che la proprietà $\mathcal{P}(x)$ è vera per ogni $x \in U \setminus \{\bar{x}\}$.

¹³Proposizione 1.36.

¹⁴Ossia, spazi vettoriali (normati) che hanno un numero infinito di vettori indipendenti (cfr. Appendice D). Per esempi di norme non equivalenti, si veda Es 1.61.

¹⁵Per informazioni generali, si veda, per esempio, <https://en.wikipedia.org/wiki/Topology>.

¹⁶Es 1.5.

¹⁷Chiaramente A è aperto se e solo se $A = \overset{\circ}{A}$ e C è chiuso se e solo se $C = \bar{C}$.

¹⁸Si noti che questa definizione è più generale (ma, sostanzialmente, equivalente) di quella data nel caso unidimensionale in [C2019].

Proposizione 1.14 Sia E un qualunque sottoinsieme di \mathbb{R}^n . Allora:

- (i) $(\mathring{E})^c = \overline{E^c}$.
- (ii) $\partial E = \overline{E} \cap \overline{E^c}$.

Dimostrazione (i): per definizione di \mathring{E} , $\mathring{E} \subseteq E$ e passando ai complementari $E^c \subseteq (\mathring{E})^c$ e poiché $(\mathring{E})^c$ è chiuso dalla definizione di chiusura d'un insieme segue che $\overline{E^c} \subseteq (\mathring{E})^c$. D'altra parte, $E^c \subseteq \overline{E^c}$ e quindi $(\overline{E^c})^c \subseteq E$; poiché $(\overline{E^c})^c$ è aperto, dalla definizione di interno d'un insieme segue che $(\overline{E^c})^c \subseteq \mathring{E}$ ossia $(\mathring{E})^c \subseteq \overline{E^c}$.

(ii): L'asserto deriva dal punto (i); infatti $\partial E = \overline{E} \setminus \mathring{E} = \overline{E} \cap (\mathring{E})^c = \overline{E} \cap \overline{E^c}$.

Sebbene in questo testo considereremo solo la topologia standard su \mathbb{R}^n , useremo anche la seguente nozione.

Definizione 1.15 Dato un qualunque insieme $E \subseteq \mathbb{R}^n$, la famiglia di insiemi

$$\mathcal{A}_E := \{A \cap E \mid A \in \mathcal{A}\} \quad (1.20)$$

viene chiamata **topologia relativa** (o, 'topologia indotta') di E .

Il fatto che \mathcal{A}_E in (1.20) formi una topologia su E nel senso della Definizione 1.12 è una immediata verifica (Es 1.5).

Osservazione 1.16 (i) La topologia standard di \mathbb{R}^n è stata definita tramite la norma euclidea. Di fatto, la topologia standard di \mathbb{R}^n non dipende dalla norma che si usa per definire le sfere (e quindi gli aperti). Per esempio, se avessimo usato la norma uniforme al posto della norma euclidea avremmo ottenuto la stessa famiglia di aperti. Infatti, denotiamo $B'_r(x) := \{y \in \mathbb{R}^n \mid |y - x|_\infty < r\}$ e denotiamo \mathcal{A}' la famiglia di insiemi A tali che, per ogni $x \in A$ esiste $r > 0$ tale che $B'_r(x) \subseteq A$ (in altre parole, \mathcal{A}' è la famiglia di aperti associati alla norma $|\cdot|_\infty$). Osserviamo che da (1.18) segue che, per ogni x e r si ha

$$B'_{\frac{r}{\sqrt{n}}}(x) \subseteq B_r(x) \subseteq B'_r(x). \quad (1.21)$$

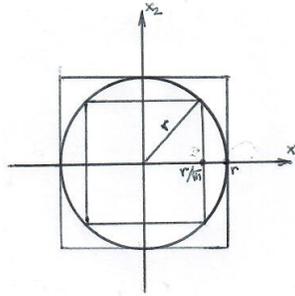


Figura 1.4: (1.21) per $n = 2$.

Quindi, se $x \in A \in \mathcal{A}$, allora esiste $r > 0$ tale che $B_r(x) \subseteq A$, e per (1.21) $B'_{\frac{r}{\sqrt{n}}}(x) \subseteq B_r(x)$ e dunque $A \in \mathcal{A}'$. Viceversa, se $x \in A \in \mathcal{A}'$, allora esiste $r > 0$ tale che $B'_r(x) \subseteq A$, e per (1.21) $B_r(x) \subseteq B'_r(x)$ e dunque $A \in \mathcal{A}$. In conclusione, $\mathcal{A} = \mathcal{A}'$.

Questo argomento non si basa sul valore esplicito delle costanti che lega la norma euclidea con quella uniforme, ma esclusivamente sulla equivalenza di tali norme: dunque, in vista del fatto che su \mathbb{R}^n tutte le norme sono equivalenti (cfr. Proposizione 1.36) si ha che qualunque norma genera la stessa topologia su \mathbb{R}^n .

(ii) Esempi di topologie diverse su \mathbb{R}^n (che non useremo mai in questo testo) sono la 'topologia banale' $\mathcal{A}_0 := \{\mathbb{R}^n, \emptyset\}$ in cui i soli insiemi aperti sono \mathbb{R}^n e \emptyset oppure la 'topologia discreta' $\mathcal{A}_1 := \{A \mid A \subseteq \mathbb{R}^n\}$ in cui qualunque sottoinsieme di \mathbb{R}^n è aperto.

3 Successioni e completezza

Molte proprietà di \mathbb{R}^n possono essere descritte in termini di *successioni*.

Definizione 1.17 (i) Una **successione** $\{x^{(k)}\}$ in \mathbb{R}^n è una funzione da \mathbb{N} in \mathbb{R}^n ossia una mappa che associa ad ogni $k \in \mathbb{N}$ un vettore¹⁹ $x^{(k)} := (x_1^{(k)}, \dots, x_n^{(k)}) \in \mathbb{R}^n$.

(ii) Una successione si dice **di Cauchy** (o ‘fondamentale’) se

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad |x^{(k)} - x^{(j)}| < \varepsilon \quad \forall k, j \geq N. \quad (1.22)$$

(iii) Una successione $\{x^{(k)}\}$ in \mathbb{R}^n **converge** a $x \in \mathbb{R}^n$ o ‘ha limite x ’ (in formule, $\lim_{k \rightarrow \infty} x^{(k)} = x$, oppure $\lim x^{(k)} = x$, o anche $x^{(k)} \rightarrow x$) se

$$\forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad |x^{(k)} - x| < \varepsilon, \quad \forall k \geq N. \quad (1.23)$$

(iv) Una **sottosuccessione** di $\{x^{(k)}\}$ è una successione della forma

$$\{x^{(k_j)}\} := \{(x_1^{(k_j)}, \dots, x_n^{(k_j)})\}$$

con $\{k_j\}$ successione a valori in \mathbb{N} e strettamente crescente.

Osservazione 1.18 Come nel caso unidimensionale, una successione $\{x^{(k)}\}$ di Cauchy è limitata, ossia, esiste $M > 0$ tale che $|x^{(k)}| \leq M$ per ogni k .

Dimostrazione Per (1.22), $\exists N \in \mathbb{N}$ tale che $|x^{(k)} - x^{(N)}| < 1$ se $k \geq N$. Sia

$$M := \max\{|x^{(1)}|, \dots, |x^{(N-1)}|, |x^{(N)}| + 1\}.$$

Allora, $|x^{(k)}| \leq M$, $\forall k \leq N$ e, se $k > N$,

$$|x^{(k)}| = |x^{(k)} - x^{(N)} + x^{(N)}| \leq |x^{(k)} - x^{(N)}| + |x^{(N)}| \leq 1 + |x^{(N)}| \leq M. \quad \blacksquare$$

Proposizione 1.19 (i) Una successione $\{x^{(k)}\}$ in \mathbb{R}^n è di Cauchy se e solo se sono di Cauchy (in \mathbb{R}) le n successioni $\{x_1^{(k)}\}, \dots, \{x_n^{(k)}\}$; $\{x^{(k)}\}$ converge a x se e solo se $\{x_i^{(k)}\}$ converge a x_i per ogni $1 \leq i \leq n$.

(ii) [Teorema di Bolzano–Weierstrass] Se $\{x^{(k)}\}$ è una successione limitata in \mathbb{R}^n allora esiste una sottosuccessione di $\{x^{(k)}\}$ convergente.

(iii) \mathbb{R}^n è completo (ossia, ogni successione di Cauchy ammette limite in \mathbb{R}^n).

Dimostrazione (i): Segue immediatamente da (1.18) (con $x^{(k)} - x$ al posto di x).

(ii): Nel caso $n = 1$ il risultato è noto²⁰. Sia $n \geq 2$. Poiché $\{x^{(k)}\}$ è limitata, per l’Osservazione 1.18, esiste $M > 0$ tale che $|x^{(k)}| \leq M$ per ogni k . Quindi, per (1.18) si ha anche $|x_j^{(k)}| \leq M$, per $1 \leq j \leq n$. Per il caso unidimensionale, esiste una successione $\{k_j\}$ strettamente crescente e $x_1 \in \mathbb{R}$ tale che $\lim x_1^{(k_j)} = x_1$. Poiché $|x_2^{(k_j)}| \leq M$, sempre per il caso unidimensionale, esiste una sottosuccessione $\{k'_j\}$ di $\{k_j\}$ e $x_2 \in \mathbb{R}$ tale che $\lim x_2^{(k'_j)} = x_2$.

Iterando (se $n \geq 3$), otteniamo una successione $\{k^*_j\}$ e numeri x_1, \dots, x_n tali che $\lim x_j^{(k^*_j)} = x_j$ per ogni $1 \leq j \leq n$ e per opportuni $x_j \in \mathbb{R}^n$. Da (1.18) segue, ora, che $\{x^{(k^*_j)}\}$ converge a x .

(iii): Se $\{x^{(k)}\}$ è di Cauchy, sempre da (1.18), segue che sono di Cauchy le n successioni $\{x_j^{(k)}\}$.

Dalla completezza di \mathbb{R} segue che esistono n numeri $x_j \in \mathbb{R}$ tali che $x_j^{(k)} \rightarrow x_j$ per ogni j e usando ancora una volta (1.18) si ha che $x^{(k)} \rightarrow x \in \mathbb{R}^n$. \blacksquare

¹⁹Per evitare i doppi indici, per le successioni, indichiamo, di solito, la dipendenza dalla variabile indipendente $k \in \mathbb{N}$ in apice e fra parentesi tonde.

²⁰Cfr. Osservazione 6.5–(i) in [C2019].

Proposizione 1.20 Sia $E \subseteq \mathbb{R}^n$.

- (i) E è chiuso se e solo se per ogni successione $\{x^{(k)}\}$ in E convergente in \mathbb{R}^n , $\lim x^{(k)} \in E$.
(ii) $\bar{E} = \{x \in \mathbb{R}^n \mid \exists \{x^{(k)}\} \text{ in } E \text{ con } x^{(k)} \rightarrow x\}$.

Dimostrazione (i): L'enunciato è equivalente (per 'contrapposizione') a: ' E non chiuso' se e solo se ' $\exists \{x^{(k)}\}$ in E tale che $x^{(k)} \rightarrow x \notin E$ '. Ora, E non chiuso $\iff E^c$ non aperto $\iff \exists \bar{x} \in E^c$ tale che, $\forall r > 0$, $B_r(\bar{x}) \cap E \neq \emptyset \iff \exists \bar{x} \in E^c$ tale che, $\forall k \in \mathbb{N}$, $B_{\frac{1}{k}}(\bar{x}) \cap E \neq \emptyset \iff \exists \{x^{(k)}\} \subseteq E$ tale che $x^{(k)} \rightarrow \bar{x} \in E^c$.

(ii): Sia $D := \{x \in \mathbb{R}^n \mid \exists \{x^{(k)}\} \text{ in } E \text{ con } x^{(k)} \rightarrow x\}$ e dimostriamo che D è chiuso. Sia $\{x^{(k)}\}$ una successione in D convergente a $y \in \mathbb{R}^n$. Dalla definizione di D segue che, per ogni $k \in \mathbb{N}$, $x^{(k)}$ esiste una successione $\{x^{(k,j)}\}$ in E tale che $\lim_{j \rightarrow +\infty} x^{(k,j)} = x^{(k)}$. Quindi, per ogni k esiste un j_k tale che $|x^{(k,j_k)} - x^{(k)}| < 1/k$, per cui:

$$|x^{(k,j_k)} - y| \leq |x^{(k,j_k)} - x^{(k)}| + |x^{(k)} - y| < \frac{1}{k} + |x^{(k)} - y| \rightarrow 0 \implies \lim_{k \rightarrow +\infty} x^{(k,j_k)} = y.$$

Ma questo significa che $y \in D$ e quindi, per il punto (i), D è chiuso.

Sia, ora, $C \subseteq \mathbb{R}^n$ un chiuso che contiene E e sia $y \in D$. Per definizione di D , esiste una successione $\{x^{(k)}\}$ in E tale che $\lim x^{(k)} = y$, ma poiché $E \subseteq C$, tale successione è anche in C e, poiché C è chiuso, per il punto (i), $y = \lim x^{(k)} \in C$. Dunque, $D \subseteq C$. Ma questo significa che D è il più piccolo chiuso che contiene E e quindi, per definizione di chiusura, $D = \bar{E}$. ■

4 Continuità e limiti

Definizione 1.21 Dato $E \subseteq \mathbb{R}^n$ e una funzione $f : E \rightarrow \mathbb{R}^m$, diremo che f è **continua** in $\bar{x} \in E$ se

$$\forall \varepsilon > 0 \exists \delta > 0 \text{ tale che } |f(x) - f(\bar{x})| < \varepsilon, \forall x \in B_\delta(\bar{x}) \cap E; \quad (1.24)$$

f si dice **continua su** E se è continua in ogni punto \bar{x} di E .

L'insieme delle funzioni continue da E in \mathbb{R}^m si denota con $C(E, \mathbb{R}^m)$.

Definizione 1.22 Sia $f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$, \bar{x} un punto di accumulazione per E . Se $\delta > 0$, poniamo:

$$E_\delta(\bar{x}) := \{x \in E \mid 0 < |x - \bar{x}| < \delta\} = E \cap \{x \in \mathbb{R}^n \mid 0 < |x - \bar{x}| < \delta\}.$$

- (i) (Limiti finiti) $\alpha \in \mathbb{R}^m$ è il **limite di f per x che tende a \bar{x}** , $\alpha = \lim_{x \rightarrow \bar{x}} f(x)$, se

$$\forall \varepsilon > 0 \exists \delta > 0 \text{ tale che } |f(x) - \alpha| < \varepsilon, \forall x \in E_\delta(\bar{x}). \quad (1.25)$$

- (ii) Nel caso $m = 1$, grazie all'ordine totale di \mathbb{R} , possiamo includere valori $\pm\infty$ del limite²¹:

$$\lim_{x \rightarrow \alpha} f(x) = \alpha \in \mathbb{R}^* \iff \forall V \text{ intorno di } \alpha, \exists \delta > 0 \mid f(x) \in V, \forall x \in E_\delta(\bar{x}). \quad (1.26)$$

Definiamo il **limite superiore** (o '**massimo limite**') e **limite inferiore** (o '**minimo limite**') di f , per x tende a \bar{x} , rispettivamente, i valori in \mathbb{R}^*

$$\begin{aligned} \limsup_{x \rightarrow \bar{x}} f(x) &:= \overline{\lim}_{x \rightarrow \bar{x}} f(x) := \inf_{\delta > 0} \sup_{x \in E_\delta} f(x), \\ \liminf_{x \rightarrow \bar{x}} f(x) &:= \underline{\lim}_{x \rightarrow \bar{x}} f(x) := \sup_{\delta > 0} \inf_{x \in E_\delta} f(x). \end{aligned} \quad (1.27)$$

Infine, se E è non limitato, diremo anche che

$$\lim_{|x| \rightarrow +\infty} f(x) = \alpha \in \mathbb{R}^* \iff \forall V \text{ intorno di } \alpha, \exists r > 0 \mid f(x) \in V, \forall x \in E, |x| > r. \quad (1.28)$$

²¹Si ricorda che $\mathbb{R}^* = \mathbb{R} \cup \{+\infty\} \cup \{-\infty\}$ e che un intorno di $+\infty$ è un intervallo della forma $(a, +\infty)$ e in intorno di $-\infty$ è un intervallo della forma $(-\infty, a)$; cfr. [C2019].

Osservazione 1.23 (i) Ovviamente, se \bar{x} non è un punto isolato di E , allora f è continua in \bar{x} se e solo se $\lim_{x \rightarrow \bar{x}} f(x) = f(\bar{x})$.

(ii) La ‘funzione vettoriale²²’ $f : x \in E \subseteq \mathbb{R}^n \mapsto f(x) := (f_1(x), \dots, f_m(x)) \in \mathbb{R}^m$ è continua in $\bar{x} \in E$ se e solo se sono continue in \bar{x} le m ‘funzioni scalari’ $f_i : E \rightarrow \mathbb{R}$.

Analoga affermazione vale per i limiti di una funzione vettoriale $f : E \rightarrow \mathbb{R}^m$: $\lim_{x \rightarrow \bar{x}} f(x) = \alpha$ se e solo se, per ogni i , $\lim_{x \rightarrow \bar{x}} f_i(x) = \alpha_i$

Dimostrazione Siano f_i continue in \bar{x} e sia $\varepsilon > 0$. Sia δ tale che $|f_i(x) - f_i(\bar{x})| < \varepsilon/\sqrt{n}$ se $x \in E \cap B_\delta(\bar{x})$: allora, per tali x , $|f(x) - f(\bar{x})| \leq \sqrt{n} |f(x) - f(\bar{x})|_\infty < \varepsilon$.

Il viceversa è ovvio, essendo $|f_i(x)| \leq |f(x)|$.

L’analoga dimostrazione per i limiti è lasciata come (facile) esercizio. \blacksquare

(iii) Nel ‘caso scalare’ ($m = 1$), dalla definizione di limite deriva immediatamente il ‘teorema di permanenza del segno’:

se $L = \lim_{x \rightarrow \bar{x}} f(x) \neq 0$, allora esiste $\delta > 0$ tale che $f(x)$ ha lo stesso segno di L , per ogni $x \in E_\delta(\bar{x})$.

La dimostrazione è identica al caso unidimensionale $n = 1$.

Nel caso generale (codominio \mathbb{R}^m), se $\lim_{x \rightarrow \bar{x}} f(x) \neq 0$, allora esiste $\delta > 0$ tale che $f(x) \neq 0$, per ogni²³ $x \in E_\delta(\bar{x})$.

(iv) Come nel caso unidimensionale ($n = 1$), vale l’algebra dei limiti su²⁴ \mathbb{R} , con sostanzialmente identiche dimostrazioni, dove al modulo andrà sostituita la norma euclidea; naturalmente, nel caso $m \geq 2$ il codominio non è un campo, ma solo uno spazio vettoriale e quindi ‘l’algebra’ che va considerata riguarda solo somma e moltiplicazioni per scalari.

Da queste osservazioni seguono anche la validità delle proprietà fondamentali delle funzioni continue che, per completezza, qui riportiamo:

Siano f, g funzioni continue in $\bar{x} \in E$ ed a valori in \mathbb{R}^m e sia $h : A \subseteq \mathbb{R}^m \rightarrow \mathbb{R}^p$ continua in $\bar{y} \in A$. Allora:

$af + bg$ è continua in \bar{x} , per ogni $a, b \in \mathbb{R}$.

Se $f(\bar{x}) = \bar{y}$, $h \circ f$ è continua in \bar{x} .

Nel caso $m = 1$, fg è continua in \bar{x} ; nel caso $m = 1$ e $f(\bar{x}) \neq 0$, $\frac{1}{f}$ è continua in²⁵ \bar{x} .

(v) Un’altra generalizzazione immediata riguarda il cosiddetto ‘teorema ponte’:

Una funzione f è continua in $\bar{x} \in E$ se e solo se

$$\forall \{x^{(k)}\} \text{ in } E \quad \text{tale che} \quad \lim_{k \rightarrow \infty} x^{(k)} = \bar{x} \quad \implies \quad \lim_{k \rightarrow \infty} f(x^{(k)}) = f(\bar{x}). \quad (1.29)$$

In particolare, se esistono due successioni $x^{(k)} \rightarrow \bar{x}$ e $\bar{x}^{(k)} \rightarrow \bar{x}$ tali che $\lim f(x^{(k)}) \neq \lim f(\bar{x}^{(k)})$, allora f non è continua in \bar{x} .

(vi) Un **cammino** in $E \subseteq \mathbb{R}^n$ è, per definizione, una funzione continua $z : t \in [a, b] \mapsto z(t) = (z_1(t), \dots, z_n(t)) \in E$. Se $f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ è continua in $\bar{x} \in E$ e $z \in C([a, b], E)$ è un cammino tale che $z(t_0) = \bar{x}$, per $t_0 \in [a, b]$, allora, dal punto (iv) qui sopra, segue che $f \circ z$ è continua in t_0 . In generale, si ha²⁶:

$f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ è continua in $\bar{x} \in \overset{\circ}{E}$ se e solo se per ogni cammino $z : [a, b] \rightarrow E$ tale che $z(t_0) = \bar{x}$ per un qualche $a < t_0 < b$, la funzione $f \circ z$ è continua in t_0 .

Usando la nozione di topologia relativa (Definizione 1.15) la continuità può essere espressa in termini puramente topologici (senza far ricorso alla metrica euclidea). Ricordiamo che se $f : X \rightarrow Y$ e $A \subseteq Y$, la **preimmagine** o (o ‘controimmagine’) di A è definita come

$$f^{-1}(A) := \{x \in X \mid f(x) \in A\}.$$

²²Chiameremo una funzione ‘vettoriale’ se il suo codominio è \mathbb{R}^m e ‘scalare’ o ‘unidimensionale’ se il codominio è \mathbb{R} .

²³Se $L = \lim_{x \rightarrow \bar{x}} f(x) \neq 0$, esiste i tale che $L_i \neq 0$ e si può applicare il caso unidimensionale alla componente $f_i(x)$.

²⁴Cfr., per esempio, [C2019, Sez. 2.6].

²⁵Poiché $f(\bar{x}) \neq 0$, $1/f$ è definita in un intorno di \bar{x} .

²⁶Es 1.14.

Le seguenti proprietà sono di immediata verifica:

$$\begin{aligned} f^{-1}(A \cap B) &= f^{-1}(A) \cap f^{-1}(B), & f^{-1}(A \cup B) &= f^{-1}(A) \cup f^{-1}(B), & (1.30) \\ f^{-1}(f(X)) &= X, & A \subseteq B &\implies f^{-1}(A) \subseteq f^{-1}(B). \end{aligned}$$

Allora, si ha:

Proposizione 1.24 *Una funzione $f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ è continua su E se e solo se per ogni aperto $U \subseteq \mathbb{R}^m$, $f^{-1}(U)$ è un aperto nella topologia relativa. In particolare, f è continua su \mathbb{R}^n se e solo se per ogni aperto $U \subseteq \mathbb{R}^m$, $f^{-1}(U)$ è aperto.*

Dimostrazione Supponiamo che f sia continua su E e sia U un aperto di \mathbb{R}^m tale che $f^{-1}(U) \neq \emptyset$ (altrimenti non c'è nulla da verificare essendo \emptyset un aperto). Dobbiamo dimostrare che ogni punto $\bar{x} \in f^{-1}(U)$ è un punto interno nella topologia relativa di E . Sia $y_0 := f(\bar{x})$. Poiché U è aperto esiste una sfera aperta $B_r(y_0)$ interamente contenuta in U . Da (1.24) con $\varepsilon = r$ segue che l'insieme $E \cap B_\varepsilon(\bar{x}) \subseteq f^{-1}(U)$ e l'insieme $E \cap B_\varepsilon(\bar{x})$ è, per definizione, un aperto nella topologia relativa di E , che è quanto si voleva dimostrare.

Viceversa, assumiamo che per ogni aperto $U \subseteq \mathbb{R}^m$, $f^{-1}(U)$ sia un aperto nella topologia relativa. Sia $\bar{x} \in E$, sia $\varepsilon > 0$ e sia $y_0 := f(\bar{x})$. Per la nostra ipotesi $B := f^{-1}(\{y : |y - y_0| < \varepsilon\})$ è un aperto nella topologia relativa di E ossia esiste A aperto di \mathbb{R}^n tale che $B = A \cap E$. Essendo $\bar{x} \in A$ e A aperto esiste $\delta > 0$ tale che $B_\delta(\bar{x}) \subseteq A$ quindi $B_\delta(\bar{x}) \cap E \subseteq B$ il che equivale a (1.24).

La seconda affermazione segue immediatamente dalla prima. ■

Osservazione 1.25 (i) La norma euclidea (o, più in generale, una qualunque norma) $x \in \mathbb{R}^n \rightarrow |x| \in \mathbb{R}$ è continua su \mathbb{R}^n .

Infatti, $\forall \bar{x} \in \mathbb{R}^n$ e $\forall \varepsilon > 0$, se $|x - \bar{x}| < \delta := \varepsilon$, da (1.15) segue che $||x| - |\bar{x}|| \leq |x - \bar{x}| < \varepsilon$.

(ii) Se $1 \leq i \leq n$, la i -esima proiezione $x \in \mathbb{R}^n \rightarrow \pi_i(x) = x_i \in \mathbb{R}$ è continua.

Infatti, per ogni $\bar{x} \in \mathbb{R}^n$ e $\varepsilon > 0$, se $|x - \bar{x}| < \delta := \varepsilon$, si ha: $|x_i - \bar{x}_i| \leq |x - \bar{x}| < \varepsilon$.

(iii) Un **monomio** su \mathbb{R}^n è una funzione della forma $x \in \mathbb{R}^n \mapsto x_1^{\alpha_1} x_2^{\alpha_2} \cdots x_n^{\alpha_n}$, con $\alpha_i \in \mathbb{N}_0$; la somma $\alpha_1 + \cdots + \alpha_n$ definisce il **grado** del monomio; un **polinomio** su \mathbb{R}^n è una combinazione lineare (a coefficienti in \mathbb{R}) di monomi su \mathbb{R}^n ; il grado massimo dei monomi che costituiscono un polinomio definisce il grado del polinomio.

Dal punto (ii) e dall'algebra dei limiti (Osservazione 1.23–(iv)) segue che ogni polinomio su \mathbb{R}^n è continuo su tutto \mathbb{R}^n .

(iv) Sia $I \subseteq \mathbb{R}$ e $g : I \rightarrow \mathbb{R}$ continua in $t_0 \in I$. Allora, la funzione $x \in I \times \mathbb{R}^{n-1} \subseteq \mathbb{R}^n \rightarrow f(x) := g(x_1) \in \mathbb{R}$ è una funzione continua in ogni $\bar{x} \in \mathbb{R}^n$ tale che $\bar{x}_1 = t_0$. Infatti, fissato $\varepsilon > 0$, se $\delta > 0$ è tale che $|g(t) - g(t_0)| < \varepsilon$ per ogni $t \in (t_0 - \delta, t_0 + \delta) \cap I$, allora si avrà $|f(x) - f(\bar{x})| = |g(x_1) - g(t_0)| < \varepsilon$ se $x \in I \times \mathbb{R}^{n-1}$ (che significa $x_1 \in I$) e $|x - \bar{x}| < \delta$ (che implica $|x_1 - \bar{x}_1| < \delta$).

Analogamente, se $1 \leq i \leq n$, la funzione $x \rightarrow f(x) := g(x_i)$ sarà continua in ogni $\bar{x} \in \mathbb{R}^n$ tale che $\bar{x}_i = t_0$.

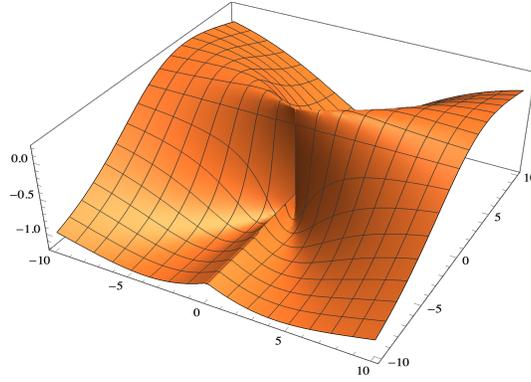
Vediamo, ora, alcuni semplici esempi. In vista dell'Osservazione 1.23–(ii), considereremo solo funzioni scalari.

Esempio 1.26 (i) Sia²⁷ $f : (x, y) \in \mathbb{R}_+ \times \mathbb{R} \mapsto f(x, y) := x^y \in \mathbb{R}$. Allora, $x^y = \exp(y \log x)$, ed essendo $t \rightarrow \log t$ continua su \mathbb{R}_+ e $t \rightarrow \exp(t)$ continua su \mathbb{R} , dall'algebra dei limiti (Osservazione 1.23–(iv)) e dall'Osservazione 1.25–(iv), segue che $f \in C(\mathbb{R}_+ \times \mathbb{R})$.

(ii) Con lo stesso argomento usato nel punto precedente si vede subito che, per ogni $n \geq 3$,

la funzione $f : x \in \mathbb{R}^n \rightarrow \frac{e^{(x_1 x_3)}}{\sqrt{1 + x_2}} + \cos(|x|^{\sqrt{2}} + 2)$ è continua sul suo insieme di definizione, ossia su $\{x \in \mathbb{R}^n : x_2 > -1\}$.

²⁷ $\mathbb{R}_+ := (0, +\infty)$.

Figura 1.5: Grafico di f in (1.31)

(iii) Sia $\alpha \in \mathbb{R}$. La funzione²⁸

$$(x, y) \in \mathbb{R}^2 \mapsto f(x, y) := \begin{cases} \frac{|x|y - y^2}{x^2 + y^2} & \text{se } (x, y) \neq (0, 0) \\ \alpha & \text{se } (x, y) = (0, 0) \end{cases} \quad (1.31)$$

è continua su $\mathbb{R}^2 \setminus \{(0, 0)\}$ ma non è continua in $(0, 0)$.

Infatti, la continuità su $\mathbb{R}^2 \setminus \{(0, 0)\}$ segue dall'argomento del punto (i) qui sopra. D'altra parte,

$$\lim_{k \rightarrow +\infty} f\left(\frac{1}{k}, \frac{1}{k}\right) = 0, \quad \lim_{k \rightarrow +\infty} f\left(0, \frac{1}{k}\right) = -1,$$

e quindi, per il teorema ponte (Osservazione 1.23–(v)), f non è continua in $(0, 0)$.

5 Compattezza

Diamo la definizione generale di compattezza:

Definizione 1.27 Un insieme K si dice **compatto** se da ogni ricoprimento di aperti si può estrarre un sottoricoprimento finito, ossia se vale la seguente proprietà:

$$(a) \quad K \subseteq \bigcup_{j \in J} E_j, \quad E_j \text{ aperto} \implies \exists j_1, \dots, j_k \in J \mid K \subseteq \bigcup_{i=1}^k E_{j_i};$$

(qui J è un insieme arbitrario, non necessariamente numerabile, di indici).

Siano, ora, (b) e (c) le seguenti proprietà:

- (b) ogni successione in K ammette una sottosuccessione convergente in K ;
- (c) K è chiuso e limitato.

La (a) viene chiamata ‘proprietà di Heine–Borel’; la (b) viene chiamata ‘compattezza per successioni’²⁹.

Il seguente fondamentale teorema mostra che, in \mathbb{R}^n , (a), (b) e (c) sono proprietà equivalenti³⁰.

²⁸La funzione f è un esempio di funzione (positivamente) omogenea; cfr. Proposizione 1.35 nella prossima sezione.

²⁹In [C2019] i compatti di \mathbb{R} sono definiti come gli insiemi compatti per successioni.

³⁰In parentesi sono riportati i nomi dei principali matematici legati alla definizione del concetto di compattezza (termine coniato da Maurice Fréchet); per maggiori informazioni, si veda https://en.wikipedia.org/wiki/Compact_space.

Teorema 1.28 (Bolzano, Borel, Fréchet, Heine, Weierstrass) *Sia K un sottoinsieme di \mathbb{R}^n . Allora, le proprietà (a), (b) e (c) sopra elencate sono equivalenti.*

Dimostrazione Dimostriamo le implicazioni: (a) \implies (b) \implies (c) \implies (a). Tutte le dimostrazioni sono ‘per assurdo’.

(a) \implies (b): assumiamo per assurdo che valga (a) e non valga (b). La negazione di (b) è: $\exists \{x^{(k)}\}$ in K che non ammette una sottosuccessione convergente in K . Questo implica che, per ogni $y \in K$, esistono una sfera $B(y)$ di centro y ed un intero $k(y)$ tale che $x^{(k)} \notin B(y)$ per ogni $k \geq k(y)$. Essendo $\{B(y) : y \in K\}$ un ricoprimento aperto di K , per (a) (con $J = K$), si ha che esistono $y^{(1)}, \dots, y^{(m)}$ in K tali che $K \subseteq \bigcup_{i=1}^m B(y^{(i)})$. Se $k_0 := \max\{k(y^{(i)}) \mid 1 \leq i \leq m\}$, per $k \geq k_0$ si ha che $x^{(k)} \notin E_{y^{(i)}}$ per ogni $1 \leq i \leq m$, il che significa che $x^{(k)} \notin K$ se $k \geq k_0$, e questo contraddice l’assunzione che $\{x^{(k)}\}$ ha valori in K .

(b) \implies (c): assumiamo per assurdo che valga (b) e non valga (c), ossia, che K è non limitato oppure non chiuso. Se K non è limitato esiste una successione $\{x^{(k)}\}$ in K tale che $|x^{(k)}| \rightarrow \infty$ e chiaramente, per ogni sottosuccessione $\{x^{(k_j)}\}$ di $\{x^{(k)}\}$, si ha che $\lim_{i \rightarrow \infty} |x^{(k_j)}| = \infty$ il che contraddice (b). Se K non è chiuso, per la Proposizione 1.20 esiste $\bar{x} \in K^c$ e una successione $\{x^{(k)}\}$ in K tale che $x^{(k)} \rightarrow \bar{x}$. Ma allora, per ogni sottosuccessione $\{x^{(k_j)}\}$ di $\{x^{(k)}\}$, $x^{(k_j)} \rightarrow \bar{x}$ il che, nuovamente, contraddice (b).

(c) \implies (a): assumiamo per assurdo che valga (c) e non valga (a), ossia che esista un ricoprimento aperto $\{E_j\}_{j \in A}$ di K da cui non si possa estrarre alcun sottoricoprimento finito di K . Poiché K è limitato esiste un ‘cubo’³¹ (chiuso) $Q_0 := \{x \in \mathbb{R}^n \mid |x|_\infty \leq R\}$ che contiene K al suo interno. Suddividiamo il cubo Q_0 in 2^n cubi uguali di lato R . Almeno uno di questi cubi, la cui chiusura verrà denotata Q_1 , è tale che $Q_1 \cap K$ è non vuoto e non può essere ricoperto da alcun sottoricoprimento finito di $\{E_j\}$. Iterando tale costruzione otteniamo una famiglia di cubi chiusi $Q_0 \supseteq Q_1 \supseteq Q_2 \supseteq \dots$, con lato di $Q_i := R \cdot 2^{-i}$, e tali che $Q_i \cap K$ è non vuoto e non è contenuto in alcun sottoricoprimento finito di $\{E_j\}$. Se per ogni i scegliamo un punto $x^{(i)} \in Q_i \cap K$, allora la successione $\{x^{(i)}\}$ è di Cauchy³² e per la completezza di \mathbb{R}^n esiste \bar{x} tale che $x^{(i)} \rightarrow \bar{x}$. Poiché K è chiuso si ha che $\bar{x} \in K$. Quindi esiste un j_0 tale che $\bar{x} \in E_{j_0}$. Ed essendo E_{j_0} aperto ne segue che esiste una sfera di centro \bar{x} contenuta in E_{j_0} . Ma allora esiste N tale che $Q_N \subseteq E_{j_0}$ e quindi anche $Q_N \cap K \subseteq E_{j_0}$ e questo contraddice il fatto che, per ogni i , $Q_i \cap K$ non è contenuto in alcun sottoricoprimento finito di $\{E_j\}$. ■

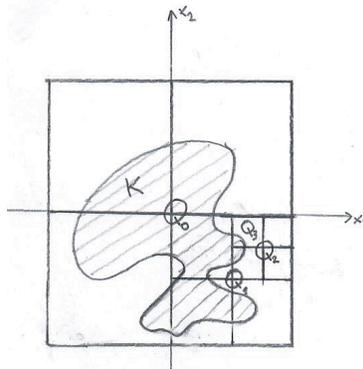


Figura 1.6: (c) \implies (a)

Osservazione 1.29 La nozione di compattezza (fondamentale in topologia, geometria e analisi) si estende a spazi più generali di \mathbb{R}^n , ma *non* si estende il Teorema 1.28; per maggiori informazioni, si veda il Complemento 1.7 a fine capitolo.

³¹Un **cubo di lato** $2r$ in \mathbb{R}^n è un insieme della forma $\{x \in \mathbb{R}^n \mid |x - \bar{x}|_\infty < r\}$ per qualche $\bar{x} \in \mathbb{R}^n$ e qualche $r > 0$; per $n = 1$ un ‘cubo’ è un intervallo aperto di centro \bar{x} e lunghezza $2r$, per $n = 2$ un ‘cubo’ è un quadrato di centro \bar{x} e lato $2r$, etc.

³²Si noti che, per (1.18), se $x^{(j)}, x^{(i)} \in Q_i$ e $j > i$, allora $|x^{(i)} - x^{(j)}| \leq \sqrt{n}|x^{(i)} - x^{(j)}|_\infty \leq \sqrt{n}R2^{-i}$.

La compattezza è un ‘invariante topologico’:

Proposizione 1.30 *Se E è compatto e $f \in C(E, \mathbb{R}^m)$, allora $f(E)$ è compatto.*

Dimostrazione (iii) Se $\{y^{(k)}\}$ una successione in $f(E)$, esistono $x^{(k)} \in E$ tali che $f(x^{(k)}) = y^{(k)}$. Per la compattezza di E e per il Teorema 1.28, esiste $x^{(k_j)} \rightarrow \bar{x} \in E$. Dal punto (i) segue allora che $y^{(k_j)} = f(x^{(k_j)}) \rightarrow f(\bar{x}) := \bar{y} \in f(E)$ il che (sempre per il Teorema 1.28) implica che $f(E)$ è compatto. ■

Teorema 1.31 (Teorema di Weierstrass) *Sia E un insieme compatto di \mathbb{R}^n e sia $f : E \rightarrow \mathbb{R}$ una funzione continua su E . Allora f assume massimo e minimo in E .*

Dimostrazione Sia $M = \sup_{x \in E} f(x)$; dalla definizione di estremo superiore segue che possiamo scegliere una successione $\{x^{(k)}\}$ in E tale che $f(x^{(k)}) \rightarrow M$. Per il Teorema 1.28 possiamo trovare una sottosuccessione $\{x^{(k_j)}\}$ convergente ad un punto $\bar{x} \in E$. Allora per la continuità di f segue che $M = \lim_{k \rightarrow \infty} f(x^{(k)}) = \lim_{j \rightarrow \infty} f(x^{(k_j)}) = f(\bar{x})$. Per il minimo si ragiona analogamente. ■

Definizione 1.32 *Una funzione $f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ si dice **uniformemente continua** su E se*

$$\forall \varepsilon > 0 \exists \delta > 0 \text{ tale che } |f(x) - f(y)| < \varepsilon, \quad \forall x, y \in E \text{ con } |x - y| < \delta. \quad (1.32)$$

Teorema 1.33 (Teorema di Heine–Cantor) *Sia E un insieme compatto di \mathbb{R}^n e sia $f : E \rightarrow \mathbb{R}^m$ una funzione continua su E . Allora f è uniformemente continua su E .*

Dimostrazione Sia $\varepsilon > 0$. Per ogni $x \in E$, f è continua in x , quindi esiste $\delta(x) > 0$ tale che

$$|f(y) - f(x)| < \frac{\varepsilon}{2}, \quad \forall y \in B_{\delta(x)}(x). \quad (1.33)$$

Chiaramente $E \subseteq \bigcup_{x \in E} B_{\frac{\delta(x)}{2}}(x)$ e, essendo E compatto, esistono $x^{(1)}, \dots, x^{(N)}$ in E tali che

$$E \subseteq \bigcup_{j=1}^N B_{\delta_j}(x^{(j)}), \quad \delta_j := \frac{\delta(x^{(j)})}{2}. \quad (1.34)$$

Sia $\delta := \min_{1 \leq j \leq N} \delta_j$ e siano $y, x \in E$ con $|y - x| < \delta$. Da (1.34), segue che esiste $1 \leq j \leq N$ tale che $x \in B_{\delta_j}(x^{(j)})$ ed, inoltre,

$$|y - x^{(j)}| \leq |y - x| + |x - x^{(j)}| < \delta + \delta_j \leq 2\delta_j \stackrel{(1.34)}{=} \delta(x^{(j)}).$$

Quindi, per (1.33), $|f(y) - f(x^{(j)})| < \varepsilon/2$. D'altra parte, sempre per (1.33), poiché $x \in B_{\delta_j}(x^{(j)})$, $|f(x) - f(x^{(j)})| < \varepsilon/2$. Da queste due relazioni e la disuguaglianza triangolare segue che $|f(x) - f(y)| < \varepsilon$. ■

Discutiamo, ora, alcune conseguenze del teorema di Weierstrass applicato alla sfera unitaria in \mathbb{R}^n

$$S^{n-1} := \{x \in \mathbb{R}^n \mid |x| = 1\} \quad (1.35)$$

che, ovviamente, è un sottoinsieme compatto di \mathbb{R}^n .

Funzioni omogenee

Definizione 1.34 Una funzione $f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$ si dice (**positivamente**) **omogenea di grado** $\alpha \in \mathbb{R}$ se

$$f(tx) = t^\alpha f(x), \quad \forall t > 0, \quad \forall x \in \mathbb{R}^n \setminus \{0\}. \quad (1.36)$$

Proposizione 1.35 Sia $f \in C(\mathbb{R}^n \setminus \{0\}, \mathbb{R})$ una funzione omogenea. Se $\alpha \leq 0$, f non può essere estesa ad una funzione continua su tutto \mathbb{R}^n a meno che $\alpha = 0$ e f sia identicamente costante. Se $\alpha > 0$, f si può estendere ad una funzione continua su \mathbb{R}^n ponendo $f(0) := 0$.

Dimostrazione Se $\alpha < 0$ e se $x \in \mathbb{R}^n \setminus \{0\}$, si ha $\lim_{t \rightarrow 0} |f(tx)| = \lim_{t \rightarrow 0} |t|^\alpha |f(x)| = +\infty$ e quindi f non è estendibile ad una funzione continua.

La funzione $f(x) := c$ è omogenea di grado 0 ed è continua. Se f è omogenea di grado 0 e non è identicamente costante, significa che esistono due punti distinti $\mathbb{R}^n \setminus \{0\}$, x_0, y_0 , tali che $f(x_0) \neq f(y_0)$. Allora, le due successioni $x^{(k)} := \frac{1}{k}x_0$ e $y^{(k)} := \frac{1}{k}y_0$ tendono a zero, e per l'omogeneità, $f(x^{(k)}) = f(x_0)$ e $f(y^{(k)}) = f(y_0)$. Quindi, $\lim_{k \rightarrow \infty} f(x^{(k)}) = f(x_0) \neq f(y_0) = \lim_{k \rightarrow \infty} f(y^{(k)})$, il che mostra (teorema ponte) che f non è estendibile con continuità in 0. Infine consideriamo il caso $\alpha > 0$. Chiaramente, $f \in C(S^{n-1}, \mathbb{R})$ e, dunque, per il teorema di Weierstrass, f assume massimo M su S^{n-1} . Se estendiamo f a \mathbb{R}^n ponendo $f(0) := 0$, si ha che $f \in C(\mathbb{R}^n)$, infatti, se $x \neq 0$,

$$|f(x)| = \left| f\left(|x| \frac{x}{|x|}\right) \right| = |x|^\alpha \left| f\left(\frac{x}{|x|}\right) \right| \leq M|x|^\alpha \rightarrow 0 \quad \text{per } |x| \rightarrow 0. \quad \blacksquare$$

Equivalenza delle norme su \mathbb{R}^n

Proposizione 1.36 In \mathbb{R}^n tutte le norme sono equivalenti³³.

Dimostrazione Sia $\|\cdot\|$ una qualunque norma su \mathbb{R}^n e sia $|\cdot|$ la norma euclidea. Facciamo vedere che $\|\cdot\|$ è equivalente a³⁴ $|\cdot|$.

Sia $x = \sum_{i=1}^n x_i e^{(i)}$ ($\{e^{(i)}\}$ base standard di \mathbb{R}^n) un punto arbitrario di \mathbb{R}^n . Dalla disuguaglianza triangolare, dall'omogeneità della norma $\|\cdot\|$ e dalla disuguaglianza di Schwarz segue che

$$\|x\| = \left\| \sum_{i=1}^n x_i e^{(i)} \right\| \leq \sum_{i=1}^n |x_i| \|e^{(i)}\| \leq \sqrt{\sum_{i=1}^n |x_i|^2} \sqrt{\sum_{i=1}^n \|e^{(i)}\|^2} := |x| c, \quad (1.37)$$

dove $c > 0$ è la norma euclidea del vettore $(\|e^{(1)}\|, \dots, \|e^{(n)}\|)$; il fatto che $c > 0$ segue dalla non degenerazione di $\|\cdot\|$. Dunque la funzione $f(x) := \|x\|$ è una funzione continua (nella topologia indotta dalla norma euclidea) poiché se $|x^{(k)} - x| \rightarrow 0$, da (1.37) e dalla disuguaglianza triangolare, segue che

$$|f(x^{(k)}) - f(x)| := \left| \|x^{(k)}\| - \|x\| \right| \leq \|x^{(k)} - x\| \rightarrow 0.$$

Dunque, dal teorema di Weierstrass, segue che la funzione f assume un minimo su $S^{n-1} := \{x \in \mathbb{R}^n : |x| = 1\}$ e tale minimo (per la non degenerazione di $\|\cdot\|$) è strettamente positivo. Quindi esiste un $x_0 \in S^{n-1}$ tale che

$$\|x\| \geq \|x_0\| := b > 0, \quad \forall x \in S^{n-1}. \quad (1.38)$$

Ed allora, per ogni $x \neq 0$

$$\left\| \frac{x}{|x|} \right\| \geq b \iff \|x\| \geq b|x|.$$

Tale relazione, assieme a (1.37), implica l'asserto. \blacksquare

³³Si ricordino le definizioni 1.8 e 1.10.

³⁴Naturalmente, 'essere equivalenti' è una relazione di equivalenza tra norme e quindi se ogni norma è equivalente alla norma euclidea, per transitività, tutte le norme sono equivalenti tra loro.

6 Connessione

Mentre, come abbiamo appena visto, la compattezza non presenta differenze sostanziali con il caso unidimensionale, la connessione in \mathbb{R}^n con $n \geq 2$ è significativamente più complicata rispetto al caso $n = 1$, dove gli insiemi connessi sono semplicemente gli intervalli.

Definizione 1.37 Un sottoinsieme E di \mathbb{R}^n si dice **sconnesso** (o ‘non connesso’) se esistono due aperti di $A, B \subseteq \mathbb{R}^n$, tali che

$$A \cap B = \emptyset, \quad A \cap E \neq \emptyset \neq B \cap E, \quad E \subseteq A \cup B; \quad (1.39)$$

in tal caso, diremo che A e B sconnettono E . Un insieme è **connesso** se non è sconnesso.

Un esempio di insieme sconnesso di \mathbb{R}^n è³⁵ \mathbb{Q}^n .

Osservazione 1.38 (i) In altri termini, un insieme $E \subseteq \mathbb{R}^n$ è connesso se non è unione disgiunta di due aperti non vuoti nella topologia relativa³⁶.

(ii) Dalla definizione segue anche immediatamente che: E è connesso se e solo se i soli sottoinsiemi di E simultaneamente aperti e chiusi (nella topologia relativa) sono \emptyset e E .

(iii) Se E è connesso, allora anche \overline{E} è connesso³⁷.

Infatti, siano A e B due aperti che sconnettono \overline{E} , sia $x \in A \cap \overline{E}$ e sia $B_r(x) \subseteq A$. Dalla Proposizione 1.20 segue che esiste $\bar{x} \in E \cap B_r(x) \subseteq E \cap A$ e quindi $A \cap E \neq \emptyset$. Analogamente, si ha anche che $B \cap E \neq \emptyset$; ma allora, $E \subseteq \overline{E} \subseteq A \cup B$ è sconnesso da A e B . ■

Esempio 1.39 (i) \mathbb{Q}^n è un insieme sconnesso di \mathbb{R}^n : ad esempio, $A = \{x \in \mathbb{R}^n : x_1 < \sqrt{2}\}$ e $B := \{x \in \mathbb{R}^n : x_1 > \sqrt{2}\}$ disconnettono \mathbb{Q}^n .

(ii) Siano $E_1 := \{(x, y) \in \mathbb{R}^2 \mid 0 \leq x \leq \pi, \text{ sen } x \leq y \leq 2 \text{ sen } x\}$, $E_2 := \{(x, y) \in \mathbb{R}^2 \mid \pi \leq x \leq 2\pi, 2 \text{ sen } x \leq y \leq \text{sen } x\}$, $E := E_1 \cup E_2$. Allora, E è connesso (Es 1.20) e \dot{E} è sconnesso (da $A := \{(x, y) \mid x < \pi\}$ e $B := \{(x, y) \mid x > \pi\}$).

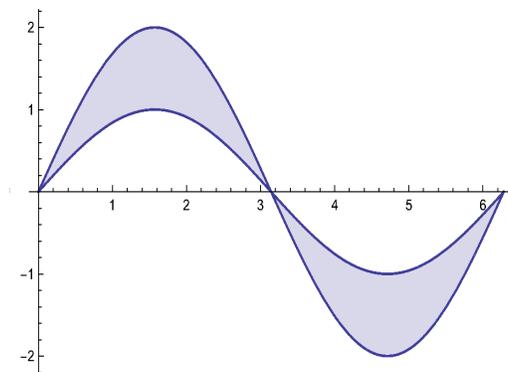


Figura 1.7: Esempio 1.39–(ii)

La connessione è un invariante topologico:

Proposizione 1.40 Se $E \subseteq \mathbb{R}^n$ è connesso e $f \in C(E, \mathbb{R}^m)$, allora $f(E)$ è connesso.

³⁵Si prenda $A := \{x \mid x_1 < \sqrt{2}\}$ e $B = \{x \mid x_1 > \sqrt{2}\}$.

³⁶Si ricordi la definizione di topologia relativa data nella Definizione 1.15.

³⁷Il viceversa è falso; cfr. Esempio 1.39–(i).

Dimostrazione Dimostriamo l'affermazione equivalente: ' $f(E)$ sconnesso $\implies E$ sconnesso'. Se $f(E)$ è sconnesso esistono due aperti I e J tali che

$$I \cap f(E) \neq \emptyset \neq J \cap f(E), \quad I \cap J = \emptyset, \quad f(E) \subseteq I \cup J. \quad (1.40)$$

Siano $A := f^{-1}(I)$ e $B := f^{-1}(J)$. Dalla prima relazione in (1.40) segue che A e B sono non vuoti. Poiché $f \in C(E)$, dalla Proposizione 1.24, da (1.30) e da (1.40) segue che:

$$A \cap B = f^{-1}(I) \cap f^{-1}(J) = f^{-1}(I \cap J) = f^{-1}(\emptyset) = \emptyset$$

e

$$E = f^{-1}(f(E)) \subseteq f^{-1}(I \cup J) = f^{-1}(I) \cup f^{-1}(J) = A \cup B,$$

il che vuol dire E è sconnesso. ■

Non è sempre immediato riconoscere che un insieme è connesso (anche perché la definizione data è 'indiretta'). Un criterio diretto per verificare la connessione è basato sulla nozione di 'connessione per curve'.

Definizione 1.41 (i) Sia $E \subseteq \mathbb{R}^n$. Una **curva** in E è un sottoinsieme di E che è immagine di un cammino³⁸, ossia un insieme della forma

$$\Gamma := \{x \in \mathbb{R}^n \mid x = z(t) \text{ per } a \leq t \leq b\}, \quad \text{dove } z \in C([a, b], E). \quad (1.41)$$

I punti $z(a)$ e $z(b)$ si chiamano gli **estremi** della curva Γ .

(ii) Un **segmento** in $E \subseteq \mathbb{R}^n$ di estremi x e y è la curva data da

$$P(x, y) := \{z \in \mathbb{R}^n \mid z = x + t(y - x), t \in [0, 1]\}; \quad (1.42)$$

(iii) una **poligonale** di estremi x e y e 'vertici successivi' $x^{(0)} := x, x^{(1)}, \dots, x^{(k)} := y$ è l'insieme

$$P(x^{(0)}, \dots, x^{(k)}) := \bigcup_{i=1}^k P(x^{(i-1)}, x^{(i)}). \quad (1.43)$$

(iv) Un insieme $E \subseteq \mathbb{R}^n$ si dice **connesso per curve** (rispettivamente, 'per poligonali') se per ogni $x, y \in E$, esiste una curva (rispettivamente, 'una poligonale') di estremi x e y contenuta in E .

(v) Un insieme $E \subseteq \mathbb{R}^n$ si dice **convesso** se presi comunque due punti in E il segmento $P(x, y)$ che li unisce è contenuto in E .

Osservazione 1.42 (i) Dalla Proposizione 1.40 segue che una curva è un insieme connesso.

(ii) Un esempio (banale) di curva in E è data da $\Gamma := \{\bar{x}\}$ con $\bar{x} \in E$ (essendo la funzione $t \in [0, 1] \rightarrow z(t) := \bar{x}$ chiaramente continua).

Un segmento è una poligonale; una poligonale è una curva³⁹.

Una curva 'strana' è il quadrato $[0, 1] \times [0, 1]$; il cammino che la realizza è detta *curva di Peano*⁴⁰.

Proposizione 1.43 (i) Un insieme di \mathbb{R}^n connesso per curve è connesso.

(ii) Un insieme aperto di \mathbb{R}^n connesso è connesso per poligonali.

Per dimostrare la parte (ii) della Proposizione 1.43 avremo bisogno del seguente semplice risultato:

Lemma 1.44 Sia $E \subseteq \mathbb{R}^n$ un insieme aperto non vuoto e sia $x \in E$. L'insieme

$$W_x := \{y \in E \text{ tale che esiste una poligonale in } E \text{ di estremi } x \text{ e } y\}$$

è un insieme aperto.

³⁸Cfr. (vi) Definizione Osservazione 1.23.

³⁹Per una 'galleria' di curve si veda https://en.wikipedia.org/wiki/Gallery_of_curves.

⁴⁰Cfr. https://it.wikipedia.org/wiki/Curva_di_Peano ed Es. 1.30.

Dimostrazione Se $y \in W_x$ esiste una poligonale $P(x, \dots, y)$ contenuta in E di estremi x e y . In particolare $y \in E$. Poiché E è aperto esiste una sfera B di centro y contenuta in E . Ma per ogni $z \in B$ il segmento $P(y, z) \subseteq B$ e quindi la poligonale $P(x, \dots, y, z) := P(x, \dots, y) \cup P(y, z)$ è contenuta in E . Questo significa che $B \subseteq W_x$ e cioè la tesi. ■

Dimostrazione (della Proposizione 1.43) (i) Si supponga, per assurdo, che $E \subseteq \mathbb{R}^n$ sia connesso per curve e sconnesso e mostriamo che si perviene ad una contraddizione. Siano A e B due aperti per cui valga (1.39). Sia $x \in A \cap E$ e $y \in B \cap E$. Per ipotesi esiste una curva $\Gamma := \{z(t) : a \leq t \leq b\}$ contenuta in E con $z(a) = x$ e $z(b) = y$. Ma allora A e B sconnetterebbero Γ che, invece, per Osservazione 1.42–(i), è connessa.

(ii) Dimostriamo l’affermazione equivalente: ‘ E aperto e non connesso per poligonali implica E sconnesso’. Dire che E non è connesso per poligonali è equivalente a dire che esistono x e y in E che non possono essere gli estremi di alcuna poligonale contenuta in E . Siano allora $A := W_x$ e $C := W_y$ gli insiemi dei punti unibili tramite poligonali in E , rispettivamente, a x ed a y : tali insiemi sono, per il Lemma 1.44, sottoinsiemi aperti di E . Chiaramente $A \cap C = \emptyset$. Sia ora $D := E \setminus (A \cup C)$. Se $D = \emptyset$ avremmo l’asserto (essendo A e C insiemi aperti disgiunti e non vuoti e $E = A \cup C$). Se $D \neq \emptyset$, fissiamo $z \in D$. Poiché $D \subseteq E$ ed E è aperto, esiste una sfera V centrata in z e contenuta in E . La sfera V non può contenere punti di A o di C (altrimenti potremmo trovare una poligonale che unisce z con x o con y contravvenendo la definizione di D come l’insieme dei punti che non possono essere uniti tramite poligonali in E con x o con y). Ma questo significa che $V \subseteq D$ e cioè che D è aperto. Ma, allora A e $B := C \cup D$ sconnettono E , poiché $A \cap B = \emptyset$, $x \in A \cap E$, $y \in B \cap E$ e $E = A \cup B$. ■

Esempio 1.45 (i) Un qualunque insieme convesso di \mathbb{R}^n è connesso (essendo connesso per poligonali). In particolare \mathbb{R}^n è connesso e⁴¹ i soli sottoinsiemi di \mathbb{R}^n simultaneamente aperti e chiusi sono \mathbb{R}^n e \emptyset .

(ii) I sottoinsiemi connessi di \mathbb{R} sono gli intervalli. Infatti, gli intervalli, essendo convessi, sono connessi; viceversa, se E non è un intervallo, allora esistono $x < z < y$ con $x, y \in E$ e $z \in E^c$, ma allora $(-\infty, z)$ e $(z, +\infty)$ sconnettono E .

(iii) L’anello $\{x \in \mathbb{R}^2 : 1 \leq |x| < 2\}$ è connesso per poligonali.

(iv) La circonferenza $\{x \in \mathbb{R}^2 : |x| = 1\}$ è un insieme connesso (essendo connesso per curve) ma non connesso per poligonali.

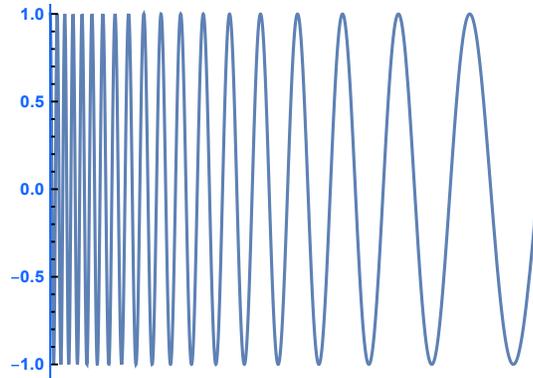


Figura 1.8: *The Topologist’s sine curve*

(v) (*Topologist’s sine curve*: un insieme connesso ma non connesso per curve⁴²)
Siano $E_1 := \{(x, y) \in \mathbb{R}^2 \mid 0 < x \leq \frac{1}{\pi}, y = \sin \frac{1}{x}\}$, $E_0 := \{(x, y) \in \mathbb{R}^2 \mid x = 0, -1 \leq y \leq 1\}$, e $E := E_0 \cup E_1$. È facile vedere che $E = \overline{E_1}$ e poiché E_1 è connesso per curve, è connesso;

⁴¹ Per l’Osservazione 1.38–(ii).

⁴² https://en.wikipedia.org/wiki/Topologist%27s_sine_curve

quindi, dall'Osservazione 1.38–(iii) segue che E è connesso. Però, E non è connesso per curve. Un argomento per dimostrare questa affermazione è, in sintesi, il seguente⁴³.

Innanzitutto, osserviamo che se $(x_0, y_0) \in E_1$, allora $E_1 \setminus \{(x_0, y_0)\}$ è sconnesso dagli aperti $A := \{(x, y) \mid x < x_0\}$ e $B := \{(x, y) \mid x > x_0\}$. Dunque qualunque curva con un estremo in $(\frac{1}{\pi}, 0)$ e l'altro estremo in E_0 deve passare per tutti i punti di E_1 . Se, per assurdo, E fosse connesso per curve, esisterebbe una curva $t \in [0, 1] \mapsto z(t) \in E$ tale che $z(0) = (\frac{1}{\pi}, 0)$ e $z(1) = (0, 0) \in E_0$. Inoltre, per quanto detto prima, si può trovare una successione strettamente crescente $\{t_k\}$ tale che $z(t_k) = (x_k, 1)$ con $x_k := (\frac{\pi}{2} + 2k\pi)^{-1}$, che sono le ascisse consecutive decrescenti in cui la curva E_1 ha ordinata 1. Ora, è facile vedere che $\sup t_k$ deve essere 1, ma questo porta ad una contraddizione poiché si avrebbe $\lim z(t_k) = \lim(x_k, 1) = (0, 1)$ per costruzione, ma per la continuità di $z(t)$ si dovrebbe avere $\lim z(t_k) = z(1) = (0, 0)$. ■

7 Spazi metrici e Lemma delle contrazioni

Come sottolineato in precedenza, molte delle strutture presentate in questo primo capitolo sono una estensione solo formale di strutture già presenti in \mathbb{R} e le dimostrazioni di molte proposizioni sono, essenzialmente, una ripetizione, *mutatis mutandis*, delle dimostrazioni dei risultati analoghi in \mathbb{R} . Ci si può, naturalmente, spingere oltre ed estendere ulteriormente gran parte delle strutture e delle proposizioni discusse qui a 'spazi astratti'.

Bisogna dire che l'astrazione 'fine a sè stessa', in genere, può non essere particolarmente interessante ed è sempre fondamentale che le teorie matematiche siano sviluppate alla luce di esempi non banali. Va anche detto, però, che l'astrazione ha il pregio di mettere in luce cosa è veramente necessario per dimostrare un certo risultato e porta in generale a dimostrazioni più sintetiche e chiare.

Una delle nozioni più importanti della matematica è certamente quella di 'distanza' tra due oggetti. Gli 'spazi metrici' sono quegli insiemi dove è possibile definire una distanza: vediamo brevemente i fondamenti.

Spazi metrici

Definizione 1.46 Dato un qualunque insieme non vuoto X , una funzione $d(\cdot, \cdot) : X \times X \rightarrow [0, \infty)$ che verifichi

- (i) $d(x, y) = d(y, x), \quad \forall x, y \in X,$ (simmetria)
- (ii) $d(x, y) = 0 \iff x = y,$ (non degenerazione)
- (iii) $d(x, y) \leq d(x, z) + d(z, y), \quad \forall x, y, z \in X,$ (disuguaglianza triangolare)

si chiama **distanza** o **metrica** su X .

Una coppia (X, d) con d distanza su $X \neq \emptyset$ prende il nome di **spazio metrico**.

Esempio 1.47 (i) \mathbb{R}^n ammette una metrica naturale indotta dalla norma euclidea; infatti (come è immediato verificare) :

ogni spazio normato $(X, \|\cdot\|)$ è anche uno spazio metrico (X, d) con distanza ('indotta dalla norma') data da⁴⁴ $d(x, y) := \|x - y\|$.

(ii) Il cerchio unitario $S^1 := \{x \in \mathbb{R}^2 \mid |x| = 1\}$ con la distanza ereditata da \mathbb{R}^2 (ossia $d(x, y) := |x - y|$ per ogni $x, y \in S^1$) è uno spazio metrico (ma non uno spazio normato, non essendo S^1 uno spazio vettoriale).

(iii) Sia ℓ^∞ l'insieme delle successioni limitate a valori in \mathbb{R} :

$$\ell^\infty := \{x = \{x_k\} : k \in \mathbb{N} \mapsto x_k \in \mathbb{R} \text{ e } \|x\|_\infty := \sup_k |x_k| < +\infty\}. \quad (1.44)$$

⁴³I dettagli sono lasciati per esercizio (Es 1.21).

⁴⁴In particolare la disuguaglianza triangolare in (iii) segue immediatamente dalla disuguaglianza triangolare per norme: $d(x, y) = \|x - y\| = \|(x - z) + (z - y)\| \leq \|x - z\| + \|z - y\| = d(x, z) + d(z, y)$.

È immediato verificare che $(\ell^\infty, \|\cdot\|_\infty)$ è uno spazio normato⁴⁵ (infinito dimensionale⁴⁶) e, dunque, che (ℓ^∞, d) con $d(x, y) := \|x - y\|_\infty$ è uno spazio metrico.

Gli spazi metrici ammettono una topologia naturale indotta dalle **sfere di centro** $x \in X$ e **raggio** $r > 0$, definite come:

$$B_r(x) := \{y \in X \mid d(y, x) < r\}. \quad (1.45)$$

Definendo un sottoinsieme $A \subseteq X$ **aperto** se⁴⁷:

$$\forall x \in A, \quad \exists r > 0, \text{ tale che } B_r(x) \subseteq A, \quad (1.46)$$

si vede immediatamente che $\mathcal{A} := \{A \subseteq X \mid A \text{ è aperto}\}$ è una topologia nel senso della Definizione 1.12.

Le Definizioni 1.13 e 1.15 si estendono *verbatim* sostituendo \mathbb{R}^n con X .

Anche le nozioni relative a successioni, completezza e continuità si estendono nel modo ovvio:

Una **successione** in X è una funzione $\{x_k\}$ da \mathbb{N} in X . Una successione $\{x_k\}$ in X si dice **convergente**, con limite $x \in X$, se $\forall \varepsilon > 0$ esiste $k_0 \in \mathbb{N}$, tale che $d(x_k, x) < \varepsilon$, per ogni $k \geq k_0$ e scriveremo $\lim x_k = x$.

Una successione $\{x_k\}$ si dice di Cauchy se $\forall \varepsilon > 0$ esiste $k_0 \in \mathbb{N}$, tale che $d(x_h, x_k) < \varepsilon$, per ogni $h, k \geq k_0$.

Un sottoinsieme $E \subseteq X$ si dice **limitato** se esiste una sfera che lo contiene.

Uno spazio metrico (X, d) si dice **completo** se ogni successione di Cauchy è convergente ad un $x \in X$.

Anche le nozioni legate a limiti e continuità si estendono immediatamente. In particolare, dati due spazi metrici (X, d_X) , (Y, d_Y) , una funzione $f : X \rightarrow Y$, si dice **continua in** $x_0 \in X$, se $\forall \varepsilon > 0$ esiste $\delta > 0$ tale che $d_Y(f(x), f(x_0)) < \varepsilon$, per ogni $x \in X$ con $d_X(x, x_0) < \delta$. Una funzione $f : X \rightarrow Y$ si dice **continua** su X se è continua in ogni $x_0 \in X$. Una funzione $f : X \rightarrow Y$ si dice **uniformemente continua** su X se $\forall \varepsilon > 0$ esiste $\delta > 0$ tale che $d_Y(f(x), f(x_0)) < \varepsilon$, per ogni $x, x_0 \in X$ con $d_X(x, x_0) < \delta$.

Osservazione 1.48 (i) Un elenco delle dimostrazioni che si generalizzano, con le loro dimostrazioni, parola-per-parola (sostituendo \mathbb{R}^n con X e $|x - y|$ con $d(x, y)$) sono:

Proposizione 1.14;

Proposizione 1.20;

Teorema ponte (caratterizzazione della continuità tramite successioni, cfr. (1.29));

Proposizione 1.24 (equivalenza di ‘continuità topologica’ e ‘continuità metrica’);

(ii) Una proprietà importante di \mathbb{R}^n che *non si estende* ad arbitrari spazi metrici (completi) è, invece, il Teorema di Bolzano–Weierstrass: non è vero, in arbitrari spazi metrici completi, che una successione limitata ammette sottosuccessioni convergenti. Si consideri, ad esempio, la successione

$$\begin{aligned} \{e^{(j)}\} \text{ con } e^{(j)} = \{e_k^{(j)}\} \in \ell^\infty \text{ definita da } e_k^{(j)} = \delta_{kj}, \text{ ossia :} & \quad (1.47) \\ e^{(1)} = (1, 0, 0, 0, 0, \dots), \quad e^{(2)} = (0, 1, 0, 0, 0, \dots), \quad e^{(3)} = (0, 0, 1, 0, 0, 0, \dots), \dots \end{aligned}$$

Chiaramente, per ogni j , $e^{(j)} \in \ell^\infty$, ossia è una successione limitata: se $0 = (0, 0, 0, \dots) \in \ell^\infty$ denota la successione identicamente nulla, $d(e^{(j)}, 0) = 1$ e quindi $e^{(j)} \in B_2(0)$ per ogni j . Ma se $j \neq i$,

$$d(e^{(j)}, e^{(i)}) = \sup_k |\delta_{jk} - \delta_{ik}| = 1, \quad (i \neq j).$$

Segue che, per qualunque sottosuccessione $\{e^{(j_n)}\}$ si ha, per $n \neq m$, $d(\{e^{(j_n)}\}, \{e^{(j_m)}\}) = 1$, ossia, non è possibile estrarre alcuna sottosuccessione di Cauchy da $\{e^{(j)}\}$ (e quindi non esistono sottosuccessioni convergenti).

⁴⁵Le operazioni di spazio vettoriale sono definite nella maniera ovvia: $\{x_k\} + \{y_k\} := \{x_k + y_k\}$ e $a\{x_k\} := \{ax_k\}$

⁴⁶Cfr. nota 14.

⁴⁷*L'insieme vuoto è un insieme aperto*: non essendovi punti in \emptyset , la negazione di (1.46) è falsa e quindi (1.46) è vera.

Il Lemma delle contrazioni

I cosiddetti ‘teoremi di punto fisso’ sono risultati (generalmente, assai utili) che garantiscono sotto opportune ipotesi, l’esistenza di un ‘punto’ lasciato invariato dall’azione di una trasformazione da un qualche spazio in se stesso⁴⁸.

Uno dei più elementari, ma anche dei più importanti, è il seguente risultato dovuto ai matematici Stefan Banach e Renato Caccioppoli⁴⁹:

Teorema 1.49 (Lemma delle Contrazioni) *Sia $E \neq \emptyset$ un sottoinsieme chiuso di uno spazio metrico completo (X, d) e sia Φ una funzione da E in E tale che, per un qualche $0 < \theta < 1$, valga*

$$d(\Phi(u), \Phi(v)) \leq \theta d(u, v), \quad \forall u, v \in E. \quad (1.48)$$

Allora esiste un unico $\bar{u} \in E$ tale che $\Phi(\bar{u}) = \bar{u}$; inoltre si ha che

$$\bar{u} = \lim_{k \rightarrow \infty} \Phi^k(u_0) := \underbrace{\Phi \circ \dots \circ \Phi}_{k \text{ volte}}(u_0), \quad \forall u_0 \in E. \quad (1.49)$$

Una funzione Φ per cui vale (1.48) si chiama *contrazione*. Si noti che *ogni contrazione è continua* cioè se $d(v_k, v) \rightarrow 0$ allora $d(\Phi(v_k), \Phi(v)) \leq \theta d(v_k, v) \rightarrow 0$.

Osservazione 1.50 (i) L’ipotesi che θ sia strettamente minore di 1 è essenziale: $\Phi : x \in \mathbb{R} \rightarrow \Phi(x) := 1 + x \in \mathbb{R}$, (\mathbb{R} con metrica euclidea) verifica (1.48) con $\theta = 1$ ma ovviamente Φ non ha alcun punto fisso in \mathbb{R} ; d’altra parte, l’identità $\Phi(x) = \text{id}(x)$ (anch’essa verifica (1.48) con $\theta = 1$) ha infiniti punti fissi.

(ii) Anche le altre ipotesi sono, in generale, necessarie:

- (a) Sia $X = \mathbb{R}$, $E := (0, 1)$ e $\Phi : x \in E \mapsto \Phi(x) := x/2$. Φ è una contrazione da E in E (con $\theta = 1/2$), ma, ovviamente, non ha punti fissi. Si noti che E non è chiuso.
- (b) Sia $X = \mathbb{Q}$, $E := \{x \in \mathbb{Q} \mid 5/4 \leq x \leq 3/2\}$ e $\Phi : x \in E \mapsto \Phi(x) := x - x^2/4 + 1/2$. Si vede facilmente che⁵⁰ $\Phi : E \rightarrow E$ e che è una contrazione con⁵¹ $\theta = 3/8$. D’altra parte, $\Phi(x) = x$ equivale a $x^2 = 2$, che non ha soluzioni in X . Qui il problema è che X non è uno spazio metrico completo.

Dimostrazione (del Teorema 1.49) Sia u_0 un punto arbitrario di E e definiamo la seguente successione per ricorrenza:

$$v_0 = u_0, v_1 = \Phi(v_0), \dots, v_k = \Phi(v_{k-1}) := \Phi^k(u_0), \dots \quad (1.50)$$

Poiché Φ manda E in se stesso tale successione è ben definita e $v_k \in E$ per ogni $k \geq 0$. Poiché Φ è una contrazione, si ha:

$$d(v_j, v_{j-1}) = d(\Phi(v_{j-1}), \Phi(v_{j-2})) \leq \theta d(v_{j-1}, v_{j-2}) \leq \theta^{j-1} d(v_1, v_0).$$

Dunque, per ogni $k > h$ si ha

$$\begin{aligned} d(v_k, v_h) &\leq d(v_k, v_{k-1}) + d(v_{k-1}, v_h) \\ &\leq \sum_{j=0}^{k-h-1} d(v_{h+j+1}, v_{h+j}) \leq \sum_{j=0}^{k-h-1} \theta^{h+j} d(v_1, v_0) \\ &\leq \theta^h d(v_1, v_0) \sum_{j=0}^{\infty} \theta^j = \theta^h \frac{d(v_1, v_0)}{1 - \theta}. \end{aligned}$$

⁴⁸Per informazioni, si veda https://en.wikipedia.org/wiki/Fixed-point_theorem.

⁴⁹Cfr. https://en.wikipedia.org/wiki/Banach_fixed-point_theorem.

⁵⁰ $\Phi'(x) > 0$ se $x \in E$, quindi $\Phi(E) = [87/64, 23/16] \subseteq E$.

⁵¹Se $x, y \in E$, $|\Phi(x) - \Phi(y)| = |x - y| \left(1 - \frac{x+y}{4}\right) \leq |x - y| \left(1 - \frac{2 \cdot 5/4}{4}\right) = \frac{3}{8} |x - y|$.

Tale quantità tende a 0 per $h \rightarrow \infty$ e ciò significa che la successione $\{v_k\}$ è di Cauchy. Poiché X è completo v_k converge a $\bar{u} \in X$ e poiché E è chiuso $\bar{u} \in E$. Quindi, per la continuità di Φ si ha che $\bar{u} = \lim v_k = \lim \Phi(v_{k-1}) = \Phi(\bar{u})$. Per dimostrare l'unicità, supponiamo, per assurdo, che $\bar{v} \in E$ sia un altro punto fisso per Φ : $\Phi(\bar{v}) = \bar{v}$. Si ha allora $d(\bar{u}, \bar{v}) = d(\Phi(\bar{u}), \Phi(\bar{v})) \leq \theta d(\bar{u}, \bar{v}) < d(\bar{u}, \bar{v})$ che è una contraddizione. ■

Complementi

Complemento 1.1: Successioni di funzioni

Definizione 1.51 Una successione di funzioni $\{f_k\}$ da $E \subseteq \mathbb{R}^n$ in \mathbb{R}^m è una famiglia di funzioni $f_k : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ con⁵² $k \in \mathbb{N}$.

Se, per ogni $x \in E$, la successione $\{f_k(x)\}$ converge (per $k \rightarrow \infty$) in \mathbb{R}^m , tale limite definisce una funzione $f : E \rightarrow \mathbb{R}^m$ che chiameremo il **limite puntuale della successione** $\{f_k\}$. In tal caso, diremo anche che $\{f_k\}$ **converge ad f puntualmente su E** e scriveremo anche $f_n \rightarrow f$ su E .

Esempio 1.52 La successione data da $f_k : x \in E := [0, 1] \mapsto x^k \in \mathbb{R}$ converge puntualmente a f dove

$$f(x) := \begin{cases} 0, & \text{se } x \in [0, 1), \\ 1, & \text{se } x = 1. \end{cases}$$

Da questo semplicissimo esempio si vede che le proprietà di f_k non vengono, in generale, conservate nel limite: f_k sono polinomi (e quindi funzioni $C^\infty(E)$), ma il limite è una funzione discontinua su E .

Una nozione di convergenza più forte è quella di *convergenza uniforme*:

Definizione 1.53 La successione $\{f_k\}$, $f_k : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$, **converge uniformemente a $f : E \rightarrow \mathbb{R}^m$** se

$$\forall \varepsilon > 0, \quad \exists k_0 \text{ t.c. } |f_k(x) - f(x)| < \varepsilon, \quad \forall x \in E, \forall k \geq k_0. \quad (1.51)$$

Osservazione 1.54 (i) La definizione di converge puntuale di f_n a f è equivalente a dire che

$$\forall x \in E, \forall \varepsilon > 0, \quad \exists k_0 \text{ t.c. } |f_k(x) - f(x)| < \varepsilon, \quad \forall k \geq k_0. \quad (1.52)$$

Si noti la differenza (nell'ordine in cui appaiono i quantificatori logici) con la convergenza uniforme: in (1.52) k_0 dipende da ε e da x mentre in (1.51) *dipende solo da ε* .

(ii) Chiaramente la (1.51) è equivalente a

$$\forall \varepsilon > 0, \quad \exists k_0 \text{ t.c. } \sup_E |f_k - f| < \varepsilon, \quad \forall k \geq k_0, \quad (1.53)$$

il che, a sua volta, è equivalente a $\lim_{k \rightarrow +\infty} \sup_E |f_k - f| = 0$.

(iii) La successione $\{f_k\}$ nell'Esempio 1.52 converge a f puntualmente ma *non* uniformemente (essendo $\sup_E |f_k - f| = 1$ per ogni⁵³ k). D'altra parte, $\{f_k\}$ converge uniformemente a 0 su $E := [0, \delta]$ per ogni $0 \leq \delta < 1$ (poiché, per ogni $x \in E$, $0 \leq x^k \leq \delta^k \rightarrow 0$).

(iv) La successione $\{f_k\}$ con

$$x \in \mathbb{R} \mapsto f_k(x) := \frac{x}{1 + kx^2} \in \mathbb{R}, \quad (1.54)$$

converge uniformemente a $f(x) \equiv 0$ su \mathbb{R} . Infatti, sia $\varphi(t) := t/(1 + t^2)$ e si noti che $|\varphi(t)| \leq 1/2 = \max_{\mathbb{R}} \varphi$. Allora, $|f_k(x)| = \varphi(\sqrt{k}|x|)/\sqrt{k} \leq 1/(2\sqrt{k})$, dal che segue immediatamente che $f_k \rightarrow 0$ uniformemente su \mathbb{R} .

È possibile caratterizzare la convergenza uniforme senza fare riferimento al limite della successione. Vale infatti il seguente semplice

⁵²In questo Complemento, per semplicità di notazione (e se non altrimenti specificato), l'indice k si riferisce al k -esimo elemento della famiglia di funzioni $\{f_k\}$ e *non* denota la k -esima componente di un vettore in \mathbb{R}^m .

⁵³Es 1.31.

Lemma 1.55 *La successione $\{f_k\}$, $f_k : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$, converge uniformemente ad una funzione f su E se e solo se le successioni $\{f_k(x)\}$ sono **uniformemente di Cauchy** su E , ossia, se*

$$\forall \varepsilon > 0, \quad \exists k_0 \text{ t.c. } |f_k(x) - f_h(x)| < \varepsilon, \quad \forall x \in E, \forall k, h \geq k_0. \quad (1.55)$$

Dimostrazione Se $f_k \rightarrow f$ uniformemente su E e $\varepsilon > 0$, vale (1.51) con ε sostituito da $\varepsilon/2$ e quindi, se $h, k \geq k_0$, si ha, per ogni $x \in E$,

$$|f_k(x) - f_h(x)| \leq |f_k(x) - f(x)| + |f_h(x) - f(x)| < 2 \frac{\varepsilon}{2} = \varepsilon,$$

e quindi vale (1.55).

Viceversa, se assumiamo (1.55), si ha che per ogni $x \in E$ le successioni $\{f_k(x)\}$ sono di Cauchy e quindi, per la completezza di \mathbb{R}^m , convergono ad un vettore che chiamiamo $f(x)$. Fissato $\varepsilon > 0$, sia k_0 come in (1.55). Prendendo il limite per $h \rightarrow \infty$ in (1.55), si ottiene $|f_k(x) - f(x)| \leq \varepsilon$, per ogni $x \in E$, il che significa che $f_k \rightarrow f$ uniformemente su E . ■

Convergenza uniforme e continuità

Il limite uniforme di funzioni continue è una funzione continua:

Proposizione 1.56 *Siano $f, f_k \in C(E, \mathbb{R}^m)$, $E \subseteq \mathbb{R}^n$. Se $f_k \rightarrow f$ uniformemente su E , allora $f \in C(E, \mathbb{R}^m)$.*

Dimostrazione Siano $x_0 \in E$, $\varepsilon > 0$ e sia k_0 tale che $|f_k(x) - f(x)| < \varepsilon/3$ per ogni $x \in E$ e per ogni $k \geq k_0$ (l'esistenza di k_0 segue dalla definizione di convergenza uniforme). Sia $\delta > 0$ tale che $|f_{k_0}(x) - f_{k_0}(x_0)| < \varepsilon/3$ per ogni $x \in E$ con $|x - x_0| < \delta$ (tale δ esiste, essendo f_{k_0} continua su E). Allora, se $x \in E$ con $|x - x_0| < \delta$,

$$|f(x) - f(x_0)| \leq |f(x) - f_{k_0}(x)| + |f_{k_0}(x) - f_{k_0}(x_0)| + |f_{k_0}(x_0) - f(x_0)| < \varepsilon. \quad \blacksquare$$

Osservazione 1.57 (i) Da questa dimostrazione segue anche che se le funzioni f_k sono uniformemente continue su E , allora lo è anche f . In tal caso, infatti, il δ nella dimostrazione precedente dipenderà solo da ε (si noti che anche k_0 dipende solo da ε).

(ii) Sia $A \subseteq \partial E$ e $f_k \in C(E \cup A)$. Se $\{f_k\}$ converge uniformemente su E , allora $\{f_k\}$ converge uniformemente su⁵⁴ $E \cup A$. Dunque, Se $f_k \in C(E \cup A)$, $A \subseteq \partial E$, e $\{f_k\}$ converge puntualmente su E , ma per $y \in A$, la successione $\{f_k(y)\}$ non converge, allora la convergenza su E non è uniforme.

(iii) Un esempio di successione di funzioni convergente puntualmente ma non uniformemente è il seguente: Sia $f : x \in E := (0, 1] \mapsto \frac{kx}{1 + kx^2}$. Allora, $f_k(x) \rightarrow 1/x$ puntualmente su E , ma non uniformemente. Infatti, se la convergenza fosse uniforme su E , lo sarebbe anche su $\bar{E} = [0, 1]$ e f sarebbe continua su $[0, 1]$, ma $f_k(0) \rightarrow 0 =: f(0)$: contraddizione ($\lim_{x \rightarrow 0^+} f(x) = +\infty$).

Convergenza uniforme e limitatezza

In generale una successione di funzioni può convergere uniformemente ad una funzione illimitata: ad esempio $\frac{1}{k} + \frac{1}{x}$ converge uniformemente a $\frac{1}{x}$ su $(0, +\infty)$. D'altra parte si ha

Proposizione 1.58 *Sia $E \subseteq \mathbb{R}^n$ e $f_k : E \rightarrow \mathbb{R}^m$ funzioni limitate e tali che $\{f_k\}$ converge uniformemente su E ad una funzione f . Allora, f è limitata.*

Dimostrazione Sia $M_k := \sup_E |f_k|$. Per ipotesi $M_k < +\infty$. Sia k_0 tale che $\sup_E |f_h - f_k| \leq 1$ per ogni $h, k \geq k_0$ (tale k_0 esiste per l'ipotesi sulla convergenza uniforme di $\{f_k\}$) e si noti che (prendendo il limite per $h \rightarrow +\infty$) $\sup_E |f - f_{k_0}| \leq 1$. Dunque:

$$\sup_E |f| = \sup_E |f - f_{k_0} + f_{k_0}| \leq \sup_E |f - f_{k_0}| + \sup_E |f_{k_0}| \leq 1 + M_{k_0} < +\infty. \quad \blacksquare$$

⁵⁴Infatti, $\sup_{E \cup A} |f_k - f_h| = \sup_E |f_k - f_h|$.

Convergenza uniforme e integrazione

L'uniforme convergenza gioca un ruolo fondamentale anche nell'integrare o derivare 'termine a termine' una successione⁵⁵.

Proposizione 1.59 *Sia E un intervallo limitato di \mathbb{R} e $f_k \in C(E, \mathbb{R})$ funzioni limitate su E . Se la successione $\{f_k\}$ converge uniformemente ad f allora*

$$\lim_{k \rightarrow \infty} \int_E |f_k(x) - f(x)| dx = 0, \quad \lim_{k \rightarrow \infty} \int_E f_k(x) dx = \int_E f(x) dx. \quad (1.56)$$

Dimostrazione Si noti che dalle Proposizioni 1.56 e 1.58 segue che f è una funzione continua e limitata su E (e quindi integrabile secondo Riemann su E). Sia $\varepsilon > 0$. Allora, esiste $k_0 > 0$ tale che, per ogni $k \geq k_0$,

$$\sup_E |f_k(x) - f(x)| < \frac{\varepsilon}{\text{mis } E}.$$

Da tale relazione segue che

$$\int_E |f_k(x) - f(x)| dx \leq \int_E \frac{\varepsilon}{\text{mis } E} dx = \varepsilon,$$

il che dimostra la prima relazione in (1.56). La seconda relazione in (1.56) è conseguenza del fatto che

$$\left| \int_E f_k(x) dx - \int_E f(x) dx \right| \leq \int_E |f_k(x) - f(x)| dx. \quad \blacksquare$$

Anche in questo caso l'uniforme convergenza è un'ipotesi essenziale, come segue dal seguente

Esempio 1.60 Sia $E = [0, 1]$ e sia $\varphi(x)$ la funzione 'a tenda'

$$\varphi(x) := \begin{cases} x, & \text{se } 0 \leq x \leq 1, \\ 2 - x, & \text{se } 1 < x \leq 2, \\ 0, & \text{altrimenti.} \end{cases} \quad (1.57)$$

Si noti che $\varphi \in C(\mathbb{R})$, $\varphi \geq 0$, $\text{supp}(\varphi) = [0, 2]$ e $\int_0^2 \varphi = 1$. Sia ora $f_k(x) := 2k \cdot \varphi(2kx)$. È facile vedere che: $f_k \in C(\mathbb{R})$, $f_k(x) \rightarrow 0$ per ogni $x \in \mathbb{R}$, ma che $\int_0^1 f_k = 1$ per ogni k . Quindi, in tal caso non valgono le relazioni in (1.56). Infatti, f_k non converge uniformemente a 0, essendo $\sup_{k \geq k_0} \|f_k\|_\infty = \infty$, per ogni k_0 .

Convergenza uniforme e derivazione

Vediamo ora un criterio (che sarà di nuovo basato sulla uniforme convergenza) che garantisca che se una successione di funzioni differenziabili f_k converge a f , allora $f' = \lim f'_k$.

Proposizione 1.61 *Sia E un intervallo aperto non vuoto di \mathbb{R} e siano $f_k : E \rightarrow \mathbb{R}$ funzioni differenziabili su E . Assumiamo che $\{f'_k\}$ converga uniformemente su E e che esista $x_0 \in E$ tale che $\lim_{k \rightarrow \infty} f_k(x_0) \in \mathbb{R}$. Allora, $\{f_k\}$ converge puntualmente a una $f \in C^1(E)$ e $\{f'_k\}$ converge uniformemente a f' su E . Inoltre, $f_k \rightarrow f$ uniformemente su un qualunque intervallo $[a, b] \subseteq E$.*

Dimostrazione Dalla Proposizione 1.56 segue che esiste $g \in C(E)$ tale che $\{f'_k\}$ converge uniformemente a g in E . Dalla Proposizione 1.59, dal Teorema Fondamentale del Calcolo⁵⁶ e dalla ipotesi su $\{f_k(x_0)\}$ segue che, per ogni $x \in E$,

$$\lim_{k \rightarrow \infty} f_k(x) = \lim_{k \rightarrow \infty} \left(f_k(x_0) + \int_{x_0}^x f'_k(t) dt \right) = \alpha + \int_{x_0}^x g(t) dt =: f(x), \quad (1.58)$$

⁵⁵Non avendo ancora sviluppato il calcolo integrale in \mathbb{R}^n per $n \geq 2$, ci limitiamo, qui, al caso, $n = 1$ (anche se enunciati e dimostrazioni si generalizzano facilmente al caso $n \geq 2$).

⁵⁶Cfr. [C2019], §8.2.

dove $\alpha := \lim f_k(x_0)$. Da (1.58) (ancora per il Teorema Fondamentale del Calcolo) segue che $f \in C^1(E)$ e che $f' = g$, il che dimostra la prima parte della tesi.

Infine, sia $[a, b] \subseteq E$ un intervallo compatto. Dato $\varepsilon > 0$, esiste k_0 tale che, per ogni $k \geq k_0$,

$$|f_k(a) - f(a)| < \frac{\varepsilon}{2}, \quad \sup_{x \in [a, b]} |f'_k(x) - f'(x)| < \frac{\varepsilon}{2(b-a)}.$$

Dal Teorema Fondamentale del Calcolo, segue che, per ogni $x \in [a, b]$ e per ogni $k \geq k_0$,

$$\begin{aligned} |f_k(x) - f(x)| &= \left| f_k(a) - f(a) + \int_a^x (f'_k(y) - f'(y)) dy \right| \\ &\leq |f_k(a) - f(a)| + \int_a^x |f'_k(y) - f'(y)| dy \\ &< \frac{\varepsilon}{2} + (b-a) \frac{\varepsilon}{2(b-a)} = \varepsilon. \quad \blacksquare \end{aligned}$$

Esempio 1.62 (i) Sia $f_k(x) = k + x$. Ovviamente f'_k converge uniformemente su \mathbb{R} ma f_k non converge per nessun x .

(ii) Sia $f_k(x) := \frac{\sin kx}{k}$. Allora, $f_k \rightarrow 0$ uniformemente su \mathbb{R} (essendo $|f_k(x)| \leq 1/k$, per ogni x), ma $f'_k(x) = \cos(kx)$ in genere non converge puntualmente (per esempio, è facile vedere che non converge su tutti i punti della forma $r\pi$ con $r \in \mathbb{Q}_+ \setminus 2\mathbb{Z}$).

(iii) Se f_k è come nell'Osservazione 1.54-(iv), si ha che $f_k \rightarrow f \equiv 0$ uniformemente su \mathbb{R} e la successione delle derivate

$$f'_k(x) = \frac{1 - kx^2}{(1 + kx^2)^2}$$

converge a 0 uniformemente su $\{x \mid |x| > \varepsilon\}$, per ogni $\varepsilon > 0$ e, in particolare $f'_k(x) \rightarrow f'(x) = 0$ per ogni $x \neq 0$; ma, per ogni k , $f'_k(0) = 1 \neq f'(0) = 0$.

Convergenza uniforme e monotonia

Proposizione 1.63 (Lemma di Dini) Sia $K \subseteq \mathbb{R}^n$ un insieme compatto, siano f_n (per $n \geq 0$) e f funzioni continue su K e tali che⁵⁷ $f_n \uparrow f$ [oppure $f_n \downarrow f$]. Allora $f_n \rightarrow f$ uniformemente su K .

Dimostrazione Per assurdo: la negazione di ' f_n converge a f uniformemente' è:

$$\exists \varepsilon > 0 : \forall n \geq 0 \exists x_n \text{ tale che } |f(x_n) - f_n(x_n)| \geq \varepsilon. \quad (1.59)$$

Poiché K è compatto esiste $x_{n_k} \rightarrow \bar{x} \in K$. Dalla convergenza puntuale di f_n segue che

$$\exists N : |f(\bar{x}) - f_n(\bar{x})| \leq \frac{\varepsilon}{8} \quad \forall n \geq N. \quad (1.60)$$

Per l'uniforme continuità di f e f_N segue che esiste $\delta > 0$ tale che

$$|f(x) - f(y)| \leq \frac{\varepsilon}{8} \quad \text{e} \quad |f_N(x) - f_N(y)| \leq \frac{\varepsilon}{8} \quad \forall x, y \in K \text{ con } |x - y| \leq \delta. \quad (1.61)$$

Dalla convergenza monotona di f_n a f , da (1.60) e (1.61) segue che, per ogni $x, y \in K$ con $|x - y| \leq \delta$ e per ogni $n \geq N$,

$$f_n(x) - f_n(y) \leq f(x) - f_N(y) \leq |f(x) - f_N(x)| + |f_N(x) - f_N(y)| \leq \frac{\varepsilon}{4}.$$

Analogamente

$$f_n(x) - f_n(y) \geq f_N(x) - f(y) \geq -|f_N(x) - f_N(y)| - |f_N(y) - f(y)| \geq -\frac{\varepsilon}{4},$$

⁵⁷ Si ricorda che una successione $\{a_n\}$ si dice monotona crescente [decrecente] se $\forall n, a_{n+1} \geq a_n$ [$a_{n+1} \leq a_n$]; $\{a_n\}$ si dice strettamente monotona crescente [decrecente] se $\forall n, a_{n+1} > a_n$ [$a_{n+1} < a_n$]; $a_n \uparrow a$ significa che la successione $\{a_n\}$ è monotona crescente e che $\lim a_n = a$; $a_n \downarrow a$ significa che la successione $\{a_n\}$ è monotona decrescente e che $\lim a_n = a$; $f_n \uparrow f$ significa $f_n(x) \uparrow f(x)$ per ogni x ed analogamente per $f_n \downarrow f$; una funzione $f : E \rightarrow \mathbb{R}$ si dice monotona crescente [decrecente] se $\forall x < y$ in $E, f(x) \leq f(y)$ [$f(x) \geq f(y)$]; una funzione $f : E \rightarrow \mathbb{R}$ si dice monotona strettamente crescente [decrecente] se $\forall x < y$ in $E, f(x) < f(y)$ [$f(x) > f(y)$].

e quindi

$$|f_n(x) - f_n(y)| \leq \frac{\varepsilon}{4}, \quad \forall x, y \in K \text{ con } |x - y| \leq \delta, \quad \forall n \geq N. \quad (1.62)$$

Sia ora k_0 tale che $\bar{n} := n_{k_0} \geq N$ e $|\bar{x} - x_{\bar{n}}| \leq \delta$. Allora, da (1.61), (1.60) ed (1.62) segue che

$$|f(x_{\bar{n}}) - f_{\bar{n}}(x_{\bar{n}})| \leq |f(x_{\bar{n}}) - f(\bar{x})| + |f(\bar{x}) - f_{\bar{n}}(\bar{x})| + |f_{\bar{n}}(\bar{x}) - f_{\bar{n}}(x_{\bar{n}})| \leq \frac{\varepsilon}{2},$$

relazione che contraddice (1.59).

Per il caso $f_n \downarrow f$ basta applicare quanto dimostrato a $-f_n$ e $-f$. ■

Proposizione 1.64 *Sia $K \subseteq \mathbb{R}$ un insieme compatto, siano f_n (per $n \geq 0$) e f funzioni continue su K e tali che $f_n(x) \rightarrow f(x)$ per ogni $x \in K$. Se le f_n sono funzioni monotone crescenti per ogni n [o monotone decrescenti per ogni n] allora $f_n \rightarrow f$ uniformemente su K .*

Dimostrazione Poiché $f \in C(K)$, f è uniformemente continua su K e dato $\varepsilon > 0$ esiste $\delta > 0$ tale che $|f(x) - f(y)| \leq \varepsilon$ se $x, y \in K$ e $|x - y| \leq \delta$. Essendo K compatto è possibile trovare insiemi a due a due disgiunti K_1, \dots, K_M tali che $K_1 \cup \dots \cup K_M = K$ e $\text{diam } K_i := \sup_{x, y \in K_i} |x - y| \leq \delta$; chiamiamo $a_i := \inf K_i$, $b_i := \sup K_i$ (poiché $K_i \subseteq K$ e K è chiuso, $a_i, b_i \in K$ per ogni i). Poiché $f_n \rightarrow f$ puntualmente, esiste N tale che per ogni $1 \leq i \leq M$ e per ogni $n \geq N$ si ha

$$|f_n(a_i) - f(a_i)| \leq \varepsilon, \quad |f_n(b_i) - f(b_i)| \leq \varepsilon.$$

Si osservi che dalla monotonia delle f_n segue la monotonia di f . Sia $x \in K$ e sia j tale che $x \in K_j$ e per ogni $n \geq N$ si ha

$$f(x) - f_n(x) \leq f(b_j) - f_n(a_j) \leq |f(b_j) - f_n(b_j)| + |f_n(b_j) - f_n(a_j)| \leq 2\varepsilon,$$

ed analogamente

$$f(x) - f_n(x) \geq f(a_j) - f_n(b_j) \geq -|f(a_j) - f_n(a_j)| - |f_n(a_j) - f_n(b_j)| \geq -2\varepsilon,$$

ossia $|f(x) - f_n(x)| \leq 2\varepsilon$ per ogni $n \geq N$ e per ogni $x \in K$ il che prova l'asserto. Per il caso in cui le f_n sono monotone decrescenti basta applicare quanto dimostrato a $-f_n$ e $-f$. ■

Complemento 1.2: Equicontinuità e Teorema di Ascoli–Arzelà

Definizione 1.65 *Una famiglia di funzioni $\{f_\alpha \mid \alpha \in A\}$ (dove $A \neq \emptyset$ è un insieme arbitrario) $f_\alpha : E \subseteq \mathbb{R} \rightarrow \mathbb{R}$ si dice **equicontinua** (su E) se $\forall \varepsilon > 0 \exists \delta > 0$ tale che $|f_\alpha(x) - f_\alpha(y)| < \varepsilon$, $\forall \alpha \in A$ e $\forall x, y \in E$ con $|x - y| < \delta$.*

Alcune proprietà elementari di successioni equicontinue sono raccolte nella seguente

Proposizione 1.66 *Sia $\{f_n\}$ una successione equicontinua su $E \subseteq \mathbb{R}$.*

- (i) *Se $f_n(x)$ converge $\forall x \in D$, con D denso in E , allora $f_n(x)$ converge $\forall x \in E$;*
- (ii) *se E è compatto e $f_n(x) \rightarrow f(x) \forall x \in E$ allora f_n converge uniformemente a f in E ;*
- (iii) *se E è compatto e $\sup_n \sup |f_n(x)| < \infty$ per ogni $x \in E$, allora $\sup_n \sup_E |f_n(x)| < \infty$.*

Dimostrazione (i): Sia $x_0 \in E$ e $\varepsilon > 0$. Dalla definizione di equicontinuità segue che

$$\exists \delta > 0 : |f_n(x) - f_n(y)| < \frac{\varepsilon}{3} \quad \forall n \quad \forall x, y \in E \text{ con } |x - y| < \delta. \quad (1.63)$$

Poiché D è denso in E esiste $\bar{x} \in D$ tale che $|\bar{x} - x_0| < \delta$. Sia $N \geq 1$ tale che $|f_n(\bar{x}) - f_m(\bar{x})| < \varepsilon/3$ per ogni $n, m \geq N$ (criterio di Cauchy per la successione convergente $\{f_n(\bar{x})\}$). Allora, per ogni $n, m \geq N$,

$$|f_n(x_0) - f_m(x_0)| \leq |f_n(x_0) - f_n(\bar{x})| + |f_n(\bar{x}) - f_m(\bar{x})| + |f_m(\bar{x}) - f_m(x_0)| < \varepsilon,$$

il che significa che $\{f_n(x_0)\}$ è una successione di Cauchy e quindi convergente.

(ii): Dato $\varepsilon > 0$ sia δ come in (1.63) e si osservi che prendendo il limite per $n \rightarrow \infty$ si ha anche che $|f(x) - f(y)| \leq \varepsilon/3$. Poiché E è compatto si possono trovare k intervalli aperti B_i di centro $x^{(i)} \in E$

di lunghezza 2δ tali che $E \subseteq B_1 \cup \dots \cup B_k$. Poiché f_n converge puntualmente su E , esiste N tale che $|f_n(x^{(i)}) - f(x^{(i)})| < \varepsilon/3$ per ogni $n \geq N$ e per ogni $1 \leq i \leq k$. Dato un qualunque punto $x \in E$ esiste un $1 \leq j \leq k$ tale che $x \in B_j$ e quindi $|x - x^{(j)}| < \delta$. Dunque, per ogni $n \geq N$ si ha

$$|f_n(x) - f(x)| \leq |f_n(x) - f_n(x^{(j)})| + |f_n(x^{(j)}) - f(x^{(j)})| + |f(x^{(j)}) - f(x)| < \varepsilon.$$

(iii): Dato ε siano δ, B_1, \dots, B_k come qui sopra. Sia $M := \sup_{1 \leq i \leq k} \sup_n |f_n(x^{(i)})|$. Sia x un qualunque punto in E e sia (come sopra) j tale che $|x - x^{(j)}| < \delta$. Allora per ogni n

$$|f_n(x)| \leq |f_n(x) - f_n(x^{(j)})| + |f_n(x^{(j)})| \leq \frac{\varepsilon}{3} + M,$$

da cui segue anche (iii). ■

Teorema 1.67 (Ascoli, Arzelà) *Sia $E \subseteq \mathbb{R}$ un insieme compatto e sia $\{f_n\}$ una successione equicontinua di funzioni su E tali che $\sup_n |f_n(x)| < \infty$ per ogni $x \in E$. Allora esiste una successione $\{n_k\}$ di interi tali che f_{n_k} converge uniformemente su E .*

Lemma 1.68 (Trucco diagonale di Cantor) *Siano $a_n^{(j)}$ numeri reali (con $n, j \in \mathbb{N}$) tali che, per ogni j , $\sup_n |a_n^{(j)}| < \infty$. Allora esiste una successione di interi n_k tale che le successioni $\{a_{n_k}^{(j)}\}_{k \geq 0}$ convergono per ogni j .*

Dimostrazione Poiché $\{a_n^{(1)}\}$ è limitata esiste una successione $\{n_k^{(1)}\}$ tale che $a_{n_k^{(1)}}^{(1)}$ converge; analogamente essendo $\{a_{n_k^{(1)}}^{(2)}\}$ limitata esiste una sottosuccessione $\{n_k^{(2)}\}$ di $\{n_k^{(1)}\}$ tale che $a_{n_k^{(2)}}^{(2)}$ converge; proseguendo indefinitivamente in tal modo si ottengono successioni di interi $\{n_k^{(i)}\}$ tali che: (a) $\{n_k^{(i+1)}\}_{k \geq 0}$ è una sottosuccessione di $\{n_k^{(i)}\}_{k \geq 0}$; (b) $a_{n_k^{(i)}}^{(j)}$ converge (per $k \rightarrow \infty$) per ogni $1 \leq j \leq i$. Definendo $n_k := n_k^{(k)}$ si ha l'asserto (infatti, per ogni j , $\{n_j, n_{j+1}, \dots\}$ è una sottosuccessione di $\{n_k^{(j)}\}$). ■

Dimostrazione (del Teorema di Ascoli–Arzelà) Sia $D := \{r_j : j \in \mathbb{N}\}$ una numerazione dei razionali in E (cosicché D è denso in E). Dal punto (iii) della Proposizione 1.66 segue che esiste M tale che $|f_n(x)| \leq M$ per ogni n e per ogni $x \in E$. Applicando il Lemma 1.68 con $a_n^{(j)} := f_n(r_j)$ si ha che esiste una successione $\{n_k\}$ tale che $f_{n_k}(r_j)$ converge (per $k \rightarrow \infty$) per ogni $r_j \in D$. Dal punto (i) della Proposizione segue che $f_{n_k}(x)$ converge per ogni $x \in E$ e per il punto (ii) della Proposizione la convergenza è uniforme in E . ■

Complemento 1.3: Integrazione di funzioni vettoriali di variabile reale

Discutiamo brevemente come integrare una funzione vettoriale continua $f : [a, b] \rightarrow \mathbb{R}^n$, con $n \geq 2$.

Sia $t \in [a, b] \rightarrow f := (f_1(t), \dots, f_n(t)) \in \mathbb{R}^n$ una funzione continua (il che significa che le n funzioni scalari $f_i(t)$ sono funzioni continue sull'intervallo chiuso $[a, b]$) e definiamo

$$\int_a^b f(t) dt := \left(\int_a^b f_1(t) dt, \dots, \int_a^b f_n(t) dt \right). \tag{1.64}$$

È evidente che da tale definizione segue che l'integrale di funzioni vettoriali continue di una variabile reale è una operazione lineare.

È immediato, anche, verificare che vale il Teorema fondamentale del calcolo: se $f \in C([a, b], \mathbb{R}^n)$ e se $F(t) := \int_{t_0}^t f(s) ds$ per un qualunque $t_0 \in [a, b]$, allora $F'(t) = f(t)$.

Inoltre, dalla teoria dell'integrazione di Riemann⁵⁸ segue che, per ogni $f \in C([a, b], \mathbb{R})$,

$$\int_a^b f(t) dt = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f(t_i), \quad t_i := a + i \frac{b-a}{N}. \tag{1.65}$$

⁵⁸Si veda, per esempio, Proposizione 8.19 di [C2019].

Dunque, se $|\cdot|$ è una qualunque norma in \mathbb{R}^n , se t_i è come in (1.65), si ha

$$\begin{aligned} \left| \int_a^b f(t) dt \right| &= \left| \left(\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f_1(t_i), \dots, \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f_n(t_i) \right) \right| = \left| \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N f(t_i) \right| \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \left| \sum_{i=1}^N f(t_i) \right| \leq \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N |f(t_i)| = \int_a^b |f(t)| dt . \end{aligned}$$

Abbiamo dimostrato: se $f \in C([a, b], \mathbb{R}^n)$ e se $|\cdot|$ è una qualunque norma su \mathbb{R}^n allora

$$\left| \int_a^b f(t) dt \right| \leq \int_a^b |f(t)| dt . \quad (1.66)$$

Osserviamo, infine, anche che se $f \in C([a, b], \mathbb{R}^n)$ e $A \in \mathbb{R}^{m \times n}$, allora⁵⁹

$$A \int_a^b f(t) dt = \int_a^b A f(t) dt . \quad (1.67)$$

Gli argomenti usati si estendono immediatamente anche al caso di **funzioni a valori matriciali**:

Sia I un intervallo di \mathbb{R} e sia $t \in I \rightarrow A(t)$ una funzione a valori nell'insieme $\mathbb{R}^{n \times m}$ delle matrici $n \times m$; ossia, per ogni $t \in I$, $A(t)$ è una matrice $n \times m$ con elementi $A_{ij}(t)$ (con $1 \leq i \leq n$ e $1 \leq j \leq m$). Diremo che la funzione $t \in I \rightarrow A(t) \in \mathbb{R}^{n \times m}$ è continua, ovvero $A \in C(I, \mathbb{R}^{n \times m})$, se sono continue le nm funzioni $A_{ij}(t)$ (per ogni i, j). Analogamente diremo che $t \in I \rightarrow A(t)$ è $C^1(I, \mathbb{R}^{n \times m})$ se sono C^1 le funzioni $t \rightarrow A_{ij}(t)$ e per definizione la **derivata di A** sarà data dalla matrice di elementi $A'_{ij}(t)$ ovvero, per ogni i, j , $(A_{ij}(t))' := A'_{ij}(t)$. Se $A \in C(I, \mathbb{R}^{n \times m})$ possiamo definire l'**integrale di A** integrando elemento per elemento, ossia, per definizione, $\int_I A(t) dt$ è una matrice $n \times m$ tale che

$$\left(\int_I A(t) dt \right)_{ij} := \int_I A_{ij}(t) dt . \quad (1.68)$$

E' immediato verificare, anche in questo caso il Teorma fondamentale del calcolo:

se $A \in C(I, \mathbb{R}^{n \times m})$ e, per ogni $t_0 \in I$, la funzione $t \rightarrow B(t) := \int_{t_0}^t A(\tau) d\tau$ appartiene a $C^1(I, \mathbb{R}^{n \times m})$ e $B'(t) = A(t)$.

Complemento 1.4: Disuguaglianze di Hölder e Minkowski. Norme p su \mathbb{R}^n

Le norme introdotte in § 1.2 su \mathbb{R}^n sono esempi particolari delle cosiddette 'norme p '.

Se $p \geq 1$ e $x \in \mathbb{R}^n$ (o $x \in \mathbb{C}^n$), definiamo

$$|x|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} . \quad (1.69)$$

La disuguaglianza di Hölder è una generalizzazione della disuguaglianza di Schwarz: siano x_i e y_i , $i = 1, \dots, n$ numeri reali (o complessi), e siano p e q due numeri reali maggiori di 1 e coniugati cioè

$$\frac{1}{p} + \frac{1}{q} = 1 , \quad p, q > 1 , \quad (1.70)$$

allora, si ha:

$$\sum_{i=1}^n |x_i y_i| \leq \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}} \left(\sum_{i=1}^n |y_i|^q \right)^{\frac{1}{q}} = |x|_p |y|_q , \quad (1.71)$$

dove $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$.

La dimostrazione di tale disuguaglianza è basata sulla convessità della funzione esponenziale, come si evince dal seguente

⁵⁹Eq. (1.67) segue immediatamente da (1.65) e dal fatto che la funzione $x \in \mathbb{R}^n \mapsto Ax$ è continua.

Lemma 1.69 Siano p e q come in (1.70). Allora per ogni s e t in $[0, \infty)$ si ha

$$st \leq \frac{s^p}{p} + \frac{t^q}{q} . \quad (1.72)$$

Dimostrazione Assumiamo che $st > 0$ altrimenti la disuguaglianza è ovvia. Dalla convessità della funzione $x \rightarrow \exp(x)$ segue

$$\begin{aligned} st &= \exp(\log s + \log t) = \exp\left(\frac{1}{p} p \log s + \frac{1}{q} q \log t\right) \\ &\leq \frac{1}{p} \exp(p \log s) + \frac{1}{q} \exp(q \log t) = \frac{s^p}{p} + \frac{t^q}{q} . \quad \blacksquare \end{aligned}$$

È facile ora dimostrare la disuguaglianza di Hölder: se $|x|_p$ o $|y|_q$ sono nulli significa che $x = 0$ o $y = 0$, nel qual caso la disuguaglianza è vera (col segno $=$). Assumiamo dunque che $|x|_p \neq 0$ e $|y|_q \neq 0$. Definiamo

$$a_i := \frac{|x_i|}{|x|_p}, \quad b_i := \frac{|y_i|}{|y|_q},$$

cosicché (1.71) risulti equivalente a

$$\sum_{i=1}^n a_i b_i \leq 1, \quad \text{con} \quad \sum_{i=1}^n a_i^p = \sum_{i=1}^n b_i^q = 1 . \quad (1.73)$$

Usando n volte la (1.72), si ha

$$\sum_{i=1}^n a_i b_i \leq \frac{1}{p} \sum_{i=1}^n a_i^p + \frac{1}{q} \sum_{i=1}^n b_i^q = \frac{1}{p} + \frac{1}{q} = 1 ,$$

che è quanto volevasi dimostrare. \blacksquare

Una conseguenza importante della disuguaglianza di Hölder è la seguente **disuguaglianza di Minkowski** (che non è altro che la disuguaglianza triangolare per $|\cdot|_p$):

$$|x + y|_p \leq |x|_p + |y|_p, \quad \forall x, y \in \mathbb{R}^n . \quad (1.74)$$

Dimostrazione Se $p = 1$ sappiamo che la disuguaglianza è vera. Assumiamo dunque che p sia un numero reale $p > 1$. Assumiamo anche che sia x che y non siano nulli (cioè che $|x|_p \neq 0$ e $|y|_p \neq 0$) e definiamo i vettori $a, b \in \mathbb{R}_+^n$ come i vettori le cui componenti siano

$$a_i := |x_i|, \quad b_i := |y_i| .$$

È evidente che

$$|x + y|_p \leq |a + b|_p, \quad |x|_p = |a|_p, \quad |y|_p = |b|_p . \quad (1.75)$$

Infine sia $q := p(p-1)^{-1}$, cosicché $\frac{1}{p} + \frac{1}{q} = 1$. Allora, essendo $(p-1)q = p$, si ha

$$\begin{aligned} |a + b|_p^p &= \sum_{i=1}^n (a_i + b_i)^p = \sum_{i=1}^n (a_i + b_i)(a_i + b_i)^{p-1} \\ &= \sum_{i=1}^n a_i (a_i + b_i)^{p-1} + \sum_{i=1}^n b_i (a_i + b_i)^{p-1} \\ &\leq (|a|_p + |b|_p) \left(\sum_{i=1}^n (a_i + b_i)^{(p-1)q} \right)^{\frac{1}{q}} \\ &= (|a|_p + |b|_p) (|a + b|_p)^{\frac{p}{q}} \end{aligned}$$

cioè

$$\left(|a + b|_p \right)^{p - \frac{p}{q}} \leq |a|_p + |b|_q$$

ed essendo $p - \frac{p}{q} = 1$ si ha l'asserto. \blacksquare

Osservazione 1.70 (i) Ovviamente, per ogni $p \geq 1$, $|x|_p = 0$ implica $x = 0$ e $|ax|_p = |a| |x|_p$ per ogni $a \in \mathbb{R}$ e $x \in \mathbb{R}^n$. Dunque dalla disuguaglianza di Minkowski (1.74) segue che $|\cdot|_p$ è una norma su \mathbb{R}^n .

(ii) Non è difficile mostrare che⁶⁰, per ogni $x \in \mathbb{R}^n$, $\lim_{p \rightarrow +\infty} |x|_p = |x|_\infty$, il che giustifica la scelta del simbolo per la norma del massimo.

Complemento 1.5: Norme ed esponenziale di matrici

L'insieme delle matrici ($n \times m$) su un campo K (di norma $K = \mathbb{R}$ o $K = \mathbb{C}$), denotato con $K^{n \times m}$, ha una struttura naturale di spazio vettoriale⁶¹ (su K): se $A = (A_{ij}), B = (B_{ij}) \in K^{m \times n}$, $a \in K$, $(A+B)_{ij} := A_{ij} + B_{ij}$ e $(aA)_{ij} = aA_{ij}$. Tale spazio vettoriale ha dimensione $n \times m$ (una base è data dalle nm matrici (δ_{ij}) dove $\delta_{ij} = 1$ se $i = j$ e $\delta_{ij} = 0$ se $i \neq j$) ed è isomorfo (come spazio vettoriale) a K^{nm} (il che giustifica la notazione). Quindi, qualunque norma su K^{nm} induce una norma su $K^{n \times m}$ (come è facile verificare). In realtà, però, vi sono norme più 'utili', le cosiddette 'norme operatoriali', che riflettono la funzione di operatori lineari che hanno le matrici.

Questo è il punto di vista di questo complemento, in cui mostreremo come tali norme operatoriali permettano, in particolare, di estendere le funzioni $1/(1-x)$ e e^x al caso in cui la variabile reale x sia sostituita da una matrice quadrata.

Norme di matrici

Proposizione 1.71 (i) Siano $|\cdot|_a$ e $|\cdot|_b$ due norme arbitrarie in, rispettivamente, K^m e K^n e per $A \in K^{n \times m}$ si definisca

$$\|A\|_{a,b} := \sup_{\substack{x \in K^m \\ |x|_a = 1}} |Ax|_b. \quad (1.76)$$

Allora $\|\cdot\|_{a,b}$ è una norma su $K^{n \times m}$ e $(K^{n \times m}, \|\cdot\|_{a,b})$ è uno spazio di Banach⁶². Inoltre,

$$|Ax|_b \leq \|A\|_{a,b} |x|_a, \quad \forall A \in K^{n \times m}, \quad \forall x \in K^m. \quad (1.77)$$

Nel caso speciale $a = b = \infty$ si ha

$$\|A\|_\infty := \|A\|_{\infty, \infty} = \max_{1 \leq i \leq n} \sum_{j=1}^m |A_{ij}|, \quad (1.78)$$

dove A_{ij} denotano gli elementi di matrice di A .

(ii) Siano $|\cdot|_a, |\cdot|_b$ e $|\cdot|_c$ tre norme arbitrarie in, rispettivamente, K^m, K^n e K^p . Allora, se $A \in K^{n \times m}$ e se $B \in K^{m \times p}$, si ha che

$$\|AB\|_{c,b} \leq \|A\|_{a,b} \|B\|_{c,a}. \quad (1.79)$$

Inoltre se I_n denota la matrice identità in $K^{n \times n}$ si ha che $\|I_n\|_{a,a} = 1$ per ogni scelta della norma $|\cdot|_a$ in K^n : $(K^{n \times n}, \|\cdot\|_{a,a})$ è una algebra di Banach.

(iii) (**Serie di Neumann**⁶³) Sia $A \in K^{n \times m}$ e $\|\cdot\| := \|\cdot\|_{a,b}$. Se $\|A\| < 1$ allora la matrice $(I_n - A)$ è invertibile e⁶⁴

$$(I_n - A)^{-1} = \sum_{k=0}^{\infty} A^k \quad \text{ovvero} \quad \lim_{N \rightarrow \infty} \|(I_n - A)^{-1} - \sum_{k=0}^N A^k\| = 0. \quad (1.80)$$

Dimostrazione (i): che (1.76) definisca una norma è immediato come è immediato verificare che anche $\|A\| := \max_{i,j} |A_{ij}|$ è una norma. Dunque, poiché dalla Proposizione 1.36 segue che le norme $\|\cdot\|$ e $\|\cdot\|_{a,b}$ sono equivalenti è sufficiente dimostrare che $K^{n \times m}$ è uno spazio di Banach rispetto alla norma $\|\cdot\|$. Ma la convergenza in norma $\|\cdot\|$ è equivalente alla convergenza elemento per elemento e dunque dalla completezza di K segue la completezza di $(K^{n \times m}, \|\cdot\|)$.

⁶⁰Es 1.4.

⁶¹Per informazioni generali di algebra lineare, si veda l'Appendice D.

⁶²Le norme (1.76) vengono dette **norme matriciali** o **norme operatoriali**.

⁶³Cfr. https://en.wikipedia.org/wiki/Neumann_series.

⁶⁴ $A^0 := I_n$ per qualunque matrice A (inclusa la matrice nulla).

Per $x = 0$ la (1.77) è verificata (con $=$). Sia $0 \neq x \in K^m$, allora, dalla definizione di $\|\cdot\|_{a,b}$, segue che

$$\left\| A \frac{x}{|x|_a} \right\| \leq \|A\|_{a,b},$$

il che è equivalente a (1.77).

Per verificare (1.78) si osservi che se $x \in K^m$ con $|x|_\infty \leq 1$ (ovvero $|x_j| \leq 1$ per ogni j),

$$|Ax|_\infty := \sup_i \left| \sum_j A_{ij} x_j \right| \leq \sup_i \sum_j |A_{ij}| |x_j| \leq \sup_i \sum_j |A_{ij}|$$

ovvero $\|A\|_{\infty, \infty} \leq \sup_{1 \leq i \leq n} \sum_{j=1}^m |A_{ij}|$. D'altra parte, sia i_0 tale che $\sup_i \sum_j |A_{ij}| = \sum_j |A_{i_0 j}|$ e sia

$x_j := \text{segno}(A_{i_0 j})$ se $A_{i_0 j} \neq 0$ e $x_j = 1$ altrimenti; allora si ha che $|x|_\infty = 1$ e $A_{i_0 j} x_j \geq 0$. In tal caso $\sup_i \sum_j |A_{ij}| = \sum_j |A_{i_0 j}| = \sum_j A_{i_0 j} x_j = \left| \sum_j A_{i_0 j} x_j \right| \leq \|A\|_{\infty, \infty}$ il che prova (1.78).

(ii): se $x \in K^p$ con $|x|_c = 1$ allora (per l'associatività del prodotto tra matrici) e per la definizione (1.76) si ha che

$$|ABx|_b \leq \|A\|_{a,b} |Bx|_a \leq \|A\|_{a,b} \|B\|_{c,a},$$

il che implica (1.79). L'asserto su $\|I_n\|_{a,b}$ è ovvio.

(iii): poiché $\|A\| < 1$, dal punto (ii) (iterato) segue che $\|A^k\| \leq \|A\|^k$ per ogni k e dunque, per ogni $M \geq N \geq 0$, $\left\| \sum_{k=N}^M A^k \right\| \leq \sum_{k=N}^M \|A\|^k$, quantità che tende a zero al crescere di N . Questo significa

che la successione 'delle ridotte' $\left\{ \sum_{k=0}^N A^k \right\}$ è una successione di Cauchy e quindi, per il punto (i), ammette limite in $(K^{n \times m}, \|\cdot\|)$. Sia B tale limite e facciamo vedere che B coincide con l'inversa di⁶⁵ $(I_n - A)$:

$$\begin{aligned} B(I_n - A) &:= \left(\lim_{N \rightarrow \infty} \sum_{k=0}^N A^k \right) (I_n - A) = \lim_{N \rightarrow \infty} \sum_{k=0}^N (A^k (I_n - A)) \\ &= \lim_{N \rightarrow \infty} I_n - A^{N+1} = I_n, \end{aligned}$$

ma dire che $B = (I_n - A)^{-1}$ è equivalente alla (1.80). ■

Esponenziale di matrici

Sia $A \in K^{n \times m}$ e consideriamo la successione di matrici $e_N(A) := \left\{ \sum_{k=1}^N \frac{A^k}{k!} \right\}$. Se $M > N \geq 0$,

$$\begin{aligned} \left\| \sum_{k=0}^M \frac{A^k}{k!} - \sum_{k=0}^N \frac{A^k}{k!} \right\| &= \left\| \sum_{k=N+1}^M \frac{A^k}{k!} \right\| \leq \sum_{k=N+1}^M \left\| \frac{A^k}{k!} \right\| \\ &\leq \sum_{k=N+1}^M \frac{\|A\|^k}{k!} \leq \sum_{k=N+1}^{\infty} \frac{\|A\|^k}{k!}, \end{aligned}$$

ma questa è la coda della serie esponenziale⁶⁶ $\exp(\|A\|) = e^{\|A\|}$ e, quindi, la successione $\{e_N(A)\}$ è di Cauchy e il suo limite è, per definizione, la matrice **esponenziale di A**:

$$\lim_{N \rightarrow \infty} e_N(A) := \lim_{N \rightarrow \infty} \sum_{k=1}^N \frac{A^k}{k!} =: \sum_{k=1}^{\infty} \frac{A^k}{k!} =: \exp(A) =: e^A. \quad (1.81)$$

⁶⁵Si noti che se B e C sono due matrici quadrate e $BC = I_n$ allora $B = C^{-1}$: infatti da $BC = I_n$ segue che $(\det B)(\det C) = 1$ e quindi sia B che C sono invertibili e moltiplicando a destra ambo i membri della relazione $BC = I_n$ si ottiene $B = C^{-1}$ (e, invertendo ambo i membri di tale relazione, $B^{-1} = C$). Si osservi anche se $\{A_k\}$ tende a A (cioè, per definizione, $\|A_k - A\| \rightarrow 0$) allora, per ogni B , $\{A_k B\}$ tende a AB ; infatti: $\|A_k B - AB\| \leq \|A_k - A\| \|B\| \rightarrow 0$.

⁶⁶Si veda, per esempio, [C2019], Cap. 5.

Dalla definizione e dal calcolo appena fatto segue immediatamente che

$$\exp(0) = I_n, \quad \|\exp(A)\| \leq \exp(\|A\|), \quad \forall A \in K^{n \times m}. \quad (1.82)$$

Estendiamo, altre proprietà dell'esponenziale al caso matriciale.

Se A, B sono due matrici in $K^{n \times m}$, si definisce il **commutatore di A e B** la matrice $[A, B] \in K^{n \times m}$ definita da

$$[A, B] := AB - BA. \quad (1.83)$$

Si dice che **due matrici quadrate commutano** se il loro commutatore è nullo. Osserviamo che per due matrici che commutino vale la formula del binomio di Newton cioè:

$$[A, B] = 0 \implies (A + B)^k = \sum_{j=0}^k \binom{k}{j} A^j B^{k-j}. \quad (1.84)$$

Per esempio $(A + B)^2 = A^2 + AB + BA + B^2$ ma poiché $AB = BA$ (essendo $[A, B] = 0$) si ottiene la formula nota del quadrato di un binomio $A^2 + 2AB + B^2$. Naturalmente se A e B non commutano la formula di Newton è, in generale, falsa: per esempio

$$\begin{aligned} A &:= \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad B := \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} \implies [A, B] = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \\ A^2 = B^2 = 0, \quad A^2 + 2AB + B^2 &= \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix} \neq (A + B)^2 = I. \end{aligned} \quad (1.85)$$

Il fatto che, per due matrici che commutano, valga la formula del binomio di Newton permette di adattare in maniera ovvia la dimostrazione della formula di addizione per la funzione reale $x \in \mathbb{R} \rightarrow e^x = \exp(x)$ e di ottenere il seguente “teorema di addizione per esponenziali di matrici”:

$$[A, B] = 0 \implies e^{A+B} = e^A e^B = e^B e^A. \quad (1.86)$$

In particolare tale relazione implica che: se A e B commutano allora commutano anche le matrici e^A e e^B . Si noti anche che, poiché A e $-A$ commutano, si ha che

$$e^A e^{-A} = e^{-A} e^A = e^0 = I,$$

ovvero la matrice $\exp(A)$ è sempre invertibile e la sua matrice inversa è $\exp(-A)$.

Calcoliamo ora la derivata di e^{At} dove A è una matrice quadrata assegnata e t è una variabile reale⁶⁷. Consideriamo dunque il rapporto incrementale in $t \in \mathbb{R}$: poiché e^{At} e e^{As} commutano per ogni t ed s in \mathbb{R} , si ha, per ogni numero $h \neq 0$,

$$\frac{e^{A(t+h)} - e^{At}}{h} = e^{At} \frac{e^{Ah} - I}{h} = \frac{e^{Ah} - I}{h} e^{At}. \quad (1.87)$$

Inoltre⁶⁸

$$\frac{e^{Ah} - I}{h} = \sum_{j=1}^{\infty} \frac{(Ah)^j}{h j!} := A + hB(h),$$

ove $B(h) := \sum_{j=0}^{\infty} h^j \frac{A^{j+2}}{(j+2)!}$. Per $|h| \leq 1$ si ha $\|B(h)\| \leq \sum_{j=0}^{\infty} \|A\|^{j+2}/(j+2)! < \infty$ (che è un numero

indipendente da h); quindi $h\|B(h)\| \rightarrow 0$ per $h \rightarrow 0$ e questo è equivalente a dire che $\lim_{h \rightarrow 0} \frac{e^{Ah} - I}{h} = A$, dunque, da (1.87) segue che

$$(e^{At})' := \lim_{h \rightarrow 0} \frac{e^{A(t+h)} - e^{At}}{h} = A e^{At} = e^{At} A. \quad (1.88)$$

⁶⁷ In generale, se X e Y sono spazi normati dotati di norme, rispettivamente, $\|\cdot\|_X$ e $\|\cdot\|_Y$, si possono considerare funzioni $f: x \in E \subseteq X \rightarrow f(x) \in Y$ ovvero funzioni con variabile in X e a valori nello spazio normato Y : f è **continua** in $x_0 \in E$ se $\forall \varepsilon > 0 \exists \delta > 0$ tale che $\|f(x) - f(x_0)\|_Y < \varepsilon$ per ogni $x \in E$ con $\|x - x_0\|_X < \delta$; $f(x)$ **tende** a y_0 per x che tende a x_0 , e si denota $y_0 = \lim_{x \rightarrow x_0} f(x)$, se $\forall \varepsilon > 0 \exists \delta > 0$ tale che $\|f(x) - y_0\|_Y < \varepsilon$ per ogni $x \in E$ con $0 < \|x - x_0\|_X < \delta$. Se lo spazio “di partenza” X è il campo degli scalari K anche la nozione di derivata di $f(t)$ si dà nel solito modo: ovvero $f'(t_0)$ è la **derivata** di f in t_0 se esiste il limite del rapporto incrementale (che è un elemento dello spazio normato Y) $(f(t) - f(t_0))/(t - t_0)$ per t che tende a t_0 .

⁶⁸ Si osservi che se $A^{(k)} \in K^{n \times m}$ è una successione di matrici convergente alla matrice A (ovvero $\|A^{(k)} - A\| \rightarrow 0$) allora per ogni scalare a si ha che $aA^{(k)} \rightarrow aA$: infatti $\|aA^{(k)} - aA\| = |a| \|A^{(k)} - A\|$.

Osservazione 1.72 In generale, se $A_k \in K^{n \times m}$ e $A_k \rightarrow A$ (rispetto ad una norma matriciale), allora $\|A_k x - Ax\| \leq \|A_k - A\| \|x\| \rightarrow 0$ per ogni $x \in K^m$. Analogamente se $A(t) \rightarrow A$ (ossia, $\|A(t) - A\| \rightarrow 0$) per $t \rightarrow t_0$, allora $A(t)x \rightarrow Ax$ uniformemente in $x \in K^m$. In particolare, da (1.88) segue che

$$(e^{At} x)' := \lim_{h \rightarrow 0} \frac{e^{A(t+h)} x - e^{At} x}{h} = A e^{At} x = e^{At} A x, \quad \forall x. \quad (1.89)$$

La forma canonica di Jordan⁶⁹ permette di calcolare esplicitamente l'esponenziale di una matrice $A \in \mathbb{C}^{n \times n}$ e più in generale della matrice At con $t \in \mathbb{R}$.

Per calcolare la matrice $e^{At} := \exp(At)$ con $A \in \mathbb{C}^{n \times n}$ e $t \in \mathbb{R}$, facciamo due osservazioni preliminari:

(a) Se $A = U^{-1} B U$, allora, essendo (come è immediato verificare), $A^k = U^{-1} B^k U$, si ha che

$$e^A := \sum_{k=0}^{\infty} \frac{A^k}{k!} = \sum_{k=0}^{\infty} \frac{U^{-1} B^k U}{k!} = U^{-1} \left(\sum_{k=0}^{\infty} \frac{B^k}{k!} \right) U = U^{-1} e^B U. \quad (1.90)$$

(b) Se $n \geq 2$, le potenze di $J_n(0)t = [0, te^{(1)}, \dots, te^{(n-1)}]$ sono date da (come è immediato verificare) da

$$(J_n(0)t)^k = \begin{cases} [\underbrace{0, \dots, 0}_{k \text{ volte}}, te^{(1)}, \dots, te^{(n-k)}], & \text{se } 1 \leq k < n, \\ 0, & \text{se } k \geq n. \end{cases} \quad (1.91)$$

Quindi,

$$e^{J_2(0)t} = \begin{pmatrix} 1 & t \\ 0 & 1 \end{pmatrix}, \quad e^{J_3(0)t} = \begin{pmatrix} 1 & t & \frac{t^2}{2} \\ 0 & 1 & t \\ 0 & 0 & 1 \end{pmatrix}, \quad (1.92)$$

e, per $n \geq 4$:

$$e^{J_n(0)t} = \sum_{k=0}^{\infty} \frac{(J_n(0)t)^k}{k!} = \sum_{k=0}^{n-1} \frac{(J_n(0)t)^k}{k!} = \begin{pmatrix} 1 & t & \frac{t^2}{2!} & \frac{t^3}{3!} & \cdots & \frac{t^{n-1}}{(n-1)!} \\ 0 & 1 & t & \frac{t^2}{2!} & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & \ddots & \frac{t^3}{3!} \\ \vdots & \vdots & & \ddots & \ddots & \frac{t^2}{2!} \\ \vdots & \vdots & & & \ddots & t \\ 0 & \cdots & \cdots & \cdots & \cdots & 1 \end{pmatrix}. \quad (1.93)$$

Quindi se U è la matrice che coniuga A alla sua forma canonica di Jordan⁷⁰ $J := \text{block}(J_1, \dots, J_k)$ con $J_i = J_{n_i}(\lambda_i)$ per $1 \leq i \leq k$, si ha che

$$e^{At} \stackrel{(a)}{=} U e^{Jt} U^{-1} = U \text{block}(e^{J_1 t}, \dots, e^{J_k t}) U^{-1}, \quad (1.94)$$

e⁷¹

$$e^{J_i t} = e^{J_{n_i}(0)t} = e^{\lambda_i t I_{n_i} + J_{n_i}(0)t} = e^{\lambda_i t} e^{J_{n_i}(0)t} \quad (1.95)$$

con $e^{J_{n_i}(0)t}$ come in (1.92) e (1.93) con n sostituito da n_i .

Complemento 1.6: Spazi di Banach e spazi di Hilbert

Come già detto, gli spazi normati (Definizione 1.8) sono spazi metrici: se $(X, \|\cdot\|)$ è uno spazio normato e se poniamo, per ogni $x, y \in X$, $d(x, y) := \|x - y\|$, è immediato verificare che d è una metrica su E .

Definizione 1.73 Uno spazio normato completo⁷² prende il nome di spazio di Banach⁷³.

⁶⁹Si veda l'Appendice D.9.

⁷⁰Cfr. (D.96).

⁷¹ $\lambda_i t I_{n_i}$ commuta con ogni matrice di ordine n_i .

⁷²Nella topologia indotta dalla norma.

⁷³Stefan Banach (1892–1945) https://en.wikipedia.org/wiki/Stefan_Banach.

Uno dei più importanti esempi di spazi di Banach è lo spazio di funzioni continue rispetto alla metrica uniforme.

Fissiamo un insieme arbitrario non vuoto $E \subseteq \mathbb{R}^n$. L'insieme delle funzioni continue su E a valori in \mathbb{R}^m , $C(E, \mathbb{R}^m)$, è uno spazio vettoriale su \mathbb{R} . Se $f \in C(E, \mathbb{R}^m)$, poniamo⁷⁴

$$\|f\|_\infty := \|f\|_{E, \infty} := \sup_E |f| := \sup_{x \in E} |f(x)|; \quad (1.96)$$

definiamo, poi,

$$\mathcal{B} := \{f \in C(E, \mathbb{R}^m) : \|f\|_\infty < \infty\}. \quad (1.97)$$

È immediato verificare che $\|\cdot\|_\infty$ è una norma su \mathcal{B} e quindi che $(\mathcal{B}, \|\cdot\|_\infty)$ è uno spazio normato.

Si noti che la convergenza in $(\mathcal{B}, \|\cdot\|_\infty)$ non è altro che la convergenza uniforme⁷⁵.

Dalle Proposizioni 1.56 e 1.58 segue poi che tale spazio normato è completo:

Teorema 1.74 $(\mathcal{B}, \|\cdot\|_\infty)$ è uno spazio di Banach.

Dimostrazione Sia $\{f_k\}$ una successione di Cauchy in \mathcal{B} . Dal Lemma 1.55 segue che $\{f_k\}$ converge uniformemente ad una funzione f . Per la Proposizioni 1.56 f è continua su E e per la Proposizione 1.58 la funzione f è limitata. Quindi $f \in \mathcal{B}$. ■

Altri esempi di spazi di Banach (infinito dimensionali) sono dati dagli spazi di successioni. Definiamo, per ogni $p \geq 1$,

$$\ell^p := \{x = \{x_k\} : k \in \mathbb{N} \mapsto x_k \in \mathbb{R} \text{ e } \|x\|_p := \left(\sum_{k=1}^{\infty} |x_k|^p\right)^{1/p} < +\infty\}. \quad (1.98)$$

Allora si ha⁷⁶:

Proposizione 1.75 $(\ell^p, \|\cdot\|_p)$ è uno spazio di Banach per ogni $p \geq 1$ o $p = \infty$.

Dimostrazione Sia $p \geq 1$. Innanzitutto osserviamo che dalla disuguaglianza di Minkowski 1.74 segue che $\|\cdot\|_p$ è una norma su ℓ^p : infatti, se $x, y \in \ell^p$, allora, per ogni $n \in \mathbb{N}$ si ha che

$$\left(\sum_{k=1}^n |x_k + y_k|^p\right)^{1/p} \stackrel{(1.74)}{\leq} \left(\sum_{k=1}^n |x_k|^p\right)^{1/p} + \left(\sum_{k=1}^n |y_k|^p\right)^{1/p} \leq \|x\|_p + \|y\|_p,$$

da cui, prendendo il limite per $n \rightarrow \infty$, segue la disuguaglianza triangolare per $\|\cdot\|_p$ su ℓ^p (la non degenerazione e la omogeneità sono ovvie).

Sia ora $\varepsilon > 0$ e sia $\{x^{(j)}\}$ una successione di Cauchy in ℓ^p . Allora esiste $j_0 \in \mathbb{N}$ tale che $\|x^{(i)} - x^{(j)}\|_p < \varepsilon$ per ogni $i, j \geq j_0$ e dunque

$$\left(\sum_{k=1}^n |x_k^{(i)} - x_k^{(j)}|^p\right)^{1/p} < \varepsilon \quad \forall n \in \mathbb{N}, \quad \forall i, j \geq j_0. \quad (1.99)$$

In particolare, per ogni k sono di Cauchy le successioni di numeri reali $\{x_k^{(j)}\}_{j \in \mathbb{N}}$ e dunque (per la completezza di \mathbb{R}) esiste $x_k \in \mathbb{R}$ tale che $x_k = \lim_{j \rightarrow \infty} x_k^{(j)}$. Prendendo prima il limite per $j \rightarrow \infty$ e poi il limite per $n \rightarrow \infty$ in (1.99) si ha che $\|x^{(i)} - x\|_p \leq \varepsilon$ per ogni $i \geq j_0$. Questa relazione implica che $x \in \ell^p$ e che $\lim x^{(i)} = x$ (che per definizione significa che $\|x^{(i)} - x\|_p \rightarrow 0$).

La dimostrazione nel caso $p = \infty$ è analoga (ma più semplice) e viene lasciata per esercizio⁷⁷. ■

⁷⁴Naturalmente, se f non è limitata, tale estremo superiore è uguale a $+\infty$.

⁷⁵Ossia, $\|f_k - f\|_\infty \rightarrow 0$ se e solo se $f \rightarrow f$ uniformemente su E : cfr. Osservazione 1.54–(ii).

⁷⁶Si ricordi la definizione di ℓ^∞ data nell'Esempio 1.47–(iii).

⁷⁷Es 1.59.

Spazi di Hilbert

Generalizziamo ora la nozione di prodotto scalare e ortogonalità in spazi astratti:

Definizione 1.76 Sia X uno spazio vettoriale su \mathbb{R} . Un **prodotto scalare** su X , è un'applicazione $\langle \cdot, \cdot \rangle : V \times V \rightarrow \mathbb{R}$ bilineare, simmetrica e definita positiva, ossia, tale che

$$\begin{aligned} \text{(i)} \quad & \langle x, y \rangle = \langle y, x \rangle, & \forall x, y \in X, & \quad \text{(simmetria)} \\ \text{(ii)} \quad & \begin{cases} \langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle, \\ \langle ax, y \rangle = a \langle x, y \rangle, \end{cases} & \begin{array}{l} \forall x, y, z \in X, \\ \forall x, y \in X, \forall a \in \mathbb{R}, \end{array} & \quad \text{(linearità)} \\ \text{(iii)} \quad & \langle x, x \rangle > 0, & \forall x \in X, x \neq 0. & \quad \text{(positività)} \end{aligned}$$

Osservazione 1.77 (i) Si osservi che da (i) e (ii) segue che il prodotto scalare è lineare anche come funzione della seconda variabile $y \in X \mapsto \langle x, y \rangle$ (il che spiega il termine ‘bilineare’). Si noti anche che dalla linearità segue che $\langle x, 0 \rangle = 0$ per ogni $x \in X$.

(ii) Su \mathbb{R}^n è possibile definire prodotti scalari diversi dal prodotto scalare euclideo: esempi di prodotti scalari su \mathbb{R}^n , infatti, sono dati da $\langle x, y \rangle := x \cdot Py$ con P matrice $(n \times n)$ definita positiva⁷⁸.

(iii) Dalla linearità segue immediatamente la seguente generalizzazione del quadrato di un binomio (cfr. (1.11)):

$$\langle x + y, x + y \rangle = \langle x, x \rangle + \langle y, y \rangle + 2\langle x, y \rangle, \quad \forall x, y \in X. \quad (1.100)$$

Dalla (1.100) si ottiene (con la stessa dimostrazione!) la seguente generalizzazione della Proposizione 1.4:

Proposizione 1.78 (Disuguaglianza di Cauchy) Sia $\langle \cdot, \cdot \rangle$ un prodotto scalare su X . Allora,

$$|\langle x, y \rangle| \leq \sqrt{\langle x, x \rangle \langle y, y \rangle}, \quad \forall x, y \in X. \quad (1.101)$$

Se, ora, poniamo

$$\|x\| := \sqrt{\langle x, x \rangle}, \quad \forall x \in X, \quad (1.102)$$

vediamo (come nel caso di \mathbb{R}^n) che dalla disuguaglianza di Cauchy segue la disuguaglianza triangolare

$$\|x + y\| \leq \|x\| + \|y\|, \quad \forall x, y \in X. \quad (1.103)$$

Dunque, $\|\cdot\|$ soddisfa gli assiomi (a), (b) e (c) di norma nella Definizione 1.8 ossia:

uno spazio vettoriale X su cui è definito un prodotto scalare $\langle \cdot, \cdot \rangle$ è uno spazio normato con norma $\|\cdot\|$ definita in (1.102).

Definizione 1.79 Uno spazio vettoriale X , su cui è definito un prodotto scalare $\langle \cdot, \cdot \rangle$, completo rispetto alla distanza indotta dalla norma (1.102) si chiama **spazio di Hilbert**.

Chiaramente \mathbb{R}^n è uno spazio di Hilbert rispetto prodotto scalare standard (essendo la norma Euclidea⁷⁹ $|x| = (x \cdot x)^{1/2}$).

Un esempio importante di spazio di Hilbert *infinito* dimensionale è dato da ℓ^2 dotato del prodotto scalare

$$\langle x, y \rangle := \sum_{k=1}^{\infty} x_k y_k. \quad (1.104)$$

Innanzitutto si osservi che se $x, y \in \ell^2$, allora, ricordando la disuguaglianza di Hölder (1.71) con $p = q = 2$, si ha che, per ogni $n \in \mathbb{N}$,

$$\sum_{i=1}^n |x_i y_i| \leq \left(\sum_{i=1}^n |x_i|^2 \right)^{\frac{1}{2}} \left(\sum_{i=1}^n |y_i|^2 \right)^{\frac{1}{2}} \leq \|x\|_2 \|y\|_2,$$

e quindi, prendendo in tale relazione il limite per $n \rightarrow \infty$ si vede che il prodotto scalare in (1.104) è ben definito su ℓ^2 . Inoltre, $\|x\|_2 = \langle x, x \rangle^{1/2}$ e dalla Proposizione 1.75 segue che $(\ell^2, \|\cdot\|_2)$ è completo e dunque che è uno spazio di Hilbert.

⁷⁸Una matrice reale, simmetrica $(n \times n)$ P si dice definita positiva se $x \cdot Px > 0$ per ogni $x \in \mathbb{R}^n \setminus \{0\}$.

⁷⁹Cfr. § 1.1 e § 3.

Complemento 1.7: La proprietà di Heine–Borel

Come visto nell'Osservazione 1.48–(ii), in spazi metrici (o normati) generali è possibile avere successioni da cui non è possibile estrarre alcuna sottosuccessione convergente. Questo fatto ha naturalmente ripercussioni sulla nozione di compattezza in spazi metrici generali⁸⁰. In uno spazio topologico (e, quindi, in particolare, in uno spazio metrico) **la proprietà di Heine–Borel è la definizione di insieme compatto**.

Come già anticipato, in spazi metrici generali, non vale il Teorema 1.28; il teorema però, si estende se la nozione di limitatezza viene sostituita dalla nozione di **totale limitatezza**:

Definizione 1.80 *Un sottoinsieme $E \subseteq X$ di uno spazio metrico (X, d) si dice **totalmente limitato** se $\forall \varepsilon > 0$ esistono n sfere di raggio ε , $B_\varepsilon(x_i)$, tali che $E \subseteq \bigcup_{i=1}^n B_\varepsilon(x_i)$.*

Osservazione 1.81 (i) Si noti che se E è totalmente limitato è limitato: infatti se $E \subseteq \bigcup_{i=1}^n B_\varepsilon(x_i)$ e se $r = \max_{1 \leq j \leq n} d(x_j, x_1)$, dalla disuguaglianza triangolare segue che, se $x \in B_\varepsilon(x_j)$ allora $d(x, x_1) \leq d(x, x_j) + d(x_j, x_1) < \varepsilon + r$ e quindi $E \subseteq B_{r+\varepsilon}(x_1)$.

(ii) L'insieme E formato dai valori della successione dell'esempio nell'Osservazione 1.48–(ii) è un insieme limitato ma non totalmente limitato⁸¹.

Teorema 1.82 (Heine–Borel) *Sia (X, d) uno spazio metrico e $\emptyset \neq E \subseteq X$. Allora le seguenti proprietà sono equivalenti:*

- (a) E è compatto⁸²;
- (b) E è compatto per successioni⁸³;
- (c) E è chiuso e totalmente limitato.

Dimostrazione Tutte le implicazioni sono dimostrate 'per assurdo'.

(a) \implies (b): Supponiamo, per assurdo, che esista una successione $\{x_j\}$ in E che non ammetta una sottosuccessione convergente e, quindi⁸⁴, l'insieme dei suoi valori $X := \{x_j \mid j \in \mathbb{N}\}$ non ha punti di accumulazione. Dunque, per ogni $x \in E$, esiste una sfera aperta di centro x , $B(x)$, tale che $\#\{j \in \mathbb{N} \mid x_j \in B(x)\} < +\infty$. Chiaramente, $E \subseteq \bigcup_{x \in E} B(x)$ e, per compattezza, esistono $x_1, \dots, x_N \in E$ tali che $E \subseteq \bigcup_{j=1}^N B(x_j)$. Ma, allora, $\mathbb{N} = \{j \in \mathbb{N} \mid x_j \in E\} = \bigcup_{k=1}^N \{j \in \mathbb{N} \mid x_j \in B(x_k)\}$, che è un insieme finito: contraddizione.

(b) \implies (c): Sia $\{x_j\}$ una successione in E convergente. Per (b), esiste $\{x_{j_k}\}$ convergente ad un punto $\bar{x} \in E$ ed, allora, $x_j \rightarrow \bar{x}$: questo significa che E è chiuso.

Supponiamo, ora, per assurdo, che E non sia totalmente limitato, ossia, che esista $\varepsilon > 0$ tale che nessuna collezione finita di sfere aperte di raggio ε ricopra E . Sia x_1 un qualunque punto di E e sia $B_1 := B_\varepsilon(x_1)$. Sia $x_2 \in E \setminus B_1$ e sia $B_2 := B_\varepsilon(x_2)$. Proseguendo in tal modo iterativamente otteniamo una sequenza di sfere B_i di raggio ε e centro x_i tali che $x_{j+1} \in E \setminus \bigcup_{k=1}^j B_k$. Quindi se $j > i$, $d(x_j, x_i) \geq \varepsilon$, e questo implica che $\{x_i\}$ non ammette sottosuccessioni convergenti: contraddizione.

(c) \implies (a): Supponiamo, per assurdo, che E non sia compatto, ossia, che esista un ricoprimento aperto di E , $\{E_j \mid j \in J\}$, da cui non sia possibile estrarre alcun sottoricoprimento finito di E . Poiché E è totalmente limitato, esistono n_1 insiemi aperti $A_j^{(1)}$ di diametro⁸⁵ non superiore ad 1 tali che $E \subseteq \bigcup_{j=1}^{n_1} A_j^{(1)}$ (in questo primo passo gli $A_j^{(1)}$ possono essere delle sfere aperte di raggio 1). Non tutti gli $A_j^{(1)}$ possono essere ricoperti da un numero finito di E_j (altrimenti esisterebbe un sottoricoprimento finito di E formato da E_j). Sia \hat{A}_1 un tale $A_j^{(1)}$. Poiché E è totalmente limitato, esistono n_2 sfere aperte $B_j^{(2)}$ di diametro $1/2$ tali che $\hat{A}_1 \subseteq \bigcup_{j=1}^{n_2} B_j^{(2)}$ e quindi $\hat{A}_1 = \bigcup_{j=1}^{n_2} A_j^{(2)}$ con $A_j^{(2)} := B_j^{(2)} \cap \hat{A}_1$: gli $A_j^{(2)}$ sono insiemi aperti di diametro non superiore a $1/2$ e $A_j^{(2)} \subseteq A_1^{(1)}$. Come prima, non tutti gli $A_i^{(2)}$ possono essere ricoperti da un numero finito di E_j . Sia \hat{A}_2 uno di tali $A_i^{(2)}$. Iterando, otteniamo una sequenza di insiemi aperti non vuoti \hat{A}_j di diametro non superiore a $1/j$,

⁸⁰Si ricordi che nella implicazione (c) \implies (a) del Teorema 1.28 si è usato il Teorema di Bolzano–Weierstrass.

⁸¹Se $\varepsilon < 1/2$, non è possibile ricoprire E con un numero finito di sfere di raggio ε (Es. 1.25).

⁸²Ossia, E soddisfa la proprietà di Heine–Borel.

⁸³Ossia, comunque presa una successione $\{x_k\}$ in E esiste una sottosuccessione di $\{x_k\}$ con limite in E .

⁸⁴Si noti che una successione $\{x_j\}$ ammette una sottosuccessione convergente se e solo se l'insieme dei suoi valori $X := \{x_j \mid j \in \mathbb{N}\}$ ammette un punto di accumulazione.

⁸⁵Il diametro di un sottoinsieme non vuoto A di uno spazio metrico è definito come $\text{diam}(A) = \sup_{x, y \in A} d(x, y)$. Dalla disuguaglianza triangolare segue che una sfera di raggio r ha diametro $2r$.

tali che $\hat{A}_1 \supseteq \hat{A}_2 \supseteq \dots$ e tali che nessun \hat{A}_i può essere ricoperto da un numero finito di E_j . Scegliamo $x_j \in \hat{A}_j$, per ogni j . La successione $\{x_j\}$ è di Cauchy (se $i > j$, $x_i \in \hat{A}_j$ e $d(x_i, x_j) < 1/i$). Quindi, essendo E chiuso, $x_j \rightarrow \bar{x} \in E$ e (essendo $\{E_j\}$ un ricoprimento di E) esiste i tale che $\bar{x} \in E_i$. Poiché E_i è aperto, esiste $\delta > 0$ tale che $B_\delta(x_i) \subseteq E_i$. Sia, ora, $N > 2/\delta$ tale che $d(x_N, \bar{x}) < \delta/2$. Allora, se $x \in \hat{A}_N$, $d(x, x_N) \leq 1/N$ (essendo $\text{diam}(\hat{A}_N) \leq 1/N$) e quindi $x \in \overline{B_{1/N}(x_N)} \subseteq B_{\delta/2}(x_N)$ (essendo $1/N < \delta/2$): ossia, $\hat{A}_N \subseteq B_{\delta/2}(x_N)$. Inoltre, $B_{\delta/2}(x_N) \subseteq B_\delta(\bar{x})$ (essendo $d(\bar{x}, x_N) < \delta/2$). In definitiva, avremmo $\hat{A}_N \subseteq E_i$: contraddizione. ■

Esempio 1.83 Sia $D \subseteq \mathbb{R}$ un insieme compatto e sia $X \subseteq C(D, \mathbb{R})$ una famiglia di funzioni limitate (nella norma del sup) e equicontinue (Definizione 1.65). Allora, dal Teorema di Heine–Borel 1.82 e dal Teorema di Ascoli–Arzelà 1.67 segue che X è compatto nello spazio di Banach $(C(D, \mathbb{R}), \|\cdot\|_\infty)$.

Esercizi

Esercizio 1.1 Dimostrare che il segno di uguaglianza in (1.12) vale se e solo se $y = 0$ oppure se $y \neq 0$ e $x = ty$ con $t \in \mathbb{R}$.

Esercizio 1.2 Si verifichi che le funzioni definite in (1.16) siano delle norme su \mathbb{R}^n .

Esercizio 1.3 Si dimostri che le costanti davanti alle norme in (1.18) sono ottimali, ossia non esistono costanti più piccole per cui valga una delle relazioni (per ogni $x \in \mathbb{R}$).

Esercizio 1.4 Si dimostri che, per ogni $x \in \mathbb{R}^n$, $\lim_{p \rightarrow +\infty} |x|_p = |x|_\infty$.

Esercizio 1.5 (i) Dimostrare che $(\mathbb{R}^n, \mathcal{A})$ con \mathcal{A} definita nella Definizione 1.11 è uno spazio topologico (Definizione 1.12).

(ii) Dimostrare che se (X, \mathcal{F}) è uno spazio topologico e $Y \subseteq X$ allora la famiglia di insiemi $\mathcal{F}_Y := \{A \cap Y \mid A \in \mathcal{F}\}$ forma una topologia su Y .

Esercizio 1.6 Dimostrare che se $g = g(t)$ è una funzione di una variabile, continua in $t_0 \in \mathbb{R}$ e $f = f(x_1, \dots, x_n) = g(x_n)$ allora $f(x)$ è continua in \bar{x} , $\forall \bar{x} \in \mathbb{R}^n$ tale che $\bar{x}_n = t_0$.

Esercizio 1.7 Sia $f(x) := \text{sen}(x_1 x_4^{100}) + \exp(|x|^2)$, $x \in \mathbb{R}^4$. Trovare un $\delta > 0$ tale che se $|x| < \delta$ allora $|f(x) - f(0)| < 1/100$.

Esercizio 1.8 Fissato $\varepsilon > 0$, si trovi un δ tale che $|f(x) - f(\bar{x})| < \varepsilon$ per ogni $|x - \bar{x}| < \delta$ nei seguenti casi:

(i) $f(x) = |x|^\alpha$, $\alpha > 0$, $x \in \mathbb{R}^4$, $\bar{x} = (0, 1, 1, 2)$ e $\bar{x} = 0$;

(ii) $f(x) = \text{sen} \frac{1}{x_1 x_3^2}$, $x \in \mathbb{R}^3$, $\bar{x} = (-1, 0, -1)$;

(iii) $f(x) = \log \left(\cos \left(\prod_{i=1}^n x_i \right) \right)$, $x \in \mathbb{R}^n$, $\bar{x} = 0$;

(iv) $f(x) = \sum_{k=0}^{\infty} e^{-k|x|^2}$, $x \in \mathbb{R}^n$, $\bar{x} = (1, \dots, 1)$;

(v) $f(x) = \tanh |x|_1$, $x \in \mathbb{R}^n$, $\bar{x} = (1, \dots, 1)$;

(vi) $f(x) = (|x|^{3/2}, \tanh |x|_1)$, $x \in \mathbb{R}^4$, $\bar{x} = (0, 1, 1, 2)$.

Esercizio 1.9 Sia⁸⁶ $f(x, y) = \begin{cases} \frac{x^2 + y^2}{\pi + 2 \arctan(y/x)}, & \text{se } x \neq 0, \\ \frac{y^2}{2\pi}, & \text{se } x = 0. \end{cases}$

(i) Si calcoli, per ogni θ fissato, il $\lim_{r \rightarrow 0^+} f(r \cos \theta, r \sin \theta)$.

(ii) È continua f nell'origine?

⁸⁶ $\arctan 0 = 0$.

Esercizio 1.10 (i) Sia $f(x, y)$ definita in un intorno di $(0, 0)$ ed a valori in \mathbb{R} . Dimostrare la seguente affermazione:

Se $\forall \varepsilon > 0, \exists \delta > 0$ tale che, per ogni $0 < r < \delta$ e per ogni θ , si ha $|f(r \cos \theta, r \sin \theta) - f(0, 0)| < \varepsilon$, allora f è continua in $(0, 0)$.

(ii) Discutere tale risultato in connessione con l'Es 1.9.

Esercizio 1.11 Sia $f(x, y) = \begin{cases} \frac{x^2 y}{x^4 + y^2}, & \text{se } (x, y) \neq (0, 0), \\ 0, & \text{se } (x, y) = (0, 0). \end{cases}$

(i) Dire se f è limitata o no sugli insiemi: $A := \{(x, y) \mid |(x, y)| \leq 1\}$ e $B := \{(x, y) \mid |(x, y)| \geq 1\}$.

(ii) Si calcoli, fissato $(a, b) \neq (0, 0)$, il $\lim_{t \rightarrow 0^+} f(ta, tb)$.

(iii) È continua f nell'origine?

Esercizio 1.12 Sia $f : (x, y) \in \mathbb{R}_+^2 \mapsto \frac{1}{x} + \frac{x}{y} + y \in \mathbb{R}$. Dimostrare che $\lim_{|(x, y)| \rightarrow +\infty} f(x, y) = +\infty$.

Esercizio 1.13 (i) Siano $x^{(0)}, x^{(1)}, \dots, x^{(k)}$ punti di \mathbb{R}^n . Scrivere esplicitamente il cammino che realizza la poligonale $P(x^{(0)}, \dots, x^{(k)})$.

(ii)* Sia $\{x^{(k)}\}$ una successione in \mathbb{R}^n con $x^{(k)} \rightarrow \bar{x}$. Mostrare che l'insieme $\Gamma := \bigcup_{k \geq 1} P(x^{(1)}, \dots, x^{(k)})$ è una curva di estremi $x^{(1)}$ e \bar{x} .

Esercizio 1.14* Sia $f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ e $\bar{x} \in \bar{E}$. Dimostrare che:

Se per ogni cammino $z : [a, b] \rightarrow E$ tale che $z(t_0) = \bar{x}$ per un qualche $a < t_0 < b$, la funzione $f \circ z$ è continua in t_0 è continua, allora f è continua in \bar{x} .

Esercizio 1.15 Si dimostri che se f e g sono due funzioni continue su $A \subseteq \mathbb{R}^n$ allora $\min\{f, g\}$ e $\max\{f, g\}$ sono funzioni continue su A .

Esercizio 1.16 Dimostrare (usando direttamente la Definizione 1.32) che la funzione $f : x \in \mathbb{R}^n \rightarrow |x|^2$ è uniformemente continua su $B_1 := \{x \in \mathbb{R}^n \mid |x| < 1\}$.

Esercizio 1.17 Sia $S^2 := \{x \in \mathbb{R}^3 \mid |x| = 1\}$, $\bar{x} := (2, 0, 0)$, $\bar{x} = (1, 0, 0)$, e si definisca $f(x) := x_1 x_2 (\sin |x - \bar{x}|)^{-1}$. Trovare δ tale che $|f(x) - f(\bar{x})| < \varepsilon$ per ogni $x \in S^2$ tale che $|x - \bar{x}| < \delta$.

Esercizio 1.18 Si dimostri direttamente che l'anello $A := \{x \in \mathbb{R}^2 \mid 1 \leq |x| < 2\}$ è connesso per poligonali, ossia, per ogni $x, y \in A$ si costruisca una poligonale che unisca x con y .

Esercizio 1.19 Si dimostri che se $E \subseteq \mathbb{R}^n$ è connesso per curve e $f \in C(E, \mathbb{R}^m)$, allora $f(E)$ è connesso per curve.

Esercizio 1.20 Sia E come nell'Esempio 1.39. Si dimostri che E è connesso per curve.

Esercizio 1.21* Si completino i dettagli relativi all'Esempio 1.45-(v).

Esercizio 1.22 Per $x \in \mathbb{R}^2$ sia $\|x\| := \sqrt{\frac{x_1^2}{100} + x_2^2}$.

(i) Dire se $\|\cdot\|$ definisce una norma su \mathbb{R}^2 ed in caso affermativo trovare due costanti positive a e b tali che $a|x|_\infty \leq \|x\| \leq b|x|_\infty$ per ogni $x \in \mathbb{R}^2$.

(ii) Calcolare il $\|\cdot\|$ -diametro del disco unitario $B^2 := \{x \in \mathbb{R}^2 \mid x_1^2 + x_2^2 \leq 1\}$, ossia, calcolare il $\sup_{x, y \in B^2} \|x - y\|$.

Esercizio 1.23 Si dimostrino le seguenti proprietà del diametro: (i) $\text{diam } A = 0$ se e solo se A è costituito da un solo punto. (ii) A è limitato se e solo il suo diametro è finito.

Esercizio 1.24 Si calcoli il diametro di $B := \{x \in \mathbb{R}^n \mid |x - x_0|_p < r\}$ per $1 \leq p \leq \infty$.

Esercizio 1.25 Sia $E \subseteq X$ l'insieme formato dai valori della successione in (1.47). Si mostri che E non è totalmente limitato.

Esercizio 1.26 Siano $\{v^{(1)}, \dots, v^{(n)}\}$ vettori in \mathbb{R}^n (o \mathbb{C}^n) e sia $A \in K^{n \times n}$. Dimostrare:

- (i) Se i vettori $v^{(i)}$ sono linearmente dipendenti, allora lo sono anche i vettori $Av^{(i)}$.
- (ii) Se A non è invertibile, allora i vettori $Av^{(i)}$ sono linearmente dipendenti.
- (iii) Dedurre da (i), da (ii) (e da (xxiii)) che se A o B sono matrici non invertibili, allora $\det(AB) = 0 = \det A \cdot \det B$.

Esercizio 1.27 (Il gruppo $SO(2)$). Sia $U \in SO(2)$.

- (i) Dimostrare che esiste $\theta \in [0, 2\pi)$ tale che

$$U = U(\theta) := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}. \quad (1.105)$$

- (ii) Dimostrare che se ad un punto $x \in \mathbb{R}^2$ facciamo corrispondere il numero complesso $z = x_1 + ix_2$, l'azione di U in (1.27) corrisponde alla moltiplicazione di z per $e^{i\theta}$.

[Più formalmente, sia $\varphi : x \in \mathbb{R}^2 \rightarrow z = \varphi(x) := x_1 + ix_2 \in \mathbb{C}$. Bisogna allora dimostrare che $\varphi(Ux) = e^{i\theta} \varphi(x)$.]

- (iii) Si dimostri che $U(\theta)^{-1} = U(-\theta)$.

Esercizio 1.28 (Il gruppo $SO(3)$). In virtù dell'esercizio precedente possiamo definire in \mathbb{R}^3 rotazioni di un angolo θ attorno ad un asse coordinato:

$$\begin{aligned} U^{(1)}(\theta) &:= \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cos \theta & -\sin \theta \\ 0 & \sin \theta & \cos \theta \end{pmatrix}, & \text{(rotazione di } \theta \text{ attorno a } e^{(1)}) \\ U^{(2)}(\theta) &:= \begin{pmatrix} \cos \theta & 0 & -\sin \theta \\ 0 & 1 & 0 \\ \sin \theta & 0 & \cos \theta \end{pmatrix}, & \text{(rotazione di } \theta \text{ attorno a } e^{(2)}) \\ U^{(3)}(\theta) &:= \begin{pmatrix} \cos \theta & -\sin \theta & 0 \\ \sin \theta & \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, & \text{(rotazione di } \theta \text{ attorno a } e^{(3)}) . \end{aligned} \quad (1.106)$$

- (i) Si dimostri che se $U \in SO(3)$ allora esistono $\theta, \varphi, \psi \in [0, 2\pi)$ tali che

$$U = U^{(3)}(\theta)U^{(1)}(\varphi)U^{(3)}(\psi). \quad (1.107)$$

- (ii) Si dimostri che $(U^{(i)}(\theta))^{-1} = U^{(i)}(-\theta)$.

Esercizio 1.29 Si dimostri che dato $x \neq 0$ in \mathbb{R}^n , il piano ortogonale a x , π_x , è un sottospazio vettoriale di \mathbb{R}^n di dimensione $(n - 1)$.

Esercizio 1.30* (Curva di Peano) Esiste una funzione continua $\varphi : [0, 1] \rightarrow \mathbb{R}^2$ tale che $\varphi([0, 1]) = Q := [0, 1] \times [0, 1]$, ossia, Q è una curva e φ è un cammino che la realizza.

Si completino i dettagli del seguente schema di dimostrazione:

1) Sia K un quadrato di lato di lunghezza ℓ e lo si suddivida in quattro quadrati di lato $\ell/2$. Si fissi un lato L di K e su di esso si fissi un punto M a distanza $\ell/4$ da uno dei vertici di L e si fissi un altro lato L' di K diverso da L . Si faccia vedere che esiste una ed una sola poligonale⁸⁷ di lunghezza 2ℓ , che passi per i centri dei quattro quadrati in cui è stato suddiviso K e con estremi M ed N dove N è un punto su L' a distanza $\ell/4$ da uno dei suoi due estremi.

2) Sia $a < b$. Si 'parametrizzi' su $[a, b]$ una qualunque poligonale in \mathbb{R}^n ossia, se gli estremi della poligonale hanno coordinate x e y , si trovi una funzione $\gamma : t \in [a, b] \rightarrow \gamma(t) \in \mathbb{R}^n$ tale che $\gamma(a) = x$, $\gamma(b) = y$ e tale che l'insieme $\{\gamma(t) : a \leq t \leq b\}$ coincida con la poligonale data. [Suggerimento: un segmento di estremi x e z è parametrizzato su $[a, b]$ da $\gamma(t) = x + \frac{t-a}{b-a}(z-x)$.]

3) La funzione (o 'curva') φ sarà ottenuta come limite (uniforme) di funzioni continue $\varphi_k : [0, 1] \rightarrow Q$. Ogni φ_k sarà una parametrizzazione di una poligonale chiusa che passa per ogni centro dei 2^{2k} quadrati di lato 2^{-k} in cui si divide Q . Denoteremo con Γ_k la poligonale parametrizzata da φ_k , cioè $\Gamma_k := \{\varphi_k(t) : 0 \leq t \leq 1\}$. *Primo passo:* si divida Q in quattro quadrati di lato $1/2$ e sia Γ_1 la poligonale formata dai quattro segmenti che uniscono i centri dei quattro quadrati. *Secondo passo:*

⁸⁷Una 'poligonale' in \mathbb{R}^n è l'unione di $N \geq 1$ segmenti aventi come estremi coppie di punti $(x^{(1)}, x^{(2)})$, $(x^{(2)}, x^{(3)})$, ..., $(x^{(N)}, x^{(N+1)})$.

Si consideri il quadrato Q_1^1 di lato $1/2$ di centro $(\frac{1}{4}, \frac{3}{4})$ e sia $\Gamma_1^1 := \Gamma_1 \cap Q_1^1$. Si divida Q_1^1 in quattro quadrati di lato $1/4$ (che chiameremo Q_2^1, \dots, Q_4^1) ed usando il punto 1) si costruisca una poligonale Γ_2^1 che passi per i quattro centri di Q_2^j con estremi a distanza $1/8$ dai due punti in $\Gamma_1^1 \cap Q_1^1$: vi sono infatti due poligonali con tali proprietà e per eliminare tale ambiguità si scelga quella con uno degli estremi in $(\frac{1}{2}, \frac{7}{8})$. Si divida Q in 16 quadrati di lato $1/4$, Q_2^j , $j = 1, \dots, 16$. Iterando il procedimento sopra descritto si costruisca la poligonale chiusa Γ_2 che passi per tutti i 16 centri dei Q_2^j (e per il punto $(\frac{1}{2}, \frac{7}{8})$: $\Gamma_2^1 = \Gamma_2 \cap Q_1^1$ etc.). *Terzo passo*: si generalizzi il procedimento descritto sopra e per ogni k si costruisca una poligonale chiusa, Γ_k , che passi per i 2^{2k} centri dei quadrati di lato 2^{-k} in cui è possibile suddividere Q . Tale costruzione è unica se si impone che la poligonale passi anche per il punto $(\frac{1}{2}, 1 - \frac{1}{2^{k+1}})$.

4) Si parametrizzi Γ_k tramite $\varphi_k(t)$ in modo tale che $\varphi_k(\frac{j}{2^{k+1}})$ coincida, per ogni $j = 1, \dots, 2^{2k}$, con il centro di Q_k^j (si noti che Γ_k dà un ordine ai quadrati di lato 2^{-k} per i cui centri passa).

5) Si dimostri che per ogni $0 \leq t \leq 1$ vale $|\varphi_k(t) - \varphi_{k+1}(t)| \leq \frac{1}{2^{k+1}}$ e da questo si deduca che $\varphi_k(t)$ converge uniformemente: il limite $\varphi(t)$ sarà dunque una funzione continua da $[0, 1]$ in Q (e $\varphi(0) = (\frac{1}{2}, 1)$).

6) Si dimostri che φ è surgettiva su Q cioè $\varphi([0, 1]) = Q$.

7) φ è anche iniettiva? [Risposta: no]

Esercizio 1.31 Siano f e f_k come nell'Esempio 1.52. Si dimostri che $\sup_{[0,1]} |f_k - f| = 1$ per ogni k .

Esercizio 1.32 Dimostrare che valgono le seguenti due affermazioni: (i) Se $u_n \in C([a, b])$ e $\sum u_n$ converge uniformemente in (a, b) , allora $\sum u_n$ converge uniformemente in $[a, b]$.

(ii) Se $u_n \in C([a, b])$, $\sum u_n$ converge in (a, b) ma $\sum u_n(a)$ non converge, allora $\sum u_n$ non converge uniformemente in (a, b) .

Esercizio 1.33 Studiare la convergenza delle seguenti serie di funzioni di x , (ossia si trovino i più grandi insiemi dove le serie convergono puntualmente, uniformemente e totalmente) al variare, qualora appaia, del parametro reale α :

$$\begin{aligned} (1) \quad & \sum_{n=0}^{\infty} e^{-\alpha n} x^n, & (2) \quad & \sum_{n=1}^{\infty} \frac{x^{\alpha n}}{n^x}, & (3) \quad & \sum_{n=2}^{\infty} \frac{(x \operatorname{sen} n)^n}{1 + n^2 x}, \\ (4) \quad & \sum_{n=1}^{\infty} \frac{(xn)^n}{x + n!}, & (5) \quad & \sum_{n=1}^{\infty} u_n(x) \text{ con } u_n(x) := \left(\sum_{j=1}^n j^x \right)^{-1}, \\ (6) \quad & \sum_{n=2}^{\infty} \frac{n^x}{(\log n)^\alpha}, & (7) \quad & \sum_{n=1}^{\infty} \log \left[1 + \arctan \left(\frac{x}{n} \right)^\alpha \right], \\ (8) \quad & \sum_{n=1}^{\infty} \frac{(-1)^n}{n^x}, & (9) \quad & \sum_{n=0}^{\infty} e^{nx} \frac{1}{1 + \operatorname{sen}^2 nx}, & (10) \quad & \sum_{n=1}^{\infty} (-1)^n e^{-(\log n)^x}. \end{aligned}$$

Esercizio 1.34 Si considerino le serie dell'esercizio precedente e si dia, per ogni x nell'insieme di convergenza, una stima superiore semplice di $|\sum u_n(x)|$ in termini di x .

Esempio: la serie $\sum_{n \geq 1} \frac{\cos x^n}{1+x^n}$ ha come insieme di convergenza $E := \{x : |x| > 1\}$. La convergenza è puntuale ed assoluta in E mentre è totale in ogni insieme della forma $(-\infty, a] \cup [b, \infty)$ con $a < -1$, $b > 1$. Per $x \in E$ si ha

$$\begin{aligned} \left| \sum_{n=1}^{\infty} \frac{\cos x^n}{1+x^n} \right| & \leq \sum_{n=1}^{\infty} \frac{1}{|1+x^n|} \leq \sum_{n=1}^{\infty} \frac{1}{|x|^n - 1} \\ & \leq \sum_{n=1}^{\infty} \frac{1}{|x|^n} \frac{1}{1 - \frac{1}{|x|}} = \frac{|x|}{(|x| - 1)^2}. \end{aligned}$$

Esercizio 1.35 (i) Siano $\sum u_n$ e $\sum v_n$ due serie di funzioni convergenti uniformemente in E . Si dimostri che, per ogni $a, b \in \mathbb{R}$, la serie $\sum (au_n + bv_n)$ converge uniformemente in E a $a \sum u_n + b \sum v_n$.

(ii) Siano f_n e g_n due successioni uniformemente convergenti in $[a, b]$ a, rispettivamente, f e g . Si dimostri che $f_n g_n$ converge uniformemente a fg .

Esercizio 1.36 Sia $u_0 = x$, e, per $n \geq 1$, $u_n = (x-1)x^n$, $x \in [0, 1]$. Dimostrare *direttamente* che la serie $\sum u_n$ non converge né totalmente, né uniformemente.

Esercizio 1.37 Sia $u_0 := u_1 := 0$ e, per $n \geq 2$, si definisca la funzione $u_n(x)$ come segue: sia

$$I_n := \left[\frac{1}{n} - \frac{1}{n^4}, \frac{1}{n} + \frac{1}{n^4}\right], \quad I'_n := \left[\frac{1}{n} - \frac{1}{n^4}, \frac{1}{n}\right], \quad I''_n := \left[\frac{1}{n}, \frac{1}{n} + \frac{1}{n^4}\right].$$

Sia ora $u_n = 0$ se $x \notin I_n$; in I'_n , u_n coincide con la retta passante per $(\frac{1}{n} - \frac{1}{n^4}, 0)$ e per $(\frac{1}{n}, \frac{1}{n})$ mentre nell'intervallo I''_n u_n coincide con la retta passante per $(\frac{1}{n} + \frac{1}{n^4}, 0)$ e per $(\frac{1}{n}, \frac{1}{n})$.

Si dimostri che $u_n \in C(\mathbb{R})$ e che la serie $\sum u_n$ converge uniformemente su \mathbb{R} ma non vi converge totalmente.

Esercizio 1.38 Sia f_n la funzione continua che vale 0 se $x \notin (0, \frac{1}{n})$, per $x = \frac{1}{2n}$ vale 1 e coincide con una retta negli intervalli $(0, \frac{1}{2n})$ e $(\frac{1}{2n}, \frac{1}{n})$.

Dimostrare le seguenti affermazioni:

(i) $\lim_{n \rightarrow \infty} f_n(x) = 0$ per ogni x ; (ii) f_n non converge uniformemente su $[0, 1]$; (iii) $\lim \int_0^1 f_n = \int_0^1 \lim f_n = 0$.

Esercizio 1.39 Trovare una successione di funzioni continue $f_n \in C([0, 1])$ convergente puntualmente ad una funzione continua $f \in C([0, 1])$ ma tale che $\lim \int_0^1 f_n \neq \int_0^1 f$.

Esercizio 1.40 Usando le proprietà di derivazione delle serie uniformemente convergenti, si calcoli il valore delle seguenti serie:

$$\sum_{n=1}^{\infty} \frac{n^3}{2^n}, \quad \sum_{n=1}^{\infty} n^3 x^n$$

Esercizio 1.41 Trovare delle formule ricorsive per calcolare

$$\sum_{n=1}^{\infty} n^p x^n.$$

Esercizio 1.42 Per ogni intero $N \geq 2$, calcolare

$$\sum_{k=1}^{N-1} (N-k)2^{k-1}.$$

Esercizio 1.43 Si dimostrino le seguenti identità

$$\sum_{k=0}^{\infty} (1-x)^k = \frac{1}{1-x},$$

$$\sum_{k=n}^{\infty} k(k-1) \cdots (k-n+1)(1-x)^{k-n} = \frac{n!}{(1-x)^{n+1}}, \quad \forall 0 < x < 1, \forall n \geq 1.$$

Esercizio 1.44 Si studi la convergenza uniforme delle serie in (A.2)÷(A.15).

Esercizio 1.45 Si costruisca una successione di funzioni continue $\delta_n \in C(\mathbb{R})$ che verifichino le seguenti proprietà: (i) $\delta_n \geq 0$; (ii) $\delta_n(x) = 0$ se $x \notin [a_n, b_n]$ dove a_n e b_n sono successioni monotone che convergono a 0 rispettivamente da sinistra e da destra; (iii) $\int_{-\infty}^{\infty} \delta_n := \int_{a_n}^{b_n} \delta_n = 1$.

Esercizio 1.46 ('Delta approssimata') Una successione di funzioni $\{\delta_n(x)\}$ che verifichi le proprietà (i), (ii), (iii) di 1.45 prende il nome di *delta approssimata*. Dimostrare che se $\{\delta_n\}$ è una delta approssimata allora, per ogni funzione continua $f \in C([a_0, b_0])$ si ha

$$\lim_{n \rightarrow \infty} \int_{-\infty}^{\infty} \delta_n(x) f(x) dx = f(0). \quad (1.108)$$

Esercizio 1.47 Si dia un esempio di una serie di funzioni C^∞ tale che $\sum u'_n$ converge uniformemente su tutto \mathbb{R} ma tale che $\sum u_n(x)$ non converge per nessun $x \in \mathbb{R}$.

Esercizio 1.48 Sia E la chiusura di un aperto limitato A e sia, per $n \geq 0$, $f_n \in C(E) \cap C^1(A)$. Si dimostri che se $\sup_{\substack{x \in (a,b) \\ n \geq 0}} |f'_n(x)| < \infty$ allora $\{f_n\}$ è una successione equicontinua su E .

Esercizio 1.49 Sia $f_n(x) := \cos(x \operatorname{sen} n)$.

- (i) Dimostrare che $\{f_n\}$ è equicontinua su $[-a, a]$ per ogni $a > 0$.
(ii) Trovare infinite successioni $\{n_k^{(i)}\}_{k \geq 0}$ tali che $\lim_{k \rightarrow \infty} f_{n_k^{(i)}} = f^{(i)}$ uniformemente in $[-a, a]$ con $f^{(i)} \neq f^{(j)}$ se $i \neq j$.

Esercizio 1.50 (Operatori alle differenze finite) Sia E un aperto di \mathbb{R} e sia $u \in C^1(E)$.

- (i) Dimostrare che per ogni x, a, b tali che $[a+x, b+x] \subseteq E$ si ha

$$\frac{d}{dx} \int_a^b u(x+t) dt = \int_a^b u'(x+t) dt. \quad (1.109)$$

- (ii) Fissiamo tre numeri α, β, r tali che $0 < r < \beta - \alpha$ e tali che $[\alpha - r, \beta + r] \subseteq E$. Definiamo ora un 'operatore' D_h che, dato un numero $h \neq 0$ con $|h| \leq r$, associ alla funzione u la funzione $D_h u$ che assume in $x \in [\alpha, \beta]$ il valore

$$D_h u(x) := \frac{u(x+h) - u(x)}{h}. \quad (1.110)$$

L'operatore D_h si chiama **operatore alle differenze finite**. L'analogia con l'operatore di derivata, $D := \frac{d}{dx}$, è evidente: dalla definizione di derivata e dalle ipotesi su u segue che

$$\lim_{h \rightarrow 0} D_h u(x) = u'(x) := Du(x), \quad \forall x \in [\alpha, \beta]. \quad (1.111)$$

- (iii) L'osservazione fatta al punto (ii) si generalizza ad ordine n come segue. Assumiamo $u \in C^n(E)$ e definiamo iterativamente l'**operatore alle differenze finite di ordine p** , per $1 \leq p \leq n$ e per $|h| \leq \frac{r}{n}$, come

$$D_h^p u(x) := D_h(D_h^{p-1} u)(x); \quad (1.112)$$

cioè

$$\begin{aligned} D_h^2 u(x) &= \frac{u(x+2h) - 2u(x+h) + u(x)}{h^2}, \\ D_h^3 u(x) &= \frac{u(x+3h) - 3u(x+2h) + 3u(x+h) - u(x)}{h^3}, \end{aligned} \quad (1.113)$$

e così via. Dimostrare che

$$\lim_{h \rightarrow 0} D_h^p u(x) = u^{(p)}(x), \quad \forall 1 \leq p \leq n, \quad \forall x \in [\alpha, \beta]. \quad (1.114)$$

- (iv) Si dimostri che

$$D_h^p u(x) = \frac{1}{h^p} \sum_{j=0}^p \binom{p}{j} (-1)^j u(x + (p-j)h). \quad (1.115)$$

Esercizio 1.51 Calcolare $\binom{\pm \frac{1}{2}}{k}$.

Esercizio 1.52 Dimostrare che $\limsup_{k \rightarrow \infty} \left| \binom{\alpha}{k} \right|^{\frac{1}{k}} = 1$, per ogni $\alpha \in \mathbb{R} \setminus \mathbb{Z}$.

Esercizio 1.53 Sia

$$b(x) := \frac{x}{e^x - 1}. \quad (1.116)$$

- (i) Si dimostri che $b(x)$ è analitica in $x = 0$.
(ii) Si definiscano i numeri di Bernoulli B_n come in (A.19) e si calcolino i primi cinque B_n .
(iii) Si dimostri l'identità

$$b(2x) + x = x \operatorname{cotanh} x. \quad (1.117)$$

- (iv) Dimostrare che $B_{2k+1} = 0$ per ogni $k \geq 1$.

Esercizio 1.54 (i) Usando la (1.117) si dimostri la (A.23).

(ii) Si dimostrino le identità elementari

$$\begin{aligned} \sinh 2x &= 2 \sinh x \cosh x, & \cosh 2x &= 2 \cosh^2 x - 1, \\ \cosh^2 x - \sinh^2 x &= 1, & 2 \operatorname{cotanh} 2x - \operatorname{cotanh} x &= \tanh x. \end{aligned}$$

(iii) Usando l'ultima relazione nel punto (ii) si dimostri la (A.22).

(iv) Si dimostrino le identità elementari

$$\tanh x = -i \tan ix, \quad \operatorname{cotanh} x = i \cotan ix. \quad (1.118)$$

Usando il punto (iv) si dimostrino (A.20) e (A.21).

Esercizio 1.55 Si consideri una serie di potenze $u = \sum_{n \geq 1} a_n x^n$ con $a_1 \neq 0$. Assumendo che la funzione inversa di u , u^{-1} sia espandibile in serie di potenze attorno a 0, e cioè che $u^{-1} := v = \sum b_n x^n$, si determinino i coefficienti b_n .

Esercizio 1.56 Sia $u = \frac{x}{1-x}$ e si calcolino i coefficienti della serie di potenze di $u^n := \sum_{k \geq n} a_k^{(n)} x^k$.

Esercizio 1.57 Sia $u = \frac{x}{1-x}$. La funzione inversa è data da $x = v(y) = \frac{y}{1+y} = \sum_{n \geq 1} b_n y^n$ con $b_n = (-1)^{n-1}$.

(i) Usando 1.55 e 1.56, si dimostri che

$$b_1 = 1, \quad b_k = - \sum_{n=1}^{k-1} b_n \binom{k-1}{n-1}, \quad (1.119)$$

ossia che

$$(-1)^{k-1} = - \sum_{n=1}^{k-1} (-1)^{n-1} \binom{k-1}{n-1}. \quad (1.120)$$

(ii) Si definiscano i numeri β_k come in (1.119) omettendo il segno meno, ossia si ponga

$$\beta_1 = 1, \quad \beta_k = \sum_{n=1}^{k-1} \beta_n \binom{k-1}{n-1}. \quad (1.121)$$

Si dimostri che $\beta_k \geq (k-1)!$ e che quindi la serie di potenze $\sum \beta_k y^k$ ha raggio di convergenza nullo.

Esercizio 1.58 Dare un sempio di una serie di potenze u tale che $\rho(u) = 1$ e $\rho(\frac{1}{u}) = \infty$.

Esercizio 1.59 Si dimostri che $(\ell^\infty, \|\cdot\|_\infty)$ (cfr. (1.44)) è uno spazio di Banach.

Esercizio 1.60 Si dimostri che $\ell^q \subsetneq \ell^p$, per ogni $1 \leq q < p \leq \infty$.

Esercizio 1.61 (i) Si dimostri che su ℓ^1 le norme $\|\cdot\|_1$, $\|\cdot\|_2$ e $\|\cdot\|_\infty$ non sono equivalenti tra loro.

(ii) Si dimostri che su ℓ^2 le norme $\|\cdot\|_2$ e $\|\cdot\|_\infty$ non sono equivalenti.

(iii) Si dimostri che per ogni $p > q$, $\|\cdot\|_p$ è una norma su ℓ^q e che le norme $\|\cdot\|_q$ e $\|\cdot\|_p$ non sono equivalenti su ℓ^q .

Esercizio 1.62 Dimostrare che i Teoremi di Weierstrass e Heine–Cantor in \mathbb{R}^n (con le dimostrazioni date) si estendono al caso di arbitrari spazi metrici.

Esercizio 1.63 (Completamento canonico di spazi metrici)

Due spazi metrici (X, d) e (X', d') si dicono **isometrici** se esiste una funzione biunivoca⁸⁸ $j : X \rightarrow X'$ tale che $d(x, y) = d'(j(x), j(y))$, $\forall x, y \in X$.

Sia (X, d) uno spazio metrico non vuoto e non completo. Sia Y la famiglia di tutte le successioni di Cauchy in X : $Y := \{\{x_k\} \mid x_k \in X \text{ e } \{x_k\} \text{ è di Cauchy}\}$. Si introduca su Y la seguente relazione:

$$\{x_k\} \sim \{y_k\} \iff \lim_{k \rightarrow \infty} d(x_k, y_k) = 0.$$

Si dimostri che \sim è una relazione d'equivalenza su Y .

⁸⁸Ossia, iniettiva e suriettiva.

Si denoti, ora, con $[\{x_k\}]$ la classe di equivalenza generata da $\{x_k\} \in Y$ e $X' := \{\xi = [\{x_k\}] \mid \{x_k\} \in Y\}$. Per ogni $\xi, \eta \in X'$, definiamo $d'(\xi, \eta) = \lim d(x_k, y_k)$, dove $\{x_k\} \in \xi$ e $\{y_k\} \in \eta$. Si dimostrino le seguenti affermazioni:

- (i) la definizione di d' è ben posta (ossia, non dipende dai particolari rappresentanti scelti in ξ e η) ed, inoltre, d' è una metrica su X' ;
- (ii) (X', d') è uno spazio metrico completo;
- (iii) $j : x \in X \mapsto [\{x\}]$, dove $\{x\}$ è la successione identicamente costante $\{x_k\}$ con $x_k := x$ per ogni k , è una funzione iniettiva che conserva la metrica: $d'(j(x), j(y)) = d(x, y)$, per ogni $x, y \in X$;
- (iv) $j(X)$ è denso in X' (ossia, $\overline{j(X)} = X'$);
- (v) X' è unico a meno di isometrie, più precisamente: se (X'', d'') è uno spazio metrico completo ed esiste una funzione iniettiva $j : X \rightarrow X''$ che conserva la metrica e tale che $j(X)$ è denso in X'' , allora X' e X'' sono isometrici.

Altri esercizi su spazi di Banach

Lo spazio $\text{Lip}(E, \mathbb{R}^m)$

Sia $E \subseteq \mathbb{R}^n$ un insieme compatto e si definisca il seguente sottoinsieme di $C(E, \mathbb{R}^m)$:

$$\text{Lip}(E, \mathbb{R}^m) := \left\{ f \in C(E, \mathbb{R}^m) : \|f\|_{\text{Lip}} := \sup_{x, y \in E, x \neq y} \frac{|f(x) - f(y)|}{|x - y|} + \sup_{x \in E} |f(x)| < \infty \right\}. \quad (1.122)$$

- (i) L'insieme $\text{Lip}(E, \mathbb{R}^m)$ è strettamente contenuto in $C(E, \mathbb{R}^m)$.
- (ii) $(\text{Lip}(E, \mathbb{R}^m), \|\cdot\|_{\text{Lip}})$ è uno spazio di Banach.

Si noti che se $f \in \text{Lip}(E, \mathbb{R}^m)$ allora f è **uniformemente lipschitziana** su E ovvero (per definizione) esiste una costante positiva C tale che

$$|f(x) - f(y)| \leq C|x - y|, \quad \forall x, y \in E, \quad (1.123)$$

inoltre la costante più piccola per cui vale (1.123) è data da $\sup_{x, y \in E, x \neq y} \frac{|f(x) - f(y)|}{|x - y|}$.

Spazi hölderiani $C^\alpha(E, \mathbb{R}^m)$ con $0 < \alpha < 1$

Sia $E \subseteq \mathbb{R}^n$ un insieme compatto e si definisca il seguente sottoinsieme di $C(E, \mathbb{R}^m)$:

$$C^\alpha(E, \mathbb{R}^m) := \left\{ f \in C(E, \mathbb{R}^m) : \|f\|_{C^\alpha} := \sup_{x, y \in E, x \neq y} \frac{|f(x) - f(y)|}{|x - y|^\alpha} + \sup_{x \in E} |f(x)| < \infty \right\}; \quad (1.124)$$

una funzione f tale che $\exists C > 0$ e $0 < \alpha < 1$ per cui $|f(x) - f(y)| \leq C|x - y|^\alpha$ per ogni $x, y \in A$ si dice **hölderiana** in A con esponente α .

- (i) L'insieme $C^\alpha(E, \mathbb{R}^m)$ è strettamente contenuto in $\text{Lip}(E, \mathbb{R}^m)$.
- (ii) $(C^\alpha(E, \mathbb{R}^m), \|\cdot\|_{C^\alpha})$ è uno spazio di Banach.

Gli spazi $C^k(E, \mathbb{R}^m)$

Sia $E \subseteq \mathbb{R}^n$ compatto e k un intero positivo. La classe $C^k(E, \mathbb{R}^m)$ è stata definita in § 2. Se $f \in C^k(E, \mathbb{R}^m)$, definiamo

$$\|f\|_{C^k} := \sum_{\beta \in \mathbb{N}, |\beta| \leq k} \sup_E |\partial_x^\beta f|. \quad (1.125)$$

- (i) $C^k(E, \mathbb{R}^m) \subsetneq C^h(E, \mathbb{R}^m) \subsetneq C^\alpha(E, \mathbb{R}^m)$, $\forall k > h \geq 1 > \alpha > 0$.

- (ii) $(C^k(E, \mathbb{R}^m), \|\cdot\|_{C^k})$ è uno spazio di Banach.

Spazi $C^\alpha(E, \mathbb{R}^m)$ con $\alpha > 0$

Sia E un compatto di \mathbb{R}^n ; sia α un numero positivo, sia $k := [\alpha]$ la sua parte intera ed $r := \{\alpha\}$ la sua parte frazionaria cosicché $\alpha = k + r$ con k intero non negativo e $0 \leq r < 1$. Se $k = 0$ o $r = 0$ gli spazi $C^\alpha(E, \mathbb{R}^m)$ sono già stati definiti. Se $k \geq 1$ e $r > 0$ poniamo

$$C^\alpha(E, \mathbb{R}^m) := \left\{ f \in C^k(E, \mathbb{R}^m) : \partial_x^\beta f \in C^r(E, \mathbb{R}^m), \forall \beta \in \mathbb{N}^n, |\beta| = k \right\},$$

(con $\alpha := k + r$, $k \in \mathbb{N}$, $0 < r < 1$).

(1.126)

Per $f \in C^\alpha(E, \mathbb{R}^m)$ [con α come in (1.126)] poniamo

$$\|f\|_{C^\alpha} := \|f\|_{C^k} + \sum_{\beta \in \mathbb{N}^n, |\beta|=k} \|\partial_x^\beta f\|_{C^r}.$$
(1.127)

Si dimostri che (anche in tal caso) $(C^\alpha(E, \mathbb{R}^m), \|\cdot\|_{C^\alpha})$ è uno spazio di Banach.

Lo spazio $C^\infty(E, \mathbb{R}^m)$

(i) (caso $n = m = 1$) Sia E un intervallo compatto di \mathbb{R} . Su $C^\infty(E, \mathbb{R})$ si consideri la seguente *metrica*

$$d(f, g) := \sum_{k \geq 0} 2^{-k} \frac{\|f^{(k)} - g^{(k)}\|_{\infty, E}}{1 + \|f^{(k)} - g^{(k)}\|_{\infty, E}}.$$
(1.128)

Dimostrare che (C^∞, d) è uno spazio metrico completo.

(ii) Sia E un compatto di \mathbb{R}^n . Su $C^\infty(E, \mathbb{R}^m)$ si consideri la seguente *metrica*:

$$d(f, g) := \sum_{k \geq 0} \frac{1}{2^k k!} \sum_{\beta \in \mathbb{N}^n, |\beta|=k} \frac{\|\partial_x^\beta f - \partial_x^\beta g\|_{\infty, E}}{1 + \|\partial_x^\beta f - \partial_x^\beta g\|_{\infty, E}}.$$
(1.129)

$(C^\infty(E, \mathbb{R}^m), d)$ è uno spazio metrico completo, ma non è uno spazio di Banach....

Si dimostri che gli spazi di successioni $\ell^p :=$ insieme delle successioni reali (o complesse) x tali che

$$\|x\|_p := \left(\sum_{i=0}^{\infty} |x_i|^p \right)^{\frac{1}{p}} < \infty \},$$

sono spazi di Banach per ogni $1 \leq p \leq \infty$.

Si dica qual è la relazione (come insiemi) di ℓ^p e ℓ^q con $1 \leq p < q \leq \infty$.

(i) Si dimostri che su ℓ^1 le norme $\|\cdot\|_1$, $\|\cdot\|_2$ e $\|\cdot\|_\infty$ non sono equivalenti tra loro; si dimostri che su ℓ^2 le norme $\|\cdot\|_2$ e $\|\cdot\|_\infty$ non sono equivalenti.

(ii) Si dimostri che per ogni $p > q$, $\|\cdot\|_p$ è una norma su ℓ^q e che le norme $\|\cdot\|_q$ e $\|\cdot\|_p$ non sono equivalenti su ℓ^q .

Capitolo 2

Differenziabilità

In questo capitolo studieremo la ‘regolarità’ delle funzioni di più variabili. Il punto di partenza è – come per la funzioni scalari – l’idea di approssimare localmente una funzione con una mappa lineare. Nelle prime tre sezioni considereremo solo funzioni da \mathbb{R}^n in \mathbb{R} .

1 Derivate e differenziali

Una funzione f di una variabile definita su un intervallo I di \mathbb{R} è derivabile in $x \in I$ se esiste il limite del rapporto incrementale

$$f'(x) := \lim_{t \rightarrow 0} \frac{f(x+t) - f(x)}{t}. \quad (2.1)$$

Nel generalizzare questa idea a \mathbb{R}^n è possibile considerare singolarmente gli n rapporti incrementali nelle n variabili indipendenti.

Per esempio, se consideriamo la funzione $f : (x, y) \in \mathbb{R}_+ \times \mathbb{R} \mapsto x^y \in \mathbb{R}$, possiamo considerare la funzione di una variabile $x \in (0, +\infty) \rightarrow f_1(x; y) := x^y$, avendo fissato l’esponente $y \in \mathbb{R}$ o considerare la funzione $y \in \mathbb{R} \rightarrow f_2(y; x) := x^y$, avendo fissato la base x ; entrambe le funzioni sono regolari e le loro derivate sono date da $f'_1(x; y) = yx^{y-1}$ e $f'_2(y; x) = (\log x) x^y$. Quest’idea di fare derivate ‘parziali’ risulterà essere molto utile, ma, se $n \geq 2$, non è ‘l’idea giusta’. Infatti, in una variabile l’esistenza della derivata dà la *la miglior approssimazione lineare di una funzione vicino ad un punto fissato*, essendo (2.1) equivalente a:

$$f(x+t) = f(x) + f'(x)t + r(t) \quad \text{con} \quad \lim_{t \rightarrow 0} \frac{r(t)}{t} = 0; \quad (2.2)$$

inoltre, come risulta immediato da (2.2), una funzione (di una variabile) derivabile in x è *continua* in x . Ma l’esistenza delle n ‘derivate parziali’ di una funzione di n variabili, in generale, non garantiscono neanche la continuità della funzione, come risulta considerando il seguente semplice

Esempio 2.1 Sia $P := \{(x, y) \in \mathbb{R}^2 \mid y = x^2, x \in \mathbb{R}, x \neq 0\}$, e sia $(x, y) \in \mathbb{R}^2 \mapsto f(x, y) := \chi_P(x, y)$ la funzione caratteristica¹ di P , ossia $f(x, y) = 1$ se $(x, y) \in P$, e $f(x, y) = 0$ se $(x, y) \notin P$. Per ogni $t \in \mathbb{R}$, $f(t, 0) = f(0, t) = 0$ e quindi

$$0 = \lim_{t \rightarrow 0} \frac{f(t, 0) - f(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{f(0, t) - f(0, 0)}{t},$$

ma, chiaramente, f non è continua in $(0, 0)$ (né, in un qualunque altro punto di \overline{P}).

¹In generale, se $E \subseteq \mathbb{R}^n$, la **funzione caratteristica** di E è la funzione χ_E tale che $\chi_E(x) = 1$ se $x \in E$ e $\chi_E(x) = 0$ se $x \notin E$.

Diamo, ora, le definizioni fondamentali legate alla nozione di derivazione di una funzione da \mathbb{R}^n in \mathbb{R} :

Definizione 2.2 Sia E un aperto di \mathbb{R}^n , $f : E \rightarrow \mathbb{R}$ e $\bar{x} \in E$.

(i) La **derivata parziale** di f rispetto a x_i (o ‘la i -esima derivata parziale’) nel punto $\bar{x} \in E$ è (qualora esista) il valore del limite²

$$D_i f(\bar{x}) := \lim_{t \rightarrow 0} \frac{f(\bar{x} + te^{(i)}) - f(\bar{x})}{t} = \lim_{t \rightarrow 0} \frac{f(\bar{x}_1, \dots, \bar{x}_i + t, \dots, \bar{x}_n) - f(\bar{x})}{t}; \quad (2.3)$$

l’elenco ordinato di tale derivate parziali (qualora esistano) forma un vettore $\nabla f(\bar{x}) \in \mathbb{R}^n$ che prende il nome di **gradiente** di f (nel punto \bar{x}).

(ii) Se $\xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n \setminus \{0\}$, la **derivata direzionale** di f rispetto a ξ nel punto \bar{x} è il valore del limite (qualora esista)

$$D_\xi f(\bar{x}) := \lim_{t \rightarrow 0} \frac{f(\bar{x} + t\xi) - f(\bar{x})}{t}; \quad (2.4)$$

(iii) la funzione f si dice **differenziabile** nel punto $\bar{x} \in \mathbb{R}^n$ se esiste un’applicazione lineare³ $L : h \in \mathbb{R}^n \mapsto Lh \in \mathbb{R}$ tale che

$$\lim_{h \rightarrow 0} \frac{f(\bar{x} + h) - f(\bar{x}) - Lh}{|h|} = 0; \quad (2.5)$$

in tal caso, L è unica⁴, si chiama il **differenziale** di f in \bar{x} e si denota con $df_{\bar{x}}$.

Le derivate parziali o direzionali di f vengono anche chiamate le **derivate prime** o ‘derivate di ordine 1’ di f .

Osservazione 2.3 (I simboli di Landau) La differenziabilità ha una formulazione standard in termini dei simboli di Landau (i cosiddetti ‘o piccoli’ e ‘O grandi’⁵).

Ricordiamo che:

Se $\bar{x} \in E \subseteq \mathbb{R}^n$ e $f, g : E \setminus \{\bar{x}\} \rightarrow \mathbb{R}^m$, si dice che $f = O(g)$ (‘ f è un O grande di g ’) vicino a \bar{x} se $\exists M > 0$ tale che $|f| \leq M|g|$ vicino a \bar{x} ;

$f = o(g)$ con $g \neq 0$ vicino a \bar{x} (‘ f è un o piccolo di g ’) vicino a \bar{x} se $\lim_{x \rightarrow \bar{x}} |f|/|g| = 0$;

infine, $f_1 = f_2 + o(g)$ o $f_1 = f_2 + O(g)$ significa, rispettivamente, $f = o(g)$ e $f = O(g)$ con $f := f_1 - f_2$.

In base a tali definizioni, possiamo dire che f è differenziabile in \bar{x} se e solo se esiste un’applicazione lineare L tale che

$$f(\bar{x} + h) = f(\bar{x}) + Lh + o(h). \quad (2.6)$$

Si noti che nella relazione (2.6), la funzione $o(h) := f(\bar{x} + h) - f(\bar{x}) - Lh$ è definita anche per $h = 0$ dove vale 0.

Esempio 2.4 (i) La funzione f dell’Esempio 2.1, pur non essendo continua in 0, ammette tutte le derivate direzionali in 0 e sono tutte nulle. Infatti sia $\xi := (\alpha, \beta) \in \mathbb{R}^2 \setminus \{0\}$. È chiaro che, se t è sufficientemente piccolo, si ha che⁶ $t\xi = (\alpha t, \beta t) \notin P$, e dunque, per t piccoli, $f(\alpha t, \beta t) - f(0, 0) = 0$, il che implica $D_\xi f(0, 0) = 0$, $\forall \xi \neq 0$.

² $\{e^{(i)}\}$ denota la base standard in \mathbb{R}^n ; cfr. (1.3).

³ Ossia, $L(ax + by) = aLx + bLy \forall x, y \in \mathbb{R}^n$ e $a, b \in \mathbb{R}$.

⁴ Se anche \tilde{L} soddisfa (2.5) (con \tilde{L} al posto di L), segue che $\lim_{h \rightarrow 0} \frac{(L - \tilde{L})h}{|h|} = 0$ e, prendendo $h = tx$ con $t \in \mathbb{R}$ e $x \in \mathbb{R}^n$ con $|x| = 1$, si ottiene $Lx = \tilde{L}x, \forall x \in \mathbb{R}^n$ con $|x| = 1$, ma essendo L e \tilde{L} lineari, questo implica che $L = \tilde{L}$.

⁵ Cfr. [Landau, E. *Handbuch der Lehre von der Verteilung der Primzahlen*. Leipzig, Germany: Teubner, 1909. Reprinted by New York: Chelsea, 1953] o, anche, §2.1, Cap. 7 di [C2019]

⁶ Se $\alpha = 0$, $t\xi = (0, \beta t) \notin P$ per ogni t ; se $\beta = 0$, $t\xi = (\alpha t, 0) \notin P$ per ogni t ; se $\alpha \neq 0 \neq \beta$, $t\xi = (\alpha t, \beta t) \notin P$ se $|t| < |\beta|/\alpha^2$.

(ii) Vediamo un esempio di funzione differenziabile.

Consideriamo la funzione $f : x \in \mathbb{R}^2 \mapsto f(x) = x_1 + \text{sen}(x_1x_2) \in \mathbb{R}$. Poiché $\text{sen } t = t + O(t^3)$ per t piccoli, si ha, per $|x|$ piccoli, $f(x) = x_1 + x_1x_2 + O(|x|^3) = x_1 + O(|x|^2)$ (essendo $|x_1x_2| \leq |x|^2$). Dunque la ‘parte lineare’ di f è data da $x \rightarrow x_1$. In base a queste osservazioni, se definiamo L come la mappa lineare: $h = (h_1, h_2) \in \mathbb{R}^2 \rightarrow Lh := h_1$, si trova, per $h \neq 0$:

$$\frac{|f(h) - f(0) - h_1|}{|h|} = \frac{|\text{sen}(h_1h_2)|}{|h|} \leq \frac{|h_1| |h_2|}{|h|} \leq \frac{|h|^2}{|h|} = |h| \rightarrow 0, \text{ per } h \rightarrow 0,$$

il che mostra che f è differenziabile in $\bar{x} = (0, 0)$ e che⁷ $df_0(h) = h_1$.

Osservazione 2.5 (i) Dalle definizioni (2.3) e (2.4) segue che la i -ma derivata parziale non è altro che la derivata direzionale rispetto all’ i -esimo versore $e^{(i)}$:

$$D_i f(\bar{x}) = D_{e^{(i)}} f(\bar{x}). \quad (2.7)$$

(ii) Una applicazione lineare $L : \mathbb{R}^n \rightarrow \mathbb{R}$ corrisponde alla moltiplicazione scalare per un dato vettore in \mathbb{R}^n : infatti, sia $\{e^{(i)}\}$ la base standard di \mathbb{R}^n come in (1.3) e sia ℓ il vettore definito da $\ell := (Le^{(1)}, \dots, Le^{(n)})$; allora, dalla linearità di L segue che

$$Lx = L \sum_{i=1}^n x_i e^{(i)} = \sum_{i=1}^n x_i L e^{(i)} =: \sum_{i=1}^n x_i \ell_i = \ell \cdot x. \quad (2.8)$$

Dunque, la definizione di differenziabilità può essere riformulata come segue:

$f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ è differenziabile in $\bar{x} \in E$ se e solo se esiste $\ell \in \mathbb{R}^n$ tale che

$$f(\bar{x} + h) = f(\bar{x}) + \ell \cdot h + r(h), \quad \lim_{h \rightarrow 0} \frac{r(h)}{|h|} = 0, \quad (2.9)$$

o, equivalentemente, usando i simboli di Landau (cfr. Osservazione 2.3),

$$f(\bar{x} + h) = f(\bar{x}) + \ell \cdot h + o(h), \quad (2.10)$$

e in tal caso ℓ è unico.

Si noti che il ‘resto’ $r(h)$ in (2.9) è definito implicitamente dalla prima uguaglianza in (2.9); inoltre, sia r che ℓ dipendono, naturalmente, anche da \bar{x} . Come già osservato⁸ $r(h)$ è definita anche per $h = 0$ dove vale 0.

(iii) Ovviamente, nel caso scalare $n = 1$, le tre nozioni introdotte nella Definizione 2.2 equivalgono a dire che f è derivabile in \bar{x} e che $D_1 f(\bar{x}) = f'(\bar{x}) = \nabla f(\bar{x})$, $D_\xi f(\bar{x}) = \xi f'(\bar{x})$ e $Lh = f'(\bar{x})h = df_{\bar{x}}(h)$.

(iv) *Notazioni:* Al posto del simbolo D spesso si usa il simbolo ∂ , nel qual caso, è più frequente l’uso di ∂_{x_i} rispetto a ∂_i ; sono anche frequenti (sempre per indicare la derivata parziale rispetto a x_i) le notazioni

$$\frac{\partial f}{\partial x_i}, \quad \partial_{x_i} f, \quad f_{x_i}; \quad (x_i \in \mathbb{R})$$

nel caso di funzioni di due o tre variabili, usando la notazione classica $f(x, y)$ o $f(x, y, z)$, le derivate parziali si denotano di solito con:

$$\frac{\partial f}{\partial x} = \partial_x f = f_x, \quad \frac{\partial f}{\partial y} = \partial_y f = f_y, \quad \frac{\partial f}{\partial z} = \partial_z f = f_z; \quad (x, y, z \in \mathbb{R}).$$

⁷Sebbene normalmente per mappe lineari non si usino le parentesi, indicando con Lh l’azione della mappa L sul vettore h , nel caso del differenziale df , per motivi estetici, si preferisce l’uso delle parentesi.

⁸Cfr. fine dell’Osservazione. 2.3.

Le derivate direzionali si possono anche denotare

$$\frac{\partial f}{\partial \xi}, \quad \partial_\xi f; \quad (\xi \in \mathbb{R}^n).$$

Infine, altre notazioni per il gradiente ∇f sono

$$\left(\frac{\partial f}{\partial x_1}, \dots, \frac{\partial f}{\partial x_n} \right), \quad \partial_x f, \quad f_x, \quad f', \quad \text{grad } f; \quad (x_i \in \mathbb{R}^n, x \in \mathbb{R}^n).$$

L'esistenza di tutte le derivate direzionali implica, ovviamente (cfr. (2.7)), l'esistenza delle derivate parziali (ma non implica la continuità, cfr. Esempio 2.4–(i)). Vediamo, ora, che la differenziabilità implica sia la continuità che l'esistenza di tutte le derivate direzionali:

Proposizione 2.6 *Sia E un aperto di \mathbb{R}^n e sia $f : E \rightarrow \mathbb{R}$ differenziabile in $\bar{x} \in E$. Allora:*

- (i) f è continua in \bar{x} ;
- (ii) f ha tutte le derivate direzionali in \bar{x} e per ogni vettore non nullo $\xi \in \mathbb{R}^n$ si ha:

$$D_\xi f(\bar{x}) = \xi \cdot \nabla f(\bar{x}) = df_{\bar{x}}(\xi). \quad (2.11)$$

Dimostrazione (i): Poiché $\ell \cdot h + o(h) = O(h)$, la continuità di f in \bar{x} deriva immediatamente dall'Osservazione 2.5–(ii), cfr. (2.10).

(ii): Sia $\xi \in \mathbb{R}^n \setminus \{0\}$ e $h = t\xi$. Allora, da (2.9) segue che

$$\begin{aligned} \frac{f(\bar{x} + t\xi) - f(\bar{x})}{t} &= \frac{\ell \cdot t\xi + r(t\xi)}{t} \\ &= \ell \cdot \xi + \frac{r(t\xi)}{|t\xi|} |\xi| \rightarrow \ell \cdot \xi \quad (\text{per } t \rightarrow 0). \end{aligned} \quad (2.12)$$

In particolare, se $\xi = e^{(i)}$, si ha⁹ $D_i f(\bar{x}) = \ell_i$ e quindi: $\ell = \nabla f(\bar{x})$ e $df_{\bar{x}}(\xi) = \nabla f(\bar{x}) \cdot \xi$. ■

Osservazione 2.7 (i) Dalla definizione di differenziabilità (e dalla linearità del limite) segue subito che se f e g sono differenziabili in x allora anche la combinazione lineare $af + bg$ (con a e b numeri reali) lo è e si ha:

$$d(af + bg)_x = a df_x + b dg_x. \quad (2.13)$$

(ii) (*Regola di Leibnitz*) Se f e g sono differenziabili in $x \in \mathbb{R}^n$ allora anche la funzioni fg lo è e si ha

$$d(fg)_x = g(x) df_x + f(x) dg_x. \quad (2.14)$$

Per dimostrare la (2.14) usiamo i simboli di Landau (cfr. nota 2.3). Sia: $\Delta f := f(x+h) - f(x)$ e $\Delta g := g(x+h) - g(x)$. Poiché f e g sono differenziabili in x segue che $\Delta f = df_x(h) + o(h)$, $\Delta g = dg_x(h) + o(h)$ (cfr. (2.9)). Naturalmente, si ha anche che $\Delta f = O(h)$ e $\Delta g = O(h)$ cosicché $\Delta f \Delta g = O(|h|^2) = o(h)$. Dunque:

$$\begin{aligned} f(x+h)g(x+h) - f(x)g(x) &= \Delta f \Delta g + g(x) \Delta f + f(x) \Delta g \\ &= g(x) df_x(h) + f(x) dg_x(h) + o(h), \end{aligned}$$

il che vuol dire (cfr. (2.9)) che fg è differenziabile in x e che il suo differenziale in x è dato da $g(x) df_x + f(x) dg_x$.

(iii) (*Derivate lungo curve o cammini*) La derivata direzionale $D_\xi f(\bar{x})$ è per definizione la derivata rispetto a t della funzione $t \mapsto f(\bar{x} + t\xi)$ in $t = 0$, ossia è 'la derivata di f lungo il cammino $t \mapsto \bar{x} + t\xi$ nel punto \bar{x} '. Dato un cammino

$$z : t \in [a, b] \mapsto z(t) := (z_1(t), \dots, z_n(t)) \in \mathbb{R}^n$$

⁹Cfr. Osservazione 2.5–(i).

diremo che il **cammino** z è **differenziabile** in $t_0 \in [a, b]$ se le n funzioni $t \mapsto z_i(t)$ sono derivabili in t_0 , ed in tal caso, porremo

$$z'(t_0) := (z'_1(t_0), \dots, z'_n(t_0)). \quad (2.15)$$

Allora vale la seguente affermazione, che generalizza la formula (2.11):

Sia $U \subseteq \mathbb{R}^n$ un intorno di $\bar{x} \in \mathbb{R}^n$ e sia $f : U \rightarrow \mathbb{R}$ differenziabile in \bar{x} . Sia $z : [a, b] \rightarrow U$ un cammino tale che $z(t_0) = \bar{x}$ e sia z differenziabile in t_0 . Allora, la funzione $t \in [a, b] \mapsto f \circ z(t)$ è derivabile in t_0 e si ha

$$(f \circ z)'(t_0) = \nabla f(\bar{x}) \cdot z'(t_0). \quad (2.16)$$

Infatti, se $\xi_0 := z'(t_0)$, dall'ipotesi di differenziabilità del cammino z segue che

$$\xi := \xi(t) := z(t) - z(t_0) = z(t) - \bar{x} = \xi_0 t + o(t).$$

Dunque, essendo f differenziabile in \bar{x} , si ha (per $t \rightarrow 0$)

$$\begin{aligned} \frac{(f \circ z)(t_0 + t) - f(\bar{x})}{t} &= \frac{f(\bar{x} + \xi) - f(\bar{x})}{t} = \frac{\nabla f(\bar{x}) \cdot \xi + o(\xi)}{t} \\ &= \nabla f(\bar{x}) \cdot \xi_0 + \frac{o(t)}{t} + \frac{o(\xi_0 t + o(t))}{t} \\ &\rightarrow \nabla f(\bar{x}) \cdot \xi_0 = \nabla f(\bar{x}) \cdot z'(t_0). \quad \blacksquare \end{aligned}$$

Concludiamo questa sezione con alcuni esempi.

Esempio 2.8 (i) La funzione norma (euclidea) di x , $x \in \mathbb{R}^n \mapsto |x|$, è differenziabile in $\mathbb{R}^n \setminus \{0\}$ ma non ammette derivate parziali in 0 dove è, come già visto, continua. Infatti, fissiamo $x \neq 0$ e osserviamo che se $t_0 > 0$, allora, per t piccoli, si ha $\sqrt{t_0 + t} = \sqrt{t_0} + \frac{1}{2\sqrt{t_0}} t + o(t)$. Dunque, ponendo $t_0 := |x|^2$, da (1.11), segue, per $|h|$ piccolo,

$$|x + h| = \sqrt{|x|^2 + |h|^2 + 2x \cdot h} = |x| + \frac{1}{2|x|}(2x \cdot h + |h|^2) + o(|h|) = |x| + \frac{x}{|x|} \cdot h + o(h),$$

il che mostra la differenziabilità della norma euclidea in $x \neq 0$ e che

$$\nabla |x| = \frac{x}{|x|} \iff D_i |x| = D_i \sqrt{\sum_{i=1}^n x_i^2} = \frac{x_i}{|x|}, \quad (2.17)$$

(relazione che può, naturalmente, essere verificata calcolando direttamente $D_i |x|$).

Si osservi, infine, che $\lim_{t \rightarrow 0} \frac{|te^{(i)}|}{t} = \lim_{t \rightarrow 0} \operatorname{sgn}(t)$ non esiste e quindi non esistono le derivate parziali in 0.

(ii) Una funzione differenziabile in 0 ma non continua in alcun intorno di 0:

$$x \in \mathbb{R}^2 \mapsto f(x) := \begin{cases} x_1^2, & \text{se } x_2 = 0, \\ 0, & \text{se } x_2 \neq 0. \end{cases} \quad (2.18)$$

Poiché $|f(x)| \leq x_1^2 \leq |x|^2$, si ha che $f(x) = o(x)$ e dunque, dall'Osservazione 2.5–(ii), segue che f è differenziabile in 0 e che $\nabla f(0) = 0$. Ma, per ogni $a \neq 0$, i punti $(a, 0)$ sono punti di discontinuità, poiché $f(a, 0) = a^2 \neq 0 = \lim_{k \rightarrow \infty} f(a, 1/k)$.

(iii) Una funzione $f \in C(\mathbb{R}^2)$ con tutte le derivate direzionali in 0 ma non differenziabile in 0:

$$x \in \mathbb{R}^2 \mapsto f(x) = \begin{cases} \frac{x_1^3 + x_2^3}{|x|^2}, & \text{se } x \neq 0, \\ 0, & \text{se } x = 0. \end{cases} \quad (2.19)$$

Per l'Osservazione 1.23–(iv), la funzione f è continua su $\mathbb{R}^2 \setminus \{0\}$. Inoltre f è omogenea di grado 1 e quindi, per la Proposizione 1.35, $f \in C(\mathbb{R}^2)$.

Ora, $\forall t \neq 0$ e $\forall \xi \in \mathbb{R}^2 \setminus \{0\}$ si ha che $f(t\xi) = tf(\xi)$; dunque

$$D_\xi f(0) := \lim_{t \rightarrow 0} \frac{f(t\xi) - f(0)}{t} = \lim_{t \rightarrow 0} \frac{f(t\xi)}{t} = f(\xi).$$

In particolare, $\nabla f(0,0) = (1,1)$ e, se $h \neq 0$,

$$g(h) := \frac{r(h)}{|h|} := \frac{f(h) - f(0) - \nabla f(0) \cdot h}{|h|} = \frac{h_1^3 + h_2^3 - (h_1 + h_2)|h|^2}{|h|^3}.$$

Poiché $g(1/k, 1/k) = (1 - \sqrt{2})/\sqrt{2}$ si ha che $g(h)$ non tende a 0 quando $h \rightarrow 0$, il che vuol dire che f non è differenziabile in 0.

Il teorema del differenziale totale

Un criterio semplice per capire quando una funzione sia differenziabile è dato dalla proposizione che segue.

Proposizione 2.9 (Teorema del differenziale totale)

Sia $E \subseteq \mathbb{R}^n$ aperto, $x \in E$ e $f \in C(E, \mathbb{R})$. Assumiamo che esistano le derivate parziali di f in E e che queste siano continue in x . Allora, f è differenziabile in x .

Dimostrazione Assumiamo $n \geq 2$ (per $n = 1$ la tesi è ovvia). Dalle ipotesi segue che per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che $B_\delta(x) \subseteq E$ e

$$|D_i f(y) - D_i f(x)| < \frac{\varepsilon}{n}, \quad \forall y \text{ t.c. } |y - x| < \delta, \quad \forall i. \quad (2.20)$$

Consideriamo dapprima il caso bidimensionale $n = 2$. Sia $\xi \in \mathbb{R}^2$ con $0 < |\xi| < \delta$, dal teorema del valor medio di Lagrange per funzioni di una variabile reale¹⁰ esistono $t_i \in (0, 1)$ tali che

$$\begin{aligned} f(x + \xi) - f(x) - \nabla f(x) \cdot \xi &= (f(x + \xi) - f(x_1 + \xi_1, x_2) - D_2 f(x) \xi_2) + (f(x_1 + \xi_1, x_2) - f(x) - D_1 f(x) \xi_1) \\ &= (D_2 f(x_1 + \xi_1, x_2 + t_2 \xi_2) - D_2 f(x)) \xi_2 + (D_1 f(x_1 + t_1 \xi_1, x_2) - D_1 f(x)) \xi_1. \end{aligned}$$

Da tale relazione, da (2.20), ed essendo $|\xi_i| \leq |\xi|$, si ha che

$$\begin{aligned} |f(x + \xi) - f(x) - \nabla f(x) \cdot \xi| &\leq |D_1 f(x_1 + t_1 \xi_1, x_2) - D_1 f(x)| |\xi_1| + |D_2 f(x_1 + \xi_1, x_2 + t_2 \xi_2) - D_2 f(x)| |\xi_2| \\ &\leq \frac{\varepsilon}{2} (|\xi_1| + |\xi_2|) \leq \varepsilon |\xi|, \end{aligned}$$

che è quanto volevasi dimostrare.

Nel caso generale (n arbitrario), se poniamo

$$\xi^{(i)} := (\xi_1, \dots, \xi_i, 0, \dots, 0), \quad \xi^{(0)} := 0,$$

si ha

$$f(x + \xi) - f(x) - \nabla f(x) \cdot \xi = \sum_{i=0}^{n-1} (f(x + \xi^{(i+1)}) - f(x + \xi^{(i)}) - D_i f(x) \xi_i),$$

e possiamo ripetere l'argomento usato nel caso $n = 2$. ■

Il 'viceversa' di questa proposizione è, in generale, falso, come mostra il seguente esempio.

¹⁰Se $f \in C([a, b], \mathbb{R})$ e derivabile in (a, b) , allora esiste $a < t < b$ tale che $f(b) - f(a) = f'(t)(b - a)$; cfr. Proposizione 7.26 in [C2019].

Esempio 2.10 Una funzione di n variabili differenziabile in 0 , che abbia derivate parziali continue in ogni punto con la sola eccezione di 0 ove le derivate parziali esistono ma non sono continue:

$$x \in \mathbb{R}^n \mapsto f(x) = \begin{cases} |x|^2 \operatorname{sen} \frac{1}{|x|}, & \text{se } x \neq 0, \\ 0, & \text{se } x = 0. \end{cases} \quad (2.21)$$

Dall'Esempio 2.8–(i) (e dall'algebra dei limiti) segue che $f \in C(\mathbb{R}^n)$ e che le derivate parziali esistono e sono continue in $\mathbb{R}^n \setminus \{0\}$. Le derivate direzionali in 0 sono nulle: se $\xi \neq 0$,

$$D_\xi f(0) := \lim_{t \rightarrow 0} \frac{f(t\xi)}{t} = \lim_{t \rightarrow 0} t^2 |\xi|^2 \operatorname{sen} \frac{1}{|t| |\xi|} = 0.$$

La funzione f è differenziabile in 0 :

$$\lim_{|h| \rightarrow 0} \frac{f(h) - f(0) - \nabla f(0) \cdot h}{|h|} = \lim_{h \rightarrow 0} |h| \operatorname{sen} \frac{1}{|h|} = 0.$$

D'altra parte le derivate parziali *non sono continue* in $x = 0$. Infatti, per la (2.17), si ha che, per $x \neq 0$,

$$D_i f(x) = 2x_i \operatorname{sen} \frac{1}{|x|} - \frac{x_i}{|x|} \cos \frac{1}{|x|},$$

e, mentre $\lim_{x \rightarrow 0} 2x_i \operatorname{sen}(1/|x|) = 0$, la funzione $\frac{x_i}{|x|} \cos \frac{1}{|x|}$ non ammette limite per $x \rightarrow 0$, come è facile verificare¹¹.

Una classe ‘naturale’ di funzioni differenziabili sono le funzioni C^1 :

Definizione 2.11 Sia E un aperto di \mathbb{R}^n . Una funzione $f \in C(E) := C(E, \mathbb{R})$ si dice **di classe C^1** su E se, per ogni $1 \leq i \leq n$ e per ogni $x \in E$, esiste la derivata parziale $D_i f(x)$ e se la funzione $D_i f \in C(E)$. L'insieme delle funzioni di classe C^1 su E si denota con $C^1(E) = C^1(E, \mathbb{R})$.

Dal teorema del differenziale totale segue che se f è di classe C^1 , allora f è differenziabile.

2 Derivate successive

2.1 Derivate seconde e matrice hessiana

Definizione 2.12 Sia $E \subseteq \mathbb{R}^n$ aperto e $f : E \rightarrow \mathbb{R}$. Supponiamo che, $\forall x \in E$, esista $D_i f(x)$.

(i) Se esiste la derivata parziale rispetto a x_j della funzione $x \in E \rightarrow D_i f$ nel punto $\bar{x} \in E$, chiamiamo $(D_j(D_i f))(\bar{x})$ la **derivata seconda** di f rispetto a x_i e x_j e denotiamo tale derivata con uno dei seguenti simboli

$$\partial_{x_j} \partial_{x_i} f(\bar{x}), \quad \frac{\partial^2 f}{\partial x_j \partial x_i}(\bar{x}), \quad D_{ij} f(\bar{x}), \quad f_{x_i x_j}(\bar{x}). \quad (2.22)$$

(ii) Se f ha tutte le derivate seconde in $\bar{x} \in E$, definiamo la **matrice hessiana**¹² di f nel punto \bar{x} la matrice $(n \times n)$ costituita dalle derivate seconde di f in \bar{x} :

$$(\mathbf{H}f)(\bar{x}) := f''(\bar{x}) := (D_{ij} f(\bar{x}))_{i,j=1,\dots,n}. \quad (2.23)$$

¹¹Si prenda, per esempio $x^{(k)} := ((2\pi k)^{-1}, 0, \dots, 0)$ e $y^{(k)} := ((2\pi k + \pi/2)^{-1}, 0, \dots, 0)$.

¹²Dal nome del matematico tedesco Ludwig Otto Hesse (1811–1874), cfr. https://en.wikipedia.org/wiki/Otto_Hesse.

Una domanda naturale è quando si ha che $D_{ij}f = D_{ji}f$, ossia quando sia possibile scambiare l'ordine nel calcolo di una derivata seconda. A questo proposito, sussiste la seguente fondamentale

Proposizione 2.13 (Lemma di Schwarz) *Sia $n \geq 2$, $E \subseteq \mathbb{R}^n$ un insieme aperto, $\bar{x} \in E$, $f : E \rightarrow \mathbb{R}$ una funzione che ammette le derivate prime $D_i f(x)$ e $D_j f(x)$ per ogni $x \in E$, con $i \neq j$. Assumiamo che esista anche la derivata seconda $D_j(D_i f)(x)$, per ogni $x \in E$, e che tale funzione sia continua in \bar{x} . Allora esiste anche $D_i(D_j f)(\bar{x})$ e si ha*

$$D_i(D_j f)(\bar{x}) = D_j(D_i f)(\bar{x}). \quad (2.24)$$

Dimostrazione Poiché le variabili x_k con k diverso da i e j non giocano nessun ruolo, possiamo direttamente considerare il caso di due variabili $f : (x, y) \in \mathbb{R}^2 \mapsto f(x, y) \in \mathbb{R}$ facendo corrispondere x a x_i , y a x_j e (\bar{x}, \bar{y}) a \bar{x} . Le ipotesi, in tal caso, diventano: esistono in E le derivate parziali prime f_x e f_y e la derivata seconda $\partial_y f_x$, la quale è continua in (\bar{x}, \bar{y}) . La tesi equivale a dimostrare che, fissato $\varepsilon > 0$, esiste $\delta > 0$ tale che, se $0 < |h| < \delta$, allora

$$\left| \frac{f_y(\bar{x} + h, \bar{y}) - f_y(\bar{x}, \bar{y})}{h} - \partial_y f_x(\bar{x}, \bar{y}) \right| \leq \varepsilon. \quad (2.25)$$

Sia $\delta > 0$ tale che $B := \{|x - \bar{x}| \leq \delta, |y - \bar{y}| \leq \delta\} \subseteq E$ e¹³

$$|\partial_y f_x(x, y) - \partial_y f_x(\bar{x}, \bar{y})| \leq \varepsilon, \quad \forall (x, y) \in B. \quad (2.26)$$

Fissiamo $h \in \mathbb{R}$ tale che $0 < |h| < \delta$. Consideriamo la funzione

$$k \in [-\delta, \delta] \setminus \{0\} \mapsto \theta(k) := \frac{(f(\bar{x} + h, \bar{y} + k) - f(\bar{x} + h, \bar{y})) - (f(\bar{x}, \bar{y} + k) - f(\bar{x}, \bar{y}))}{hk}. \quad (2.27)$$

Si noti che

$$\lim_{k \rightarrow 0} \theta(k) = \frac{f_y(\bar{x} + h, \bar{y}) - f_y(\bar{x}, \bar{y})}{h}. \quad (2.28)$$

Si fissi un k con $0 < |k| < \delta$ e si noti anche che

$$\theta(k) = \frac{g(\bar{x} + h) - g(\bar{x})}{h}, \quad \text{con } g(x) := \frac{f(x, \bar{y} + k) - f(x, \bar{y})}{k}. \quad (2.29)$$

Ora, per il teorema di Lagrange¹⁴ applicato alla funzione $x \rightarrow g(x)$ nell'intervallo di estremi \bar{x} e $\bar{x} + h$, si ha che esiste $t \in (0, 1)$ tale che

$$\theta(k) \stackrel{(2.29)}{=} \frac{g(\bar{x} + h) - g(\bar{x})}{h} = g'(\bar{x} + th) \stackrel{(2.29)}{=} \frac{f_x(\bar{x} + th, \bar{y} + k) - f_x(\bar{x} + th, \bar{y})}{k}. \quad (2.30)$$

Per ipotesi, la funzione $y \rightarrow f_x(\bar{x} + th, y)$ è derivabile sull'intervallo (chiuso) di estremi \bar{y} e $\bar{y} + k$ e quindi possiamo nuovamente applicare il teorema di Lagrange ottenendo da (2.30)

$$\theta(k) = \partial_y f_x(\bar{x} + th, \bar{y} + sk). \quad (2.31)$$

per un opportuno $s \in (0, 1)$. Dunque, in vista di (2.26), possiamo concludere che, per ogni $0 < |k| < \delta$,

$$|\theta(k) - \partial_y f_x(\bar{x}, \bar{y})| \leq \varepsilon.$$

Prendendo il limite per k che tende a 0 in tale relazione, ricordando (2.28), otteniamo la (2.25). ■

¹³Tale δ esiste perché E è aperto e $\partial_y f_x$ è continua in (\bar{x}, \bar{y}) .

¹⁴Cfr. nota 10.

Osservazione 2.14 Sia $E \subseteq \mathbb{R}^n$ un aperto e $f : E \rightarrow \mathbb{R}$. Se le ipotesi del Lemma di Schwarz valgono per ogni $i < j$, allora la matrice hessiana $f''(\bar{x})$ è simmetrica. In particolare, se tutte le derivate seconde di f esistono e sono continue si ha che $D_{ij}f(x) = D_{ji}f(x)$, per ogni $x \in E$, per ogni i, j .

Le ipotesi del Lemma di Schwarz sono ‘ottimali’, come mostra il seguente esempio.

Esempio 2.15 Una funzione $(x, y) \in \mathbb{R}^2 \rightarrow f(x, y) \in \mathbb{R}$ con derivate prime continue su \mathbb{R}^2 , con derivate seconde continue in $\mathbb{R}^2 \setminus \{(0, 0)\}$, con derivate seconde in $(0, 0)$ ma con $f_{yx}(0, 0) \neq f_{xy}(0, 0)$.

$$f(x, y) = \begin{cases} \frac{x^3y - y^3x}{x^2 + y^2}, & \text{se } (x, y) \neq (0, 0), \\ 0, & \text{se } (x, y) = (0, 0). \end{cases} \quad (2.32)$$

La funzione f è omogenea di grado 2 ed è quindi continua in \mathbb{R}^n (Proposizione 1.35). Per ogni $h \in \mathbb{R}$, $f(h, 0) = 0 = f(0, h)$ e dunque $f_x(0, 0) = f_y(0, 0) = 0$; inoltre per ogni $(x, y) \neq (0, 0)$,

$$\begin{aligned} f_x &= \frac{3x^2y - y^3}{x^2 + y^2} - 2x \frac{x^3y - y^3x}{(x^2 + y^2)^2}, \\ f_y &= \frac{x^3 - 3xy^2}{x^2 + y^2} - 2y \frac{x^3y - y^3x}{(x^2 + y^2)^2}, \end{aligned} \quad (2.33)$$

e tali funzioni sono omogenee di grado 1 e quindi tendono a zero per $|(x, y)| \rightarrow 0$, e quindi le derivate prime sono continue su \mathbb{R}^2 . Però

$$\frac{f_x(0, h)}{h} = -1, \quad \frac{f_y(h, 0)}{h} = 1, \quad (2.34)$$

da cui segue che $f_{xy}(0, 0) = -1 \neq 1 = f_{yx}(0, 0)$. D'altra parte, come è facile controllare, le derivate seconde miste sono funzioni omogenee di grado 0 (e non identicamente costanti) e quindi non sono continue in $(0, 0)$.

Le funzioni che ammettono derivate prime e derivate seconde continue si dicono di classe C^2 :

Definizione 2.16 Sia E un aperto di \mathbb{R}^n e $f \in C^1(E, \mathbb{R})$. Se, per ogni coppia di indici i, j , $1 \leq i, j \leq n$, e per ogni $x \in E$, esiste la derivata parziale $D_i(D_j f(x))$ e la funzione $D_i(D_j f) \in C(E)$ diremo che f è di classe C^2 su E . L'insieme delle funzioni di classe C^2 su E si denota con $C^2(E) = C^2(E, \mathbb{R})$.

Dal lemma di Schwarz segue che se $f \in C^2(E)$, allora, per ogni coppia di indici i, j , $1 \leq i, j \leq n$, e per ogni $x \in E$, $D_i(D_j f(x)) = D_j(D_i f(x))$, ossia $D_{ji}f = D_{ij}f$.

2.2 Punti critici

Come per funzioni di una variabile, le derivate prime e seconde danno informazione sui massimi e minimi locali di una funzione da \mathbb{R}^n in \mathbb{R} .

Definizione 2.17 Sia E un insieme aperto di \mathbb{R}^n .

(i) Sia $f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$. Un punto $\bar{x} \in E$ è un punto di **minimo locale** (o ‘relativo’) per f se esiste una sfera B di centro \bar{x} tale che

$$f(\bar{x}) \leq f(x), \quad \forall x \in E \cap B. \quad (2.35)$$

Se in (2.35) vale la disuguaglianza stretta per $x \neq \bar{x}$, si dice che \bar{x} è un minimo locale **stretto**. In modo analogo (scambiando il verso delle disuguaglianze) si definisce un **massimo locale**

(eventualmente, stretto).

(ii) Sia $f \in C^1(E, \mathbb{R})$. Un punto \bar{x} tale che $\nabla f(\bar{x}) = 0$, si chiama **punto critico** (o ‘punto stazionario’) per f .

Un punto critico che non sia un minimo o massimo locale viene detto **punto di sella**.

L’idea fondamentale nello studio locale di una funzione – che è un’idea ricorrente nello studio delle funzioni di più variabili (si pensi, per esempio, alla definizione di derivate parziali e derivate direzionali) – è quello di guardare il comportamento di f vicino a \bar{x} lungo rette di \mathbb{R}^n per \bar{x} ossia di guardare il comportamento della funzione di una variabile $t \rightarrow F(t) := f(\bar{x} + ty)$ con $y \in \mathbb{R}^n \setminus \{0\}$ e t piccoli.

Su tale idea si basa la seguente

Proposizione 2.18 Sia A un aperto convesso di \mathbb{R}^n .

(i) **(Teorema del valor medio e formula di Taylor al prim’ordine)**

Sia $f \in C^1(A, \mathbb{R})$, e siano $\bar{x}, \bar{x} + y \in A$. Allora, la funzione $t \in [0, 1] \mapsto F(t) := f(\bar{x} + ty) \in \mathbb{R}$ è derivabile e si ha:

$$F'(t) = \nabla f(\bar{x} + ty) \cdot y, \quad \forall t \in [0, 1], \quad (2.36)$$

e valgono i seguenti ‘teoremi del valor medio’¹⁵:

$$\begin{aligned} f(\bar{x} + y) - f(\bar{x}) &= \sum_{i=1}^n \left(\int_0^1 D_i f(\bar{x} + ty) dt \right) y_i \\ &= \left(\int_0^1 \nabla f(\bar{x} + ty) dt \right) \cdot y; \end{aligned} \quad (2.37)$$

$$f(\bar{x} + y) - f(\bar{x}) = \nabla f(\bar{x} + sy) \cdot y, \quad (2.38)$$

dove la (2.38) vale per un opportuno $s \in (0, 1)$. Se $\sigma = \{\bar{x} + ty \mid t \in [0, 1]\}$ denota il segmento che unisce \bar{x} e $\bar{x} + y$, si ha

$$|f(\bar{x} + y) - f(\bar{x})| \leq M |y|, \quad M := \max_{\sigma} |\nabla f(x)|. \quad (2.39)$$

Infine, vale la seguente ‘formula di Taylor al prim’ordine’:

$$f(\bar{x} + y) = f(\bar{x}) + \nabla f(\bar{x}) \cdot y + r_1(y; \bar{x}) \quad \text{con} \quad r_1(y; \bar{x}) = o(|y|). \quad (2.40)$$

(ii) **(Formula di Taylor al second’ordine)** Sia $f \in C^2(A, \mathbb{R})$, $\bar{x}, \bar{x} + y \in A$. Allora, la funzione $t \in [0, 1] \mapsto F(t) := f(\bar{x} + ty) \in \mathbb{R}$ è derivabile due volte e si ha¹⁶:

$$F''(t) = f''(\bar{x} + ty) y \cdot y = \sum_{i,j=1}^n D_{ij} f(\bar{x} + ty) y_i y_j, \quad \forall t \in [0, 1]. \quad (2.41)$$

Inoltre, vale la seguente ‘formula di Taylor al second’ordine’:

$$f(\bar{x} + y) = f(\bar{x}) + \nabla f(\bar{x}) \cdot y + \frac{1}{2} f''(\bar{x}) y \cdot y + r_2(y; \bar{x}) \quad (2.42)$$

dove $f''(\bar{x}) = (Hf)(\bar{x})$ è la matrice hessiana di f in \bar{x} e

$$\begin{aligned} r_2(y; \bar{x}) &:= \sum_{i,j=1}^n \left(\int_0^1 (1-t) (D_{ij} f(\bar{x} + ty) - D_{ij} f(\bar{x})) dt \right) y_i y_j \\ &= \left(\int_0^1 (1-t) (f''(\bar{x} + ty) - f''(\bar{x})) dt \right) y \cdot y \\ &= o(|y|^2). \end{aligned} \quad (2.43)$$

¹⁵Si ricorda che se $g \in C([a, b], \mathbb{R}^n)$, allora $\int_a^b g(t) dt := \left(\int_a^b g_1(t) dt, \dots, \int_a^b g_n(t) dt \right)$ e vale $\left| \int_a^b g(t) dt \right| \leq \int_a^b |g(t)| dt$; cfr. Complemento 1.3.

¹⁶ $\sum_{i,j=1}^n a_{ij}$ significa $\sum_{i=1}^n \sum_{j=1}^n a_{ij} = \sum_{j=1}^n \sum_{i=1}^n a_{ij}$.

Dimostrazione Se $f \in C^1(B)$, per il teorema del differenziale totale (Proposizione 2.9), f è differenziabile su B e quindi (2.36) segue da (2.11) (con $\bar{x} + ty$ al posto di \bar{x}):

$$\begin{aligned} F'(t) &= \lim_{s \rightarrow 0} \frac{F(t+s) - F(t)}{s} := \lim_{s \rightarrow 0} \frac{f(\bar{x} + ty + sy) - f(\bar{x} + ty)}{s} =: D_y f(\bar{x} + ty) \\ &\stackrel{(2.11)}{=} \nabla f(\bar{x} + ty) \cdot y. \end{aligned}$$

Le identità (2.37) e (2.38) seguono immediatamente dal teorema del valor medio di Lagrange e dal Teorema fondamentale del calcolo¹⁷.

La (2.39) segue da (2.38), dalla disuguaglianza di Cauchy e dal teorema di Weierstrass.

Per dimostrare la (2.40), osserviamo che, grazie a (2.37),

$$\begin{aligned} r_2(y; \bar{x}) &= f(\bar{x} + y) - f(\bar{x}) - \nabla f(\bar{x}) \cdot y \\ &= \left(\int_0^1 \nabla f(\bar{x} + ty) dt \right) \cdot y - \nabla f(\bar{x}) \cdot y \\ &= \left(\int_0^1 \nabla f(\bar{x} + ty - \nabla f(\bar{x})) dt \right) \cdot y. \end{aligned} \quad (2.44)$$

Ora, poiché $f \in C^1(A)$, ∇f è continuo in \bar{x} , quindi, fissato $\varepsilon > 0$ esiste $\delta > 0$ tale che $B_\delta(\bar{x}) \subseteq A$ e $|\nabla f(x) - \nabla f(\bar{x})| < \varepsilon$ per ogni $|x - \bar{x}| < \delta$. Da questo e dalla disuguaglianza di Cauchy, segue che, se $|y| < \delta$,

$$|r_2(y; \bar{x})| \leq \left| \int_0^1 \nabla f(\bar{x} + ty - \nabla f(\bar{x})) dt \right| |y| \leq \left(\int_0^1 |\nabla f(\bar{x} + ty - \nabla f(\bar{x}))| dt \right) |y| \leq \varepsilon |y|.$$

ossia, che $r_2 = o(|y|)$.

Se $f \in C^2(A)$, applicando due volte la (2.36) ed usando il lemma di Schwarz, otteniamo

$$\begin{aligned} F''(t) &\stackrel{(2.36)}{=} \left(\nabla f(\bar{x} + ty) \cdot y \right)' = \left(\sum_{i=1}^n D_i f(\bar{x} + ty) y_i \right)' = \sum_{i=1}^n (D_i f(\bar{x} + ty))' y_i \\ &\stackrel{(2.36)}{=} \sum_{i=1}^n \left(\sum_{j=1}^n D_{ij} f(\bar{x} + ty) y_j \right) y_i = \sum_{i,j=1}^n D_{ij} f(\bar{x} + ty) y_i y_j = f''(\bar{x} + ty) y \cdot y. \end{aligned}$$

Dimostriamo, ora, la formula di Taylor (2.42).

Poiché $F \in C^2([0, 1])$, si ha¹⁸

$$F(1) = F(0) + F'(0) + \int_0^1 (1-t) F''(t) dt, \quad (2.45)$$

che, in vista di (2.36) e (2.41), diviene¹⁹

$$\begin{aligned} f(\bar{x} + y) &= f(\bar{x}) + \nabla f(\bar{x}) \cdot y + \sum_{i,j=1}^n \left(\int_0^1 (1-t) D_{ij} f(\bar{x} + ty) dt \right) y_i y_j \\ &= f(\bar{x}) + \nabla f(\bar{x}) \cdot y + \frac{1}{2} f''(\bar{x}) y \cdot y + r_2(y; \bar{x}) \end{aligned} \quad (2.46)$$

con r_2 definito in²⁰ (2.43). Ora, poiché $|y_i y_j| \leq |y|^2$ per ogni i, j e $\sum_{i,j=1}^n 1 = n^2$, se definiamo

$$R := \max_{i,j} \sup_{t \in [0,1]} |D_{ij} f(\bar{x} + ty) - D_{ij} f(\bar{x})|$$

¹⁷Cfr., rispettivamente, Proposizione 7.26 e Corollario 8.25 in [C2019].

¹⁸Infatti, integrando per parti, $\int_0^1 (1-t) F''(t) dt = [(1-t) F'(t)]_0^1 + \int_0^1 F'(t) dt = -F'(0) + \int_0^1 F'(t) dt = F(1) - F(0) - F'(0)$, per il Teorema fondamentale del calcolo.

¹⁹Si noti che $f''(\bar{x}) y \cdot y = \sum_{i,j=1}^n D_{ij} f(\bar{x}) y_j y_i$.

²⁰Si noti che $\int_0^1 (1-t) dt = \frac{1}{2}$ e dunque $\frac{1}{2} f''(\bar{x}) y \cdot y := \sum_{i,j=1}^n D_{ij} f(\bar{x}) y_i y_j = \sum_{i,j=1}^n \left(\int_0^1 (1-t) D_{ij} f(\bar{x}) dt \right) y_i y_j$.

da (2.43) segue che

$$|r_2(y; \bar{x})| \leq |y|^2 \sum_{i,j=1}^n \sup_{t \in [0,1]} |D_{ij}f(\bar{x} + ty) - D_{ij}f(\bar{x})| \leq |y|^2 n^2 R. \quad (2.47)$$

Ma essendo $f \in C^2(A)$ segue che, per ogni $\varepsilon > 0$ esiste $0 < \delta_0 < \delta$ tale che se $|y| < \delta_0$, allora $\sup_{t \in [0,1]} |D_{ij}f(\bar{x} + ty) - D_{ij}f(\bar{x})| < \varepsilon/n^2$ per ogni i, j e, dunque, da (2.47) segue che $|r_2(y; \bar{x})| < \varepsilon|y|^2, \forall |y| < \delta_0$. ■

Un corollario immediato del Teorema del valor medio è:

Corollario 2.19 *Se A è un aperto connesso in \mathbb{R}^n , $f \in C^1(A, \mathbb{R})$ e $f'(x) := \nabla f(x) = 0 \forall x \in A$, allora f è costante su A .*

Dimostrazione Dalla Proposizione 1.43–(ii), segue che A è connesso per poligonalità. Dunque, basta far vedere che preso un qualunque segmento σ contenuto in A , il valore di f nei suoi estremi coincide. Ma questo segue immediatamente da (2.38) e dall'ipotesi $\nabla f = 0$. ■

Discutiamo, ora, il comportamento locale di una funzione vicino ad un punto critico.

Proposizione 2.20 *Sia E un aperto di \mathbb{R}^n .*

- (i) *Se $\bar{x} \in E$ è un punto di minimo (o massimo) locale per una funzione $f \in C^1(E)$, allora \bar{x} è un punto critico per f .*
- (ii) *Se $f \in C^2(E)$, \bar{x} è un punto critico per f e e^{21} $(Hf)(\bar{x}) = f''(\bar{x}) > 0$ (rispettivamente, $f''(\bar{x}) < 0$), allora f ha un minimo locale (rispettivamente, un 'massimo locale') stretto in \bar{x} .*
- (iii) *Se $f \in C^2(E)$, \bar{x} è un punto critico per f e $f''(\bar{x})$ ha un autovalore²² positivo ed uno negativo, allora \bar{x} è una sella per f .*

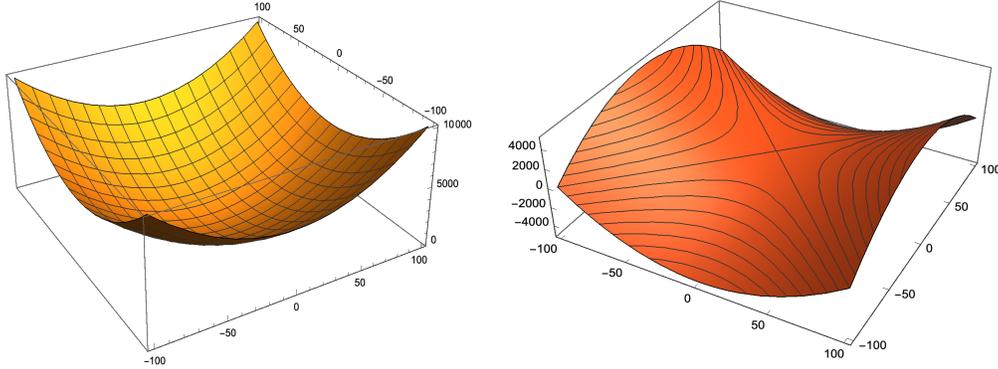


Figura 2.1: Un minimo locale e una sella

Dimostrazione Poiché E è aperto esiste $\delta > 0$ tale che $B = B_\delta(\bar{x}) \subseteq E$.

(i): Assumiamo che f abbia un minimo (o massimo) locale in \bar{x} . Allora, le n funzioni di una variabile $t \rightarrow F_i(t) := f(\bar{x} + te^{(i)})$ (definita per $|t| < \delta$) hanno un minimo (o massimo) locale in $t = 0$. Quindi, per il teorema di Fermat sui punti critici di funzioni di una variabile²³ $F'_i(0) = D_i f(\bar{x}) = 0$, per ogni i , ossia $\nabla f(\bar{x}) = 0$.

(ii): Dalle ipotesi e dalla formula di Taylor al second'ordine (Proposizione 2.18) segue che, per $|y| < \delta$,

$$f(\bar{x} + y) = f(\bar{x}) + \frac{1}{2} f''(\bar{x})y \cdot y + r_2(y; x), \quad \text{con } r_2(y; x) = o(|y|^2). \quad (2.48)$$

²¹Una matrice A reale e simmetrica ($n \times n$) si dice **definita positiva** (rispettivamente, **negativa**) se $Ay \cdot y > 0$ (rispettivamente, $Ay \cdot y < 0$), $\forall y \in \mathbb{R}^n \setminus \{0\}$; per richiami di algebra lineare, si veda l'Appendice D.

²² λ è un autovalore di A se $\exists y$ tale che $Ay = \lambda y$ o, equivalentemente se $\det(A - \lambda I) = 0$ dove I è la matrice identità ($n \times n$).

²³Proposizione 7.22 in [C2019].

Il polinomio di secondo grado in n variabili $y \in \mathbb{R}^n \mapsto Q(y) := \frac{1}{2}f''(\bar{x})y \cdot y$ è continuo sulla sfera unitaria $S^{n-1} = \{y \mid |y| = 1\}$, dove assume valori positivi (essendo $f''(\bar{x})$ una matrice definita positiva) e dunque $\min_{S^{n-1}} Q = \lambda_0 > 0$ (il minimo esiste per il teorema di Weierstrass). Per (2.43), esiste $0 < \delta_0 < \delta$ tale che $|r_2(y; x)|/|y|^2 \leq \lambda_0/2$, per ogni $0 < |y| < \delta_0$. Dunque:

$$\begin{aligned} f(\bar{x} + y) &\stackrel{(2.48)}{=} f(\bar{x}) + |y|^2 \left(\frac{1}{2}f''(\bar{x}) \frac{y}{|y|} \cdot \frac{y}{|y|} + \frac{r_2(y; x)}{|y|^2} \right) \\ &= f(\bar{x}) + |y|^2 \left(Q\left(\frac{y}{|y|}\right) + \frac{r_2(y; x)}{|y|^2} \right) \\ &\geq f(\bar{x}) + |y|^2 \frac{\lambda_0}{2} \\ &> f(\bar{x}), \quad \forall 0 < |y| < \delta_0. \end{aligned}$$

Nel caso $f''(\bar{x}) < 0$, la tesi segue da quanto appena dimostrato applicato alla funzione $-f$.

(iii): Siano $\lambda > 0 > \mu$ i due autovalori di $H := f''(\bar{x})$ menzionati nell'ipotesi e siano v e w due corrispondenti autovettori di norma 1: $Hv = \lambda v$ e $Hw = \mu w$. Essendo \bar{x} un punto critico per f , dalla formula di Taylor al secondo ordine (2.42) con $y = tv$, per t sufficientemente piccoli, si ha²⁴

$$f(\bar{x} + tv) = f(\bar{x}) + \frac{1}{2}\lambda t^2|v|^2 + o(t^2) = f(\bar{x}) + t^2 \left(\frac{\lambda}{2}|v|^2 + \frac{o(t^2)}{t^2} \right) > f(\bar{x})$$

Dunque, sulla retta $\{y = \bar{x} + tv \mid t \in \mathbb{R}\}$ vi sono punti y arbitrariamente vicini a \bar{x} tali che $f(y) > f(\bar{x})$. Analogamente, si vede che sulla retta $\{y = \bar{x} + tw \mid t \in \mathbb{R}\}$ vi sono punti y arbitrariamente vicini a \bar{x} tali che $f(y) < f(\bar{x})$ e quindi \bar{x} è una sella per f . ■

Esempio 2.21 Sia $f : (x, y) \in \mathbb{R}_+^2 \mapsto \frac{1}{x} + \frac{x}{y} + y \in \mathbb{R}$. Vogliamo calcolare $\sup f$ e $\inf f$ e determinare, se esistono, massimi/minimi (globali o locali) di f .

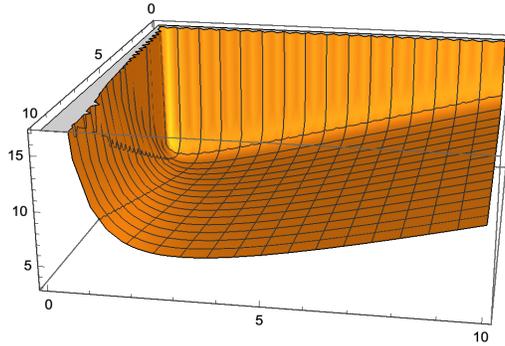


Figura 2.2: Grafico di $f : (x, y) \in \mathbb{R}_+^2 \mapsto \frac{1}{x} + \frac{x}{y} + y$

Osserviamo che $f(x, y) > 0, \forall (x, y) \in \mathbb{R}_+^2$. È, poi, chiaro che, per ogni $y > 0$ $\lim_{x \rightarrow 0^+} f(x, y) = +\infty$, quindi²⁵ $\sup f = +\infty$ e f non ha massimo. Cerchiamo i punti critici. Calcolando le derivate parziali, si ha

$$f_x(x, y) = -\frac{1}{x^2} + \frac{1}{y}, \quad f_y(x, y) = -\frac{x}{y^2} + 1.$$

²⁴Se $y = tv$, $y \cdot Hy = t^2 v \cdot Hv = t^2 \lambda v \cdot v = \lambda t^2 |v|^2$ e $o(|y|^2) = o(t^2)$.

²⁵In effetti, non è difficile mostrare che si ha anche che $\lim_{|(x, y)| \rightarrow +\infty} f(x, y) = +\infty$; cfr. Es 1.12.

Quindi $\nabla f(x, y) = 0$ se e solo se $y = x^2$ e $y^2 = x$. Tali relazioni individuano due curve in \mathbb{R}_+^2 , ossia $\{(x, y) \in \mathbb{R}_+^2 \mid y = x^2\}$ e $\{(x, y) \in \mathbb{R}_+^2 \mid y = \sqrt{x}\}$. Tali curve hanno unica intersezione in $(1, 1)$. Quindi, l'unico punto critico di f è $(1, 1)$. Calcolando la matrice hessiana in tale punto, si trova

$$\begin{pmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{pmatrix} \Big|_{(x,y)=(1,1)} = \begin{pmatrix} \frac{2}{x^3} & -\frac{1}{y^2} \\ -\frac{1}{y^2} & \frac{2x}{y^3} \end{pmatrix} \Big|_{(x,y)=(1,1)} = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix} =: H$$

Poiché²⁶ $\det(H - \lambda I_2) = (2 - \lambda)^2 - 1$ si trova che gli autovalori di H sono dati dalle due soluzioni di $(2 - \lambda)^2 - 1 = 0$, ossia $\lambda_1 = 1$ e $\lambda_2 = 3$. Dunque, $H > 0$ e, dalla Proposizione 2.20–(ii), segue che $(1, 1)$ è un punto di minimo locale stretto per f . Poiché $(1, 1)$ è l'unico punto critico di f , $(1, 1)$ deve essere il punto di minimo globale. Dunque, $\inf_{\mathbb{R}_+^2} f = f(1, 1) = 3$.

2.3 Funzioni C^k

Definizione 2.22 Sia E un aperto di \mathbb{R}^n e $f : E \rightarrow \mathbb{R}$.

(i) Un **indice di derivazione** è un elemento di \mathbb{N}_0^n , ossia un vettore a n componenti intere non negative. Su \mathbb{N}_0^n introduciamo un ordine parziale ponendo, $\forall \alpha, \beta \in \mathbb{N}_0^n$, $\alpha \leq \beta$ se $\alpha_i \leq \beta_i$ per ogni $1 \leq i \leq n$; $\alpha < \beta$ significa $\alpha \leq \beta$ e $\alpha \neq \beta$. Se α è un indice di derivazione, chiamiamo

l'**ordine di α** , il numero intero non negativo $|\alpha|_1 := \sum_{i=1}^n \alpha_i$.

(ii) Dato $k \in \mathbb{N}$, diremo che $f : E \rightarrow \mathbb{R}$ è di classe C^k su E se, per ogni $\alpha \in \mathbb{N}_0^n$ con $|\alpha|_1 \leq k$, esistono funzioni continue $f_\alpha : E \rightarrow \mathbb{R}$ tali che, $f_0 = f$ e per ogni β con $|\beta|_1 < k$, f_β ha tutte le derivate parziali di ordine 1 continue su E e valgono le seguenti 'relazioni di compatibilità': se $\alpha = \beta + e^{(i)}$, allora $f_\alpha = D_i f_\beta$.

Le funzioni f_α per $|\alpha|_1 > 0$ si chiamano le **α -derivate parziali di f di ordine $|\alpha|_1$** e si denotano con uno dei seguenti simboli:

$$D^\alpha f, \quad \partial^\alpha f, \quad \partial_x^\alpha f, \quad \frac{\partial^{|\alpha|_1} f}{\partial x_1^{\alpha_1} \dots \partial x_n^{\alpha_n}}, \quad \alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}_0^n. \quad (2.49)$$

La funzione f è di classe C^∞ su E se f è di classe C^k su E per ogni $k \in \mathbb{N}$.

$C^0(E) := C^0(E, \mathbb{R}) := C(E, \mathbb{R})$ e, per $k \geq 1$, $C^k(E) := C^k(E, \mathbb{R})$ denota l'insieme delle funzioni di classe C^k su E .

Questa definizione richiede alcuni commenti.

Osservazione 2.23 (i) Nel caso $k = 1$, gli indici di derivazione di ordine 1 sono i versori $e^{(i)}$, si pone $f_{e^{(i)}} := D_i f$ e la definizione data qui coincide con la Definizione 2.11.

Nel caso $k = 2$, gli indici di derivazione di ordine 2 sono dati da $\alpha = e^{(i)} + e^{(j)}$ e, in tal caso, possiamo porre $f_\alpha := D_{ji} f := D_i(D_j f)$ e le relazioni di compatibilità sono vere per il lemma di Schwarz, essendo $D_i f_{e^{(j)}} = D_i(D_j f) = D_j(D_i f) = D_j f_{e^{(i)}}$. Anche in questo caso, dunque, la definizione data qui coincide con la Definizione 2.16.

(ii) In tutti i casi, la definizione è ben posta per il lemma di Schwarz. Infatti, in generale, se $|\alpha|_1 \geq 2$ ci possono essere più modi di rappresentare α come somma di $e^{(i)}$ con un $\beta \in \mathbb{N}_0^n$ con $|\beta|_1 = |\alpha|_1 - 1$. Per esempio, se $\alpha = (3, 2, 0, 1)$, ci sono tre modi: $\alpha = e^{(1)} + (2, 2, 0, 1) = e^{(2)} + (3, 1, 0, 1) = e^{(4)} + (3, 2, 0, 0)$; in generale, il numero dei modi di rappresentare α come somma di un versore $e^{(i)}$ con un $\beta \in \mathbb{N}_0^n$ con $|\beta|_1 = |\alpha|_1 - 1$ è dato dal numero di componenti non nulle di α .

Ora, se per un α si ha $\alpha = e^{(i)} + \beta = e^{(j)} + \gamma$ con $i \neq j$ e $\beta, \gamma \in \mathbb{N}_0^n$, allora $\alpha = e^{(i)} + e^{(j)} + \bar{\beta}$ con $\bar{\beta} := \beta - e^{(j)} = \gamma - e^{(i)}$ e dunque, per il lemma di Schwarz, si ha:

$$f_\alpha = D_i f_\beta = D_i f_{e^{(j)} + \bar{\beta}} = D_i(D_j f_{\bar{\beta}}) = D_j(D_i f_{\bar{\beta}}) = D_j(f_{e^{(i)} + \bar{\beta}}) = D_j f_\gamma,$$

²⁶ I_2 denota la matrice identità (2×2) .

il che mostra che le relazioni di compatibilità valgono *indipendentemente dal modo di rappresentare* α come somma di un versore con un $\beta \in \mathbb{N}_0^n$.

(iii) Dunque, se $\alpha = \sum_{j=1}^k e^{(i_j)}$ con $1 \leq i_j \leq n$, si ha che $f_\alpha = D_{j_1}(\cdots(D_{j_k}f)\cdots)$ e l'ordine delle derivate parziali non conta, e, in particolare, si ha

$$f_\alpha =: D^\alpha f = D_n^{\alpha_n}(D_{n-1}^{\alpha_{n-1}}(\cdots(D_1^{\alpha_1}f)\cdots)), \quad (2.50)$$

dove, naturalmente, D_j^k è definito, per $k \geq 2$, iterativamente: $D_j^k f := D_j(D_j^{k-1}f)$.

Esempio 2.24 Sia $f : x \in \mathbb{R}^4 \mapsto f(x) := x_1^2 x_2^5 \log(1 + x_4^2) \in \mathbb{R}$, $\bar{x} = (1, 1, \sqrt{2}, 0)$, $\alpha = (1, 3, 0, 1)$. Allora,

$$D^{(1,3,0,1)}f(\bar{x}) := \frac{\partial^5 f}{\partial x_1 \partial x_2^3 \partial x_4}(\bar{x}) = D_2^3 D_1 D_4 f|_{\bar{x}} = 240 \frac{x_1 x_2^2 x_4}{1 + x_4^2} \Big|_{x=(1,1,\sqrt{2},0)} = 240.$$

3 Differenziale di funzioni vettoriali

In questa sezione estendiamo la nozione di differenziabilità a funzioni vettoriali e dimostriamo che la composizione di funzioni differenziabili è differenziabile e che la composizione di funzioni C^k è C^k .

3.1 Differenziali e matrici jacobiane

Definizione 2.25 Sia E un aperto di \mathbb{R}^n , $f : E \rightarrow \mathbb{R}^m$. Diremo che f è **differenziabile** nel punto $\bar{x} \in E$ se esiste un'applicazione lineare $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tale che $f(\bar{x}+h) - f(\bar{x}) - Lh = o(h)$, cioè:

$$\lim_{h \rightarrow 0} \frac{f(\bar{x}+h) - f(\bar{x}) - Lh}{|h|} = 0, \quad (2.51)$$

in tal caso, L è unica²⁷, si chiama il **differenziale** di f in \bar{x} e si denota con $df_{\bar{x}}$.

Osservazione 2.26 (i) Come nel caso scalare²⁸, da (2.51) segue che se f è differenziabile in \bar{x} , allora $\lim_{h \rightarrow 0} f(\bar{x}+h) = f(\bar{x})$, ossia:

una funzione $f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ differenziabile in $\bar{x} \in E$ è ivi continua.

(ii) se $\xi \in \mathbb{R}^n \setminus \{0\}$, ponendo $h = t\xi$ in (2.51) e prendendo il limite per t che tende a 0, si ottiene

$$\lim_{t \rightarrow 0} \frac{f(\bar{x} + t\xi) - f(\bar{x})}{t} = df_{\bar{x}}(\xi). \quad (2.52)$$

(iii) Poiché l'esistenza del limite di una funzione vettoriale è equivalente all'esistenza dei limiti della componenti (Osservazione 1.23–(ii)), dalle Definizioni 2.2–(iii) e 2.25, segue immediatamente che

$f = (f_1, \dots, f_m) : E \rightarrow \mathbb{R}^m$ è differenziabile in $\bar{x} \in E$ se e solo se sono differenziabili in \bar{x} le m funzioni $f_i : E \rightarrow \mathbb{R}$.

(iv) Osservando che – fissate le basi standard $\{e^{(j)}\}$ in \mathbb{R}^n e $\{\hat{e}^{(i)}\}$ in \mathbb{R}^m – la matrice $J \in K^{m \times n}$ associata al differenziale $df_{\bar{x}}$ ha elementi di matrice³⁰

²⁷Cfr. nota 4.

²⁸Cfr. Proposizione 2.6.

²⁹Qui, usiamo simboli diversi per sottolineare $e^{(j)} \in \mathbb{R}^n$, mentre gli $\hat{e}^{(i)} \in \mathbb{R}^m$;

³⁰ Infatti, se $\xi = \sum_{j=1}^n \xi_j e^{(j)} \in \mathbb{R}^n$, la i -esima componente del vettore $df_{\bar{x}}(\xi)$ è data da

$$(df_{\bar{x}}(\xi)) \cdot \hat{e}^{(i)} = \left(df_{\bar{x}} \left(\sum_{j=1}^n \xi_j e^{(j)} \right) \right) \cdot \hat{e}^{(i)} = \sum_{j=1}^n \xi_j \left(df_{\bar{x}} \left(e^{(j)} \right) \right) \cdot \hat{e}^{(i)} = \sum_{j=1}^n J_{ij} \xi_j = (J\xi)_i.$$

$$J_{ij} = (df_{\bar{x}}(e^{(j)})) \cdot \hat{e}^{(i)}, \quad (1 \leq i \leq m, 1 \leq j \leq n), \quad (2.53)$$

vediamo che

$$\begin{aligned} J_{ij} &\stackrel{(2.53)}{=} df_{\bar{x}}(e^{(j)}) \cdot \hat{e}^{(i)} \\ &\stackrel{(2.52)}{=} \left(\lim_{t \rightarrow 0} \frac{f(\bar{x} + te^{(j)}) - f(\bar{x})}{t} \right) \cdot \hat{e}^{(i)} \\ &= \lim_{t \rightarrow 0} \frac{(f(\bar{x} + te^{(j)}) - f(\bar{x})) \cdot \hat{e}^{(i)}}{t} \\ &= \lim_{t \rightarrow 0} \frac{f_i(\bar{x} + te^{(j)}) - f_i(\bar{x})}{t} \\ &\stackrel{(2.3)}{=} \frac{\partial f_i}{\partial x_j}(\bar{x}). \end{aligned} \quad (2.54)$$

Definizione 2.27 Data una funzione $f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ differenziabile in $\bar{x} \in E$, la matrice $J := J_{f, \bar{x}} \in K^{m \times n}$, con elementi di matrice dati da

$$J_{ij} = (df_{\bar{x}}(e^{(j)})) \cdot \hat{e}^{(i)} = \frac{\partial f_i}{\partial x_j}(\bar{x}), \quad (1 \leq i \leq m, 1 \leq j \leq n) \quad (2.55)$$

prende il nome di **matrice jacobiana** (o ‘jacobiano’) di f nel punto \bar{x} . Tale matrice si denota anche con uno dei seguenti simboli

$$\frac{\partial f}{\partial x}(\bar{x}), \quad \partial_x f(\bar{x}), \quad f_x(\bar{x}), \quad f'(\bar{x}). \quad (2.56)$$

Il Teorema del valor medio e la formula di Taylor al prim’ordine per funzioni $f \in C^1(A, \mathbb{R})$ con A aperto convesso di \mathbb{R}^n (Proposizione 2.18–(i)) si generalizzano immediatamente a funzioni vettoriali $f \in C^1(A, \mathbb{R}^m)$:

Teorema 2.28 (Teorema del valor medio e formula di Taylor al prim’ordine per funzioni vettoriali)

Sia A un aperto convesso di \mathbb{R}^n e $f \in C^1(A, \mathbb{R}^m)$. Allora, per ogni $\bar{x}, \bar{x} + y \in A$ si ha³¹

$$f(\bar{x} + y) - f(\bar{x}) = \left(\int_0^1 f'(\bar{x} + ty) dt \right) y, \quad (2.57)$$

$$|f(\bar{x} + y) - f(\bar{x})| \leq M |y|, \quad M := \max_{\sigma} \|f'\|, \quad (2.58)$$

dove σ denota il segmento che unisce \bar{x} e $\bar{x} + y$. Inoltre,

$$f(\bar{x} + y) = f(\bar{x}) + f'(\bar{x})y + r_1(y; \bar{x}) \quad \text{con} \quad r_1(y; \bar{x}) = o(|y|). \quad (2.59)$$

Dimostrazione Per definizione, la (2.57) è equivalente a

$$f_j(\bar{x} + y) - f_j(\bar{x}) = \sum_{k=1}^n \left(\int_0^1 \frac{\partial f_j}{\partial x_k}(\bar{x} + ty) dt \right) y_k, \quad \forall 1 \leq j \leq m, \quad (2.60)$$

e tale relazione è conseguenza del Teorema del valor medio per funzioni scalari (Proposizione 2.18–(i)) applicato alle m funzioni $f_j \in C^1(A, \mathbb{R})$.

La (2.58) segue dalla definizione di norma di matrice e dalle proprietà degli integrali di matrici

³¹ Si ricorda che se $t \in [a, b] \mapsto G(t) = (G_{ij}(t)) \in C([a, b], \mathbb{R}^{n \times m})$, allora $:= \int_a^b G(t) dt$ è la matrice $(m \times n)$ con elementi dati da $\int_a^b G_{ij}(t) dt$ e vale $\left\| \int_a^b G(t) dt \right\| \leq \int_a^b \|G(t)\| dt$; cfr. Complemento 1.3.

(cfr. nota 31).

Analogamente, prendendo la j -esima componente della relazione vettoriale (2.59), si ha

$$\begin{aligned} f_j(\bar{x} + y) &= f_j(\bar{x}) + \sum_{k=1}^m \frac{\partial f_j}{\partial x_k}(\bar{x}) y_k + r_{1j}(y; \bar{x}) \\ &= f_j(\bar{x}) + \nabla f_j(\bar{x}) \cdot y + r_{1j}(y; \bar{x}), \quad 1 \leq j \leq m, \end{aligned}$$

e il fatto che $r_1(y; \bar{x}) = o(|y|)$ segue ora da (2.40) applicato alle m funzioni f_j . ■

3.2 Differenziali di funzioni composte. Regola della catena

Proposizione 2.29 (Regola della catena per differenziali) *Siano $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}^m$ due insiemi aperti, e siano $f : A \rightarrow B$ e $g : B \rightarrow \mathbb{R}^p$. Se f è differenziabile in $\bar{x} \in A$ e g è differenziabile in $\bar{y} := f(\bar{x}) \in B$, allora $g \circ f$ è differenziabile in \bar{x} e si ha*

$$d(g \circ f)_{\bar{x}} = dg_{\bar{y}} \circ df_{\bar{x}}, \quad (2.61)$$

$$J_{g \circ f, \bar{x}} = J_{g, \bar{y}} J_{f, \bar{x}}. \quad (2.62)$$

Dimostrazione Innanzitutto, osserviamo che la (2.62) è una conseguenza immediata della (2.61) e della definizione di matrice jacobiana come matrice associata al differenziale (rispetto alle basi standard).

Dimostriamo, ora, che $g \circ f$ è differenziabile in \bar{x} e che vale (2.61).

Dalle ipotesi segue che

$$f(\bar{x} + h) = f(\bar{x}) + Lh + \alpha(h) = \bar{y} + Lh + \alpha(h), \quad \lim_{h \rightarrow 0} \frac{\alpha(h)}{|h|} = 0, \quad (2.63)$$

$$g(\bar{y} + k) = g(\bar{y}) + Mk + \beta(k), \quad \lim_{k \rightarrow 0} \frac{\beta(k)}{|k|} = 0, \quad (2.64)$$

dove $L := df_{\bar{x}}$ e $M := dg_{\bar{y}}$ denotano i differenziali f e g .

A priori, la funzione β è definita vicino a 0, ma, poiché $\lim_{k \rightarrow 0} \beta = 0$ possiamo estenderla in $k = 0$, ponendo $\beta(0) := 0$, ottenendo una funzione continua in un intorno di 0.

Dalla definizioni date e dalla linearità di M , segue che

$$\begin{aligned} g \circ f(\bar{x} + h) &\stackrel{(2.63)}{=} g(\bar{y} + Lh + \alpha(h)) \\ &\stackrel{(2.64)}{=} g(\bar{y}) + (M \circ L)h + M\alpha(h) + \beta(Lh + \alpha(h)). \end{aligned} \quad (2.65)$$

Dimostrare che $g \circ f$ è differenziabile in \bar{x} e che vale (2.61) è, dunque, equivalente a mostrare che

$$\gamma(h) := M\alpha(h) + \beta(Lh + \alpha(h)) = o(h). \quad (2.66)$$

Fissiamo $0 < \varepsilon < 1$ e denotiamo con $\|L\|$ e $\|M\|$ le norme³², rispettivamente, di L e M . Dai limiti in (2.63) e (2.64) segue che esistono $\rho > 0$ e $0 < \delta < \rho(1 + \|L\|)^{-1}$ tali che

$$|\beta(k)| \leq \frac{\varepsilon}{2(1 + \|L\|)} |k|, \quad \forall 0 \leq |k| < \rho, \quad (2.67)$$

$$|\alpha(h)| < \frac{\varepsilon}{2(1 + \|M\|)} |h| < |h|, \quad \forall 0 < |h| < \delta. \quad (2.68)$$

Allora, se $0 < |h| < \delta$, si ha che³³

$$|M\alpha(h)| \leq \|M\| |\alpha(h)| \stackrel{(2.68)}{\leq} \|M\| \frac{\varepsilon}{2(1 + \|M\|)} |h| < \frac{\varepsilon}{2} |h|. \quad (2.69)$$

³²Per definizione, $\|L\| = \sup\{|Lh| : h \in \mathbb{R}^n, |h| = 1\}$ e $\|M\| = \sup\{|Mk| : k \in \mathbb{R}^m, |k| = 1\}$

³³Si ricordi che dalla definizione di norma di una applicazione lineare, segue che $|Mk| \leq \|M\| |k|$ per ogni $k \in \mathbb{R}^m$.

Ricordando che δ è anche tale che $(\|L\| + 1)\delta < \rho$, si ha:

$$|Lh + \alpha(h)| \leq \|L\| |h| + |\alpha(h)| \stackrel{(2.68)}{<} (\|L\| + 1) |h| < (\|L\| + 1) \delta < \rho.$$

Quindi, possiamo applicare la (2.67), ottenendo:

$$\begin{aligned} |\beta(Lh + \alpha(h))| &\stackrel{(2.67)}{<} \frac{\varepsilon}{2(1 + \|L\|)} |Lh + \alpha(h)| \\ &\stackrel{(2.68)}{<} \frac{\varepsilon}{2(1 + \|L\|)} (\|L\| |h| + |h|) \\ &= \frac{\varepsilon}{2}. \end{aligned} \tag{2.70}$$

Mettendo (2.69) e (2.70) assieme otteniamo che $|\gamma(h)| < \varepsilon|h|$, che implica la (2.66). \blacksquare

Osservazione 2.30 Si noti che la relazione in (2.62) può scriversi in una delle seguenti forme equivalenti:

$$\left\{ \begin{array}{l} (g \circ f)'(\bar{x}) = g'(\bar{y}) f'(\bar{x}), \\ \frac{\partial(g \circ f)}{\partial x}(\bar{x}) = \frac{\partial g}{\partial y}(\bar{y}) \frac{\partial f}{\partial x}(\bar{x}), \\ (g \circ f)_x(\bar{x}) = g_y(\bar{y}) f_x(\bar{x}), \\ \frac{\partial(g \circ f)_i}{\partial x_j}(\bar{x}) = \sum_{k=1}^m \frac{\partial g_i}{\partial y_k}(\bar{y}) \frac{\partial f_k}{\partial x_j}(\bar{x}), \quad (1 \leq i \leq p, 1 \leq j \leq n). \end{array} \right. \tag{2.71}$$

Ovviamente, essendo i prodotti in tali formule, prodotti di matrici, l'ordine è importante: a sinistra dell'uguaglianze c'è una matrice $(p \times n)$ e a destra un prodotto di una matrice $(p \times m)$ con una matrice $(m \times n)$.

3.3 Funzioni $C^k(E, \mathbb{R}^m)$

Estendiamo, qui, la definizione di funzione vettoriale di classe C^k e dimostriamo che la composizione di funzioni C^k è C^k .

Definizione 2.31 Sia E un aperto di \mathbb{R}^n , $f = (f_1, \dots, f_m) : E \rightarrow \mathbb{R}^m$ e $k \in \mathbb{N}_0 \cup \{\infty\}$. Diremo che f è di classe C^k se³⁴ $f_i \in C^k(E)$, per ogni $1 \leq i \leq m$. L'insieme delle funzioni a valori in \mathbb{R}^m di classe C^k si denota con $C^k(E, \mathbb{R}^m)$.

Osservazione 2.32 (i) Se $f \in C^1(E, \mathbb{R}^m)$, dal teorema del differenziale totale (Proposizione 2.9), segue che le funzioni f_i sono differenziabili su E e quindi, per l'Osservazione 2.26–(iii), f è differenziabile su E .

(ii) (Il caso speciale $n = 1$) Poiché in una dimensione 'derivabilità' e 'differenziabilità' coincidono, dire che $\varphi : t \in (a, b) \rightarrow (\varphi_1(t), \dots, \varphi_m(t)) \in E \subseteq \mathbb{R}^m$ è differenziabile in $t_0 \in (a, b)$ equivale a dire che le m funzioni φ_i sono derivabili in t_0 . Ora, se $g : E \rightarrow \mathbb{R}$ è differenziabile in $\bar{x} = \varphi(t_0)$, dalla Proposizione 2.29 e da (2.71) segue che la funzione $t \in (a, b) \rightarrow g \circ \varphi(t)$ è derivabile in t_0 e si ha

$$\frac{d(g \circ f)}{dt}(t_0) = \nabla g(\bar{y}) \cdot \varphi'(t_0) \tag{2.72}$$

dove $\varphi'(t_0)$ denota $(\varphi'_1(t_0), \dots, \varphi'_m(t_0))$.

³⁴Si ricordi la Definizione 2.22–(ii).

Per esempio, se $g : \mathbb{R}^2 \rightarrow \mathbb{R}$ è differenziabile in (x_0, y_0) e se $h : \mathbb{R} \rightarrow \mathbb{R}$ è derivabile in x_0 , con $h(x_0) = y_0$, allora la funzione di una variabile $x \rightarrow g(x, h(x))$ è derivabile in x_0 e

$$\frac{d}{dx} g(x, h(x))|_{x_0} = g_x(x_0, y_0) + h'(x_0)g_y(x_0, y_0).$$

Tale formula segue dalla (2.72) con $\varphi(t) := (x_0 + t, h(x_0 + t))$ e $t_0 = 0$.

(iii) Se $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}^m$ sono aperti, $f : A \rightarrow B$ è di classe C^1 e $g \in C^1(B, \mathbb{R}^p)$, allora, per il punto (i), f e g sono differenziabili, rispettivamente, su A e B . Dunque dalla Proposizione 2.29 segue che $g \circ f$ è differenziabile su A e vale (2.71) per ogni $\bar{x} \in A$. Ma poiché la somma e il prodotto di funzioni continue è continua, (2.71) implica che tutte le derivate parziali $\frac{\partial(g \circ f)_i}{\partial x_j}$

sono continue in A , dunque:

Se $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}^m$, $f \in C^1(A, \mathbb{R}^m)$, $f(A) \subseteq B$ e $g \in C^1(B, \mathbb{R}^p)$ allora $g \circ f \in C^1(A, \mathbb{R}^p)$.

Il punto (iii) della precedente osservazione si generalizza immediatamente. Vale infatti la seguente

Proposizione 2.33 *Sia $k \in \mathbb{N}_0 \cup \{\infty\}$. Se $A \subseteq \mathbb{R}^n$, $B \subseteq \mathbb{R}^m$, $f \in C^k(A, \mathbb{R}^m)$, $f(A) \subseteq B$ e $g \in C^k(B, \mathbb{R}^p)$ allora $g \circ f \in C^k(A, \mathbb{R}^p)$.*

Dimostrazione Dimostriamo l'asserto per induzione su k . I casi $k = 0, 1$ corrispondono, rispettivamente, all'Osservazione 1.23–(iv) e al punto (iii) dell'Osservazione 2.32. Sia $k \geq 2$, assumiamo l'asserto vero per $0, 1, \dots, k-1$ e siano f e g funzioni di classe C^k . Allora, $g \circ f \in C^1(A, \mathbb{R}^p)$ e, per il punto (i) dell'Osservazione 2.32, la funzione $g \circ f$ è differenziabile e vale (2.71) su A , ossia:

$$\frac{\partial(g \circ f)_i}{\partial x_j}(x) = \sum_{k=1}^m \frac{\partial g_i}{\partial y_k}(f(x)) \frac{\partial f_k}{\partial x_j}(x), \quad \forall x \in A, \forall i, j.$$

Ora, $\frac{\partial g_i}{\partial y_k}$ e f sono funzioni C^{k-1} e, dunque, per l'ipotesi induttiva, $\frac{\partial g_i}{\partial y_k} \circ f$ è di classe $C^{k-1}(A)$. Anche $\frac{\partial f_k}{\partial x_j}$ è una funzione $C^{k-1}(A)$. Poiché somme e prodotti di funzioni C^{k-1} sono C^{k-1} , segue che $\frac{\partial(g \circ f)_i}{\partial x_j} \in C^{k-1}(A)$, il che equivale a dire che $g \circ f \in C^k(A, \mathbb{R}^p)$. ■

Complementi

Complemento 2.1: Teorema di Eulero sulle funzioni omogenee

Discutiamo brevemente un paio di proprietà delle derivate delle funzioni omogenee.

Proposizione 2.34 (Teorema di Eulero) *Sia $f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$ una funzione positivamente omogenea di grado $\alpha \in \mathbb{R}$, sia $x \neq 0$ e sia $\nu := \frac{x}{|x|}$. Allora esiste $D_\nu f(x)$ e risulta*

$$|x|D_\nu f(x) = \alpha f(x). \quad (2.73)$$

Inoltre, se f è differenziabile in x , si ha

$$|x|D_\nu f(x) = x \cdot \nabla f(x) = \alpha f(x). \quad (2.74)$$

Dimostrazione Per definizione di derivata direzionale, si ha:

$$\begin{aligned} D_\nu f(x) &:= \lim_{t \rightarrow 0} \frac{f(x + t \frac{x}{|x|}) - f(x)}{t} = \lim_{t \rightarrow 0} \frac{f((1 + \frac{t}{|x|})x) - f(x)}{t} \\ &= f(x) \lim_{t \rightarrow 0} \frac{((1 + \frac{t}{|x|})^\alpha - 1)}{t} = \alpha \frac{f(x)}{|x|}. \end{aligned}$$

La (2.74) deriva dalla Proposizione 2.6 e (2.11). ■

Proposizione 2.35 Sia $f : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}$ una funzione positivamente omogenea di grado $\alpha \in \mathbb{R}$. Se f ha derivate parziali su $\mathbb{R}^n \setminus \{0\}$, allora, per ogni i , $D_i f$ risulta positivamente omogenea di grado $\alpha - 1$.

Dimostrazione Sia $t > 0$ e $x \neq 0$. Usando la definizione di derivata parziale

$$\begin{aligned} D_i f(tx) &:= \lim_{h \rightarrow 0} \frac{f(tx + he^{(i)}) - f(tx)}{h} = \lim_{h \rightarrow 0} t^\alpha \frac{f(x + \frac{h}{t} e^{(i)}) - f(x)}{h} \\ &= t^{\alpha-1} \lim_{\rho \rightarrow 0} \frac{f(x + \rho e^{(i)}) - f(x)}{\rho} = t^{\alpha-1} D_i f(x). \quad \text{Dove } \rho := \frac{h}{t}. \quad \blacksquare \end{aligned}$$

Complemento 2.2: Derivazione ‘sotto segno di integrale’

Integrando una funzione di due variabili x e t rispetto a t si ottiene una funzione della sola variabile x ; il problema che qui ci poniamo è: sotto quali ipotesi tale funzione è derivabile rispetto a x e qual è il valore di tale derivata.

Proposizione 2.36 Sia $-\infty < a < b < +\infty$, $x_0 \in \mathbb{R}$ e I un intorno di x_0 . Siano $f = f(x, t)$ e la sua derivata parziale rispetto a x , f_x , funzioni di classe $C(I \times [a, b])$. Allora la funzione

$$F(x) := \int_a^b f(x, t) dt \quad (2.75)$$

è derivabile in x_0 e si ha

$$F'(x_0) = \int_a^b f_x(x_0, t) dt. \quad (2.76)$$

Dimostrazione Senza perdita di generalità, possiamo assumere che $I := [x_0 - \delta, x_0 + \delta]$ per un qualche $\delta > 0$, cosicché f e f_x sono uniformemente continue su $I \times [a, b]$. Se $\Delta(h)$ denota il rapporto incrementale della funzione F in x_0 , dal teorema fondamentale del calcolo segue che:

$$\begin{aligned} \Delta(h) &= \frac{F(x_0 + h) - F(x_0)}{h} := \int_a^b \frac{f(x_0 + h, t) - f(x_0, t)}{h} dt \\ &= \int_a^b \left(\int_0^1 f_x(x_0 + sh, t) ds \right) dt. \end{aligned} \quad (2.77)$$

Per la uniforme continuità di f_x in $I \times [a, b]$, dato un qualunque $\varepsilon > 0$, esiste $0 < r < \delta$ tale che se $|(x, t) - (\bar{x}, \bar{t})| < r$, allora $|f_x(x, t) - f_x(\bar{x}, \bar{t})| < \varepsilon$. Quindi, se $|h| < r$, da (2.77) segue che

$$\begin{aligned} \left| \Delta(h) - \int_a^b f_x(x_0, t) dt \right| &= \left| \int_a^b \left(\int_0^1 [f_x(x_0 + sh, t) - f_x(x_0, t)] ds \right) dt \right| \\ &\leq \int_a^b \left(\int_0^1 |f_x(x_0 + sh, t) - f_x(x_0, t)| ds \right) dt \\ &\leq \varepsilon (b - a). \end{aligned} \quad (2.78)$$

Dall'arbitrarietà di ε segue l'asserto. \blacksquare

Naturalmente, questo risultato si estende immediatamente a funzioni $f(x, t)$ con $x \in \mathbb{R}^n$:

Proposizione 2.37 Siano $-\infty < a, b < +\infty$ e siano $f = f(x, t)$ e la sua derivata parziale rispetto a x_j , $D_j f$, funzioni di classe $C(U \times [a, b])$ con $U \subseteq \mathbb{R}^n$ intorno di x_0 . Allora la funzione

$$F(x) := \int_a^b f(x, t) dt \quad (2.79)$$

ammette derivata parziale rispetto a x_j in x_0 e si ha

$$D_j F(x_0) = \int_a^b D_j f(x_0, t) dt. \quad (2.80)$$

Esempio 2.38 Siano a, b ed f come nell'enunciato della Proposizione 2.36; siano $\alpha(t)$, $\beta(t)$ e $g(t)$ funzioni continue su $J = [t_0 - \varepsilon, t_0 + \varepsilon]$ tali che $a \leq \alpha(t), \beta(t) \leq b$, $g(t) \in I$ per ogni $t \in J$. Assumiamo che α e β siano derivabili in t_0 e siano $\alpha_0 = \alpha(t_0)$, $\beta(t_0) = \beta_0$ e $x_0 = g(t_0)$. Allora la funzione

$$F(t) := \int_{\alpha(t)}^{\beta(t)} f(g(t), s) ds$$

è derivabile in t_0 e si ha:

$$F'(t_0) = \beta'(t_0)f(x_0, \beta_0) - \alpha'(t_0)f(x_0, \alpha_0) + g'(t_0) \int_a^b f_x(x_0, s) ds. \quad (2.81)$$

Infatti, per il Teorema fondamentale del calcolo e per la Proposizione 2.36, la funzione

$$G(x, y, z) := \int_y^z f(x, s) ds$$

ha derivate parziali date da

$$G_x = \int_y^z f_x(x, s) ds, \quad G_y = -f(x, y), \quad G_z = f(x, z),$$

e, per le ipotesi fatte, tali funzioni sono continue su $D := I \times [a, b] \times [a, b]$. In particolare, per il Teorema del differenziale totale, $G \in C^1(D)$ e quindi, la (2.81) segue dalla (2.16) (con $n = 3$) con $G(x, y, z)$ al posto di f e $z(t) := (g(t), \alpha(t), \beta(t))$, essendo $F(t) = (G \circ z)(t)$. ■

Un'altra applicazione interessante della Proposizione 2.36 è il seguente risultato sulle 'primitive di ordine n ':

Proposizione 2.39 Sia I un intervallo di \mathbb{R} e $f \in C(I)$. Fissato t_0 , sia, per $n \in \mathbb{N}_0$, $F_n(t; t_0)$ la 'primitiva ennesima di f rispetto al punto base t_0 ', ossia

$$F_n(t; t_0) := \begin{cases} f(t), & \text{se } n = 0, \\ \int_{t_0}^t F_{n-1}(s; t_0) ds, & \text{se } n \geq 1. \end{cases} \quad (2.82)$$

Allora,

$$F_n(t; t_0) = \int_{t_0}^t \frac{(t-s)^{n-1}}{(n-1)!} f(s) ds, \quad \forall n \in \mathbb{N}. \quad (2.83)$$

Si noti che, per il Teorema fondamentale del calcolo, F_n verifica, per ogni $n \geq 1$:

$$F_n'(t; t_0) = F_{n-1}(t; t_0), \quad F_n^{(n)}(t; t_0) = f(t), \quad F_n(t_0, t_0) = 0. \quad (2.84)$$

Dimostrazione (della Proposizione 2.39). Per induzione su $n \in \mathbb{N}$: per $n = 1$, (2.83) vale per definizione di F_1 . Assumiamo (2.83) e dimostriamola con $(n+1)$ al posto di n . Definiamo

$$G(t) := F_{n+1}(t; t_0) - \int_{t_0}^t \frac{(t-s)^n}{n!} f(s) ds \stackrel{(2.82)}{=} \int_{t_0}^t F_n(s; t_0) ds - \int_{t_0}^t \frac{(t-s)^n}{n!} f(s) ds.$$

Dal Teorema fondamentale del calcolo e dalla Proposizione 2.36 (si veda anche l'Esempio 2.38), segue che

$$G'(t) = F_n(t; t_0) - \int_{t_0}^t \partial_t \frac{(t-s)^n}{n!} f(s) ds = F_n(t; t_0) - \int_{t_0}^t \frac{(t-s)^{n-1}}{(n-1)!} f(s) ds \stackrel{(2.83)}{=} 0.$$

Dunque G è costante su I ed essendo $G(t_0) = 0$ si ha che $G \equiv 0$ su I , ossia, vale la (2.83) con $(n+1)$ al posto di n . ■

Complemento 2.3: Formula di Taylor in \mathbb{R}^n

Per generalizzare la formula di Taylor (Proposizione 2.18) ad ordine arbitrario, abbiamo bisogno di introdurre alcune notazioni.

Definizione 2.40 (i) Se $\alpha \in \mathbb{N}_0^n$ e $y \in \mathbb{R}^n$, poniamo:

$$\alpha! := \alpha_1! \cdots \alpha_n! := \prod_{j=1}^n \alpha_j!, \quad y^\alpha := y_1^{\alpha_1} \cdots y_n^{\alpha_n} := \prod_{j=1}^n y_j^{\alpha_j}. \quad (2.85)$$

(ii) Sia $k \in \mathbb{N}_0$ e $f : B_\delta(\bar{x}) \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione che ammetta tutte le derivate parziali fino ad ordine k in \bar{x} . Chiamiamo **polinomio di Taylor di f in \bar{x} di ordine k** il polinomio di grado al più k in $y \in \mathbb{R}^n$ dato da:

$$\bar{x} \in \mathbb{R}^n \mapsto T_k(y; \bar{x}) := T_{f,k}(y; \bar{x}) := \sum_{|\alpha|_1 \leq k} \frac{D^\alpha f(\bar{x})}{\alpha!} y^\alpha, \quad (2.86)$$

dove, gli indici α sono, naturalmente, indici di derivazione, ossia elementi di \mathbb{N}_0^n (specifica che, normalmente, ometteremo).

Chiamiamo **resto di Taylor di ordine k in \bar{x}** la funzione

$$y \in B_\delta(\bar{x}) \mapsto R_k(y; \bar{x}) := R_{f,k}(y; \bar{x}) := f(\bar{x} + y) - T_k(y; \bar{x}). \quad (2.87)$$

Proposizione 2.41 (Formula di Taylor, Peano, Lagrange) Siano $\bar{x}, y \in \mathbb{R}^n$, $k \in \mathbb{N}$, $\delta > 0$, $f \in C^k(B)$ con $B := B_\delta(\bar{x})$ e $|y| < \delta$. Allora, si ha:

$$f(\bar{x} + y) = T_k(y; \bar{x}) + R_k(y; \bar{x}), \quad \text{con } R_k(y; \bar{x}) = o(|y|^k). \quad (2.88)$$

Inoltre, per il resto di ordine $k - 1$ valgono le seguenti formule

$$R_{k-1}(y; \bar{x}) = \sum_{|\alpha|_1 = k} \frac{y^\alpha}{\alpha!} \int_0^1 (1-t)^{k-1} D^\alpha f(\bar{x} + ty) dt; \quad (2.89)$$

$$R_{k-1}(y; \bar{x}) = \sum_{|\alpha|_1 = k} \frac{D^\alpha f(\bar{x} + t_0 y)}{\alpha!} y^\alpha, \quad \text{con } t_0 \in (0, 1); \quad (2.90)$$

$$|R_{k-1}(y; \bar{x})| \leq M_k \frac{|y|_1^k}{k!}, \quad \text{con } M_k := \sup_{\substack{|\alpha|_1 = k \\ |y| < \delta}} |D^\alpha f(\bar{x} + y)|. \quad (2.91)$$

Facciamo alcuni commenti:

Osservazione 2.42 (i) La prima formula in (2.88) è vera banalmente per definizione di R_k (vedi (2.87) e (2.86)); l'affermazione non banale in (2.88) è che $R_k(y; \bar{x}) = o(|y|^k)$, affermazione nota anche come ‘Teorema di Peano’.

(ii) La formula $f(\bar{x} + y) = T_{k-1}(y; \bar{x}) + R_{k-1}(y; \bar{x})$ (anche questa vera per definizione di R_{k-1}) con R_{k-1} dato in (2.89) o (2.90) è nota come formula di Taylor con, rispettivamente, ‘resto in forma integrale’ o ‘resto di Lagrange’.

(iii) La stima del resto in (2.91) segue facilmente da (2.90). In base a tale stima, possiamo dire che

$$f(\bar{x} + y) = T_{k-1}(y; \bar{x}) + O(|y|^k); \quad (2.92)$$

questa è un'altra forma standard della formula di Taylor per funzioni C^k .

La dimostrazione dell'Proposizione 2.41 si basa su alcuni lemmi. Il primo risultato è una versione particolare della formula di Taylor per funzioni di una variabile con resto integrale e resto di Lagrange e deriva facilmente dal Teorema fondamentale del calcolo³⁵.

³⁵Cfr. Corollario 8.25-(iii) [C2019].

Lemma 2.43 Sia³⁶ $F \in C^k([0, 1], \mathbb{R})$. Allora, per un opportuno $t_0 \in (0, 1)$,

$$F(1) = \sum_{j=0}^{k-1} \frac{F^{(j)}(0)}{j!} + \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} F^{(k)}(t) dt \quad (2.93)$$

$$= \sum_{j=0}^{k-1} \frac{F^{(j)}(0)}{j!} + \frac{F^{(k)}(t_0)}{k!}. \quad (2.94)$$

Dimostrazione Mostriamo la (2.93) per induzione su k . Se $k = 1$, (2.93) segue dal Teorma fondamentale del calcolo. Sia $k \geq 1$, e assumiamo che valga (2.93). Allora, integrando per parti, si ha che:

$$\begin{aligned} & \sum_{j=0}^k \frac{F^{(j)}(0)}{j!} + \int_0^1 \frac{(1-t)^k}{k!} F^{(k+1)}(t) dt \\ &= \sum_{j=0}^k \frac{F^{(j)}(0)}{j!} + \left[F^{(k)} \frac{(1-t)^k}{k!} \right]_0^1 + \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} F^{(k)}(t) dt \\ &= \sum_{j=0}^{k-1} \frac{F^{(j)}(0)}{j!} + \int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} F^{(k)}(t) dt \\ &\stackrel{(2.93)}{=} F(1). \end{aligned}$$

Per dimostrare la (2.94), facendo il cambio di variabile $s = (1-t)^k$ nell'integrale in (2.93) ed usando il teorema della media integrale (in una variabile³⁷), si ottiene, per un opportuno $t_0 \in (0, 1)$:

$$\int_0^1 \frac{(1-t)^{k-1}}{(k-1)!} F^{(k)}(t) dt = \frac{1}{k!} \int_0^1 F^{(k)}(1 - \sqrt[k]{s}) ds = \frac{F^{(k)}(t_0)}{k!}. \quad \blacksquare$$

L'idea – come nella dimostrazione della Proposizione 2.18 – è di applicare questo lemma alla funzione

$$t \in [0, 1] \mapsto F(t) := f(\bar{x} + ty), \quad (2.95)$$

che, nelle ipotesi della Proposizione 2.41, è di classe $C^k([0, 1], \mathbb{R})$. Si ricordi³⁸ che la derivata rispetto a t di F corrisponde alla derivata direzionale

$$D_y := y \cdot \nabla := \sum_{i=1}^n y_i D_i$$

su f , ossia,

$$F'(t) = D_y f(\bar{x} + ty) = y \cdot \nabla f(\bar{x} + ty) = \sum_{i=1}^n y_i D_i f(\bar{x} + ty).$$

Iterando tale relazione $j \leq k$ volte, otteniamo

$$\begin{aligned} F^{(j)}(t) &= D_y^j f(\bar{x} + ty) = (y \cdot \nabla)^j f(\bar{x} + ty) = \sum_{i_1, \dots, i_j=1}^n (D_{i_1} \cdots D_{i_j} f)(\bar{x} + ty) \\ &= \sum_{i_1, \dots, i_j=1}^n D^{e^{(i_1)} + \dots + e^{(i_j)}} f(\bar{x} + ty). \end{aligned} \quad (2.96)$$

Le derivate parziali che appaiono nell'ultimo termine di (2.96) sono derivate parziali di ordine j e i termini ottenuti permutando gli indici (i_1, \dots, i_j) danno luogo alla stessa derivata parziale $D^\alpha f$. Per riscrivere la relazione (2.96) in modo più compatto useremo il seguente risultato che generalizza la formula del binomio di Newton³⁹:

³⁶ $F \in C^k([0, 1], \mathbb{R})$ significa che esiste $\varepsilon > 0$ tale che $F \in C^k((-\varepsilon, 1 + \varepsilon))$.

³⁷ Cfr. Es. 8.5 in [C2019].

³⁸ Cfr. (2.36).

³⁹ Cfr. Proposizione 1.44 [C2019].

Lemma 2.44 (Teorema multinomiale) Sia $y = (y_1, \dots, y_n) \in \mathbb{R}^n$ e $k \in \mathbb{N}_0$. Allora,

$$\left(\sum_{i=1}^n y_i \right)^k = (y_1 + \dots + y_n)^k = \sum_{|\alpha|_1=k} \binom{k}{\alpha} y^\alpha, \quad (2.97)$$

dove

$$\binom{k}{\alpha} := \frac{k!}{\alpha!} = \frac{k!}{\alpha_1! \cdots \alpha_n!}, \quad (2.98)$$

è il coefficiente multinomiale ‘ k sopra α ’.

Dimostrazione Dimostriamo la (2.97) per induzione su $n \in \mathbb{N}$, $\forall k \in \mathbb{N}$. Per $n = 1$, (2.97) è ovviamente vera. Per $n = 2$, è la nota formula del binomio di Newton⁴⁰. Riscriviamo la (2.97) nella forma equivalente

$$\frac{1}{k!} \left(\sum_{i=1}^n y_i \right)^k = \sum_{|\alpha|_1=k} \frac{y^\alpha}{\alpha!}, \quad (2.99)$$

ed assumiamola vera per un $n \geq 2$ fissato e per ogni $k \in \mathbb{N}_0$ e dimostriamola per $n+1$. Per la formula del binomio di Newton e per la (2.99) con $k = j$, si ha che

$$\begin{aligned} \frac{1}{k!} \left(\sum_{i=1}^{n+1} y_i \right)^k &= \frac{1}{k} \left(\sum_{i=1}^n y_i + y_{n+1} \right)^k \\ &= \frac{1}{k!} \sum_{j=0}^k \binom{k}{j} \left(\sum_{i=1}^n y_i \right)^j y_{n+1}^{k-j} \\ &= \sum_{j=0}^k \frac{1}{j!} \left(\sum_{i=1}^n y_i \right)^j \frac{y_{n+1}^{k-j}}{(k-j)!} \\ &\stackrel{(2.99)}{=} \sum_{j=0}^k \sum_{|\alpha|_1=j} \frac{y^\alpha}{\alpha!} \cdot \frac{y_{n+1}^{k-j}}{(k-j)!} \\ &= \sum_{\substack{\alpha \in \mathbb{N}_0^{n+1} \\ |\alpha|_1=k}} \frac{y^\alpha}{\alpha!}, \end{aligned}$$

dove, nell’ultima uguaglianza, abbiamo posto $\alpha_{n+1} = k - j$ e osservato che $|\alpha_1, \dots, \alpha_n|_1 + k - j = j + (k - j) = k$. ■

Si noti ora che, per ogni i, j si ha⁴¹

$$(y_i D_i)(y_j D_j) = (y_j D_j)(y_i D_i), \quad (2.100)$$

ossia, che gli ‘operatori differenziali $(y_i D_i)$ e $(y_j D_j)$ commutano’ e dunque, come nel Teorema multinomiale, si ha

$$(y \cdot \nabla)^j = \left(\sum_{i=1}^n y_i D_i \right)^j = \sum_{|\alpha|_1=k} \binom{j}{\alpha} \prod_{i=1}^n (y_i D_i)^{\alpha_i} = j! \sum_{|\alpha|_1=k} \frac{y^\alpha}{\alpha!} D^\alpha. \quad (2.101)$$

Da (2.96) e da (2.101) segue, dunque, il seguente

Lemma 2.45 Sia F come in (2.95) con $f \in C^k$ come nella Proposizione 2.41. Allora,

$$\frac{F^{(j)}(t)}{j!} = \sum_{|\alpha|_1=j} \frac{D^\alpha f(\bar{x} + ty)}{\alpha!} y^\alpha, \quad \forall j \leq k; \quad (2.102)$$

$$\frac{F^{(j)}(0)}{j!} = \sum_{|\alpha|_1=j} \frac{D^\alpha f(\bar{x})}{\alpha!} y^\alpha, \quad \forall j \leq k. \quad (2.103)$$

⁴⁰Essendo $\sum_{\alpha \in \mathbb{N}_0^2: |\alpha|_1=k} c(\alpha) = \sum_{j=0}^k c(j, k-j)$.

⁴¹Se $i = j$ (2.100) è ovvia; se $i \neq j$, si ha $(y_i D_i)(y_j D_j) = y_i y_j D_i D_j = y_i y_j D_j D_i = (y_j D_j)(y_i D_i)$ (essendo y_j una costante rispetto alla derivazione D_i e viceversa).

Dimostrazione (della Proposizione 2.41) Dimostriamo (2.90):

$$\begin{aligned}
f(\bar{x} + y) &\stackrel{(2.95)}{=} F(1) \\
&\stackrel{(2.94)}{=} \sum_{j=0}^{k-1} \frac{F^{(j)}(0)}{j!} + \frac{F^{(k)}(t_0)}{k!} \\
&\stackrel{(2.102),(2.103)}{=} \sum_{j=0}^{k-1} \sum_{|\alpha|_1=j} \frac{D^\alpha f(\bar{x})}{\alpha!} y^\alpha + \sum_{|\alpha|_1=j} \frac{D^\alpha f(\bar{x} + t_0 y)}{\alpha!} y^\alpha \\
&= \sum_{|\alpha|_1 \leq k-1} \frac{D^\alpha f(\bar{x})}{\alpha!} y^\alpha + \sum_{|\alpha|_1=j} \frac{D^\alpha f(\bar{x} + t_0 y)}{\alpha!} y^\alpha \\
&\stackrel{(2.86)}{=} T_{k-1}(y; \bar{x}) + \sum_{|\alpha|_1=j} \frac{D^\alpha f(\bar{x} + t_0 y)}{\alpha!} y^\alpha,
\end{aligned}$$

da cui, ricordando ((2.87)) che $R_{k-1}(y; \bar{x}) = f(\bar{x} + y) - T_{k-1}(y; \bar{x})$, segue (2.90).

In maniera del tutto analoga, usando il resto in forma integrale (2.93) anziché il resto in forma di Lagrange (2.94), si ottiene la (2.89).

La (2.91) segue immediatamente da (2.90) e dal teorema multinomiale: infatti, per ogni $|y| < \delta$, dalla definizione di M_k in (2.91), segue che

$$|R_{k-1}(y; \bar{x})| \stackrel{(2.90)}{\leq} M_k \sum_{|\alpha|_1=k} \frac{|y_1|^{\alpha_1} \cdots |y_n|^{\alpha_n}}{\alpha!} \stackrel{(2.99)}{=} M_k \frac{|y|_1^k}{k!}.$$

Dimostriamo, infine, la (2.88), ossia che $\lim_{y \rightarrow 0} R_k(y; \bar{x})/|y|^k = 0$. Sia $\varepsilon > 0$. Poiché $f \in C^k(B)$, esiste $0 < \rho < \delta$ tale che

$$\sum_{|\alpha|_1=k} \frac{|D^\alpha f(\bar{x} + y) - D^\alpha f(\bar{x})|}{\alpha!} < \varepsilon, \quad \forall |y| < \rho. \quad (2.104)$$

Allora, se $|y| < \rho$,

$$\begin{aligned}
\frac{1}{|y|^k} |R_k(y; \bar{x})| &\stackrel{(2.87)}{=} \frac{1}{|y|^k} |f(\bar{x} + y) - T_k(y; \bar{x})| \\
&\stackrel{(2.87)}{=} \frac{1}{|y|^k} |T_{k-1}(y; \bar{x}) + R_{k-1}(y; \bar{x}) - T_k(y; \bar{x})| \\
&= \frac{1}{|y|^k} \left| R_{k-1}(y; \bar{x}) - \sum_{|\alpha|_1=k} \frac{D^\alpha f(\bar{x})}{\alpha!} y^\alpha \right| \\
&\stackrel{(2.90)}{=} \frac{1}{|y|^k} \left| \sum_{|\alpha|_1=k} \frac{D^\alpha f(\bar{x} + t_0 y) - D^\alpha f(\bar{x})}{\alpha!} y^\alpha \right| \\
&\leq \frac{1}{|y|^k} \sum_{|\alpha|_1=k} \frac{|D^\alpha f(\bar{x} + t_0 y) - D^\alpha f(\bar{x})|}{\alpha!} (|y_1|^{\alpha_1} \cdots |y_n|^{\alpha_n}) \\
&\leq \sum_{|\alpha|_1=k} \frac{|D^\alpha f(\bar{x} + t_0 y) - D^\alpha f(\bar{x})|}{\alpha!} \stackrel{(2.104)}{<} \varepsilon. \quad \blacksquare
\end{aligned}$$

Esempio 2.46 (i) Sia $P(x) := \sum_{|\alpha|_1 \leq k} a_\alpha x^\alpha$ un polinomio in n variabili $x \in \mathbb{R}^n$ di grado al più $k \in \mathbb{N}_0$.

Poiché, per ogni $\alpha, \beta \in \mathbb{N}_0^n$,

$$D^\beta x^\alpha = \begin{cases} \frac{\alpha!}{(\alpha - \beta)!} x^{\alpha - \beta}, & \text{se } \alpha \geq \beta, \\ 0, & \text{se } \exists i \text{ t.c. } \beta_i > \alpha_i, \end{cases} \quad (2.105)$$

si ha che

$$D^\beta x^\alpha(0) = \begin{cases} \alpha! , & \text{se } \alpha = \beta \\ 0 & , \text{ altrimenti} \end{cases} . \quad (2.106)$$

Dunque,

$$\frac{D^\alpha P(0)}{\alpha!} = \begin{cases} a_\alpha , & \text{se } |\alpha|_1 \leq k , \\ 0 , & \text{se } |\alpha|_1 > k , \end{cases}$$

e, quindi,

$$T_{P,k} = P ,$$

ossia, il polinomio di Taylor di ordine k di P coincide con sé stesso (come ci si aspettava).

(ii) Calcoliamo il polinomio di Taylor di ordine 5 in $(\bar{x}, \bar{y}) = (0, 0)$ della funzione $(1+x)^y$. Calcolando le derivate fino all'ordine 5 si trova $(1+x)^y = T_5(x, y; 0, 0) + O(|(x, y)|^6)$ con

$$T_5(x, y; 0, 0) = 1 + xy - \frac{x^2 y}{2} + \frac{x^2 y^2}{2} + \frac{x^3 y}{3} - \frac{x^4 y}{4} - \frac{x^3 y^2}{2} . \quad (2.107)$$

Osservazione 2.47 Da (2.105) segue immediatamente che se $P = \sum_{|\alpha|_1 \leq k} a_\alpha x^\alpha$ è un polinomio nelle n variabili x e se $P \equiv 0$, allora $a_\alpha = 0$ per ogni α . Equivalentemente, due polinomi in $x \in \mathbb{R}^n$ coincidono se e solo se hanno gli stessi coefficienti: più precisamente, se $P(x) = \sum_{|\alpha|_1 \leq k} a_\alpha x^\alpha$ e $Q(x) = \sum_{|\beta|_1 \leq h} b_\beta x^\beta$ e $P(x) = Q(x)$, $\forall x$, allora $h = k$ e $a_\alpha = b_\alpha$ per ogni α .

Infine, osserviamo che il polinomio di Taylor di una funzione C^k è unico, nel senso che segue.

Proposizione 2.48 Siano $\bar{x}, y \in \mathbb{R}^n$, $k \in \mathbb{N}$, $\delta > 0$, $f \in C^k(B)$ con $B := B_\delta(\bar{x})$ e $|y| < \delta$. Esistono numeri $a_\alpha \in \mathbb{R}$, per $|\alpha|_1 \leq k$, tali che

$$f(\bar{x} + y) = \sum_{|\alpha|_1 \leq k} a_\alpha y^\alpha + o(|y|^k) , \quad (2.108)$$

se e solo se

$$a_\alpha = \frac{D^\alpha f(\bar{x})}{\alpha!} . \quad (2.109)$$

La dimostrazione si basa sul seguente

Lemma 2.49 Sia $k \in \mathbb{N}$ e $a_\alpha \in \mathbb{R}$ per $\alpha \in \mathbb{N}_0^n$ con $|\alpha|_1 \leq k$. Se $P(y) := \sum_{|\alpha|_1 \leq k} a_\alpha y^\alpha = o(|y|^k)$, allora $a_\alpha = 0$ per ogni α .

Dimostrazione Poiché $\lim_{y \rightarrow 0} o(|y|^k) = 0$, si ha che $a_0 = 0$. Ora, nel caso $n = 1$, l'affermazione è ovvia⁴². Fissiamo $y \neq 0$ e consideriamo il polinomio in $t \in \mathbb{R}$ dato da

$$p(t) := \sum_{j=1}^k A_j t^j , \quad \text{con } A_j := \sum_{|\alpha|_1 = j} a_\alpha y^\alpha . \quad (2.110)$$

Allora, $P(ty) = p(t)$ e dall'ipotesi segue che $p(t) = o(t^{k+1})$ e quindi, dal caso unidimensionale, segue che $A_j = 0$ per ogni j e per ogni y , ossia, che il polinomio omogeneo in y , $A_j = \sum_{|\alpha|_1 = j} a_\alpha y^\alpha$, è identicamente nullo. Questo implica che i coefficienti a_α sono nulli se $|\alpha|_1 = j$ e, poiché questo vale per ogni j , si ha la tesi. ■

Dimostrazione (della Proposizione 2.48) Il 'se' è il Teorema di Peano (ossia, la prima parte della Proposizione 2.41).

⁴²Infatti, $a_1 y + a_2 y^2 + \dots + a_k y^k = o(y^k)$ implica (dividendo per y) che $a_1 + a_2 y + \dots + a_k y^{k-1} = o(y^{k-1})$ e, quindi, che $a_1 = 0$, etc.

Dimostriamo il ‘solo se’, ossia, che (2.108) implica (2.109). Dal Teorema di Peano segue che $f(\bar{x}+y) = T_{f,k}(y; \bar{x}) + o(|y|^k)$, il che, assieme a (2.108), implica che

$$T_{f,k}(y; \bar{x}) - \sum_{|\alpha|_1 \leq k} a_\alpha y^\alpha = o(|y|^k).$$

Dunque, per il Lemma 2.49 $T_{f,k}(y; \bar{x}) = \sum_{|\alpha|_1 \leq k} a_\alpha y^\alpha$, ma allora i coefficienti dei polinomi devono coincidere. il che equivale a (2.109). ■

Esempio 2.50 Calcoliamo il polinomio di Taylor di ordine 1000 in $\bar{x} = 0$ di

$$f(x) = x_1 x_4^2 + \frac{x_1 + x_3^2}{1 - x_1^2 - x_2^2}; \quad \text{dove } x \in \mathbb{R}^4, x_1^2 + x_2^2 < 1.$$

Poiché

$$\frac{1}{1-t} = \sum_{j=0}^{\infty} t^j = \sum_{j=0}^{500} t^j + O(|t|^{501}),$$

si ha che

$$\begin{aligned} \frac{1}{1 - (x_1^2 + x_2^2)} &= \sum_{j=0}^{500} (x_1^2 + x_2^2)^j + O(|x|^{1002}) \\ &= \sum_{j=0}^{500} \sum_{i=0}^j \binom{j}{i} x_1^{2i} x_2^{2(j-i)} + O(|x|^{1002}) \\ &= \sum_{\substack{h=(h_1, h_2) \in \mathbb{N}^2 \\ h_1+h_2 \leq 500}} \frac{(h_1+h_2)!}{h_1! h_2!} x_1^{2h_1} x_2^{2h_2} + O(|x|^{1002}). \end{aligned}$$

Quindi

$$f(x) = x_1 x_4^2 + \sum_{\substack{h=(h_1, h_2) \in \mathbb{N}^2 \\ h_1+h_2 \leq 500}} \frac{(h_1+h_2)!}{h_1! h_2!} \left(x_1^{2h_1+1} x_2^{2h_2} + x_1^{2h_1} x_2^{2h_2} x_3^2 \right) + O(|x|^{1002}).$$

Da tale relazione e dalla Proposizione 2.48 segue che $f(x) = P_{1000}(x; 0) + O(|x|^{1002})$ con $P_{1000}(x; 0) := \sum_{|\alpha|_1 \leq 1000} a_\alpha x^\alpha$ e coefficienti a_α dati da

$$a_\alpha = \begin{cases} 1, & \alpha_1 = 1, \alpha_2 = \alpha_3 = 0, \alpha_4 = 2, \\ \frac{((\alpha_1 + \alpha_2 - 1)/2)!}{((\alpha_1 - 1)/2)! (\alpha_2/2)!}, & \alpha_1 \text{ dispari, } \alpha_2 \text{ pari, } \alpha_3 = \alpha_4 = 0, \\ \frac{((\alpha_1 + \alpha_2)/2)!}{(\alpha_1/2)! (\alpha_2/2)!}, & \alpha_1 \text{ pari, } \alpha_2 \text{ pari, } \alpha_3 = 2, \alpha_4 = 0, \\ 0, & \text{altrimenti.} \end{cases}$$

Esercizi

Esercizio 2.1 Calcolare $D_i |x|^a$, $\forall x \in \mathbb{R}^n \setminus \{0\}$ e $\forall a \in \mathbb{R}$.

Esercizio 2.2 A seconda del contesto, il simbolo f_x può significare: derivata parziale rispetto alla variabili scalare $x \in \mathbb{R}$, gradiente rispetto alla variabile vettoriale $x \in \mathbb{R}^n$, oppure matrice jacobiana

$(m \times n)$ di una funzione di n variabili $x \in \mathbb{R}^n$ e a valori in \mathbb{R}^m con $m > 1$. Si calcoli f_x nei seguenti casi:

- (i) $f = \cos(xy) - z^3$, $(x, y, z) \in \mathbb{R}^3$;
- (ii) $f = e^{-|x|^2}(1 + x_1x_3)$, $x \in \mathbb{R}^3$;
- (iii) $f = \left(\tanh\left(\frac{x_1}{x_2}\right), x_1x_2^3, |x| \right)$, $x \in \mathbb{R}^2, x_2 \neq 0$;
- (iv) $f = (x_1, x_1x_2, \dots, x_1 \cdots x_n)$, $x \in \mathbb{R}^n$;
- (v) $f = \left(e^{|x|^3}, \left(\sum_{i=1}^n x_i \right)^3 \right)$, $x \in \mathbb{R}^n \setminus \{0\}$.

Esercizio 2.3 Calcolare $\frac{\partial f}{\partial x}$ per:

- (i) $f(x) = \prod_{i=1}^n x_i^2$,
- (ii) $f(x) = (x_1 + x_2^2, \cos(x_1, x_2))$,
- (iii) $f(x, y, z) = \frac{y}{z - x^2}$.

Esercizio 2.4 Sia $x \in \mathbb{R}^2 \mapsto f(x) := e^{x_1^2 + x_2} + \sinh x_1$. Calcolare:

- (i) $\frac{D^\alpha f}{\alpha!}$ con $\alpha = (2, 3)$; (ii) f' ; (iii) f'' .

Esercizio 2.5 Calcolare $D^\alpha f(\bar{x})$ con $f(x_1, x_2) = x_1^2 x_2^4$, $\alpha = (2, 1)$.

Esercizio 2.6 Sia $P := \{(x, y) \in \mathbb{R}^2 \setminus \{(0, 0)\} \mid y = x^2\}$ e sia

$$f(x, y) = \begin{cases} 0, & \text{se } (x, y) \notin P, \\ x^{-1}, & \text{se } (x, y) \in P. \end{cases}$$

Si calcolino le derivate direzionali $D_{(\xi, \eta)} f(0, 0)$, per ogni $(\xi, \eta) \in \mathbb{R}^2 \setminus \{(0, 0)\}$. Si calcolino $\sup_B f$ e $\inf_B f$ dove B è un qualunque aperto che contiene l'origine.

Esercizio 2.7 Si considerino le seguenti funzioni:

$$f(x_1, x_2, x_3) = \sin(x_1 e^{x_2 + x_3}), \quad \varphi(t) = (e^t, \sin t, t)$$

Scrivere $f \circ \varphi$; calcolarne la derivata in $t = 0$ in maniera diretta e applicando la regola di derivazione per funzioni composte.

Esercizio 2.8 Calcolare $D_i g(|x|)$ dove $x \in \mathbb{R}^n \setminus \{0\}$ e g è una funzione $C^1((0, \infty))$.

Esercizio 2.9 Fissati $\alpha, \beta, \delta \in \mathbb{R}$, si consideri la seguente funzione:

$$f : (x, y) \in \mathbb{R}^2 \mapsto f(x, y) = \begin{cases} \frac{|x|^\alpha |y|^\beta}{|x^2 + y^2|^{\delta/2}}, & \text{se } (x, y) \neq 0, \\ 0 & \text{se } (x, y) = 0. \end{cases}$$

Trovare condizioni necessarie e sufficienti affinché: (i) f sia continua nell'origine; (ii) f abbia derivate direzionali nell'origine; (iii) f sia differenziabile nell'origine; (iv) f sia $C^1(B_r(0))$, con un $r > 0$.

Esercizio 2.10 Studiare la regolarità di

$$f = \begin{cases} x^2 + y^2, & \text{se } x = 0, \\ y, & \text{se } x \neq 0. \end{cases}$$

Esercizio 2.11 Sia

$$f(x) := \begin{cases} \frac{\log(1+|x|)}{|x|}, & \text{se } x \in \mathbb{R}^n \setminus \{0\}, \\ 1, & \text{se } x = 0. \end{cases}$$

- (i) Discutere la regolarità di f in $x = 0$ (continuità, derivabilità etc.).
 (ii) Trovare $\delta > 0$ tale che $|f(x)| < 1/10$ per $|x| < \delta$.

Esercizio 2.12 Si studi, al variare di $\alpha \in \mathbb{R}$ la continuità della funzione $f(x) := x_1|x|^\alpha$ per $x \in \mathbb{R}^n \setminus \{0\}$ e $f(0) := 0$.

Esercizio 2.13 Sia $f(0) := 0$ e per $x \in \mathbb{R}^2 \neq 0$ sia $f(x) = |x|^{2p} \operatorname{sen}|x|^{-1}$. Si dica per quali p , f_{x_1} esiste ed è continua.

Esercizio 2.14 Sia $f(0,0) = 0$ e, per $(x,y) \neq (0,0)$, sia $f(x,y) = (x^2 + y^2)^{-1} \operatorname{sen}(x^3 + y^3)$. Si studi la regolarità di f in \mathbb{R}^2 (continuità, differenziabilità, derivabilità, etc.)

Esercizio 2.15 Si dimostri che non esiste alcuna $f \in C^2(\mathbb{R}^2)$ tale che $\nabla f = (x \operatorname{sen} y, y \operatorname{sen} x)$.

Esercizio 2.16 Sia $f := e^{-|x|^2} (1 + x_1 x_3)$ con $x \in \mathbb{R}^3$. Calcolare $f''(0)$ e scrivere la formula di Taylor (in 0) al secondo ordine per tale f .

Esercizio 2.17 Scrivere la matrice hessiana della funzione $f(x,y) = e^y \operatorname{sen} x$ e calcolare $f''(0)\xi \cdot \eta$, dove $\xi = (3, 2)$ e $\eta = (-1, 3)$.

Esercizio 2.18 Si trovino i punti stazionari e si dica se si tratta di massimi o minimi (relativi o assoluti) di $f(x,y) = (x + 3y)e^{-xy}$.

Esercizio 2.19 Sia

$$f(x,y) = \int_{2+\operatorname{sen} y^2}^4 \operatorname{senh}(t^2 x + y) dt$$

- (a) Calcolare f_x e f_y ;
 (b) dimostrare che f non ha punti critici in \mathbb{R}^2 ;
 (c) determinare l'estremo superiore e l'estremo inferiore di f su \mathbb{R}^2 .

Esercizio 2.20 Trovare i massimi e minimi di $F(x) = \int_0^1 \log(1 + x^2 + y^2) dy$.

Esercizio 2.21 Per $x \in \mathbb{R}^n$, sia

$$f(x) = \begin{cases} 1 - |x|^2, & |x| < 1, \\ 0, & |x| \geq 1. \end{cases}$$

- (a) Dimostrare che $f \in C(\mathbb{R}^n)$ ma $f \notin C^1(\mathbb{R}^n)$ e discutere la differenziabilità di f ;
 (b) trovare un numero positivo δ per il quale se $|x - x_0| < \delta$ allora $|f(x) - f(x_0)| < 1/100$, con x_0 tale che $|x_0| = 1$;
 (c) f è differenziabile nel punto $(\sqrt{2}, 1, 1, \dots, 1)$?

Esercizio 2.22* Sia $f \in C^1(\mathbb{R}^n, \mathbb{R})$ e supponiamo che $f'(\bar{x}) = 0$ se e solo se $x = \bar{x}$, e che $\sup_{\mathbb{R}^n} f > f(\bar{x}) > \inf_{\mathbb{R}^n} f$.

- (i) Nel caso $n = 1$ dimostrare che $x = \bar{x}$ non può essere né un minimo né un massimo locale per f .
 (ii) È vera l'affermazione fatta al punto (i) nel caso di $n \geq 2$?

Esercizio 2.23 Sia $x \in \mathbb{R}^2 \mapsto f(x)$ così definita: $f(x,y) := 0$ se $x = 1$ e

$$f(x,y) := y \exp\left(-\left(\frac{y}{x-1}\right)^2\right), \quad \text{se } x \neq 1.$$

- (i) Dimostrare che $f \in C^\infty(\mathbb{R}^2 \setminus \{(1,0)\})$.
 (ii) Studiare la continuità e regolarità di f in $(1,0)$.
 (iii) Studiare i punti critici di f .
 (iv) Trovare $\delta > 0$ tale che $|f(x,y) - f(x_0,y_0)| < 1/100$ per $|(x,y) - (x_0,y_0)| < \delta$ con $(x_0,y_0) = (1,1)$ e (qualora f fosse continua in $(1,0)$) con $(x_0,y_0) = (1,0)$.

Esercizio 2.24 Si calcoli il polinomio di Taylor di grado 2 della funzione y^x nell'intorno di $(1, 1)$.

Esercizio 2.25 Sia f la funzione:

$$f(y, x) = x + x^3 y^2 + y + \frac{\sin xy - xy}{x}.$$

Dire se è possibile applicare il Teorema delle funzioni implicite per $y = y(x)$ in $(0, 0)$ ed in caso affermativo calcolare $y'(0)$ e $y''(0)$.

Esercizio 2.26 Data la funzione $f(y, x) = x^2 + x^4 y^2 + \sin xy$ determinare i punti dove è possibile applicare il Teorema delle Funzioni Implicite.

Esercizio 2.27 Sia $f(x, y) := (\sin(x - y^2), x^4 + \tan y)$. Si dimostri che f è invertibile in un intorno di $(0, 0)$ e si trovi un aperto su cui è definita la funzione inversa.

Esercizio 2.28 Si consideri la funzione $f : y \in \mathbb{R}^2 \rightarrow f(y) = (f_1(y), f_2(y)) \in \mathbb{R}^2$ definita come

$$f_1 = y_1 + y_1^2 \cos y_2, \quad f_2 = y_2 + y_1^2.$$

Si dica se la funzione f è invertibile in un intorno di $y_0 = (0, 0)$ e, in caso affermativo, si trovi un aperto su cui è definita la funzione inversa.

Esercizio 2.29 Si consideri la funzione $f : x \in \mathbb{R}^2 \rightarrow f(x) = (f_1(x), f_2(x)) \in \mathbb{R}^2$ definita come

$$f_1 = x_1 + x_1^2 \cos x_2, \quad f_2 = x_2 + x_1^2. \quad (2.111)$$

Si dica se la funzione f è invertibile in un intorno di $x_0 = (0, 0)$ e, in caso affermativo, si trovi un aperto su cui è definita la funzione inversa.

Esercizio 2.30 Sia $f : x \in \mathbb{R}^2 \rightarrow y \in \mathbb{R}^2$ data da $f(x) := (x_1 + \sin(x_1 x_2), x_2 + \sin x_1^2)$. Si trovi una sfera $B_r(0) := \{y \in \mathbb{R}^2 : |y| < r\}$ su cui sia definita e regolare la funzione inversa di f .

Esercizio 2.31 Sia $(x, y) \in \mathbb{R}^2 \mapsto f(x, y) := (\sin(x - y^2), x^4 + \tan y)$. (i) Si dimostri che f è invertibile in un intorno di $(0, 0)$ e si trovi una sfera su cui è definita la funzione inversa. (ii) Su tale sfera si calcoli la jacobiana della funzione inversa.

Esercizio 2.32 Sia $\varepsilon \geq 0$ e $f(x) := x + \varepsilon g(x)$ con $x \in \mathbb{R}^n$ e $g \in C^1(D, \mathbb{R}^n)$, dove $D := \{x \in \mathbb{R}^n : |x| \leq 1\}$. Trovare $\varepsilon_0 > 0$ tale che f sia invertibile (con inversa C^1) su D .

Esercizio 2.33 (i) Si dica quante funzioni continue $(x, y) \rightarrow z(x, y)$ soddisfano l'equazione $x \cos y + y \cos z + z \cos x = 1$ in un intorno di $(x, y) = (0, 0)$. (ii) Si calcolino le derivate parziali delle funzioni di cui al punto (i) (sempre in un intorno dell'origine).

Esercizio 2.34 Sia $z = z(x, y)$ la funzione definita implicitamente dalle relazioni:

$$z^3 - 2xy + y = 0, \quad z(1, 1) = 1.$$

Si calcoli il polinomio di Taylor di grado 2 nell'intorno di $(1, 1)$.

Esercizio 2.35 Si studi il luogo dei punti $\{(x, y) \in \mathbb{R}^2 : \frac{y^2}{2} - \cos x = E\}$ al variare del parametro reale E .

Esercizio 2.36 Trovare il massimo e minimo di

$$(x, y) \in \mathbb{R}^2 \mapsto f(x, y) := \frac{1 + x - y}{\sqrt{1 + x^2 + y^2}}, \quad \text{su } D := \{x^2 + y^2 \leq 4\}.$$

Esercizio 2.37 Determinare l'estremo superiore ed inferiore (specificando se si tratta di massimi o minimi) della funzione $f(x, y) := 1/(x^2 + y^2)$ sull'insieme $A := \{(x, y) \in \mathbb{R}^2 \mid xy + \frac{1}{2} \sin(xy) > 1\}$.

Esercizio 2.38 Si trovino il massimo ed il minimo della funzione $(x, y) \in \mathbb{R}^2 \mapsto f(x, y) = xy$ sull'insieme $D := \{(x, y) \in \mathbb{R}^2 : |x| + y^2 \leq 3\}$.

Esercizio 2.39 Trovare il massimo ed il minimo della funzione

$$f(x) = \left(\sum_{i=1}^4 |x_i|^5 \right)^{\frac{1}{5}}$$

sul bordo sferico $\{x \in \mathbb{R}^4 : |x| = 1\}$.

Esercizio 2.40 Trovare i massimi e i minimi (se esistono) della funzione

$$(x, y) \in \mathbb{R}^2 \mapsto f(x, y) := \exp(x^2 + y^2) - \frac{x^2}{2} - y^2$$

sull'insieme $\{x \geq 1\} \cap \{4x^2 + y^2 \leq 8\}$.

Esercizio 2.41 Sia

$$(x, y) \in \mathbb{R}^2 \times \mathbb{R} \mapsto f(x, y) := |x|^2 + y^2 - 2x_1 + 4x_2 - 6y - 11.$$

(i) Quante soluzioni g di classe C^∞ nell'intorno del punto $(1, -2)$ esistono dell'equazione $f(x, g(x)) = 0$?

(ii) Si verifichi che se $g(1, -2) > 0$ allora g ha un massimo relativo stretto in $(1, -2)$.

Esercizio 2.42 Si discutano i massimi e minimi relativi ed assoluti (qualora esistano) della funzione

$$(x, y) \in \{(x, y) \in \mathbb{R}^2 \mid x + y \geq 1\} \mapsto f(x, y) = xy^2(x + y - 1).$$

Esercizio 2.43 Trovare il valore massimo e quello minimo (qualora esistano) della somma degli spigoli di parallelepipedi rettangoli di volume unitario.

Esercizio 2.44 Si scriva esplicitamente il polinomio di Taylor di ordine 2, 3 e 4 per una funzione arbitraria di due e tre variabili nell'intorno dell'origine (assumendo, naturalmente che tali funzioni siano sufficientemente derivabili).

Esercizio 2.45 Si calcoli il polinomio di Taylor $T_k(\xi; \bar{x})$ nei seguenti casi:

- (i) $k = 3, n = 2, \bar{x} = 0, f(x) = e^{x_1 x_2} + x_1^4,$
- (ii) $k = 3, (n \text{ arbitrario}), \bar{x} = 0, f(x) = \cos |x|,$
- (iii) $k = 2, \bar{x} = 0, f(x) = x_2^4 - \sinh(x_1 + \dots + x_n),$
- (iv) $k = 4, n = 2, \bar{x} = (\pi, 1), f(x) = \tan(x_1 x_2) - \log x_2.$

Esercizio 2.46 Calcolare il polinomio di Taylor di ordine 10 in $x_0 = 0$ della funzione

$$x \in \mathbb{R}^3 \mapsto f(x) := \log(1 + x_1 x_2 x_3).$$

Esercizio 2.47 Si dia una stima dei resti $|R_k|$ nell'Es 2.45.

Esercizio 2.48 Calcolare il polinomio di Taylor di grado 3 della funzione $f(x, y) = (2x + y)e^{x^2 - y^2}$ nel punto $(0, 0)$.

Esercizio 2.49 (i) Sia $f(y, x) = y^2 + x^2 - 1$, con $y \in \mathbb{R}, x \in \mathbb{R}$ e si fissi $(y_0, x_0) = (1, 0)$. Si trovino ρ ed r tali che (3.34) e (3.35) valgano.

(ii) Si dimostri che il valore $r = (1/\sqrt{2})$ è il valore massimo che si può ottenere anche se si usano (3.37) e (3.38) facendo variare $\alpha \in (0, 1)$.

Esercizio 2.50 Sia $f : (y, x) \in \mathbb{R}^2 \times \mathbb{R} \rightarrow f(y, x) \in \mathbb{R}^2$ definita come

$$\begin{aligned} f_1(y_1, y_2, x) &= \sin x + e^x y_1 + x \sin(y_1 y_2), \\ f_2(y_1, y_2, x) &= 3|x| + y_2 + y_1^4, \end{aligned} \tag{2.112}$$

e sia $(y_0, x_0) := (y_{01}, y_{02}, x_0) = (0, 0, 0)$. Si dica se il Teorema 3.19 è applicabile in un intorno di (y_0, x_0) e se sì si trovino dei valori di ρ e r per cui (3.34) e (3.35) vengano soddisfatte.

Esercizio 2.51 Sia $E_0 := \{(x, y) \in \mathbb{R}^2 : (x - 1)^2 + (y - 1)^2 = 2\}$ e sia K_r il quadrato chiuso di lato $2r$ centrato nell'origine. Dimostrare che nell'intorno di $(0, 0)$ è applicabile il teorema delle funzioni implicite ma che non esiste $r > 0$ tale che $E_0 \cap K_r$ possa essere espresso come grafico di una funzione delle x o delle y .

Capitolo 3

Funzioni inverse e funzioni implicite

1 Teorema della Funzione Inversa

Consideriamo il caso di una *funzione lineare*: $f(x) = Jx + \bar{y}$ con $x, \bar{y} \in \mathbb{R}^n$ e $J \in \mathbb{R}^{n \times n}$. Invertire f significa, dato y , risolvere l'equazione in x data da: $Jx + \bar{y} = y$. Dall'algebra lineare sappiamo che tale soluzione ha una e una sola soluzione x se e solo se $\det J \neq 0$, in qual caso, la matrice J è invertibile e¹ $x = J^{-1}(y - \bar{y})$. Dunque, la funzione $y \in \mathbb{R}^n \mapsto g(y) := J^{-1}(y - \bar{y})$ è la funzione inversa di f . Si noti che J è la matrice jacobiana $f'(x)$.

Ora, sapendo che vicino ad ogni punto \bar{x} dove una arbitraria funzione f è differenziabile, f è ben approssimata dalla mappa lineare $x \mapsto \bar{y} + J(x - \bar{x})$ con $\bar{y} := f(\bar{x})$ e $J := f'(\bar{x})$, l'enunciato del seguente Teorema della funzione Inversa potrebbe non sorprendere.

Teorema 3.1 (Teorema della Funzione Inversa) *Siano $r_0 > 0$, $x_0 \in \mathbb{R}^n$ e sia f una funzione di classe $C^1(B_{r_0}(x_0), \mathbb{R}^n)$ tale che $\det f'(x_0) \neq 0$. Allora:*

- (i) *esiste $0 < r \leq r_0$ tale che f è iniettiva su $B := B_r(x_0)$ e la matrice jacobiana f' è invertibile su B ;*
- (ii) *$V := f(B)$ è un insieme aperto, la funzione inversa $f^{-1} \in C^1(V)$ e, per ogni $y \in V$,*

$$(f^{-1})'(y) = (f'(x))^{-1}, \quad \text{dove } x = f^{-1}(y). \quad (3.1)$$

La dimostrazione di questo importante teorema si basa su due lemmi: il primo è un risultato elementare di algebra lineare; il secondo si basa sul teorema del valor medio per funzioni vettoriali e sul lemma delle contrazioni su \mathbb{R}^n .

Lemma 3.2 *Siano $T, S \in \mathbb{R}^{n \times n}$ e sia $\det T \neq 0$. Se $\|I - TS\| < 1$, allora S è invertibile e*

$$\|S^{-1}\| \leq \frac{\|T\|}{1 - \|I - TS\|}. \quad (3.2)$$

Dimostrazione Si osservi che $TS = I - (I - TS)$. Ricordando il punto (iii) della Proposizione 1.71 nel Complemento 1.5, si ha che TS è invertibile con inversa data da

$$(TS)^{-1} = \sum_{k=0}^{\infty} (I - TS)^k, \quad (3.3)$$

¹Cfr. punto (xxxiv) nell'Appendice D.

relazione che implica che $\det S \neq 0$ (se fosse $\det S = 0$, si avrebbe $\det TS = \det T \det S = 0$, contraddicendo l'invertibilità di TS). Quindi, da (3.3) segue che

$$S^{-1} = \left(\sum_{k=0}^{\infty} (I - TS)^k \right) T, \quad (3.4)$$

da cui segue

$$\|S^{-1}\| \leq \|T\| \sum_{k=0}^{\infty} \|I - TS\|^k = \frac{\|T\|}{1 - \|I - TS\|}. \quad \blacksquare \quad (3.5)$$

Il Teorema della Funzione Inversa sarà grave una semplice conseguenza della sua seguente formulazione quantitativa:

Teorema 3.3 *Sia $B \subseteq \mathbb{R}^n$ un insieme aperto, convesso e $x_0 \in B$. Sia $f \in C^1(B, \mathbb{R}^n)$ tale che $\det f'(x_0) \neq 0$ e supponiamo che*

$$\theta := \sup_{x \in B} \|I - Tf'(x)\| < 1, \quad \text{dove } T := (f'(x_0))^{-1}. \quad (3.6)$$

Allora, $\det f'(x) \neq 0$, per ogni $x \in B$ e

$$\|(f'(x))^{-1}\| \leq \lambda := \frac{\|T\|}{1 - \theta}; \quad (3.7)$$

inoltre f è iniettiva su B e la sua funzione inversa $f^{-1} : V := f(B) \xrightarrow{\text{su}} B$ soddisfa

$$\sup_{y, \bar{y} \in V} \frac{|f^{-1}(y) - f^{-1}(\bar{y})|}{|y - \bar{y}|} \leq \lambda. \quad (3.8)$$

Infine, se $B := B_r(x_0)$ e se $\rho := r/\lambda$, $y_0 := f(x_0)$, allora si ha

$$B_\rho(y_0) \subseteq f(B). \quad (3.9)$$

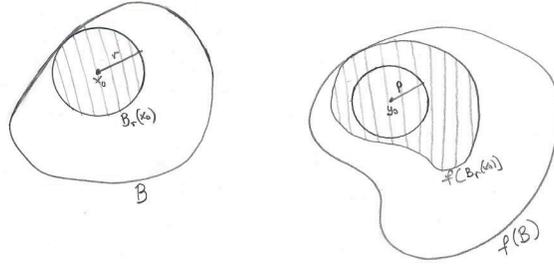


Figura 3.1: Teorema della funzione inversa

Dimostrazione Dal Lemma 3.2 e da (3.6) segue immediatamente (3.7).

Poiché B è convesso, per ogni $x, \bar{x} \in B$, per il Teorema del valor medio 2.28, si ha

$$f(x) - f(\bar{x}) = \left(\int_0^1 f'(\bar{x} + t(x - \bar{x})) dt \right) (x - \bar{x}) =: S_{\bar{x}, x} (x - \bar{x}), \quad (3.10)$$

con $S_{\bar{x}, x} \in \mathbb{R}^{n \times n}$. Da (3.6) segue che

$$\begin{aligned} \|I - TS_{\bar{x}, x}\| &= \left\| \int_0^1 (I - Tf'(\bar{x} + t(x - \bar{x}))) dt \right\| \\ &\leq \int_0^1 \|I - Tf'(\bar{x} + t(x - \bar{x}))\| dt \stackrel{(3.6)}{\leq} \theta < 1. \end{aligned}$$

Dal Lemma 3.2 segue che la matrice $S_{\bar{x},x}$ è invertibile per ogni $\bar{x}, x \in B$ e, quindi, per (3.10), si ha che:

$$x - \bar{x} = S_{\bar{x},x}^{-1}(f(x) - f(\bar{x})). \quad (3.11)$$

Tale relazione implica immediatamente l'iniettività di f . Ponendo $x = g(y)$ e $\bar{x} = g(\bar{y})$, da (3.5) e (3.11) segue poi (3.8).

Per dimostrare la (3.9) faremo uso del Lemma delle Contrazioni (Teorema 1.49).

Sia $y_0 := f(x_0)$ e sia $\bar{y} \in B_\rho(y_0)$: dobbiamo far vedere che esiste $\bar{x} \in B_r(x_0)$ tale che $f(\bar{x}) = \bar{y}$. Se $\bar{y} = y_0$, poniamo $\bar{x} = x_0$ e abbiamo finito. Se $\bar{y} \neq y_0$, poniamo $\bar{\rho} := |\bar{y} - y_0|$ e $\bar{r} := \lambda\bar{\rho}$ cosicché $0 < \bar{\rho} < \rho$ (poiché $\bar{y} \in B_\rho(y_0)$) e $\bar{r} = \lambda\bar{\rho} < \lambda\rho = r$ (per definizione di λ e ρ). Sia $E := \overline{B_{\bar{r}}(x_0)}$. Allora, $E \subseteq B_r(x_0)$ ed è un sottoinsieme chiuso e non vuoto dello spazio metrico \mathbb{R}^n . Definiamo la funzione

$$\Phi : x \in E \mapsto \Phi(x) := x - T(f(x) - \bar{y}) \in \mathbb{R}^n, \quad (3.12)$$

e osserviamo che $\Phi'(x) = I_n - Tf'(x)$. Dal Teorema del valor medio segue che, per ogni $x, x' \in E$,

$$\|\Phi(x) - \Phi(x')\| = \left\| \left(\int_0^1 \Phi'(x' + t(x - x')) dt \right) (x - x') \right\| \leq \theta |x - x'|, \quad (3.13)$$

ossia, Φ è una contrazione. Mostriamo, ora, che $\Phi : E \rightarrow E$: infatti, se $x \in E$, si ha²

$$\begin{aligned} |\Phi(x) - x_0| &\leq |\Phi(x) - \Phi(x_0)| + |\Phi(x_0) - x_0| \\ &\stackrel{(3.13)}{\leq} \theta \bar{r} + \mu |\bar{y} - y_0| \leq \theta \bar{r} + \mu \bar{\rho} = \theta \bar{r} + \mu \frac{\bar{r}}{\lambda} = \bar{r}. \end{aligned}$$

Quindi, per il Lemma delle contrazioni, Φ ha un unico punto fisso $\bar{x} \in E$, ma $\Phi(\bar{x}) = \bar{x}$ è equivalente a $f(\bar{x}) = \bar{y}$ ed essendo $E \subseteq B_r(x_0)$ si ha che $\bar{x} \in B_r(x_0)$. ■

Dimostrazione (del Teorema 3.1)

Poiché $f \in C^1(B_{r_0}(x_0))$, la funzione $x \rightarrow u(x) := \|I_n - Tf'(x)\| \in [0, +\infty)$ è continua su $B_{r_0}(x_0)$ e $u(x_0) = 0$. Dunque, per ogni $0 < \theta < 1$, esiste $0 < r \leq r_0$ tale che $u(x) = \|I_n - Tf'(x)\| \leq \theta < 1$. Il punto (i) segue, immediatamente, dal Teorema 3.3 applicato a $B = B_r(x_0)$. Si noti che in particolare vale (3.8).

Se $\bar{y} \in V = f(B)$ esiste un (unico) $\bar{x} \in B$ tale che $f(\bar{x}) = \bar{y}$ e un $0 < \hat{r}$ tale che $B_{\hat{r}}(\bar{x}) \subseteq B$. Allora, possiamo applicare il Teorema 3.3 con \bar{x}, \bar{y} e \hat{r} al posto, rispettivamente, di x_0, y_0 e r : da (3.9) segue allora che $B_\rho(\bar{y}) \subseteq V$, il che vuol dire che V è un insieme aperto.

Mostriamo, ora, che $g := f^{-1} : V \rightarrow B$ è differenziabile su V . Siano $\bar{y}, y \in V$ e siano $\bar{x} = g(\bar{y}), x = g(y) \in B$. Poiché $f \in C^1(B)$, f è differenziabile in B . Quindi,

$$f(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + R(x; \bar{x}), \quad \text{e} \quad R(x; \bar{x}) = o(|x - \bar{x}|). \quad (3.14)$$

Tale relazione è equivalente a:

$$y = \bar{y} + f'(\bar{x})(g(y) - g(\bar{y})) + R(g(y), g(\bar{y})),$$

o, anche, moltiplicando per la matrice³ $T_{\bar{x}} = (f'(\bar{x}))^{-1}$, a:

$$g(y) = g(\bar{y}) + T_{\bar{x}}(y - \bar{y}) - T_{\bar{x}}R(g(y), g(\bar{y})) =: g(\bar{y}) + T_{\bar{x}}(y - \bar{y}) + \bar{R}(y; \bar{y}).$$

Da (3.8) segue ora che

$$|y - \bar{y}| \geq \frac{1}{\lambda} |x - \bar{x}|.$$

²Si osservi che $|\Phi(x_0) - x_0| = |T(x_0)(f(x_0) - y_0)| = |T(x_0)(y_0 - \bar{y})| \leq |T(x_0)|\bar{\rho} \leq \mu\bar{\rho}$.

³Si ricordi la (3.7).

Quindi, da (3.7), segue che

$$\frac{|\bar{R}(y; \bar{y})|}{|y - \bar{y}|} = \frac{|T_{\bar{x}}R(x; \bar{x})|}{|y - \bar{y}|} \leq \lambda \frac{|R(x; \bar{x})|}{|x - \bar{x}|}. \quad (3.15)$$

Ma se y tende a \bar{y} , allora x tende a \bar{x} , e, quindi, da (3.14) e (3.15) segue che $\bar{R}(y; \bar{y}) = o(|y - \bar{y}|)$, ossia, che $g = f^{-1}$ è differenziabile in \bar{y} e che

$$g'(\bar{y}) = T_{\bar{x}} = f'(g(\bar{y}))^{-1}, \quad (3.16)$$

(che è equivalente alla (3.1)).

Infine, poiché l'inversa di una matrice è una funzione algebrica dei suoi elementi, dal fatto che $f \in C^1$ e g è continua, dalla espressione esplicita (3.16), segue, infatti, che $g \in C^1$. ■

Osservazione 3.4

- (i) Ovviamente, se $f : B \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ è invertibile (ossia, è iniettiva) la funzione inversa $f^{-1} : f(B) \rightarrow B$ è unica (segue immediatamente dall'iniettività).
- (ii) Se $f : B \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ è invertibile, la funzione inversa, per definizione, è tale che $f^{-1} \circ f(x) = x$ per ogni $x \in B$. Dall'iniettività di f , segue anche che $f \circ f^{-1}(y) = y$ per ogni $y \in V := f(B)$. In definitiva:

se $f : B \rightarrow \mathbb{R}^n$ è invertibile, allora

$$\begin{cases} f^{-1} \circ f = \text{id}, & \text{su } B, \\ f \circ f^{-1} = \text{id}, & \text{su } V = f(B). \end{cases} \quad (3.17)$$

- (iii) Una conseguenza immediata del Teorema della funzione inversa è il seguente

Teorema 3.5 ('della applicazione aperta') Sia $A \subseteq \mathbb{R}^n$ e $f \in C^1(A, \mathbb{R}^n)$ ed assumiamo che $\det f' \neq 0$ su A . Allora f è una 'applicazione aperta', ovvero, per ogni aperto $U \subseteq A$ l'insieme $f(U)$ è un insieme aperto.

Dimostrazione Sia U un sottoinsieme aperto di A , e sia $y_0 \in f(U)$. Allora esiste $x_0 \in U$ tale che $y_0 = f(x_0) \in f(U)$. Per ipotesi, $\det f'(x_0) \neq 0$ e quindi, per il Teorema della funzione inversa 3.1, l'insieme $f(B_r(x_0))$ è aperto e contiene y_0 . Ma questo significa che $f(U)$ è aperto. ■

- (iv) Un'altra conseguenza immediata del Teorema della funzione inversa è una caratterizzazione dei 'diffeomorfismi C^1 ', ossia, delle funzioni invertibili di classe C^1 con inversa C^1 :

Corollario 3.6 Sia A un aperto di \mathbb{R}^n e $f \in C^1(A, \mathbb{R}^n)$. Allora, f è invertibile su A con inversa $C^1 \iff f$ è iniettiva e $\det f' \neq 0$ su A .

Dimostrazione ' \implies ': se $g = f^{-1} \in C^1$, allora si ha $g \circ f(x) = x$ e prendendo lo jacobiano di tale relazione (cosa che possiamo fare essendo una composizione di funzioni C^1) si ha che $g' \circ f(x) f'(x) = I_n$ per ogni $x \in A$, e questo implica che $\det f' \neq 0$.

' \impliedby ': dal Teorema della funzione inversa segue che nell'intorno di ogni punto $x_0 \in A$ la funzione inversa di f è C^1 . ■

Le ipotesi del Corollario 3.6 non possono essere indebolite, come mostrano i seguenti esempi:

Esempio 3.7 La funzione $f : x \in (-1, 1) \rightarrow x^3 \in (-1, 1)$ è biunivoca su $(-1, 1)$, C^∞ ma il suo jacobiano (che in questo caso coincide con la derivata $f'(x) = 3x^2$) non è invertibile in $x = 0$. Infatti, la funzione inversa $g(y) = \sqrt[3]{y}$ è continua ma non è derivabile in $y = 0$

Esempio 3.8 La funzione⁴

$$f : A := \{x \in \mathbb{R}^2 : 0 < |x|^2 < 1\} \rightarrow (x_1^2 - x_2^2, 2x_1x_2) \in \mathbb{R}^2$$

è $C^\infty(A, \mathbb{R}^2)$, ha jacobiano $f'(x) = \begin{pmatrix} 2x_1 & -2x_2 \\ 2x_2 & 2x_1 \end{pmatrix}$ invertibile su tutto A , $f(A) = A$, ma f non è iniettiva su A e quindi non invertibile *globalmente* su A (infatti, per ogni $y \in A$ si hanno esattamente due soluzioni di $f(x) = y$ con $x \in A$).

(v) Si noti che nella versione quantitativa del Teorema della Funzione Inversa (Teorema 3.3) la scelta specifica della norma in \mathbb{R}^n è irrilevante, ma la norma delle matrici deve essere quella ‘operatoriale’ associata, cioè $\|A\| = \sup_{x \neq 0} |Ax|/|x|$.

In casi concreti è spesso più semplice usare la norma $|\cdot|_\infty$ in \mathbb{R}^n e la relativa norma operatoriale $\|\cdot\|_\infty := \|\cdot\|_{\infty, \infty}$ in (1.78); si veda l’Esempio 3.9 qui di seguito.

Esempio 3.9 Consideriamo la funzione

$$x = (x_1, x_2) \in \mathbb{R}^2 \rightarrow f(x) := (\text{sen}(x_1 + x_2), e^{x_2}(1 + x_1^4)) \in \mathbb{R}^2.$$

Allora, $f(0) = (0, 1)$ e

$$f'(x)|_{x=0} = \left(\begin{array}{cc} \cos(x_1 + x_2) & \cos(x_1 + x_2) \\ 4e^{x_2}x_1^3 & e^{x_2}(1 + x_1^4) \end{array} \right) \Big|_{x=0} = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}. \quad (3.18)$$

Dunque dal Teorema della funzione Inversa 3.1 segue che f è invertibile in un intorno dell’origine.

Cerchiamo ora, tramite il Teorema 3.3, di trovare $r > 0$ tale che f sia invertibile su $B := B_r(0)$ e $\rho > 0$ tale che $f(B) \supseteq B_\rho((0, 1))$.

L’inversa di $f'(0)$ è data da $T := \begin{pmatrix} 1 & -1 \\ 0 & 1 \end{pmatrix}$ e dunque

$$\begin{aligned} I_2 - Tf'(x) &= \begin{pmatrix} 1 - \cos(x_1 + x_2) + 4e^{x_2}x_1^3 & \cos(x_1 + x_2) - e^{x_2}(1 + x_1^4) \\ 4e^{x_2}x_1^3 & 1 - e^{x_2}(1 + x_1^4) \end{pmatrix} \\ &= \begin{pmatrix} \alpha + \varepsilon & -\alpha + \beta - \frac{1}{4}\varepsilon x_1 \\ \varepsilon & \beta - \frac{1}{4}\varepsilon x_1 \end{pmatrix}, \end{aligned}$$

dove abbiamo posto

$$\alpha := 1 - \cos(x_1 + x_2), \quad \beta := 1 - e^{x_2}, \quad \varepsilon := 4e^{x_2}x_1^3.$$

Poniamo, qui, $|\cdot| := |\cdot|_\infty$, chiamiamo $\|\cdot\|$ la relativa norma operatoriale su $\mathbb{R}^{2 \times 2}$ (ossia, $\|\cdot\| := \|\cdot\|_{\infty, \infty}$) e stimiamo il valore di $\|I_2 - Tf'(x)\|$. Assumiamo che $r \leq 1/2$, cosicché $x \in B \implies |x| \leq 1/2$. Usando le stime elementari⁵

$$|1 - \cos t| \leq t^2, \quad |e^t - 1| \leq 2t, \quad e^t \leq 2, \quad \forall |t| \leq \frac{1}{2}, \quad (3.19)$$

⁴Identificando \mathbb{R}^2 col campo complesso \mathbb{C} , ponendo $z = x_1 + ix_2$, $f(x)$ corrisponde alla mappa $z \rightarrow z^2$.

⁵Vedi, per esempio, [C2019], §5.

otteniamo

$$\begin{aligned} |\alpha + \varepsilon| + |\varepsilon| &\leq 4r^2 + 16r^3 = r(4r + 16r^2) \leq 6r, \\ \left| -\alpha + \beta - \frac{1}{4}\varepsilon x_1 \right| + \left| \beta - \frac{1}{4}\varepsilon x_1 \right| &\leq r(4r + 4 + 4r^3) < \frac{13}{2}r, \end{aligned}$$

e, quindi,

$$\sup_{B_r(0)} \|I_2 - Tf'(x)\| < 7r.$$

In definitiva, dal Teorema 3.3, segue che se prendiamo $r = 1/8$, $\theta = 7/8$, si ha che f è invertibile sul quadrato di lato⁶ $1/4$ centrato nell'origine; la costante λ può essere presa come $\lambda = \|T\|(1 - \theta)^{-1} = 16$ e l'immagine $f(B_{1/2}(0))$ contiene un quadrato di lato $1/64$ centrato in $(0, 1)$.

2 Teorema delle Funzioni Implicite

Il Teorema della funzione Inversa permette, tra l'altro, di risolvere un problema classico, ossia, di determinare l'esistenza (e le proprietà) di 'funzioni definite implicitamente da un sistema di equazioni'.

Consideriamo il problema di trovare le soluzioni di un sistema di n equazioni (nelle incognite x_1, \dots, x_n) che contengano m parametri. In altre parole vogliamo trovare $x \in \mathbb{R}^n$ che soddisfi il sistema di n equazioni

$$f(x, \xi) = 0, \quad \text{ovvero} \quad \begin{cases} f_1(x_1, \dots, x_n, \xi_1, \dots, \xi_m) = 0 \\ f_2(x_1, \dots, x_n, \xi_1, \dots, \xi_m) = 0 \\ \vdots \\ f_n(x_1, \dots, x_n, \xi_1, \dots, \xi_m) = 0 \end{cases} \quad (3.20)$$

al variare dei "parametri" $\xi \in \mathbb{R}^m$ avendo assegnato la funzione $f : A \subseteq \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n$. Esempi semplici di natura geometrica sono dati dai grafici di funzioni.

(a) Sia $f(x, \xi) = x^3 - \xi$, ($n = m = 1$); in tal caso le soluzioni, al variare di $\xi \in \mathbb{R}$, di $f(x, \xi) = 0$ sono $x = \sqrt[3]{\xi}$ cioè l'insieme delle soluzioni di $f(x, \xi) = 0$, $\{(x, \xi) \in \mathbb{R}^2 : f(x, \xi) = 0\}$, coincide con il grafico di $\sqrt[3]{\xi}$ ovvero con l'insieme $\{(\sqrt[3]{\xi}, \xi) : \xi \in \mathbb{R}\}$.

(b) Sia $f(x, \xi) = x^2 - \xi$ (ancora $n = m = 1$). Se $\xi < 0$ non esistono soluzioni di $f(x, \xi) = 0$; se $\xi = 0$, l'unica soluzione è data da $x = 0$; se $\xi > 0$ vi sono due soluzioni date da $x = \pm\sqrt{\xi}$. Quindi, fissato $\xi_0 > 0$, in un intorno (sufficientemente piccolo) di $(\xi_0, x_0) := (\xi_0, \sqrt{\xi_0})$ vi è una unica funzione di ξ che soddisfi identicamente $f(x(\xi), \xi) = 0$, tale funzione essendo, ovviamente, $x(\xi) = +\sqrt{\xi}$ mentre in un intorno di $(\xi_0, x_0) := (\xi_0, -\sqrt{\xi_0})$, avremo la soluzione $x(\xi) = -\sqrt{\xi}$. In un intorno di $(\xi_0, x_0) = (0, 0)$ invece non è possibile definire in maniera univoca le soluzioni di $f(x, \xi) = 0$ in termini di funzioni di ξ . Il punto $(0, 0)$ è dunque, rispetto al nostro problema, "singolare": il numero di soluzioni varia tra 0 e 2 quando $\xi \in (-\varepsilon, \varepsilon)$, comunque piccolo sia $\varepsilon > 0$. Notiamo anche che in tale punto singolare si ha $f_x = 0$.

Già da questi semplicissimi esempi si vede che una formulazione più propria del 'problema delle funzioni implicite' è:

Dato un punto (x_0, ξ_0) tale che $f(x_0, \xi_0) = 0$, è possibile trovare una funzione $x = g(\xi)$ tale che $g(\xi_0) = x_0$ e tale che il grafico di g rappresenti localmente tutte le soluzioni di⁷ $f(x, \xi) = 0$?

(c) Cambiamo punto di vista e diamo un esempio analitico con n ed m arbitrari. Consideriamo il sistema di equazioni *lineari* nelle incognite x_1, \dots, x_n con coefficienti che dipendono da $\xi \in \mathbb{R}^m$:

$$A(\xi)x + b(\xi) = 0, \quad (3.21)$$

⁶Poiché stiamo usando la norma $|\cdot|_\infty$ in \mathbb{R}^2 le 'sfere' di raggio r sono dei quadrati di lato $2r$.

⁷Ovvero $f(g(\xi), \xi) = 0$ in un intorno di ξ_0 e l'insieme delle coppie $(g(\xi), \xi)$ al variare di ξ in tale intorno sono *tutte* le soluzioni di $f(x, \xi) = 0$ in un intorno di (x_0, ξ_0) .

dove $A(\xi) \in \mathbb{R}^{n \times n}$ e $b(\xi) \in \mathbb{R}^n$. È evidente che tale sistema è unicamente risolubile se $\det A(\xi) \neq 0$, nel qual caso si ha

$$x = -A(\xi)^{-1}b(\xi) \quad (3.22)$$

e se $A(\xi)$ e $b(\xi)$ sono funzioni continue di ξ , così risulterà anche la funzione $x(\xi)$ definita dalla (3.22). Si noti che, se poniamo $f(x, \xi) = A(\xi)x + b(\xi)$, $A(\xi)$ coincide con la matrice jacobiana f_x e dunque la condizione di risolubilità è che tale matrice jacobiana sia invertibile.

Il Teorma fondamentale in questo contesto è il seguente

Teorema 3.10 (Teorema delle Funzioni Implicite o ‘Teorema del Dini’)

Sia A un aperto di $\mathbb{R}^n \times \mathbb{R}^m$, $f : (x, \xi) \in A \mapsto f(x, \xi) \in C^1(A, \mathbb{R}^n)$ e $(x_0, \xi_0) \in A$ tali che

$$\begin{cases} f(x_0, \xi_0) = 0, \\ \det f_x(x_0, \xi_0) \neq 0. \end{cases} \quad (3.23)$$

Allora, esiste $\rho > 0$ ed una unica funzione $g \in C^1(U, \mathbb{R}^n)$, con⁸ $U := B_\rho^m(\xi_0) \subseteq \mathbb{R}^m$, tale che $(g(\xi), \xi) \in A$, per ogni $\xi \in U$, e

$$\begin{cases} g(\xi_0) = x_0; \\ f(g(\xi), \xi) = 0, \quad \forall \xi \in U. \end{cases} \quad (3.24)$$

Inoltre la matrice $(n \times n)$ f_x è invertibile su $B_\rho^{n+m}(x_0, \xi_0)$ e la matrice jacobiana di g è data da

$$g_\xi(\xi) = -(f_x(g(\xi), \xi))^{-1} f_\xi(g(\xi), \xi), \quad \forall \xi \in U. \quad (3.25)$$

Dimostrazione Poniamo

$$\begin{aligned} X &:= (x, \xi) \in \mathbb{R}^n \times \mathbb{R}^m, & Y &:= (y, \eta) \in \mathbb{R}^n \times \mathbb{R}^m, & X_0 &:= (x_0, \xi_0), \\ F(X) &:= (f(X), \xi). \end{aligned} \quad (3.26)$$

Allora $F \in C^1(A, \mathbb{R}^n \times \mathbb{R}^m)$ e calcolando lo jacobiano di F troviamo:

$$F' = \begin{pmatrix} f_x & f_\xi \\ 0 & I_m \end{pmatrix}, \quad (3.27)$$

dove $f_x \in \mathbb{R}^{n \times n}$, $f_\xi \in \mathbb{R}^{n \times m}$, I_m è la matrice identità $(m \times m)$ (e lo 0 è una matrice $(m \times n)$ di tutti 0). Dunque $\det F'(X_0) = \det f_x(X_0) = \det f_x(x_0, \xi_0) \neq 0$ e le ipotesi del Teorema della Funzione Inversa 3.1 sono soddisfatte da F . Quindi, esiste una sfera $B := B_r^{n+m}(x_0, \xi_0)$ tale che: F è invertibile su B con inversa C^1 sull'aperto $V := F(B)$; inoltre, la matrice jacobiana F' è invertibile su B con inversa data da⁹

$$(F^{-1})' = \begin{pmatrix} (f_x)^{-1} & -(f_x)^{-1}f_\xi \\ 0 & I_m \end{pmatrix}, \quad (3.28)$$

dove la matrice $(F^{-1})'$ è calcolata in Y , mentre le matrici $(f_x)^{-1}$ e f_ξ sono calcolate in $X = f^{-1}(Y)$.

Poiché $V \subseteq \mathbb{R}^{n+m}$ è un aperto che contiene $Y_0 := F(X_0) = (0, \xi_0)$, esiste $\rho > 0$ tale che $B_\rho^n(0) \times B_\rho^m(\xi_0) \subseteq V$ e $G := F^{-1} : B_\rho^n(0) \times B_\rho^m(\xi_0) \rightarrow B$. Poniamo $(\alpha(Y), \beta(Y)) := G(Y) \in \mathbb{R}^n \times \mathbb{R}^m$ e osserviamo che $G(0, \xi_0) = (x_0, \xi_0)$, cioè $\alpha(0, \xi_0) = x_0$ e $\beta(0, \xi_0) = \xi_0$. Ora, la relazione $F \circ G = \text{id}$ equivale a

$$\begin{cases} f(\alpha(Y), \beta(Y)) = y, \\ \beta(Y) = \eta, \end{cases} \quad \implies \quad f(\alpha(y, \eta), \eta) = y, \quad \implies \quad f(\alpha(0, \eta), \eta) = 0,$$

⁸In questa sezione indicheremo le dimensioni delle sfere con un suffisso.

⁹Cfr. Appendice D.

relazioni valide per ogni $\eta \in B_\rho^m(\xi_0)$. Ponendo $g(\xi) := \alpha(0, \xi)$ (avendo posto $\eta = \xi$), si ha la¹⁰ (3.24).

Infine, osservando che il gradiente della funzione $\xi \rightarrow f(g(\xi), \xi)$ è identicamente nullo, si ha che¹¹

$$0 = \partial_\xi f(g(\xi), \xi) = f_x(g(\xi), \xi) g_\xi(\xi) + f_\xi(g(\xi), \xi),$$

da cui (essendo $f_x(g(\xi), \xi)$ invertibile su U) segue (3.25). ■

Osservazione 3.11 (i) Nel Complemento 3.1 daremo una dimostrazione indipendente (ossia non basata sul Teorema della Funzione Inversa) di questo teorema – interamente basata sul Lemma della Contrazioni (su un sottoinsieme chiuso dello spazio di Banach delle funzioni continue) – sotto ipotesi di regolarità più generali. Inoltre, il Teorema delle Funzioni Implicite discusso nel Complemento 3.1 è *quantitativo* e, in particolare, viene data una stima esplicita del raggio ρ che appare nell’enunciato del Teorema 3.10.

(ii) Qui sopra abbiamo dedotto il Teorema delle Funzioni Implicite dal Teorema della Funzione Inversa, ma in effetti, è vero anche il viceversa, ossia, è possibile dedurre il Teorema della Funzione Inversa dal Teorema delle Funzioni Implicite. Supponiamo, infatti, che $f \in C^1(B_{r_0}(x_0), \mathbb{R}^n)$ sia tale che $\det f'(x_0) \neq 0$. Allora, se poniamo, per $y \in \mathbb{R}^n$, $F(x, y) := f(x) - y$ e $y_0 := f(x_0)$, si ha che $F \in C^1(B_{r_0}(x_0) \times B_R(y_0), \mathbb{R}^n)$, per qualunque $R > 0$ e $\det F_x(x_0, y_0) = \det f'(x_0) \neq 0$. Quindi, dal Teorema 3.10, segue che esiste una unica funzione g di classe $C^1(B_\rho(y_0), B_r(x_0))$, per un opportuno $0 < \rho \leq R$, tale che $g(y_0) = x_0$ e $F(g(y), y) = 0$ per ogni $y \in B_\rho(y_0)$, ossia, $f \circ g(y) = y$ per ogni $y \in B_\rho(y_0)$, il che vuol dire che g è la funzione inversa di f .

3 Problemi vincolati

Vediamo, in questo paragrafo, un’interessante applicazione del teorema delle funzioni implicite che permette di risolvere, in certi casi, *problemi di massimo o minimo vincolati*.

Supponiamo che un insieme compatto in \mathbb{R}^n sia dato come l’insieme $K := \{x : \phi(x) \leq 0\}$, con ϕ continua, e sia f una funzione continua su K . Dal teorema di Weierstrass sappiamo che f ammette massimo e minimo su K ed il problema che ci poniamo è quello di sviluppare dei metodi che permettano di determinare i punti estremali ossia i punti di massimo o minimo. Per far questo assumeremo un po’ di regolarità ossia assumiamo che f e ϕ siano di classe C^1 . Sappiamo che se un punto estremo è interno a K allora vi si annulla il gradiente di f e questo permette di selezionare i punti interni candidati ad essere estremali. Cosa possiamo dire se gli estremali di f sono sulla frontiera di K , che è contenuta nell’insieme $E_0 := \{x : \phi(x) = 0\}$? A tale proposito sussiste il seguente risultato dovuto a Lagrange. Prima una definizione:

Definizione 3.12 Sia A un aperto di \mathbb{R}^n , $\phi \in C^1(A, \mathbb{R})$ e

$$E_0 := \{x \in A : \phi(x) = 0\}. \quad (3.29)$$

Un punto $x_0 \in E_0$ si dice **regolare per E_0** se $\nabla\phi(x_0) \neq 0$.

A volte, l’insieme $E_0 = \{\phi = 0\}$ viene chiamato **vincolo**.

Proposizione 3.13 Siano ϕ e E_0 come nella Definizione 3.12 e sia f una funzione $C^1(A, \mathbb{R})$. Se $x_0 \in E_0$ è un punto regolare per E_0 ed è un massimo o un minimo relativo di f su E_0 , allora esiste un numero $\lambda \in \mathbb{R}$ tale che $\nabla f(x_0) = \lambda \nabla\phi(x_0)$.

¹⁰Si osservi che $g(\xi_0) = \alpha(0, \xi_0) = 0$, che la posizione $\xi = \eta$ è un irrilevante cambio di nomi e, infine, che l’unicità di g segue dall’unicità della funzione inversa.

¹¹Usando la notazione di matrici a blocchi: $\partial_\xi f(g(\xi), \xi) = \partial_\xi f \circ G(0, \xi) = f'(G(0, \xi)) \partial_\xi G(0, \xi) = (f_x \ f_\xi) \begin{pmatrix} g_\xi \\ I_m \end{pmatrix} = f_x g_\xi + f_\xi$.

Dimostrazione Poiché x_0 è regolare per E_0 , una delle derivate parziali di ϕ è non nulla in $x_0 = (x_{01}, \dots, x_{0n})$; supponiamo, per fissare le idee, che sia la prima variabile: $D_1\phi(x_0) \neq 0$. Poniamo $y := x_1$, $\xi = (x_2, \dots, x_n)$, $y_0 := x_{01}$ e $\xi_0 := (x_{02}, \dots, x_{0n})$. Allora, $\phi_y(y_0, \xi_0) \neq 0$ e per il Teorema delle Funzioni Implicite 3.10 esiste un intorno $U \subseteq \mathbb{R}^{n-1}$ di ξ_0 e una funzione $g \in C^1(U, \mathbb{R})$ tale che $g(\xi_0) = x_{01}$ e $\phi(g(\xi), \xi) = 0$, per ogni $\xi \in U$. Se x_0 è, per esempio, un massimo locale per f su E_0 , esiste $r > 0$ tale che $B_r^n(x_0) \subseteq A$ e $f(y_0, \xi_0) \geq f(g(\xi), \xi)$ per ogni $\xi \in U \cap B_r^{n-1}(\xi_0)$. Essendo $U \cap B_r^{n-1}(\xi_0)$ un intorno di ξ_0 , per il punto (i) della Proposizione 2.20, il gradiente della funzione $\xi \rightarrow f(g(\xi), \xi)$ è nullo in ξ_0 :

$$\frac{\partial}{\partial \xi} f(g(\xi), \xi)|_{\xi=\xi_0} = 0,$$

ossia,

$$\frac{\partial f}{\partial y}(y_0, \xi_0) \frac{\partial g}{\partial \xi}(\xi_0) + \frac{\partial f}{\partial \xi}(y_0, \xi_0) = 0, \quad (3.30)$$

relazione che, in vista di (3.25), diviene

$$\frac{\partial f}{\partial \xi}(y_0, \xi_0) = \frac{\partial f}{\partial y}(y_0, \xi_0) \left(\frac{\partial \phi}{\partial y}(y_0, \xi_0) \right)^{-1} \frac{\partial \phi}{\partial \xi}(y_0, \xi_0). \quad (3.31)$$

Ma questo significa che per $i = 2, \dots, n$

$$D_i f(x_0) = \lambda \frac{\partial \phi}{\partial x_i}, \quad \text{con } \lambda := \frac{\partial f}{\partial y}(x_0) \left(\frac{\partial \phi}{\partial y}(x_0) \right)^{-1}. \quad (3.32)$$

Per $i = 1$ la (3.32) è ovviamente vera. ■

Osservazione 3.14 (Metodo dei moltiplicatori di Lagrange) Si noti che:

$\exists x_0 \in E_0, \lambda_0 \in \mathbb{R}$ tali che $\nabla f(x_0) = \lambda_0 \nabla \phi(x_0) \iff (x_0, \lambda_0)$ è un punto critico per la funzione di $(n+1)$ variabili $F(x, \lambda) := f(x) - \lambda \phi(x)$.

Infatti, $F_\lambda(x_0, \lambda_0) = 0$ equivale a $x_0 \in E_0$, e $F_x = 0$ equivale a $\nabla f(x_0) = \lambda_0 \nabla \phi(x_0)$.

Il cosiddetto ‘metodo dei moltiplicatori di Lagrange’ consiste nel cercare di individuare punti estremali di f su di un vincolo $\{\phi = 0\}$ tra i punti critici della funzione $F(x, \lambda) = f(x) - \lambda \phi(x)$.

Esempio 3.15 Sia $f(x) = \prod_{i=1}^n x_i$, $M > 0$ e sia $B := \{x \in \mathbb{R}^n \mid x_i \geq 0 \text{ e } \sum_{i=1}^n x_i = M\}$. L’insieme B è compatto e si vuole determinare il massimo ed il minimo (assoluti) di f su B . Poiché $f \geq 0$ su B , i punti della frontiera in cui una delle coordinate è nulla (e sui quali $f = 0$) sono punti di minimo assoluto. Poiché $f_{x_j} = \prod_{i \neq j} x_i$, ∇f non si annulla mai all’interno di B ; dunque il massimo di f è raggiunto su $A := \partial B \cap \{x \in \mathbb{R}^n : x_i > 0 \text{ per ogni } i \text{ e } \sum_{i=1}^n x_i = M\}$. Per trovare i punti critici su A usiamo il metodo dei moltiplicatori di Lagrange. Sia $F(x, \lambda) := f(x) - \lambda \phi(x)$ con $\phi(x) = \sum_{i=1}^n x_i - M$. Poiché $F_{x_j} = \prod_{i \neq j} x_i - \lambda$ (e $x_i > 0$ su A) $F_{x_j} = 0$ su A se e solo se $f(x) = \lambda x_j$ per ogni j . Dunque $x_1 = \dots = x_n$. Il massimo è dunque raggiunto su x_0 dove $x_{0j} = M/n$. Vale dunque

$$\left(\prod_{i=1}^n x_i \right)^{\frac{1}{n}} \leq \frac{1}{n} \sum_{i=1}^n x_i, \quad x_i \geq 0, \quad (3.33)$$

relazione nota come *disuguaglianza aritmetico-geometrica*.

Le tecniche sopra esposte si generalizzano facilmente al caso di più vincoli ossia nel caso in cui la funzione ϕ che definisce il vincolo sia una funzione vettoriale $\phi \in C(\mathbb{R}^n, \mathbb{R}^m)$ (con $m < n$) in modo tale che

$$E_0 := \{\phi = 0\} = \bigcap_{i=1}^m \{\phi_i = 0\}.$$

L’analogo della Definizione 3.12 è:

Definizione 3.16 Sia A un aperto di \mathbb{R}^n , $\phi \in C^1(A, \mathbb{R}^m)$ e $E_0 := \{x \in A : \phi(x) = 0\}$. Un punto $x_0 \in E_0 := \{x \in A : \phi(x) = 0\}$ si dice regolare per E_0 se la matrice jacobiana $\phi_x(x_0)$ ha rango massimo (e cioè uguale ad m).

La Proposizione 3.13 si generalizza, allora, come segue:

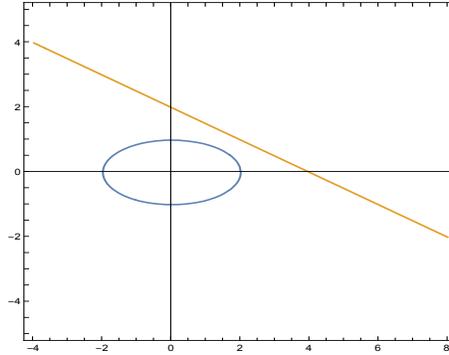
Proposizione 3.17 Siano ϕ e E_0 come nella Definizione 3.16 e sia f una funzione $C^1(A, \mathbb{R})$. Se x_0 è un punto regolare per E_0 ed è un massimo o un minimo relativo di f su E_0 , allora esiste un vettore $\lambda \in \mathbb{R}^m$ tale che¹² $f'(x_0) = \lambda^T \phi_x(x_0)$.

La dimostrazione è una ripetizione parola per parola della dimostrazione della Proposizione 3.13 (con la giusta interpretazione dei simboli!).

Il metodo dei moltiplicatori di Lagrange, in questo caso, consiste nel cercare di individuare punti estremali di f su di un vincolo $\{\phi = 0\} = \bigcap_{i=1}^m \{\phi_i = 0\}$ tra i punti critici della funzione $F(x, \lambda) = f(x) - \lambda \cdot \phi(x)$.

Esempio 3.18 Si calcoli la distanza d tra l'ellisse $\mathcal{E} := \{x \in \mathbb{R}^2 : \frac{x_1^2}{4} + x_2^2 = 1\}$ e la retta $R := \{x \in \mathbb{R}^2 : x_1 + 2x_2 = 4\}$. Si ricorda che la distanza tra due insiemi \mathcal{E} ed R è data da

$$\text{dist}(\mathcal{E}, R) := \inf_{\substack{x \in \mathcal{E} \\ y \in R}} |x - y|.$$



Per calcolare d usiamo il metodo dei moltiplicatori di Lagrange su \mathbb{R}^4 con:

$$\begin{aligned} f(x, y) &:= |x - y|^2, \quad (x, y \in \mathbb{R}^2), \\ \phi_1(x, y) &:= \frac{x_1^2}{4} + x_2^2 - 1, \quad \phi_2(x, y) := y_1 + 2y_2 - 4, \\ F(x, y, \lambda) &:= f(x, y) - \lambda_1 \phi_1(x, y) - \lambda_2 \phi_2(x, y). \end{aligned}$$

Eguagliando a zero le derivate rispetto ad x e y di F si ottiene

$$\begin{aligned} 2(x_1 - y_1) - \lambda_1 \frac{x_1}{2} &= 0 \\ x_2 - y_2 - \lambda_1 x_2 &= 0 \\ 2(y_1 - x_1) - \lambda_2 &= 0 \\ y_2 - x_2 - \lambda_2 &= 0. \end{aligned}$$

Dalle ultime due equazioni otteniamo $2(x_1 - y_1) = x_2 - y_2$ e quindi dalle prime due equazioni segue che $x_1 = 2x_2$ (avendo osservato che $\lambda_1 \neq 0$ altrimenti si avrebbe $x = y$ il che è impossibile essendo $\mathcal{E} \cap R = \emptyset$). Da tale relazione, assieme a $\phi_1 = 0$, si ottiene che $x =$

¹² In questa formula, f' è una matrice $(1 \times n)$ e λ^T una matrice $(1 \times m)$ (il simbolo di 'trasposto' è presente in quanto solitamente i vettori in \mathbb{R}^m si interpretano come vettori 'colonna', ossia, matrici $(m \times 1)$).

$(2/\sqrt{2}, 1/\sqrt{2})$ (avendo scartato, per ovvi motivi geometrici, l'altra soluzione nel quadrante $\{x_i < 0\}$).

Infine da $2y_1 - y_2 = 2x_1 - x_2 = 3/\sqrt{2}$ e da $\phi_2 = 0$ si ottiene $y = \left(\frac{2}{5}(2 + \frac{3}{\sqrt{2}}), \frac{1}{5}(8 - \frac{3}{\sqrt{2}})\right)$ e quindi $d = 2 \frac{\sqrt{5}}{5} (2 - \sqrt{2})$.

Complementi

Complemento 3.1: Teorema quantitativo delle Funzioni Implicite

Teorema 3.19 Sia $X_0 \subseteq \mathbb{R}^m$ un insieme compatto, $Y_0 := \{y \in \mathbb{R}^n : |y - y_0| \leq \rho\}$ una sfera di raggio ρ in \mathbb{R}^n e $F : (y, x) \in Y_0 \times X_0 \rightarrow F(y, x) \in \mathbb{R}^n$ una funzione continua su $Y_0 \times X_0$ assieme alle sue derivate parziali rispetto alle y_j . Assumiamo che esista una matrice $T \in \mathbb{R}^{n \times n}$ per cui valgano le seguenti due condizioni:

$$\sup_{X_0} |F(y_0, x)| \leq \frac{\rho}{2\|T\|}, \quad (3.34)$$

$$\sup_{Y_0 \times X_0} \|I - T \frac{\partial F}{\partial y}(y, x)\| \leq \frac{1}{2} \quad (3.35)$$

dove $I \in \mathbb{R}^{n \times n}$ è la matrice identità. Allora esiste un'unica $g \in C(X_0, Y_0)$ che soddisfa $F(g(x), x) = 0$ per ogni $x \in X_0$. Inoltre, il grafico $\{(y, x) = (g(x), x) : x \in X_0\}$ rappresenta tutte e sole le soluzioni di $F(y, x) = 0$ in $Y_0 \times X_0$.

Osservazione 3.20 (i) Sotto le ipotesi enunciate nella Proposizione 3.10, è sempre possibile trovare ρ , X_0 e T per cui le due condizioni del Teorema 3.19 valgano. Infatti, poniamo

$$X_0 = \{x \in \mathbb{R}^m : |x - x_0| \leq r\}, \quad T := \left(\frac{\partial F}{\partial y}(y_0, x_0)\right)^{-1}, \quad (3.36)$$

e vediamo come, grazie alle ipotesi della Proposizione 3.10, sia possibile determinare r e ρ in modo tale da soddisfare (3.34) ed (3.35). Poiché $x \rightarrow F(y_0, x)$ è continua in x_0 e $F(y_0, x_0) = 0$, si ha che, per ogni $\rho > 0$, è sempre possibile scegliere $r = r(\rho)$ per il quale (3.34) è soddisfatta (nelle notazioni standard ρ corrisponde ad ε e r a δ); dopodiché, essendo anche $(y, x) \rightarrow \|I - T \frac{\partial F}{\partial y}(y, x)\|$ continua in (y_0, x_0) dove vale 0, esistono ρ_1 e r_1 tali che $(y, x) \rightarrow \|I - T \frac{\partial F}{\partial y}(y, x)\| \leq 1/2$ per ogni y e x tali che $|y - y_0| \leq \rho_1$ e $|x - x_0| \leq r_1$ (qui ε corrisponde a $1/2$, ρ_1 e r_1 a δ). Dunque, scegliendo

$$\rho := \rho_1, \quad r := \min\{r_1, r(\rho_1)\}$$

le (3.34) e (3.35) saranno soddisfatte.

Viceversa l'enunciato 'quantitativo' nel Teorema 3.19 è più generale dell'enunciato della Proposizione 3.10: infatti non viene assunto che $F(y_0, x_0) = 0$ ossia non viene assunta l'esistenza di una soluzione esatta di $F = 0$; d'altra parte si assume che esistano, per valori dei parametri x in X_0 , delle soluzioni approssimate di $F = 0$ (si veda (3.34) dove ρ dovrà essere 'piccolo' per poter soddisfare (3.35)).

(ii) L'enunciato del Teorema 3.19 potrebbe, a prima vista, apparire più generale dell'enunciato della Proposizione 3.10 anche perché non viene assunto esplicitamente che la matrice jacobiana $\frac{\partial F}{\partial y}$ è invertibile per un qualche punto di $Y_0 \times X_0$; ma, in effetti, la (3.35) implica che $\frac{\partial F}{\partial y}$ è invertibile per ogni punto di $Y_0 \times X_0$. Infatti, se poniamo $A := I - T \frac{\partial F}{\partial y}$, da (iii) della Proposizione 1.71 segue che la matrice $I - A = T \frac{\partial F}{\partial y}$ è invertibile, e quindi, sia T che $\frac{\partial F}{\partial y}$, per ogni y e x , sono invertibili.

(iii) (i) Il fattore $\frac{1}{2}$ che compare nelle condizioni (3.34) e (3.35) è alquanto arbitrario ed in effetti dalla dimostrazione del teorema segue facilmente che le condizioni (3.34) e (3.35) possono venire sostituite, rispettivamente, con

$$\sup_{X_0} |F(y_0, x)| \leq \alpha \frac{\rho}{\|T\|}, \quad T := \left(\frac{\partial F}{\partial y}(y_0, x_0)\right)^{-1}; \quad (3.37)$$

e da

$$\sup_{Y_0 \times X_0} \|I - T \frac{\partial F}{\partial y}(y, x)\| \leq (1 - \alpha) \quad (3.38)$$

con $0 < \alpha < 1$ prefissato.

(iv) Le norme usate nel Teorema 3.19 sono state fissate per semplicità ma è immediato verificare nella dimostrazione che segue che il Teorema 3.19 vale per qualunque scelta delle norme in \mathbb{R}^n e \mathbb{R}^m e con la relativa norma (1.76) su $\mathbb{R}^{n \times n}$. Per esempio, è spesso più conveniente nelle applicazioni usare le norme $|\cdot|_\infty$ in \mathbb{R}^n e \mathbb{R}^m con la relativa norma $\|\cdot\|_{\infty, \infty}$, (1.78), sulle matrici $n \times n$.

(v) Il nome ‘funzioni implicite’ deriva dal fatto che la relazione (3.20), sotto le ipotesi del Teorema 3.10, definisce implicitamente la funzione $y = g(x)$.

Dimostrazione (del Teorema 3.19) Poniamo $G(y, x) = y - TF(y, x)$, cosicché $\frac{\partial G}{\partial y} = I - T \frac{\partial F}{\partial y}$. Allora, per ogni $y, z \in Y_0$ e $x \in X_0$, in virtù del Teorema fondamentale del calcolo di (1.66), di (1.77) e di (3.35), si ha

$$\begin{aligned} |G(y, x) - G(z, x)| &= \left| \int_0^1 \frac{\partial G}{\partial y}(z + t(y - z), x)[y - z] dt \right| \\ &\leq \sup_{Y_0 \times X_0} \left\| \frac{\partial G}{\partial y} \right\| |y - z| \leq \frac{1}{2} |y - z|. \end{aligned} \quad (3.39)$$

Sia, ora, $E := C(X_0, Y_0)$. Si noti che E è un sottoinsieme chiuso dello spazio di Banach¹³ $C(X_0, \mathbb{R}^n)$ (dotato della norma uniforme $\|\cdot\|_{\infty, X_0}$). Sull'insieme E definiamo la seguente mappa Φ che associa ad una funzione $u \in C(X_0, Y_0)$ la funzione $\Phi(u)$ definita da

$$\Phi(u)(x) := G(u(x), x) = u(x) - TF(u(x), x). \quad (3.40)$$

Si noti che tale definizione è ben posta poiché i valori $u(x)$ cadono nella sfera chiusa Y_0 su cui $F(\cdot, x)$ è continua. Per ogni $u, v \in E$, e per ogni $x \in X_0$, da (3.39) (con $y = u(x)$ e $z = v(x)$) segue che

$$|\Phi(u)(x) - \Phi(v)(x)| \leq \frac{1}{2} |u(x) - v(x)| \quad (3.41)$$

e, prendendo l'estremo superiore su X_0 , vediamo che

$$\|\Phi(u) - \Phi(v)\|_{\infty, X_0} \leq \frac{1}{2} \|u - v\|_{\infty, X_0}. \quad (3.42)$$

Cioè Φ è una contrazione sullo spazio di Banach $C(X_0)$. Chiamiamo $u_0(x)$ la funzione costante $u_0(x) := y_0$; tale funzione è ovviamente in E e da (3.42), la definizione di Φ e da (3.34), si ottiene¹⁴, per ogni $u \in E$,

$$\begin{aligned} \|\Phi(u) - y_0\|_\infty &\leq \|\Phi(u) - \Phi(u_0)\|_\infty + \|\Phi(u_0) - y_0\|_\infty \leq \\ &\leq \frac{1}{2} \|u - u_0\|_\infty + \|TF(y_0, x)\|_\infty \\ &\leq \frac{1}{2} \|u - y_0\|_\infty + \|T\| \frac{\rho}{2\|T\|} \leq \rho, \end{aligned} \quad (3.43)$$

cioè $\Phi(u)(x) \in Y_0$ per ogni $x \in X_0$, il che equivale a dire che $\Phi(u) \in C(X_0, Y_0) = E$. Quindi Φ manda E in se stesso ed è una contrazione. Per il Lemma delle contrazioni 1.49 possiamo concludere che esiste una unica funzione $g \in C(X_0, Y_0)$ tale che $\Phi(g) = g$, ossia $g - TF(g, x) = g$, ma questa relazione è equivalente a

$$F(g(x), x) = 0, \quad \forall x \in X_0. \quad (3.44)$$

Dimostriamo ora che il grafico $\{(y, x) = (g(x), x) : x \in X_0\}$ è l'insieme di tutte le soluzioni di $F(y, x) = 0$ in $Y_0 \times X_0$. Per dimostrare tale affermazione, supponiamo, per assurdo, che esista $\bar{x} \in X_0$ ed un $\bar{y} \in Y_0$ tale che $\bar{y} \neq \bar{g} := g(\bar{x})$ e tale che $F(\bar{y}, \bar{x}) = 0$. Allora, dalla definizione di G e da (3.39) seguirebbe che

$$|\bar{y} - \bar{g}| = |G(\bar{y}, \bar{x}) - G(\bar{g}, \bar{x})| \leq \frac{1}{2} |\bar{y} - \bar{g}| \quad (3.45)$$

il che è assurdo (avendo noi assunto che $\bar{y} \neq \bar{g}$). ■

Dimostrazione (della Proposizione 3.10) Come detto in (i) dell'Osservazione 3.20, le ipotesi della Proposizione 3.10 implicano quelle del Teorema 3.19. Quindi l'unica cosa che resta da dimostrare è

¹³ $u_k \in E$ significa $u_k \in C(X_0, \mathbb{R}^n)$ e $|u_k(x) - y_0| \leq \rho$ per ogni $x \in X_0$; se $\{u_k\}$ è di Cauchy allora, poiché $(C(X_0, \cdot), \|\cdot\|_{\infty, X_0})$ è uno spazio di Banach, esiste $u \in C(X_0, \mathbb{R}^n)$ tale che $\sup_{x \in X_0} |u_k(x) - u(x)| \rightarrow 0$ e quindi, per ogni $x \in X_0$, $|u(x) - y_0| = \lim_{k \rightarrow \infty} |u_k(x) - y_0| \leq \rho$ e quindi $u \in E$.

¹⁴ Qui $\|\cdot\|_\infty$ denota, per brevità, $\|\cdot\|_{\infty, X_0}$.

che $F(y_0, x_0) = 0$ implica che $g(x_0) = y_0$. Ma questo deriva dall'ultima affermazione nell'enunciato del Teorema 3.19 che assicura che tutte le soluzioni di $F = 0$ in $Y_0 \times X_0$ (e $(y_0, x_0) \in Y_0 \times X_0$) sono date dal grafico su X_0 di g il che implica che $g(x_0) = y_0$. ■

Esempio 3.21 Sia $F(y, x_1, x_2) = e^y - \text{sen}(yx_1 + \frac{x_2\pi}{2})$, con $y \in \mathbb{R}$ e $x = (x_1, x_2) \in \mathbb{R}^2$, e sia $(y_0, x_0) := (y_0, x_{01}, x_{02}) = (0, 1, 1)$. Mostriamo che la Proposizione 3.10 è applicabile in un intorno di (y_0, x_0) e troviamo dei valori di ρ e r per cui (3.34) e (3.35) vengano soddisfatte con $X_0 = \{x : |x - x_0| \leq r\}$.

$F(0, 1, 1) = 0$ e $\frac{\partial F}{\partial y} = e^y - x_1 \cos(yx_1 + \frac{x_2\pi}{2})$, e $\frac{\partial F}{\partial y}(0, 1, 1) = 1$. Quindi il teorema è applicabile nell'intorno di (y_0, x_0) . Diamo ora delle stime concrete su ρ e r ; per semplificare tali stime useremo le seguenti disuguaglianze elementari¹⁵:

$$\begin{aligned} |1 - \cos y| &\leq \frac{y^2}{2}, \quad \forall y \in \mathbb{R}, \\ |e^y - 1| &\leq (e - 1)|y| < 2|y|, \quad \forall |y| \leq 1. \end{aligned} \quad (3.46)$$

Poiché $F(y_0, x) = 1 - \text{sen} \frac{x_2\pi}{2}$, scrivendo $x_2 = 1 + a$ con $|a| \leq r$ ed usando la prima disuguaglianza in (3.46) otteniamo

$$\begin{aligned} \sup_{X_0} |F(y_0, x)| &= \sup_{|x_2-1| \leq r} \left| 1 - \text{sen} \frac{x_2\pi}{2} \right| \\ &= \sup_{|a| \leq r} \left| 1 - \text{sen} \frac{(1+a)\pi}{2} \right| = \sup_{|a| \leq r} \left| 1 - \cos \frac{a\pi}{2} \right| \\ &\leq \frac{1}{2} \left(\frac{\pi r}{2} \right)^2 < 2r^2. \end{aligned} \quad (3.47)$$

Quindi, poiché $T = 1$ (e quindi $\|T\| = 1$), la condizione (3.34) è soddisfatta se scegliamo $r = \rho/2$. Passiamo alla condizione (3.35). Per ogni $|y| \leq \rho$ e $|x - (1, 1)| \leq r$, scrivendo ancora $x_2 = 1 + a$ con $|a| \leq r$, usando la seconda disuguaglianza in (3.46) e maggiorando $|\text{sen } t|$ con $|t|$, si ha

$$\begin{aligned} \left| 1 - \frac{\partial F}{\partial y} \right| &= \left| e^y - 1 - x_1 \cos(yx_1 + \frac{x_2\pi}{2}) \right| \leq |e^y - 1| + |x_1| \left| \text{sen}(yx_1 + \frac{a\pi}{2}) \right| \\ &\leq 2|y| + |x_1| (|y||x_1| + \frac{|a|\pi}{2}) \leq 2\rho + (1+r)[\rho(1+r) + \frac{r\pi}{2}] \end{aligned} \quad (3.48)$$

quindi, per la definizione di r , si ottiene

$$\sup_{Y_0 \times X_0} \left| 1 - T \frac{\partial F}{\partial y} \right| \leq 2\rho + (1 + \frac{\sqrt{\rho}}{2})[\rho(1 + \frac{\sqrt{\rho}}{2}) + \frac{\sqrt{\rho}\pi}{4}]$$

e quindi (3.35) è soddisfatta se prendiamo, per esempio, $\rho = (1/16)$ (e quindi $r = (1/8)$).

Esempio 3.22 Sia $F(y_1, y_2, x) := (\text{senh}(y_2 + x) + y_1^2 + y_2^2, \frac{1}{8} + y_1 + (xy_1)^2)$ e si cerchino soluzioni di $F(g(x), x) = 0$ con g funzione regolare di una variabile reale ed a valori in \mathbb{R}^2 . Si noti che non vi sono soluzioni ovvie di $F(y_0, x_0) = 0$ e quindi non è possibile applicare immediatamente il teorema delle funzioni implicite nella versione classica della Proposizione 3.10. Proviamo ad applicare il Teorema 3.19 cercando una 'buona' soluzione *approssimata* di $F(y_0, x_0) = 0$ (e fissando le norme $|\cdot|_\infty$). Scegliamo $y_0 = (-\frac{1}{8}, 0)$ e $x_0 = 0$, $T := \left(\frac{\partial F}{\partial y}(y_0, x_0) \right)^{-1}$ (che è sempre la scelta più naturale), $X_0 := \{|x| \leq r\}$ e cerchiamo di determinare r e ρ cosicché siano soddisfatte (3.34) e (3.35). Lo jacobiano $\frac{\partial F}{\partial y}$ è dato da

$$\frac{\partial F}{\partial y}(y, x) = \begin{pmatrix} 2y_1 & \cosh(y_2 + x) + 2y_2 \\ 1 + 2y_1x^2 & 0 \end{pmatrix} \implies T = \begin{pmatrix} 0 & 1 \\ 1 & \frac{1}{4} \end{pmatrix}. \quad (3.49)$$

Ora, per $|x| \leq r$ (che assumeremo comunque minore di 1)

$$|F(y_0, x)|_\infty = \left| \left(\frac{1}{64} + \text{senh } x, \frac{x^2}{64} \right) \right|_\infty \leq \frac{1}{64} + \text{senh } r, \quad (3.50)$$

¹⁵ $|1 - \cos y| = \left| \int_0^y \text{sen } t \, dt \right| \leq \int_0^{|y|} |\text{sen } t| \, dt \leq \int_0^{|y|} t \, dt = \frac{|y|^2}{2}$.

Per $|y| \leq 1$, $\left| \frac{e^y - 1}{y} \right| \leq \sum_{k \geq 1} \frac{|y|^{k-1}}{k!} \leq \sum_{k \geq 1} \frac{1}{k!} = e - 1 < 2$.

e poiché $\|T\| = \|T\|_{\infty, \infty} = 5/4$ la (3.34) è implicata da

$$\frac{5}{128} + \frac{5}{2} \sinh r \leq \rho, \quad (3.51)$$

(si noti che tale relazione ci dice che ρ non potrà esser scelto più piccolo di $5/128$). Per $|x| \leq r$, $|y_1 + \frac{1}{8}| \leq \rho$ e $|y_2| \leq \rho$ (e assumendo che $\rho + r < 1$) si ha che¹⁶

$$\begin{aligned} \|I - T \frac{\partial F}{\partial y}\| &= \left\| \begin{pmatrix} 2y_1 x^2 & 0 \\ 2y_1 + \frac{1}{4} + \frac{y_1 x^2}{2} & \cosh(y_2 + x) - 1 + 2y_1 \end{pmatrix} \right\| \\ &\leq \max \left\{ 2|y_1| r^2 + 2\rho + \frac{|y_1| r^2}{2}, (\rho + r)^2 + 2|y_1| \right\}. \end{aligned} \quad (3.53)$$

Osservando che $|y_1| \leq \frac{1}{8} + \rho$ e scegliendo $\rho = \frac{1}{10}$ e $r = \frac{1}{50}$, da (3.51) e da (3.53) segue che (3.34) ed (3.35) sono soddisfatte e che, dunque, valgono le tesi del Teorema 3.19.

Corollario 3.23 (1) *Sia k un intero positivo o $+\infty$ e sia f una funzione C^k in un intorno di un punto $y_0 \in \mathbb{R}^n$ ed a valori in \mathbb{R}^n . Se la matrice jacobiana $\frac{\partial f}{\partial y}(y_0)$ è invertibile, allora esiste (ed è unica) la funzione inversa g di f . Tale funzione g è C^k in un intorno di $x_0 = f(y_0)$ e vale*

$$\frac{\partial g}{\partial x}(x_0) = T, \quad \text{con} \quad T := \left(\frac{\partial f}{\partial y}(y_0) \right)^{-1}. \quad (3.54)$$

(2) *Più precisamente, sia ρ tale che*

$$\sup_{Y_0} \|I - T \frac{\partial f}{\partial y}(y)\| \leq \frac{1}{2} \quad (3.55)$$

con $Y_0 := \{y \in \mathbb{R}^n : |y - y_0| \leq \rho\}$. Allora, se poniamo $r := \rho(2\|T\|)^{-1}$ e $X_0 := \{x \in \mathbb{R}^n : |x - x_0| \leq r\}$, esiste una unica $g \in C^k(X_0, Y_0)$ tale che $g(x_0) = y_0$, $f(g(x)) = x$ in X_0 e $g(f(y)) = y$ in¹⁷ $g(X_0)$.

Dimostrazione Basta porre $F(y, x) = f(y) - x$. Per definizione di x_0 , $F(y_0, x_0) = 0$ e $F \in C^1(Y_0 \times X_0)$. Inoltre $\frac{\partial F}{\partial y} = \frac{\partial f}{\partial y}$ e $F(y_0, x) = f(y_0) - x = x_0 - x$, quindi la condizione (3.34) è automaticamente soddisfatta grazie alla definizione di r . Da queste osservazioni, dal Teorema 3.19 seguono tutte le tesi del corollario tranne quella sull'identità $g(f(y)) = y$. Sia $\bar{y} \in g(X_0)$, ossia sia $\bar{y} = g(\bar{x})$ per qualche $\bar{x} \in X_0$. Allora, essendo $f(g(x)) = x$ per $x \in X_0$, si ha che $f(\bar{y}) = \bar{x}$, da cui $g(f(\bar{y})) = g(\bar{x}) = \bar{y}$. ■

¹⁶ In tali disuguaglianze si è usato che, se $|t| \leq 1$ allora

$$0 \leq \cosh t - 1 \leq t^2, \quad (3.52)$$

infatti $\cosh t - 1 = t^2 \sum_{j=1}^{\infty} \frac{t^{2(j-1)}}{(2j)!} \leq t^2 \sum_{j=2}^{\infty} \frac{1}{j!} < t^2$.

¹⁷ $g(X_0)$ è l'insieme dei valori $g(x)$ per x che varia in X_0 .

Capitolo 4

Teoria classica dell'integrazione

1 Integrale di Riemann e misura di Peano–Jordan

In questo capitolo discutiamo la teoria classica dell'integrazione secondo Riemann e la collegata teoria della misura secondo Peano–Jordan¹.

1.1 Rettangoli, insieme elementari, funzioni a scalini

1.1 Rettangoli in \mathbb{R}^n

Si ricorda che un *intervallo* di \mathbb{R} è un sottoinsieme connesso di \mathbb{R} ossia un sottoinsieme $I \subseteq \mathbb{R}$ tale che se $x, y \in I$ allora $(1-t)x + ty \in I$ per ogni $0 \leq t \leq 1$; $\sup I$ e $\inf I$ denotano, rispettivamente, l'estremo destro e sinistro di I . Se un intervallo I è limitato e $a \leq b$ sono i suoi estremi, chiamiamo *lunghezza o misura di I* il numero non negativo $(b-a)$. Si osservi che se I, J sono intervalli di \mathbb{R} , allora $I \setminus J$ è dato dall'unione disgiunta di due intervalli (eventualmente vuoti²).

Un *pluri-rettangolo* (o 'rettangolo') $R \subseteq \mathbb{R}^n$ è il prodotto cartesiano

$$R = I_1 \times \cdots \times I_n := \bigotimes_{i=1}^n I_i$$

di n intervalli I_j ('spigoli o lati del pluri-rettangolo'); un *cubo* è un rettangolo i cui lati hanno lunghezza uguale.

Un rettangolo si dice *degenere* se uno dei suoi lati ha lunghezza nulla.

Esempi di rettangoli in \mathbb{R}^2 sono:

$$\begin{aligned} [0, 1]^2 &:= [0, 1] \times [0, 1] := \{x = (x_1, x_2) \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, 0 \leq x_2 \leq 1\}; \\ [-\pi, \pi) \times (-1, +\infty) &:= \{x \in \mathbb{R}^2 : -\pi \leq x_1 < \pi, x_2 > -1\}; \\ [0, 1] \times \{3\} &:= \{x \in \mathbb{R}^2 : 0 \leq x_1 \leq 1, x_2 = 3\}. \end{aligned}$$

L'ultimo esempio è degenere.

In questo capitolo chiameremo *rettangolo standard* un rettangolo chiuso, limitato e non degenere, ossia un insieme della forma

$$R := [a_1, b_1] \times \cdots \times [a_n, b_n], \quad \text{con} \quad -\infty < a_i < b_i < \infty. \quad (4.1)$$

Se $n \geq 2$, il termine *faccia* del rettangolo $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$, sta ad indicare uno dei $2n$ rettangoli $(n-1)$ -dimensionali che formano la frontiera di R : per ogni j tra 1 ed n vi sono

¹In questo capitolo, i termini 'integrale' e 'misura' stanno per 'integrale di Riemann' e 'misura di Peano–Jordan'.

²Esercizio 4.1.

due facce opposte date da

$$\begin{aligned} \{a_j\} \times \bigotimes_{i \neq j} [a_i, b_i] &:= \{x \in \mathbb{R}^n : x_j = a_j, a_i \leq x_i \leq b_i, \forall i \neq j\}, \text{ e da} \\ \{b_j\} \times \bigotimes_{i \neq j} [a_i, b_i] &:= \{x \in \mathbb{R}^n : x_j = b_j, a_i \leq x_i \leq b_i, \forall i \neq j\}. \end{aligned} \quad (4.2)$$

1.2 Insiemi elementari e partizioni

Un *insieme elementare* E è un insieme unione di una famiglia finita di rettangoli $\{R_i\}$ limitati a due a due disgiunti, ossia $1 \leq i \leq N \in \mathbb{N}$, $R_i \cap R_j = \emptyset$ se $i \neq j$ e $E = \cup R_i$; l'insieme $\mathcal{R} = \{R_i\}$ di tali rettangoli è una *partizione* dell'insieme elementare E .

Denotiamo con \mathcal{E}_n la famiglia degli insiemi elementari di \mathbb{R}^n .

Il *diametro* di una partizione \mathcal{R} di $E \in \mathcal{E}_n$ è il numero non negativo³

$$\text{diam } \mathcal{R} := \sup_{R \in \mathcal{R}} \text{diam } R. \quad (4.3)$$

1.3 Raffinamenti di partizioni

Se $P = (P_1, \dots, P_n)$ e $P' = (P'_1, \dots, P'_n)$ sono partizioni di E , diremo che $P \subseteq P'$ se, per ogni i , $P_i \subseteq P'_i$; in tal caso, diremo che P' è un *raffinamento* di P .

Chiaramente, se P' è un raffinamento di P , ogni rettangolo R' di P' è interamente contenuto in un rettangolo R di P e ogni rettangolo R di P è dato dall'unione di tutti i rettangoli di P' contenuti in R .

Date due partizioni di E , $P' = (P'_1, \dots, P'_n)$ e $P'' = (P''_1, \dots, P''_n)$, definiamo la *partizione unione* di P' e P'' , $P := P' \cup P''$, la partizione $P := (P_1, \dots, P_n)$ con P_i formato dall'unione di tutti i punti di P'_i e P''_i . In particolare,

$$P' \subseteq P' \cup P'', \quad P'' \subseteq P' \cup P'' \quad (4.4)$$

qualunque siano le partizioni P' e P'' .

1.4 Misura di insiemi elementari

Se un rettangolo E è non-degenere ed uno dei suoi lati è illimitato, poniamo $\text{mis}_n E = +\infty$; se un rettangolo E è degenere, poniamo $\text{mis}_n E = 0$. Se $E = I_1 \times \dots \times I_n \subseteq \mathbb{R}^n$ è un rettangolo limitato, definiamo la *misura* di E come il prodotto delle lunghezze degli intervalli unidimensionali I_i , ossia, se gli estremi di I_i sono $a_i \leq b_i$, definiamo la misura di E come⁴

$$\text{mis}_n E := \prod_{i=1}^n (b_i - a_i) = (b_1 - a_1) \cdots (b_n - a_n). \quad (4.5)$$

La misura di $E = I_1 \times \dots \times I_n$ non dipende dal fatto che gli intervalli I_i siano aperti o chiusi a sinistra o destra. Si noti anche che l'insieme vuoto è un particolare rettangolo (aperto e con tutti i lati degeneri $a_i = b_i$ per ogni i) dunque, per definizione, la *misura dell'insieme vuoto* è uguale a 0.

1.5 Integrale di Riemann

Sia $E \subseteq \mathbb{R}^n$ un rettangolo standard e sia $f : E \rightarrow \mathbb{R}$ una funzione limitata (cioè $\sup_E |f| < \infty$). Data una partizione P di E si definiscono *somma inferiore* e *somma superiore* di f su E rispetto a P i numeri

$$\begin{aligned} \underline{S}_E(f, P) &:= \sum_{R \in \mathcal{R}(P)} \left(\inf_{x \in R} f(x) \right) \text{mis } R, \\ \overline{S}_E(f, P) &:= \sum_{R \in \mathcal{R}(P)} \left(\sup_{x \in R} f(x) \right) \text{mis } R. \end{aligned} \quad (4.6)$$

³Si ricorda che dato un qualunque insieme non vuoto $A \subseteq \mathbb{R}^n$ si definisce il 'diametro di A (rispetto alla norma euclidea)' la quantità $\text{diam } A := \sup_{x, y \in A} |x - y|$.

⁴Normalmente, quando non vi sia ambiguità, ometteremo l'indice della dimensione n dal simbolo mis_n , e denoteremo semplicemente $\text{mis } E$ la misura n -dimensionale del rettangolo n -dimensionale E .

Dalle definizioni date segue che se $P \subseteq P'$

$$-\infty < \operatorname{mis} E \inf_E f \leq \underline{S}_E(f, P) \leq \underline{S}_E(f, P') \leq \overline{S}_E(f, P') \leq \overline{S}_E(f, P) \leq \operatorname{mis} E \sup f < \infty. \quad (4.7)$$

Si definiscono, rispettivamente, *integrale inferiore* e *integrale superiore di Riemann* di f su E le quantità

$$\begin{aligned} \underline{\sigma}_E(f) &:= \sup_{\{P\}} \underline{S}_E(f, P), \\ \overline{\sigma}_E(f) &:= \inf_{\{P\}} \overline{S}_E(f, P) \end{aligned} \quad (4.8)$$

dove, come sopra, l'estremo superiore è preso su tutte le partizioni di E . Chiaramente⁵

$$-\infty < \sup_{\{P\}} \underline{S}_E(f, P) \leq \inf_{\{P\}} \overline{S}_E(f, P) < \infty. \quad (4.9)$$

Si dice che la funzione f è *integrabile (secondo Riemann)* su E se in (4.9) vale il segno di uguaglianza e chiameremo, in tal caso, *l'integrale di Riemann* di f su E tale numero:

$$\int_E f := \underline{\sigma}_E(f) = \overline{\sigma}_E(f). \quad (4.10)$$

Si usano anche le seguenti notazioni equivalenti

$$\int_E f := \int_E f(x) dx := \int_E f(x) dx_1 \cdots dx_n := \int_E f(y) dy_1 \cdots dy_n.$$

1.6 Somme parziali di Riemann

Sia $B \subseteq E$ un insieme misurabile e $f : B \rightarrow \mathbb{R}$ una funzione limitata. Definiamo la funzione f_B come la funzione che coincide con f su B e vale 0 fuori di B . Diremo che f è *integrabile su B* se la funzione f_B è integrabile su E e *definiamo* l'integrale di f su B come l'integrale di f_B su E :

$$\int_B f dx := \int_E f_B dx.$$

Data una partizione P di E chiameremo una *scelta di punti* di B associati alla partizione P un insieme della forma $Q := \{x_R : R \in \mathcal{R}(P)\}$ dove, per ogni $R \in \mathcal{R}(P)$, x_R è un punto di $R \cap B$. Data una partizione P ed una scelta di punti Q di B , si definisce la *somma parziale di Riemann* di f su B relativa a P e Q il numero

$$S_B(f, P, Q) := \sum_{R \in \mathcal{R}'_B(P)} f(x_R) \operatorname{mis} R. \quad (4.11)$$

1.7 Oscillazione di funzioni

La teoria dell'integrazione di Riemann è intimamente connessa alla nozione di continuità⁶.

Sia $f : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ e $D \subseteq \mathbb{R}^n$ un insieme aperto o la chiusura di un aperto tale che $D \cap A \neq \emptyset$; si definisce l'*oscillazione di f sull'insieme D* la quantità

$$\operatorname{osc}(f, D) := \sup_{D \cap A} f - \inf_{D \cap A} f. \quad (4.12)$$

Se $x \in A$ si definisce l'*oscillazione di f in x* la quantità

$$\operatorname{osc}(f, x) := \inf_{\delta > 0} \operatorname{osc}(f, B_\delta(x)) := \inf_{\delta > 0} \left(\sup_{y \in A: |y-x| < \delta} f(y) - \inf_{y \in A: |y-x| < \delta} f(y) \right). \quad (4.13)$$

⁵Si osservi che se P e P' sono due partizioni di E allora $\underline{S}_E(f, P) \leq \underline{S}_E(f, P \cup P') \leq \overline{S}_E(f, P \cup P') \leq \overline{S}_E(f, P')$, ossia 'le classi dei numeri delle somme inferiori e superiori sono separate'.

⁶Vedremo in seguito, infatti, che le funzioni integrabili secondo Riemann coincidono con le funzioni che non hanno 'troppi punti di discontinuità'

Dunque, la continuità di f in x è equivalente a: per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che $\text{osc}(f, B_\delta(x)) < \varepsilon$; o anche: f è continua in x se e solo se $\text{osc}(f, x) = 0$. D'altra parte, l'integrabilità di f su E è equivalente a: per ogni $\varepsilon > 0$ esiste una partizione P di E tale che

$$\sum_{R \in \mathcal{R}(P)} \text{osc}(f, R) \text{mis } R \leq \varepsilon. \quad (4.14)$$

1.8 Funzioni a scalini

Una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$ si dice *funzione a scalini* se⁷

$$f = \sum_{j=1}^N c_j \chi_{R_j}$$

con $c_j \in \mathbb{R}$ e $\{R_j \mid 1 \leq j \leq N\}$ è una famiglia (finita) di rettangoli limitati a due a due disgiunti. Dato un rettangolo $E \subseteq \mathbb{R}^n$, $S(E)$ denota la famiglia di funzioni a scalini f con⁸ $\text{supp}(f) \subseteq E$.

1.9 Funzioni continue a tratti

Una funzione limitata $f : B \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ si dice *continua a tratti* se $B = \bigcup_{i=1}^N B_i$ con $B_i \subseteq \mathbb{R}^n$ misurabili e tali che $\overset{\circ}{B}_i \cap \overset{\circ}{B}_j = \emptyset$ per $i \neq j$ e se f è continua su $\overset{\circ}{B}_i$ per ogni i .

1.2 Integrale di Riemann

Discutiamo brevemente alcune proprietà elementari, ma fondamentali, dell'integrale di Riemann e della misura di Peano–Jordan.

In questo paragrafo E denota un rettangolo standard in \mathbb{R}^n , f, g funzioni limitate da E in \mathbb{R} e B un sottoinsieme di E .

1.10 \mathcal{E}_n è un'algebra: se $E, F \in \mathcal{E}_n$, allora $E \cap F, E \setminus F, E \cup F \in \mathcal{E}_n$.

Dimostrazione Sia $E = \bigcup_i R_i$ e $F = \bigcup_j S_j$ con $1 \leq i \leq N, 1 \leq j \leq M$ e R_i rettangoli limitati a due a due disgiunti e analogamente per gli S_j .

$$E \cap F = \bigcup_{i,j} (R_i \cap S_j),$$

da cui segue immediatamente che $E \cap F \in \mathcal{E}_n$.

Chiaramente, se si ha anche $R_i \cap S_j = \emptyset$ per ogni $i \neq j$, allora⁹

$$E \dot{\cup} F \in \mathcal{E}_n, \quad (R_i \cap S_j = \emptyset, \forall i \neq j). \quad (4.15)$$

Osserviamo che se $R = I_1 \times \cdots \times I_n$ e $S = J_1 \times \cdots \times J_n$ sono rettangoli limitati in \mathbb{R}^n , allora $R \setminus S \in \mathcal{E}_n$: infatti, (si ricordi il punto 1.1), per ogni j , $I_j \setminus J_j = I'_j \dot{\cup} I''_j$ con I'_j e I''_j intervalli limitati e disgiunti; dunque:

$$\begin{aligned} R \setminus S &= \bigcup_j R \cap \{x \mid x_j \notin J_j\} = \bigcup_j R \cap \{x \mid x_j \in I_j \setminus J_j\} \\ &= \bigcup_j R \cap \{x \mid x_j \in I'_j\} \dot{\cup} \bigcup_j R \cap \{x \mid x_j \in I''_j\}, \end{aligned}$$

e, poiché insiemi della forma $R \cap \{x \mid x_j \in I\}$ con I intervallo sono rettangoli limitati, da (4.15) segue che $R \setminus S \in \mathcal{E}_n$. Quindi anche $R \cup S = (R \setminus S) \dot{\cup} (R \cap S) \dot{\cup} (S \setminus R) \in \mathcal{E}_n$ e la tesi segue facilmente da relazioni insiemistiche elementari.

⁷ Si ricorda che, se B è un qualunque insieme, χ_B denota la funzione caratteristica (o indicatrice) di B ossia la funzione che vale 1 su ogni punto di B e zero altrimenti.

⁸ Il supporto di una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$, $\text{supp}(f)$ è la chiusura dell'insieme $\{x \in \mathbb{R}^n \mid f(x) \neq 0\}$.

⁹ $A = A \dot{\cup} C$ significa $A = B \cup C$ e $B \cap C = \emptyset$.

1.11 Una funzione a scalini s con supporto in E si può rappresentare come

$$s(x) = \sum_{R \in \mathcal{P}(P)} c_R \chi_R(x) \quad (4.16)$$

dove $\mathcal{P}(P)$ è una partizione propria di E , c_R sono numeri reali.

1.12 Data una qualunque partizione P di E si ha

$$\sum_{R \in \mathcal{R}(P)} \text{mis } R = \text{mis } E. \quad (4.17)$$

Dimostrazione Consideriamo per semplicità il caso $n = 2$. Se $P = (P_1, P_2)$ con P_i come in (??), si ha

$$\begin{aligned} \sum_{R \in \mathcal{R}(P)} \text{mis } R &= \sum_{j=(j_1, j_2)} \text{mis } R_j \\ &= \sum_{j_1, j_2} (\xi_{j_1}^{(1)} - \xi_{j_1-1}^{(1)}) (\xi_{j_2}^{(2)} - \xi_{j_2-1}^{(2)}) \\ &= \left(\sum_{j_1} (\xi_{j_1}^{(1)} - \xi_{j_1-1}^{(1)}) \right) \left(\sum_{j_2} (\xi_{j_2}^{(2)} - \xi_{j_2-1}^{(2)}) \right) \\ &= (\xi_{N_1}^{(1)} - \xi_0^{(1)}) (\xi_{N_2}^{(2)} - \xi_0^{(2)}) \\ &= (b_1 - a_1)(b_2 - a_2) \\ &=: \text{mis } R, \end{aligned}$$

essendo le somme unidimensionali somme ‘telescopiche’. ■

1.13 f è integrabile su E se e solo se per ogni $\varepsilon > 0$ esiste una partizione P di E tale che

$$\begin{aligned} \overline{S}_E(f, P) - \underline{S}_E(f, P) &= \sum_{R \in \mathcal{R}(P)} \left(\sup_R f - \inf_R f \right) \text{mis } R \\ &= \sum_{R \in \mathcal{R}(P)} \left(\sup_{x, y \in R} |f(x) - f(y)| \right) \text{mis } R \\ &\leq \varepsilon. \end{aligned} \quad (4.18)$$

Dimostrazione Innanzitutto si osservi che, in generale, si ha che

$$\sup_B f - \inf_B f = \sup_{x, y \in B} |f(x) - f(y)| \quad (4.19)$$

per qualunque B e $f : B \rightarrow \mathbb{R}$. Infatti, per ogni $x', y' \in B$ si ha che

$$f(x') - f(y') \leq |f(x') - f(y')| \leq \sup_{x, y \in B} |f(x) - f(y)|$$

e prendendo l'estremo superiore su x' e su y' si ha che

$$\sup_B f - \inf_B f \leq \sup_{x, y \in B} |f(x) - f(y)|.$$

Viceversa, per ogni $x, y \in B$,

$$\inf_B f - \sup_B f \leq f(x) - f(y) \leq \sup_B f - \inf_B f$$

che implica che $|f(x) - f(y)| \leq \sup_B f - \inf_B f$ e prendendo l'estremo superiore su $x, y \in B$ si ottiene anche che

$$\sup_{x, y \in B} |f(x) - f(y)| \leq \sup_B f - \inf_B f,$$

e quindi, la validità di (4.19), che a sua volta implica la seconda uguaglianza in (4.18).

La (4.18) implica immediatamente l'integrabilità di f su E . Viceversa, se f è integrabile, dato $\varepsilon > 0$, esistono due partizioni P' e P'' di E tali che

$$0 \leq \overline{S}(f, P') - \underline{S}(f, P'') \leq \varepsilon,$$

e se poniamo $P = P' \cup P''$, da (4.7) segue che

$$\overline{S}(f, P) - \underline{S}(f, P) \leq \overline{S}(f, P') - \underline{S}(f, P'') \leq \varepsilon$$

ossia la (4.18). ■

1.14 B è misurabile se e solo se χ_B è integrabile su E .

Dimostrazione Sia P una partizione di E e si osservi che, per ogni $R \in \mathcal{R}(P)$,

$$\inf_R \chi_B = \begin{cases} 1, & \text{se } R \subseteq B \\ 0, & \text{altrimenti} \end{cases}, \quad \sup_R \chi_B = \begin{cases} 1, & \text{se } R \cap B \neq \emptyset \\ 0, & \text{altrimenti} \end{cases}$$

Quindi,

$$\begin{aligned} \underline{S}_E(\chi_B, P) &= \sum_{R \in \mathcal{R}(P)} \inf_R \chi_B \operatorname{mis} R = \sum_{R \in \mathcal{R}'_B(P)} \operatorname{mis} R \\ \overline{S}_E(\chi_B, P) &= \sum_{R \in \mathcal{R}(P)} \sup_R \chi_B \operatorname{mis} R = \sum_{R \in \mathcal{R}_B(P)} \operatorname{mis} R \end{aligned}$$

da cui segue immediatamente l'asserto. ■

1.15 B è misurabile se e solo se per ogni $\varepsilon > 0$ esiste una partizione P tale che

$$\sum_{R \in \mathcal{R}_B \setminus \mathcal{R}'_B} \operatorname{mis} R \leq \varepsilon, \quad (4.20)$$

dove $\mathcal{R}_B := \mathcal{R}_B(P)$ e $\mathcal{R}'_B := \mathcal{R}'_B(P)$.

Dimostrazione Segue subito da 1.13 e 1.14. ■

1.16 Se f integrabile su B , per ogni $\varepsilon > 0$, esiste una partizione P tale che per ogni scelta di punti Q su B relativa a P si ha

$$\left| \int_B f(x) dx - S_B(f, P, Q) \right| \leq \varepsilon. \quad (4.21)$$

In particolare¹⁰, esistono partizioni P_k tali che per ogni scelta di punti Q_k su B relativa a P_k si ha

$$\lim_{k \rightarrow \infty} S_B(f, P_k, Q_k) = \int_B f(x) dx. \quad (4.22)$$

Dimostrazione Segue subito da 1.13, 1.15 e dalle definizioni 1.5 e 1.6. ■

1.17 Siano f e g due funzioni integrabili su E .

(i) Per ogni $a, b \in \mathbb{R}$, la funzione $af + bg$ è integrabile su E e si ha

$$\int_E (af + bg) = a \int_E f + b \int_E g. \quad (4.23)$$

(ii) fg è integrabile su E .

¹⁰Si scelga, per esempio, $\varepsilon = 1/k$ in (4.21).

Dimostrazione (i): Si osservi che $\sup_{x,y \in R} |af(x) - af(y)| = |a| \sup_{x,y \in R} |f(x) - f(y)|$ e quindi, da 1.13, segue l'integrabilità di af ; analogamente l'integrabilità di $f + g$ segue osservando che

$$\sup_{x,y \in R} |f(x) + g(x) - f(y) - g(y)| \leq \sup_{x,y \in R} |f(x) - f(y)| + \sup_{x,y \in R} |g(x) - g(y)|$$

ed invocando nuovamente 1.13. La (4.23) segue osservando che le somme parziali di Riemann $S_E(\cdot, P, Q)$ (cfr. (4.11)) sono lineari¹¹ ed usando la 1.16.

(ii): Poiché f e g sono integrabili sono limitate, quindi esiste $M > 0$ tale che $\sup_E |f| \leq M$ e $\sup_E |g| \leq M$. Per ogni x, y si ha che

$$\begin{aligned} |f(x)g(x) - f(y)g(y)| &\leq |f(x)| |g(x) - g(y)| + |g(y)| |f(x) - f(y)| \\ &\leq M(|f(x) - f(y)| + |g(x) - g(y)|) \end{aligned}$$

e l'integrabilità di fg segue, ancora, da 1.13. ■

1.18 (i) Se $f \geq 0$ è integrabile, allora $\int_E f \geq 0$.

(ii) Se $f \geq g$ sono integrabili allora $\int_E f \geq \int_E g$.

(iii) Una funzione f è integrabile su E se e solo se¹² f_+ e f_- sono integrabili su E .

(iv) Se f è integrabile su E , allora $|f|$ è integrabile su E e

$$\left| \int_E f \right| \leq \int_E |f|. \quad (4.26)$$

Dimostrazione (i) segue da (4.22) e (ii) da $f - g \geq 0$, da (4.23) e da (i).

(iii): Si osservi che¹³ $|f_{\pm}(x) - f_{\pm}(y)| \leq |f(x) - f(y)|$; dunque, se f è integrabile, per 1.13, si ha che anche f_+ e f_- sono integrabili. Viceversa se f_{\pm} sono integrabili, l'integrabilità di f segue dalla relazione $f = f_+ - f_-$ e da 1.17-(i).

(iv): $|f| = f_+ + f_-$ e da 1.17-(i) segue l'integrabilità di $|f|$ e (4.26) segue da

$$\left| \int_E f \right| = \left| \int_E f_+ - \int_E f_- \right| \leq \int_E f_+ + \int_E f_- = \int_E |f|. \quad \blacksquare$$

1.19 L'integrale di Riemann è 'invariante per traslazioni':

Sia B è un insieme misurabile, $x_0 \in \mathbb{R}^n$ e $x \rightarrow \tau(x) := x_0 + x$ una traslazione. Allora¹⁴ $\tau(B)$ è misurabile e se f è integrabile su $\tau(B)$ allora $f \circ \tau$ è integrabile su B e

$$\int_{\tau(B)} f(y) dy = \int_B f \circ \tau(x) dx. \quad (4.27)$$

In particolare (prendendo $f := 1$) si ha che $\text{mis}(\tau(B)) = \text{mis}(B)$.

Dimostrazione Basta osservare che il traslato $\tau(R)$ di un rettangolo è un rettangolo e quindi che se P è una partizione di $E \supseteq B$, $x_0 + P$ è una partizione di $\tau(E) \supseteq \tau(B)$. La (4.27) segue, per esempio, da (4.22). ■

¹¹Cioè $S_E(af + bg, P, Q) = aS_E(f, P, Q) + bS_E(g, P, Q)$ per ogni a, b, f, g, P e Q .

¹²Data una funzione $f: A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ si chiamano, rispettivamente, parte positiva e parte negativa di f , le funzioni

$$f_+(x) = \max\{f(x), 0\}, \quad f_-(x) = -\min\{f(x), 0\} = \max\{-f(x), 0\}. \quad (4.24)$$

Si noti che f_+ e f_- sono ambedue funzioni non negative e che

$$f = f_+ - f_-, \quad |f| = f_+ + f_-. \quad (4.25)$$

¹³Se $f(x)f(y) \geq 0$, vale l'uguaglianza; se, per esempio $f(x) > 0 > f(y)$ il membro di sinistra è uguale a $f(x)$ e il membro di destra a $f(x) - f(y) = f(x) + |f(y)|$; il caso f_- segue osservando che $f_- = (-f)_+$.

¹⁴ $\tau(B) = x_0 + B = \{y : y = x_0 + x \text{ con } x \in B\}$.

1.20 Se A e B sono misurabili allora:

- (i) $A \cap B$ e $A \cup B$ sono misurabili;
- (ii) $\text{mis}(A \cup B) \leq \text{mis} A + \text{mis} B$;
se $\text{mis}(A \cap B) = 0$ allora $\text{mis}(A \cup B) = \text{mis} A + \text{mis} B$;
- (iii) se $A \subseteq B$, $\text{mis}(B \setminus A) = \text{mis} B - \text{mis} A$;
- (iv) se f è integrabile su A e su B lo è anche su $A \cup B$ e se $\text{mis}(A \cap B) = 0$ allora

$$\int_{A \cup B} f = \int_A f + \int_B f. \quad (4.28)$$

Dimostrazione Segue subito da 1.14 e 1.17 osservando che

$$\chi_{A \cap B} = \chi_A \chi_B, \quad \chi_{A \cup B} = \chi_A + \chi_B - \chi_{A \cap B}. \quad \blacksquare \quad (4.29)$$

- 1.21** (i) L 'intersezione finita di rettangoli è un rettangolo.
- (ii) L 'unione finita di rettangoli chiusi è un insieme elementare¹⁵.
- (iii) Gli insiemi elementari sono misurabili.

Dimostrazione (i): Segue dal fatto che l'intersezione di due intervalli è un intervallo.
(ii): Diamo la dimostrazione nel caso $n = 2$ (il caso generale è del tutto analogo). Siano $R_k = [a_k, b_k] \times [c_k, d_k]$ rettangoli in $E = [a, b] \times [c, d]$ con $1 \leq k \leq N$ e sia P_1 la partizione di $[a, b]$ che contiene tutti gli a_k e b_k e P_2 la partizione di $[c, d]$ che contiene tutti i c_k e d_k e $P = (P_1, P_2)$. Se $\mathcal{R}_* = \{R \in \mathcal{R}(P) : R \subseteq \cup R_k\}$ allora $\cup R_k = \cup_{R \in \mathcal{R}_*} R$, il che mostra che $\cup R_k$ è un insieme elementare.
(iii): Segue dal punto (i) di 1.20. \blacksquare

1.22 Dato un rettangolo limitato chiuso R ed $\varepsilon > 0$ esistono N cubi chiusi $K_i \subseteq R$ tali che:

- (i) $\overset{\circ}{K}_i \cap \overset{\circ}{K}_j = \emptyset$, per ogni $i \neq j$;

- (ii) $\text{mis} \left(R \setminus \bigcup_{i=1}^N K_i \right) \leq \varepsilon$.

Dimostrazione Sia $R = [a_1, b_1] \times \dots \times [a_n, b_n]$, sia $L = \max(b_i - a_i)$ e $\delta = \varepsilon / (2^n L^{n-1})$, $P_i = \{a_i, a_i + \delta, \dots, a_i + m_i \delta, b_i\}$ con $a_i + m_i \delta \leq b_i < a_i + (m_i + 1) \delta$ (cioè $m_i = (b_i - a_i) / \delta$). Allora i K_i della tesi saranno tutti i rettangoli di $\mathcal{R}(P)$ che hanno tutti i lati di lunghezza δ (ossia lati della forma $[a_i + jd, a_i + (j+1)\delta]$ con $j \leq m_i - 1$): infatti $\text{mis} \left(R \setminus \bigcup_{i=1}^N K_i \right) \leq n \delta L^{n-1} = \varepsilon$.

\blacksquare

1.23 (i) B è misurabile se e solo se, per ogni $\varepsilon > 0$, esistono due insiemi misurabili A e C tali che $A \subseteq B \subseteq C$ e $\text{mis} C - \text{mis} A \leq \varepsilon$.

(ii) B è misurabile se e solo se, per ogni $\varepsilon > 0$, esistono due insiemi elementari A e C tali che $A \subseteq B \subseteq C$ e $\text{mis} C - \text{mis} A \leq \varepsilon$; tali insiemi elementari possono essere scelti come unione di cubi K_i tali che $\overset{\circ}{K}_i \cap \overset{\circ}{K}_j = \emptyset$.

(iii) Se $\text{mis est } B = 0$ allora B è misurabile e $\text{mis} B = 0$.

Dimostrazione (i): Se B è misurabile, dato ε sia P come in (4.20) e si prenda $A := \cup_{R \in \mathcal{R}'_B(P)} R$ e $C := \cup_{R \in \mathcal{R}_B(P)} R$.

Viceversa, si scelga una partizione P tale che $\sum_{R \in \mathcal{R}'_A} \text{mis} R \geq \text{mis} A - \varepsilon$ e $\sum_{R \in \mathcal{R}_C} \text{mis} R \leq \text{mis} C + \varepsilon$ e si usi 1.15.

(ii): Il 'se' deriva dal punto (i) (poiché gli insiemi elementari sono misurabili). Il 'solo se' deriva dalla dimostrazione di (i) (poiché gli insiemi scelti nella dimostrazione nel 'solo se' di (i) sono elementari). L'ultima affermazione segue facilmente da 1.22.

(iii) Segue da (i) prendendo $A = \emptyset$ e dalla definizione di misura esterna. \blacksquare

¹⁵Cfr. 1.2.

1.24 (i) Se B è misurabile lo sono anche \overline{B} , $\overset{\circ}{B}$ e ∂B . Inoltre

$$\text{mis } \overline{B} = \text{mis } \overset{\circ}{B} = \text{mis } B. \tag{4.30}$$

(ii) B è misurabile se e solo se B è limitato, ∂B è misurabile e $\text{mis } \partial B = 0$.

Dimostrazione (i): La misurabilità di \overline{B} e $\overset{\circ}{B}$ e la (4.30) seguono da 1.23–(i) prendendo $A := \bigcup_{R \in \mathcal{R}'_B} \overset{\circ}{R}$ e $C := \bigcup_{R \in \mathcal{R}_B} R$.

(ii): Assumiamo che B sia limitato, ∂B misurabile e $\text{mis } \partial B = 0$. Sia E un rettangolo standard che contiene \overline{B} (e quindi ∂B). Poiché $\text{mis } \partial B = 0$, per ogni $\varepsilon > 0$ esiste una partizione P di E tale che $\partial B \subseteq \bigcup_{R \in \mathcal{R}_{\partial B}} R$ e $\sum_{R \in \mathcal{R}_{\partial B}} \text{mis } R \leq \varepsilon$. Se $R \in \mathcal{R}$ non interseca ∂B allora o $R \subseteq \overset{\circ}{B}$ oppure $R \subseteq E \setminus \overline{B}$. Quindi, $\mathcal{R}_B = \mathcal{R}_{\partial B} \cup \mathcal{R}'_B$, che implica (4.20).

Il viceversa segue dal punto (i). ■

1.25 Sia $Q \subseteq \mathbb{R}^n$ un insieme che ha misura nulla¹⁶ e A un sottoinsieme limitato in \mathbb{R}^m . Allora $Q \times A \subseteq \mathbb{R}^{n+m}$ è misurabile e $\text{mis}_{n+m} Q \times A = 0$.

Dimostrazione Sia $\varepsilon > 0$ e sia $D \subseteq \mathbb{R}^n$ un insieme elementare tale che $D \supseteq Q$ e $\text{mis}_n D \leq \varepsilon$. Poiché A è limitato, esiste $L > 0$ tale che $A \subseteq [-L, L]^m$. Allora $D \times [-L, L]^m$ è un insieme elementare di misura non superiore a $L^m \varepsilon$. Dall'arbitrarietà di ε segue la tesi. ■

1.26 Sia $F : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ una funzione uniformemente Lipschitziana con costante di Lipschitz $L > 0$ rispetto alla norma del sup¹⁷, allora

- (i) per ogni $B \subseteq E$ misurabile $\text{mis est}_m F(B) \leq L^m \text{mis}_n B$.
- (ii) se Q ha misura nulla allora $\text{mis}_m F(Q) = 0$.
- (iii) Se $m > n$, allora $\text{mis}_m F(B) = 0$.

Dimostrazione Sia $\varepsilon > 0$ e sia $D \subseteq E$ un insieme elementare unione di cubi K_i tali che $\overset{\circ}{K}_i \cap \overset{\circ}{K}_j = \emptyset$ se $i \neq j$ e tale che $B \subseteq D$ e $\text{mis}_n D - \text{mis}_n B \leq \varepsilon/L^m$ (cfr. 1.23–(ii)). Sia $x^{(i)}$ e r_i , rispettivamente il centro ed il lato di K_i : $K_i = \{x \in \mathbb{R}^n : |x - x^{(i)}| \leq r_i/2\}$. Per ipotesi, $|F(x) - F(x^{(i)})| \leq L|x - x^{(i)}|$ il che implica che $F(K_i)$ è un sottoinsieme del cubo m -dimensionale K'_i di centro $F(x^{(i)})$ e lato Lr_i e tale cubo ha misura $(Lr_i)^m = L^m \text{mis}_n K_i$. Quindi, poiché l'immagine dell'unione coincide con l'unione delle immagini, segue che

$$F(B) \subseteq F(D) = \bigcup_i F(K_i) \subseteq \bigcup_i K'_i$$

e quindi

$$\text{mis est}_m F(B) \leq \sum_i \text{mis}_m K'_i = L^m \sum_i \text{mis}_n K_i = L^m \text{mis}_n D \leq L^m \text{mis}_n B + \varepsilon.$$

Dall'arbitrarietà di ε segue che $\text{mis est}_m F(B) \leq L^m \text{mis}_n B$.

(ii) segue immediatamente da (i) e da 1.23–(iii).

(iii) Sia $\widehat{B} := B \times [0, 1]^{m-n} \subseteq \mathbb{R}^m$ e sia \widehat{F} l'estensione di F a \widehat{B} ponendo, per ogni $(x, y) \in Q \times [0, 1]^{m-n}$, $\widehat{F}(x, y) = F(x)$. Chiaramente \widehat{F} è Lipschitziana su \widehat{B} e $\widehat{F}(\widehat{B}) = F(B)$. Poiché $\text{mis}_m \widehat{B} = 0$ (per 1.25) da (ii) segue che $\text{mis}_m F(B) = 0$. ■

I punti (ii) e (iii) non valgono, in generale, se assumiamo F semplicemente continua: nell'Esercizio 1.30 di fine sezione viene dato un esempio (la cosiddetta 'curva di Peano') di una funzione continua $F : I := [0, 1] \times \{0\} \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}^2$ la cui immagine sia tutto il cubo unitario $[0, 1]^2$ e quindi trasforma il segmento I che ha misura (bidimensionale) nulla, nel cubo (quadrato) unitario che ha misura (bidimensionale) uguale ad uno.

¹⁶Ossia Q è misurabile secondo Peano–Jordan e $\text{mis}_n Q = 0$.

¹⁷Ossia, esiste $L > 0$ tale che $|F(x) - F(y)| \leq L|y - x|$ per ogni $x, y \in E$ e $|\cdot| = |\cdot|_\infty$.

1.27 (i) Se $s = \sum_{R \in \mathcal{P}(P)} c_R \chi_R$ è una funzione a scalini (cfr. ??), allora s è integrabile e

$$\int s = \sum_{R \in \mathcal{P}(P)} c_R \text{mis } R. \quad (4.31)$$

(ii) f è integrabile su E se e solo se, per ogni $\varepsilon > 0$, esistono due funzioni a scalini su E , $s_1(x)$ e $s_2(x)$, tali che

$$s_1(x) \leq f(x) \leq s_2(x), \quad \forall x \in E, \quad \text{e} \quad \int_E (s_2 - s_1) \leq \varepsilon; \quad (4.32)$$

tali funzioni possono essere espresse intermini dello stesso partizione propria di E .

(iii) f è integrabile su E se e solo se esistono due funzioni integrabili su E , g_1 e g_2 tali che

$$g_1(x) \leq f(x) \leq g_2(x), \quad \forall x \in E, \quad \text{e} \quad \int_E (g_2 - g_1) \leq \varepsilon; \quad (4.33)$$

Dimostrazione (i): Un rettangolo limitato è misurabile e la sua misura è l'integrale della sua funzione caratteristica; dunque da 1.18 e 1.17 segue la tesi.

(ii): Sia f integrabile ed $\varepsilon > 0$. Sia P una partizione come in (4.18) e sia $\mathcal{P}(P)$ una qualunque partizione propria associato a P . Chiaramente, per ogni $R \in \mathcal{P}$, si ha $\sup_R f - \inf_R f \leq \sup_{\bar{R}} f - \inf_{\bar{R}} f$ e quindi se $\bar{c}_R = \sup_R f$, $\underline{c}_R = \inf_R f$, $s_2 = \sum_{R \in \mathcal{P}} \bar{c}_R \chi_R$ e $s_1 = \sum_{R \in \mathcal{P}} \underline{c}_R \chi_R$,

segue la (4.32).

Assumiamo, ora, che valga la (4.32). Prendendo la partizione unione P associata alle due rappresentazioni si ottiene subito che s_1 e s_2 possono essere espresse in termini dello stesso partizione propria di E associato ad una stessa partizione P come in (?). Sia $0 < \delta < \min_{i,j} (\xi_{j+1}^{(i)} - \xi_j^{(i)})/2$ e consideriamo la partizione $P^\delta = (P_1^\delta, \dots, P_n^\delta)$ con $P_i^\delta = \{\xi_0^{(i)}, \xi_0^{(i)} + \delta, \xi_1^{(i)} - \delta, \xi_1^{(i)}, \xi_1^{(i)} + \delta, \dots\}$. I rettangoli di $\mathcal{R}(P)^\delta$ si suddividono in due famiglie: la prima, \mathcal{R}_1 , formata da rettangoli R contenuti all'interno di un rettangolo di $\mathcal{R}(P)$ ed una seconda famiglia, \mathcal{R}_2 , formata da rettangoli in cui almeno un lato ha misura δ . Si osservi che se $R \in \mathcal{R}_1$, $\sup_R f - \inf_R f \leq \sup_R s_2 - \inf_R s_1$ e che, quindi,

$$\sum_{R \in \mathcal{R}_1} \sup_R f - \inf_R f \text{mis } R \leq \sum_{R \in \mathcal{R}_1} \sup_R s_2 - \inf_R s_1 \text{mis } R \leq \sum_{R \in \mathcal{R}} \sup_R s_2 - \inf_R s_1 \text{mis } R \leq \varepsilon.$$

D'altra parte, esiste una costante $c > 0$ indipendente da δ , tale che $\sum_{R \in \mathcal{R}_2} \text{mis } R < c\delta$. Dall'arbitrarietà di δ segue l'asserto.

(iii) Deriva facilmente dai punti (ii) e (i) precedenti. ■

1.28 Se $f : A \rightarrow \mathbb{R}$ è una funzione continua e limitata su di un insieme misurabile $A \subseteq \mathbb{R}^n$ allora f è integrabile su A .

Dimostrazione Sia $\varepsilon > 0$ e sia E un rettangolo standard contenente A e sia P una partizione di E per cui valga (4.20) con $B = A$. Poiché f è continua sul compatto $K := \bigcup_{R \in \mathcal{R}'_A} R \subseteq A$ esiste $\delta > 0$ tale che $|f(x) - f(x')| \leq \varepsilon$ per ogni coppia di punti x e x' in K con $|x - x'| \leq \delta$. Sia P' un raffinamento di P di diametro non superiore a δ e chiamiamo \mathcal{R}_1 i rettangoli di $\mathcal{R}(P')$ che appartengono a K e $\mathcal{R}_2 := \mathcal{R}_A(P') \setminus \mathcal{R}_1$. Chiaramente da (4.20) segue che $\sum_{R \in \mathcal{R}_2} \text{mis } R \leq \varepsilon$ e che $\text{osc}(f_A, R) := \text{osc}(f, R) \leq \varepsilon$ per ogni $R \in \mathcal{R}_1$. Dunque

$$\sum_{R \in \mathcal{R}(P')} \text{osc}(f_A, R) \text{mis } R \leq 2 \sup_A |f| \sum_{R \in \mathcal{R}_2} \text{mis } R + \varepsilon \sum_{R \in \mathcal{R}_1} \text{mis } R \leq \varepsilon (2 \sup_A |f| + \text{mis } A).$$

Da 1.7 e dall'arbitrarietà di ε segue la tesi. ■

1.29 Sia $F : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione uniformemente Lipschitziana con costante di Lipschitz $L > 0$ rispetto alla norma del sup. Allora, per ogni $\delta > 0$, per ogni partizione P di E con $\text{diam } P \leq \delta$ e per ogni scelta di punti Q relativa a P si ha¹⁸

$$\left| \int_E f(x) \, dx - S(f, P, Q) \right| \leq \delta L \, \text{mis } E. \quad (4.34)$$

Dimostrazione Segue subito osservando che se il diametro della partizione P è minore di δ , l'oscillazione di F su un rettangolo qualunque della partizione è minore di $L\delta$. ■

1.30 Se $f : B = \bigcup_i B_i \rightarrow \mathbb{R}$ è continua a tratti (cfr. 1.9), allora f è integrabile su B e

$$\int_B f = \sum_{i=1}^N \int_{B_i} f.$$

Dimostrazione Da 1.24 segue che B è misurabile e che (essendo $\text{mis } \partial B_i = 0$ per ogni i) $\text{mis } B = \sum_{i=1}^N \text{mis } B_i$; da 1.28 segue la tesi. ■

Integrali iterati

La seguente proposizione, di fondamentale importanza nella pratica, permette, sotto opportune ipotesi sul dominio di integrazione, di ridurre il calcolo di un integrale n -dimensionale (ossia di una funzione di n variabili) al calcolo successivo di un integrale unidimensionale e di un integrale $(n-1)$ -dimensionale.

Proposizione 4.1 Sia $n \geq 2$ e $A \subseteq \mathbb{R}^{n-1}$ un insieme misurabile; siano α e β due funzioni integrabili su A e sia $B \subseteq \mathbb{R}^n$ l'insieme definito come

$$B := \{(x, y) \in A \times \mathbb{R} : \alpha(x) \leq y \leq \beta(x)\}. \quad (4.35)$$

Allora: (i) B è misurabile e

$$\text{mis } B = \int_A (\beta(x) - \alpha(x)) \, dx. \quad (4.36)$$

(ii) Se f è una funzione integrabile su B e $y \rightarrow f(x, y)$ è integrabile su $[\alpha(x), \beta(x)]$ per ogni $x \in A$ allora la funzione

$$g(x) := \int_{\alpha(x)}^{\beta(x)} f(x, y) \, dy \quad (4.37)$$

è integrabile su A e, (iii) ,

$$\int_B f = \int_A g := \int_A \left(\int_{\alpha(x)}^{\beta(x)} f(x, y) \, dy \right) dx. \quad (4.38)$$

Osservazione 4.2 (i) Un insieme B della forma (4.35) prende il nome di **insieme normale** rispetto all'asse delle¹⁹ y ; nel caso A sia misurabile e le funzioni α e β siano continue e limitate su A l'insieme B si dice C^0 -normale. Ovviamente un rettangolo E è normale (rispetto a qualunque asse e le funzioni α e β sono costanti).

(ii) Un caso assai importante nelle applicazioni è quando α , β e f sono continue e limitate (rispettivamente su A e su B): in tal caso infatti, per 1.28, le ipotesi della Proposizione 4.1 sono soddisfatte.

¹⁸Questa formula è alla base del calcolo approssimato (con un semplice controllo dell'errore) degli integrali di funzioni regolari su rettangoli.

¹⁹Naturalmente il ruolo di x_1, \dots, x_{n-1} e y è del tutto arbitrario e analogamente si definirà un insieme normale rispetto ad un qualunque asse.

(iii) È chiaro che se α , β e f sono continue e limitate (su A e su B) e se A è a sua volta un insieme C^0 -normale in \mathbb{R}^{n-1} la proposizione può essere riapplicata. Possiamo quindi costruire induttivamente la classe degli in insiemi C^0 -normali in \mathbb{R}^n , denotata con \mathcal{N}^n , usando la (4.35) con $A \in \mathcal{N}^{n-1}$. Iterando, ove possibile, si ridurrà il calcolo dell'integrale di f su B al calcolo successivo di n integrali unidimensionali. In particolare se $E := [a_1, b_1] \times \cdots \times [a_n, b_n]$ è un rettangolo standard e $f \in C(E)$ e se σ è una qualunque una permutazione²⁰ di $\{1, \dots, n\}$, allora

$$\int_E f = \int_{a_{\sigma_n}}^{b_{\sigma_n}} \left(\cdots \left(\int_{a_{\sigma_1}}^{b_{\sigma_1}} f(x) dx_{\sigma_1} \right) \cdots \right) dx_{\sigma_n}. \quad (4.39)$$

(iv) Le ipotesi della Proposizione 4.1 non possono essere indebolite; si vedano a tal proposito E 4.2 e E 4.3 di fine sezione.

Dimostrazione Sia E un rettangolo standard in \mathbb{R}^{n-1} che contiene A e sia $a := \inf_A \alpha < b := \sup_A \beta$, cosicchè $B \subseteq E' := E \times [a, b]$.

Cominciamo col dimostrare la misurabilità di B costruendo, dato $\varepsilon > 0$, due insiemi elementari B_1 e B_2 tali che $B_1 \subseteq B \subseteq B_2$ e $\text{mis}_n B_2 \setminus B_1 \leq c\varepsilon$ con una costante $c > 0$ indipendente da ε (cfr. 1.23–(ii)). Dalle ipotesi segue che esiste una partizione P di E tale che

$$\sum_{R \in \mathcal{R}_A(P) \setminus \mathcal{R}'_A(P)} \text{mis } R \leq \varepsilon, \quad (4.40)$$

$$\sum_{R \in \mathcal{R}(P)} \text{osc}(\alpha_A, R) \text{mis } R \leq \varepsilon, \quad \sum_{R \in \mathcal{R}(P)} \text{osc}(\beta_A, R) \text{mis } R \leq \varepsilon, \quad (4.41)$$

dove $\text{mis} = \text{mis}_{n-1}$. Definiamo le seguenti famiglie disgiunte di rettangoli di $\mathcal{R}(P)$:

$$\begin{aligned} \mathcal{R}_0 &:= \mathcal{R}_A(P) \setminus \mathcal{R}'_A(P), \\ \mathcal{R}_1 &:= \{R \in \mathcal{R}'_A(P) : \sup_R \alpha \geq \inf_R \beta\}, \\ \mathcal{R}_2 &:= \mathcal{R}'_A(P) \setminus \mathcal{R}_1 = \{R \in \mathcal{R}'_A(P) : \sup_R \alpha < \inf_R \beta\}. \end{aligned}$$

Si noti che \mathcal{R}_1 potrebbe essere vuota, ma se non lo è, allora per ogni $R \in \mathcal{R}_1$ si ha che

$$\begin{aligned} \sup_R \beta - \inf_R \alpha &= (\sup_R \beta - \inf_R \beta) - (\sup_R \alpha - \inf_R \beta) + (\sup_R \alpha - \inf_R \alpha) \\ &\leq (\sup_R \beta - \inf_R \beta) + (\sup_R \alpha - \inf_R \alpha) \\ &= \text{osc}(\beta_A, R) + \text{osc}(\alpha_A, R). \end{aligned} \quad (4.42)$$

Definiamo ora

$$\begin{aligned} B_1 &:= \bigcup_{R \in \mathcal{R}_2} R \times [\sup_R \alpha, \inf_R \beta], \\ B_2^{(0)} &:= \bigcup_{R \in \mathcal{R}_0} R \times [a, b], \\ B_2^{(1)} &:= \bigcup_{R \in \mathcal{R}_1} R \times [\inf_R \alpha, \sup_R \beta], \\ B_2^{(2)} &:= \bigcup_{R \in \mathcal{R}_2} R \times [\inf_R \alpha, \sup_R \beta], \\ B_2 &:= B_2^{(0)} \cup B_2^{(1)} \cup B_2^{(2)}. \end{aligned}$$

Chiaramente questi insiemi sono elementari e

$$B_1 \subseteq B \subseteq B_2, \quad B_2 \setminus B_1 = B_2^{(0)} \cup B_2^{(1)} \cup (B_2^{(2)} \setminus B_1).$$

²⁰Una permutazione σ dell'insieme $I = \{1, \dots, n\}$ è una applicazione biunivoca $\sigma : i \in I \rightarrow \sigma_i \in I$.

Ora, da (4.40) segue che

$$\text{mis}_n B_2^{(0)} = \left(\sum_{R \in \mathcal{R}_0} \text{mis}_n R \right) (b - a) \leq \varepsilon (b - a) ; \quad (4.43)$$

mentre da (4.42) e da (4.41) segue che

$$\text{mis}_n B_2^{(1)} = \sum_{R \in \mathcal{R}_1} (\text{mis } R) (\sup_R \beta - \inf_R \alpha) \leq \sum_{R \in \mathcal{R}_1} (\text{mis } R) \left(\text{osc}(\beta_A, R) + \text{osc}(\alpha_A, R) \right) \leq 2\varepsilon . \quad (4.44)$$

Infine,

$$\begin{aligned} \text{mis}_n (B_2^{(2)} \setminus B_1) &= \sum_{R \in \mathcal{R}_2} (\text{mis } R) \left((\sup_R \beta - \inf_R \beta) + (\sup_R \alpha - \inf_R \alpha) \right) \\ &\leq \sum_{R \in \mathcal{R}(P)} \text{osc}(\alpha_A, R) \text{mis } R + \sum_{R \in \mathcal{R}(P)} \text{osc}(\beta_A, R) \text{mis } R \\ &\leq 2\varepsilon . \end{aligned} \quad (4.45)$$

Mettendo assieme (4.43) ÷ (4.45) si ha che $\text{mis } B_2 \setminus B_1 \leq c\varepsilon$ con $c = 4 + (b - a)$ da cui segue la misurabilità di B .

Dimostriamo l'integrabilità di g su A e simultaneamente la (4.38) (dalla quale deriva immediatamente (4.36) scegliendo $f := 1$). Innanzitutto g è limitata su A infatti, per ogni $x \in A$, si ha

$$|g(x)| \leq (\sup_A |\beta - \alpha|) (\sup_B |f|) \leq \left(\sup_A |\beta| + \sup_A |\alpha| \right) (\sup_B |f|) .$$

Sia ora P' una qualunque partizione di E' e si noti che $P' = (P, P_n)$ con P partizione di E e P_n partizione di $[a, b]$ e che i rettangoli R' di $\mathcal{R}(P')$ sono dati dai prodotti cartesiani dei rettangoli $R \in \mathcal{R}(P)$ per rettangoli (intervalli) $I \in \mathcal{R}(P_n)$. Si noti anche che dalle ipotesi segue che la funzione di una variabile $y \rightarrow f_B(x, y)$ è integrabile su I . Quindi, dato $R' = R \times I \in \mathcal{R}(P')$ e fissato $x \in R$, integrando su I la relazione

$$\inf_{R'} f_B \leq f_B(x, y) \leq \sup_{R'} f_B , \quad (4.46)$$

si ha che

$$(\inf_{R'} f_B) \text{mis } I \leq \int_{I'} f_B(x, y) dy \leq (\sup_{R'} f_B) \text{mis } I . \quad (4.47)$$

Sommando le relazioni in (4.47) su tutti gli $I \in \mathcal{R}(P_n)$ (tenendo fisso il rettangolo R di $\mathcal{R}(P)$), otteniamo per ogni $x \in R$,

$$\begin{aligned} \sum_{I \in \mathcal{R}(P_n)} (\inf_{R'} f_B) \text{mis } I &\leq \int_I f_B(x, y) dy = \int_{\alpha(x)}^{\beta(x)} f_B(x, y) dy = g_A(x) \\ \sum_{I \in \mathcal{R}(P_n)} (\sup_{R'} f_B) \text{mis } I &\geq \int_I f_B(x, y) dy = \int_{\alpha(x)}^{\beta(x)} f_B(x, y) dy = g_A(x) . \end{aligned} \quad (4.48)$$

Prendendo, rispettivamente, l'estremo inferiore su R nella prima riga di (4.48) e quello superiore nella seconda, otteniamo

$$\sum_{I \in \mathcal{R}(P_n)} (\inf_{R'} f_B) \text{mis } I \leq \inf_R g_A , \quad \sum_{I \in \mathcal{R}(P_n)} (\sup_{R'} f_B) \text{mis } I \geq \sup_R g_A . \quad (4.49)$$

Quindi,

$$\begin{aligned}
\underline{S}_{E'}(f_B, P') &= \sum_{R' \in \mathcal{R}(P')} (\inf_{R'} f_B) \text{mis}(R') = \sum_{\substack{R \in \mathcal{R}(P) \\ I \in \mathcal{R}(P_n)}} (\inf_{R'} f_B) \text{mis } R \cdot \text{mis } I \\
&= \sum_{R \in \mathcal{R}(P)} \text{mis } R \sum_{I \in \mathcal{R}(P_n)} (\inf_{R'} f_B) \text{mis } I \\
&\leq \sum_{R \in \mathcal{R}(P)} \text{mis } R \inf_R g_A = \underline{S}_E(g_A, P)
\end{aligned} \tag{4.50}$$

e (ragionando in maniera analoga per le somme superiori ed usando la seconda delle (4.50)) si ottiene $\overline{S}_{E'}(f_B, P') \geq \overline{S}_E(g_A, P)$. Dall'integrabilità di f_B su E' segue dunque l'asserto. ■

1.3 Misura di Peano–Jordan

Sia $E \subseteq \mathbb{R}^n$ un rettangolo standard, P una partizione di E e $B \subseteq E$. Denotiamo con $\mathcal{R}_B(P)$ e $\mathcal{R}'_B(P)$ i seguenti insiemi di rettangoli di $\mathcal{R}(P)$:

$$\begin{aligned}
\mathcal{R}_B(P) &:= \{R \in \mathcal{R}(P) : R \cap B \neq \emptyset\}, \\
\mathcal{R}'_B(P) &:= \{R \in \mathcal{R}(P) : R \subseteq B\}.
\end{aligned} \tag{4.51}$$

Sia P' un raffinamento di P . Ricordando che ogni rettangolo R di $\mathcal{R}(P)$ è dato dall'unione dei rettangoli di $\mathcal{R}(P')$ contenuti in R si ha che

$$\sum_{R \in \mathcal{R}'_B(P)} \text{mis } R \leq \sum_{R \in \mathcal{R}'_B(P')} \text{mis } R, \tag{4.52}$$

$$\sum_{R \in \mathcal{R}_B(P')} \text{mis } R \leq \sum_{R \in \mathcal{R}_B(P)} \text{mis } R.$$

Da tali relazioni segue subito che, se P e P' sono due partizioni arbitrarie di E ,

$$\sum_{R \in \mathcal{R}'_B(P)} \text{mis } R \leq \sum_{R \in \mathcal{R}'_B(P \cup P')} \text{mis } R \leq \sum_{R \in \mathcal{R}_B(P \cup P')} \text{mis } R \leq \sum_{R \in \mathcal{R}_B(P')} \text{mis } R. \tag{4.53}$$

Si definiscono, rispettivamente, la *misura interna* e la *misura esterna* (secondo Peano–Jordan) di B le quantità

$$\text{mis int } B := \sup_{\{P\}} \sum_{R \in \mathcal{R}'_B(P)} \text{mis } R, \quad \text{mis est } B := \inf_{\{P\}} \sum_{R \in \mathcal{R}_B(P)} \text{mis } R, \tag{4.54}$$

dove gli estremi inferiori e superiori sono presi su tutte le partizioni P di E . Da (4.53), segue che

$$0 \leq \text{mis int } B \leq \text{mis est } B \leq \text{mis } E. \tag{4.55}$$

L'insieme B si dice *misurabile* secondo Peano–Jordan se $\text{mis int } B = \text{mis est } B$ e tale comune valore viene chiamato la *misura di Peano–Jordan* di B e si denota con $\text{mis}_n B$ o $\text{mis } B$.

1.4 Teorema del cambio di variabili

Fissiamo le seguenti notazioni: per $x \in \mathbb{R}^n$ e $T \in \mathbb{R}^{n \times n}$ denotiamo con $|x| = \max |x_i|$ la norma del massimo e con $\|T\| = \sup_{x \neq 0} |Tx|/|x|$ la relativa norma matriciale. Denotiamo anche con $K_r(x) = \{y \in \mathbb{R}^n \mid |y - x| \leq r\}$ il cubo chiuso di centro x e lato $2r$.

Cominciamo con alcuni elementari risultati preliminari.

Il primo lemma è un risultato elementare topologico che descrive come si comporta l'immagine della frontiera di un insieme rispetto alla frontiera dell'immagine. Si ricorda che, per un qualunque insieme $A \subseteq \mathbb{R}^n$,

$$\partial A := \overline{A} \setminus \overset{\circ}{A} = \{x \in \mathbb{R}^n \mid \forall U \text{ intorno di } x, U \cap A \neq \emptyset \neq U \cap A^c\}. \quad (4.56)$$

Lemma 4.3 (i) *Sia A un insieme limitato in \mathbb{R}^n e $\phi \in C(\overline{A}, \mathbb{R}^m)$, allora $\overline{\phi(A)} = \phi(\overline{A})$; se, inoltre, A e $\phi(A)$ sono aperti, allora*

$$\partial\phi(A) \subseteq \phi(\partial A). \quad (4.57)$$

(ii) *Sia $\phi : A \subseteq \mathbb{R}^n \rightarrow B := \phi(A) \subseteq \mathbb{R}^m$ una mappa continua e iniettiva e sia E tale che $\overline{E} \subseteq A$. Allora²¹:*

$$\phi(\partial E) \subseteq \partial\phi(E). \quad (4.58)$$

Inoltre, se si ha anche che $\overline{\phi(E)} \subseteq B$, allora

$$\phi(\partial E) = \partial\phi(E). \quad (4.59)$$

Dimostrazione (i) Sia $y \in \overline{\phi(A)}$. Allora, esistono $y_k \in \phi(A)$ tali che $y_k \rightarrow y$ e $x_k \in A$ tali che $\phi(x_k) = y_k$. Poiché \overline{A} è compatto, esistono $x_{k_j} \in \overline{A}$ tali che $x_{k_j} \rightarrow \bar{x}$. Dunque, $\phi(\bar{x}) = \lim \phi(x_{k_j}) = \lim y_{k_j} = y$. Abbiamo mostrato che $\overline{\phi(A)} \subseteq \phi(\overline{A})$. Sia, ora, $\bar{y} = \phi(\bar{x})$ con $\bar{x} \in \overline{A}$. Allora esistono $x_k \in A$ tali che $x_k \rightarrow \bar{x}$. Quindi $\phi(x_k) \rightarrow \phi(\bar{x})$ e $\bar{y} := \phi(\bar{x}) \in \overline{\phi(A)}$, ossia $\phi(\overline{A}) \subseteq \overline{\phi(A)}$. Assumiamo ora che A e $\phi(A)$ siano aperti. Allora, per quanto dimostrato, $\partial\phi(A) = \phi(\overline{A}) \setminus \phi(A)$. Quindi, se $y \in \partial\phi(A)$, esiste $x \in \overline{A} \setminus A = \partial A$ tale che $\phi(x) = y$, ossia, $y \in \phi(\partial A)$.

(ii) Essendo ϕ iniettiva da A su B è definita la funzione inversa $\phi^{-1} : B \rightarrow A$. Si noti che per ipotesi $\partial E \subseteq A$. Sia $F := \phi(E)$.

Dimostriamo la (4.58). Sia $x \in \partial E$, $y = \phi(x)$ e V un intorno di y . Allora, $U = \phi^{-1}(V)$ è un intorno di x e poiché $x \in \partial E$ esistono $x_1 \in E \cap U$ e $x_2 \in E^c \cap U$. Quindi, se $y_i = \phi(x_i)$, si ha che²² $y_1 \in V \cap F$ e $y_2 \in V \cap \phi(E^c) = V \cap F^c$, ossia $y \in \partial F$ e la (4.58) vale.

Dimostriamo la (4.59). Sia $y \in \partial F \subseteq B$ e sia V un intorno di y . Allora, esistono $y_1 \in V \cap F$ e $y_2 \in V \cap F^c$. Siano $x = \phi^{-1}(y)$, $x_i = \phi^{-1}(y_i)$, $U = \phi^{-1}(V)$. Allora, U è un intorno di x e $x_1 \in U \cap E$ e $x_2 \in U \cap E^c$, ossia, $x \in \partial E$. Quindi $\partial\phi(E) \subseteq \phi(\partial E)$ e, poiché vale anche (4.58), segue la (4.59) vale. ■

Si ricorda che dato un insieme $A \subseteq \mathbb{R}^n$ si ha che²³

$$A \text{ è PJ-misurabile} \iff \forall \varepsilon > 0 \exists E \text{ elementare} \mid \partial A \subseteq E \text{ e } \text{mis } E < \varepsilon \quad (4.60)$$

$$\iff \forall \varepsilon > 0 \exists N \text{ cubi } K_j \mid \partial A \subseteq \bigcup_{j=1}^N K_j \text{ e } \sum_{j=1}^N \text{mis } K_j < \varepsilon \quad (4.61)$$

nella (4.61) i cubi possono essere presi sia aperti che chiusi.

La proprietà (4.61) viene preservata da trasformazioni uniformemente lipschitziane²⁴:

Lemma 4.4 *Sia $Q \subseteq \mathbb{R}^n$ un insieme che gode della seguente proprietà:*

$$\forall \varepsilon > 0 \exists \{K_j \mid 1 \leq j \leq N\}, K_j \text{ cubi} \mid Q \subseteq \bigcup_{j=1}^N K_j, \sum_{j=1}^N \text{mis } K_j \leq \varepsilon. \quad (4.62)$$

Sia $\phi : Q \rightarrow \mathbb{R}^m$ una mappa uniformemente lipschitziana su Q . Allora, $\phi(Q)$ gode della stessa proprietà di Q (ossia, vale la (4.62) con $\phi(Q)$ al posto di Q).

²¹Ovviamente, nella (4.58) può valere l'inclusione propria come nel seguente esempio: $E = A = \mathbb{R}$, $B = (-1, 1)$, $\phi(x) = \tanh x$, nel qual caso, $\partial A = \emptyset$ e $\partial\phi(A) = \{0, 1\}$.

²²Mappe iniettive conservano le unioni, le intersezioni e i complementari.

²³'PJ-misurabile' è una abbreviazione di 'misurabile secondo Peano-Jordan'.

²⁴Una mappa $\phi : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ si dice uniformemente lipschitziana (su E) se esiste $L > 0$ tale che $|\phi(x) - \phi(y)| \leq L|x - y|$ per ogni $x, y \in E$.

Dimostrazione Fissiamo $\varepsilon > 0$. Per l'ipotesi su Q , esistono N cubi $K_j = K_{r_j}(x_j)$ tali che $Q \subseteq \bigcup K_j$ e tali che $\sum \text{mis } K_j < \varepsilon / (2L)^n$, dove L è la costante di Lipschitz di ϕ . Per ogni j fissiamo²⁵ $\bar{x}_j \in K_j \cap Q$. I cubi $K'_j := K_{2Lr_j}(\phi(\bar{x}_j))$ formano un ricoprimento di $\phi(Q)$: infatti, se $y \in \phi(Q)$, esiste $x \in Q$ tale che $\phi(x) = y$ ed esiste j tale che $x \in K_j$. Dunque, poiché $x, \bar{x}_j \in K_j$,

$$|y - \phi(\bar{x}_j)| = |\phi(x) - \phi(\bar{x}_j)| \leq L|x - \bar{x}_j| \leq L(|x - x_j| + |x_j - \bar{x}_j|) \leq 2Lr_j,$$

ossia $y \in K'_{2Lr_j}(\phi(\bar{x}_j))$. Inoltre

$$\sum_{j=1}^N \text{mis } K'_j = \sum_{j=1}^N (2Lr_j)^n = (2L)^n \sum_{j=1}^N \text{mis } K_j < \varepsilon. \quad \blacksquare$$

Lemma 4.5 Sia $A \subseteq \mathbb{R}^n$ un insieme aperto limitato e $\phi : A \rightarrow B := \phi(A) \subseteq \mathbb{R}^m$ una mappa iniettiva uniformemente lipschitziana su \bar{A} .

- (i) Se E è un insieme PJ-misurabile tale che $\bar{E} \subseteq A$ e $\overline{\phi(E)} \subseteq B$, allora $\phi(E)$ è PJ-misurabile.
(ii) Se A è PJ-misurabile e B è aperto, allora B è PJ-misurabile.

Dimostrazione (i) segue dal Lemma 4.3–(ii) (cfr. (4.59)) e dal Lemma 4.4.

(ii) segue dal Lemma 4.3–(i) e dal Lemma 4.4. \blacksquare

Infine ricordiamo che dal teorema della funzione inversa segue che²⁶

$$A \subseteq \mathbb{R}^n \text{ aperto, } \phi \in C^1(A, \mathbb{R}^m), \det \phi' \neq 0 \text{ su } A \implies \phi(A) \text{ aperto.} \quad (4.63)$$

Teorema 4.6 Sia $A \subseteq \mathbb{R}^n$ un insieme aperto PJ-misurabile e $\phi \in C^1(\bar{A}, \mathbb{R}^m)$ tale che

$$\phi \text{ è iniettiva in } A, \quad \det \phi' \neq 0 \quad \text{su } A. \quad (4.64)$$

Allora $B := \phi(A)$ è un insieme aperto PJ-misurabile e

$$\text{mis}(\phi(A)) = \int_A |\det \phi'|. \quad (4.65)$$

Inoltre, se $f \in \mathcal{R}(B)$ allora $f \circ \phi \in \mathcal{R}(A)$ e si ha

$$\int_B f = \int_A f \circ \phi |\det \phi'|. \quad (4.66)$$

Osserviamo subito che dal punto (ii) del Lemma 4.5 e da (4.63) segue immediatamente che $B = \phi(A)$ è un aperto PJ-misurabile.

Lemma 4.7 Sia $A \subseteq \mathbb{R}^n$ PJ-misurabile e $T \in \mathbb{R}^{n \times n}$. Allora TA è PJ-misurabile e

$$\text{mis}(TA) = |\det T| \text{mis}(A). \quad (4.67)$$

Dimostrazione (1) Se $\det T = 0$, i vettori colonna che formano la matrice T sono linearmente dipendenti e quindi l'insieme limitato TA è contenuto in un sottospazio vettoriale di dimensione $(n-1)$ e quindi TA ha misura di Peano–Jordan nulla e la (4.67) vale banalmente.

(2) D'ora in poi assumiamo che $\det T \neq 0$. Si noti che in questo caso la mappa lineare $x \rightarrow Tx$ è un diffeomorfismo di \mathbb{R}^n e dal Lemma 4.5–(i) segue che TA è PJ-misurabile.

Osserviamo che la (4.67) è additiva: ossia se A_1 e A_2 sono due insiemi PJ-misurabili tali che $\text{mis}(A \cap B) = 0$ anche TA_1 e TA_2 sono tali e dunque, se vale (4.67), allora vale anche

$$\text{mis}(T(A_1 \cup A_2)) = |\det T| (\text{mis}(A_1) + \text{mis}(A_2)), \quad \text{mis}(A_1 \cap A_2) = 0. \quad (4.68)$$

²⁵Ovviamente possiamo non considerare cubi tali che $K_j \cap Q = \emptyset$.

²⁶Si ricorda che $\phi' = \phi'(x)$ denota lo jacobiano di ϕ ossia la matrice con elementi $\frac{\partial \phi_k}{\partial x_j}$.

Ne segue che se (4.67) vale per rettangoli vale anche per insiemi elementari.

(3) Dimostriamo ora la (4.67) con A rettangolo $I_1 \times \cdots \times I_n = \prod I_j$ con lati $I_j = [0, L_j]$ e T matrice elementare; cfr Appendice A, di cui useremo le notazioni.

Se $T = R_\lambda^i$, allora $TA = A' = \prod I'_j$ dove $I'_j = I_j$ se $i \neq j$ e I'_i intervallo chiuso di estremi 0 e λL_i . Dunque, $\text{mis}(TA) = |\lambda| \prod_j L_j = |\det T| \text{mis} A$ e la (4.67) è verificata.

Se $T = T^{ij}$, TA è dato dal rettangolo con I_i e I_j scambiati e poiché $\det T^{ij} = 1$, la (4.67) è ovviamente verificata.

Se $T = S^{ij}$, la (4.67) è conseguenza immediata del teorema sugli integrali iterati (Fubini): infatti

$$S^{ij}x = \sum_{k \neq j} x_k e^{(k)} + (x_i + x_j)e^{(j)}, \quad S^{ij}A = \{0 \leq x_k \leq L_k \mid k \neq i, j \text{ \& } (x_i, x_j) \in P_{ij}\},$$

dove P_{ij} denota l'immagine della mappa $(x_i, x_j) \mapsto (x_i, x_i + x_j)$ del rettangolo $[0, L_i] \times [0, L_j]$, ossia, il 'parallelogramma' $\{(x_i, x_j) \mid 0 \leq x_i \leq L_i, x_i \leq x_j \leq x_i + L_j\}$. Dunque, poiché $\det S^{ij} = 1$, per il teorema sugli integrali iterati,

$$\text{mis}(S^{ij}A) = \int_{\mathbb{R}^n} \chi_{S^{ij}A} = \left(\prod_{k \neq i, j} L_k \right) \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \chi_{P_{ij}}(x_i, x_j) dx_j \right) dx_i = \prod_k L_k = |\det S^{ij}| \text{mis} A.$$

(4) Poiché la misura di Peano-Jordan è invariante per traslazioni, dai punti precedenti segue che la (4.67) vale per T elementare e per qualunque rettangolo limitato. Per additività (cfr. punto (2)), la (4.67) vale per matrici elementari ed insiemi A elementari e dunque per ogni A PJ-misurabile.

(5) Poiché ogni matrice T si può scrivere come prodotto finito di matrici elementari T_i (cfr. Proposizione A in Appendice) e poiché $T_i B$ è PJ-misurabile se B lo è, per la moltiplicatività del determinante, la (4.67) segue per una qualunque matrice T invertibile. ■

Osservazione 4.8 Una 'mappa affine' è una funzione da \mathbb{R}^n in \mathbb{R}^n della forma $x \in \mathbb{R}^n \mapsto Lx = x_0 + Tx$ con $x_0 \in \mathbb{R}^n$ e $T \in \mathbb{R}^{n \times n}$, il cui jacobiano è dato da $L' = T$. Dunque, per l'invarianza per traslazioni della misura di Peano-Jordan segue che se A è PJ-misurabile allora lo è anche LA e

$$\text{mis}(LA) = |\det L'| \text{mis}(A). \quad (4.69)$$

Il prossimo lemma è una versione parziale della (4.65):

Lemma 4.9 *Siano ϕ ed A come nel Teorema 4.6 e sia $E \subseteq A$ un insieme PJ-misurabile. Allora*

$$\text{mis}(\phi(E)) \leq \int_E |\det \phi'|. \quad (4.70)$$

Dimostrazione Cominciamo col dimostrare la (4.70) nel caso di E cubo chiuso contenuto in A . Sia $0 < \varepsilon < 1$ e siano $K_j := K_\delta(x_j)$ N cubi tali che $\text{mis}(K_i \cap K_j) = 0$ se $i \neq j$ e tali che $E = \bigcup K_j$. Poniamo

$$\Lambda_j x := \phi(x_j) + \phi'(x_j)(x - x_j), \quad \Lambda'_j = \phi'(x_j), \quad \Lambda_j^{-1} y = x_j + \phi'(x_j)^{-1}(y - \phi(x_j)). \quad (4.71)$$

Si noti che $\Lambda_j x$ non è altro che il polinomio (vettoriale) di Taylor di ordine 1 di ϕ in x_j . Sia

$$\psi_j(x) := \Lambda_j^{-1} \circ \phi(x) = x_j + \phi'(x_j)^{-1}(\phi(x) - \phi(x_j)), \quad \psi'_j(x) = \phi'(x_j)^{-1} \phi'(x). \quad (4.72)$$

Poiché ϕ è iniettiva su E si ha che²⁷

$$\phi(E) = \phi\left(\bigcup K_j\right) = \bigcup \phi(K_j), \quad \text{mis} \phi(E) = \sum \text{mis}(\phi(K_j)). \quad (4.73)$$

²⁷Si noti che $\text{mis}(K_i \cap K_j) = 0$ e $\text{mis}(\phi(K_i) \cap \phi(K_j)) = 0$ se $i \neq j$.

Per il Lemma 4.5–(i) gli insiemi $\phi(K_j)$ e gli insiemi $\psi_j(K_j)$ sono PJ–misurabili e

$$\begin{aligned} \text{mis}(\phi(E)) &\stackrel{(4.73)}{=} \sum \text{mis}(\phi(K_j)) = \sum \text{mis}(\Lambda_j(\Lambda_j^{-1}\phi(K_j))) \\ &\stackrel{(4.72)}{=} \sum \text{mis}(\Lambda_j(\psi_j(K_j))) \stackrel{(4.69,4.71)}{=} \sum |\det \phi'(x_j)| \text{mis}(\psi_j(K_j)). \end{aligned} \quad (4.74)$$

Dal teorema del valor medio, dalla (4.72) e dalla uniforme continuità di $\|\phi'(x)^{-1}\phi'(y)\|$ su E , segue che esiste δ_0 tale che se $0 < \delta \leq \delta_0$, allora

$$|\psi_j(x) - \psi_j(y)| \leq \left(1 + \frac{\varepsilon}{4n}\right)|x - y|, \quad \forall j, \forall x, y \in K_j.$$

Quindi

$$\psi_j(K_j) = \psi_j(K_\delta(x_j)) \subseteq K_{(1+\varepsilon/4n)\delta}(\phi(x_j)),$$

da cui segue che

$$\text{mis}(\psi_j(K_j)) \leq \left(1 + \frac{\varepsilon}{4n}\right)^n (2\delta)^n \leq \left(1 + \frac{\varepsilon}{4}\right)(2\delta)^n = \left(1 + \frac{\varepsilon}{4}\right) \text{mis}(K_j). \quad (4.75)$$

Dalla uniforme continuità di $|\det \phi'(x)|$ su E e da (4.64) segue anche che esiste $\delta_1 \leq \delta_0$ tale che se $0 < \delta \leq \delta_1$, allora $|\det \phi'(x_j)|/|\det \phi'(x)| < (1 + \varepsilon/4)$ per ogni j e per ogni $x \in K_j$. Dunque:

$$\text{mis}|\det \phi'(x_j)| \text{mis}(\psi_j(K_j)) = \int_{K_j} |\det \phi'(x_j)| \leq \left(1 + \frac{\varepsilon}{4}\right) \int_{K_j} |\det \phi'|, \quad \forall j. \quad (4.76)$$

Mettendo assieme (4.74), (4.75), (4.76), si ha che, se $\delta \leq \delta_1$, allora

$$\text{mis}(\phi(E)) \leq \left(1 + \frac{\varepsilon}{4}\right)^2 \sum \int_{K_j} |\det \phi'| < (1 + \varepsilon) \sum \int_{K_j} |\det \phi'| = (1 + \varepsilon) \int_E |\det \phi'|.$$

Dall'arbitrarietà di ε segue la (4.70) nel caso E sia un cubo chiuso contenuto in A .

La generalizzazione a insiemi $E \subseteq A$ misurabili secondo Peano–Jordan è routine: poiché per ogni rettangolo chiuso esistono rettangoli R_j unione finita disgiunta di cubi tali che $R_j \subseteq R_{j+1} \subseteq R$ e $\cup R_j = R$, per additività la (4.70) vale per rettangoli chiusi in A e quindi (sempre per additività) per insiemi elementari la cui chiusura è contenuta in A ; infine, poiché se E è PJ–misurabile allora esistono $E_k \subseteq \overset{\circ}{E}$ tali che $E_k \subseteq E_{k+1}$ e $\sup \text{mis} E_k = \text{mis} E$, la (4.70) vale per ogni insieme misurabile $E \subseteq A$ (in particolare per $E = A$). ■

Dimostrazione della (4.65) Cominciamo col dimostrare che vale la disuguaglianza inversa di (4.70) con E cubo chiuso contenuto in A , ossia,

$$\int_E |\det \phi'| \leq \text{mis}(\phi(E)), \quad \forall \text{ cubo chiuso } E \subseteq A. \quad (4.77)$$

Come sopra, sia $0 < \varepsilon < 1$ e sia $K_j := K_\delta(x_j)$ N cubi tali che $\text{mis}(K_i \cap K_j) = 0$ se $i \neq j$ e tali che $E = \cup K_j$. Sia Λ_j come in (4.71) e

$$\varphi_j(x) := \Lambda_j \circ \phi^{-1}(x) = \phi(x_j) + \phi'(x_j)(\phi^{-1}(x) - x_j), \quad \varphi_j'(x) = \phi'(x_j)\phi'(x)^{-1}. \quad (4.78)$$

Per l'uniforme continuità di $|\det \phi'|$ su E , esiste $0 < \delta_0$ tale che se $0 < \delta \leq \delta_0$, allora $|\det \phi'(x)|/|\det \phi'(x_j)| < (1 + \varepsilon/4)$ per ogni j e per ogni $x \in K_j$. Quindi

$$\begin{aligned} \int_E |\det \phi'| &= \sum \int_{K_j} |\det \phi'| \leq \left(1 + \frac{\varepsilon}{4}\right) \sum |\det \phi'(x_j)| \text{mis}(K_j) \\ &\stackrel{(4.67)}{=} \left(1 + \frac{\varepsilon}{4}\right) \sum \text{mis}(\Lambda_j K_j) \stackrel{(4.78)}{=} \left(1 + \frac{\varepsilon}{4}\right) \sum \text{mis}(\varphi_j \circ \phi(K_j)) \end{aligned} \quad (4.79)$$

Osserviamo che, per ogni j , $\phi(K_j) \subseteq \phi(E) \subseteq B$ è PJ-misurabile; inoltre, φ_j è iniettiva su B , $\varphi_j \in C^1(B)$ (per il teorema della funzione inversa) e $\det \varphi'_j \neq 0$ su B . quindi possiamo applicare il Lemma 4.9 con φ_j e $\phi(K_j)$ al posto di ϕ ed E , ottenendo da (4.79)

$$\int_E |\det \phi'| \leq (1 + \frac{\varepsilon}{4}) \sum \int_{\phi(K_j)} |\det \varphi'_j|. \quad (4.80)$$

Per l'uniforme continuità di $|\det \phi'(y)\phi'(x)^{-1}|$ sul compatto $\phi(E) \times \phi(E)$, segue che esiste $0 < \delta_1 \leq \delta_0$ tale che $|\det \varphi'_j(y)| < 1 + \varepsilon/4$ per ogni j e per ogni $y \in \phi(K_j)$. Quindi, da (4.80), concludiamo che

$$\int_E |\det \phi'| \leq (1 + \frac{\varepsilon}{4})^2 \sum \text{mis}(\phi(K_j)) < (1 + \varepsilon) \sum \text{mis}(\phi(K_j)) = (1 + \varepsilon) \text{mis}(\phi(E)).$$

Dall'arbitrarietà di ε segue la (4.79), dalla quale, ragionando come all fine della dimostrazione del Lemma 4.9, segue che

$$\int_A |\det \phi'| \leq \text{mis}(\phi(A)),$$

relazione che assieme alla (4.79) con $E = A$ mostra la validità di (4.65). ■

Dimostrazione della (4.66) Sia $f = \sum c_j \chi_{R_j} \in S(B)$ e osserviamo che, per ogni j ,

$$\chi_{\phi^{-1}(R_j)} = \chi_{R_j} \circ \phi. \quad (4.81)$$

Dunque, applicando la (4.65) con l'insieme PJ-misurabile $\phi^{-1}(R_j)$ al posto di A , otteniamo

$$\begin{aligned} \int_B f &= \sum c_j \text{mis}(R_j) = \sum c_j \text{mis}(\phi \circ \phi^{-1}(R_j)) \\ &\stackrel{(4.65)}{=} \sum c_j \int_{\phi^{-1}(R_j)} |\det \phi'| \stackrel{(4.81)}{=} \sum c_j \int_A \chi_{R_j} \circ \phi |\det \phi'| \\ &= \int_A f \circ \phi |\det \phi'|. \end{aligned}$$

Quindi la (4.66) vale per funzioni semplici e questo implica che vale anche per funzioni f integrabili secondo Riemann. ■

Applicazioni

(Coordinate polari in \mathbb{R}^2) Sia $B := (0, R) \times (0, 2\pi)$ e $F : B \rightarrow \mathbb{R}^2$ definita da

$$F : (\rho, \theta) \in B \rightarrow x = (\rho \cos \theta, \rho \sin \theta). \quad (4.82)$$

(i) Dimostrare che F e B verificano le ipotesi del Teorema 4.6 e calcolare, in particolare, lo jacobiano $\frac{\partial F}{\partial y}$ ed il suo determinante (qui $y := (\rho, \theta)$).

(ii) Si definisca e si discuta la trasformazione inversa ϕ .

(iii) Qual è l'immagine, secondo F , di un rettangolino $[r, r + s] \times [\theta, \theta + \sigma]$?

(Coordinate polari in \mathbb{R}^3) Sia $B := (0, R) \times (0, 2\pi) \times (0, \pi)$ e $F : B \rightarrow \mathbb{R}^3$ definita da

$$F : (\rho, \theta, \psi) \in B \rightarrow x = (\rho \cos \theta \sin \psi, \rho \sin \theta \sin \psi, \rho \cos \psi). \quad (4.83)$$

(i) Dimostrare che F e B verificano le ipotesi del Teorema 4.6 e calcolare, in particolare, lo jacobiano $\frac{\partial F}{\partial y}$ ed il suo determinante (qui $y := (\rho, \theta, \psi)$).

(ii) Si definisca e si discuta la trasformazione inversa ϕ .

(Coordinate cilindriche in \mathbb{R}^3) Sia $B := (0, R) \times (0, 2\pi) \times (-h, h)$ e $F : B \rightarrow \mathbb{R}^3$ definita da

$$F : (\rho, \theta, z) \in B \rightarrow x = (\rho \cos \theta, \rho \sin \theta, z). \quad (4.84)$$

- (i) Dimostrare che F e B verificano le ipotesi del Teorema 4.6 e calcolare, in particolare, lo jacobiano $\frac{\partial F}{\partial y}$ ed il suo determinante (qui $y := (\rho, \theta, z)$).
- (ii) Si definisca e si discuta la trasformazione inversa ϕ .

(Solidi di rotazione in \mathbb{R}^3) Sia $E_0 \subseteq (0, \infty) \times \mathbb{R}$ un insieme misurabile di \mathbb{R}^2 e α un numero in $(0, 2\pi]$. Si dimostri che

$$E := \{(x, y, z) \in \mathbb{R}^3 : x = \rho \cos \theta, y = \rho \sin \theta, \text{ con } (\rho, z) \in E_0, \text{ e } \theta \in [0, \alpha]\}$$

è un insieme misurabile in \mathbb{R}^3 e se ne calcoli il volume in termini di un integrale doppio su E_0 .

Si calcoli il volume del toro solido ottenuto ruotando un disco di raggio r attorno ad un asse a distanza $R > r$ dal centro del disco.

Si verifichi che se definiamo i numeri L, M, N tramite la relazione

$$\frac{\partial \varphi}{\partial u_1} \times \frac{\partial \varphi}{\partial u_2} := Le^{(1)} - Me^{(2)} + Ne^{(3)} \quad (4.85)$$

ed i numeri E, F, G tramite

$$E := \left| \frac{\partial \varphi}{\partial u_1} \right|, \quad F := \left| \frac{\partial \varphi}{\partial u_2} \right|, \quad G := \frac{\partial \varphi}{\partial u_1} \cdot \frac{\partial \varphi}{\partial u_2}, \quad (4.86)$$

allora

$$\left| \frac{\partial \varphi}{\partial u_1} \times \frac{\partial \varphi}{\partial u_2} \right|^2 = L^2 + M^2 + N^2 = E^2 F^2 - G^2. \quad (4.87)$$

(Tali notazioni sono usate in vari testi).

Dimostrare che se $\mathcal{S} := \{(x, f(x)) : x \in U\}$, con U aperto limitato di \mathbb{R}^2 , è il grafico di una funzione $f \in C^1(\bar{U}, \mathbb{R})$, allora

$$\text{Area}(\mathcal{S}) = \int_U \sqrt{1 + |\nabla f|^2} du. \quad (4.88)$$

(Superfici di rotazione in \mathbb{R}^3) Sia $\Gamma := \{(u(t), v(t)) : t \in (a, b)\}$ un elemento di curva regolare (eventualmente chiusa) in $(0, \infty) \times \mathbb{R}$ e α un numero in $(0, 2\pi]$. Si dimostri che

$$\mathcal{S} := \{(x, y, z) \in \mathbb{R}^3 : x = u(t) \cos \theta, y = u(t) \sin \theta, z = v(t) \text{ con } t \in (a, b), \text{ e } \theta \in (0, \alpha)\}$$

è un elemento di superficie regolare in \mathbb{R}^3 e se ne calcoli l'area in termini di un integrale su (a, b) .

Sia $\Gamma := \varphi((a, b))$ un elemento di curva regolare in \mathbb{R}^2 , sia $f \in C(\bar{\Gamma}, (0, \infty))$ e sia $\mathcal{S} := \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in \Gamma \text{ e } 0 < z < f(x, y)\}$. Si dimostri che \mathcal{S} è un elemento di superficie regolare in \mathbb{R}^3 e si dimostri che

$$\text{Area}(\mathcal{S}) = \int_{\Gamma} f d\sigma_1.$$

Esempio 4.10 (Numeri diofantini) Un numero $\omega \in \mathbb{R}$ si dice **diofantino** se esistono $\gamma > 0$ e $\tau \geq 1$ tali che

$$|\omega q - p| \geq \frac{\gamma}{|q|^\tau}, \quad \forall p, q \in \mathbb{Z} \quad \text{con } q \neq 0. \quad (4.89)$$

Si noti che numeri diofantini sono irrazionali e che la (4.89) dà una stima sulla ‘velocità’ con cui è possibile approssimare un numero diofantino ω con dei numeri razionali (si divida la (4.89) per q).

Sia $\mathcal{D}_{\gamma, \tau}$ l'insieme dei numeri diofantini per cui vale (4.89) e sia \mathcal{D} l'insieme di tutti numeri diofantini con $\tau > 1$ (ossia $\mathcal{D} := \bigcup_{\gamma > 0} \bigcup_{\tau > 1} \mathcal{D}_{\gamma, \tau}$).

Se $\gamma \leq 1$ allora $(\mathcal{D}_{\gamma,\tau}^c \cap [-R, R]) \subseteq \cup_{p,q \neq 0} I_{p,q}$ dove $I_{p,q}$ è l'intervallo aperto di centro $p/q \in [-R-1, R+1]$ di lunghezza $2\gamma/|q|^{\tau+1}$. Dunque

$$\begin{aligned} m(\mathcal{D}_{\gamma,\tau}^c \cap [-R, R]) &\leq m\left(\bigcup_{p,q \neq 0} I_{p,q}\right) \leq \sum_{q \neq 0} \sum_{|p| \leq (R+1)|q|} \frac{2\gamma}{|q|^{\tau+1}} \\ &\leq 2(R+2)\gamma \sum_{q \neq 0} \frac{1}{|q|^\tau} := \gamma c(\tau, R). \end{aligned}$$

Da questo segue che \mathcal{D}^c è un insieme trascurabile.

2 I teoremi classici del calcolo integrale

2.1 Curve e superfici

2.2 Integrali su curve e superfici

Se Γ è un elemento di curva regolare, esiste una parametrizzazione ‘canonica’ in termini della lunghezza. Se $\Gamma = \varphi((a, b))$ definiamo la seguente funzione, chiamata **ascissa curvilinea**:

$$s(t) := \int_a^t |\varphi'(u)| du \quad (4.90)$$

ossia $s(t)$ è la lunghezza dell'elemento di curva di estremi $\varphi(a)$ e $\varphi(t)$: in particolare $s(a) = 0$ e $s(b) = L(\Gamma)$. Essendo $s'(t) = |\varphi'(t)| \neq 0$ tale funzione è strettamente crescente in $[a, b]$ e quindi invertibile. Sia $t(s)$ la funzione inversa. Poniamo allora

$$\tilde{\varphi}(s) := \varphi(t(s)). \quad (4.91)$$

È chiaro che tale inclusione *non dipende da* φ ²⁸. Dato un elemento di curva regolare Γ e fissato un suo estremo x_0 si chiama **parametrizzazione secondo l'ascissa curvilinea** con punto iniziale x_0 la parametrizzazione $\gamma(s) := \tilde{\varphi}(s)$ dove $\varphi : [a, b] \rightarrow \Gamma$ è una qualunque inclusione su Γ tale che $\varphi(a) = x_0$ e $\tilde{\varphi}(s)$ è definita in (4.91).

2.3 1–forme differenziali. Lavoro e circuitazione

In questo paragrafo discuteremo alcune conseguenze del Teorema di Gauss e cioè ‘i teoremi classici di Green e Stokes’; per far ciò, generalizzando il concetto di differenziale di una funzione (scalare) di n variabili, introdurremo le 1–forme differenziali su \mathbb{R}^n . Sia $f \in C^1(A, \mathbb{R})$ con A aperto in \mathbb{R}^n . Sappiamo che, per ogni $x \in A$, il differenziale df è un'applicazione lineare da \mathbb{R}^n in \mathbb{R} che agisce nel seguente modo

$$df(\xi) := df_x(\xi) = \nabla f(x) \cdot \xi, \quad (4.92)$$

per ogni vettore $\xi \in \mathbb{R}^n$; più concisamente²⁹ $df(\xi) = \nabla f \cdot \xi$. In particolare se $f(x) = x_i$ è la funzione lineare che ad x associa la sua i -esima coordinata, si ha

$$dx_i(\xi) = (\nabla x_i) \cdot \xi = e^{(i)} \cdot \xi = \xi_i \quad (4.93)$$

ossia il differenziale della funzione x_i non è altro che la proiezione sulla i -esima coordinata di ξ .

²⁸Cioè se $\psi : [c, d] \rightarrow \mathbb{R}^n$ è un'altra inclusione su Γ tale che $\psi(c) = \varphi(a)$ allora $\tilde{\varphi}(s) = \tilde{\psi}(s)$.

²⁹Non riportando esplicitamente nella notazione la dipendenza *non lineare* da $x \in A$ (cosa che faremo spesso in seguito).

Definizione 4.11 I differenziali dx_i sono esempi di 1-forme differenziali (o in breve ‘1-forme’) e verranno chiamati **1-forme elementari**. Le 1-forme differenziali in A sono combinazioni lineari (aventi funzioni per coefficienti) delle n 1-forme elementari. Più precisamente ω^1 è una **1-forma differenziale su A** se

$$\omega^1 = \sum_{i=1}^n g_i dx_i \quad (4.94)$$

con $g_i \in C(A, \mathbb{R})$; se $g_i \in C^k(A, \mathbb{R})$ diremo che ω^1 è di classe C^k .

Fissato $x \in A$, l’azione di una 1-forma su vettori in $\xi \in \mathbb{R}^n$ è definita (per linearità) nella maniera ovvia:

$$\omega_x^1(\xi) = \sum_{i=1}^n g_i(x) dx_i(\xi) = \sum_{i=1}^n g_i(x) \xi_i. \quad (4.95)$$

Come sopra, normalmente sottintenderemo la dipendenza (*non lineare*) della 1-forma dal punto $x \in A$ e scriveremo semplicemente $\omega^1(\xi)$ al posto di $\omega_x^1(\xi)$. Chiaramente, poichè i differenziali sono funzioni lineari su \mathbb{R}^n , anche le 1-forme lo sono:

$$\omega^1(a_1 \xi^{(1)} + a_2 \xi^{(2)}) = a_1 \omega^1(\xi^{(1)}) + a_2 \omega^1(\xi^{(2)}). \quad (4.96)$$

Quindi una 1-forma è, per ogni $x \in A$ fissato, la più generale applicazione lineare da \mathbb{R}^n in \mathbb{R} ed inoltre la dipendenza da x è regolare. In particolare, se f è una funzione regolare su A il suo differenziale df è una 1-forma su A :

$$df = \sum_{i=1}^n f_{x_i} dx_i. \quad (4.97)$$

Dare una 1-forma in $A \subseteq \mathbb{R}^n$ (in cui penseremo sempre fissata la base standard $e^{(i)}$) equivale dunque ad assegnare n funzioni regolari, o, equivalentemente un ‘campo vettoriale’ $F \in C^1(A, \mathbb{R}^n)$. Data $F \in C^1(A, \mathbb{R}^n)$ chiameremo ω_F^1 la 1-forma i cui coefficienti g_i sono le componenti di F :

$$\omega_F^1 := \sum_{i=1}^n F_i dx_i. \quad (4.98)$$

Si noti che da (4.97) e (4.98) segue che

$$\omega_{\nabla f}^1 = df. \quad (4.99)$$

Osservazione 4.12 (i) Non si confonda il significato degli ‘indici’ x e F in (4.95) e (4.98): usando il simbolismo in (4.95) dovremmo scrivere $\omega_{F,x}^1(\xi) = \sum_{i=1}^n F_i(x) dx_i(\xi)$, ma normalmente ometteremo la x in tali scritte.

(ii) Una 1-forma ω^1 ‘di classe $C^p(A)$ ’, è per definizione una 1-forma in cui i coefficienti g_i sono funzioni $C^p(A)$ (anche nel caso A non sia aperto).

Nel caso $n = 1$ siamo abituati a vedere il simbolo ‘ dx ’ all’interno di un integrale ed ora, infatti, definiremo l’integrale di una 1-forma. Per far questo dobbiamo prima formalizzare il concetto di *orientamento su curve*. Sia $\Gamma \subseteq \mathbb{R}^n$ un elemento di curva e sia φ un’inclusione tale che $\varphi((a, b)) = \Gamma$, definiamo l’**orientamento su Γ** indotto da φ come

$$\text{Or}(\varphi) := \left\{ \psi : \psi \text{ è un'inclusione su } \Gamma : \frac{d}{dt}(\psi^{-1} \circ \varphi) > 0 \right\}. \quad (4.100)$$

Da questa definizione segue che: (i) la relazione $\psi \in \text{Or}(\varphi)$ definisce una *relazione di equivalenza*; (ii) se definiamo l’**inclusione inversa di φ** come

$$\bar{\varphi}(t) := \varphi(-t), \quad t \in [-b, -a] \quad (4.101)$$

allora $\bar{\varphi} \notin \text{Or}(\varphi)$; (iii) se ψ è una inclusione su Γ allora o $\psi \in \text{Or}(\varphi)$ oppure $\psi \in \text{Or}(\bar{\varphi})$. L'orientamento $\text{Or}(\bar{\varphi})$ si chiama **orientamento opposto** (o inverso) di $\text{Or}(\varphi)$ e scriveremo $\text{Or}(\bar{\varphi}) = -\text{Or}(\varphi)$. Diremo anche che $(\Gamma, \text{Or}(\varphi))$ è la *curva Γ orientata nel verso che va da $\varphi(a)$ a $\varphi(b)$* . Allo stesso modo, prendendo un'inclusione continua e regolare a tratti³⁰, si definisce l'*orientamento su curve regolari a tratti*: naturalmente in questo caso la relazione differenziale in (4.100) verrà richiesta nei punti in cui φ è derivabile.

Sia Γ un elemento di curva regolare orientato in A : $\Gamma = \varphi((a, b))$ e $\text{Or} = \text{Or}(\varphi)$ e sia $\omega^1 := \omega_F^1$ una 1-forma in A . Definiamo, allora, l'**integrale di ω^1 sulla curva orientata Γ** come

$$\int_{\Gamma, \text{Or}} \omega^1 := \int_a^b F \circ \varphi(t) \cdot \varphi'(t) dt = \sum_{i=1}^n \int_a^b F_i \circ \varphi \varphi'_i dt, \quad (4.102)$$

Dal punto (vii) dell'Osservazione C.3 segue che tale integrale *non dipende dalla particolare inclusione ma solo da Γ e dall'orientamento scelto*. Si noti che (4.102) può essere riscritta come

$$\int_{\Gamma, \text{Or}} \omega^1 = \int_a^b \omega^1(\varphi') dt \quad (4.103)$$

dove la scrittura completa dell'integrando a destra è $\omega_{F, \varphi(t)}^1(\varphi'(t))$. Come ci si aspetta, cambiando l'orientamento l'integrale in (4.102) cambia segno:

$$\int_{\Gamma, -\text{Or}} \omega^1 = - \int_{\Gamma, \text{Or}} \omega^1 \quad (4.104)$$

Comunque, spesso, quando non vi siano ambiguità, *ometteremo l'indicazione esplicita dell'orientamento nel simbolo di integrale di 1-forme*.

Tale definizione si generalizza in maniera ovvia al caso di curve chiuse o di curve regolari a tratti contenute in A .

Nel caso $n = 1$ possiamo dunque interpretare l'integrale $\int_a^b f dx$ come l'integrale della 1-forma $f dx$ su $\Gamma = (a, b)$ (si ricordi l'esempio (E1) con $n = 1$ della sezione C.1).

2.4 Lemma di Poincarè. Campi conservativi e irrotazionali

Se $\omega^1 = df$ è il differenziale di una funzione $f \in C^1(A)$ e $\Gamma = \varphi((a, b))$ è un elemento di curva orientato in A di estremi $x_0 = \varphi(a)$ ed $x = \varphi(b)$, otteniamo (sottintendendo $\text{Or} = \text{Or}(\varphi)$)

$$\begin{aligned} \int_{\Gamma} \omega^1 &:= \int_{\Gamma} df = \int_a^b \nabla f \cdot \varphi' dt \\ &= \int_a^b \frac{d}{dt} f \circ \varphi dt = f(x) - f(x_0), \end{aligned} \quad (4.105)$$

Integrando sui tratti di regolarità di una curva regolare a tratti³¹, si vede subito che la (4.105) vale anche per curve regolari a tratti. In particolare, se Γ è una curva (regolare a tratti e) chiusa (cioè $x_0 = x$) si ha

$$\int_{\Gamma} df = 0, \quad (\Gamma \text{ curva chiusa regolare a tratti}). \quad (4.106)$$

Chiaramente, *dire che una 1-forma ω^1 ha integrale nullo su ogni curva chiusa è equivalente a dire che l'integrale (4.103) dipende solo dagli estremi di Γ* . Infatti vale anche il viceversa:

³⁰Si ricordi il punto (iii) dell'Osservazione C.5.

³¹I valori nei punti interni dove l'inclusione non è differenziabile si cancellano a due a due.

Proposizione 4.13 *Sia A un aperto connesso di \mathbb{R}^n . Una 1-forma ω^1 su A è il differenziale di una funzione f se e solo se per ogni curva chiusa Γ regolare a tratti si ha*

$$\int_{\Gamma} \omega^1 = 0. \quad (4.107)$$

Dimostrazione Il ‘solo se’ lo abbiamo già dimostrato. Assumiamo ora che valga (4.107) e sia $\omega^1 := \omega_F^1$. Per l’osservazione fatta dopo (4.106), l’integrale di ω^1 dipende solo dagli estremi di una curva Γ . Fissiamo un qualunque punto $x_0 \in A$ e, per ogni $x \in A$, consideriamo una qualunque curva $\Gamma(x) \subseteq A$ regolare a tratti con primo estremo in x_0 e secondo estremo³² in x . Per ipotesi, risulta ben definita la funzione

$$f(x) := \int_{\Gamma(x)} \omega^1. \quad (4.108)$$

Facciamo ora vedere che $f_{x_i} = F_i$: sia $h \in \mathbb{R}$ tale che il segmento $P := P(x, x + h e^{(i)})$ sia interamente contenuto in A . Sia $P = \{z = z(t) := x + t h e^{(i)}, t \in [0, 1]\}$. Allora,

$$\begin{aligned} \frac{1}{h}(f(x + h e^{(i)}) - f(x)) &= \frac{1}{h} \int_P \omega_F^1 \\ &= \frac{1}{h} \int_0^1 F(x + t h e^{(i)}) \cdot z' dt = \int_0^1 F(x + t h e^{(i)}) \cdot e^{(i)} dt \\ &= \int_0^1 F_i(x + t h e^{(i)}) dt \end{aligned} \quad (4.109)$$

e tale quantità tende, quando $h \rightarrow 0$, a $F_i(x)$. ■

Definizione 4.14 *Sia $\omega^1 := \omega_F^1$ una 1-forma su $A \subseteq \mathbb{R}^n$; ω^1 si dice esatta se $F = df$ per qualche $f \in C^2(A)$; tale f viene chiamata ‘primitiva di F ’. ω_F^1 si dice chiusa se*

$$\frac{\partial F_i}{\partial x_j} = \frac{\partial F_j}{\partial x_i}, \quad \forall i, j. \quad (4.110)$$

Chiaramente, per il lemma di Schwarz, una 1-forma esatta è chiusa. In generale non è vero il viceversa. Infatti, sia $A = \mathbb{R}^2 \setminus \{0\}$ e sia

$$\omega^1 = \frac{x dy - y dx}{x^2 + y^2} \quad (4.111)$$

(cosicché $\omega^1 \in C^\infty(A)$). È immediato controllare che ω^1 è chiusa ma d’altra parte è anche facile vedere che se Γ è il cerchio unitario orientato in senso antiorario allora

$$\int_{\Gamma} \omega^1 = 2\pi, \quad (4.112)$$

e quindi, per la Proposizione 4.13 ω^1 non è esatta su³³ A .

Un viceversa parziale si ottiene facendo opportune ipotesi sul dominio A . Un insieme A si dice **stellato** se esiste un punto $x_0 \in A$ tale che per ogni $x \in A$ il segmento di estremi x_0 e x è interamente contenuto in A ; in tal caso si dirà che A è stellato rispetto a³⁴ x_0 . Vale allora la seguente³⁵

³²Questo significa che se $\Gamma = \varphi((a, b))$, e $\text{Or} = \text{Or}(\varphi)$, $x_0 = \varphi(a)$ e $x = \varphi(b)$; si ricordi anche che un aperto connesso è connesso per poligonali (ed una poligonale è una curva regolare a tratti).

³³Si noti però che $\omega^1 = d(\arctan(y/x)) = d(-\arctan(x/y))$ in $\{(x, y) : y \neq 0\}$ il che mostra che ω^1 è esatta su domini della forma $\mathbb{R}^2 \setminus P$ dove P è una qualunque semiretta chiusa che parte dall’origine (esercizio 4.52).

³⁴Per esempio un insieme connesso è stellato rispetto ad un suo qualunque punto.

³⁵Questo risultato e le sue generalizzazioni sono note come ‘Lemma di Poincaré’.

Proposizione 4.15 Sia $A \subseteq \mathbb{R}^n$ un insieme stellato. Una 1-forma su A è chiusa se e solo se è esatta.

Dimostrazione Assumiamo per semplicità che $x_0 = 0$ (caso a cui si può sempre ridurre tramite il cambio di variabili $x = x_0 + y$). Sia $\omega^1 := \omega_F^1$ con F che verifichi (4.110) e si consideri la funzione

$$f(x) := \sum_{i=1}^n x_i \int_0^1 F_i(tx) dt, \quad (4.113)$$

che per le ipotesi su A è ben definita. Dai risultati noti sulla derivazione sotto segno di integrale, da (4.110) e dal Teorema fondamentale del calcolo segue che

$$\begin{aligned} \frac{\partial f}{\partial x_j} &= \int_0^1 F_j(tx) dt + \sum_{i=1}^n x_i \int_0^1 t \frac{\partial F_i}{\partial x_j}(tx) dt \\ &= \int_0^1 F_j(tx) dt + \sum_{i=1}^n x_i \int_0^1 t \frac{\partial F_j}{\partial x_i}(tx) dt \\ &= \int_0^1 \frac{d}{dt} (tF_j(tx)) dt = F_j(x). \end{aligned}$$

Ovvero $F = \nabla f$ e cioè ω^1 è esatta. Il viceversa, come già osservato, è conseguenza immediata del lemma di Schwarz. ■

Osservazione 4.16 Si noti che la formula (4.108) fornisce una ricetta per costruire una primitiva di una 1-forma esatta. Per esempio $\omega^1 := (\sin(xy) + xy \cos(xy)) dx + x^2 \cos(xy) dy$ è esatta su \mathbb{R}^2 (come deriva dal Lemma di Poincarè, essendo ω^1 chiusa su \mathbb{R}^2). Dato (x, y) scegliamo $\Gamma(x, y)$ come la poligonale orientata di vertici (in ordine) $(0, 0)$, $(x, 0)$ e (x, y) : $\Gamma(x, y) = \Gamma_1 \cup \Gamma_2$ con $\Gamma_1 = \{(tx, 0) : 0 \leq t \leq 1\}$ e $\Gamma_2 = \{(x, ty) : 0 \leq t \leq 1\}$. Si ha allora $\int_{\Gamma_1} \omega^1 = 0$ e $\int_{\Gamma_2} \omega^1 = yx^2 \int_0^1 \cos(txy) dt = x \sin(xy)$; quindi $x \sin(xy)$ è una primitiva di ω^1 .

Nella definizione della divergenza di una funzione $f \in C^1(A, \mathbb{R}^n)$ (con $A \subseteq \mathbb{R}^n$) il simbolo $\nabla \cdot f$ può essere interpretato come il prodotto scalare tra l'operatore differenziale vettoriale $(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n})$ e il campo vettoriale f . Nel caso $n = 3$ vi è un altro prodotto notevole, il prodotto vettoriale, e il prodotto vettoriale tra l'operatore differenziale vettoriale $(\frac{\partial}{\partial x_1}, \dots, \frac{\partial}{\partial x_n})$ e il campo vettoriale f definisce un nuovo campo vettoriale che prende il nome di **rotore** di f :

$$\nabla \times F := \text{rot } F := \left(\frac{\partial F_3}{\partial x_2} - \frac{\partial F_2}{\partial x_3}, \frac{\partial F_1}{\partial x_3} - \frac{\partial F_3}{\partial x_1}, \frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right). \quad (4.114)$$

Osservazione 4.17 (i) Si noti che la divergenza è definita per n arbitrario al contrario del rotore che è definito solo per $n = 3$.

(ii) Un campo $F \in C^1(A, \mathbb{R}^3)$ tale che $\nabla \times F = 0$ si chiama **irrotazionale**. Ricordando la definizione di 1-forma chiusa si ha che $F \in C^1(A, \mathbb{R}^3)$ è irrotazionale se e solo se ω_F^1 è chiusa.

Proposizione 4.18 Sia A un aperto di \mathbb{R}^3 . Allora

- (i) $\nabla \times (\nabla f) = 0, \forall f \in C^2(\mathbb{R}^3, \mathbb{R})$;
- (ii) $\nabla \cdot (\nabla \times G) = 0, \forall G \in C^2(A, \mathbb{R}^3)$;

Sia ora $A \subseteq \mathbb{R}^3$ un aperto stellato e $H, F \in C^1(A, \mathbb{R}^3)$. Allora

- (iii) $\nabla \times F = 0 \iff \exists f \in C^2(A, \mathbb{R})$ tale che $F = \nabla f$;
- (iv) $\nabla \cdot H = 0 \iff \exists G \in C^2(A, \mathbb{R}^3)$ tale che $H = \nabla \times G$.

Dimostrazione La (i) e la (ii) seguono immediatamente dalle definizioni date e dal Lemma di Schwarz.

La (iii), in vista del punto (ii) dell'Osservazione 4.17, non è altro che la Proposizione 4.15 nel caso $n = 3$.

(iv): l'implicazione ' \Leftarrow ' è un caso speciale del punto (ii). La dimostrazione dell'implicazione ' \Rightarrow ' è assai simile alla dimostrazione della Proposizione 4.15. Infatti, assumiamo (senza perdita di generalità) che A sia stellato rispetto a $x_0 = 0$ e definiamo la funzione $\tilde{H} : A \rightarrow \mathbb{R}^3$ ponendo

$$\tilde{H}_i := \int_0^1 t H_i(tx) dt.$$

Allora, una verifica diretta mostra che la funzione cercata G è data dal prodotto vettoriale tra \tilde{H} e la trasformazione identica $x \rightarrow x$ ossia da

$$G := (\tilde{H}_2 x_3 - \tilde{H}_3 x_2, \tilde{H}_3 x_1 - \tilde{H}_1 x_3, \tilde{H}_1 x_2 - \tilde{H}_2 x_1).$$

Verifichiamo, a titolo di esempio, che $(\nabla \times G)_1 = H_1$ ossia che $\frac{\partial G_3}{\partial x_2} - \frac{\partial G_2}{\partial x_3} = H_1$: dalle definizioni date segue che

$$\begin{aligned} \frac{\partial G_3}{\partial x_2} &= \frac{\partial \tilde{H}_1}{\partial x_2} x_2 - \frac{\partial \tilde{H}_2}{\partial x_2} x_1 + \tilde{H}_1, & \frac{\partial G_2}{\partial x_3} &= \frac{\partial \tilde{H}_3}{\partial x_3} x_1 - \frac{\partial \tilde{H}_1}{\partial x_3} x_3 - \tilde{H}_1, \\ \frac{\partial \tilde{H}_i}{\partial x_j} &= \int_0^1 t^2 \frac{\partial H_i}{\partial x_j}(tx) dt. \end{aligned}$$

Quindi, poiché $\nabla \cdot H = 0$ (e quindi $\frac{\partial H_2}{\partial x_2} + \frac{\partial H_3}{\partial x_3} = -\frac{\partial H_1}{\partial x_1}$), si ha

$$\begin{aligned} \frac{\partial G_3}{\partial x_2} - \frac{\partial G_2}{\partial x_3} &= \int_0^1 \left\{ t^2 \left(\frac{\partial H_1}{\partial x_2} x_2 + \frac{\partial H_1}{\partial x_3} x_3 - x_1 \left[\frac{\partial H_2}{\partial x_2} + \frac{\partial H_3}{\partial x_3} \right] \right) + 2tH_2 \right\} dt \\ &= \int_0^1 \left\{ t^2 \left(\frac{\partial H_1}{\partial x_2} x_2 + \frac{\partial H_1}{\partial x_3} x_3 + x_1 \frac{\partial H_1}{\partial x_1} \right) + 2tH_1 \right\} dt \\ &= \int_0^1 \frac{d}{dt} \left(t^2 H_1(tx) \right) dt := H_1. \quad \blacksquare \end{aligned}$$

2.5 Il teorema della divergenza

2.6 I teoremi classici di Gauss, Green e Stokes

Teorema 4.19 (Gauss, o 'della divergenza') *Sia $\mathcal{S} \subseteq \mathbb{R}^n$ una n -superficie orientabile con bordo e $F \in C^1(\mathcal{S}, \mathbb{R}^n)$. Allora*

$$\begin{aligned} \int_{\mathcal{S}} \nabla \cdot F dx &= \int_{\partial \mathcal{S}} F \cdot \nu ds, & (n = 2), \\ \int_{\mathcal{S}} \nabla \cdot F dx &= \int_{\partial \mathcal{S}} F \cdot \nu d\sigma, & (n = 3), \end{aligned} \quad (4.115)$$

dove gli integrali a sinistra sono integrali standard di Riemann mentre il vettore ν negli integrali a destra è la normale esterna.

Per formulare i teoremi classici di Green e Stokes abbiamo ancora bisogno di alcune definizioni. Dato un aperto regolare $A \subseteq \mathbb{R}^2$, si chiama *orientamento positivo* su ∂A , (o anche 'orientamento antiorario') l'orientamento Or_+ indotto da una parametrizzazione $\varphi : [a, b] \rightarrow \mathbb{R}^2$ di ∂A tale che

$$\frac{\varphi'(t)}{|\varphi'(t)|} = (-\nu_2, \nu_1), \quad (4.116)$$

dove $\nu := (\nu_1, \nu_2)$ è la normale esterna a A nel punto $x = \varphi(t) \in \partial A$; in tal caso la coppia $(\partial A, \text{Or}_+)$ si denota anche con ∂A^+ . Una **superficie elementare con bordo ed orientabile** in \mathbb{R}^3 è un elemento di 2-varietà regolare $\mathcal{S} \subseteq \mathbb{R}^3$ dato dall'immagine di una inclusione

$\psi \in C^2(A, \mathbb{R}^3)$ con A aperto regolare di \mathbb{R}^2 : $\mathcal{S} := \psi(A)$; il **bordo** di \mathcal{S} , denotato³⁶ $\partial\mathcal{S}$, è la curva regolare chiusa $\psi(\partial A)$. Ad una inclusione (ψ, A) che realizza una superficie elementare con bordo ed orientabile $\mathcal{S} = \psi(A)$ possiamo associare un versore ν^+ normale a³⁷ \mathcal{S} definito, per ogni $x = \psi(u) \in \mathcal{S}$, ($u \in A$), da

$$\nu^+ := \frac{\frac{\partial\psi}{\partial u_1} \times \frac{\partial\psi}{\partial u_2}}{\left| \frac{\partial\psi}{\partial u_1} \times \frac{\partial\psi}{\partial u_2} \right|}, \quad (4.117)$$

ed un orientamento su $\partial\mathcal{S}$ definito da $\text{Or}(\psi \circ \varphi)$ dove φ è tale che $\partial A^+ = (\partial A, \text{Or}(\varphi))$; la coppia $(\nu^+, \partial\mathcal{S}^+)$, con $\partial\mathcal{S}^+ := (\partial\mathcal{S}, \text{Or}(\psi \circ \varphi))$ viene detta **orientata positivamente**. Si hanno allora i seguenti risultati

Teorema 4.20 (Teorema di Green) *Sia $A \subseteq \mathbb{R}^2$ un aperto regolare e sia $F \in C^1(\bar{A}, \mathbb{R}^2)$. Allora*

$$\int_A \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx = \int_{\partial A^+} \omega_F^1. \quad (4.118)$$

Teorema 4.21 (Teorema del rotore di Stokes) *Sia \mathcal{S} una superficie elementare con bordo ed orientabile in \mathbb{R}^3 e $F \in C^1(\bar{\mathcal{S}}, \mathbb{R}^3)$. Allora*

$$\int_{\mathcal{S}} (\nabla \times F) \cdot \nu^+ d\sigma_2 = \int_{\partial\mathcal{S}^+} \omega_F^1, \quad (4.119)$$

dove ν^+ è un versore normale a \mathcal{S} tale che la coppia $(\nu^+, \partial\mathcal{S}^+)$ è orientata positivamente.

Osservazione 4.22 (i) Il Teorema di Green può essere utile per calcolare aree di figure piane tramite integrali di 1-forme. Infatti, se applichiamo la (4.118), rispettivamente, ai campi $F := (0, x_1)$, $F := (-x_2, 0)$ oppure $F := \frac{1}{2}(-x_2, x_1)$, (osservando che in tutti questi casi $\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} = 1$) si ottiene la formula

$$m(A) := \text{Area}(A) = \int_{\partial A^+} x_1 dx_2 = - \int_{\partial A^+} x_2 dx_1 = \frac{1}{2} \int_{\partial A^+} (x_1 dx_2 - x_2 dx_1). \quad (4.120)$$

(ii) L'integrale di una 1-forma ω_F^1 su di una curva chiusa orientata Γ prende il nome di **circuitazione del campo F su Γ** . Dunque il Teorema di Stokes può essere formulato dicendo che *il flusso totale del rotore di un campo F attraverso una superficie elementare \mathcal{S} con bordo orientabile, nel verso della normale ν^+ , coincide con la circuitazione di F su $\partial\mathcal{S}^+$* .

(iii) Non è difficile mostrare che *il Teorema della divergenza ed i teoremi di Green e Stokes continuano a valere nel caso in cui le superfici su cui si integra siano solo regolari a tratti*³⁸. Per esempio il Teorema della divergenza è applicabile a cubi, coni, etc.

Dimostrazione (del Teorema di Green) Applicando il Teorema della divergenza su A al campo vettoriale $(F_2, -F_1)$, se $\varphi : [a, b] \rightarrow \mathbb{R}^2$ è tale che $\varphi([a, b]) = \partial A$ e $\text{Or}_+ = \text{Or}(\varphi)$, si ha

$$\begin{aligned} \int_A \nabla \cdot (F_2, -F_1) dx &= \int_A \left(\frac{\partial F_2}{\partial x_1} - \frac{\partial F_1}{\partial x_2} \right) dx \\ &= \int_{\partial A} (F_2, -F_1) \cdot \nu d\sigma_1 = \int_a^b F \cdot \frac{\varphi'}{|\varphi'|} |\varphi'| dt \\ &= \int_{\partial A^+} \omega_F^1. \quad \blacksquare \end{aligned}$$

³⁶Si noti che, in questo contesto, il simbolo ' $\partial\mathcal{S}$ ' non denota la frontiera insiemistica di \mathcal{S} , che, infatti, coincide con $\bar{\mathcal{S}}$ (essendo $\bar{\mathcal{S}} = \mathcal{S}$).

³⁷'Versore' significa di norma unitaria ($|\nu^+| = 1$) e 'normale a \mathcal{S} ' significa 'ortogonale ad ogni vettore tangente a \mathcal{S} ' ($\nu^+ \cdot \xi = 0$ per ogni ξ tangente a \mathcal{S} nel punto x).

³⁸Bisogna, naturalmente, adattare (nella maniera ovvia) le varie definizioni date. La dimostrazione di questa estensione del Teorema della divergenza può essere fatta approssimando in modo regolare la frontiera ∂A di un dominio A 'regolare a tratti' (usando, per esempio, la convoluzione con funzioni C_0^∞ con supporto in un intorno opportuno delle 'singolarità' di ∂A): se $\partial A^{(j)}$ denotano tali approssimazioni regolari, è allora facile vedere che $\lim_{j \rightarrow \infty} \int_{A^{(j)}} \nabla \cdot F = \int_A \nabla \cdot F$ e che (essendo la parte singolare di $\partial A^{(j)}$ un 'insieme di misura $(n-1)$ -dimensionale nulla') anche $\lim_{j \rightarrow \infty} \int_{\partial A^{(j)}} F \cdot \nu d\sigma_{n-1} = \int_{\partial A} F \cdot \nu d\sigma_{n-1}$.

Dimostrazione (del Teorema di Stokes) Sia $\mathcal{S} = \psi(A)$ con A aperto regolare in \mathbb{R}^2 e sia $\varphi : [a, b] \rightarrow \mathbb{R}^2$ tale che $\varphi([a, b]) = \partial A$ e $\text{Or}_+ = \text{Or}(\varphi)$. Sia

$$G(u) := \left(F \circ \psi \cdot \frac{\partial \psi}{\partial u_1}, F \circ \psi \cdot \frac{\partial \psi}{\partial u_2} \right), \quad \text{con } \psi := \psi(u), \quad u \in A. \quad (4.121)$$

Applicando il Teorema di Green a G , si ottiene

$$\begin{aligned} \int_A \left(\frac{\partial G_2}{\partial u_1} - \frac{\partial G_1}{\partial u_2} \right) du &= \int_{\partial A^+} \omega_G^1 \\ &= \sum_{i=1}^3 \sum_{j=1}^2 \int_a^b F_i(\psi \circ \varphi) \frac{\partial \psi_i}{\partial u_j} \varphi_j'(t) dt \\ &= \int_a^b \sum_{i=1}^3 F_i(\psi \circ \varphi) \left(\frac{\partial \psi_i}{\partial u_1}(\varphi) \varphi_1' + \frac{\partial \psi_i}{\partial u_2}(\varphi) \varphi_2' \right) \\ &= \int_a^b \sum_{i=1}^3 F_i(\psi \circ \varphi) \frac{d}{dt} \psi_i \circ \varphi = \int_{\partial \mathcal{S}^+} \omega_F^1. \end{aligned} \quad (4.122)$$

D'altra parte

$$\int_{\mathcal{S}} (\nabla \times F) \cdot \nu^+ d\sigma_2 = \int_A \nabla \times F(\psi) \cdot \left(\frac{\partial \psi}{\partial u_1} \times \frac{\partial \psi}{\partial u_2} \right) du. \quad (4.123)$$

Ed infine

$$\begin{aligned} \frac{\partial G_2}{\partial u_1} - \frac{\partial G_1}{\partial u_2} &= \sum_{i=1}^3 \frac{\partial}{\partial u_1} \left(F_i \circ \psi \frac{\partial \psi_i}{\partial u_2} \right) - \frac{\partial}{\partial u_2} \left(F_i \circ \psi \frac{\partial \psi_i}{\partial u_1} \right) \\ &= \sum_{i,j} \frac{\partial F_i}{\partial x_j} \left(\frac{\partial \psi_j}{\partial u_1} \frac{\partial \psi_i}{\partial u_2} - \frac{\partial \psi_j}{\partial u_2} \frac{\partial \psi_i}{\partial u_1} \right) \\ &= \sum_{i>j} \left(\frac{\partial F_i}{\partial x_j} - \frac{\partial F_j}{\partial x_i} \right) \left(\frac{\partial \psi_j}{\partial u_1} \frac{\partial \psi_i}{\partial u_2} - \frac{\partial \psi_j}{\partial u_2} \frac{\partial \psi_i}{\partial u_1} \right) \\ &:= (\nabla \times F) \cdot \left(\frac{\partial \psi}{\partial u_1} \times \frac{\partial \psi}{\partial u_2} \right). \end{aligned} \quad (4.124)$$

Mettendo assieme (4.122) ÷ (4.124) si ha l'asserto. \blacksquare

3 Integrale di Lebesgue

3.1 Insiemi trascurabili e Teorema di Vitali–Lebesgue

Definizione 4.23 Un insieme $Q \subseteq \mathbb{R}^n$ si dice *trascurabile* se per ogni $\varepsilon > 0$ esiste una famiglia numerabile di rettangoli aperti $\{R_j\}_{j \in \mathbb{N}}$ tale che³⁹

$$Q \subseteq \bigcup_{j \geq 1} R_j, \quad \text{e} \quad \sum_{j \geq 1} \text{mis}(R_j) \leq \varepsilon. \quad (4.125)$$

Osservazione 4.24 (i) Si noti, che non si è richiesto che Q sia limitato né tantomeno che Q sia misurabile secondo Peano–Jordan. Sia infatti Q un qualunque insieme numerabile e denso⁴⁰ in \mathbb{R}^n ; si può prendere, per esempio, l'insieme $Q = \mathbb{Q}^n$ dei punti $x \in \mathbb{R}^n$ a componenti

³⁹Normalmente, una famiglia (non necessariamente numerabile) di insiemi aperti, la cui unione contenga un insieme Q , si chiama un *ricoprimento (aperto) di Q* ; si ricorda che \mathbb{N} denota l'insieme degli interi positivi $\{1, 2, 3, \dots\}$.

⁴⁰Si ricordi che un insieme $Q \subseteq \mathbb{R}^n$ si dice **denso in \mathbb{R}^n** se la sua chiusura \overline{Q} coincide con tutto \mathbb{R}^n . Analogamente, se C è un insieme chiuso di \mathbb{R}^n , si dice che $Q \subseteq C$ è denso in C se la sua chiusura coincide con C .

razionali. Poiché Q è numerabile $Q = \{r_j\}_{j \geq 1}$. Scegliamo ora R_j come il cubo aperto centrato in r_j di lato $(\varepsilon j^{-2})^{\frac{1}{n}}$. Allora $\text{mis}(R_j) = \varepsilon j^{-2}$ e

$$\sum_{j \geq 1} \text{mis}(R_j) \leq \left(\sum_{j \geq 1} \frac{1}{j^2} \right) \varepsilon, \tag{4.126}$$

il che è chiaramente equivalente a dire che Q è trascurabile. D'altra parte Q non è limitato né, per lo stesso motivo visto sopra⁴¹, è misurabile $Q \cap E$ qualunque sia il rettangolo E .

- (ii) Ovviamente un sottoinsieme di un insieme trascurabile è trascurabile.
- (iii) Un'unione numerabile di insiemi trascurabili è trascurabile.

Dimostrazione Sia $\{Q_j\}_{j \in \mathbb{N}}$ una famiglia numerabile di sottoinsiemi di \mathbb{R}^n trascurabili e sia $\varepsilon > 0$. Allora per ogni $j \geq 1$ esistono rettangoli aperti $R_i^{(j)}$ tali che $Q_j \subseteq \bigcup_{i \in \mathbb{N}} R_i^{(j)}$ e $\sum_{i \geq 1} \text{mis} R_i^{(j)} \leq \varepsilon 2^{-j}$. La famiglia $\{R_i^{(j)}\}_{i,j \in \mathbb{N}}$, come è ben noto, è numerabile ed una numerazione può essere fatta come segue: definiamo $R_1 := R_1^{(1)}$ e, per ogni $r \geq 1$, definiamo⁴²

$$R_{r^2+1} := R_{r+1}^{(1)}, R_{r^2+2} := R_{r^2+1}^{(2)}, \dots, R_{r^2+r+1} := R_{r+1}^{(r+1)}, R_{r^2+r+2} := R_r^{(r+1)}, \dots, R_{r^2+2r+1} := R_1^{(r+1)}.$$

Si noti che secondo tale numerazione, per ogni $N \geq 1$, $\{R_1, \dots, R_{N^2}\} := \{R_j^{(i)} : 1 \leq i, j \leq N\}$. Allora, per ogni $N \geq 1$,

$$\begin{aligned} \sum_{r=1}^N \text{mis} R_r &\leq \sum_{r=1}^{N^2} \text{mis} R_r = \sum_{i,j \leq N} \text{mis} R_j^{(i)} = \sum_{j=1}^N \sum_{i=1}^N \text{mis} R_j^{(i)} \\ &\leq \sum_{j=1}^N \sum_{i=1}^{\infty} \text{mis} R_j^{(i)} \leq \sum_{j=1}^N \frac{\varepsilon}{2^j} \leq \varepsilon. \quad \blacksquare \end{aligned}$$

(iv) A volte è utile, nella definizione di insieme trascurabile, sostituire i rettangoli con cubi; vale infatti la seguente affermazione:

Un insieme $Q \subseteq \mathbb{R}^n$ è trascurabile se e solo se per ogni $\varepsilon > 0$ esiste un ricoprimento numerabile di Q fatto da cubi aperti la somma delle cui misure non eccede ε .

Per dimostrare tale affermazione osserviamo innanzitutto che

dato un qualunque rettangolo limitato $E \subseteq \mathbb{R}^n$ e dato comunque $\sigma > 0$ è possibile ricoprire E con un numero finito di cubi aperti la cui somma non ecceda⁴³ $\text{mis} E + \sigma$. Sia ora $\varepsilon > 0$. Poiché Q è trascurabile esiste un ricoprimento numerabile di rettangoli aperti R_j la somma delle cui misure non eccede $\varepsilon/2$. Per ogni j siano $K_1^{(j)}, \dots, K_{N_j}^{(j)}$ cubi aperti che ricoprono R_j e la somma delle cui misure non ecceda $\text{mis}(R_j) + \varepsilon/2^{j+1}$. Allora l'insieme dei cubi $\{K_i^{(j)}\}$, con $j \geq 1$ e $1 \leq i \leq N_j$, è un insieme numerabile e l'unione di tali cubi ricopre Q ed inoltre

$$\sum_{j \geq 1} \sum_{i=1}^{N_j} \text{mis} (K_i^{(j)}) \leq \sum_{j \geq 1} \text{mis}(R_j) + \varepsilon \sum_{j \geq 1} \frac{1}{2^{j+1}} \leq \varepsilon. \quad \blacksquare$$

La 1.26 si estende facilmente agli insiemi trascurabili non necessariamente misurabili secondo Peano–Jordan. Vale infatti la seguente

⁴¹E cioè che in ogni rettangolo R di una qualunque partizione di E cadono sia punti di Q (essendo Q denso in E), sia punti di $E \setminus Q$ (si ricordi che ogni intervallo di numeri reali è non numerabile).

⁴²La seguente numerazione corrisponde alla numerazione di $\{(i, j) \in \mathbb{N}^2\}$ fatta considerando i quadrati di vertici $(1, 1)$, $(r+1, 1)$, $(r+1, r+1)$ e $(1, r+1)$ e numerando successivamente i lati 'esterni' nel verso che va da $(r+1, 1)$ a $(r+1, r+1)$ e poi da $(r+1, r+1)$ a $(1, r+1)$.

⁴³Sia infatti $E = [a_1, b_1] \times \dots \times [a_n, b_n]$ (i casi in cui E non è chiuso derivano banalmente dal caso in cui E è chiuso); sia $\delta = \sigma / (2nL^{n-1})$ dove $L = \max(b_i - a_i)$; sia $k_i = [(b_i - a_i)\delta^{-1}] + 1$ (dove $[\cdot]$ denota la funzione 'parte intera'); sia, infine, $b'_i = a_i + \delta k_i$. Allora $E' = [a_1, b'_1] \times \dots \times [a_n, b'_n] \supseteq E$ e $\text{mis} E' \leq \text{mis} E + n\delta L^{n-1} \leq \text{mis} E + \sigma/2$. Sia, ora, P la partizione di E' con $P_i = \{a_i + j\delta \text{ con } 0 \leq j \leq k_i\}$. Chiaramente $\mathcal{A}(P')$ è formata da cubi di lato δ che denoteremo \tilde{K}_i per $i = 1, \dots, N = \text{cardinalità di } \mathcal{A}(P')$. Siano $K_i(r)$ i cubi aperti con lo stesso centro di \tilde{K}_i e lato $\delta + r$ cosicché $\tilde{K}_i(0) = \tilde{K}_i$. Ovviamente per ogni $r > 0$ $E \subseteq E' \subseteq \bigcup_i K_i(r)$ e la funzione $f(r) = \sum_{i=1}^N \text{mis} K_i(r)$ è una funzione continua di r e $f(0) = \text{mis} E' \leq \text{mis} E + \sigma/2$. Dunque esiste r_0 tale che per ogni $0 \leq r \leq r_0$ si ha $f(r) \leq \text{mis} E + \sigma$ e l'asserto si ottiene prendendo $K_i = K_i(r)$ per un qualunque $0 < r < r_0$.

Proposizione 4.25 *Sia $Q \subseteq \mathbb{R}^n$ un insieme trascurabile e sia $F : Q \rightarrow \mathbb{R}^n$ una funzione uniformemente lipschitziana su Q . Allora $F(Q)$ è un insieme trascurabile.*

Dimostrazione Per il punto (iv) dell'osservazione precedente esiste un ricoprimento di cubi aperti, $\{K_j\}$, di Q per cui $\sum \text{mis} K_j \leq \varepsilon$ e possiamo assumere che $K_j \cap Q \neq \emptyset$ per ogni j (altrimenti eliminiamo semplicemente K_j). Fissiamo un $x^{(j)} \in K_j \cap Q$ e chiamiamo r_j il lato di K_j cosicché $\text{mis}(K_j) = r_j^n$. Definiamo K'_j il cubo aperto di centro $F(x^{(j)})$ e lato $r'_j := 2Lr_j$, dove L è la costante di Lipschitz di F su Q rispetto la norma $|\cdot|_\infty$ su \mathbb{R}^n , cioè: $|F(x) - F(y)|_\infty \leq L|x - y|_\infty$ per ogni $x, y \in Q$. Allora $F(K_j \cap Q) \subseteq K'_j$, infatti se $x \in K_j \cap Q$, si ha che $|x - x^{(j)}|_\infty \leq r_j$ e quindi $|F(x) - F(x^{(j)})| \leq L|x - x^{(j)}| \leq (Lr_j) = r'_j/2$, ossia $F(x) \in K'_j$. Quindi $\{K'_j\}$ è un ricoprimento di $F(Q)$: se $y \in F(Q)$ allora esiste $x \in Q$ tale che $y = F(x)$, ma, poiché Q è ricoperto dai cubi K_j , esisterà un cubo K_{j_0} che contiene x , ma allora K'_{j_0} contiene y . Inoltre

$$\sum_j \text{mis}(K'_j) = \sum_j (2Lr_j)^n = (2L)^n \sum_j r_j^n = (2L)^n \sum_j \text{mis}(K_j) \leq (2L)^n \varepsilon, \quad (4.127)$$

e dall'arbitrarietà di ε segue l'asserto. \blacksquare

L'insieme delle funzioni integrabili secondo Riemann può essere ora completamente caratterizzato⁴⁴.

Teorema 4.26 (Vitali–Lebesgue) *Una funzione $f : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ (E rettangolo standard di \mathbb{R}^n) è integrabile secondo Riemann su E se e solo se l'insieme dei punti di E in cui f è discontinua è un insieme trascurabile.*

Dimostrazione Chiamiamo $D(f)$ l'insieme dei punti di E in cui f è discontinua ed osserviamo che se $Q_j := \{x \in E : \text{osc}(f, x) \geq 1/j\}$ allora $D(f) = \cup_j Q_j$. Ricordando il punto (iii) dell'Osservazione 4.24, è sufficiente dunque dimostrare:

' f è integrabile su $E \iff Q_j$ è un insieme trascurabile per ogni j '.

Si noti che Q_j è chiuso (e quindi, essendo $Q_j \subseteq E$, è compatto).

Dimostriamo ' \implies ': Siano $j, \varepsilon > 0$ allora esiste una partizione P di E tale che

$$\sum_{R \in \mathcal{R}} \text{osc}(f, R) \text{mis} R \leq \varepsilon/j.$$

Sia $\tilde{\mathcal{R}} := \{R \in \mathcal{R} : R \cap Q_j \neq \emptyset\}$ (cosicché, se $R \in \tilde{\mathcal{R}}$ allora $\text{osc}(f, R) \geq 1/j$). Allora

$$\sum_{R \in \tilde{\mathcal{R}}} \text{mis} R = j \sum_{R \in \tilde{\mathcal{R}}} \frac{\text{mis} R}{j} \leq j \sum_{R \in \tilde{\mathcal{R}}} \text{osc}(f, R) \text{mis} R \leq \varepsilon.$$

Se $N := \#\tilde{\mathcal{R}}$, prendendo dei rettangoli aperti $K_R \supseteq R$ tali che $\text{mis} K_R \leq \text{mis} R + \varepsilon/N$ si ha che $Q_j \subseteq \cup_{R \in \tilde{\mathcal{R}}} K_R$ e $\sum_{R \in \tilde{\mathcal{R}}} \text{mis} K_R \leq 2\varepsilon$. Dunque Q_j è un insieme trascurabile.

Dimostriamo ora ' \impliedby ': Sia $\varepsilon > 0$ e sia $j > 2/\varepsilon$ e sia $Q := Q_j$. Poiché Q è trascurabile esiste un ricoprimento numerabile di Q formato da cubi aperti K_i tali che $\sum \text{mis} K_i \leq \varepsilon$. Essendo Q compatto esistono K_{i_1}, \dots, K_{i_s} che ricoprono Q (e ovviamente $\sum_{j=1}^s \text{mis} K_{i_j} \leq \varepsilon$). Sia D la chiusura di $(E \setminus \cup_{j=1}^s K_{i_j})$. Per ogni $x \in D$ si ha $\text{osc}(f, x) \leq 1/j < \varepsilon/2$. Per ogni $x \in D$ sia $C(x)$ un cubo chiuso di centro x tale che $\text{osc}(f, x) \leq \varepsilon$ (l'esistenza di tali cubi deriva dalla definizione di oscillazione e dal fatto che $\text{osc}(f, x) < \varepsilon/2$). Chiaramente $\cup_{x \in D} \mathring{C}(x) \supseteq D$ e quindi esistono C_1, \dots, C_N , con $C_i := C(x^{(i)})$ per qualche $x^{(i)} \in D$, tali che $D \subseteq \mathring{C}_1 \cup \dots \cup \mathring{C}_N$. Sia P la partizione di E tale che P_i contenga la i -esima coordinata di tutti i vertici dei

⁴⁴In effetti tale descrizione è stata fatta alquanto dopo i contributi di Riemann ed è essenzialmente basata sulla moderna teoria dell'integrazione dovuta a Lebesgue, Vitali, Caratheodory etc.

K_{i_j} e degli C_i . Chiaramente se chiamiamo $\mathcal{R}_1 := \{R \in \mathcal{R} : R \subseteq C_i \text{ per qualche } i\}$ e $\mathcal{R}_2 := \{R \in \mathcal{R} : R \subseteq K_{i_j} \text{ per qualche } j\}$, si ha $\mathcal{R} = \mathcal{R}_1 \cup \mathcal{R}_2$. Allora $\sum_{R \in \mathcal{R}_2} \text{mis } R \leq \varepsilon$ e

$$\begin{aligned} \sum_{R \in \mathcal{R}} \text{osc}(f, R) \text{mis } R &\leq \sum_{R \in \mathcal{R}_1} \text{osc}(f, R) \text{mis } R + \sum_{R \in \mathcal{R}_2} \text{osc}(f, R) \text{mis } R \\ &\leq \varepsilon \text{mis } E + 2 \sup_E |f| \varepsilon. \quad \blacksquare \end{aligned}$$

Una conseguenza immediata di questo risultato, di 1.14 e del fatto che l'insieme di discontinuità di χ_B coincide con la frontiera di B è la seguente

Proposizione 4.27 *Un insieme limitato di \mathbb{R}^n è misurabile secondo Peano–Jordan se e solo se la sua frontiera è un insieme trascurabile.*

Concludiamo questo capitolo con un esempio di una funzione sul cubo unitario $K := [0, 1]^n \subseteq \mathbb{R}^n$ il cui insieme di discontinuità coincide con $K \cap \mathbb{Q}^n$. Sia $\{r_j\}_{j \in \mathbb{N}} = K \cap \mathbb{Q}^n$ e sia $f : K \rightarrow [0, 1]$ la funzione che vale zero se $x \in K \setminus \mathbb{Q}^n$ e $f(r_j) = 1/j$. Chiaramente $\text{osc}(f, x) = 1/j > 0$. Facciamo vedere che f è continua in ogni $x \in K \setminus \mathbb{Q}^n$ ($:= \{x \in K : f(x) = 0\}$). Sia $\varepsilon > 0$; sia $j_0 > 1/\varepsilon$ e sia $\delta := \max_{1 \leq j \leq j_0} |x - r_j|$ (poiché $x \notin \mathbb{Q}^n$, $\delta > 0$). Se $x' \in B_\delta(x)$ o $x' \notin \mathbb{Q}^n$, nel qual caso $|f(x') - f(x)| = 0$, oppure $x' \in \mathbb{Q}^n$, nel qual caso $x = r_j$ per qualche $j \geq j_0$, e quindi $|f(x') - f(x)| = 1/j \leq 1/j_0 < \varepsilon$.

3.2 Funzioni \mathcal{L}^1

In questa sezione ‘costruiremo’, tramite opportuni limiti di funzioni a scalini in $S(E)$, la classe delle funzioni integrabili secondo Lebesgue che denoteremo con \mathcal{L}^1 (o più precisamente con $\mathcal{L}^1(E)$ indicando esplicitamente l'insieme di riferimento su cui si considera l'integrazione). Tale classe estende l'insieme delle funzioni integrabili secondo Riemann e soprattutto permette di introdurre uno spazio funzionale completo (uno spazio di Banach che denoteremo con L^1) rispetto alla norma data dall'integrale del modulo.

Definizione 4.28 *Diremo che una proprietà \mathcal{P} vale ‘quasi ovunque in $A \subseteq \mathbb{R}^n$ ’ (in breve: q.o. in A) o ‘per quasi tutti gli $x \in A$ ’ se esiste un insieme $Q \subseteq A$ trascurabile tale che \mathcal{P} vale per ogni $x \in A \setminus Q$; l'insieme Q viene detto ‘eccezionale’ (rispetto alla proprietà \mathcal{P}).*

Ad esempio⁴⁵ $1 - \chi_{\mathbb{Q}^n}$ è una funzione q.o. strettamente positiva e l'insieme eccezionale è dato da \mathbb{Q}^n .

Il seguente risultato è alla base della costruzione delle funzioni integrabili secondo Lebesgue.

Fissiamo un rettangolo di riferimento non degenere $E \subseteq \mathbb{R}^n$ (ad esempio $E = \mathbb{R}^n$).

Proposizione 4.29 *Sia $\{f_k\}_{k \geq 1}$ una successione monotona non decrescente⁴⁶ di funzioni in $S(E)$ e tale che $\sup_{k \geq 1} \int_E f_k < \infty$. Allora f_k converge quasi ovunque in⁴⁷ E .*

Dimostrazione Innanzitutto, sostituendo eventualmente la successione $\{f_k\}$ con la successione $\{f_k - f_1\}$ possiamo assumere che $f_k \geq 0$ per ogni $k \geq 1$. Sia $Q := \{x \in E : \sup_k f_k(x) = +\infty\}$: la tesi è equivalente a dimostrare che Q è un insieme trascurabile. Denotiamo le funzioni a scalini f_k con

$$f_k := \sum_{i=1}^{N_k} c_i^{(k)} \chi_{R_i^{(k)}}, \quad (4.128)$$

⁴⁵A volte la funzione caratteristica dei vettori a componenti irrazionali $1 - \chi_{\mathbb{Q}^n}$ viene chiamata ‘funzione di Dirichlet’.

⁴⁶Ossia $f_{k+1}(x) \geq f_k(x)$ per ogni $x \in E$.

⁴⁷Ossia esiste un insieme $Q \subseteq E$ trascurabile tale che la successione numerica $\{f_k(x)\}$ converge per ogni $x \in E \setminus Q$.

dove $R_i^{(k)} \subseteq E$ sono rettangoli limitati a due a due disgiunti. Sia $Q' := \bigcup_{i,k} \partial R_i^{(k)}$. Poiché Q' è un insieme trascurabile (essendo unione numerabile di facce di rettangoli ossia di rettangoli degeneri), basterà dimostrare che $\tilde{Q} := Q \setminus Q'$ è un insieme trascurabile. Sia $M := \sup_k \int_E f_k$. Fissato $\varepsilon > 0$, sia $E_k := \{x \in E \setminus Q' : f_k(x) \geq M/\varepsilon\}$. Chiaramente E_k è unione finita di rettangoli aperti (e limitati) e, se chiamiamo \mathcal{R}_k l'insieme di tali rettangoli, si ha $\tilde{Q} \subseteq \bigcup_k E_k = \bigcup_k \bigcup_{R \in \mathcal{R}_k} R$. La funzione f_k è costante su ogni $R \in \mathcal{R}_k$ e, se chiamiamo $c_R^{(k)}$ tale valore costante, dalle definizioni date segue che

$$\frac{M}{\varepsilon} \sum_{R \in \mathcal{R}_k} \text{mis } R \leq \sum_{R \in \mathcal{R}_k} c_R^{(k)} \text{mis } R \leq \int_E f_k \leq M. \quad (4.129)$$

Essendo $\{f_k\}$ non decrescente, si ha che $E_k := \bigcup_{R \in \mathcal{R}_k} R \subseteq \bigcup_{R \in \mathcal{R}_{k+1}} R := E_{k+1}$. Applicando i risultati dimostrati nella sezione 1.2, segue che $A_{k+1} := E_{k+1} \setminus E_k$ è unione disgiunta di rettangoli limitati: chiamiamo \mathcal{R}'_{k+1} l'insieme di tali rettangoli. Dunque (4.129) implica, per ogni $k \geq 1$,

$$\sum_{R \in \mathcal{R}_1} \text{mis } R + \sum_{j=2}^k \sum_{R \in \mathcal{R}'_j} \text{mis } R \leq \sum_{R \in \mathcal{R}_k} \text{mis } R \leq \varepsilon. \quad (4.130)$$

In conclusione $\{\mathcal{R}_1, \mathcal{R}'_2, \mathcal{R}'_3, \dots\}$ è una collezione numerabile di rettangoli aperti la cui unione contiene \tilde{Q} e la somma delle cui misure [per (4.130)] non eccede ε . Dunque \tilde{Q} è un insieme trascurabile. ■

Proposizione 4.30 *Siano f, g in $S(E)$, allora valgono le seguenti:*

- (i) *Per ogni $a, b \in \mathbb{R}$, $af + bg, f, g, \sup f, g$ sono elementi di $S(E)$ e vale la (4.23).*
- (ii) *Se $f \geq 0$ allora $\int_E f \geq 0$.*
- (iv) *Se f è in $S(E)$, allora $|f|$ è in $S(E)$ e vale la (4.26).*

Introduciamo ora una classe di funzioni ‘intermedia’, definita da limiti q.o. di funzioni a scalini soddisfacenti le ipotesi della Proposizione 4.29, e su tale classe definiamo l'integrale di Lebesgue:

Definizione 4.31 $\mathcal{F}(E)$ è la classe di funzioni f per le quali esiste una successione monotona non decrescente di funzioni in $S(E)$ tali che $\sup_k \int_E f_k < \infty$ e $f(x) = \lim_{k \rightarrow \infty} f_k(x)$ per x quasi ovunque in E ; per tali funzioni f definiamo integrale di Lebesgue su E la quantità

$$\int_E f := \lim_{k \rightarrow \infty} \int_E f_k = \sup_k \int_E f_k. \quad (4.131)$$

Affinché tale definizione sia ben posta bisogna verificare che il numero definito in (4.131) non dipenda dalla particolare successione $\{f_k\}$ ma soltanto dal suo limite (limite che è definito quasi ovunque ossia a meno di un insieme trascurabile). Tale verifica segue immediatamente dal seguente

Lemma 4.32 *Siano $\{f_k\}$ e $\{g_k\}$ successioni non decrescenti di funzioni in $S(E)$ convergenti q.o., rispettivamente, a due funzioni⁴⁸ f e g . Se $\sup_k \int_E g_k < \infty$ e $f \leq g$ q.o. allora*

$$\lim_{k \rightarrow \infty} \int_E f_k \leq \lim_{k \rightarrow \infty} \int_E g_k. \quad (4.132)$$

⁴⁸Ossia esiste $Q \subseteq E$ trascurabile tale che $\lim_{k \rightarrow \infty} f_k(x) = f$ e $\lim_{k \rightarrow \infty} g_k(x) = g$ per ogni $x \in E \setminus Q$.

Assumendo la validità di tale Lemma, se⁴⁹ $f \in \mathcal{F}$ e $S \ni f_k \uparrow f$ e $S \ni g_k \uparrow f$ (con $\sup_k \int_E f_k$ e $\sup_k \int_E g_k$ finiti) possiamo applicare il Lemma (con $f = g$ q.o.) scambiando il ruolo di f_k e g_k , per cui

$$\lim_{k \rightarrow \infty} \int_E f_k \leq \lim_{k \rightarrow \infty} \int_E g_k \leq \lim_{k \rightarrow \infty} \int_E f_k,$$

e da tale relazione segue l'uguaglianza dei limiti che definiscono l'integrale di f . ■

Il Lemma 4.32 è a sua volta una semplice conseguenza del seguente

Lemma 4.33 *Sia $\{f_k\}$ una successione non crescente di funzioni non negative in $S(E)$. Allora*

$$f_k \downarrow 0 \quad \text{q.o. in } E \iff \lim_{k \rightarrow \infty} \int_E f_k = 0. \quad (4.133)$$

Dimostrazione (del Lemma 4.33) Cominciamo con ' \implies ': Siano f_k , date come in (4.128), le funzioni a scalini tali che $f_k \downarrow 0$ q.o. in E . Sia $E_0 := \bigcup_{i=1}^{N_1} R_i^{(1)}$. Poiché E_0 è limitato esiste $d > 0$ tale che $E_0 \subseteq [-d/2, d/2]^n$. Inoltre tutti i rettangoli $R_i^{(k)}$ su cui le f_k possono essere strettamente positive sono contenuti in E_0 . Osserviamo anche che se $M := \sup_{\{1 \leq i \leq N_1\}} c_i^{(1)} = \sup f_1$ allora $f_k \leq M$ per ogni $k \geq 1$. Sia Q_0 l'insieme trascurabile su cui f_k non converge a 0 e sia $Q := Q_0 \cup \bigcup_{i,k} \partial R_i^{(k)}$, che, per il punto (ii) dell'Osservazione 4.24, è anch'esso un insieme trascurabile. Fissiamo $\varepsilon > 0$. Essendo Q trascurabile esiste una collezione \mathcal{R}_1 di cubi aperti la cui unione ricopre Q e la somma delle cui misure non eccede ε . Per ogni $x \in E \setminus Q$, $f_k(x) \downarrow 0$, quindi esiste un intero $k(x)$ tale che $f_{k(x)}(x) \leq \varepsilon$. Sia $R(x)$ il più grande rettangolo aperto contenente x su cui $y \rightarrow f_{k(x)}(y)$ sia costante e sia \mathcal{R}_2 l'insieme dei rettangoli $R(x)$ al variare di $x \in E \setminus Q$. Chiaramente $\bigcup_{R \in \mathcal{R}_1 \cup \mathcal{R}_2} R$ è un ricoprimento aperto di E_0 e, per compattezza, esistono N rettangoli $R_i \in \mathcal{R}_1 \cup \mathcal{R}_2$ che ricoprono E_0 . Cambiando eventualmente nome a tali rettangoli possiamo assumere che per $1 \leq i \leq j$, $R_i \in \mathcal{R}_1$ mentre per $j+1 \leq i \leq N$, $R_i \in \mathcal{R}_2$. Si osservi che $\sum_{i=1}^j \text{mis } R_i \leq \varepsilon$, mentre i rettangoli R_i per $i > j$ sono della forma $R_i = R(x^{(i)})$ per un opportuno $x^{(i)} \in E_0 \setminus Q$. Sia $k_0 := \sup_{\{j+1 \leq i \leq N\}} k(x^{(i)})$ e si noti che $f_{k_0}(x) \leq \varepsilon$ per ogni $x \in R_{j+1} \cup \dots \cup R_N$. In conclusione, per ogni $k \geq k_0$ si ha

$$\begin{aligned} \int_E f_k &\leq \int_{E_0} f_{k_0} \leq \int_{R_1 \cup \dots \cup R_j} f_{k_0} + \int_{R_{j+1} \cup \dots \cup R_N} f_{k_0} \\ &\leq M \sum_{i=1}^j \text{mis } R_i + \varepsilon \int_{E_0} 1 \leq \varepsilon (M + d^n). \end{aligned}$$

La dimostrazione di ' \impliedby ' è molto simile alla dimostrazione della Proposizione 4.29: se Q' denota l'insieme di misura nulla $\bigcup_{i,k} \partial R_i^{(k)}$ basta dimostrare che l'insieme $\{x \in E \setminus Q' : \limsup f_k(x) > 0\}$ è trascurabile. Tale insieme coincide con $\bigcup_j Q_j$ dove $Q_j := \{x \in E \setminus Q' : \limsup f_k(x) \geq 1/j\}$ e dunque basterà mostrare che Q_j è trascurabile per ogni $j \geq 1$. Dato $\varepsilon > 0$ sia \bar{k} tale che $\int f_k \leq \varepsilon/j$ per ogni $k \geq \bar{k}$. Chiaramente

$$Q_j \subseteq \{x \in E \setminus Q' : \exists k \geq \bar{k} \text{ con } f_k(x) \geq 1/(2j)\} = \bigcup_{k \geq \bar{k}} \bigcup_{R \in \mathcal{R}_k} R,$$

dove \mathcal{R}_k è una collezione finita di rettangoli aperti e disgiunti contenuti in E e su cui f_k assume un valore costante non inferiore a $1/(2j)$. Quindi per ogni $k \geq \bar{k}$

$$\frac{1}{2j} \sum_{R \in \mathcal{R}_k} \text{mis } R \leq \int_E f_k \leq \frac{\varepsilon}{j} \implies \sum_{R \in \mathcal{R}_k} \text{mis } R \leq 2\varepsilon. \quad (4.134)$$

⁴⁹Qualora non vi sia ambiguità ometteremo l'indicazione esplicita dell'insieme E su cui si sta integrando nella notazione di $\mathcal{F}(E)$ e di altre classi funzionali che considereremo in seguito. D'ora in poi useremo anche le seguenti **notazioni standard**: ' $f_k \uparrow f$ ' denota una successione monotona non decrescente di funzioni che converge alla funzione f ossia $f_k(x) \leq f_{k+1}(x) \forall k \geq 1$ e $\lim_{k \rightarrow \infty} f_k(x) = f(x)$; ' $f_k \uparrow f$ q.o. in E ' significa, come già detto, che la successione monotona non decrescente $\{f_k(x)\}$ converge per $x \in E \setminus Q$ per un opportuno insieme trascurabile Q ; analogamente $f_k \downarrow f$ denota una successione monotona non crescente di funzioni che convergono alla funzione f .

Essendo $\{f_k\}$ non crescente, si ha che $E_k := \bigcup_{R \in \mathcal{R}_k} R \supseteq \bigcup_{R \in \mathcal{R}_{k+1}} R := E_{k+1}$. Applicando i risultati dimostrati nella sezione 1.2, segue che $A_{k+1} := E_k \setminus E_{k+1}$ è unione disgiunta di rettangoli limitati: chiamiamo \mathcal{R}'_{k+1} l'insieme di tali rettangoli. Dunque (4.134) implica, per ogni $k \geq 1$,

$$\sum_{R \in \mathcal{R}_1} \text{mis } R + \sum_{j=2}^k \sum_{R \in \mathcal{R}'_j} \text{mis } R \leq \sum_{R \in \mathcal{R}_k} \text{mis } R \leq 2\varepsilon. \quad (4.135)$$

In conclusione $\{\mathcal{R}_1, \mathcal{R}'_2, \mathcal{R}'_3, \dots\}$ è una collezione numerabile di rettangoli aperti la cui unione contiene Q_j e la somma delle cui misure [per (4.135)] non eccede ε . Dunque Q_j è un insieme trascurabile. ■

Dimostrazione (del Lemma 4.32) Per ogni j fissato definiamo la successione di funzioni a scalini $h_k := \sup\{f_j - g_k, 0\}$. Poiché (q.o.) $g \geq f \geq f_j$, si ha che $h_k \downarrow 0$ q.o. e quindi, per il Lemma 4.33, $\lim \int_E h_k = 0$. Da ciò segue

$$0 = \lim_{k \rightarrow \infty} \int_E h_k \geq \lim_{k \rightarrow \infty} \int_E (f_j - g_k) = \int_E f_j - \lim_{k \rightarrow \infty} \int_E g_k,$$

e cioè, per ogni j , $\int f_j \leq \lim \int g_k$. Prendendo il limite per j che tende ad infinito in tale relazione si ha l'asserto. ■

Osservazione 4.34 (i) Dal Lemma 4.32 segue che se $f_k \uparrow f$ (q.o.) con $f_k \in S(E)$ e se $\sup_k \int f_k = +\infty$ allora $f \notin \mathcal{F}$.

Infatti se per assurdo fosse $f \in \mathcal{F}$ allora per definizione esisterebbe una successione di funzioni a scalini $g_k \uparrow f$ q.o. con $\sup_k \int g_k < +\infty$ ma allora dal Lemma 4.32 (con $g = f$) seguirebbe che $\lim \int f_k \leq \lim \int g_k < \infty$. ■

(ii) Dalla definizione di \mathcal{F} e dalla Proposizione 4.30 segue immediatamente che se $f, g \in \mathcal{F}$ e se a è un numero positivo allora le funzioni af , $f + g$, $\sup\{f, g\}$ appartengono anch'esse alla classe \mathcal{F} ed inoltre

$$\int_E (af) = a \int_E f, \quad (a \geq 0); \quad \int_E (f + g) = \int_E f + \int_E g. \quad (4.136)$$

[Ad esempio se $f_k \uparrow f$ e $g_k \uparrow g$ (con $f_k, g_k \in S$) allora $h_k := \sup\{f_k, g_k\} \in S$ e $h_k \uparrow \sup\{f, g\}$.]

(iii) Se $f \in \mathcal{F}$ e $f \geq 0$ allora esiste una successione di funzioni a scalini $f_k \uparrow f$ con $f_k \geq 0$.

[Infatti se $S \ni g_k \uparrow f$ è una successione che definisce f basta prendere $f_k = \sup\{g_k, 0\}$.]

(iv) Le funzioni integrabili secondo Riemann su E (rettangolo limitato di \mathbb{R}^n) coincidono q.o. con le funzioni f tali che sia f che $-f$ appartengono alla classe $\mathcal{F}(E)$ o più precisamente se $f \in \mathcal{R}(E)$ allora f e $-f \in \mathcal{F}(E)$; viceversa se f e $-f \in \mathcal{F}(E)$ allora f coincide q.o. con una funzione $g \in \mathcal{R}(E)$. Inoltre, sulle funzioni integrabili secondo Riemann, i valori degli integrali secondo Riemann e secondo Lebesgue coincidono.

Dimostrazione Se $f : E \rightarrow \mathbb{R}$ è integrabile secondo Riemann, dalla (4.10) segue che, per ogni intero positivo k , esistono due funzioni $g_k, h_k \in S(E)$ tali che $g_k \leq f \leq h_k$ in E e $\int_E (h_k - g_k) \leq 1/k$. Senza perdita di generalità, possiamo assumere che $\{g_k\}$ sia monotona non decrescente e $\{h_k\}$ sia monotona non crescente⁵⁰. Per la Proposizione 4.29, poiché, per ogni k , $\int g_k \leq \int h_1 < \infty$ (essendo $g_k \leq f \leq h_1$), esiste un insieme trascurabile Q_1 ed una funzione $g \in \mathcal{F}$ tale che $g_k \uparrow g$ in $E \setminus Q_1$ e tale che l'integrale secondo Lebesgue di g è dato da $\lim \int g_k$; ma per come era stata scelta la successione $\{g_k\}$ tale integrale coincide anche con quello di Riemann. D'altra parte, per il Lemma 4.33 applicato alla successione non crescente di funzioni a scalini non negative $(h_k - g_k)$ i cui integrali tendono a zero, si ha che esiste un insieme trascurabile Q_2 tale che $(h_k - g_k) \rightarrow 0$ in $E \setminus Q_2$. Se Q denota l'insieme trascurabile $Q_1 \cup Q_2$,

⁵⁰Basta infatti sostituire eventualmente g_k e h_k con le successioni \tilde{g}_k e \tilde{h}_k definite da: $\tilde{g}_1 = g_1$, $\tilde{g}_k = \sup\{g_k, \tilde{g}_{k-1}\}$ e $\tilde{h}_1 = h_1$, $\tilde{h}_k = \inf\{h_k, \tilde{h}_{k-1}\}$.

per $x \in E \setminus Q$, i numeri $g(x)$ e $f(x)$ appartengono all'intervallo $[g_k(x), h_k(x)]$ la cui ampiezza tende a 0 per $k \rightarrow \infty$. Dunque, mandando k ad infinito, si deduce che $f(x) = g(x)$ per ogni $x \in E \setminus Q$ e quindi che $f = g$ q.o. in E , il che significa che $f \in \mathcal{F}$. D'altra parte poiché anche $-f$ è integrabile secondo Riemann si ha anche che $-f \in \mathcal{F}$. Viceversa se f è una funzione tale che $f, -f \in \mathcal{F}$ allora esistono f_k, g_k a scalini ed un insieme Q trascurabile tale che $f_k(x) \uparrow f(x)$ e $g_k(x) \downarrow f(x)$ per ogni $x \in E \setminus Q$. Si noti che poiché $g_k - f_k \downarrow 0$ q.o., per il Lemma 4.33, $\int (g_k - f_k) \rightarrow 0$. Poniamo $\tilde{f}_k := f_k$ su $E \setminus Q$, $\tilde{f}_k = g_k$ su Q ; $\tilde{f} = f$ su $E \setminus Q$ e $\tilde{f} = \sup_k \tilde{f}_k$ su Q ; chiaramente $\tilde{f} = f$ q.o. Inoltre $\tilde{f}_k \leq \tilde{f} \leq g_k$ su E e $\int (g_k - \tilde{f}_k) = \int (g_k - f_k) \rightarrow 0$ ossia $\tilde{f} \in \mathcal{R}(E)$. ■

(v) Daremo ora un **esempio** di una funzione $f \in \mathcal{F}$ che non è integrabile secondo Riemann anche se modificata su di un insieme trascurabile⁵¹. Si noti che (per il punto precedente) $-f \notin \mathcal{F}$ e dunque \mathcal{F} non è uno spazio vettoriale.

Siano $E := [-1, 1]^n$ e $E_0 := [-1/2, 1/2]^n$. Sia $0 < \gamma < 1$ e sia $\{r_j\}$ una numerazione dei vettori in \mathbb{R}^n a componenti razionali in E_0 : $\{r_j : j \geq 1\} = \mathbb{Q}^n \cap E_0$. Sia K_j il cubo aperto centrato in r_j e di lato $\gamma 2^{-j}$ e sia $A := \cup_j K_j \subseteq E$. Poniamo $f := \chi_A :=$ funzione caratteristica di A .

Dimostriamo che $f \in \mathcal{F}(E)$: sia $A_k := \cup_{j=1}^k K_j$ e $f_k := \chi_{A_k}$. Chiaramente $A = \cup_k A_k$, $f_k \in S(E)$, $f_k \uparrow f$ in E ed inoltre

$$\int_E f_k \leq \sum_{j=1}^k \text{mis } K_j = \sum_{j=1}^k \left(\frac{\gamma}{2^j}\right)^n \leq \frac{\gamma^n}{2^n - 1}. \tag{4.137}$$

Dunque $f \in \mathcal{F}(E)$.

Dimostriamo che f non è integrabile secondo Riemann su E : se $g \in S(E)$ e $g \geq f$ su E allora $g \geq 1$ su E_0 (poiché $g \geq 1$ su $\mathbb{Q}^n \cap E_0$). In particolare se $S \ni g \geq f$ allora $\int g \geq 1$. D'altra parte se $S \ni h \leq f$, poiché $\gamma < 1 \leq (2^n - 1)^{1/n}$, da (4.137) segue che

$$\int_E h \leq \int_E f \leq \frac{\gamma^n}{2^n - 1} < 1.$$

Dunque $\sup_{\{h \leq f, h \in S\}} \int h \leq \gamma^n / (2^n - 1) < 1 \leq \inf_{\{g \geq f, g \in S\}} \int g$ il che significa che f non è integrabile secondo Riemann. ■

(vi) \mathcal{F} non è nemmeno un'algebra: in generale, $f \in \mathcal{F}$ non implica che $f^2 \in \mathcal{F}$ come segue dal seguente **esempio**:

Consideriamo la funzione $f(x) := 1/\sqrt{x}$ su $E := (0, 1]$. Definiamo, per ogni intero k positivo, la seguente funzione a scalini: dividiamo $(0, 1]$ in 2^k intervalli $I_i^{(k)} := ((i-1)/2^k, i/2^k]$ ($i = 1, \dots, 2^k$) e su $I_i^{(k)}$ definiamo il valore di f_k come $\inf_{I_i^{(k)}} f$. Chiaramente $f_k \uparrow f$ in E . Inoltre, per ogni $k \geq 1$

$$\int_E f_k \leq \int_{I_1^{(k)}} f_k + \int_{\frac{1}{2^k}}^1 \frac{1}{\sqrt{x}} dx \leq 3.$$

Dunque $f \in \mathcal{F}(E)$. D'altra parte se $g := f^2 = 1/x$ e se g_k sono funzioni a scalini su E definite in maniera analoga a quanto fatto sopra (cioè il valore di g_k su $I_i^{(k)}$ è dato dallo $\inf_{I_i^{(k)}} g$) si ha di nuovo che $g_k \uparrow g$ su E , ma

$$\int_E g_k \geq \int_{\frac{1}{2^{k-1}}}^1 \frac{1}{x} dx = (k-1) \log 2.$$

Quindi il $\sup_k \int_E g_k = \infty$ e, per il punto (i) di questa Osservazione, $g = f^2 \notin \mathcal{F}(E)$.

Siamo pronti a definire le funzioni integrabili secondo Lebesgue che non sono altro che differenze q.o. di funzioni in \mathcal{F} :

Definizione 4.35 $\mathcal{L}^1(E)$ è la classe di funzioni $f : E \rightarrow \mathbb{R}$ per le quali esistono due funzioni $f_1, f_2 \in \mathcal{F}(E)$ tali che $f = f_1 - f_2$ q.o. in E . Per tali funzioni f definiamo l'integrale di Lebesgue su E come

$$\int_E f := \int_E f_1 - \int_E f_2. \tag{4.138}$$

⁵¹La funzione di Dirichlet $f = 1 - \chi_{\mathbb{Q}}$ su $[0, 1]$ non è integrabile secondo Riemann su $[0, 1]$ ma, d'altra parte, $f = 1$ q.o. in $[0, 1]$.

Di nuovo dobbiamo controllare, affinché la definizione sia ben posta, che il numero definito in (4.138) non dipenda dalle particolari f_1 e f_2 ma solo dalla loro differenza (q.o.). Ma questo è ovvio: se f è anche data (q.o.) dalla differenza $g_1 - g_2$ con $g_i \in \mathcal{F}$, allora si ha che (per il punto (ii) dell'Osservazione 4.34) $f_1 + g_2 = f_2 + g_1 \in \mathcal{F}$ (l'uguaglianza, come al solito, vale q.o. in E) e quindi $\int(f_1 + g_2) = \int(f_2 + g_1)$. Poiché (nuovamente per il punto (ii) dell'Osservazione 4.34) l'integrale della somma è uguale alla somma degli integrali si ha che $\int f_1 - \int f_2 = \int g_1 - \int g_2$. Questo argomento mostra anche che:

l'integrale di Lebesgue di due funzioni in $\mathcal{L}^1(E)$ che coincidano quasi ovunque assume lo stesso valore.

Raccogliamo nella seguente Proposizione alcune semplici proprietà della classe delle funzioni integrabili secondo Lebesgue.

Proposizione 4.36 *Nelle seguenti affermazioni \mathcal{L}^1 denota $\mathcal{L}^1(E)$.*

(i) \mathcal{L}^1 è uno spazio vettoriale e l'integrale di Lebesgue è un operatore lineare e positivo su \mathcal{L}^1 : $\forall a \in \mathbb{R}, \forall f, g \in \mathcal{L}^1, af \in \mathcal{L}^1, f + g \in \mathcal{L}^1$ e vale $\int(f + g) = \int f + \int g$; se $\mathcal{L}^1 \ni f \geq 0$ (q.o.) allora $\int f \geq 0$.

(ii) Le parti positive e negative⁵² f_+, f_- e il modulo $|f|$ di una funzione $f \in \mathcal{L}^1$ appartengono a \mathcal{L}^1 . Se $f, g \in \mathcal{L}^1$ allora $\sup\{f, g\}, \inf\{f, g\} \in \mathcal{L}^1$.

(iii) Se $f \in \mathcal{L}^1$ e $\int_E |f| = 0$ allora $f = 0$ q.o. in E .

(iv) Per ogni $f \in \mathcal{L}^1$ esiste una successione di funzioni a scalini $\{f_k\}$ tale che $f_k \rightarrow f$ q.o. e $\int_E |f - f_k| \rightarrow 0$.

(v) L'integrale di Lebesgue è invariante per traslazioni ossia per ogni $x_0 \in \mathbb{R}^n$ e per ogni $f \in \mathcal{L}^1(\mathbb{R}^n)$

$$\int_{\mathbb{R}^n} f(x) dx = \int_{\mathbb{R}^n} f(x - x_0) dx. \quad (4.139)$$

Osservazione 4.37 (i) Si noti che dalla linearità e la positività dell'integrale segue che se $f, g \in \mathcal{L}^1(E)$

$$f \leq g \text{ (q.o.)} \implies \int_E f \leq \int_E g.$$

Dunque, poiché $\pm f \leq |f|$, si ha anche che (per ogni $f \in \mathcal{L}^1$)

$$\left| \int_E f \right| \leq \int_E |f|. \quad (4.140)$$

(ii) Se $f \in \mathcal{L}^1$ e $f \geq 0$ (q.o.) allora, per ogni $\varepsilon > 0$, esistono $f_i \in \mathcal{F}$ tali che: $f_i \geq 0$ (q.o.), $\int f_2 \leq \varepsilon$ e $f = f_1 - f_2$.

[Infatti, dalla definizione di \mathcal{L}^1 segue che esistono due funzioni $g_i \in \mathcal{F}$ tali che $f = g_1 - g_2$ e dalla definizione di \mathcal{F} segue che esistono $g_i^{(k)} \rightarrow g_i$ (q.o.) con $g_i^{(k)} \in S$ e tali che $\int g_i = \sup_k \int g_i^{(k)}$. Sia \bar{k} tale che $\int g_2 - g_2^{(\bar{k})} \leq \varepsilon$. Le funzioni f_i saranno allora date da $f_i := g_i - g_2^{(\bar{k})}$.]

(iii) Se E' è un rettangolo contenuto in E e se $f \in \mathcal{L}^1(E)$ allora $f \in \mathcal{L}^1(E')$, $f \cdot \chi_{E'} \in \mathcal{L}^1(E)$ e

$$\int_{E'} f = \int_E f \chi_{E'}.$$

[Se $g \in \mathcal{F}(E)$ e se $S(E) \ni g_k \uparrow g$, allora $\chi_{E'} g_k \in S(E')$ e $\chi_{E'} g_k \uparrow f$ su E' (e $\chi_{E'} g_k \uparrow \chi_{E'} f$ su E).]

(iv) Il punto (iv) della Proposizione 4.36 è ovviamente equivalente a

$$\forall f \in \mathcal{L}^1(E), \forall \varepsilon > 0 \exists g \in S(E) : \int_E |f - g| \leq \varepsilon. \quad (4.141)$$

Infatti, si può anche assumere che $\text{supp}(g)$ sia contenuto in⁵³ E (basta sostituire i rettangoli su cui g è diversa da zero con opportuni rettangoli leggermente più piccoli).

⁵²Si ricorda che $f_+ = \sup\{f, 0\}$ e $f_- = -\inf\{f, 0\}$ cosicché $f_{\pm} \geq 0$, $f = f_+ - f_-$ e $|f| = f_+ + f_-$.

⁵³Per chiarezza: se $E = (0, 1)$, la funzione $\chi_{(0,1)} \in S(E)$ ma il suo supporto $[0, 1]$ non è contenuto in E .

Dimostrazione (della Proposizione 4.36 escluso il punto (iii) che verrà dimostrato al punto (iii) dell'Osservazione 4.39) (i): La tesi è conseguenza immediata della definizione di \mathcal{L}^1 e delle proprietà di \mathcal{F} . [Ad esempio se $f = f_1 - f_2$ con $f_i \in \mathcal{F}$ e se $a \in \mathbb{R}$ allora $af = g_1 - g_2$ dove $g_i = af_i$ se $a \geq 0$ mentre se $a < 0$, $g_1 = (-a)f_2$ e $g_2 = (-a)f_1$].
(ii): Se $\mathcal{L}^1 \ni f = f_1 - f_2$ con $f_i \in \mathcal{F}$ allora $\sup\{f_1, f_2\} \in \mathcal{F}$ e

$$f_+ = \sup\{f_1, f_2\} - f_2, \quad f_- = \sup\{f_1, f_2\} - f_1$$

da cui segue che $f_{\pm} \in \mathcal{L}^1$ e quindi (per (i) ed essendo $|f| = f_+ + f_-$) anche $|f| \in \mathcal{L}^1$. Inoltre se $f, g \in \mathcal{L}^1$, osservando che

$$\sup\{f, g\} = (f - g)_+ + g, \quad \inf\{f, g\} = -\sup\{-f, -g\},$$

si che anche $\sup\{f, g\}$ e $\inf\{f, g\}$ appartengono a \mathcal{L}^1 .

(iv): Sia $f = g - h$ con $g, h \in \mathcal{F}$ e siano $g_k \uparrow g$ e $h_k \uparrow h$ (con $g_k, h_k \in S$). Se $f_k := g_k - h_k$ allora $|f - f_k| = g - g_k + h - h_k$ e

$$\int |f - f_k| = \int g - g_k + \int h - h_k \rightarrow 0.$$

(v): Chiaramente basta dimostrare l'asserto per funzioni $f \in \mathcal{F}(\mathbb{R}^n)$ e per tali funzioni l'asserto è conseguenza della Proposizione 4.30. ■

3.3 Teoremi sull'integrazione di limiti

La potenza (e la bellezza) dell'integrale di Lebesgue è particolarmente apprezzabile nello scambio tra l'operazione di limite e quella di integrale. Cominciamo con un risultato fondamentale che, in particolare, mostra come sostituendo le funzioni a scalini con funzioni più complicate (in \mathcal{L}^1) non si ottiene una classe funzionale più ampia di \mathcal{L}^1 :

Teorema 4.38 (Convergenza monotona di Beppo Levi) *Sia $\{f_k\}$ una successione non decrescente di funzioni in $\mathcal{L}^1(E)$ tale che $\sup_k \int f_k < \infty$. Allora esiste $f \in \mathcal{L}^1(E)$ tale che $f_k \uparrow f$ q.o. ed inoltre $\int_E |f - f_k| \rightarrow 0$.*

Naturalmente da $\int |f - f_k| \rightarrow 0$ e da (4.140) segue immediatamente che

$$\lim_{k \rightarrow \infty} \int f_k = \int f := \lim_{k \rightarrow \infty} \int f_k.$$

Dimostrazione (del Teorema 4.38) Nel corso della dimostrazione tutte le disuguaglianze vanno intese quasi ovunque in E . Dimostriamo prima il teorema nel caso particolare in cui $f_k \in \mathcal{F}$ (nel qual caso otterremo che $f \in \mathcal{F}$). Assumendo, dunque, che $f_k \in \mathcal{F}$, per ogni k esistono $g_j^{(k)} \uparrow f_k$ (per $j \rightarrow \infty$) con $g_j^{(k)} \in S$ e $\int f_k = \sup_j \int g_j^{(k)}$. Sia $g_k := \sup_{j \leq k} g_k^{(j)}$. Chiaramente $\{g_k\}$ è una successione non decrescente di funzioni a scalini. Inoltre, per ogni $j \leq k$, si ha che $g_k^{(j)} \leq f_j \leq f_k$ e dunque (prendendo il massimo su $j \leq k$) $g_k \leq f_k$ e $\int g_k \leq \int f_k \leq \sup_k \int f_k < \infty$. Da queste osservazioni segue che esiste $f \in \mathcal{F}$ tale che $g_k \uparrow f$ e $\sup \int g_k = \int f$. D'altra parte, per $k \leq j$, $g_j^{(k)} \leq g_j$ e quindi (prendendo il limite per $j \rightarrow \infty$) $f_k \leq f$. Quindi, essendo $g_k \leq f_k \leq f$ si ha che $f_k \uparrow f$ (q.o.) e $\int g_k \leq \int f_k \leq \int f$ da cui (prendendo l'estremo superiore su k)

$$\int f \leq \lim \int f_k \leq \int f,$$

Poiché $f_k \leq f$, $\int |f - f_k| = (\int f - \int f_k) \rightarrow 0$.

Passiamo al caso generale in cui $f_k \in \mathcal{L}^1$. Sia $g_k := f_{k+1} - f_k$ cosicché $g_k \geq 0$ e $f_k = f_1 + \sum_{j=1}^{k-1} g_j$. Dalle ipotesi segue che

$$\int \sum_{j=1}^{k-1} g_j = \int f_k - \int f_1 \leq \sup_j \int f_j + \int |f_1| < \infty.$$

Dunque la serie numerica a termini positivi $\sum \int g_j$ è convergente. Per il punto (ii) dell'Osservazione 4.37 si ha che, per ogni k , esistono due funzioni $\varphi_k, \psi_k \in \mathcal{F}$ non negative tali che $\int \psi_k \leq 2^{-k}$ e $g_k = \varphi_k - \psi_k$. Quindi $\sum \int \psi_k < \infty$ e, per quanto osservato prima (ed essendo $\sum_1^k \varphi_j = \sum_1^k g_j + \sum_1^k \psi_j$), anche $\sum_{j \geq 1} \int \varphi_j < \infty$. Ora $\sum_1^k \varphi_j$ e $\sum_1^k \psi_j$ appartengono a \mathcal{F} (punto (ii), Osservazione 4.34) e quindi per la prima parte della dimostrazione $\sum_1^k \varphi_j \uparrow \varphi \in \mathcal{F}$ e $\sum_1^k \psi_j \uparrow \psi \in \mathcal{F}$ ed inoltre $\sum_{j \geq 1} \int \varphi_j = \int \varphi$ e $\sum_{j \geq 1} \int \psi_j = \int \psi$. In conclusione, $f_k = f_1 + \sum_{k=1}^{k-1} \varphi_j - \sum_{k=1}^{k-1} \psi_j \rightarrow f_1 + \varphi - \psi \in \mathcal{L}^1$ e

$$\int f_k = \int f_1 + \sum_1^{k-1} \int \varphi_j - \sum_1^{k-1} \int \psi_j \rightarrow \int f_1 + \int \varphi - \int \psi = \int f,$$

il che (essendo $f - f_k \geq 0$) è equivalente a $\int |f - f_k| \rightarrow 0$. ■

Osservazione 4.39 (i) Chiaramente, un enunciato equivalente al Teorema di convergenza monotona è: Se $f_k \in \mathcal{L}^1$ e $f_k \uparrow f \in \mathcal{L}^1$, allora $\int |f - f_k| \rightarrow 0$.

(ii) Un corollario immediato del teorema di convergenza monotona è il seguente utile criterio sull'integrazione termine a termine di serie di funzioni:

Sia $u_k \in \mathcal{L}^1$ e supponiamo che la serie numerica $\sum \int |u_k|$ converga. Allora $\sum_{j=1}^k u_j$ converge (per $k \rightarrow \infty$) q.o. e $\sum \int u_j = \int \sum u_j$.

[Infatti applicando il teorema di convergenza monotona alle funzioni⁵⁴ $f_k^\pm := \sum_1^k (u_j)^\pm$ si ha che $f_k^\pm \uparrow f^\pm \in \mathcal{L}^1$ e $\int f_k^\pm \rightarrow \int f^\pm$ e quindi $\sum u_j := \sum (u_j^+ - u_j^-)$ converge q.o. e la tesi segue.]

(iii) **Dimostrazione del punto (iii) della Proposizione 4.36:** Sia $f \in \mathcal{L}^1$ con $\int |f| = 0$ e poniamo $u_k := f$ per ogni $k \in \mathbb{N}$. Poiché $\sum \int |u_k|$ converge (a zero), per il punto precedente $\sum_{j=1}^k u_j = kf$ converge q.o. e quindi $f = 0$ q.o. (nei punti in cui $f(x) \neq 0$, $k|f(x)| \rightarrow \infty$). ■

(iv) Se $f_k \in \mathcal{L}^1$ e $f_k \uparrow f$ (q.o.) e $\sup \int f_k = \infty$ allora⁵⁵ $f \notin \mathcal{L}^1$.

Il teorema forse più importante e con più applicazioni dell'intera teoria è il seguente teorema dovuto a Lebesgue.

Teorema 4.40 (Convergenza dominata) Assumiamo che la successione $\{f_k\}$ di funzioni in $\mathcal{L}^1(E)$ converga q.o. in E e che esista una funzione $g \in \mathcal{L}^1(E)$ tale che $|f_k| \leq g$ (q.o.). Allora $f_k \rightarrow f \in \mathcal{L}^1(E)$ e $\int_E |f - f_k| \rightarrow 0$.

Dimostrazione Dati k ed j , sia $g_k^{(j)} := \sup\{f_j, \dots, f_{j+k}\}$ e $h_k^{(j)} := \inf\{f_j, \dots, f_{j+k}\}$. Per il punto (ii) della Proposizione 4.36, $g_k^{(j)}, h_k^{(j)} \in \mathcal{L}^1$ e le successioni $\{g_k^{(j)}\}_{k \geq 1}$ e $\{-h_k^{(j)}\}_{k \geq 1}$ sono monotone non decrescenti. Inoltre da $|g_k^{(j)}| \leq g$ e $|h_k^{(j)}| \leq g$ segue che $\sup_k \int g_k^{(j)} \leq \int g$ e $\sup_k \int (-h_k^{(j)}) \leq \int g$. Quindi, per il teorema di convergenza monotona esistono $g_j, h_j \in \mathcal{L}^1$ tali che (per $k \rightarrow \infty$ e j fissato) $g_k^{(j)} \uparrow g_j \in \mathcal{L}^1$ e $-h_k^{(j)} \uparrow -h_j \in \mathcal{L}^1$ (e, ovviamente $\int |g_j|, \int |h_j| \leq \int g$). Dunque $g_j = \sup\{f_k : k \geq j\}$ e $h_j = \inf\{f_k : k \geq j\}$ sono funzioni di \mathcal{L}^1 che, al variare di j , formano delle successioni, rispettivamente, non crescente e non decrescente. Infatti, $\lim g_j = \limsup f_j = \lim f_j$ (q.o.) e $\lim h_j = \liminf f_j = \lim f_j$ (q.o.). Dunque, per il

⁵⁴Come al solito, $(u_j)^\pm$ denotano la parte positiva e la parte negativa di u_j ; si noti che $0 \leq f_k^\pm \uparrow$ e che $\int f_k^\pm \leq \sum_j \int |u_j| < \infty$.

⁵⁵Si confronti tale affermazione con il punto (i) dell'Osservazione 4.34.

teorema di convergenza monotona (essendo $h_j \uparrow f$, $\sup h_j \leq \int g < \infty$) $f \in \mathcal{L}^1$ e $\int |f - h_j| \rightarrow 0$. Applicando lo stesso ragionamento a $-g_j \uparrow -f$ si ottiene che anche $\int |f - g_j| \rightarrow 0$ e quindi (osservando che $g_j \geq f_j \geq h_j$)

$$0 \leq \int (g_j - h_j) = \int (g_j - f_j) + \int (f_j - h_j) \rightarrow 0.$$

Poiché sia $f_j(x)$ sia $f(x)$ appartengono (q.o.) all'intervallo $[h_j(x), g_j(x)]$ segue che

$$\int |f - f_j| \leq \int (g_j - h_j) \rightarrow 0. \quad \blacksquare$$

Alcuni corollari immediati del Teorema di convergenza dominata (o della sua dimostrazione) sono i seguenti.

Lemma di Fatou Sia $\{f_k\}$ una successione di funzioni non negative di \mathcal{L}^1 tale che $\sup \int f_k < \infty$ e $f_k \rightarrow f$ (q.o.). Allora $f \in \mathcal{L}^1$ e $\int f \leq \liminf \int f_k$.

Dimostrazione Sia h_k definita come nella dimostrazione del Teorema di convergenza dominata, ossia $h_k = \inf\{f_j, j \geq k\}$. Ripetendo l'argomento usato prima, essendo $0 \leq h_k \leq f_k$, si ha che $h_k \in \mathcal{L}^1$, $h_k \uparrow f$. Inoltre, per ogni $j \geq 0$, $h_k \leq f_{k+j}$ e dunque $\int h_k \leq \int f_{k+j}$ da cui segue che

$$\int h_k \leq \liminf_{j \rightarrow \infty} \int f_{k+j} = \liminf_{k \rightarrow \infty} \int f_k.$$

Dal Teorema di convergenza monotona segue allora che $f \in \mathcal{L}^1$ e che $\int f = \lim \int h_k \leq \liminf \int f_k$. \blacksquare

La seguente affermazione dà un criterio sull'appartenenza a \mathcal{L}^1 di limiti q.o. di funzioni $f_k \in \mathcal{L}^1$ (senza fare alcuna ipotesi sulla successione ma solo sul limite):

Proposizione 4.41 Sia $f_k \in \mathcal{L}^1$, $f_k \rightarrow f$ (q.o.) e $|f| \leq g \in \mathcal{L}^1$. Allora $f \in \mathcal{L}^1$.

Dimostrazione Basta applicare il Teorema di convergenza dominata alla successione definita come $\varphi_k := \inf\{g, \sup\{f_k, -g\}\}$ (osservando che $\varphi_k = g$ se $f_k > g$, $\varphi_k = f_k$ se $|f_k| \leq g$ e $\varphi_k = -g$ se $f_k < -g$ e dunque che $\varphi_k \rightarrow f$ q.o.). \blacksquare

Il Teorema di convergenza dominata ha innumerevoli applicazioni. A titolo di esempio, concludiamo questo paragrafo con un risultato sulla derivazione sotto segno di integrale che estende la Proposizione 2.36.

Proposizione 4.42 Sia $E \subseteq \mathbb{R}^n$ un rettangolo, $I \subseteq \mathbb{R}$ un intervallo aperto e $f : E \times I \rightarrow \mathbb{R}$ una funzione che verifichi le seguenti ipotesi: per ogni $y \in I$, $f(\cdot, y) \in \mathcal{L}^1(E)$; per quasi tutti gli $x \in E$, $f(x, \cdot) \in C^1(I)$; esiste una funzione $g \in \mathcal{L}^1(E)$ tale che $|f_y(x, y)| \leq g(x)$. Allora la funzione $F(y) := \int_E f(x, y) dx$ è di classe $C^1(I)$, $f_y(\cdot, y) \in \mathcal{L}^1(E)$ e $F'(y) = \int_E f_y(x, y) dx$.

Dimostrazione Fissiamo $y \in I$ e sia $\{h_j\}$ una qualunque successione di numeri diversi da 0 tale che $h_j \rightarrow 0$. Sia $f_j(x) := \left(f(x, y + h_j) - f(x, y) \right) / h_j$. Dalle ipotesi su f segue che $f_j(\cdot) \rightarrow f_y(\cdot, y)$ q.o. in A e (per quasi tutti gli x) $|f_j(x)| = \left| \int_0^1 f_y(x, y + th_j) dt \right| \leq g(x)$. La derivabilità di F è conseguenza immediata del Teorema di convergenza dominata. Sia ora $I \ni y_n \rightarrow y \in I$. Come sopra si ha che $|F'(y_n) - F'(y)| = \left| \int_E (f_y(x, y) - f_y(x, y_n)) dx \right| \leq 2 \int_E g(x) dx$, e dunque, ancora per convergenza dominata, $F'(y_n) \rightarrow F'(y)$ ossia $F \in C^1(I)$. \blacksquare

3.4 Lo spazio di Banach L^1

Abbiamo visto che $\mathcal{L}^1 := \mathcal{L}^1(E)$ è uno spazio vettoriale. L'integrale del modulo definisce su \mathcal{L}^1 una seminorma⁵⁶ che denotiamo

$$\|f\|_{L^1} := \int_E |f|, \quad (f \in \mathcal{L}^1). \quad (4.142)$$

⁵⁶Si ricorda che una funzione $\|\cdot\|$ da uno spazio vettoriale V in $[0, \infty)$ si dice **seminorma** se verifica (i) e (iii) di (1.8) (ossia se valgono la omogeneità e la disuguaglianza triangolare ma non necessariamente la non degenerazione).

Infatti, per il punto (iii) della Proposizione 4.36, si ha che, da $\|f\|_{L^1} = 0$ (con $f \in \mathcal{L}^1$), segue solo che $f = 0$ quasi ovunque e non che f sia *identicamente* nulla. È però chiaro come ovviare a tale problema: basta ‘*identificare funzioni che differiscono su di un insieme trascurabile*’. Per far questo introduciamo in \mathcal{L}^1 una relazione d’equivalenza. Se $f, g \in \mathcal{L}^1$, diremo che $f \sim g$ se $f = g$ quasi ovunque. Ovviamente \sim è una relazione d’equivalenza in \mathcal{L}^1 .

Definizione 4.43 $L^1(E)$ è l’insieme delle classi di equivalenza indotte su $\mathcal{L}^1(E)$ dalla relazione d’equivalenza \sim appena introdotta. Se $\tilde{f} \in L^1(E)$, definiamo integrale di Lebesgue di \tilde{f} su E la quantità $\int_E f$ dove $f \in \tilde{f}$ è un qualunque rappresentante della classe d’equivalenza \tilde{f} .

Poichè l’integrale di Lebesgue di due funzioni che differiscono solo su di un insieme trascurabile coincide, tale definizione è ben posta. È chiaro che⁵⁷ L^1 è uno spazio vettoriale⁵⁸ e su L^1 possiamo definire $\|\cdot\|_{L^1}$ ponendo, per $\tilde{f} \in L^1$, $\|\tilde{f}\|_{L^1} := \|f\|_{L^1}$ dove f è un qualunque elemento di \tilde{f} . Chiaramente $\|\cdot\|_{L^1}$ è una norma su L^1 e quindi L^1 è uno spazio normato. Uno dei maggior successi della teoria dell’integrazione secondo Lebesgue è che L^1 è completo:

Teorema 4.44 (Riesz, Fischer) L^1 è uno spazio di Banach.

Dimostrazione Osserviamo innanzitutto che dalla definizione di L^1 (e dal fatto che l’integrale di Lebesgue assume lo stesso valore su funzioni \mathcal{L}^1 che differiscono su di un insieme trascurabile) segue che l’enunciato del teorema è equivalente alla seguente affermazione:

Per ogni successione $\{f_k\}$ di Cauchy di funzioni in \mathcal{L}^1 esiste una funzione $f \in \mathcal{L}^1$ tale che $\|f - f_k\|_{L^1} \rightarrow 0$.

Dimostriamo dunque tale asserzione. Da teoremi noti sulle successioni segue che è sufficiente provare che $\|f_{k_j} - f\|_{L^1} \rightarrow 0$ per qualche sottosuccessione⁵⁹ di $\{f_k\}$. Dalla successione $\{f_k\}$ possiamo estrarre una sottosuccessione $\{g_j\} := \{f_{k_j}\}$ tale che⁶⁰ $\|g_{j+1} - g_j\|_{L^1} \leq 2^{-j}$. Le funzioni non negative di \mathcal{L}^1 $G_k := \sum_{j=1}^k |g_{j+1} - g_j|$ formano una successione monotona non decrescente ed inoltre $\int G_j \leq \sum_{j \geq 1} 2^{-j} = 1$. Quindi, per il Teorema di convergenza monotona, esiste $G \in \mathcal{L}^1$ tale che $G_j \uparrow G$ quasi ovunque. Dunque anche $g_j = g_1 + \sum_{i=1}^{j-1} (g_{i+1} - g_i)$ converge quasi ovunque ad una funzione f . Inoltre, essendo $|g_j| \leq |g_1| + G \in \mathcal{L}^1$, dal Teorema di convergenza dominata segue che $\|g_j - f\|_{L^1} \rightarrow 0$. ■

Osservazione 4.45 (i) Nel corso della dimostrazione abbiamo anche dimostrato la seguente affermazione:

Se $f_k, f \in \mathcal{L}^1$ sono tali che $\|f_k - f\|_{L^1} \rightarrow 0$, allora esiste una sottosuccessione $\{f_{k_j}\}$ che converge ad f quasi ovunque.

Questa affermazione non può essere migliorata (si veda C.31).

(ii) Lo spazio ‘giusto’ da considerare (dal punto di vista ‘funzionale’) è dunque L^1 . Si faccia però attenzione al fatto che gli elementi di L^1 non sono funzioni; ad esempio, se $\tilde{f} \in L^1$ non ha senso, in generale, parlare di ‘valore di \tilde{f} in $x \in E$ ’ oppure⁶¹ scrivere $\tilde{f}(x)$.

Ciononostante ci sono dei casi in cui è possibile scegliere un rappresentante *naturale* $f \in \tilde{f}$ e quindi di *identificare* \tilde{f} con tale funzione. Questo è, ad esempio, il caso in cui \tilde{f} contiene una

⁵⁷Come per \mathcal{L}^1 non indicheremo, qualora non ve ne sia necessità, esplicitamente il rettangolo di riferimento E .

⁵⁸Se $\tilde{f}, \tilde{g} \in L^1$ e $a, b \in \mathbb{R}$, $a\tilde{f} + b\tilde{g}$ è, per definizione data dalla classe di equivalenza di $af + bg$ dove f e g sono elementi arbitrari, rispettivamente, di \tilde{f} e \tilde{g} : è immediato verificare che tale definizione è ben posta ossia che non dipende dai particolari elementi f e g ma solo dalle classi d’equivalenza.

⁵⁹In generale, se $\{x_k\}$ è una successione di Cauchy di uno spazio metrico X e se $\lim_{j \rightarrow \infty} x_{k_j} = x \in X$ allora $x_k \rightarrow x$. Infatti: sia $\varepsilon > 0$ e sia k_0 tale che $d(x_{k_j}, x) \leq \varepsilon$ e $d(x_k, x_h) \leq \varepsilon \forall j, k, h \geq k_0$ (e $d(\cdot, \cdot)$ è la distanza in X). Allora se j_0 è tale che $k_{j_0} \geq k_0$, $\forall k \geq k_0$ si ha $d(x_k, x) \leq d(x_k, x_{k_{j_0}}) + d(x_{k_{j_0}}, x) \leq 2\varepsilon$.

⁶⁰Anche questo è un fatto generale che vale per qualunque successione di Cauchy $\{x_k\}$ di un qualunque spazio metrico X : Infatti per ogni $j \geq 1$ sia h_j tale che $d(x_k, x_h) \leq 2^{-j}$ per ogni $k, h \geq h_j$. Definiamo ora $k_1 = h_1$ e, ricorsivamente per $j \geq 2$, $k_j = \sup\{k_{j-1} + 1, h_j\}$. È allora chiaro che la sottosuccessione x_{k_j} è tale che $d(x_{k_{j+1}}, x_{k_j}) \leq 2^{-j}$.

⁶¹Tale espressione non ha senso anche se la interpretassimo come $f(x)$ dove $f \in \tilde{f}$. Infatti, fissato un punto x , poichè $\{x\}$ è un insieme trascurabile siamo sempre liberi di scegliere un altro rappresentante che assuma nel punto x un valore del tutto arbitrario.

funzione f continua. Infatti, sia $\tilde{f} \in L^1(E)$ ed assumiamo che esista una funzione f continua su E che appartenga a \tilde{f} . Allora, fissato $x \in E$, se K_r denota il cubo aperto di centro x e lato r , si ha che⁶²

$$\lim_{r \downarrow 0} \frac{1}{r^n} \int_{K_r} \tilde{f} = f(x). \quad (4.143)$$

[Infatti, poiché l'integrale in (4.143) non dipende dal particolare rappresentante segue che $\int_{K_r} \tilde{f} = \int_{K_r} f$ dove f è il rappresentante continuo in \tilde{f} . Fissiamo $\varepsilon > 0$ e sia $r_0 > 0$ così piccolo che $|f(y) - f(x)| \leq \varepsilon$ per ogni $y \in K_{r_0}$. Allora per ogni $0 < r \leq r_0$ si ha

$$\left| \left(\frac{1}{r^n} \int_{K_r} f \right) - f(x) \right| = \left| \frac{1}{r^n} \int_{K_r} [f(y) - f(x)] dy \right| \leq \frac{1}{r^n} \int_{K_r} |f(y) - f(x)| dy \leq \varepsilon,$$

da cui segue (4.143).]

In tal senso penseremo a spazi di funzioni regolari (quali $C^k(E)$ o $C_0^k(E) :=$ l'insieme delle funzioni di classe C^k con supporto compatto contenuto in E) come 'sottospazi' di L^1 .

Per quanto detto, il punto (iv) della Proposizione 4.36 può essere riformulato dicendo che $S(E)$ è denso⁶³ in $L^1(E)$. Infatti vale anche la seguente

Proposizione 4.46 $C_0^\infty(E)$ è denso in $L^1(E)$.

Tale risultato è un immediato corollario della densità delle funzioni a scalini e del seguente

Lemma 4.47 Data $f \in S(E)$ e dato $\varepsilon > 0$, esiste $\varphi \in C_0^\infty(E)$ tale che $\|f - \varphi\|_{L^1} \leq \varepsilon$.

Dimostrazione (a) Dall'Esempio A.38 segue che per ogni $a < b$ e per ogni $0 < \varepsilon < (b - a)/2$ esiste $g \in C_0^\infty(\mathbb{R})$ tale che: $0 \leq g \leq 1$; $\text{supp } g \subseteq [a, b]$ e $g := 1$ su $[a + \varepsilon, b - \varepsilon]$.

(b) Dato un rettangolo limitato e non degenere $R \subseteq \mathbb{R}^n$ ed $\varepsilon > 0$ è allora facile costruire una funzione $\psi \in C_0^\infty(\mathbb{R}^n)$ tale che $0 \leq \psi \leq 1$, $\text{supp } \psi \subseteq R$ e $\|1 - \psi\|_{L^1} \leq \varepsilon$. Infatti sia $R = I_1 \times \cdots \times I_n$ e siano $a_i < b_i$ gli estremi di I_i . Fissato arbitrariamente $0 < \sigma < \min(b_i - a_i)/2$, siano g_i le funzioni definite al punto (a) con ε sostituito da σ e a, b sostituiti da a_i, b_i per $i = 1, \dots, n$. Se R_σ denota il rettangolo $[a_1 + \sigma, b_1 - \sigma] \times \cdots \times [a_n + \sigma, b_n - \sigma]$ e se $\psi = g_1 \cdots g_n$ si ha che $\int_{\mathbb{R}^n} |1 - \psi| = \int_{R \setminus R_\sigma} (1 - \psi) \leq \text{mis } R - \text{mis } R_\sigma$. È dunque chiaro che prendendo $\sigma = \sigma(\varepsilon)$ sufficientemente piccolo si ottiene la funzione ψ cercata.

(c) Sia $f \in S(E)$. Possiamo scrivere $f = f_1 + f_2$ con f_1 come in (4.128) con gli R_i limitati, disgiunti e non degeneri e con $c_i \neq 0$, mentre $f_2 = 0$ q.o. (in altri termini in f_2 abbiamo isolato i possibili contributi relativi a rettangoli degeneri o a $c_i = 0$). Dato $\varepsilon > 0$, per ogni $i \leq N$ scegliamo ψ_i come in (b) con R sostituito da R_i e con ε sostituito da $\varepsilon/(N \sup_i |c_i|)$. Allora, ponendo $\varphi := \sum_{i=1}^N c_i \psi_i$, si ha che $\varphi \in C_0^\infty(E)$ e che

$$\begin{aligned} \int_E |f - \varphi| &\leq \int_E |f_1 - \varphi| + \int_E |f_2| = \int_E |f_1 - \varphi| \\ &\leq \sum_{i=1}^N |c_i| \int_{R_i} |\chi_{R_i} - \psi_i| \leq \varepsilon. \quad \blacksquare \end{aligned}$$

3.5 Integrali iterati e Teorema di Fubini

Lo scopo di questo paragrafo è di mostrare come sia possibile, sotto opportune ipotesi, ridurre il calcolo di un integrale su \mathbb{R}^{n+m} al calcolo 'iterato' di due integrali, rispettivamente, n ed m -dimensionali. Applicando più volte tale metodo si potrà ridurre il calcolo di un integrale in \mathbb{R}^n a n integrali unidimensionali.

⁶²Grazie al punto (iii) dell'Osservazione 4.37, gli integrali in (4.143) sono ben definiti; ' $r \downarrow 0$ ' significa, naturalmente, che il limite è fatto per valori positivi di r .

⁶³Si ricorda che un sottoinsieme X_0 di uno spazio metrico X si dice denso in X se per ogni $x \in X$ e per ogni $\varepsilon > 0$ esiste un elemento $x_0 \in X_0$ tale che $d(x_0, x) \leq \varepsilon$.

In questo paragrafo denoteremo punti di $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$ con (x, y) con $x \in \mathbb{R}^n$ e $y \in \mathbb{R}^m$. Se $A \subseteq \mathbb{R}^{n+m}$ e se x è un punto fissato di \mathbb{R}^n chiameremo la x -**sezione di A** , e la denoteremo con A_x , l'insieme

$$A_x := \{y \in \mathbb{R}^m : (x, y) \in A\};$$

analogamente denoteremo la y -**sezione di A** con

$$A^y := \{x \in \mathbb{R}^n : (x, y) \in A\}.$$

Osservazione 4.48 (i) Sia $E = E_1 \times E_2$ un rettangolo in \mathbb{R}^{n+m} con E_1 rettangolo di \mathbb{R}^n e E_2 rettangolo di \mathbb{R}^m . Sappiamo che $\chi_E = \chi_{E_1} \cdot \chi_{E_2}$ e che⁶⁴ $\text{mis}_{n+m} E = \text{mis}_n E_1 \cdot \text{mis}_m E_2$. Sia ora $f \in S(E)$, allora, per ogni $x \in E_1$, la funzione $y \in E_2 \rightarrow f(x, y)$ appartiene a $S(E_2)$ e per ogni $y \in E_2$, la funzione $x \in E_1 \rightarrow f(x, y)$ appartiene a $S(E_1)$. Inoltre anche le funzioni, rispettivamente, di y e x date da $\int_{E_1} f(x, y) dx$ e $\int_{E_2} f(x, y) dy$ sono funzioni a scalini e si ha⁶⁵

$$\int_E f = \int_{E_1} \left(\int_{E_2} f(x, y) dy \right) dx = \int_{E_2} \left(\int_{E_1} f(x, y) dx \right) dy. \quad (4.144)$$

(ii) Se $Q \subseteq \mathbb{R}^n$ è un insieme trascurabile, allora l'insieme $Q \times A \subseteq \mathbb{R}^n \times \mathbb{R}^m$ è un insieme trascurabile in \mathbb{R}^{n+m} per ogni insieme $A \subseteq \mathbb{R}^m$.

Infatti: sia $A_j := A \cap (-j, j)^m$ e si noti che basta dimostrare l'asserto per A_j per ogni j (poiché $Q \times A = \cup_j Q \times A_j$). Dato $\varepsilon > 0$ sia $\{K_i\}$ un ricoprimento aperto e numerabile di Q ($K_i \subseteq \mathbb{R}^n$) tale che $\sum \text{mis}_n(K_i) \leq \varepsilon/(2j)^m$. Allora $\{K_i \times (-j, j)^m\}$ è un ricoprimento aperto di $Q \times A_j$ e $\sum \text{mis}_{n+m}(K_i \times (-j, j)^m) = (2j)^m \sum \text{mis}_n K_i \leq \varepsilon$. ■

Lemma 4.49 Sia $Q \subseteq \mathbb{R}^{n+m}$ un insieme trascurabile. Allora per quasi tutti gli $x \in \mathbb{R}^n$, Q_x è un insieme trascurabile in \mathbb{R}^m e per quasi tutti gli $y \in \mathbb{R}^m$, Q^y è un insieme trascurabile in \mathbb{R}^n .

Dimostrazione Dimostriamo che per quasi tutti gli $x \in \mathbb{R}^n$, Q_x è di misura nulla (l'altro caso è naturalmente analogo). Per il punto (v) dell'Osservazione 4.24, esiste un ricoprimento numerabile di Q formato da cubi aperti, $K^{(i)} := K_1^{(i)} \times K_2^{(i)} \subseteq \mathbb{R}^n \times \mathbb{R}^m$ (con $i \geq 1$), la somma delle cui misure è finita e tale che ogni punto di Q appartiene ad un numero infinito di cubi in $\{K^{(i)}\}$. Per il punto (i) dell'Osservazione 4.48 si ha che

$$\infty > \sum_{i \geq 1} \text{mis} K^{(i)} = \sum_{i \geq 1} \int_{\mathbb{R}^n} \left(\int_{\mathbb{R}^m} \chi_{K^{(i)}} dy \right) dx,$$

e quindi, per il Teorema di convergenza monotona (o, più precisamente, per il punto (ii) dell'Osservazione 4.39), la serie $\sum \int_{\mathbb{R}^m} \chi_{K^{(i)}} dy$ converge per quasi tutti gli x in \mathbb{R}^n . Per ogni x tale che $Q_x \neq \emptyset$, siano $K_2^{(i_j)}$ tutti i cubi in $\{K_2^{(i)}\}$ tali che $Q_x \cap K_2^{(i_j)} \neq \emptyset$; si noti che $x \in K_1^{(i_j)}$ per ogni j e che se $y \in Q_x$ allora esistono infiniti cubi in $\{K_2^{(i_j)}\}$ che contengono y . Ora, se x è tale che $Q_x \neq \emptyset$, dalla definizione di $K^{(i_j)}$ segue che $\chi_{K_1^{(i_j)}}(x) = 1$ per ogni j e dunque

$$\sum_{i \geq 1} \int_{\mathbb{R}^m} \chi_{K^{(i)}} = \sum_{i \geq 1} \chi_{K_1^{(i)}}(x) \cdot \text{mis} K_2^{(i)} \geq \sum_{j \geq 1} \chi_{K_1^{(i_j)}}(x) \cdot \text{mis} K_2^{(i_j)} = \sum_{j \geq 1} \text{mis} K_2^{(i_j)}.$$

Questo significa che, per quasi tutti gli x , Q_x è ricoperto dai $K_2^{(i_j)}$ la somma delle cui misure è finita e ogni $y \in Q_x$ appartiene ad numero infinito di cubi in $\{K_2^{(i_j)}\}$. Dunque (ancora per il punto (v) dell'Osservazione 4.24) per quasi tutti gli $x \in \mathbb{R}^n$, Q_x è un insieme trascurabile. ■

⁶⁴Si usi, in tale contesto, la convenzione che $\infty \cdot 0 = 0$.

⁶⁵Si noti che se $f = \sum_{i=1}^N c_i \chi_{R_i}$ e se $R_i = R_i^{(1)} \times R_i^{(2)} \subseteq \mathbb{R}^n \times \mathbb{R}^m$, allora $\int_{E_2} f = \sum_{i=1}^N c_i^{(1)} \chi_{R_i^{(1)}}(x)$ con $c_i^{(1)} = c_i \text{mis}_m(R_i^{(2)})$.

Teorema 4.50 (Fubini) Sia $E := E_1 \times E_2 \subseteq \mathbb{R}^n \times \mathbb{R}^m$ un rettangolo e sia $f \in \mathcal{L}^1(E)$. Allora, per quasi tutti gli $x \in E_1$, $y \rightarrow f(x, y) \in \mathcal{L}^1(E_2)$ e $x \rightarrow \int_{E_2} f(x, y) dy \in \mathcal{L}^1(E_1)$; per quasi tutti gli $y \in E_2$, $x \rightarrow f(x, y) \in \mathcal{L}^1(E_1)$ e $y \rightarrow \int_{E_1} f(x, y) dx \in \mathcal{L}^1(E_2)$. Inoltre si ha

$$\int_E f = \int_{E_1} \left(\int_{E_2} f(x, y) dy \right) dx = \int_{E_2} \left(\int_{E_1} f(x, y) dx \right) dy. \quad (4.145)$$

Dimostrazione Chiaramente (per linearità) basta dimostrare l'enunciato del teorema avendo sostituito \mathcal{L}^1 con \mathcal{F} ; inoltre dimostreremo solo la parte dell'enunciato relativa alla prima uguaglianza in (4.145) (poiché l'altra parte è del tutto analoga).

(a) Data $f \in \mathcal{F}(E)$ esiste un insieme trascurabile $Q \subseteq E$ e funzioni a scalini f_k tali che $f_k \uparrow f$ per ogni $x \in E \setminus Q$ ed inoltre $\int f_k \uparrow \int f$.

(b) Sia $F_k(x) := \int_{E_2} f_k(x, y) dy$. Per l'Osservazione 4.48, F_k è una funzione a scalini (ed ovviamente $\{F_k\}$ è una successione non decrescente) e

$$\int_{E_1} F_k = \int_E f_k \leq \int_E f < \infty, \quad (4.146)$$

e quindi (Proposizione 4.29 e definizione di \mathcal{F}) esiste $F \in \mathcal{F}(E_1)$ tale che $F_k \uparrow F$ (q.o. in E_1) e $\int_{E_1} F_k \uparrow \int_{E_1} F$. Sia $N_1 \subseteq E_1$ l'insieme trascurabile per cui $F_k(x) \uparrow F(x)$ per ogni $x \in E_1 \setminus N_1$.

(c) Per il Lemma 4.49, esiste un insieme di misura nulla $Q_1 \subseteq E_1$ tale che, per ogni $x \in E_1 \setminus Q_1$, Q_x è trascurabile. Sia $M_1 := N_1 \cup Q_1$ che, per quanto detto, è un insieme trascurabile.

(d) Fissato $x \in E_1 \setminus M_1$, $f_k(x, y) \uparrow f(x, y)$ per quasi tutti gli $y \in E_2$ (e cioè per $y \in E_2 \setminus Q_x$ poiché $y \notin Q_x \implies (x, y) \notin Q$); inoltre $\int_{E_2} f_k(x, y) dy := F_k(x) \uparrow F(x) < \infty$ (poiché $x \notin N_1$). Dunque, se $x \in E_1 \setminus M_1$, (per la Proposizione 4.29 e per la definizione di \mathcal{F}) $y \rightarrow f(x, y) \in \mathcal{F}(E_1)$ e $\int_{E_2} f_k(x, y) dy = F_k(x) \uparrow \int_{E_2} f(x, y) dy$. Quindi, per $x \in E_1 \setminus M_1$,

$$F(x) = \lim F_k(x) := \lim \int_{E_2} f_k(x, y) dy = \int_{E_2} f(x, y) dy.$$

Ovvero F e $\int_{E_2} f(\cdot, y) dy$ coincidono quasi ovunque (ed appartengono a $\mathcal{F}(E_1)$) e, ricordando il punto (a) e la (4.146),

$$\int_{E_1} \left(\int_{E_2} f(x, y) dy \right) dx = \int_{E_1} F = \lim \int_{E_1} F_k = \lim \int_E f_k = \int_E f. \quad \blacksquare$$

3.6 Funzioni ed insiemi misurabili. Misura di Lebesgue

In questo paragrafo verrà discussa la teoria della misura secondo Lebesgue (che dedurremo dall'analisi sinora fatta). Tale teoria, in particolare, ci permetterà (nel prossimo paragrafo) di estendere l'integrale di Lebesgue su insiemi assai più generali che non i rettangoli.

Definizione 4.51 Una funzione⁶⁶ $f : E \rightarrow \mathbb{R}$ si dice **misurabile** se esistono $f_k \in S(E)$ tali che $f = \lim f_k$ quasi ovunque in E . Un insieme $A \subseteq E$ si dice **misurabile** se la sua funzione caratteristica χ_A è una funzione misurabile. La famiglia delle funzioni o degli insiemi misurabili su E verrà denotato con $\mathcal{M}(E)$. Se A è un insieme misurabile si definisce la **misura di Lebesgue** di A la quantità

$$m(A) := \begin{cases} \int_E \chi_A, & \text{se } \chi_A \in \mathcal{L}^1(E), \\ \infty, & \text{se } \chi_A \notin \mathcal{L}^1(E). \end{cases} \quad (4.147)$$

⁶⁶Come al solito E è un rettangolo arbitrario che potrebbe coincidere con tutto \mathbb{R}^n . Come già fatto, ometteremo spesso nelle notazioni il simbolo del rettangolo E di riferimento.

Osservazione 4.52 (i) Si noti che nella definizione di misurabilità non entra alcuna condizione di monotonia.

(ii) Ovviamente un rettangolo R è misurabile e la sua misura di Lebesgue coincide con la sua misura n -dimensionale mis_n .

(iii) Se $A \in \mathcal{M}$ e A è limitato allora $m(A) < \infty$ [infatti se A è limitato esiste un rettangolo limitato $E \supseteq A$ e la funzione $1 \in \mathcal{L}^1(E)$ e l'affermazione fatta è conseguenza immediata del Teorema di convergenza dominata]. Il viceversa non è vero⁶⁷.

Proposizione 4.53

(i) Sia $f \in \mathcal{M}$. Allora $f \in \mathcal{L}^1$ se e solo se esiste $g \in \mathcal{L}^1$ tale che $|f| \leq g$ (q.o.).

(ii) Il modulo, le combinazioni lineari (a coefficienti in \mathbb{R}), il prodotto, l'involuppo superiore ed inferiore⁶⁸ di funzioni misurabili sono funzioni misurabili. Il reciproco di una funzione misurabile e q.o. diversa da 0 è una funzione misurabile.

(iii) Se $f_k \in \mathcal{M}$ e $f_k \rightarrow f$ q.o. allora $f \in \mathcal{M}$.

(iv) Il complementare di un insieme misurabile è misurabile.

(v) L'intersezione numerabile⁶⁹ di insiemi misurabili è misurabile.

(vi) L'unione numerabile di insiemi misurabili è misurabile.

(vii) Se $A_i \in \mathcal{M}$ ($i \in \mathbb{N}$) sono insiemi disgiunti, allora $m(\cup_i A_i) = \sum_i m(A_i)$.

(viii) Se A_i ($i \in \mathbb{N}$) sono insiemi misurabili e $A_i \subseteq A_{i+1}$, allora $m(\cup_i A_i) = \lim m(A_i)$.

(ix) Se A_i ($i \in \mathbb{N}$) sono insiemi misurabili, $A_{i+1} \subseteq A_i$ e $m(A_1) < \infty$, allora $m(\cap_i A_i) = \lim m(A_i)$.

(x) $f \in \mathcal{M}$ se e solo se $f^{-1}(U) \in \mathcal{M}$ per ogni aperto $U \subseteq \mathbb{R}$.

(xi) Gli insiemi aperti e gli insiemi chiusi sono misurabili. Le intersezioni numerabili di insiemi aperti e le unioni numerabili di insiemi chiusi sono misurabili.

(xii) Q è trascurabile se e solo se $Q \in \mathcal{M}$ e $m(Q) = 0$.

(xiii) La misura di Lebesgue (su \mathbb{R}^n) è invariante per traslazioni: se $A \in \mathcal{M}(\mathbb{R}^n)$ e $x_0 \in \mathbb{R}^n$ allora $x_0 + A \in \mathcal{M}(\mathbb{R}^n)$ e $m(x_0 + A) = m(A)$.

Dimostrazione (i): Il 'solo se' è banale (si prenda $g := |f|$); il 'se' è conseguenza della Proposizione 4.41.

(ii): Si consideri l'affermazione sul reciproco (le altre sono ovvie conseguenze della Proposizione 4.30) e sia $f_k \rightarrow f$ (q.o.). Definiamo $g_k(x) = 0$ se $f_k(x) = 0$ e $g_k(x) := 1/f_k(x)$ altrimenti. Allora $g_k \rightarrow 1/f(x)$ (q.o.) e dunque $1/f \in \mathcal{M}$.

(iii): Fissiamo una funzione $h \in \mathcal{L}^1$ con $h(x) > 0$ per ogni x (si veda la nota 75). Sia ora $g_k := hf_k/(h + |f_k|)$. Allora: g_k è misurabile [per il punto (ii)]; $|g_k| < h$ e quindi, per il punto (i), $g_k \in \mathcal{L}^1$; $g_k \rightarrow g := hf/(h + |f|)$ (q.o.) e [convergenza dominata] $g \in \mathcal{L}^1$. In particolare $g \in \mathcal{M}$ e poiché $f = hg/(h - |g|)$, $f \in \mathcal{M}$ [punto (ii)].

(iv): La funzione caratteristica di $E \setminus A$ è data da $1 - \chi_A$.

(v): Caso finito: $\chi_{A_1 \cap \dots \cap A_k} = \prod_{i=1}^k \chi_{A_i}$ [e si usi il punto (ii)]. Caso numerabile: $\chi_{\cap_i A_i} = \lim_k \prod_{i=1}^k \chi_{A_i}$ [e punto (iii)].

(vi): Caso finito: $\chi_{A_1 \cup \dots \cup A_k} = 1 - \prod_{i=1}^k (1 - \chi_{A_i}) = \sup_{1 \leq i \leq k} \chi_{A_i}$ [e punto (ii)]; Caso numerabile: $\chi_{\cup_i A_i} = \lim_k \sup_{1 \leq i \leq k} \chi_{A_i}$.

(vii): Sia $A = \cup_i A_i$ che, per il punto precedente, è misurabile e si osservi che $\chi_A = \sum_i \chi_{A_i}$. Ci sono due casi: (a) $m(A) < \infty$, (b) $m(A) = \infty$. Caso (a): sia $f := \chi_A$ e $f_k = \sum_{i=1}^k \chi_{A_i}$. Allora $f_k \uparrow f$ e $m(A) < \infty$ significa che $f \in \mathcal{L}^1$ e quindi $\int f_k \leq \int f < \infty$. Dunque [per convergenza monotona ed osservando che, per ogni k , $\chi_{A_1 \cup \dots \cup A_k} = \sum_{i=1}^k \chi_{A_i} \in \mathcal{L}^1 \implies m(A_1 \cup \dots \cup A_k) = \sum_{i=1}^k m(A_i)$] $\lim \int f_k = \sum_{i=1}^{\infty} m(A_i) = \int f = m(A)$. Caso (b): sia $B_i := A_i \cap \{x : |x|_{\infty} \leq i\}$ cosicché i B_i sono disgiunti, $A = \cup_i B_i$ e $\chi_{B_1 \cup \dots \cup B_k} \in \mathcal{L}^1$ per ogni k . La tesi è allora conseguenza immediata del punto (iii) dell'Osservazione 4.39.

⁶⁷Esercizio 4.22.

⁶⁸Si ricordi la Proposizione 4.30.

⁶⁹Naturalmente questo include il caso finito (prendendo, da un certo punto in poi, \mathbb{R}^n).

(viii) Sia $B_1 := A_1$ e, per $i \geq 2$, $B_i := A_i \setminus A_{i-1}$ e si usi il punto precedente.

(ix) La tesi è conseguenza del teorema di convergenza dominata.

[Poiché $f_k := \chi_{A_1 \cap \dots \cap A_k} \leq \chi_{A_1} \in \mathcal{L}^1$ e $f_k \rightarrow \chi_{\cap_i A_i}$].

(x): Consideriamo prima il ‘se’. Dall’ipotesi segue, in particolare, che $f^{-1}((a, \infty)) \in \mathcal{M}$ per ogni $a \in \mathbb{R}$ e quindi anche⁷⁰ $f^{-1}((-\infty, a]) = [f^{-1}((a, \infty))]^c \in \mathcal{M}$. Dunque $f^{-1}([a, b]) \in \mathcal{M}$ per ogni $a < b$. Per ogni $k \geq 1$ e per ogni $i \in \mathbb{Z}$ sia $E_i := f^{-1}([\frac{i-1}{k}, \frac{i}{k}))$ e definamo

$$f_k := \sum_{i=-\infty}^{\infty} \frac{i-1}{k} \chi_{E_i} := \lim_{j \rightarrow \infty} \sum_{i=-j}^j \frac{i-1}{k} \chi_{E_i}. \quad (4.148)$$

Innanzitutto per ogni x solo un termine della serie è diverso da 0 (e quindi la serie converge per ogni x). Poi essendo E_i misurabile lo è anche f_k (essendo il limite di funzioni misurabili). Inoltre, per ogni x esiste un unico E_i che contiene x ed allora $0 \leq f(x) - f_k(x) = f(x) - \frac{i-1}{k} < 1/k$. Dunque $f_k \uparrow f$ per ogni x e quindi $f \in \mathcal{M}$.

Per dimostrare il ‘solo se’ ci serviremo del seguente elementare fatto topologico:

Lemma 4.54 *Ogni aperto non vuoto di \mathbb{R}^n è una unione numerabile di cubi aperti.*

Dimostrazione Sia $\emptyset \neq A \subseteq \mathbb{R}^n$ un aperto e sia $N := \{x^{(i)} : i \geq 1\}$ un insieme denso e numerabile in A (per esempio $N = A \cap \mathbb{Q}^n$). Si consideri la famiglia numerabile \mathcal{K} formata da tutti i cubi aperti contenuti in A , di centro $x^{(i)}$ (per qualche i) e con la lunghezza del lato razionale. È allora chiaro che $A = \cup_{K \in \mathcal{K}} K$. ■

Dimostriamo ora il ‘solo se’ del punto (x): osserviamo che, per il Lemma, è sufficiente dimostrare che sono misurabili gli insiemi della forma⁷¹ $\{f \leq a\}$ per ogni a . Sia $\{a_k\}$ una successione (strettamente) decrescente con limite a e, denotando con $\varphi_b := \inf\{f, b\}$, sia $f_k := (\varphi_{a_k} - \varphi_a)/(a_k - a)$. Allora $f_k \in \mathcal{M}$ e⁷² $\lim f_k = \chi_{f^{-1}((-\infty, a])}$. Dunque $f^{-1}((-\infty, a]) \in \mathcal{M}$.

(xi): Dal Lemma 4.54 segue che ogni aperto $A = \cup_i K_i$ con K_i cubo aperto. Se $A_k := \cup_{i=1}^k K_i$, la funzione $f_k := \chi_{A_k}$ è una funzione a scalini tale che $f_k \uparrow \chi_A$ e dunque $A \in \mathcal{M}$. Il resto della tesi segue dai punti (iv), (v) e (vi).

(xii): Se Q è trascurabile allora $\chi_Q = 0$ (q.o.) e dunque $Q \in \mathcal{M}$ e $m(Q) = 0$. Viceversa $Q \in \mathcal{M}$ e $m(Q) = 0$ significa che $\int \chi_Q = 0$ e quindi [(iii) della Proposizione 4.36] $\chi_Q = 0$ quasi ovunque ossia l’insieme Q su cui χ_Q vale 1 è un insieme trascurabile.

(xiii): Se $S \ni f_k \rightarrow \chi_A$ allora $x \rightarrow f_k(x - x_0) \in S$ e $f_k(x - x_0) \rightarrow \chi_{x_0+A}(x)$ e la tesi segue dal punto (v) della Proposizione 4.36. ■

Osservazione 4.55 (i) Dato uno spazio metrico X , una famiglia, \mathcal{A} , di sottoinsiemi di X che contenga l’insieme vuoto e che sia chiusa per complementazione e per unioni numerabili⁷³ prende il nome di σ -algebra su X . Dunque gli insiemi misurabili secondo Lebesgue sono una σ -algebra su \mathbb{R}^n che contiene gli insiemi aperti.

(ii) Sia \mathcal{A} una σ -algebra su di uno spazio metrico X . Un’applicazione $\mu : \mathcal{A} \rightarrow [0, \infty]$ tale che $\mu(\emptyset) = 0$ e tale che $\mu(\cup_i A_i) = \sum_i \mu(A_i)$ per ogni famiglia numerabile $\{A_i\}$ di insiemi in \mathcal{A} a due a due disgiunti, prende il nome di **misura** (o ‘misura σ -additiva’) su \mathcal{A} . Dunque la misura di Lebesgue m è una misura σ -additiva su \mathcal{M} .

(iii) Una misura μ su \mathcal{A} tale che ogni sottoinsieme di un insieme $A \in \mathcal{A}$ di misura μ nulla (ossia $\mu(A) = 0$) appartiene ad \mathcal{A} si dice **completa**. Dunque la misura di Lebesgue su \mathcal{M} è completa.

⁷⁰Come al solito A^c indica il complementare di A .
⁷¹Il simbolo $\{f \leq a\}$ denota l’insieme $\{x \in \mathbb{R}^n : f(x) \leq a\} = f^{-1}((-\infty, a])$ (analogamente $\{a < f < b\}$ denota l’insieme $\{x : a < f(x) < b\}$ etc.); l’affermazione nel testo deriva dall’osservazione che $\{a < f < b\} = \bigcap_{n \geq 1} \{f \leq a\}^c \cap \{f \leq b - 1/n\} \implies \{a < f < b\} \in \mathcal{M} \implies$ (per il Lemma) $f^{-1}(U) \in \mathcal{M}$ per ogni aperto U .
⁷²Infatti: $f(x) \leq a \implies f(x) \leq a_k \implies \varphi_{a_k}(x) = \varphi_a(x) = f(x) \implies f_k(x) = 0 \forall k$; se $f(x) > a \exists k_0$ t.c. $f(x) \geq a_k \forall k \geq k_0 \implies \varphi_{a_k}(x) = a_k$ e $\varphi_a(x) = a \implies f_k(x) = 1$.
⁷³Ovvero se $A \in \mathcal{A}$ allora $A^c = X \setminus A \in \mathcal{A}$ e se $A_i \in \mathcal{A}$ per $I \in \mathbb{N}$ allora $\cup_i A_i \in \mathcal{A}$.

(iv) *Non tutti i sottoinsiemi di \mathbb{R}^n sono misurabili secondo Lebesgue* come si evince dal seguente **esempio** dovuto a Vitali⁷⁴.

Sia $E := [0, 1]$ ed introduciamo su E la seguente relazione: diremo che $x \sim y$ se $x - y \in \mathbb{Q}$. È chiaro che \sim è una relazione d'equivalenza. Denotiamo con \tilde{x} la classe d'equivalenza indotta da x ($\tilde{x} := \{y \in E : \exists r \in \mathbb{Q} \text{ t.c. } y = x + r\}$) e con \tilde{V} l'insieme di tali classi d'equivalenza. Usando l'assioma della scelta possiamo definire V come il sottoinsieme di E ottenuto scegliendo per ogni $\alpha \in \tilde{V}$ un unico elemento $x_\alpha \in \alpha$. V è un sottoinsieme di E ed ha le seguenti proprietà: (a) se r_1 e r_2 sono due razionali distinti allora $r_1 + V \cap r_2 + V = \emptyset$ [infatti: se $x + r_1 = y + r_2$ con $x, y \in V$ e $r_1, r_2 \in \mathbb{Q}$ allora $x - y \in \mathbb{Q}$ e quindi x e y appartengono alla stessa classe di equivalenza ma allora, per la costruzione di V , $x = y$ da cui $r_1 = r_2$]; (b) $E \subseteq \cup_{\{r \in \mathbb{Q}_1\}} (r + V)$ dove $\mathbb{Q}_1 := \mathbb{Q} \cap [-1, 1]$ [infatti: dato $x \in E$, sia $\alpha := \tilde{x}$, allora, per costruzione esiste $x_\alpha \in V$ tale $x = x_\alpha + r$ con $r \in \mathbb{Q}$ e d'altra parte poiché sia x che x_α appartengono a $[0, 1]$ segue che $|r| \leq 1$]. Da (a) e da (b) (e dalla σ -additività e l'invarianza per traslazioni della misura di Lebesgue) segue che V non può essere misurabile cioè $V \notin \mathcal{M}$: se per assurdo V fosse misurabile la sua misura di Lebesgue non potrebbe essere nulla (se così fosse anche la misura di $r + V$ sarebbe nulla e quindi $[0, 1]$ sarebbe contenuto in un'unione numerabile di insiemi trascurabile ed avrebbe dunque misura nulla, il che non è) e, d'altra parte, la sua misura di Lebesgue non potrebbe neanche essere positiva (se così fosse la misura di Lebesgue di $A := \cup_{\{r \in \mathbb{Q}_1\}} (r + V)$ sarebbe infinita, essendo gli insiemi a due a due disgiunti, ma questa contraddice il fatto che $A \subseteq [-1, 2]$).

(v) Analogamente a quanto fatto per le funzioni caratteristiche, se $f \in \mathcal{M}(E)$ e $f \geq 0$ (q.o.) possiamo estendere la definizione di integrale ponendo

$$\int_E f := \begin{cases} \int_E f, & \text{se } f \in \mathcal{L}^1(E), \\ \infty, & \text{se } f \notin \mathcal{L}^1(E). \end{cases} \quad (4.149)$$

Come più volte osservato, nel secondo caso di (4.149), se $\mathcal{L}^1 \ni f_k \uparrow f$ (q.o.) allora $\int f_k \uparrow \infty$.

(vi) Nella dimostrazione del punto (x) si è visto che

$$f^{-1}(U) \in \mathcal{M}, \quad \forall \text{ aperto } U \subseteq \mathbb{R} \iff \{f \leq a\} \in \mathcal{M}, \quad \forall a \in \mathbb{R}. \quad (4.150)$$

Per altre condizioni equivalenti si veda 4.21.

(vii) *La composizione di funzioni misurabili con funzioni continue è misurabile:*

$$f \in \mathcal{M}(\mathbb{R}^n), \quad g \in C(\mathbb{R}, \mathbb{R}) \implies g \circ f \in \mathcal{M}(\mathbb{R}^n). \quad (4.151)$$

[Per ogni aperto $U \subseteq \mathbb{R}$, $g^{-1}(U)$ è un aperto di \mathbb{R} e quindi $(g \circ f)^{-1}(U) = f^{-1}(g^{-1}(U)) \in \mathcal{M}$.]

In relazione al Teorema di Fubini è utile il seguente

Teorema 4.56 (Tonelli) *Sia $E := E_1 \times E_2 \subseteq \mathbb{R}^n \times \mathbb{R}^m$ un rettangolo e sia $(x, y) \in E \rightarrow f(x, y)$ misurabile. Allora per quasi tutti gli $x \in E_1$ la funzione $y \rightarrow f(x, y)$ è misurabile su E_2 e se $g(x) := \int_{E_2} |f(x, y)| dy < \infty$ per quasi tutti gli $x \in E_1$ allora $g \in \mathcal{M}(E_1)$; in tal caso*

$$\int_{E_1} \left(\int_{E_2} |f(x, y)| dy \right) dx < \infty \iff f \in \mathcal{L}^1(E). \quad (4.152)$$

Naturalmente un risultato analogo è valido scambiando il ruolo della x con quello della y e quindi da (4.152) segue che

$$\int_{E_1} \left(\int_{E_2} |f(x, y)| dy \right) dx < \infty \iff \int_{E_2} \left(\int_{E_1} |f(x, y)| dx \right) dy < \infty, \quad (4.153)$$

⁷⁴Illustreremo l'esempio di Vitali in \mathbb{R} ma è immediato generalizzare tale esempio a \mathbb{R}^n (esercizio 4.28).

ed in tal caso $f \in \mathcal{L}^1(E)$ e dunque, per il Teorema di Fubini vale la (4.145).

Dimostrazione (del Teorema 4.56) Da $f \in \mathcal{M}(E)$ segue che esiste $Q \subseteq E$ trascurabile e $S(E) \ni f_k(x, y) \rightarrow f(x, y)$ per $(x, y) \in E \setminus Q$. Dal Lemma 4.49 segue che per quasi tutti gli $x \in E_1$, $S(E_2) \ni f_k(x, \cdot) \rightarrow f(x, \cdot)$ ossia $f(x, \cdot) \in \mathcal{M}(E_2)$. Supponiamo ora che $\int_{E_2} |f(x, y)| dy < \infty$ (x q.o.) e sia $f_k := \inf\{|f|, h_k\}$ dove $h_k(x, y) := k \cdot h(x, y)$ ed h è una funzione in $\mathcal{L}^1(E)$ e strettamente positiva⁷⁵. Allora $f_k \in \mathcal{L}^1(E)$ e $f_k \uparrow |f|$. Inoltre per il Teorema di convergenza monotona $g_k(x) := \int_{E_2} f_k(x, y) dy$ (essendo tale quantità limitata da $\int_{E_2} |f(x, y)| dy < \infty$) converge a $\int_{E_2} |f(x, y)| dy$ che dunque è in $\mathcal{M}(E_1)$. Applicando il Teorema di Fubini a $f_k(x, y)$ otteniamo allora

$$\int_{E_1 \times E_2} f_k = \int_{E_1} \left(\int_{E_2} f_k(x, y) dy \right) dx \leq \int_{E_1} \left(\int_{E_2} |f(x, y)| dy \right) dx,$$

e (4.152) segue o (direttamente) per convergenza monotona applicata a f_k (nel caso l'integrale iterato sia finito) oppure dal punto (iii) dell'Osservazione 4.39. ■

3.7 Integrazione su insiemi misurabili. Insiemi normali

La teoria dell'integrazione secondo Lebesgue su un rettangolo E si può ora estendere sostituendo E con un qualunque insieme misurabile in \mathbb{R}^n (ad esempio un qualunque insieme aperto o un qualunque insieme chiuso).

Definizione 4.57 Sia E un insieme misurabile di \mathbb{R}^n [$E \in \mathcal{M}(\mathbb{R}^n)$] e sia $f : E \rightarrow \mathbb{R}$. Diremo che f è misurabile su E [$f \in \mathcal{M}(E)$] se la funzione⁷⁶

$$f_E(x) := \begin{cases} f(x), & \text{se } x \in E, \\ 0, & \text{se } x \notin E, \end{cases} \quad (4.154)$$

è misurabile su \mathbb{R}^n [$f_E \in \mathcal{M}(\mathbb{R}^n)$]. Analogamente diremo che $f \in \mathcal{L}^1(E)$ se $f_E \in \mathcal{L}^1(\mathbb{R}^n)$ ed in tal caso porremo

$$\int_E f := \int_{\mathbb{R}^n} f_E. \quad (4.155)$$

Osservazione 4.58 (i) Ovviamente tale definizione estende quelle date (cioè coincide con le precedenti definizioni nel caso in cui E sia un rettangolo).

(ii) Se A e B sono due insiemi misurabili e disgiunti e $f \in \mathcal{L}^1(A) \cap \mathcal{L}^1(B)$ allora [essendo $f_{A \cup B} = f_A + f_B$] $f \in \mathcal{L}^1(A \cup B)$ e

$$\int_{A \cup B} f = \int_A f + \int_B f. \quad (4.156)$$

Tale relazione vale anche nel caso più generale in cui $A \cap B \neq \emptyset$ ma $m(A \cap B) = 0$.

(iii) Se E è un qualunque insieme misurabile, definiamo $S(E)$, le funzioni 'a scalini' su E , come l'insieme $\{f_E : f \in S(\mathbb{R}^n)\}$: tali funzioni avranno dunque la forma $f = \sum_{i=1}^N c_i \chi_{R_i \cap E}$ dove gli R_i sono rettangoli limitati e disgiunti. Le funzioni a scalini su E possono essere, in effetti, abbastanza complicate (si pensi, ad esempio, al caso in cui l'insieme E è di misura positiva ma con interno vuoto). Tale definizione permette (banalmente) di estendere tutti i teoremi sopra dimostrati (assieme alle loro dimostrazioni) al caso di $E \in \mathcal{M}$. Ad esempio si ha la seguente generalizzazione del Teorema di convergenza monotona:

Teorema 4.59 Sia $E \in \mathcal{M}(\mathbb{R}^n)$, $\mathcal{M}(E) \ni f_k \uparrow f$ q.o. in E . Allora $\lim \int_E f_k = \int_E f$.

⁷⁵ Se E è limitato si prenda $h = 1$; se E non è limitato si può prendere la funzione h che vale $1/k^{n+2}$ su $\{z \in E : k - 1 \leq |z|_\infty < k\}$ (per $k \geq 1$): $\int h = \sum_{k \geq 1} [(2k)^n - (2(k-1))^n] / k^{n+2} \leq 2^n \sum_{k \geq 1} 1/k^2 < \infty$.

⁷⁶ Si noti che con la convenzione in (4.154) la funzione caratteristica di E si può denotare con 1_E .

[Nel caso in cui $\sup_E f_k < \infty$ è il Teorema di convergenza monotona, altrimenti segue da (v) dell'Osservazione 4.55.]

(iv) Se $f \in \mathcal{M}(\mathbb{R}^n)$ allora $f \in \mathcal{M}(\mathbb{R}^n \times \mathbb{R}^m)$: se $f_k(x)$ è una funzione a scalini in \mathbb{R}^n e $f_k \rightarrow f$ q.o. in \mathbb{R}^n allora, ponendo $R_k := [-k, k]^m$, la funzione $f_k(x, y) := f_k(x) \cdot \chi_{R_k}(y)$ è una funzione a scalini in $\mathbb{R}^n \times \mathbb{R}^m$ e $f_k(x, y) \rightarrow f(x)$ q.o. in $\mathbb{R}^n \times \mathbb{R}^m$ (per il punto (ii) dell'Osservazione 4.48). Dunque se $E_1 \subseteq \mathbb{R}^n$ e $E_2 \subseteq \mathbb{R}^m$ sono insiemi misurabili allora $E_1 \times E_2$ è un insieme misurabile di $\mathbb{R}^n \times \mathbb{R}^m$.

Discuteremo ora un corollario del Teorema di Fubini particolarmente utile nelle applicazioni.

Proposizione 4.60 Sia $n \geq 2$ e $A \subseteq \mathbb{R}^{n-1}$ un insieme misurabile; siano α e β due funzioni in $\mathcal{M}(A)$ con $\alpha \leq \beta$ (q.o. in A) e sia $B \subseteq \mathbb{R}^n$ l'insieme definito come

$$B := \{(x, y) \in \mathbb{R}^{n-1} \times \mathbb{R} : \alpha(x) \leq y \leq \beta(x), \forall x \in A\}. \quad (4.157)$$

Allora: (i) $B \in \mathcal{M}(\mathbb{R}^n)$ e

$$\text{mis } B = \int_A (\beta(x) - \alpha(x)) dx. \quad (4.158)$$

(ii) Se $f \in \mathcal{L}^1(B)$ allora, per quasi tutti gli $x \in A$, la funzione $y \rightarrow f(x, y)$ appartiene a $\mathcal{L}^1([\alpha(x), \beta(x)])$, la funzione $x \rightarrow g(x) := \int_{\alpha(x)}^{\beta(x)} f(x, y) dy$ appartiene a $\mathcal{L}^1(A)$ e si ha

$$\int_B f = \int_A g := \int_A \left(\int_{\alpha(x)}^{\beta(x)} f(x, y) dy \right) dx. \quad (4.159)$$

Un insieme B della forma (4.157) (con, eventualmente, uno o tutti e due i ' \leq ' sostituiti da '<') prende il nome di **insieme normale** rispetto all'asse delle⁷⁷ y . La dimostrazione della Proposizione è un immediato corollario del Teorema di Fubini e del seguente

Lemma 4.61 Gli 'epigràfi' $\mathcal{E}_f := \{(x, y) \in \mathbb{R}^{n-1} \times \mathbb{R} : f(x) \leq y\}$ e $\mathcal{E}'_f := \{(x, y) \in \mathbb{R}^{n-1} \times \mathbb{R} : f(x) < y\}$ di una funzione misurabile $f : x \in \mathbb{R}^{n-1} \rightarrow f(x) \in \mathbb{R}$ sono insiemi misurabili di \mathbb{R}^n .

Dimostrazione Analogamente a quanto fatto nella dimostrazione del punto (x) della Proposizione 4.53, poniamo $\varphi(x, y) := \inf\{f(x), y\}$. Dal punto (iv) della Osservazione 4.58, segue che le funzioni $(x, y) \rightarrow f(x)$ e $(x, y) \rightarrow y$ sono misurabili in \mathbb{R}^n e quindi $\varphi \in \mathcal{M}(\mathbb{R}^n)$. Se $\varphi_k(x, y) := k \cdot (\varphi(x, y + \frac{1}{k}) - \varphi(x, y))$ si ha che⁷⁸ $\varphi_k \in \mathcal{M}(\mathbb{R}^n)$ e $\varphi_k \rightarrow \chi_{\mathcal{E}'_f}$ e dunque \mathcal{E}_f è misurabile (punti (iii) e (ii) della Proposizione 4.53). Ragionando nello stesso modo avendo sostituito f e y con $-f$ e $-y$ si ottiene che $\{(x, y) : f(x) \geq y\}$ è misurabile, da cui, passando al complementare, segue che anche \mathcal{E}'_f è misurabile. ■

Dimostrazione (della Proposizione 4.60) Poiché $B = \{(x, y) : \alpha_A(x) \leq y\} \cap \{(x, y) : \beta_A(x) < y\}^c \cap \{A \times \mathbb{R}\}$, dal Lemma segue che B è misurabile. La (4.158) segue dal Teorema di Tonelli⁷⁹ e la (4.159) segue dal Teorema di Fubini⁸⁰. ■

Esercizi

Esercizio 4.1 Si dimostri che se I, J sono intervalli di \mathbb{R} , allora $I \setminus J$ è dato dall'unione disgiunta di due intervalli I_j eventualmente vuoti; tramite esempi si faccia vedere che tutti i casi sono possibili.

⁷⁷Naturalmente il ruolo di x_1, \dots, x_{n-1} e y è del tutto arbitrario e analogamente si definirà un insieme normale rispetto ad un qualunque asse. In alcuni testi l'insieme B viene detto normale rispetto al piano $\{y = 0\}$: si faccia attenzione poiché nel caso $n = 2$ le due convenzioni potrebbero creare confusioni.

⁷⁸Come nella dimostrazione del punto (x) della Proposizione 4.53 prendendo $a_k = y + 1/k$.

⁷⁹Con $E_1 = \mathbb{R}^{n-1}$, $E_2 = \mathbb{R}$ e $f = \chi_B(x, y) = \chi_A(x) \chi_{[\alpha_A(x), \beta_A(x)]}(y)$.

⁸⁰ $\int_B f = \int_{\mathbb{R}^n} f_B = \int_{\mathbb{R}^{n-1}} \left(\int_{\mathbb{R}} f_B dx \right) dy = \int_A \left(\int_{\alpha(x)}^{\beta(x)} f dy \right) dx.$

Esercizio 4.2 Sia $f : [0, 1]^2 \rightarrow \mathbb{R}$ la funzione che vale 1 se $x_1 \neq 1/2$ e, per $x_1 = 1/2$ la funzione $x_2 \in [0, 1] \rightarrow f(1/2, x_2)$ vale 1 se x_2 è irrazionale e 0 se x_2 è razionale. Si dimostri che f è integrabile su $[0, 1]^2$ (mentre, come sappiamo, $x_2 \in [0, 1] \rightarrow f(1/2, x_2)$ non è integrabile su $[0, 1]$).

Esercizio 4.3 L'ipotesi che f sia integrabile su B nella proposizione 4.1 non può essere rimossa, anche se assumessimo che $g(x)$ fosse integrabile su E . Si consideri, infatti, la seguente funzione $f : K \rightarrow \mathbb{R}$ con $K := [0, 1]^2$

$$f(x_1, x_2) = \begin{cases} 1, & \text{se } x_1 \in [0, 1] \text{ è irrazionale,} \\ 2x_2 & \text{se } x_1 \text{ è razionale.} \end{cases} \quad (4.160)$$

La funzione $x_2 \rightarrow f(x_1, x_2)$ coincide o con la funzione costante 1 o con $2x_2$ che sono entrambe funzioni integrabili su $[0, 1]$ ed inoltre

$$\int_0^1 f(x_1, x_2) dx_2 = 1, \quad \forall x_1 \in [0, 1].$$

Dunque anche $g(x) := g(x_1) := 1$ è una funzione integrabile. D'altra parte (esercizio)

$$\sup \underline{S}_K(f, P) = \frac{3}{4}, \quad \inf \overline{S}_K(f, P) = \frac{5}{4}. \quad (4.161)$$

Esercizio 4.4 (Un esempio di insieme aperto non misurabile secondo Peano–Jordan) Sia Q l'insieme di tutti i punti a coordinate razionali nel cubo unitario $[0, 1]^n$. Sia $c > 0$; sia $\{r_j\}_{j \in \mathbb{N}}$ una numerazione di Q [ossia $Q = \{r_j : j \geq 1 \text{ e } j \text{ intero}\}$]; sia B_j la sfera aperta di centro r_j e raggio c/j^2 e sia B l'insieme aperto $B := \cup_{j \in \mathbb{N}} B_j$.

- (i) Si faccia vedere che $\text{mis est } B \geq 1$.
- (ii) Si faccia vedere che $\text{mis int } B \leq 2c \sum j^{-2}$ e si concluda che, se c è sufficientemente piccola, $\text{mis int } B < \text{mis est } B$ e che quindi B non è misurabile secondo Peano–Jordan.

Esercizio 4.5 (Integrali generalizzati I) Sia $A \subseteq \mathbb{R}^n$ un insieme *non limitato*. Supponiamo che per ogni $r > 0$, l'insieme $A_r := A \cap \{x \in \mathbb{R}^n : |x| \leq r\}$ sia misurabile. Sia $f : A \rightarrow \mathbb{R}$ e supponiamo che f sia integrabile su A_r per ogni $r > 0$.

Definizione 4.62 *Se esiste finito il limite*

$$\lim_{r \rightarrow \infty} \int_{A_r} |f| < \infty \quad (4.162)$$

si dice che f è integrabile su A e si pone

$$\int_A f = \lim_{r \rightarrow \infty} \int_{A_r} f. \quad (4.163)$$

- (i) Si controlli che tale definizione è ben posta.
- (ii) Dire se $f = e^{-|x|}$ è integrabile su \mathbb{R}^n (nel senso della definizione appena data) e, in caso affermativo, si stimi $\int_{\mathbb{R}^n} f$. [Suggerimento: si noti che $\exp(-|x|) \leq \exp(-\frac{1}{\sqrt{n}}|x_1|)$.]
- (iii) Sia $n = 2$ e $A = \{x \in \mathbb{R}^2 : |x_1| \leq 1\}$. Si dica se è integrabile su A la funzione $f = x_1^{100}(\cos x_2)(1 + |x_2|)^{-\frac{3}{2}}$ e, in caso affermativo, si stimi $\int_A f$.

Esercizio 4.6 (Integrali ingeneralizzati II) Sia A un insieme misurabile in \mathbb{R}^n e sia $\bar{x} \in A$. Sia $A' := A \setminus \{\bar{x}\}$ e sia $f : A' \rightarrow \mathbb{R}$. Sia B_r una sfera di raggio r centrata in \bar{x} . Supponiamo che f sia integrabile su $A'_r := A' \setminus B_r$ per ogni $r > 0$.

Definizione 4.63 *Se esiste finito il limite*

$$\lim_{r \rightarrow 0} \int_{A'_r} |f| < \infty \quad (4.164)$$

si dice che f è integrabile su A e si pone

$$\int_A f = \lim_{r \rightarrow 0} \int_{A'_r} f. \quad (4.165)$$

- (i) Si controlli che tale definizione è ben posta.
 (ii) Sia A la sfera unitaria in \mathbb{R}^n per $n \leq 3$, $\bar{x} = 0$ e $f = \frac{1}{|x|^\alpha}$. Dire per quali α , f è integrabile su A e per tali α stimare $\int_A f$.
 (iii)* Svolgere il punto (ii) nel caso di n arbitrario.

Esercizio 4.7 Dimostrare che se $A \in \mathcal{N}^k$, $B \in \mathcal{N}^h$ allora $A \times B \in \mathcal{N}^{k+h}$.

Esercizio 4.8 Dimostrare che se A è un insieme misurabile secondo Peano–Jordan, lo sono anche \bar{A} , \dot{A} e ∂A .

Esercizio 4.9 Dimostrare che $\mathbb{Q}^n \cap E$ (con E rettangolo qualunque) è un insieme trascurabile non misurabile secondo Peano–Jordan.

Esercizio 4.10 Dimostrare che se A è misurabile secondo Peano–Jordan allora A è trascurabile se e solo se $\text{mis}_n A = 0$.

Esercizio 4.11 Sia $B \in \mathcal{N}^n$ è come in (4.35). Dimostrare che è PJ misurabile.

Esercizio 4.12 Sia R un rettangolo chiuso in \mathbb{R}^n . (i) Si dimostri che, per ogni $\varepsilon > 0$, esistono N cubi aperti R_1, \dots, R_N tali che $R \subseteq E := R_1 \cup \dots \cup R_N$ e che $m(E \setminus R) \leq \varepsilon$. (ii) Si dia una stima di N in termini di ε e delle lunghezze dei lati di R .

Esercizio 4.13 Sia $R := [a_1, b_1] \times \dots \times [a_n, b_n]$ e si dimostri che $R = \cup K_j$ con K_j cubi chiusi tali che $m(K_i \cap K_j) = 0$ se $i \neq j$.

Esercizio 4.14 Sia $A \subseteq \mathbb{R}^n$ un aperto e sia $\phi \in C^1(A, \mathbb{R}^n)$. Dimostrare che ϕ è invertibile con inversa C^1 se e solo se ϕ è iniettiva e $\det \phi' \neq 0$ su A .

Esercizio 4.15 Sia $F(y)$ la trasformazione inversa della ϕ del Teorema 4.6 cioè

$$F : y \in B := \phi(A) \rightarrow x = F(y) \in A. \quad (4.166)$$

- (i) Dimostrare che la coppia B e F verifica le ipotesi del Teorema 4.6.
 (ii) Dimostrare che se g è una funzione integrabile su $A = F(B)$, allora si ha

$$\int_A g(x) dx = \int_B g \circ F(y) \cdot \left| \det \frac{\partial F}{\partial y}(y) \right| dy. \quad (4.167)$$

Esercizio 4.16 (i) Si dimostri che

$$\int_{-\infty}^{\infty} e^{-x^2} dx = \sqrt{\pi} \quad (4.168)$$

(ii) Si dimostri che

$$\int_{\mathbb{R}^n} e^{-|x|^2} dx = \pi^{\frac{n}{2}}. \quad (4.169)$$

Esercizio 4.17 Per ogni $m \in \mathbb{N}$, si calcoli $I_m := \int_0^{\infty} x^m e^{-x^2} dx$.

Esercizio 4.18 Sia $A := B'_R := \{x \in \mathbb{R}^2 : 0 < |x| < R\}$ e sia

$$\phi(x) := (x_1^2 - x_2^2, 2x_1x_2).$$

(i) Si dimostri che $\delta := |\det \phi'| > 0$ per ogni $x \in A$ ma che

$$\int_A \delta dx = 2\pi R^4 \\ \phi(A) = B'_{R^2} \quad \implies \quad m(\phi(A)) = \pi R^4.$$

- (ii) Si spieghi perchè non vale la tesi del Teorema 4.6.
 (iii) Si trovi un insieme A' su cui, invece, valga la tesi del Teorema 4.6.

Esercizio 4.19 Dimostrare in dettaglio la Proposizione 4.30.

Esercizio 4.20 Dimostrare in dettaglio i punti (i) e (ii) della Proposizione 4.30 nel caso in cui $S(E)$ venga sostituito da $\mathcal{F}(E)$.

Esercizio 4.21 Si dimostri che ‘ $f^{-1}(U) \in \mathcal{M}$ per ogni aperto $U \subseteq \mathbb{R}$ ’ è equivalente ad ognuna delle seguenti proprietà:

- (i) $\{f \geq a\} \in \mathcal{M}, \forall a \in \mathbb{R}$; (ii) $\{f > a\} \in \mathcal{M}, \forall a \in \mathbb{R}$;
 (iii) $\{a \leq f < b\} \in \mathcal{M}, \forall a, b \in \mathbb{R}$; (iv) $\{a < f \leq b\} \in \mathcal{M}, \forall a, b \in \mathbb{R}$.

Esercizio 4.22 Si dimostri che l’insieme $A := \{(x, y) \in \mathbb{R}^2 : 0 < x < 1 \text{ e } 0 < y < 1/\sqrt{x}\}$ è misurabile secondo Lebesgue e se ne calcoli la sua misura.

Esercizio 4.23 Dimostrare che ogni insieme misurabile secondo Peano–Jordan è misurabile secondo Lebesgue.

Esercizio 4.24 Sia f come nel punto (v) dell’Osservazione 4.34, sia $Q \subseteq \mathbb{R}^n$ un insieme di misura nulla, sia f_0 una qualunque funzione che vale 0 al di fuori di Q . Si dimostri che $f + f_0$ non è integrabile secondo Riemann.

Esercizio 4.25 Siano E_0 e A come nel punto (v) dell’Osservazione 4.34 e si dimostri che l’insieme $E_0 \setminus A$ non è un insieme trascurabile.

Esercizio 4.26 Sia $A \in \mathcal{M}(\mathbb{R}^n)$ con $m(A) = \infty$ e siano $E_j \in \mathcal{M}$ tali che $E_j \subseteq E_{j+1}$, $m(E_j) < \infty$ e $\cup_i E_j = \mathbb{R}^n$. Si dimostri che $\infty > m(A \cap E_j) \uparrow \infty$.

Esercizio 4.27 Si dia un esempio di insiemi misurabili A_k in \mathbb{R}^n tali che $A_{k+1} \subseteq A_k$ ma $m(\cap A_k) \neq \lim m(A_k)$.

Esercizio 4.28 Si generalizzi l’esempio di Vitali (punto (iv) dell’Osservazione 4.55) a \mathbb{R}^n .

Esercizio 4.29 Si dimostri la (4.156) nel caso in cui $A, B \in \mathcal{M}$ e $m(A \cap B) = 0$.

Esercizio 4.30 (L’insieme ternario di Cantor) L’insieme ternario di Cantor K è un sottoinsieme di $[0, 1]$ con le seguenti proprietà: (a) K è chiuso, non ha punti isolati ed è totalmente sconnesso (ossia non contiene nessun intervallo); (b) K ha la potenza del continuo (ossia esiste un’applicazione 1–1 da K in \mathbb{R}); (c) la misura di Lebesgue di K è zero.

Costruzione di K : (i) ad un qualunque intervallo chiuso (limitato e di lunghezza positiva) I associamo due intervalli chiusi I_1^k e I_2^k ottenuti rimuovendo da I l’intervallo aperto che ha lo stesso centro di I e lunghezza uguale ad un terzo di quella di I .

(ii) Definiamo ora intervalli I_j^k per $k \geq 0$ e $1 \leq j \leq 2^k$: $I_1^0 := [0, 1]$ e per $k \geq 1$ gli I_j^k sono gli intervalli (ordinati) ottenuti dagli I_j^{k-1} come descritto nel punto (i). Per costruzione gli intervalli I_j^k sono, per ogni k fissato, chiusi, disgiunti e di lunghezza $1/3^k$ cosicché $\sum_{j=1}^{2^k} \text{mis}(I_j^k) = (2/3)^k$.

(iii) Sia $I^k := \cup_{j=1}^{2^k} I_j^k$. Chiaramente $I^{k+1} \subseteq I^k$. Definiamo $K := \cap_{k \geq 0} I^k$.

Si dimostrino le proprietà (a), (b) e (c) sopra enunciate.

Esercizio 4.31 (Funzioni semplici) Una funzione $f \in \mathcal{M}(\mathbb{R}^n)$ si dice **semplice** se assume un numero finito di valori. Esempi di funzioni semplici sono le funzioni a scalini oppure combinazioni lineari finite di funzioni caratteristiche di insiemi misurabili. Dimostrare le seguenti asserzioni:

- (i) Una funzione è misurabile se e solo se è limite (q.o.) di funzioni semplici.
 (ii) Per ogni funzione $f \in \mathcal{L}^1$ esiste una successione di funzioni semplici $f_k \in \mathcal{L}^1$ tale che $f_k \rightarrow f$ (q.o.) e $\|f - f_k\|_{\mathcal{L}^1} \rightarrow 0$.

Esercizio 4.32 Sia $f \in \mathcal{L}^1(E)$. Dimostrare: (i) $\lim_{r \rightarrow \infty} m(\{|f| \geq r\}) = 0$; (ii) $\lim_{r \rightarrow \infty} \int_{\{|f| \geq r\}} |f| = 0$.

Esercizio 4.33 Sia $f \in \mathcal{L}^1(E)$. Dimostrare che per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che per ogni $A \in \mathcal{M}(E)$ con $m(A) \leq \delta$ si ha che $\int_A |f| \leq \varepsilon$.

Esercizio 4.34 Sia $f \in \mathcal{L}^1$ e $B_j \in \mathcal{M}$ tali che $m(B_j) \rightarrow 0$. Dimostrare che $\int_{B_j} f \rightarrow 0$.

Esercizio 4.35 Dimostrare che se $f \in \mathcal{L}^1(\mathbb{R})$ allora esiste un insieme illimitato di misura positiva E tale che $\lim f(x_k) = 0$ per ogni successione $x_k \in E$ tale che $\lim |x_k| = \infty$.

Esercizio 4.36 Trovare $f \in \mathcal{L}^1(\mathbb{R})$ ed insiemi illimitati di misura positiva E_{\pm} tale che $\lim f(x_k) = \pm\infty$ per ogni successione $x_k \in E_{\pm}$ tale che $\lim |x_k| = \infty$.

Esercizio 4.37 Dimostrare che il grafico $G_f := \{(x, y) \in \mathbb{R}^{n-1} \times \mathbb{R} : y = f(x)\}$ di una funzione misurabile $f : \mathbb{R}^{n-1} \rightarrow \mathbb{R}$ è un insieme trascurabile.

Esercizio 4.38 Trovare funzioni misurabili f_k su $[0, 1]$ tali che: per ogni k , f_k è q.o. continua; $\sup_k f_k(x) = \infty$ per ogni $x \in \mathbb{Q} \cap [0, 1]$; esiste $M > 0$ tale che $|f_k(x)| \leq M$ per ogni $x \in [0, 1] \setminus \mathbb{Q}$.

Complementi

Complemento 4.1: Prodotto vettoriale in \mathbb{R}^3

Il prodotto vettoriale in \mathbb{R}^3 è una *operazione bilineare ed antisimmetrica* che associa ad una coppia di vettori $v, w \in \mathbb{R}^3$ un terzo vettore $v \times w \in \mathbb{R}^3$ di coordinate⁸¹

$$(v \times w)_i := \det[e^{(i)}, v, w], \quad (4.170)$$

ossia

$$v \times w = \sum_{i=1}^3 e^{(i)} \det[e^{(i)}, v, w], \quad (e_j^{(i)} := \delta_{ij}). \quad (4.171)$$

Si noti che, per definizione di i -esima coordinata, si ha

$$e^{(i)} \cdot (v \times w) = \det[e^{(i)}, v, w]. \quad (4.172)$$

Dati due vettori linearmente indipendenti v e w definiamo $\pi_{v,w}$ il piano generato da v e w

$$\pi_{v,w} := V(v, w) := \{y \in \mathbb{R}^3 : y = av + bw \text{ con } a, b \in \mathbb{R}\}. \quad (4.173)$$

Ed infine, dati tre vettori indipendenti $v^{(i)} \in \mathbb{R}^3$, definiamo

$$\text{Or}(v^{(1)}, v^{(2)}, v^{(3)}) := \left\{ (y^{(1)}, y^{(2)}, y^{(3)}) \in \mathbb{R}^3 \times \mathbb{R}^3 \times \mathbb{R}^3 : \det[y^{(1)}, y^{(2)}, y^{(3)}] \det[v^{(1)}, v^{(2)}, v^{(3)}] > 0 \right\}. \quad (4.174)$$

È facile vedere che⁸² l'appartenenza a $\text{Or}(v^{(1)}, v^{(2)}, v^{(3)})$ è una *relazione d'equivalenza*; una base (ordinata) $(y^{(1)}, y^{(2)}, y^{(3)})$ in $\text{Or}(v^{(1)}, v^{(2)}, v^{(3)})$ si dice che *definisce lo stesso orientamento di* (o anche *che è equiorientata con*) $(v^{(1)}, v^{(2)}, v^{(3)})$.

Si verificano facilmente le seguenti proprietà del prodotto vettoriale $v \times w$, di $\pi_{v,w}$ e dell'orientamento $\text{Or}(v^{(1)}, v^{(2)}, v^{(3)})$:

Proprietà del prodotto vettoriale

- 1) $v \times w = (v_2w_3 - v_3w_2)e^{(1)} - (v_1w_3 - v_3w_1)e^{(2)} + (v_1w_2 - v_2w_1)e^{(3)}$.
- 2) L'operazione $(v, w) \rightarrow v \times w$ è bilineare⁸³ ed antisimmetrica⁸⁴.
- 3) $e^{(\sigma_1)} = \varepsilon_{\sigma} e^{(\sigma_2)} \times e^{(\sigma_3)}$, dove σ è una permutazione di $\{1, 2, 3\}$ e ε_{σ} è il segno di tale permutazione⁸⁵.
- 4) $w \cdot v^{(1)} \times v^{(2)} = \det[w, v^{(1)}, v^{(2)}]$ e quindi $w \cdot v^{(1)} \times v^{(2)} = v^{(1)} \cdot v^{(2)} \times w = v^{(2)} \cdot w \times v^{(1)}$.
- 5) $v \times w \perp v$, $v \times w \perp w$.

⁸¹In alcuni testi (soprattutto di fisica) il prodotto vettoriale tra v e w viene denotato con il simbolo ' $v \wedge w$ '.

⁸²Esercizio 4.40.

⁸³Cioè lineare sia in v che in w .

⁸⁴Cioè $v \times w = -w \times v$.

⁸⁵Si ricordi la nota ?? di §4.

6) $v \times w \neq 0$ se e solo se $\text{rango}[v, w] = 2$ se e solo se v e w sono linearmente indipendenti (cioè ‘non appartengono alla stessa retta’ o equivalentemente ‘non sono paralleli’).

$$7) |v \times w|^2 = |v|^2|w|^2 - (v \cdot w)^2.$$

8) Se U è una matrice ortogonale (3×3), $|Uv \times Uw| = |v \times w|$.

9) Se $U \in SO(3)$, allora $Uv \times Uw = U(v \times w)$.

10) $(v \times w, v, w) \in \text{Or}(e^{(1)}, e^{(2)}, e^{(3)})$.

11) Se v e w sono indipendenti, allora $\pi_{v,w} = \pi_{v \times w}$.

12) $|v \times w|$ coincide con l’area del parallelogramma generato da v e w (per una formulazione più precisa si veda 4.39).

Dimostrazione della 9): usando la 4) e la (4.172), per ogni i si ha:

$$\begin{aligned} e^{(i)} \cdot U(v \times w) &= U^T e^{(i)} \cdot (v \times w) = \det[U^T e^{(i)}, v, w] = \det[U^{-1} e^{(i)}, v, w] \\ &= \det(U^{-1}[e^{(i)}, Uv, Uw]) = \det[e^{(i)}, Uv, Uw] = e^{(i)} \cdot (Uv \times Uw). \quad \blacksquare \end{aligned}$$

La definizione del prodotto vettoriale può essere facilmente ricordata se si nota che la seguente *identità formale*:

$$v \times w = \det \begin{pmatrix} e^{(1)} & e^{(2)} & e^{(3)} \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{pmatrix} \quad (4.175)$$

dove il determinante formale verrà pensato sviluppato secondo la prima ‘riga’.

Complemento 4.2: Rette e piani tangenti

Complemento 4.3: Volumi, aree e curvatura

Complemento 4.4: Dimostrazione del Teorema della divergenza in \mathbb{R}^n

Il Teorema della divergenza è equivalente alla seguente affermazione:

Sia $A := \{x \in \mathbb{R}^n : \phi(x) < 0\}$ un insieme regolare in \mathbb{R}^n . Allora

$$\int_A \frac{\partial f}{\partial x_i} dx = \int_{\partial A} f \nu_i d\sigma_{n-1}, \quad (\forall f \in C^1(\bar{A}, \mathbb{R}), \forall 1 \leq i \leq n). \quad (4.176)$$

Infatti se vale tale affermazione il Teorema della divergenza verrà ottenuto applicando la (4.176) alle componenti F_i e sommando su i , mentre dal Teorema della divergenza segue immediatamente la (4.176) ponendo $F_j = 0$ se $j \neq i$ e $F_i := f$.

Nel corso della dimostrazione faremo uso del seguente notevole risultato

Lemma 4.64 (Partizione dell’unità) Sia D un compatto di \mathbb{R}^n e sia $\{V_i\}_{i \in J}$ un ricoprimento aperto⁸⁶ di D . Allora esistono N funzioni f_1, \dots, f_N di classe⁸⁷ $C_0^\infty(\mathbb{R}^n)$, che godono delle seguenti proprietà:

$$\begin{aligned} 1) \quad & \forall 1 \leq j \leq N, \exists i \in J : \text{supp}(f_j) \subseteq V_i; \\ 2) \quad & 0 \leq f_j(x) \leq 1, \forall x \in \mathbb{R}^n, 1 \leq j \leq N; \\ 3) \quad & \sum_{j=1}^N f_j(x) = 1, \forall x \in D; \quad \sum_{j=1}^N f_j(x) \leq 1, \forall x \in \mathbb{R}^n. \end{aligned} \quad (4.177)$$

⁸⁶Cioè V_i sono insiemi aperti e $\cup_{i \in J} V_i \supseteq D$.

⁸⁷Si ricorda che $C_0^\infty(\mathbb{R}^n)$ denota l’insieme delle funzioni $C^\infty(\mathbb{R}^n)$ a supporto compatto e che il supporto di f , $\text{supp}(f)$, è la chiusura dell’insieme $\{x \in \mathbb{R}^n : f(x) \neq 0\}$.

Una famiglia $\{f_j\}_{j=1,\dots,N}$ che goda delle 3 proprietà sopra elencate si chiama una **partizione dell'unità su D subordinata al ricoprimento $\{V_i\}$** .

Dimostrazione Per ogni $x \in D$ scegliamo un indice $i := i(x)$ in J tale che $x \in V_{i(x)}$. Essendo gli insiemi V_i aperti, è possibile scegliere, per ogni $x \in D$, un numero positivo $r := r(x)$ tale che, se $K_r(x)$ denota il cubo aperto di centro x e lato $2r$ (cioè l'insieme $\{y \in \mathbb{R}^n : |y_i - x_i| < r, \forall 1 \leq i \leq n\}$),

$$\overline{K_{r(x)}}(x) \subseteq K_{R(x)}(x) \subseteq \overline{K_{R(x)}}(x) \subseteq V_{i(x)}, \quad \text{dove } R(x) := 2r(x). \quad (4.178)$$

Gli insiemi $K_{r(x)}(x)$, al variare di $x \in D$ formano un ricoprimento aperto di D . Essendo tale insieme compatto è possibile estrarre un sottoricoprimento finito, esistono cioè N punti in D , $x^{(1)}, \dots, x^{(N)}$, tali che

$$D \subseteq K_1 \cup \dots \cup K_N, \quad K_j := K_{r(x^{(j)})}(x^{(j)}). \quad (4.179)$$

Inoltre, se denotiamo $K'_j := K_{R(x^{(j)})}(x^{(j)})$, e $i_j := i(x^{(j)})$, da (4.178) segue che

$$\overline{K_j} \subseteq K'_j \subseteq \overline{K'_j} \subseteq V_{i_j}. \quad (4.180)$$

Usando le funzioni χ dell'Esempio A.38, possiamo costruire N funzioni $\psi_j \in C_0^\infty(\mathbb{R}^n)$ tali che

$$0 \leq \psi_j \leq 1, \quad \psi_j := 1 \text{ su } K_j, \quad \text{supp}(\psi_j) = \overline{K'_j}. \quad (4.181)$$

Poniamo ora

$$\begin{aligned} f_1 &:= \psi_1, \\ f_2 &:= (1 - \psi_1)\psi_2, \\ &\dots\dots\dots \\ f_N &:= (1 - \psi_1)(1 - \psi_2) \dots (1 - \psi_{N-1})\psi_N. \end{aligned} \quad (4.182)$$

Da tale definizione segue immediatamente che $f_j \in C_0^\infty(\mathbb{R}^n)$ e che $\text{supp}(f_j) \subseteq \overline{K'_j} \subseteq V_{i_j}$ e quindi valgono le prime due proprietà, 1) e 2), enunciate nella proposizione. Inoltre da un elementare fatto algebrico segue che⁸⁸

$$\sum_{j=1}^N f_j = 1 - \prod_{j=1}^N (1 - \psi_j) \quad (4.183)$$

quindi se $x \in D$, x appartiene a qualche K_j e da (4.183) segue anche la proprietà 3). ■

Dimostrazione (del Teorema C.9) In vista dell'Osservazione C.10 punto (iii), dimostreremo la (4.176) per una $f \in C^1(\overline{A}, \mathbb{R})$. Fissato $\tilde{x} \in E := \partial A$, siano $j, g, K_{r,\rho}$ e \hat{K}_r come al punto (ii) di (E6), § C.1. Introduciamo anche le seguenti notazioni: $\tilde{x} := (x_1, \dots, \tilde{x}_j, \dots, x_n)$ cosicché $g : \tilde{x} \in \hat{K}_r \subseteq \mathbb{R}^{n-1} \rightarrow \mathbb{R}$; $G(\tilde{x}) := (x_1, \dots, g(\tilde{x}), \dots, x_n)$ (qui ed in seguito useremo questa notazione sottintendendo che la funzione g sta al posto di x_j), dunque l'insieme $\partial A \cap K_{r,\rho} = \{G(\tilde{x}) : \tilde{x} \in \hat{K}_r\}$. Infine, eventualmente diminuendo il valore di r possiamo assumere che $\frac{\partial \phi}{\partial x_j}(x) \neq 0$ per ogni $x \in K_{r,\rho}$. Divideremo la dimostrazione in vari passi.

1) *Calcoliamo la normale esterna $\nu(x)$ per $x \in K_{r,\rho} \cap \partial A$ in termini della funzione $g(\tilde{x})$.* Prendendo la derivata parziale rispetto a x_i per $i \neq j$ nella relazione $\phi(x_1, \dots, g, \dots, x_n) = 0$ (valida in \hat{K}_r) si ottiene

$$\frac{\partial \phi}{\partial x_i}(G(\tilde{x})) + \frac{\partial \phi}{\partial x_j}(G(\tilde{x})) \frac{\partial g}{\partial x_i}(\tilde{x}) = 0 \quad \implies \quad \frac{\partial \phi}{\partial x_i} = - \frac{\partial \phi}{\partial x_j} \frac{\partial g}{\partial x_i} \quad (\tilde{x} \in \hat{K}_r, i \neq j).$$

Dunque, su $\partial A \cap K_{r,\rho}$, si ha⁸⁹

$$\nu := \frac{\nabla \phi}{|\nabla \phi|}(G(\tilde{x})) = \text{segno} \left(\frac{\partial \phi}{\partial x_j} \right) \frac{\left(-\frac{\partial g}{\partial x_1}, \dots, 1, \dots, -\frac{\partial g}{\partial x_n} \right)}{\sqrt{1 + |g'|^2}}, \quad (4.184)$$

⁸⁸Per induzione su N : per $N = 1$ la (4.183) è vera. Assumiamo la (4.183) vera per $N - 1$ e dimostriamola per N : $\sum_{j=1}^N f_j = \sum_{j=1}^{N-1} f_j + f_N = 1 - \prod_{j=1}^{N-1} (1 - \psi_j) + \left(\prod_{j=1}^{N-1} (1 - \psi_j) \right) \psi_N = 1 - \left(\prod_{j=1}^{N-1} (1 - \psi_j) \right) (1 - \psi_N) = 1 - \prod_{j=1}^N (1 - \psi_j)$.

⁸⁹Se α è un numero reale diverso da zero, $\text{segno}(\alpha) = \alpha/|\alpha|$; naturalmente nella formula (4.184) ν e ϕ sono calcolate in $G(\tilde{x}) = (x_1, \dots, g(\tilde{x}), \dots, x_m)$ e g è calcolata in $\tilde{x} \in \hat{K}_r$: quando non vi sia ambiguità ometteremo di indicare gli argomenti delle varie funzioni.

dove, naturalmente, l'1 sta al j° posto e $g' := \partial_{\tilde{x}}g$.

2) Calcoliamo la quantità in (C.10) su $\partial A \cap K_{r,\rho}$ (cosa necessaria per trovare un'espressione "esplicita" di σ_{n-1} su $\partial A \cap K_{r,\rho}$). Qui, come già detto, l'inclusione φ coincide con G e la variabile $(n-1)$ -dimensionale u con $\tilde{x} \in \hat{K}_r$. In generale, nel caso $k = n-1$, si ha

$$\sum_{1 \leq i_1 < \dots < i_k \leq n} \left| \det \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial u} \right|^2 = \sum_{i=1}^n \left| \det \frac{\partial(\varphi_1, \dots, \hat{\varphi}_i, \dots, \varphi_n)}{\partial u} \right|^2. \quad (4.185)$$

Inoltre, nel caso presente (e con le solite convenzioni), la matrice $\frac{\partial \varphi}{\partial u}$ è una matrice $n \times (n-1)$ avente la seguente forma

$$\begin{aligned} \frac{\partial(\varphi_1, \dots, \varphi_n)}{\partial u} &:= \frac{\partial G}{\partial \tilde{x}} := \frac{\partial(x_1, \dots, g, \dots, x_n)}{\partial(x_1, \dots, \hat{x}_j, \dots, x_n)} = \\ &= \begin{pmatrix} 1 & 0 & \dots & 0 & 0 & \dots & 0 \\ 0 & 1 & & 0 & 0 & \dots & 0 \\ \vdots & & \ddots & 0 & 0 & \dots & 0 \\ 0 & 0 & & 1 & 0 & \dots & 0 \\ \frac{\partial g}{\partial x_1} & \dots & \frac{\partial g}{\partial x_{j-1}} & \frac{\partial g}{\partial x_{j+1}} & \dots & \frac{\partial g}{\partial x_n} \\ 0 & 0 & \dots & 0 & 1 & & 0 \\ \vdots & \vdots & & \vdots & & \ddots & \\ 0 & 0 & \dots & 0 & 0 & & 1 \end{pmatrix}, \end{aligned} \quad (4.186)$$

(la riga che contiene le derivate di g è la j -esima). Quindi la matrice $\frac{\partial(\varphi_1, \dots, \hat{\varphi}_j, \dots, \varphi_n)}{\partial \tilde{x}}$ è la matrice identità ed ha determinante uguale ad uno, mentre se $i \neq j$, scambiando la i -ma riga con la j -esima riga si trova

$$\left| \det \frac{\partial(\varphi_1, \dots, \hat{\varphi}_i, \dots, \varphi_n)}{\partial \tilde{x}} \right| = \left| \frac{\partial g}{\partial x_i} \right|, \quad (i \neq j).$$

Dunque

$$\begin{aligned} \sum_{1 \leq i_1 < \dots < i_k \leq n} \left| \det \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial u} \right|^2 &= \sum_{i=1}^n \left| \det \frac{\partial(\varphi_1, \dots, \hat{\varphi}_i, \dots, \varphi_n)}{\partial u} \right|^2 \\ &= 1 + \sum_{i \neq j} \left| \frac{\partial g}{\partial x_i} \right|^2 := 1 + |g'|^2. \end{aligned} \quad (4.187)$$

3) Supponiamo che $f \in C^1(\bar{A}, \mathbb{R})$ abbia supporto contenuto in $K_{r,\rho}$, rettangolo "centrato" in $\bar{x} \in \partial A$ [come in 1)] e dimostriamo che vale (4.176). Vi sono due casi: o $\frac{\partial \phi}{\partial x_j} > 0$ (su $K_{r,\rho}$) oppure $\frac{\partial \phi}{\partial x_j} < 0$; i due casi si trattano in maniera simile e per fissare le idee assumeremo $\frac{\partial \phi}{\partial x_j} > 0$. In tal caso $\nu_j > 0$ [vedi (4.184)] il che significa⁹⁰ che i punti interni di A si trovano "sotto" il grafico di g o, più precisamente, $A \cap K_{r,\rho} = \{x \in \mathbb{R}^n : \bar{x}_j - \rho < x_j < g(\tilde{x}), \tilde{x} \in \hat{K}_r\}$. Dunque poiché stiamo assumendo che $\text{supp } f \subseteq K_{r,\rho}$, integrando prima rispetto a x_j , troviamo

$$\int_A \frac{\partial f}{\partial x_i} dx = \int_{\hat{K}_r} \left(\int_{\bar{x}_j - \rho}^{g(\tilde{x})} \frac{\partial f}{\partial x_i} dx_j \right) d\tilde{x}. \quad (4.188)$$

Se $i = j$, dal Teorema fondamentale del calcolo (in una variabile) e dal fatto che

$$f(x_1, \dots, \bar{x}_j - \rho, \dots, x_n) = 0$$

per⁹¹ $\tilde{x} \in \hat{K}_r$, si ha che

$$\begin{aligned} \int_{\bar{x}_j - \rho}^{g(\tilde{x})} \frac{\partial f}{\partial x_j} dx_j &= f(x_1, \dots, g(\tilde{x}), \dots, x_n) - f(x_1, \dots, \bar{x}_j - \rho, \dots, x_n) \\ &= f(x_1, \dots, g(\tilde{x}), \dots, x_n). \end{aligned} \quad (4.189)$$

⁹⁰Si ricordi la discussione nella nota 16.

⁹¹Si noti che i punti della forma $(x_1, \dots, \bar{x}_j - \rho, \dots, x_n)$ con $\tilde{x} \in \hat{K}_r$ appartengono alla frontiera di $K_{r,\rho}$ e dunque, poiché $\text{supp } f \subseteq K_{r,\rho}$, $f(x_1, \dots, \bar{x}_j - \rho, \dots, x_n) = 0$.

Poiché (per (4.184) e per le nostre ipotesi) $\nu_j = (1 + |g'|)^{-\frac{1}{2}}$, da (4.187), dalla definizione (C.12), da (4.188) e (4.189) segue che

$$\begin{aligned} \int_A \frac{\partial f}{\partial x_j} dx &= \int_{\tilde{K}_r} f(x_1, \dots, g(\tilde{x}), \dots, x_n) d\tilde{x} \\ &= \int_{\tilde{K}_r} f(x_1, \dots, g(\tilde{x}), \dots, x_n) \frac{1}{\sqrt{1 + |g'|^2}} \sqrt{1 + |g'|^2} d\tilde{x} \\ &= \int_{\partial A \cap K_{r,\rho}} f \nu_j d\sigma_{n-1} = \int_{\partial A} f \nu_j d\sigma_{n-1}. \end{aligned} \quad (4.190)$$

Sia ora $i \neq j$. In tal caso

$$\partial_{x_i} \int_{\tilde{x}_j - \rho}^{g(\tilde{x})} f(x) dx_j = f(x_1, \dots, g(\tilde{x}), \dots, x_n) \frac{\partial g}{\partial x_i}(\tilde{x}) + \int_{\tilde{x}_j - \rho}^{g(\tilde{x})} \frac{\partial f}{\partial x_i}(x) dx_j. \quad (4.191)$$

D'altra parte, ancora per il Teorema fondamentale del calcolo e per il fatto che $\text{supp } f \subseteq K_{r,\rho}$, se denotiamo con $x_{\pm r}$ un punto $x \in K_{r,\rho}$ con la i -ma componente x_i fissata ed uguale a $\tilde{x}_i \pm r$ e con $\tilde{x}_{\pm r}$ un punto $\tilde{x} := (x_1, \dots, \hat{x}_j, \dots, x_n) \in \tilde{K}_r$ con x_i fissato ed uguale a $\tilde{x}_i \pm r$, si trova⁹²

$$\int_{\tilde{x}_i - r}^{\tilde{x}_i + r} \partial_{x_i} \left(\int_{\tilde{x}_j - \rho}^{g(\tilde{x})} f(x) dx_j \right) dx_i = \int_{\tilde{x}_j - \rho}^{g(\tilde{x}_r)} f(x_r) dx_j - \int_{\tilde{x}_j - \rho}^{g(\tilde{x}_{-r})} f(x_{-r}) dx_j = 0, \quad (4.192)$$

che, insieme a (4.191), implica

$$\int_{\tilde{x}_i - r}^{\tilde{x}_i + r} \left(\int_{\tilde{x}_j - \rho}^{g(\tilde{x})} \frac{\partial f}{\partial x_i}(x) dx_j \right) dx_i = - \int_{\tilde{x}_i - r}^{\tilde{x}_i + r} f(x_1, \dots, g(\tilde{x}), \dots, x_n) \frac{\partial g}{\partial x_i}(\tilde{x}) dx_i. \quad (4.193)$$

Integrando su A prima rispetto a x_j , poi rispetto a x_i e poi rispetto alle altre componenti, da (4.193) segue, per $i \neq j$, che⁹³

$$\begin{aligned} \int_A \frac{\partial f}{\partial x_i} dx &= \int_{A \cap K_{r,\rho}} \frac{\partial f}{\partial x_i} dx = - \int_{\tilde{K}_r} f(x_1, \dots, g(\tilde{x}), \dots, x_n) \frac{\partial g}{\partial x_i}(\tilde{x}) d\tilde{x} \\ &= \int_{\tilde{K}_r} f(x_1, \dots, g, \dots, x_n) \frac{\left(-\frac{\partial g}{\partial x_i}\right)}{\sqrt{1 + |g'|^2}} \sqrt{1 + |g'|^2} d\tilde{x} \\ &= \int_{\tilde{K}_r} f(x_1, \dots, g, \dots, x_n) \nu_i \sqrt{1 + |g'|^2} d\tilde{x} \\ &= \int_{\partial A \cap K_{r,\rho}} f \nu_i d\sigma_{n-1} = \int_{\partial A} f \nu_i d\sigma_{n-1}. \end{aligned}$$

Dunque (4.176) vale quando il supporto di f è contenuto in un rettangolo (con le proprietà sopra elencate) centrato su un punto della frontiera di A .

4) Sia ora K un cubo aperto la cui chiusura è contenuta in A e dimostriamo (4.176) per funzioni $f \in C^1(\bar{A}, \mathbb{R})$ che abbiano supporto contenuto in K . In questo caso f è nulla sul bordo di A e quindi il secondo membro di (4.176) è nullo. Quanto al membro di sinistra, se \bar{x} è il centro del cubo K e $2r$ il suo lato, e se, per i fissato, denotiamo con \hat{K} il cubo $\{(x_1, \dots, \hat{x}_i, \dots, x_n) \in \mathbb{R}^{n-1} : |x_h - \bar{x}_h| < r, \text{ per } h \neq i\}$ si ha

$$\begin{aligned} \int_A \frac{\partial f}{\partial x_i} dx &= \int_K \frac{\partial f}{\partial x_i} dx = \int_{\hat{K}} \left(\int_{\tilde{x}_i - r}^{\tilde{x}_i + r} \frac{\partial f}{\partial x_i} dx_i \right) dx_1 \cdots \widehat{dx}_i \cdots dx_n \\ &= \int_{\hat{K}} \left(f(x_1, \dots, \tilde{x}_i + r, \dots, x_n) - f(x_1, \dots, \tilde{x}_i - r, \dots, x_n) \right) dx_1 \cdots \widehat{dx}_i \cdots dx_n \\ &= 0. \end{aligned}$$

⁹²Si noti che la variabile x_i appare nell'argomento di g e di f ; inoltre i punti $x_{\pm r}$ appartengono alla frontiera di $K_{r,\rho}$ e dunque la funzione f vi si annulla.

⁹³Come già osservato, su $A \cap K_{r,\rho}$ la variabile x_j varia tra $\tilde{x}_j - \rho$ e $g(\tilde{x})$, mentre le variabili x_i , per $i \neq j$, variano tra $\tilde{x}_i - r$ e $\tilde{x}_i + r$; si ricordino anche le osservazioni che precedono (4.190).

5) Completiamo ora la dimostrazione di (4.176) per funzioni arbitrarie $f \in C^1(\bar{A}, \mathbb{R})$. Se $x \in A$ sia $K(x)$ un cubo chiuso (non degenero) centrato in x e contenuto in A (tali cubi esistono perché A è aperto); se $x \in \partial A$ sia $K(x)$ un rettangolo che goda delle proprietà elencate all'inizio della dimostrazione [prima del punto 1)]. L'unione di $\tilde{K}(x)$ forma un ricoprimento del compatto \bar{A} e per il Lemma 4.64 esiste una partizione dell'unità $\{f_j : 1 \leq j \leq N\}$ subordinata al ricoprimento $\{\tilde{K}(x) : x \in \bar{A}\}$. Per ogni j , il prodotto $f_j f$ è una funzione di classe $C^1(\bar{A})$ e con supporto contenuto o in un rettangolo $K(x)$ con $x \in \partial A$ o in un cubo $K(x)$ contenuto in A . Dunque per i punti 3) e 4) si ha, per ogni $1 \leq j \leq N$, che

$$\int_A \frac{\partial(f_j f)}{\partial x_i} dx = \int_{\partial A} (f_j f) \nu_i d\sigma_{n-1},$$

e dunque (ricordando che $\sum_{j=1}^N f_j(x) = 1$ per ogni $x \in \bar{A}$) troviamo

$$\begin{aligned} \int_A \frac{\partial f}{\partial x_i} dx &= \int_A \frac{\partial}{\partial x_i} \left(f \sum_{j=1}^N f_j \right) dx = \sum_{j=1}^N \int_A \frac{\partial}{\partial x_i} (f f_j) dx \\ &= \sum_{j=1}^N \int_{\partial A} f_j f \nu_i d\sigma_{n-1} = \int_{\partial A} \left(\sum_{j=1}^N f_j \right) f \nu_i d\sigma_{n-1} \\ &= \int_{\partial A} f \nu_i d\sigma_{n-1}. \quad \blacksquare \end{aligned}$$

Come applicazione del Teorema della divergenza discutiamo brevemente l'**integrazione in coordinate polari in \mathbb{R}^n** . Useremo le seguenti notazioni:

$$\begin{aligned} D_{r,R}^n &:= \{x \in \mathbb{R}^n : r \leq |x| \leq R\}, & B_r^n &:= \{x \in \mathbb{R}^n : |x| < r\}, \\ S^{n-1} &:= \{x \in \mathbb{R}^n : |x| = 1\}, \end{aligned} \quad (4.194)$$

dove $0 \leq r < R$ e, come al solito, $|\cdot|$ denota la norma euclidea. Supponiamo ora che f sia una funzione integrabile su $D_{r,R}^n$, allora la funzione $(\rho, y) \in [r, R] \times S^{n-1} \rightarrow f(\rho y)$ è integrabile su $[r, R] \times S^{n-1}$ e si ha

$$\int_{D_{r,R}^n} f(x) dx = \int_r^R \rho^{n-1} \left(\int_{S^{n-1}} f(\rho y) d\sigma_{n-1} \right) d\rho. \quad (4.195)$$

Dimostrazione (di (4.195) nel caso in cui $f \in C^1(D_{r,R}^n)$) Per ogni $\rho > 0$, facendo il cambio di variabile $x = \rho y$, si ha che

$$\int_{B_\rho^n} f(x) dx = \rho^n \int_{B_1^n} f(\rho y) dy, \quad (4.196)$$

e derivando rispetto a ρ si ha

$$\begin{aligned} \frac{\partial}{\partial \rho} \left(\int_{B_\rho^n} f(x) dx \right) &= n\rho^{n-1} \int_{B_1^n} f(\rho y) dy + \rho^n \int_{B_1^n} (\nabla f)(\rho y) \cdot y dy \\ &= \rho^{n-1} \int_{B_1^n} \left(n f(\rho y) dy + \rho y \cdot (\nabla f)(\rho y) \right) dy, \end{aligned} \quad (4.197)$$

e la quantità nell'ultimo integrale non è altro che la divergenza della funzione vettoriale $y \rightarrow y f(\rho y)$. Dunque da (4.197) e dal Teorema della divergenza segue che⁹⁴

$$\frac{\partial}{\partial \rho} \left(\int_{B_\rho^n} f(x) dx \right) = \rho^{n-1} \int_{B_1^n} \nabla \cdot (y f(\rho y)) dy = \rho^{n-1} \int_{S^{n-1}} f(\rho y) d\sigma_{n-1}, \quad (4.198)$$

ed integrando tale relazione tra r e R si ottiene⁹⁵ (4.195). \blacksquare

Prendendo $f := 1$, (4.195) fornisce la relazione tra la misura n -dimensionale della sfera unitaria con la misura $(n-1)$ -dimensionale della superficie sferica S^{n-1} :

$$n m(B_1^n) = \sigma_{n-1}(S^{n-1}). \quad (4.199)$$

⁹⁴Si noti che la normale esterna a S^{n-1} nel punto y coincide con y .

⁹⁵Naturalmente: $\int_{B_R^n} f(x) dx - \int_{B_r^n} f(x) dx = \int_{D_{r,R}^n} f(x) dx$.

Denotiamo, ora, con Ω_n il volume della sfera euclidea in \mathbb{R}^n ossia $\Omega_n := m(B_1^n)$. Applicando la (4.195) a $f(x) := e^{-|x|^2}$ con $r = 0$ e R arbitrario e mandando R ad infinito, ricordando (4.169), si trova

$$\pi^{\frac{n}{2}} = \int_{\mathbb{R}^n} e^{-|x|^2} dx = n\Omega_n \int_0^\infty \rho^{n-1} e^{-\rho^2} d\rho = \frac{n\Omega_n}{2} \int_0^\infty t^{\frac{n}{2}-1} e^{-t} dt := \frac{n\Omega_n}{2} \Gamma\left(\frac{n}{2}\right),$$

dove $\Gamma(z) := \int_0^\infty t^{z-1} e^{-t} dt$ è la ‘funzione Γ di Eulero’ (si veda C.9), da cui

$$\Omega_n = \frac{2 \pi^{\frac{n}{2}}}{n \Gamma(n/2)}. \quad (4.200)$$

Esercizi

Esercizio 4.39 (i) Siano v e w due vettori in \mathbb{R}^3 indipendenti. Si dimostri che esiste $U \in SO(3)$ tale che $U(v \times w) = \lambda e^{(3)}$ con $\lambda > 0$.

(ii) Si dimostri che $Uv, Uw \in \pi_{e^{(1)}, e^{(2)}}$.

Sia $Uv = (\bar{v}, 0)$, $Uw = (\bar{w}, 0)$ con $\bar{v}, \bar{w} \in \mathbb{R}^2$. Definiamo l’area del parallelogramma $P := \{a_1 v + a_2 w$ con $a_i \in [0, 1]\}$ come la misura (bidimensionale) di $\Pi(\bar{v}, \bar{w})$.

(iii)* Si dimostri che la definizione in (ii) *non dipende* dalla particolare scelta di U (che non è unica).

(iv) Si dimostri che $|v \times w|$ coincide con l’area del parallelogramma P .

Esercizio 4.40 Si dimostri che $\text{Or}(v^{(1)}, v^{(2)}, v^{(3)})$ è una classe di equivalenza nell’insieme delle triple ordinate di vettori indipendenti in \mathbb{R}^3 .

Esercizio 4.41 Dimostrare che se $(v^{(1)}, v^{(2)}, v^{(3)}) \in \text{Or}(e^{(1)}, e^{(2)}, e^{(3)})$ allora

$$v^{(3)} \cdot v^{(1)} \times v^{(2)} > 0. \quad (4.201)$$

Ora si mediti sulla ‘regola della mano destra’.

Esercizio 4.42 (Vettori normali in \mathbb{R}^3) Sia $\mathcal{S} = \varphi(U)$ un elemento di 2–superficie in \mathbb{R}^3 . Dalle proprietà del prodotto vettoriale deriva che il vettore $n = \frac{\partial \varphi}{\partial u_1} \times \frac{\partial \varphi}{\partial u_2}$ è ortogonale ai vettori tangenti $\frac{\partial \varphi}{\partial u_1}$ e $\frac{\partial \varphi}{\partial u_2}$, e, poiché tali vettori formano una base per lo spazio tangente a \mathcal{S} in $\varphi(u)$, segue che il vettore n è ortogonale all’intero spazio $T\mathcal{S}_{\varphi(u)}$. Il vettore unitario (‘versore’)

$$\nu := \frac{\frac{\partial \varphi}{\partial u_1} \times \frac{\partial \varphi}{\partial u_2}}{\left| \frac{\partial \varphi}{\partial u_1} \times \frac{\partial \varphi}{\partial u_2} \right|} \quad (4.202)$$

verrà chiamato **la normale ‘positiva’** a \mathcal{S} in $x = \varphi(u)$ e verifica

$$|\nu| = 1, \quad \nu \perp T\mathcal{S}_{\varphi(u)} \quad \left(\nu, \frac{\partial \varphi}{\partial u_1}, \frac{\partial \varphi}{\partial u_2} \right) \in \text{Or}(e^{(1)}, e^{(2)}, e^{(3)}). \quad (4.203)$$

La parola ‘positiva’ si riferisce all’ultima proprietà in (4.203).

Esercizio 4.43* Sia \mathcal{S} la porzione cilindrica in \mathbb{R}^3 data da $\{(x, y, z) \in \mathbb{R}^3 : 0 < z < 1, x^2 + y^2 = 1, \text{ e } y > 0\}$. Si dimostri che per ogni $M > 0$ (comunque grande) e per ogni $\delta > 0$ (comunque piccolo) esistono N triangoli con interni a due a due disgiunti e con vertici su \mathcal{S} tali che la somma delle loro aree supera M .

Esercizio 4.44 Calcolare l’area superficiale dell’ellissoide di rotazione

$$\left(\frac{x_1}{R}\right)^2 + \left(\frac{x_2}{R}\right)^2 + \left(\frac{x_3}{r}\right)^2 = 1 \quad (4.204)$$

dove $R \neq r$ sono due numeri positivi.

Esercizio 4.45 Si calcoli l’area della superficie del toro T^2 ottenuto ruotando una circonferenza di raggio r attorno ad un asse a distanza $R > r$ dal centro della circonferenza.

Si noti che T^2 ammette la parametrizzazione:

$$T^2 = \{x = (R + r \cos t) \cos \theta, y = (R + r \cos t) \sin \theta, z = r \sin t, \text{ con } 0 \leq t, \theta \leq 2\pi\}.$$

Esercizio 4.46* Sia $A := \mathbb{R}^2 \setminus \{0\}$ e sia ω^1 una 1-forma chiusa su A . Dimostrare che se Γ è un cerchio centrato nell'origine e se $\int_{\Gamma} \omega^1 = 0$ allora ω^1 è esatta in A .

Esercizio 4.47 Sia T il triangolo in \mathbb{R}^2 con vertici in $(0, 0)$, $(0, 1)$ e $(2, 0)$ e sia $\omega := \sinh(x+y)dx + xdy$. Si calcoli

$$\int_{\partial T^+} \omega$$

sia direttamente che usando la formula di Green.

Esercizio 4.48 Sia $\omega^1 = \frac{(x+2y)dx + ydy}{(x+y)^2}$ definita su $A := \{(x, y) \in \mathbb{R}^2 : x+y \neq 0\}$.

(i) Dire se ω^1 è esatta su A ed in caso affermativo, calcolare la primitiva $U(x, y)$ tale che $U(0, e) = U(0, -e) = 1$. (ii) Calcolare $\int_{\Gamma} \omega^1$ dove $\Gamma := \{(x, y) = e(\cos^{100} t, \sin^{100} t) : 0 \leq t \leq \pi/2\}$ orientata nel verso che va da $(e, 0)$ a $(0, e)$.

Esercizio 4.49 Sia $\omega^1 = ydx + xdy$; sia $\Gamma_1 \subseteq \mathbb{R}^2$ il segmento orientato che va da $(2, 0)$ a $(1, 1)$ e sia Γ_2 l'arco orientato della circonferenza unitaria di centro $(1, 0)$ che va da $(2, 0)$ a $(1, 1)$. Si calcoli $\int_{\Gamma_i} \omega^1$. Si trovi poi una funzione f tale che $df = \omega^1$.

[Risposta: $f = xy$.]

Esercizio 4.50 Sia $A = \mathbb{R}^2 \setminus \{0\}$ e

$$\omega^1 = \frac{xdy - ydx}{x^2 + y^2} \quad (4.205)$$

cosicchè $\omega^1 \in C^\infty(A)$. Sia Γ la circonferenza di raggio R centrata nell'origine orientata in senso antiorario e si calcoli $\int_{\Gamma} \omega^1$.

[Risposta: 2π .]

Esercizio 4.51 Far vedere che $\omega = xdy + ydx$ è esatta in \mathbb{R}^2 e verificare direttamente che $\int_{\Gamma_1} \omega = \int_{\Gamma_2} \omega = \int_{\Gamma_3} \omega$, dove Γ_1 è il segmento di estremi $(0, 0)$ e $(1, 1)$; Γ_2 è la poligonale di vertici (in ordine) $(0, 0)$, $(1, 0)$ e $(1, 1)$; Γ_3 è l'arco di circonferenza di raggio unitario, centro $(0, 1)$, estremi $(0, 0)$ e $(1, 1)$ e lunghezza $\frac{\pi}{4}$ [e in tutti e tre e casi l'orientamento è quello corrispondente al verso che va da $(0, 0)$ a $(1, 1)$].

Esercizio 4.52 (i) Si dimostri che la forma $\bar{\omega}^1$ definita in (4.111) è chiusa su $\mathbb{R}^2 \setminus \{0\}$ e che vale (4.112). (ii) Si dimostri che ω^1 è esatta su un qualunque dominio della forma $\mathbb{R}^2 \setminus P$ con P una qualunque semiretta chiusa con estremo nell'origine.

Esercizio 4.53 Si calcoli il flusso esterno di $F(x) = x$ ($x \in \mathbb{R}^3$) attraverso la calotta sferica $\mathcal{S} := \{x \in \mathbb{R}^3 : |x| = 1 \text{ e } x_3 > \frac{1}{2}\}$ dove l'orientamento è quello indotto da $\varphi(u) = (u, \sqrt{1-|u|^2})$, $u = (u_1, u_2)$.

Appendice A

Serie

A.1 Serie di funzioni

Naturalmente tutte le definizioni e affermazioni fatte per successioni di funzioni nel Complemento 1.1 valgono anche per serie di funzioni, intendendo, come va fatto, una serie come successione delle sue somme parziali. Per esempio, data una famiglia di funzioni $u_j : E \subseteq \mathbb{R}^n \rightarrow$

\mathbb{R}^m , $j \in \mathbb{N}$, la serie di funzioni $\sum_{j=1}^{\infty} u_j(x)$ converge uniformemente (o puntualmente)

su E , se la successione delle somme parziali $f_k(x) := \sum_{j=1}^k u_j(x)$, converge uniformemente (o puntualmente) su E .

Vi è, però, nel caso di serie, una nozione utile in più:

Definizione A.1 Una serie di funzioni $\sum u_j$, $u_j : E \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^m$ converge totalmente su E se converge la serie numerica a termini positivi

$$\sum_{j=1}^{\infty} \sup_{x \in E} |u_j(x)|. \quad (\text{A.1})$$

La convergenza totale è (strettamente) più forte della convergenza uniforme: se $k > h$

$$\left| \sum_{j=1}^k u_j(x) - \sum_{j=1}^h u_j(x) \right| = \left| \sum_{j=h+1}^k u_j(x) \right| \leq \sum_{j=h+1}^k |u_j(x)| \leq \sum_{j=h+1}^k \sup_{x \in E} |u_j(x)|,$$

e, dunque, se $\sum_{j=1}^k \sup_E |u_j|$ è di Cauchy, allora $\sum_{j=1}^k u_j$ è uniformemente di Cauchy su E .

Il viceversa, in generale, è falso. Per esempio, la serie $\sum \frac{(-x)^k}{k}$ converge uniformemente su $[0, 1]$ ma non vi converge totalmente. Infatti, chiaramente la serie converge puntualmente su $[0, 1]$, ed inoltre, per ogni $x \in [0, 1]$, si ha¹

$$\begin{aligned} \left| \sum_{j=k}^{\infty} \frac{(-x)^j}{j} \right| &= \left(\frac{x^k}{k} - \frac{x^{k+1}}{k+1} \right) + \left(\frac{x^{k+2}}{k+2} - \frac{x^{k+3}}{k+3} \right) + \dots \\ &\leq \left(\frac{1}{k} - \frac{1}{k+1} \right) + \left(\frac{1}{k+2} - \frac{1}{k+3} \right) + \dots \\ &= (-1)^k \sum_{j=k}^{\infty} \frac{(-1)^j}{j}, \end{aligned}$$

¹La disuguaglianza segue dal fatto che le funzioni $x \in [0, 1] \rightarrow \left(\frac{x^j}{j} - \frac{x^{j+1}}{j+1} \right)$ sono positive e crescenti in $[0, 1]$.

da cui segue facilmente la convergenza uniforme della serie. D'altra parte,

$$\sum \sup_{[0,1]} \left| \frac{(-x)^k}{k} \right| = \sum \frac{1}{k} = +\infty.$$

Esempio A.2 (i) La funzione di Riemann $\zeta(x) := \sum_{n=1}^{\infty} \frac{1}{n^x}$ converge puntualmente per $x \in (1, +\infty)$ (come noto²); converge totalmente (e quindi uniformemente) su $[a, +\infty)$ per ogni $a > 1$, essendo $\sum_{x \geq a} \sup n^{-x} \leq \sum n^{-a} < +\infty$. Dunque $\zeta(x)$ è una funzione continua su $(1, +\infty)$. Ma non converge uniformemente su $(1, +\infty)$ (per $x = 1$, la serie diverge).

(ii) La serie $\sum \frac{1}{k^2} \operatorname{sen} kx$ converge totalmente su \mathbb{R} e quindi definisce una funzione continua su \mathbb{R} (e periodica di periodo 2π).

Sviluppi in serie delle funzioni elementari

Dimostriamo, ora, gli sviluppi (puntuali) in serie delle principali funzioni elementari lasciando ai lettori l'interessante esercizio di determinare dove le convergenze sono uniformi³. Nelle seguenti identità $x \in \mathbb{R}$ (o, più in generale, $x \in \mathbb{C}$).

$$\frac{x^m}{(1-x)^n} = \sum_{k=m}^{\infty} \binom{n-1-m+k}{n-1} x^k, \quad n, m \in \mathbb{N}, n \geq 1, |x| < 1; \quad (\text{A.2})$$

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}, \quad \forall x; \quad (\text{A.3})$$

$$\log(1+x) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k}, \quad |x| < 1; \quad (\text{A.4})$$

$$\log\left(\frac{1+x}{1-x}\right) = 2 \sum_{k=0}^{\infty} \frac{x^{2k+1}}{2k+1}, \quad |x| < 1; \quad (\text{A.5})$$

$$(1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k, \quad \alpha \in \mathbb{R} \setminus \mathbb{Z}, |x| < 1; \quad (\text{A.6})$$

$$\operatorname{sen} x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}, \quad \forall x; \quad (\text{A.7})$$

$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}, \quad \forall x; \quad (\text{A.8})$$

$$\operatorname{Arcsen} x = \sum_{k=0}^{\infty} \frac{(2k-1)!!}{(2k)!!} \frac{1}{2k+1} x^{2k+1}, \quad |x| < 1; \quad (\text{A.9})$$

²Si veda, per esempio, [C2019].

³Es 1.44.

$$\operatorname{Arccos} x = \frac{\pi}{2} - \sum_{k=0}^{\infty} \frac{(2k-1)!!}{(2k)!!} \frac{1}{2k+1} x^{2k+1}, \quad |x| < 1; \quad (\text{A.10})$$

$$\operatorname{Arctan} x = - \sum_{k=1}^{\infty} (-1)^k \frac{x^{2k-1}}{2k-1}, \quad |x| < 1; \quad (\text{A.11})$$

$$\operatorname{senh} x = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!}, \quad \forall x; \quad (\text{A.12})$$

$$\operatorname{cosh} x = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!}, \quad \forall x; \quad (\text{A.13})$$

$$\operatorname{Arcsenh} x = \sum_{k=0}^{\infty} (-1)^k \frac{(2k-1)!!}{(2k)!!} \frac{x^{2k+1}}{2k+1}, \quad |x| < 1; \quad (\text{A.14})$$

$$\operatorname{Arctanh} x = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{2k+1}, \quad |x| < 1; \quad (\text{A.15})$$

dove per α numero reale non intero (e k intero non negativo)

$$\binom{\alpha}{0} := 1, \quad \binom{\alpha}{1} := \alpha,$$

e per $k \geq 2$,

$$\binom{\alpha}{k} := \frac{\alpha(\alpha-1)\cdots(\alpha-k+1)}{k!}; \quad (\text{A.16})$$

il doppio fattoriale $n!!$ (per n intero) è definito come

$$0!! = 1, \quad 1!! := 1, \quad \text{e per } n \geq 2 \quad n!! := n(n-2)!!; \quad (\text{A.17})$$

si noti la scelta del ‘ramo principale’ per le funzioni circolari e iperboliche inverse⁴.

Dimostrazione La (A.2) si ottiene facilmente dalla serie geometrica.

La (A.3) è ben nota⁵.

La (A.4) si ottiene dalla relazione

$$\int_0^x \frac{dt}{1+t} = \log(1+x)$$

espandendo in serie $(1+t)^{-1}$ ed usando il⁶ Teorema 1.59.

La (A.5) si ottiene dalla relazione precedente osservando che $\log\left(\frac{1+x}{1-x}\right) = \log(1+x) - \log(1-x)$.

La (A.6) si ottiene calcolando la serie di Taylor ed usando il usare il criterio del rapporto per serie⁷.

⁴ $\operatorname{Arcsen} 0 = 0$, $\operatorname{Arccos} 0 = \pi/2$, $\operatorname{Arctan} 0 = 0$, $\operatorname{Arcsenh} 0 = 0$, $\operatorname{Arctanh} 0 = 0$.

⁵Cfr., per esempio, Teorema 5.3 in [C2019].

⁶ $(1+t)^{-1} = \sum_{k=0}^{\infty} (-1)^k t^k$ e la convergenza è uniforme in ogni insieme $[-a, a]$ con $0 < a < 1$.

⁷Dal criterio del rapporto per serie numeriche si ottiene immediatamente che: Se esiste il $\lim_{n \rightarrow \infty} \left| \frac{a_{n+1}}{a_n} \right|$, tale limite coincide con $1/R$ dove R è il raggio di convergenza della serie di potenze $\sum_{n \geq 0} a_n (x - x_0)^n$.

Le (A.7) e (A.8) valgono per definizione.

La (A.9) si ottiene dalla relazione

$$\operatorname{Arcsen} x = \int_0^x (1-t^2)^{-\frac{1}{2}} dt,$$

usando la (A.6).

La (A.10) si ottiene dalla (A.9) osservando che $\operatorname{Arccos} x = \frac{\pi}{2} - \operatorname{Arcsen} x$.

La (A.11) si ottiene dalla relazione

$$\operatorname{Arctan} x = \int_0^x (1+t^2)^{-\frac{1}{2}} dt,$$

usando la (A.6).

Le (A.12) e (A.13) derivano dalla serie esponenziale (A.2); le (A.13) e (A.14) seguono (analogamente al caso delle funzioni circolari inverse), rispettivamente, dalle identità

$$\operatorname{Arcsenh} x = \int_0^x (1+t^2)^{-\frac{1}{2}} dt, \quad \operatorname{Arctan} x = \int_0^x (1-t^2)^{-\frac{1}{2}} dt. \quad \blacksquare \quad (\text{A.18})$$

Le espansioni in serie delle funzioni $\tan x$, $\cotan x$, $\tanh x$, $\cotanh x$ sono date in termini dei cosiddetti **numeri di Bernoulli**. Per definizione i numeri di Bernoulli B_n ($n \geq 0$) sono i coefficienti dello sviluppo attorno a $x = 0$ della funzione $\frac{x}{e^x-1}$ moltiplicati per $n!$:

$$\frac{x}{e^x-1} := \sum_{n=0}^{\infty} B_n \frac{x^n}{n!}. \quad (\text{A.19})$$

Valgono allora le seguenti identità per $|x|$ sufficientemente piccolo⁸:

$$\tan x = \sum_{k=1}^{\infty} \frac{2^{2k}(2^{2k}-1)}{(2k)!} |B_{2k}| x^{2k-1}, \quad (\text{A.20})$$

$$\cotan x = \frac{1}{x} - \sum_{k=1}^{\infty} \frac{2^{2k}}{(2k)!} |B_{2k}| x^{2k-1}, \quad (\text{A.21})$$

$$\tanh x = \sum_{k=1}^{\infty} \frac{2^{2k}(2^{2k}-1)}{(2k)!} B_{2k} x^{2k-1}, \quad (\text{A.22})$$

$$\cotanh x = \frac{1}{x} + \sum_{k=1}^{\infty} \frac{2^{2k}}{(2k)!} B_{2k} x^{2k-1}. \quad (\text{A.23})$$

Per la dimostrazione si vedano gli Es 1.53 e 1.54.

Un'altra funzione elementare trascendente che compare spesso nelle applicazioni è la cosiddetta *funzione di errore*

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt. \quad (\text{A.24})$$

L'espansione in serie di tale funzione si ottiene immediatamente calcolando la serie esponenziale con $x = t^2$ e integrando da 0 a x :

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \sum_{k=0}^{\infty} \frac{(-1)^k}{k!} \frac{x^{2k+1}}{2k+1}, \quad \forall x. \quad (\text{A.25})$$

⁸La determinazione esatta dei domini su cui valgono le (A.20)÷(A.23) non è immediata: si può dimostrare che le (A.20) e (A.22) valgono per $x^2 < \pi^2/4$ mentre (A.21) e (A.23) valgono per $x^2 < \pi^2$.

A.2 Serie di potenze

Serie di potenze

Si consideri la ‘serie di potenze’ $\sum_{n \geq 0} a_n(x - x_0)^n$ dove x è una variabile reale e x_0 e a_n sono numeri reali dati. Il problema è studiare la convergenza della successione $\sum_{n=0}^N a_n(x - x_0)^n$ per $N \rightarrow \infty$ per valori di x in un intorno di x_0 .

Teorema A.3 Sia⁹ $R := (\limsup |a_n|^{\frac{1}{n}})^{-1}$. La serie $\sum a_n(x - x_0)^n$ converge assolutamente se $|x - x_0| < R$ e non converge se $|x - x_0| > R$.

Dimostrazione Si ha:

$$\limsup (|a_n| |x - x_0|^n)^{\frac{1}{n}} = |x - x_0| \limsup |a_n|^{\frac{1}{n}} = |x - x_0| \frac{1}{R}$$

e il teorema segue dal criterio della radice per serie numeriche¹⁰. ■

Definizione A.4 Il valore $R =: \rho(f) \in [0, \infty]$ definito nel Teorema A.3 prende il nome di **raggio di convergenza** della serie di potenze $f = \sum a_n(x - x_0)^n$.

Esempio A.5 È immediato verificare che¹¹

$$(1) \quad \rho\left(\sum_{n=1}^{\infty} \frac{x^n}{n^n}\right) = \infty, \quad (2) \quad \rho\left(\sum_{n=1}^{\infty} n^n x^n\right) = 0, \quad (3) \quad \rho\left(\sum_{n=1}^{\infty} n^\alpha x^n\right) = 1$$

e, nel caso (3), per $\alpha = 0$ la serie non converge per $|x| = 1$, per $\alpha < -1$ la serie converge per ogni $|x| = 1$, e per $-1 \leq \alpha < 0$ la serie diverge per $x = 1$ e converge per $x = -1$ (criterio di Leibnitz). Quindi, la convergenza per $|x - x_0|$ uguale al raggio di convergenza va discussa di volta in volta.

Osservazione A.6 Per ogni $r < \rho(f)$ la serie $\sum a_n(x - x_0)^n$ converge *totalmente* per $|x - x_0| \leq r$.

Ricordando che la convergenza totale implica quella uniforme, applichiamo la Proposizione 1.61 e consideriamo le serie ottenute derivando termine a termine la serie f :

$$\begin{aligned} f_1 &= \sum_{n=0}^{\infty} a_n((x - x_0)^n)' = \sum_{n=0}^{\infty} a_{n+1}(n+1)(x - x_0)^n, \\ f_k &= \sum_{n=0}^{\infty} a_n((x - x_0)^n)^{(k)} \\ &= \sum_{n=0}^{\infty} a_{n+k}(n+k)(n+k-1) \cdots (n+1)(x - x_0)^n \\ &= \sum_{n=0}^{\infty} a_{n+k} \frac{(n+k)!}{n!} (x - x_0)^n, \quad (k \geq 2). \end{aligned} \tag{A.26}$$

⁹Adottiamo qui la convenzione che $R = 0$ se il limite superiore è $+\infty$ ed $R = +\infty$ se il limite superiore è 0.

¹⁰Si ricorda il criterio della radice per serie numeriche $\sum_{n=0}^{\infty} a_n$: se $\limsup_{n \rightarrow \infty} |a_n|^{\frac{1}{n}} < 1$ allora la serie $\sum_{n=0}^{\infty} a_n$ converge assolutamente; se $\limsup_{n \rightarrow \infty} |a_n|^{\frac{1}{n}} > 1$ allora $\sum_{n=0}^{\infty} a_n$ non converge.

¹¹Qui $x_0 = 0$.

Poiché $\limsup_{n \rightarrow \infty} (n+h)^{\frac{1}{n}} = 1$ per ogni h , si vede immediatamente che $\rho(f_k) = \rho(f)$ per ogni $k \geq 1$. Analogamente se consideriamo le serie ottenute integrando termine a termine la serie f :

$$\begin{aligned} F_1 &= \sum_{n=0}^{\infty} a_n \int_{x_0}^x (t-x_0)^n dt = \sum_{n=1}^{\infty} \frac{a_{n-1}}{n} (x-x_0)^n, \\ F_k &= \sum_{n=k}^{\infty} \frac{a_{n-k}}{n(n-1) \cdots (n-k+1)} (x-x_0)^n, \quad k \geq 2. \end{aligned} \quad (\text{A.27})$$

si vede subito che $\rho(F_k) = \rho(f)$ per ogni $k \geq 1$. Quindi dalle Proposizioni 1.61 e 1.59 e dall'Osservazione A.6 segue immediatamente il seguente risultato.

Teorema A.7 Sia $f = \sum a_n (x-x_0)^n$ e siano f_k e F_k le serie di potenze definite in (A.26) e (A.27). Se R è il raggio di convergenza della serie f , si ha $\rho(f_k) = R = \rho(F_k)$ (per ogni $k \geq 1$) ed inoltre f è una funzione $C^\infty(\{x : |x-x_0| < R\})$ e valgono le relazioni

$$\begin{aligned} f^{(k)} &= f_k, \quad (k \geq 1), \\ \int_{x_0}^x f(t) dt &= F_1(x), \dots, \int_{x_0}^x F_{k-1}(t) dt = F_k(x), \quad (k \geq 2). \end{aligned}$$

Come applicazione consideriamo il seguente

Esempio A.8 Sappiamo che la serie geometrica di ragione x con $|x| < 1$ ha somma $(1-x)^{-1}$ cioè

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}. \quad (\text{A.28})$$

Derivando k volte la funzione $(1-x)^{-1}$ si ottiene

$$\frac{d^k}{dx^k} \frac{1}{1-x} = \frac{k!}{(1-x)^{k+1}}. \quad (\text{A.29})$$

Dunque, da (A.26), da (A.29) e dal Teorema A.7 otteniamo la relazione

$$\frac{k!}{(1-x)^{k+1}} = \sum_{n=0}^{\infty} \frac{(n+k)!}{n!} x^n \quad (\text{A.30})$$

ossia

$$\frac{1}{(1-x)^{k+1}} = \sum_{n=0}^{\infty} \binom{n+k}{k} x^n, \quad (\text{A.31})$$

che dunque rappresenta l'espansione in serie della funzione $(1-x)^{-(k+1)}$.

Osservazione A.9 (i) Sia $f(x) := \sum_{n \geq 0} a_n (x-x_0)^n$ una serie di potenze con raggio di convergenza $R > 0$. Dal Teorema A.7 sappiamo che $f \in C^\infty(\{x : |x-x_0| < R\})$ e calcolando la k -esima derivata in $x = x_0$ otteniamo

$$f^{(k)}(x_0) = a_k k!, \quad \text{cioè} \quad a_n = \frac{f^{(n)}(x_0)}{n!} \quad (n \geq 0). \quad (\text{A.32})$$

(ii) Se $f = \sum a_n (x-x_0)^n$ e $g = \sum b_n (x-x_0)^n$ sono due serie di potenze convergenti per $|x-x_0| < r$ e se $f(x) = g(x)$ per ogni $|x-x_0| < r$, chiaramente $f^{(n)}(x_0) = g^{(n)}(x_0)$ per ogni $n \geq 0$. Quindi da (A.32) segue che $a_n = b_n$ per ogni $n \geq 0$. Riassumendo: *due serie di potenze hanno la stessa somma nell'intorno di un punto se e solo se hanno tutti i coefficienti coincidenti.*

La convergenza di una serie di potenze si caratterizza facilmente in termini del tasso di crescita dei suoi coefficienti:

Teorema A.10 *Se la serie di potenze $\sum a_n(x - x_0)^n$ ha raggio di convergenza R positivo allora, per ogni $r < R$ esiste $M > 0$ tale che*

$$|a_n| \leq M r^{-n}, \quad \forall n \geq 0. \quad (\text{A.33})$$

Viceversa se vale (A.33), allora la serie di potenze $\sum a_n(x - x_0)^n$ ha raggio di convergenza $R \geq r$.

Dimostrazione Assumiamo che $\sum a_n(x - x_0)^n$ abbia raggio di convergenza $R > 0$. Dalla definizione di R (e dalla definizione di estremo superiore) segue che per ogni $r < R$ esiste $N > 0$ tale che se $n \geq N$ allora $|a_n|^{\frac{1}{n}} \leq r^{-1}$ ossia $|a_n| \leq r^{-n}$. Quindi (A.33) vale se scegliamo

$$M_0 := \max_{0 \leq n \leq N} |a_n| r^n, \quad \text{e} \quad M := \max\{M_0, 1\}. \quad (\text{A.34})$$

Assumiamo ora che valga (A.33) ossia (prendendo la radice n -esima) $|a_n|^{\frac{1}{n}} \leq M^{\frac{1}{n}} r^{-1}$. Prendendo il limite superiore per n che tende ad ∞ di tale relazione otteniamo immediatamente che $R^{-1} \leq r^{-1}$ ossia che il raggio di convergenza R di $\sum a_n(x - x_0)^n$ verifica $R \geq r > 0$. ■

Il ‘comportamento’ di una serie di potenze $\sum a_n(x - x_0)^n$ attorno a x_0 è legato al primo coefficiente non nullo.

Proposizione A.11 *Sia $m \geq 0$, sia $u = \sum_{n \geq m} a_n(x - x_0)^n$ con $a_m \neq 0$ una serie di potenze con raggio di convergenza positivo. Per ogni $\varepsilon > 0$ esiste $r_0 > 0$ tale che, per ogni $|x - x_0| \leq r_0$,*

$$(1 - \varepsilon)|a_m||x - x_0|^m \leq \left| \sum_{n \geq m} a_n(x - x_0)^n \right| \leq (1 + \varepsilon)|a_m||x - x_0|^m. \quad (\text{A.35})$$

Dimostrazione Poniamo $y := x - x_0$. Per il Teorema A.10 esistono M, r tali che (A.33) vale. Allora, per ogni $|y| < r$, si ha

$$\begin{aligned} \left| \sum_{n \geq m} a_n y^n - a_m y^m \right| &\leq \sum_{n=m+1}^{\infty} |a_n| |y|^n \leq M \sum_{n=m+1}^{\infty} \left(\frac{|y|}{r} \right)^n \\ &= |a_m| |y|^m \left(\frac{M}{|a_m| r^m} \frac{|y|}{r - |y|} \right). \end{aligned} \quad (\text{A.36})$$

Se ora imponiamo che $|y|$ sia così piccolo che la quantità in parentesi nell’ultima riga di (A.36) sia minore o uguale ad ε otteniamo l’asserto. È immediato verificare che ciò si ottiene richiedendo che $|y| \leq r_0$ con

$$r_0 := \frac{\varepsilon r^{m+1} |a_m|}{M + \varepsilon r^m |a_m|}. \quad \blacksquare \quad (\text{A.37})$$

A.3 Serie di Fourier

Polinomi trigonometrici

Una funzione f di variabile reale si dice **periodica di periodo** $T > 0$ se $f(x + T) = f(x)$ per ogni $x \in \mathbb{R}$. Se f è una funzione periodica di periodo T , la funzione \tilde{f} definita da $\tilde{f}(x) := f(x \frac{T}{2\pi})$ è periodica di periodo 2π ; pertanto considereremo solo funzioni di periodo¹² 2π .

Definizione A.12 $C_{\text{per}}^p := \{f \in C^p(\mathbb{R}) : f(x + 2\pi) = f(x), \quad \forall x \in \mathbb{R}\}$.

¹²Per le formule relative al caso generale si veda A.11.

Esempi naturali di funzioni C_{per}^∞ (o meglio C_{per}^ω) sono le costanti, $\text{sen } x$, $\cos x$, $\text{sen } 2x$, $\cos 2x, \dots$ e loro combinazioni lineari:

$$s_N(x) = \frac{a_0}{2} + \sum_{n=1}^N a_n \cos nx + b_n \text{sen } nx, \quad a_n, b_n \in \mathbb{R}; \quad (\text{A.38})$$

(il ruolo, del tutto formale, del fattore $1/2$ davanti alla costante a_0 apparirà chiaro in seguito). Una tale espressione si chiama *polinomio trigonometrico (reale) di grado N* . Una classe assai più vasta di funzioni periodiche si ottiene considerando limiti, per N che tende ad infinito, di polinomi trigonometrici:

Proposizione A.13 *Siano $\{a_n\}_{n \geq 0}, \{b_n\}_{n \geq 1}$ due successioni di numeri reali tali che*

$$\sum_{n=1}^{\infty} (|a_n| + |b_n|) n^p < \infty. \quad (\text{A.39})$$

Allora, $s_N(x)$ (definito in (A.38)) converge, per N che tende ad infinito, uniformemente su \mathbb{R} ad una funzione $f \in C_{\text{per}}^p$.

Il limite di polinomi trigonometrici ossia un'espressione della forma

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nx + b_n \text{sen } nx) \quad (\text{A.40})$$

prende il nome di *serie di Fourier*; dunque il risultato appena esposto può essere riformulato dicendo: 'se i numeri a_n e b_n soddisfano (A.39) allora la serie di Fourier (A.40) ad essi associata converge uniformemente su \mathbb{R} e definisce una funzione di classe C_{per}^p '.

Dimostrazione (della Proposizione A.13) Grazie a (A.39), per ogni $0 \leq k \leq p$, la serie delle derivate $\sum u_n^{(k)}$, con $u_0 := a_0/2$ e (per $n \geq 1$) $u_n := a_n \cos nx + b_n \text{sen } nx$, converge totalmente in \mathbb{R} e la Proposizione 1.61 implica che $\sum u_n$ converge ad una funzione $f \in C^p(\mathbb{R})$. La periodicità di f deriva dalla periodicità di s_N :

$$f(x + 2\pi) = \lim_{N \rightarrow \infty} s_N(x + 2\pi) = \lim_{N \rightarrow \infty} s_N(x) = f(x). \quad \blacksquare$$

L'ipotesi (A.39) implica, in particolare, che $\max\{|a_n|, |b_n|\} \leq M/n^p$.

Più interessante è la questione inversa: *data una funzione $f \in C_{\text{per}}^p$, trovare una successione di polinomi trigonometrici che converga ad f* .

Prima di analizzare tale questione, si noti che $\cos nx$ e $\text{sen } nx$ sono, rispettivamente, la parte reale ed immaginaria di e^{inx} : questa osservazione suggerisce una riscrittura più compatta, in forma complessa, dei polinomi trigonometrici (A.38). Ponendo

$$c_0 := \frac{a_0}{2}, \quad c_{\pm n} := \frac{a_n \mp ib_n}{2} \quad (\forall n \geq 1), \quad (\text{A.41})$$

ed osservando che la 'trasformazione inversa' di tale relazione (ossia la relazione che dà gli a_n e b_n in termini dei $c_{\pm n}$) è data da

$$a_0 = 2c_0, \quad \text{e per } n \geq 1, \quad a_n = c_n + c_{-n}, \quad b_n = ic_n - ic_{-n}, \quad (\text{A.42})$$

si ha che¹³

$$\begin{aligned} \sum_{\substack{n \in \mathbb{Z} \\ |n| \leq N}} c_n e^{inx} &= c_0 + \sum_{n=1}^N (c_n + c_{-n}) \cos nx + (ic_n - ic_{-n}) \sin nx \\ &:= \frac{a_0}{2} + \sum_{n=1}^N a_n \cos nx + b_n \sin nx := s_N(x). \end{aligned} \quad (\text{A.44})$$

Osservazione A.14 (i) Nel caso che stiamo considerando a_n e b_n sono numeri reali e quindi $c_{-n} = \bar{c}_n$ (dove, come al solito al barra denota ‘complesso coniugato’). Naturalmente avrebbe senso considerare il caso (apparentemente più generale) in cui a_n e b_n sono numeri complessi, nel qual caso i polinomi trigonometrici $s_N(x)$ sarebbero delle funzioni complesse di variabile reale. In questo capitolo considereremo *polinomi trigonometrici e serie di Fourier reali* anche se useremo spesso la notazione complessa che presenta dei vantaggi dal punto di vista algebrico

(ii) Si noti che per ogni coppia di numeri reali α e β si ha¹⁴

$$\sqrt{\alpha^2 + \beta^2} \leq |\alpha| + |\beta| \leq \sqrt{2} \sqrt{\alpha^2 + \beta^2}. \quad (\text{A.45})$$

Si noti anche che, se a_n , b_n e c_n sono legati da (A.41) (e $b_0 := 0$), si ha

$$|c_n| = |c_{-n}| = \frac{1}{2} \sqrt{a_n^2 + b_n^2}, \quad \forall n \geq 0. \quad (\text{A.46})$$

Dunque (se a_n , b_n e c_n sono legati da (A.41)) la condizione (A.39) è equivalente a

$$\sum_{n \in \mathbb{Z}} |c_n| |n|^p < \infty, \quad (\text{A.47})$$

e la Proposizione A.13 può essere riformulata come segue

se $\bar{c}_n = c_{-n}$ e $\sum_{n \in \mathbb{Z}} |c_n| |n|^p < \infty$ allora $f(x) := \sum_{n \in \mathbb{Z}} c_n e^{inx} \in C_{\text{per}}^p$.

Coefficienti di Fourier

Sia $f(x) = \sum_{n \in \mathbb{Z}} c_n e^{inx}$ una serie di Fourier con $\sum |c_n| < \infty$ [cosicché $f \in C_{\text{per}}$]. È possibile calcolare i numeri c_n (o, equivalentemente, i numeri a_n e b_n) dalla sola conoscenza della funzione $f(x)$? La risposta (affermativa) è contenuta nella seguente

Proposizione A.15 *Siano c_n , per $n \in \mathbb{Z}$, numeri complessi tali che $\sum_{n \in \mathbb{Z}} |c_n| < \infty$ e sia $f(x) := \sum_{n \in \mathbb{Z}} c_n e^{inx}$. Allora*

$$c_n = \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx. \quad (\text{A.48})$$

Dalle relazioni (A.41) ne consegue che i numeri a_n e b_n sono dati da

$$a_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos nx dx \quad (n \geq 0), \quad b_n = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin nx dx \quad (n \geq 1). \quad (\text{A.49})$$

¹³Se z_n , per $n \in \mathbb{Z}$, sono numeri complessi, i simboli

$$\sum_{\substack{n \in \mathbb{Z} \\ |n| \leq N}} z_n, \quad \sum_{n=-N}^N z_n, \quad \sum_{-N}^N z_n \quad (\text{A.43})$$

denotano, per N intero positivo, la somma $z_{-N} + z_{-N+1} + \dots + z_{N-1} + z_N$; naturalmente il simbolo $\sum_{n=-N}^M z_n$ o (qualora non vi sia ambiguità) $\sum_{-N}^M z_n$, denoterà la somma $z_{-N} + z_{-N+1} + \dots + z_{M-1} + z_M$; i simbolo $\sum_{n \in \mathbb{Z}} z_n$, $\sum_{n=-\infty}^{\infty} z_n$ o $\sum_{-\infty}^{\infty} z_n$ denotano il limite, qualora esista, $\lim_{N \rightarrow \infty} \sum_{n=-N}^N z_n$.

¹⁴La prima disuguaglianza si verifica elevando al quadrato e la seconda deriva dalla disuguaglianza di Cauchy (1.13) con $y_i = 1$.

Dimostrazione (della Proposizione A.15) Dalla Proposizione 1.59 segue che¹⁵

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx &= \frac{1}{2\pi} \lim_{N \rightarrow \infty} \int_0^{2\pi} \sum_{|m| \leq N} c_m e^{imx} e^{-inx} dx \\ &= \frac{1}{2\pi} \lim_{N \rightarrow \infty} \sum_{|m| \leq N} c_m \int_0^{2\pi} e^{i(m-n)x} dx = c_n. \quad \blacksquare \end{aligned}$$

Motivati da questa discussione (e cambiando punto di vista!) diamo la seguente

Definizione A.16 Sia f una funzione periodica di periodo 2π ed integrabile su $[0, 2\pi]$. Si chiamano coefficienti di Fourier di f i numeri complessi

$$\hat{f}_n := \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx, \quad \forall n \in \mathbb{Z}. \quad (\text{A.50})$$

Osservazione A.17 (i) Anche i numeri a_n e b_n definiti in (A.49) (per f una funzione periodica di periodo 2π ed integrabile su $[0, 2\pi]$) vengono, a volte, chiamati coefficienti di Fourier di f (o coefficienti ‘reali’ di Fourier di f).

(ii) Naturalmente il ruolo dell’intervallo $[0, 2\pi]$ può esser giocato da un qualunque intervallo lungo 2π poiché la conoscenza di f su di un qualunque intervallo lungo quanto il suo periodo permette di ricostruire f su tutto \mathbb{R} . Per esempio si ha: se $f(x)$ è una funzione periodica ed integrabile su $[0, 2\pi]$ allora l’integrale su un qualunque intervallo lungo 2π coincide con l’integrale tra 0 e 2π . Infatti, per ogni $a < b$ e per ogni $k \in \mathbb{Z}$, $\int_{a+2k\pi}^{b+2k\pi} f(x) dx = \int_a^b f(t) dt$ (si ponga $t = x + 2k\pi$), dunque, per ogni x_0 ,

$$\begin{aligned} \int_{x_0}^{x_0+2\pi} f(x) dx &= \int_{x_0}^{2\pi} f(x) dx + \int_{2\pi}^{x_0+2\pi} f(x) dx \\ &= \int_{x_0}^{2\pi} f(x) dx + \int_0^{x_0} f(x) dx = \int_0^{2\pi} f(x) dx. \end{aligned}$$

In particolare, se f è periodica ed integrabile su $[0, 2\pi]$ tale è $f(x) e^{-inx}$ e dunque il calcolo dei coefficienti di Fourier di f può essere fatto sostituendo $[0, 2\pi]$ con un qualunque intervallo lungo 2π . Per esempio, è, a volte, utile usare la formula

$$\hat{f}_n = \frac{1}{2\pi} \int_{-\pi}^{\pi} f(x) e^{-inx} dx. \quad (\text{A.51})$$

Alcune proprietà dei coefficienti di Fourier sono contenute nella seguente

Proposizione A.18 Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ una funzione periodica di periodo 2π a quadrato sommabile (ossia, con $|f|^2$ integrabile) su $[0, 2\pi]$. Allora

- (i) $|\hat{f}_n| \leq \frac{1}{2\pi} \int_0^{2\pi} |f(x)| dx \quad (\forall n \in \mathbb{Z});$
- (ii) $\bar{\hat{f}}_n = \hat{f}_{-n}, \quad (\forall n \in \mathbb{Z});$
- (iii) $\frac{1}{2\pi} \int_0^{2\pi} \left| f(x) - \sum_{|n| \leq N} \hat{f}_n e^{inx} \right|^2 dx = \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx - \sum_{|n| \leq N} |\hat{f}_n|^2;$
- (iv) $\sum_{n \in \mathbb{Z}} |\hat{f}_n|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx;$

¹⁵ (i) Qui stiamo usando la versione della Proposizione 1.59 per funzioni complesse di variabile reali: (ii) Si ricordi anche che $\int_0^{2\pi} e^{ikx} dx$ è nullo se $0 \neq k \in \mathbb{Z}$ ed è uguale a 2π se $k = 0$. (iii) Qui ed in seguito espressioni quali

$\sum_{|m| \leq N} z_m$ sono abbreviazioni per $\sum_{\substack{m \in \mathbb{Z} \\ |m| \leq N}} z_m$.

- (v) $\lim_{n \rightarrow \pm\infty} |\hat{f}_n| = 0$;
- (vi) $\widehat{(f^{(k)})}_n = (in)^k \hat{f}_n$, (per $f \in C_{\text{per}}^p$, $\forall 0 \leq k \leq p$, $\forall n$);
- (vii) se $f \in C_{\text{per}}^p$, allora: $\sum_{n \in \mathbb{Z}} |\hat{f}_n|^2 |n|^{2k} \leq \frac{1}{2\pi} \int_0^{2\pi} |f^{(k)}(x)|^2 dx$, $\forall 0 \leq k \leq p$, $\forall n$.

Dimostrazione (della Proposizione A.18)

La (i) è conseguenza immediata di (4.26) e del fatto che $|e^{inx}| = 1$.

La (ii) è conseguenza immediata delle definizioni date.

(iii): Espandendo l'integrando si trovano, oltre $\frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx$, i seguenti due termini¹⁶

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} \left| \sum_{|n| \leq N} \hat{f}_n e^{inx} \right|^2 dx &= \frac{1}{2\pi} \int_0^{2\pi} \left(\sum_{|n| \leq N} \hat{f}_n e^{inx} \right) \left(\sum_{|m| \leq N} \overline{\hat{f}_m} e^{-imx} \right) dx \\ &= \frac{1}{2\pi} \sum_{|n| \leq N} \hat{f}_n \sum_{|m| \leq N} \overline{\hat{f}_m} \int_0^{2\pi} e^{i(n-m)x} dx \\ &= \sum_{|n| \leq N} |\hat{f}_n|^2, \end{aligned}$$

e¹⁷

$$\begin{aligned} \frac{1}{2\pi} \int_0^{2\pi} 2 \operatorname{Re} \left(f(x) \sum_{|n| \leq N} \hat{f}_n e^{inx} \right) dx &= 2 \operatorname{Re} \left(\sum_{|n| \leq N} \hat{f}_n \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{inx} dx \right) \\ &:= 2 \operatorname{Re} \left(\sum_{|n| \leq N} \hat{f}_n \hat{f}_{-n} \right) = 2 \sum_{|n| \leq N} |\hat{f}_n|^2. \end{aligned}$$

Mettendo assieme tali termini si ottiene l'asserto.

(iv): La relazione in (iii) mostra, tra l'altro, che $\sum_{|n| \leq N} |\hat{f}_n|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx$ per ogni N e prendendo il limite per $N \rightarrow \infty$ si ottiene la (iv).

La (v) è implicata dalla convergenza della serie in (iv).

(vi): Si noti che se $f \in C_{\text{per}}^1$ allora¹⁸ $f' \in C_{\text{per}}$. Dunque

$$\widehat{(f')} _0 := \frac{1}{2\pi} \int_0^{2\pi} f'(x) dx = \frac{1}{2\pi} [f(x)]_0^{2\pi} := \frac{1}{2\pi} (f(2\pi) - f(0)) = 0,$$

che è la (vi) per $k = 1$ e $n = 0$; poiché tale relazione può essere iterata, la (vi) è vera per $n = 0$ e per ogni $k \leq p$. Sia ora $n \neq 0$. Essendo $f \in C_{\text{per}}^1$, integrando per parti si ottiene:

$$\begin{aligned} \hat{f}_n &:= \frac{1}{2\pi} \int_0^{2\pi} f(x) e^{-inx} dx = \frac{1}{2\pi} \int_0^{2\pi} f(x) \left(\frac{e^{-inx}}{-in} \right)' dx \\ &= \frac{1}{2\pi} \left[\frac{f(x) e^{-inx}}{-in} \right]_0^{2\pi} + \frac{1}{2\pi in} \int_0^{2\pi} f'(x) e^{-inx} dx := \frac{1}{in} \widehat{(f')} _n, \end{aligned}$$

che è la (vi) per $k = 1$ e $n \neq 0$. Iterando tale relazione si ottiene l'asserto.

(vii): Applicando la (iv) a $f^{(k)}$ (che per le ipotesi è integrabile) ed usando la (vi) si ottiene la relazione (vii). ■

¹⁶Si ricordi che, se z e w sono numeri complessi, allora $|z + w|^2 = (z + w)(\bar{z} + \bar{w}) = |z|^2 + |w|^2 + 2 \operatorname{Re}(z\bar{w})$; si usi il punto (ii) della nota 15.

¹⁷Si usi il punto (ii) e si osservi anche che se $u : [0, 2\pi] \rightarrow \mathbb{C}$ è integrabile allora $\int_0^{2\pi} \operatorname{Re} u(x) dx = \operatorname{Re} \int_0^{2\pi} u(x) dx$.

¹⁸ $f'(x + 2\pi) = \lim_{h \rightarrow 0} \frac{f(x + 2\pi + h) - f(x + 2\pi)}{h} = \lim_{h \rightarrow 0} \frac{f(x + h) - f(x)}{h} = f'(x)$.

Osservazione A.19 (i) La relazione in (iv) viene, a volte, chiamata la *disuguaglianza di Bessel*.

(ii) I vari asseriti della Proposizione A.18 possono essere letti in termini dei coefficienti di Fourier ‘reali’ a_n e b_n (A.49) facendo uso delle relazioni (A.41) o (A.42) con $c_n := \hat{f}_n$. Per esempio, se denotiamo con $a_n^{(k)}$ e $b_n^{(k)}$ i coefficienti ‘reali’ di Fourier di $f^{(k)}$, la (vi) diventa

$$\begin{aligned} a_n^{(k)} &= \begin{cases} (-1)^{\frac{k}{2}} n^k a_n, & \text{se } k \text{ è pari,} \\ (-1)^{\frac{k-1}{2}} n^k b_n, & \text{se } k \text{ è dispari,} \end{cases} \\ b_n^{(k)} &= \begin{cases} (-1)^{\frac{k}{2}} n^k b_n, & \text{se } k \text{ è pari,} \\ (-1)^{\frac{k+1}{2}} n^k a_n, & \text{se } k \text{ è dispari.} \end{cases} \end{aligned} \quad (\text{A.52})$$

(iii) Dal punto (vii) segue che se $f \in C_{\text{per}}^p$, $|\hat{f}_n|^2 |n|^{2p} \leq \frac{1}{2\pi} \int_0^{2\pi} |f^{(p)}(x)|^2 dx$ per ogni $n \in \mathbb{Z}$ e quindi (maggiorando l’integrando con $\sup_{\mathbb{R}} |f^{(p)}(x)|^2$ e prendendo la radice quadrata) si ottiene

$$f \in C_{\text{per}}^p \implies |\hat{f}_n| \leq \frac{M_p}{|n|^p} \quad (\forall n \in \mathbb{Z}), \quad \text{con } M_p := \sup_{\mathbb{R}} |f^{(p)}(x)|. \quad (\text{A.53})$$

(iv) Se f è integrabile su $[0, 2\pi]$ allora $\hat{f}_n \rightarrow 0$ quando $n \rightarrow \infty$ (‘Lemma di Riemann–Lebesgue¹⁹’). Infatti, sia $f_k(x)$ definita come $f(x)$ per $x \in [0, 2\pi - \frac{1}{k})$ e 0 altrimenti. Tale funzione f_k è chiaramente a quadrato sommabile (essendo limitata) ed inoltre dalla definizione di integrale di Riemann segue che $\int_0^{2\pi} |f - f_k| \rightarrow 0$ per $k \rightarrow \infty$. Dato $\varepsilon > 0$ sia k tale che $\int_0^{2\pi} |f - f_k| < \varepsilon/2$. Essendo f_k a quadrato sommabile, dalla Proposizione A.18, punto (v) segue che esiste N tale che per ogni $n \in \mathbb{Z}$ con $|n| > N$ si ha $|\hat{f}_{k,n}| < \varepsilon/2$. Allora, per ogni $|n| > N$ si ha: $|\hat{f}_n| \leq |\hat{f}_n - \hat{f}_{k,n}| + |\hat{f}_{k,n}| = \left| \int_0^{2\pi} (f(x) - f_k(x)) e^{-inx} dx \right| + |\hat{f}_{k,n}| \leq \int_0^{2\pi} |f(x) - f_k(x)| dx + |\hat{f}_{k,n}| \leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$. Dall’arbitrarietà di ε segue l’asserto. ■

(v) Discutiamo un po’più in dettaglio la relazione fondamentale che c’è tra *regolarità* di una funzione periodica ed il *decadimento* (per $|n| \rightarrow \infty$) dei suoi coefficienti di Fourier. Il punto (vii) della Proposizione A.18 è un *parziale* ‘viceversa’ della Proposizione A.13. Infatti, come già osservato nell’Osservazione A.14 punto (ii), la Proposizione A.13 è equivalente a:

se $\bar{c}_n = c_{-n}$ e $\sum_{n \in \mathbb{Z}} |c_n| |n|^p < \infty$ allora $f(x) := \sum_{n \in \mathbb{Z}} c_n e^{inx} \in C_{\text{per}}^p$.

Mentre il punto (vii) implica

se $f \in C_{\text{per}}^p$ allora $\sum_{n \in \mathbb{Z}} |\hat{f}_n|^2 |n|^{2p} < \infty$;

ma quest’ultima condizione è *più debole* della condizione $\sum_{n \in \mathbb{Z}} |\hat{f}_n| |n|^p < \infty$. Quindi queste relazioni non ci permettono di *caratterizzare esattamente* le funzioni C_{per}^p (con $p < \infty$) in termini del decadimento dei loro coefficienti di Fourier. Invece dalla discussione fatta segue che

$f \in C_{\text{per}}^\infty$ se e solo se i coefficienti di Fourier $|\hat{f}_n|$ decadono più rapidamente (per $|n| \rightarrow \infty$) di ogni potenza inversa di²⁰ $|n|$.

Un’altra classe di funzioni che è possibile caratterizzare esattamente tramite una condizione sul decadimento dei coefficienti di Fourier è la classe C_{per}^ω ossia le funzioni periodiche e reali analitiche su \mathbb{R} .

¹⁹Naturalmente una funzione f può essere integrabile ma non a quadrato sommabile come, per esempio, $1/\sqrt{2\pi-x}$ su $[0, 2\pi]$.

²⁰Ossia: $\forall p > 0 \exists C_p > 0$ tale che $|\hat{f}_n| \leq C_p |n|^{-p} \forall n \neq 0$.

Proposizione A.20 $f \in C_{\text{per}}^\omega$ se e solo se $f \in C_{\text{per}}^0$ e

$$\exists C, \sigma > 0 \quad \text{tali che} \quad |\hat{f}_n| \leq C e^{-\sigma|n|} \quad \forall n \in \mathbb{Z}. \quad (\text{A.54})$$

Nel corso della dimostrazione faremo uso della seguente stima

Lemma A.21 Sia $\alpha > 0$ e sia k un intero non negativo, allora

$$\sum_{j=0}^{\infty} e^{-\alpha j} j^k \leq \left(1 + \frac{1}{\alpha}\right) \alpha^{-k} k! \quad (\text{A.55})$$

Dimostrazione Se $k = 0$, poiché $\alpha \leq \sum_{j=1}^{\infty} \alpha^j / j!$, si ha

$$\sum_{j=0}^{\infty} e^{-\alpha j} = \frac{1}{1 - e^{-\alpha}} = 1 + \frac{1}{e^{\alpha} - 1} = 1 + \left(\sum_{j=1}^{\infty} \frac{\alpha^j}{j!}\right)^{-1} \leq 1 + \frac{1}{\alpha},$$

che è (A.55) nel caso $k = 0$. Sia ora $k \geq 1$ e si consideri la funzione $f(t) := e^{-\alpha t} t^k$ per $t \geq 0$. Tale funzione è positiva per $t > 0$, ha un unico massimo in $t_* := k/\alpha$, è strettamente crescente per $0 \leq t \leq t_*$ e strettamente decrescente per $t \geq t_*$. Sia $N := [t_*]$ ($[\cdot]$:= ‘parte intera’) cosicché $N \leq t_* < N + 1$. Allora, poiché $\min_{N \leq t \leq N+1} f(t) = \min\{f(N), f(N+1)\}$, $\int_N^{N+1} f(t) dt \geq \min\{f(N), f(N+1)\}$ e quindi

$$\begin{aligned} f(N) + f(N+1) &:= \min\{f(N), f(N+1)\} + \max\{f(N), f(N+1)\} \\ &\leq f(t_*) + \int_N^{N+1} f. \end{aligned}$$

Dunque, essendo f crescente tra 0 e N e decrescente per $t \geq N + 1$ si ha

$$\sum_{j=1}^{N-1} f(j) \leq \int_0^N f(t) dt, \quad \sum_{j=N+2}^{\infty} f(j) \leq \int_{N+1}^{\infty} f(t) dt,$$

e quindi²¹

$$\begin{aligned} \sum_{j=0}^{\infty} e^{\alpha j} j^k &= \sum_{j=1}^{\infty} f(j) \leq \int_0^{\infty} f(t) dt + f(k/\alpha) \\ &= \frac{k!}{\alpha^{k+1}} + \left(\frac{k}{\alpha e}\right)^k \leq k! \alpha^{-k} \left(1 + \frac{1}{\alpha}\right). \quad \blacksquare \end{aligned}$$

Dimostrazione (della Proposizione A.20) Sia $f \in C_{\text{per}}^\omega$. Dal Teorema A.34 segue che $\forall x_0 \in [0, 2\pi] \exists M_{x_0}, r_{x_0} > 0$ tali che

$$\sup_{|x-x_0| < r_{x_0}} |f^{(k)}(x)| \leq M_{x_0} r_{x_0}^{-k} k!, \quad \forall k \geq 0.$$

Gli intervalli I_{x_0} di centro x_0 e lunghezza $2r_{x_0}$ formano un ricoprimento aperto dell’insieme compatto $[0, 2\pi]$ ed è dunque possibile trovare $0 \leq x_1 < \dots < x_N \leq 2\pi$ tali che $[0, 2\pi] \subseteq I_{x_1} \cup \dots \cup I_{x_N}$. Definendo $M := \max_{1 \leq j \leq N} M_{x_j}$ e $r := \min_{1 \leq j \leq N} r_{x_j}$ si ha allora

$$|f^{(k)}(x)| \leq M r^{-k} k! \quad \forall x \in \mathbb{R}, \quad \forall k \geq 0. \quad (\text{A.56})$$

Dal punto (vi) della Proposizione A.18 e dalla (A.56) segue dunque, per $n \neq 0$, che

$$|\hat{f}_n| = \frac{1}{2\pi|n|^p} \left| \int_0^{2\pi} f^{(p)}(x) e^{-inx} dx \right| \leq \frac{1}{|n|^p} M r^{-p} p! \leq M \left(\frac{p}{|n|r}\right)^p. \quad (\text{A.57})$$

²¹Si usi la stima $k! \geq (k/e)^k$ che si ottiene facilmente per induzione su $k \geq 1$ (ricordando che $(1 + 1/k)^k < e$).

Quindi, se $|n| \geq 2/r$, scegliendo $p := \lceil |n| r/2 \rceil$ si ha che

$$M \left(\frac{p}{|n|r} \right)^p \leq M 2^{-\lceil |n|r/2 \rceil} \leq M 2^{-\{(|n|r/2)-1\}} = (2M) e^{-\{(r|n| \log 2)/2\}}. \quad (\text{A.58})$$

Da (A.57) ed (A.58) segue (A.54) se poniamo:

$$\sigma := (r \log 2)/2, \quad C_* := \max_{0 \leq |n| \leq 2/r} |\hat{f}_n| e^{\sigma|n|}, \quad C := \max\{2M, C_*\}.$$

Assumiamo ora che valga (A.54). Poiché $\sum_{n \in \mathbb{Z}} |n|^p e^{-\sigma|n|} < \infty$ per ogni $p > 0$, dal punto (ii) dell'Osservazione A.14 segue che $f \in C_{\text{per}}^\infty$. Il punto (vi) della Proposizione A.18, la (A.54) ed il Lemma A.21 implicano, per ogni $x \in [0, 2\pi]$ e per ogni intero $k \geq 1$,

$$\begin{aligned} |f^{(k)}(x)| &= \left| \sum_{n \in \mathbb{Z}} \hat{f}_n (in)^k e^{inx} \right| \leq \sum_{n \in \mathbb{Z}} |\hat{f}_n| |n|^k \\ &\leq C \sum_{n \in \mathbb{Z}} e^{-\sigma|n|} |n|^k = (2C) \sum_{j=1}^{\infty} e^{-\sigma j} j^k \\ &\leq (2C) \left(1 + \frac{1}{\sigma}\right) \sigma^{-k} k!. \end{aligned} \quad (\text{A.59})$$

Inoltre, sempre da (A.54), segue che²²

$$|f(x)| \leq \sum_{n \in \mathbb{Z}} |\hat{f}_n| \leq C \sum_{n \in \mathbb{Z}} e^{-\sigma|n|} = C \frac{1 + e^{-\sigma}}{1 - e^{-\sigma}} \leq (2C) \left(1 + \frac{1}{\sigma}\right). \quad (\text{A.60})$$

Quindi $|f^{(k)}(x)| \leq M r^{-k} k!$ per ogni $x \in \mathbb{R}$ e per ogni $k \geq 0$ se scegliamo $r := \sigma e$ $M := (2C)(1 + \frac{1}{\sigma})$. Dal Teorema A.34 segue dunque l'asserto. ■

Convergenza di serie di Fourier

Passiamo a rispondere alla domanda fatta all'inizio di questo capitolo che ora può essere così riformulata: *sotto quali condizioni f (periodica) coincide con la sua serie di Fourier?*

Cominciamo con un risultato preliminare che dà condizioni sufficienti affinché il troncamento di una serie di Fourier converga a f in un punto:

Lemma A.22 (Dini) *Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ periodica ed integrabile su $[0, 2\pi]$. Fissato $x \in [0, 2\pi)$, se il rapporto incrementale $y \rightarrow F(y) := (f(x+y) - f(x))/y$ è integrabile (come funzione di y) in un intorno di²³ $y = 0$ allora*

$$\lim_{N \rightarrow \infty} \sum_{|n| \leq N} \hat{f}_n e^{inx} = f(x). \quad (\text{A.61})$$

Osservazione A.23 Una ipotesi che implica la validità del 'test di Dini' (:= 'integrabilità del rapporto incrementale') è che f sia *derivabile* in x . Infatti se f è derivabile in x esiste $\delta > 0$ tale che $|F(y) - f'(x)| < 1$ per $|y| < \delta$. Dunque

$$|F(y)| \leq |F(y) - f'(x)| + |f'(x)| \leq 1 + |f'(x)| := M. \quad (\text{A.62})$$

Dunque per ogni $0 < \rho < \delta$ si ha $\int_{\rho \leq |y| \leq \delta} |F(y)| dy \leq 2\delta M$ e quindi F è integrabile su $[-\delta, \delta]$.

²²Se $\delta = \sum_{j=2}^{\infty} \sigma^j / j!$, $(1 + e^{-\sigma}) / (1 - e^{-\sigma}) = (e^\sigma + 1) / (e^\sigma - 1) = (2 + \sigma + \delta) / (\sigma + \delta) \leq (2 + \sigma) / \sigma \leq 2(1 + \frac{1}{\sigma})$.

²³Più precisamente: se F è integrabile su $[-x, -x + 2\pi) \cap \{0 < |y| < r\}$ per r sufficientemente piccolo.

Esempio A.24 Sia f la funzione di periodo 2π che vale x^2 nell'intervallo $[-\pi, \pi)$. Si noti che la funzione f è di classe C_{per}^0 , è derivabile in ogni $x \neq \pi + 2k\pi$ ma non è C_{per}^1 . Calcolando i coefficienti di Fourier di f [usando (A.51) ed osservando che $e^{in\pi} = (-1)^n$ per ogni $n \in \mathbb{Z}$] si trova, per $n \neq 0$,

$$\begin{aligned}\hat{f}_n &= \frac{1}{2\pi} \int_{-\pi}^{\pi} x^2 e^{-inx} dx = \frac{1}{2\pi} \left[x^2 \frac{e^{-inx}}{-in} \right]_{-\pi}^{\pi} + \frac{1}{\pi in} \int_{-\pi}^{\pi} x e^{-inx} dx \\ &= \frac{1}{\pi in} \int_{-\pi}^{\pi} x e^{-inx} dx = \frac{1}{\pi in} \left[x \frac{e^{-inx}}{-in} \right]_{-\pi}^{\pi} - \frac{1}{\pi n^2} \int_{-\pi}^{\pi} e^{-inx} dx \\ &= \frac{1}{\pi n^2} \left[x e^{-inx} \right]_{-\pi}^{\pi} = (-1)^n \frac{2}{n^2},\end{aligned}$$

mentre

$$\hat{f}_0 = \frac{1}{2\pi} \int_{-\pi}^{\pi} x^2 dx = \frac{\pi^2}{3}.$$

Dunque dal Lemma A.22 segue che, per ogni $-\pi < x < \pi$,

$$f(x) = x^2 = \frac{\pi^2}{3} + 2 \sum_{\substack{n \neq 0 \\ n \in \mathbb{Z}}} \frac{(-1)^n}{n^2} e^{inx} = \frac{\pi^2}{3} + 4 \sum_{n=1}^{\infty} \frac{(-1)^n}{n^2} \cos nx. \quad (\text{A.63})$$

La serie in tale relazione converge totalmente e dunque, come nella Proposizione A.13, definisce una funzione C_{per} . Poiché, come già osservato, anche la funzione f è di classe C_{per} si ha che la relazione (A.63) vale sull'intervallo chiuso $[-\pi, \pi]$ e calcolando tale relazione in $x = \pi$ si perviene alla notevole identità

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}. \quad (\text{A.64})$$

Dimostrazione (del Lemma A.22) Sia $G(y) := \frac{f(x-y) - f(x)}{e^{-iy} - 1}$. Dalle ipotesi segue che tale funzione (complessa di variabile reale) è periodica di periodo 2π ed è integrabile su $[0, 2\pi]$: infatti, chiaramente $G(y)$ è integrabile su $[\varepsilon, 2\pi - \varepsilon]$, per $\varepsilon > 0$ e sufficientemente piccolo²⁴, e se $F(y) := (f(x+y) - f(x))/y$ e $g(y) := (e^{iy} - 1)/y$, si ha che $G(y) = F(-y)/g(-y)$ che è integrabile su²⁵ $[0, \varepsilon] \cup [2\pi - \varepsilon, 2\pi)$. Dunque $f(x-y) - f(x) = G(y)e^{-iy} - G(y)$ ed essendo ambo i membri di tale identità (come funzioni di y) periodici ed integrabili su $[0, 2\pi]$ possiamo calcolarne i coefficienti di Fourier (ossia moltiplicare ambo i membri per $e^{-iny}/(2\pi)$ ed integrare da 0 a 2π) ottenendo le relazioni²⁶

$$\hat{f}_0 = \hat{G}_1 - \hat{G}_0 + f(x), \quad \hat{f}_{-n} e^{-inx} = \hat{G}_{n+1} - \hat{G}_n \quad (\text{per } n \neq 0). \quad (\text{A.65})$$

Dunque, per ogni $N \geq 1$,

$$\begin{aligned}\sum_{|n| \leq N} \hat{f}_n e^{inx} &= \sum_{|n| \leq N} \hat{f}_{-n} e^{-inx} = f(x) + \sum_{|n| \leq N} (\hat{G}_{n+1} - \hat{G}_n) \\ &= \hat{G}_{N+1} - \hat{G}_{-N} + f(x).\end{aligned} \quad (\text{A.66})$$

Dal punto (iv) dell'Osservazione A.19 (e da quanto sopra detto su G) segue che $\lim_{N \rightarrow \pm\infty} \hat{G}_N = 0$ e dunque, prendendo il limite per $N \rightarrow \infty$ nella (A.66), si ottiene la (A.61). ■

²⁴ $f(x-y) - f(x)$ è integrabile (in y) su $[0, 2\pi]$ per ipotesi (essendo f integrabile su $[0, 2\pi]$ e periodica).

²⁵Essendo G periodica di periodo 2π , $\int_0^\varepsilon G(y) dy + \int_{2\pi-\varepsilon}^{2\pi} G(y) dy = \int_{-\varepsilon}^\varepsilon G(y) dy$ e F è integrabile su di un intorno di 0 per ipotesi e $g(0) = i \neq 0$ ed è continua.

²⁶Per il punto (ii) dell'Osservazione A.17 si trova che $\int_0^{2\pi} f(x-y) dy = \int_{x-2\pi}^x f(t) dt = \int_0^{2\pi} f(t) dt = 2\pi \hat{f}_0$.

Teorema A.25 *Assumiamo che $f \in C_{\text{per}}^p$ per un qualche $p \geq 1$. Allora, per ogni $\varepsilon > 0$, esiste N_0 tale che, se $N \geq N_0$, si ha*

$$\sup_{\mathbb{R}} \left| f(x) - \sum_{|n| \leq N} \hat{f}_n e^{inx} \right| \leq \frac{\varepsilon}{N^{p-\frac{1}{2}}}. \quad (\text{A.67})$$

Inoltre, vale la seguente ‘uguaglianza di Parseval’

$$\frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx = \sum_{n \in \mathbb{Z}} |\hat{f}_n|^2. \quad (\text{A.68})$$

In particolare, la (A.67) mostra che $\sum_{|n| \leq N} \hat{f}_n e^{inx}$, per $f \in C_{\text{per}}^1$ converge uniformemente a f .

Dimostrazione Dal Lemma A.22 e dall'Osservazione A.23 segue che

$$f(x) - \sum_{|n| \leq N} \hat{f}_n e^{inx} = \lim_{M \rightarrow \infty} \sum_{N < |n| \leq M} \hat{f}_n e^{inx}, \quad \forall x \in \mathbb{R}. \quad (\text{A.69})$$

Dalla (vii) della Proposizione A.18 (con $k = p$) segue che, dato $\varepsilon > 0$, esiste $N_0 > 1$ tale che, per ogni $N \geq N_0$, si ha

$$\sum_{n > N} |\hat{f}_n|^2 |n|^{2p} \leq \varepsilon^2 \left(p - \frac{1}{2} \right). \quad (\text{A.70})$$

Usando la disuguaglianza di Cauchy–Schwarz (1.13) e la (A.70), si ottiene, per ogni $M > N \geq N_0$,

$$\begin{aligned} \left| \sum_{N < |n| \leq M} \hat{f}_n e^{inx} \right| &\leq \sum_{N < |n| \leq M} |\hat{f}_n| = \sum_{N < |n| \leq M} (|\hat{f}_n| |n|^p) |n|^{-p} \\ &\leq \left(\sum_{N < |n| \leq M} |\hat{f}_n|^2 |n|^{2p} \right)^{\frac{1}{2}} \left(\sum_{N < |n| \leq M} |n|^{-2p} \right)^{\frac{1}{2}} \\ &\leq \left(\sum_{|n| > N} |\hat{f}_n|^2 |n|^{2p} \right)^{\frac{1}{2}} \left(2 \sum_{n=N+1}^{\infty} n^{-2p} \right)^{\frac{1}{2}} \\ &\leq \varepsilon \left(p - \frac{1}{2} \right)^{\frac{1}{2}} \left(2 \int_N^{\infty} \frac{1}{x^{2p}} dx \right)^{\frac{1}{2}} = \frac{\varepsilon}{N^{p-\frac{1}{2}}}. \end{aligned}$$

Prendendo, in tale relazione, il limite per $M \rightarrow \infty$, e ricordando la (A.69) si ottiene la (A.67). In particolare, (A.67) mostra che la convergenza di $\sum_{|n| \leq N} \hat{f}_n e^{inx}$ a f è uniforme; quindi si può passare al limite, per $N \rightarrow \infty$, nella formula (iii) della Proposizione A.18, ottenendo l'identità di Parseval. ■

Proposizione A.26 *Se $f \in C_{\text{per}}^0$ è C^1 a tratti²⁷ allora $\sum_{n \in \mathbb{Z}} |\hat{f}_n| < \infty$ e quindi la serie di Fourier di f converge totalmente a f .*

Lemma A.27 *Se $f \in C_{\text{per}}^0$ è C^1 a tratti allora $(\widehat{f'})_n = in \hat{f}_n$.*

²⁷ Una funzione $f : E \rightarrow \mathbb{R}$, con E intervallo di \mathbb{R} , è C^1 a tratti se esistono N punti in E $a_1 < \dots < a_N$ per cui $f \in C^1((a_{k-1}, a_k))$, per $2 \leq k \leq N$, ed esistono (finiti) i limiti, rispettivamente, da destra e da sinistra di f e f' in a_{k-1} ed a_k .

Dimostrazione Siano $0 \leq a_1 < \dots < a_N < 2\pi$ i punti di salto di f' , e poniamo $a_0 := 0$ e $a_{N+1} := 2\pi$. Allora

$$\begin{aligned} (\widehat{f'})_n &:= \frac{1}{2\pi} \int_0^{2\pi} f' e^{-inx} dx = \frac{1}{2\pi} \sum_{k=1}^{N+1} \int_{a_{k-1}}^{a_k} f' e^{-inx} dx \\ &= \frac{1}{2\pi} \sum_{k=1}^{N+1} [f e^{-inx}]_{a_{k-1}}^{a_k} + \frac{in}{2\pi} \sum_{k=1}^{N+1} \int_{a_{k-1}}^{a_k} f e^{-inx} dx \\ &= \frac{in}{2\pi} \int_0^{2\pi} f e^{-inx} dx := in \hat{f}_n, \end{aligned}$$

(si noti che poiché f è continua, la somma $\sum_{k=1}^{N+1} [f e^{-inx}]_{a_{k-1}}^{a_k}$ è una somma ‘telescopica’ di cui resta la differenza tra l’ultimo ed il primo termine e tali termini si cancellano per la periodicità). ■

Dimostrazione (della Proposizione A.26) Poiché f' è periodica e continua a tratti, f' è integrabile su $[0, 2\pi]$ e dunque dalla disuguaglianza di Bessel (punto (iv), Proposizione A.18) segue che $\sum |\widehat{f'}_n|^2 \leq \frac{1}{2\pi} \int_0^{2\pi} |f(x)|^2 dx < \infty$. Dunque dal Lemma e dalla disuguaglianza di Cauchy–Schwarz per successioni segue che

$$\begin{aligned} \sum_{n \neq 0} |\hat{f}_n| &= \sum_{n \neq 0} (|\hat{f}_n| |n|) |n|^{-1} \leq \left(\sum_{n \neq 0} |\hat{f}_n|^2 |n|^2 \right)^{\frac{1}{2}} \left(\sum_{n \neq 0} |n|^{-2} \right)^{\frac{1}{2}} \\ &= \left(\sum_{n \neq 0} |\widehat{f'}_n|^2 \right)^{\frac{1}{2}} \left(\sum_{n \neq 0} |n|^{-2} \right)^{\frac{1}{2}} < \infty. \quad \blacksquare \end{aligned}$$

La seguente versione del Lemma A.22 di Dini dà, sotto opportune ipotesi, informazioni sulla convergenza della serie di Fourier in punti dove $f(x)$ ha un salto. Ricordiamo prima le notazioni

$$f(x+) = \lim_{h \downarrow 0} f(x+h) := \lim_{y \rightarrow x, y > x} f(y), \quad f(x-) = \lim_{h \uparrow 0} f(x+h) := \lim_{y \rightarrow x, y < x} f(y).$$

Lemma A.28 Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ periodica di periodo 2π ed integrabile su $[0, 2\pi]$. Se $x \in [0, 2\pi]$ è tale che esistono i limiti $f(x\pm)$ e tale che la funzione

$$F(y) := \begin{cases} \frac{f(x+y) - f(x+)}{y}, & \text{se } y > 0, \\ \frac{f(x+y) - f(x-)}{y}, & \text{se } y < 0, \\ 0, & \text{se } y = 0, \end{cases} \quad (\text{A.71})$$

è integrabile in un intorno di zero, allora

$$\lim_{N \rightarrow \infty} \sum_{|n| \leq N} \hat{f}_n e^{inx} = \frac{f(x+) + f(x-)}{2}. \quad (\text{A.72})$$

Dimostrazione Sia

$$G(y) := \begin{cases} \frac{f(x-y) - f(x-)}{e^{-iy} - 1}, & \text{se } y > 0, \\ \frac{f(x-y) - f(x+)}{e^{-iy} - 1}, & \text{se } y < 0, \\ 0, & \text{se } y = 0. \end{cases}$$

Dalle ipotesi fatte (ragionando in maniera analoga a quanto fatto nella dimostrazione del Lemma A.22) segue che G è periodica ed integrabile su $[0, 2\pi]$. Seguendo lo schema della dimostrazione del Lemma A.22 troviamo la relazione

$$f(x-y) - \chi_+(y)f(x-) - \chi_-(y)f(x+) = G(y)e^{-iy} - G(y) \quad (\text{A.73})$$

dove

$$\chi_+(y) := \begin{cases} 1, & \text{se } y > 0, \\ 0, & \text{se } y \leq 0, \end{cases} \quad \chi_-(y) := \begin{cases} 0, & \text{se } y \geq 0, \\ 1, & \text{se } y < 0. \end{cases}$$

Moltiplicando per $e^{-iny}/(2\pi)$ ed integrando tra $-\pi$ e π ambo i membri della (A.73), troviamo

$$\begin{aligned} \hat{f}_0 &= \hat{G}_1 - \hat{G}_0 + \frac{f(x+) + f(x-)}{2}, \\ \hat{f}_{-n}e^{-inx} &= \hat{G}_{n+1} - \hat{G}_n + f(x-) \frac{1 - e^{-in\pi}}{2\pi in} + f(x+) \frac{e^{in\pi} - 1}{-2\pi in} \quad (n \neq 0). \end{aligned}$$

Osservando che $e^{\pm in\pi} = (-1)^n$ si ha che

$$\sum_{0 < |n| \leq N} \frac{1 - e^{-in\pi}}{n} = \sum_{\substack{|n| \leq N \\ n \text{ dispari}}} \frac{2}{n} = 0, \quad \sum_{0 < |n| \leq N} \frac{e^{in\pi} - 1}{n} = \sum_{\substack{|n| \leq N \\ n \text{ dispari}}} \frac{-2}{n} = 0.$$

Dunque $\sum_{|n| \leq N} \hat{f}_n e^{inx} = \frac{f(x+) + f(x-)}{2} + \hat{G}_{N+1} - \hat{G}_N$ da cui (come nella dimostrazione del Lemma A.22) segue l'asserto. ■

Complementi

Complemento A.1: Funzioni reali-analitiche

Abbiamo visto che una serie di potenze $u = \sum a_n x^n$ è derivabile infinite volte nell'intervallo aperto $\{x : |x| < R\}$ dove $R = \rho(u)$ denota il raggio di convergenza di u . Una domanda naturale è quindi: 'Esistono funzioni C^∞ che non siano rappresentabili tramite serie di potenze?' La risposta è affermativa e quindi le serie di potenze sono una classe funzionale più piccola delle funzioni indefinitamente differenziabili. Prima di motivare tale risposta, diamo due definizioni.

Definizione A.29 Sia $E \subseteq \mathbb{R}$, k un intero positivo o $+\infty$ ed f una funzione definita su E . Diremo che $f \in C^k(E)$, se esiste un aperto $A \supseteq E$ tale che $f \in C^k(A)$.

Tale definizione generalizza la nozione di funzione C^k e permette di parlare di funzioni derivabili su insiemi chiusi²⁸

Definizione A.30 Sia $f \in C^\infty(\{x_0\})$, si chiama serie di Taylor di f la seguente serie di potenze

$$\sum a_n (x - x_0)^n, \quad \text{dove} \quad a_n = \frac{f^{(n)}(x_0)}{n!}. \quad (\text{A.74})$$

Per il teorema sulla formula di Taylor, i troncamenti all'ordine N della serie di Taylor di²⁹ $f \in C^\infty(\{x_0\})$ approssimano la funzione f a meno di una quantità di ordine $O(|x - x_0|^{N+1})$.

²⁸Per esempio, $f \in C^\infty(\{0\})$ significa che esiste $r > 0$ tale che $f \in C^\infty(-r, r)$. Si noti che tale definizione è data per $k > 0$ ma non per $k = 0$; infatti esistono funzioni continue in un punto ma che non sono continue su alcun intorno di tale punto.

²⁹Ossia i polinomi di ordine N $\sum_{n=0}^N a_n (x - x_0)^n$ con $a_n = f^{(n)}(x_0)/n!$.

Esempio A.31 Si consideri ora la seguente funzione di una variabile reale

$$g(x) = \begin{cases} 0 & \text{se } x \leq 0 \\ \exp(-\frac{1}{x}) & \text{se } x > 0 \end{cases} \quad (\text{A.75})$$

Dimostriamo che $g \in C^\infty$. Cominciamo col dimostrare che per $x > 0$, la derivata k -esima di g ha la forma

$$g^{(k)}(x) = P_k\left(\frac{1}{x}\right) \exp\left(-\frac{1}{x}\right) \quad (\text{A.76})$$

dove $P_k(y)$ è un polinomio in y di grado $2k$. Infatti per $k = 1$, si ha che

$$g'(x) = \frac{1}{x^2} e^{-\frac{1}{x}}, \quad (x > 0) \quad (\text{A.77})$$

e quindi l'asserto è vero con $P_1(y) := y^2$. Assumiamo l'asserto vero per $0, \dots, k-1$ e dimostriamolo per k .

$$\begin{aligned} g^{(k)}(x) &= \left(g^{(k-1)}(x)\right)' \\ &= \left[P'_{k-1}\left(\frac{1}{x}\right)\left(-\frac{1}{x^2}\right) + P_{k-1}\left(\frac{1}{x}\right)\left(\frac{1}{x^2}\right)\right] e^{-\frac{1}{x}} \\ &:= P_k\left(\frac{1}{x}\right) e^{-\frac{1}{x}} \end{aligned}$$

il che implica che l'asserto è vero anche per k . Poiché per ogni $k > 0$ si ha che

$$\lim_{y \rightarrow \infty} y^k \exp(-y) = 0,$$

da (A.76) segue che³⁰, per ogni intero $k, p \geq 0$

$$\lim_{x \downarrow 0} g^{(k)}(x) x^{-p} = 0. \quad (\text{A.78})$$

In particolare (poiché $g^{(k)}(x) := 0$ per ogni $x < 0$) segue che $g^{(k)}(0+) = g^{(k)}(0-) = 0$, per ogni $k \geq 0$. Da tale relazione segue anche che $g^{(k)}(0) = 0$ per ogni k . Infatti per $k = 0$ è vero per definizione. Assumiamo che $k \geq 1$ e che $g^{(k-1)}(0) = 0$. Allora

$$g^{(k)}(0) = \lim_{h \rightarrow 0} \frac{g^{(k-1)}(h)}{h} = 0.$$

Dunque poiché g è chiaramente $C^\infty(0, \infty)$ e $C^\infty(-\infty, 0)$ segue che $g \in C^\infty(\mathbb{R})$. Ma allora (essendo $g^{(k)}(0) = 0$) la serie di Taylor di g è la serie banale ($a_n := 0$) che ha raggio di convergenza infinito. D'altra parte non può esistere nessun intorno di 0 in cui la somma della serie di Taylor (e cioè 0) eguagli g poiché $g(x) > 0$ per ogni $x > 0$.

Definizione A.32 Una funzione $f \in C^\infty(\{x_0\})$ a valori reali si dice **reale-analitica in x_0** se la sua serie di Taylor ha raggio di convergenza positivo e se $f(x) = \sum a_n(x-x_0)^n$ (con a_n definiti in (A.74)) in un intorno di x_0 . Una funzione $f \in C^\infty(E)$ si dice **reale-analitica su E** se f è (reale) analitica in ogni punto x_0 di E ; la classe di tali funzione si denota con $C^\omega(E)$.

Osservazione A.33 (i) Una funzione reale analitica è, dunque, una funzione che localmente si rappresenta come una serie di potenze coincidente con la sua serie di Taylor. Quindi la funzione g dell'Esempio A.31 è $C^\infty(\mathbb{R})$ ma non $C^\omega(\mathbb{R})$ [più precisamente $f \in C^\infty(\mathbb{R}) \cap C^\omega(\mathbb{R} \setminus \{0\})$ ma $f \notin C^\omega(\{0\})$].

(ii) Dal Capitolo 2 segue:

la somma e il prodotto di due funzioni $C^\omega(E)$ appartengono a $C^\omega(E)$; se $f \in C^\omega(E)$ e $f(x) \neq 0$ per ogni $x \in E$ allora $1/f \in C^\omega(E)$; se $f \in C^\omega(\{x_0\})$ e $g \in C^\omega(\{y_0\})$ con $y_0 := f(x_0)$ allora $g \circ f \in C^\omega(\{x_0\})$.

Dimostriamo, per esempio, l'affermazione relativa al reciproco. Fissiamo $x_0 \in E$; dalla definizione di C^ω segue che $f = \sum_{n \geq 0} a_n(x-x_0)^n$ in un intorno di x_0 e $a_n = f^{(n)}(x_0)/n!$. Si dimostra che $1/f$ è

³⁰ Si ricorda che $\lim_{x \downarrow x_0} f(x)$ è il limite di $f(x)$ per x che tende a x_0 con $x > x_0$; analogamente $\lim_{x \uparrow x_0} f(x)$ è il limite di $f(x)$ per x che tende a x_0 con $x < x_0$.

una serie di potenze in un intorno di x_0 e dall'Osservazione A.9 punto (i) si ha che tale serie coincide con la serie di Taylor di $1/f$ e questo significa che $1/f$ è analitica in intorno di x_0 . Ragionando in maniera analoga si dimostrano anche le altre affermazioni. ■

(iii) Dal Teorema A.7, segue che una serie di potenze $f = \sum a_n(x - x_0)^n$ con $R := \rho(f) > 0$ (e $a_n \in \mathbb{R}$) è di classe $C^\omega((x_0 - R, x_0 + R))$; il 'viceversa' di tale affermazione, in generale, non è però vero: la funzione $f(x) := 1/(1 + x^2)$, per il punto (ii), è $C^\omega(\mathbb{R})$ ma la serie di Taylor di f in 0 è data da

$$\frac{1}{1 + x^2} = \sum_{k=0}^{\infty} (-1)^k x^{2k},$$

che, come si verifica immediatamente, ha raggio di convergenza 1.

In vista di queste osservazioni è naturale aspettarsi che vi sia una caratterizzazione delle funzioni analitiche simile a quella che si ha per le serie di potenze. Infatti vale il seguente

Teorema A.34 (Caratterizzazione delle funzioni analitiche) *La funzione f è analitica in x_0 se e solo se $f \in C^\infty(\{x_0\})$ ed esistono due costanti positive M, r tali che*

$$\sup_{|x-x_0|<r} |f^{(n)}(x)| \leq M r^{-n} n!, \quad \forall n \geq 0. \quad (\text{A.79})$$

Dimostrazione Cominciamo con il dimostrare il 'se'. Supponiamo, quindi, che valga (A.79) e sia $0 < R < r$, allora

$$\sum_{n \geq 0} \left| \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \right| \leq M \sum_{n \geq 0} r^{-n} R^n < \infty. \quad (\text{A.80})$$

Sia, ora, x tale che $|x - x_0| < r$. Per la formula di Taylor (con resto in forma di Lagrange) si ha, per un qualche $x^{(N)}$ tra x e x_0 ,

$$\begin{aligned} \lim_{N \rightarrow \infty} \left| f(x) - \sum_{n=0}^N \frac{f^{(n)}(x_0)}{n!} (x - x_0)^n \right| &= \lim_{N \rightarrow \infty} \left| \frac{f^{(N+1)}(x^{(N)})}{(N+1)!} \right| |x - x_0|^{N+1} \\ &\leq \lim_{N \rightarrow \infty} M (|x - x_0| r^{-1})^{N+1} = 0. \end{aligned}$$

Quindi $f = \sum a_n(x - x_0)^n$ è analitica.

Viceversa, assumiamo che f sia analitica e che la sua serie di Taylor converga per $|x - x_0| < R$; fissiamo un r tale che $0 < r < R/2$. Allora, se a_n denotano i coefficienti della serie di Taylor di f , ricordando il Teorema A.7 ed osservando che

$$\binom{m}{k} \leq \sum_{h=0}^m \binom{m}{h} = 2^m \quad (\text{A.81})$$

per ogni $0 \leq k \leq m$, otteniamo:

$$\begin{aligned} \sup_{|x-x_0|<r} |f^{(n)}(x)| &= \sup_{|x-x_0|<r} \left| \sum_{k=0}^{\infty} a_{k+n} (n+k) \cdots (1+k) (x - x_0)^k \right| \\ &\leq \frac{1}{r^n} \sum_{k=0}^{\infty} |a_{k+n}| \frac{(n+k)!}{k!} r^{k+n} = \frac{1}{r^n} n! \sum_{k=0}^{\infty} |a_{k+n}| \binom{n+k}{k} r^{k+n} \\ &\leq \frac{1}{r^n} n! \sum_{k=0}^{\infty} |a_{k+n}| (2r)^{k+n} \leq \frac{1}{r^n} n! \sum_{k=0}^{\infty} |a_k| (2r)^k \\ &= r^{-n} n! M, \quad \text{con} \quad M := \sum_{k \geq 0} |a_k| (2r)^k. \quad \blacksquare \end{aligned}$$

Una classe interessante di funzioni C^∞ ma non C^ω sono funzioni che sono identicamente nulle al di fuori di un intervallo limitato.

Definizione A.35 *Il supporto di una funzione f , denotato $\text{supp}(f)$, è la chiusura dell'insieme $\{x : f(x) \neq 0\}$. Una funzione f si dice 'a supporto compatto' se il suo supporto è un insieme limitato (e quindi compatto). La classe delle funzioni $C^k(\mathbb{R})$ a supporto compatto si denota con $C_0^k(\mathbb{R})$.*

È immediato costruire funzioni $C_0 := C_0^0$; per esempio $f(x) := 1 - |x|$ per $|x| < 1$ e $f(x) := 0$ per $|x| \geq 1$. Più interessanti sono le funzioni C_0^∞ .

Esempio A.36 Sia g la funzione dell'Esempio A.31. La funzione $g_0(x) := g(x)g(1-x)$ è una funzione $C^\infty(\mathbb{R})$. Infatti $g_0 \in C^\infty$ essendo il prodotto di due funzioni C^∞ ed inoltre $g_0(x) = 0$ se $x < 0$ oppure se $x > 1$. D'altra parte $g_0(x) > 0$ se $x \in (0, 1)$. Quindi $\text{supp}(g_0) = [0, 1]$ e $g_0 \in C_0^\infty$.

Esempio A.37 Un altro modo di costruire funzioni C_0^∞ (sempre legato alla funzione esponenziale) è il seguente. Se g è la funzione definita nell'Esempio A.31, definiamo, per ogni $\varepsilon > 0$, la funzione

$$\varphi_\varepsilon(x) := g(\varepsilon^2 - x^2) = \begin{cases} \exp\left(-\frac{1}{\varepsilon^2 - x^2}\right), & \text{se } |x| < \varepsilon, \\ 0, & \text{se } |x| \geq \varepsilon. \end{cases} \quad (\text{A.82})$$

Tale funzione, essendo composizione di funzioni $C^\infty(\mathbb{R})$, è $C^\infty(\mathbb{R})$. Inoltre è chiaro che $\varphi_\varepsilon \geq 0$, che $\varphi_\varepsilon(x) > 0$ se $|x| < \varepsilon$ e che $\varphi_\varepsilon(x) := 0$ se $|x| \geq \varepsilon$. Quindi φ_ε è una funzione a supporto compatto e $\text{supp}(\varphi_\varepsilon) = \{x : |x| \leq \varepsilon\}$.

Esempio A.38 Siano $R > r > 0$ e sia $x_0 \in \mathbb{R}$. Costruiamo una funzione $\chi \in C_0^\infty$ che abbia le seguenti proprietà:

(i) $0 \leq \chi \leq 1$; (ii) $\chi(x) := 1$ se $|x - x_0| < r$; (iii) $\text{supp}(\chi) = \{x : |x - x_0| \leq R\}$.

Sia $g_1(x) := c \int_{-\infty}^x g_0(t) dt$ dove $c := \left(\int_0^1 g_0(t) dt\right)^{-1}$ con g_0 definita nell'Esempio A.36. La funzione g_1 è chiaramente C^∞ , è monotona non decrescente (la sua derivata è $cg_0(x)$ che è una funzione non negativa), vale 0 se $x < 0$ e vale 1 se $x > 1$.

È ora elementare controllare che la funzione cercata può essere definita come

$$\chi(x) := g_1\left(\frac{x_0 - x + R}{R - r}\right) g_1\left(\frac{x - x_0 + R}{R - r}\right). \quad (\text{A.83})$$

Complemento A.2: Trasformata di Fourier

Definizione A.39 Sia $f \in \mathcal{R}_1(\mathbb{R})$. Si definisce la trasformata di Fourier³¹ di f come

$$\hat{f}(\xi) := \int_{-\infty}^{\infty} f(x) e^{-ix\xi} dx, \quad dx := \frac{dx}{\sqrt{2\pi}}. \quad (\text{A.84})$$

Proposizione A.40 La trasformata di Fourier \hat{f} di una funzione $f \in \mathcal{R}_1(\mathbb{R})$ gode delle seguenti proprietà:

(i) \hat{f} è uniformemente continua su \mathbb{R} e

$$\sup_{\mathbb{R}} |\hat{f}| \leq \|f\|_1 := \int_{-\infty}^{\infty} |f(x)| dx. \quad (\text{A.85})$$

(ii) Sia p un intero positivo. Se $x \rightarrow x^k f(x) \in \mathcal{R}_1(\mathbb{R})$ per ogni intero $0 \leq k \leq p$, allora $\hat{f} \in C^p(\mathbb{R})$ e

$$\partial_\xi^k \hat{f}(\xi) = (-i)^k \widehat{(x^k f)}(\xi), \quad \forall k \leq p; \quad (\text{A.86})$$

inoltre le funzioni $\partial_\xi^k \hat{f}$ sono, per $0 \leq k \leq p$, uniformemente continue su \mathbb{R} .

(iii) Sia p un intero positivo. Se $f \in C^p(\mathbb{R})$ e $f^{(k)} \in \mathcal{R}_1(\mathbb{R})$ per ogni $k \leq p$, allora

$$\widehat{f^{(k)}}(\xi) = (i\xi)^k \hat{f}(\xi), \quad \forall k \leq p. \quad (\text{A.87})$$

(iv) Sia p un intero positivo. Se $f \in C^p(\mathbb{R})$ e $f^{(k)} \in \mathcal{R}_1(\mathbb{R})$ per ogni $k \leq p$, allora

$$|\hat{f}(\xi)| \leq \frac{\|f^{(p)}\|_1}{|\xi|^p}, \quad \forall \xi \neq 0, \quad (\text{A.88})$$

e

$$|\hat{f}(\xi)| \leq \frac{M}{1 + |\xi|^p}, \quad M := 2 \max\{\|f\|_1, \|f^{(p)}\|_1\}. \quad (\text{A.89})$$

³¹Jean-Baptiste Joseph Fourier, 1768 (Auxerre) -1830 (Parigi).

Nel corso della dimostrazione useremo il seguente semplice

Lemma A.41 *Siano g_j ($j \geq 1$), g e G funzioni in $\mathcal{R}_1(\mathbb{R})$ tali che*

$$|g_j| \leq G. \quad (\text{A.90})$$

Sia $\{A_k\}$ una successione di insiemi, misurabili secondo Peano–Jordan, su cui g_j, g e G siano integrabili secondo Riemann e tali che $A_k \subseteq A_{k+1}$, $\bigcup_k A_k = \mathbb{R}$. Se, per ogni k , $g_j \rightarrow g$ uniformemente su A_k , allora $\|g_j - g\|_1 \rightarrow 0$.

Dimostrazione Sia $\varepsilon > 0$. Dalla definizione di integrale generalizzato segue che esiste k tale che

$$\int_{A_k^c} |g| \leq \frac{\varepsilon}{4}, \quad \int_{A_k^c} |G| \leq \frac{\varepsilon}{4}. \quad (\text{A.91})$$

Da tale relazione e da (A.90) segue anche che, per ogni $j \geq 1$,

$$\int_{A_k^c} |g_j| \leq \frac{\varepsilon}{4}. \quad (\text{A.92})$$

Poiché g_j converge a g uniformemente su A_k , esiste j_0 tale che

$$|g_j(x) - g(x)| < \frac{\varepsilon}{2 \operatorname{mis}(A_k)}$$

per ogni $x \in A_k$ e per ogni $j \geq j_0$. Dunque, per ogni $j \geq j_0$,

$$\int_{\mathbb{R}} |g_j - g| dx = \int_{A_k^c} |g_j - g| dx + \int_{A_k} |g_j - g| dx \leq \int_{A_k^c} (|g_j| + |g|) dx + \frac{\varepsilon}{2} \leq \varepsilon. \quad \blacksquare$$

Dimostrazione (della Proposizione A.40) La (A.85) è ovvia. Dimostriamo l'uniforme continuità di \hat{f} . Sia $\varepsilon > 0$. Poiché $f \in \mathcal{R}_1(\mathbb{R})$, esiste un insieme A misurabile secondo Peano–Jordan, su cui f è Riemann integrabile e tale che

$$\int_{A^c} |f| dx < \frac{\varepsilon}{4}.$$

Sia δ_0 tale che

$$|e^{it} - 1| < \frac{\varepsilon}{2(1 + \|f\|_1)}, \quad \forall |t| < \delta_0; \quad (\text{A.93})$$

sia $M > 0$ tale che $A \subseteq [-M, M]$ e sia $\delta := \delta_0/M$. Allora per ogni $\xi \in \mathbb{R}$ e per ogni $h \in \mathbb{R}$ tale che $|h| < \delta$ si ha che

$$\begin{aligned} |\hat{f}(\xi + h) - \hat{f}(\xi)| &= \left| \int_{\mathbb{R}} f(x) e^{-ix\xi} (e^{-ixh} - 1) dx \right| \\ &\leq \int_{\mathbb{R}} |f(x)| |e^{-ixh} - 1| dx \\ &= \int_{A^c} |f(x)| |e^{-ixh} - 1| dx + \int_A |f(x)| |e^{-ixh} - 1| dx \\ &\leq 2 \int_{A^c} |f(x)| dx + \int_A |f(x)| |e^{-ixh} - 1| dx \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2(1 + \|f\|_1)} \int_A |f(x)| dx \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2(1 + \|f\|_1)} \|f\|_1 \leq \varepsilon, \end{aligned}$$

avendo usato il fatto che se $|x| \leq M$ allora si ha che $|-ixh| < \delta_0$ il che permette l'uso della (A.93).

(ii) Dimostriamo dapprima la (A.86) per $p = 1$. Sia $h_j \neq 0$ una qualunque successione convergente a 0 e sia

$$g_j(x) := f(x) e^{-ix\xi} \frac{e^{-ixh_j} - 1}{h_j}, \quad g(x) := (-i)x f(x) e^{-ix\xi}.$$

Osserviamo, ora, che: $\frac{e^{-ixh_j}-1}{h_j}$ converge a $(-ix)$ uniformemente su³² $[-a, a]$ per ogni $a > 0$; che³³

$$\left| \frac{e^{-ixh_j}-1}{h_j} \right| \leq |x|;$$

che $|g_j|, |g| \leq G(x) := |x||f(x)| \in \mathcal{R}_1(\mathbb{R})$. Sia ora A_k una successione di insiemi misurabili secondo Peano-Jordan su cui $xf(x)$ sia Riemann integrabile e tali che $\sup_k \int_{A_k} |xf| < \infty$. Poiché gli A_k sono limitati, dalle osservazioni fatte segue che

$$|g_j(x) - g(x)| \leq \left(\sup_{A_k} |xf| \right) \left| \frac{e^{-ixh_j}-1}{h_j} + ix \right|$$

tende uniformemente a zero su A_k e dunque la (A.86) per $p = 1$ segue dal lemma. Il caso con p arbitrario segue per induzione: supponiamo che $x^p f \in \mathcal{R}_1(\mathbb{R})$ e che la (A.86) sia vera per $p-1$. Allora poiché $x(x^{p-1}f) \in \mathcal{R}_1(\mathbb{R})$, usando la (A.86) con $k = 1$, si ha che

$$\partial_\xi(\widehat{x^{p-1}f})(\xi) = -i(\widehat{x^p f})(\xi)$$

che, insieme alla (A.86) con $k = p-1$, implica la (A.86) anche con $k = p$. L'uniforme continuità di $\partial_\xi^k \hat{f}$ segue, ora, dalla (A.86) e dal punto (i).

(iii) Sia $p = 1$ e assumiamo, dapprima, che $f(x) \rightarrow 0$ quando $|x| \rightarrow \infty$. Allora,

$$\begin{aligned} \int_{\mathbb{R}} f'(x) e^{-ix\xi} dx &= \lim_{R \rightarrow \infty} \int_{-R}^R f'(x) e^{-ix\xi} dx \\ &= \lim_{R \rightarrow \infty} \left[\frac{f(x) e^{-ix\xi}}{\sqrt{2\pi}} \right]_{-R}^R + i\xi \lim_{R \rightarrow \infty} \int_{-R}^R f e^{-ix\xi} dx \\ &= i\xi \hat{f}(\xi). \end{aligned}$$

In generale da $f \in \mathcal{R}_1(\mathbb{R})$ non segue che $|f(x)| \rightarrow 0$ per $|x| \rightarrow \infty$ ma è sempre possibile trovare due successioni $R_j \rightarrow \infty$ e $R'_j \rightarrow -\infty$ tali che $f(R_j)$ e $f(R'_j)$ tendano a 0 per³⁴ $j \rightarrow \infty$: questo è sufficiente per ripetere l'argomento dato. Il caso con p arbitrario si ottiene per iterazione.

La (A.88) segue immediatamente da (A.87) insieme a (A.85). Per la (A.89) si usi la (A.85) nell'intervallo $|\xi| \leq 1$ e la (A.88) per $|\xi| \geq 1$. ■

Osservazione A.42 Chiaramente se³⁵ $f \in C_0^p(\mathbb{R})$ allora $f^{(k)} \in C_0(\mathbb{R}) \subseteq \mathcal{R}_1(\mathbb{R})$ per ogni $k \leq p$ e dunque, per tali f vale la (A.89). In particolare se $f \in C_0^\infty(\mathbb{R})$ la sua trasformata di Fourier decade più rapidamente di qualunque potenza.

Il prossimo risultato spiega come ricostruire la funzione f a partire dalla sua trasformata di Fourier.

Proposizione A.43 (Teorema di inversione per funzioni C_0^2) Sia $f \in C_0^2(\mathbb{R})$, allora

$$f(x) = \int_{-\infty}^{\infty} \hat{f}(\xi) e^{ix\xi} d\xi, \quad \forall x \in \mathbb{R}. \quad (\text{A.94})$$

Vale inoltre la seguente identità di Parseval³⁶:

$$\int_{-\infty}^{\infty} |\hat{f}(\xi)|^2 d\xi = \int_{-\infty}^{\infty} |f(x)|^2 dx. \quad (\text{A.95})$$

³²Dato $\varepsilon > 0$, sia δ tale che $\left| \frac{e^z-1}{z} - 1 \right| < \varepsilon/a$ per ogni $z \in \mathbb{C}$ con $0 < |z| < \delta$ e sia j_0 tale che $|h_j| < \delta/a, \forall j \geq j_0$. Allora, $\forall 0 < |x| \leq a$, si ha che $\left| \frac{e^{-ixh_j}-1}{h_j} + ix \right| = |x| \left| \frac{e^{-ixh_j}-1}{-ixh_j} - 1 \right| \leq a\varepsilon/a$.

³³ Per ogni $0 \neq t \in \mathbb{R}$, $\left| \frac{e^{it}-1}{t} \right| \leq 1$. Infatti $|e^{it}-1| = \left| \int_0^t e^{is} ds \right| \leq \int_0^{|t|} |e^{is}| ds = |t|$.

³⁴Per esempio, $\sum_{k=1}^{\infty} \int_{k-1}^k |f| = \int_0^{\infty} |f| < \infty$ e dunque per ogni $j \geq 1$ esiste $k = k_j$ tale che $\int_{k_j-1}^{k_j} |f| < 1/j$. Da questo segue che esiste $R_j \in [k_j-1, k_j]$ tale che $|f(R_j)| < 1/j$.

³⁵ Si ricorda che, per un aperto $E \subseteq \mathbb{R}^n$, $C_0^p(E)$ denota la classe delle funzioni C^p con supporto compatto contenuto in E .

³⁶Marc-Antoine Parseval des Chênes, 1755 (Rosières-aux-Salines) -1836 (Parigi).

La dimostrazione, oltre che sulle proprietà della trasformata e delle serie di Fourier è basata sulle approssimazioni discrete dell'integrale di Riemann su \mathbb{R} ; il seguente risultato sarà sufficiente per i nostri scopi.

Lemma A.44 *Sia $\varphi \in C(\mathbb{R})$ tale che esistano due costanti $M > 0$ e $\alpha > 1$ tali che*

$$|\varphi(x)| \leq \frac{M}{1 + |x|^\alpha}, \quad \forall x \in \mathbb{R}. \quad (\text{A.96})$$

Allora $\varphi \in \mathcal{R}_1(\mathbb{R})$ e

$$\lim_{\substack{\delta \rightarrow 0 \\ \delta > 0}} \delta \sum_{n \in \mathbb{Z}} \varphi(n\delta) = \int_{-\infty}^{\infty} \varphi(x) dx. \quad (\text{A.97})$$

Dimostrazione Da (A.96) segue immediatamente che $\varphi \in \mathcal{R}_1(\mathbb{R})$. Dimostriamo, ora, che dalle ipotesi segue che per ogni $\varepsilon > 0$ esiste $R > 2$ tale che³⁷

$$\int_{\{|x| \geq R-1\}} |\varphi(x)| dx < \varepsilon; \quad \delta \sum_{|n| \geq [R/\delta]} |\varphi(\delta n)| < \varepsilon, \quad \forall 0 < \delta < 1. \quad (\text{A.98})$$

Infatti, dalla (A.96) segue che

$$\int_{\{|x| \geq R-1\}} |\varphi(x)| dx \leq 2M \int_{R-1}^{\infty} \frac{1}{x^\alpha} dx = \frac{2M}{\alpha-1} \frac{1}{(R-1)^{\alpha-1}},$$

che implica la prima delle (A.98) con $R \geq R_0(\varepsilon) > 1$. Analogamente, per ogni $\delta \in (0, 1)$ e per ogni $R > 2$ si ha che

$$\begin{aligned} \delta \sum_{|n| \geq [R/\delta]} |\varphi(\delta n)| &\leq \delta M \sum_{|n| \geq [R/\delta]} \frac{1}{(|n|\delta)^\alpha} = \frac{2M}{\delta^{\alpha-1}} \sum_{j \geq [R/\delta]} \frac{1}{j^\alpha} \\ &\leq \frac{2M}{\delta^{\alpha-1}} \int_{[R/\delta]-1}^{\infty} \frac{dx}{x^\alpha} = 2M \int_{\delta[R/\delta]-\delta}^{\infty} \frac{dt}{t^\alpha} \\ &\leq 2M \int_{R-2}^{\infty} \frac{dt}{t^\alpha} = \frac{2M}{\alpha-1} \frac{1}{(R-2)^{\alpha-1}}, \end{aligned}$$

da cui segue la seconda delle (A.98) per R sufficientemente grande; nell'ultima disuguaglianza abbiamo usato il fatto che $\delta[R/\delta] - \delta \geq R - 2\delta \geq R - 2$.

Sia ora $\varepsilon > 0$ e sia R tale che

$$\int_{\{|x| > R-1\}} |\varphi(x)| dx < \frac{\varepsilon}{3}, \quad \delta \sum_{|n| \geq [R/\delta]} |\varphi(\delta n)| < \frac{\varepsilon}{3}. \quad (\text{A.99})$$

Sia $0 < \delta_0 < 1$ tale che

$$|\varphi(x) - \varphi(y)| < \frac{\varepsilon}{6R}, \quad \forall x, y \in [-R, R], \quad |x - y| < \delta_0. \quad (\text{A.100})$$

Per ogni $0 < \delta < \delta_0$ poniamo

$$N_\delta := [R/\delta], \quad R_\delta := \delta N_\delta, \quad (\text{A.101})$$

cosicch 

$$\frac{R}{\delta} - 1 < N_\delta \leq \frac{R}{\delta}, \quad R - \delta < R_\delta \leq R. \quad (\text{A.102})$$

Quindi,

$$\begin{aligned} &\left| \int_{-\infty}^{\infty} \varphi(x) dx - \delta \sum_{n \in \mathbb{Z}} \varphi(\delta n) \right| \\ &\leq \left| \int_{-R_\delta}^{R_\delta} \varphi(x) dx - \delta \sum_{n=-N_\delta}^{N_\delta-1} \varphi(\delta n) \right| + \int_{\{|x| \geq R_\delta\}} |\varphi(x)| dx + \delta \sum_{|n| \geq N_\delta} |\varphi(\delta n)| \\ &= \left| \sum_{n=-N_\delta}^{N_\delta-1} \int_{\delta n}^{\delta n + \delta} (\varphi(x) - \varphi(\delta n)) dx \right| + \int_{\{|x| \geq R_\delta\}} |\varphi(x)| dx + \delta \sum_{|n| \geq N_\delta} |\varphi(\delta n)| \\ &\leq \sum_{n=-N_\delta}^{N_\delta-1} \int_{\delta n}^{\delta n + \delta} |\varphi(x) - \varphi(\delta n)| dx + \int_{\{|x| \geq R-1\}} |\varphi(x)| dx + \delta \sum_{|n| \geq N_\delta} |\varphi(\delta n)| \\ &\leq 2N_\delta \delta \frac{\varepsilon}{6R} + \frac{2}{3}\varepsilon = R_\delta \frac{\varepsilon}{3R} + \frac{2}{3}\varepsilon \leq \varepsilon. \quad \blacksquare \end{aligned}$$

³⁷ Come al solito $[x]$ denota il pi  grande intero $m \leq x$.

Dimostrazione (della Proposizione A.43) Sia $T_0 > 0$ tale che $\text{supp } f \subseteq [-T_0/2, T_0/2]$ e, per $T \geq T_0$, sia f_T la funzione periodica di periodo T che coincide con f in $[-T/2, T/2]$. Allora $f_T \in C^2(\mathbb{R})$ e dai risultati sulle serie di Fourier³⁸ segue che, per ogni $|x| \leq T/2$,

$$f(x) = f_T(x) = \sum_{n \in \mathbb{Z}} \hat{f}_{T,n} e^{i \frac{2\pi}{T} n x}, \quad (|x| \leq T/2), \quad (\text{A.103})$$

dove la serie converge totalmente e

$$\begin{aligned} \hat{f}_{T,n} &:= \frac{1}{T} \int_{-T/2}^{T/2} f(x) e^{-i \frac{2\pi}{T} n x} dx \\ &= \frac{1}{\sqrt{2\pi}} \frac{2\pi}{T} \int_{-\infty}^{\infty} f(x) e^{-i \frac{2\pi}{T} n x} dx \\ &:= \frac{1}{\sqrt{2\pi}} \frac{2\pi}{T} \hat{f}\left(\frac{2\pi}{T} n\right). \end{aligned} \quad (\text{A.104})$$

Dunque, per $T \geq 2|x|$,

$$f(x) = \frac{1}{\sqrt{2\pi}} \sum_{n \in \mathbb{Z}} \frac{2\pi}{T} \hat{f}\left(\frac{2\pi}{T} n\right) e^{i \frac{2\pi}{T} n x}.$$

Ma dal punto (iv) della Proposizione A.40 segue che

$$|\hat{f}(\xi)| \leq \frac{M}{1 + |\xi|^2},$$

per qualche $M > 0$ e quindi la (A.94) segue dal Lemma A.44 applicato alla funzione

$$\varphi(\xi) := \hat{f}(\xi) \exp(ix\xi)$$

(e con $\delta = 2\pi/T$).

Analogamente (e usando le stesse notazioni), dalla formula di Parseval per serie di Fourier e dalla (A.104) segue che

$$\frac{1}{T} \int_{-T/2}^{T/2} |f_T|^2 dx = \sum_{n \in \mathbb{Z}} |\hat{f}_{T,n}|^2 = \frac{2\pi}{T^2} \sum_{n \in \mathbb{Z}} \left| \hat{f}\left(\frac{2\pi}{T} n\right) \right|^2.$$

Moltiplicando per T tale relazione e mandando $T \rightarrow \infty$ si ottiene, per il³⁹ Lemma A.44, la relazione (A.95). ■

La Proposizione A.43 si generalizza in vari modi; per esempio vale la seguente

Proposizione A.45 *Sia $f \in C^2(\mathbb{R})$ tale che $f^{(k)} \in \mathcal{R}_1(\mathbb{R})$ per $k = 0, 1, 2$. Allora vale la formula (A.94). Se, inoltre, $f \in \mathcal{R}_2(\mathbb{R})$, allora vale anche l'identità di Parseval (A.95).*

Dimostrazione L'idea della dimostrazione è basata sull'approssimare f con funzioni C^2 a supporto compatto. Sia $\varphi \in C^\infty$ una funzione monotona non crescente tale che:

$$\varphi(x) := \begin{cases} 1, & \text{se } x \leq 0, \\ 0, & \text{se } x \geq 1. \end{cases} \quad (\text{A.105})$$

A partire da φ , per $R > 0$, costruiamo una funzione ψ_R pari, di classe C^∞ , con supporto in $[-R-1, R+1]$ e che valga 1 per $|x| \leq R$, ponendo:

$$\psi_R(x) := \begin{cases} \varphi(x-R), & \text{se } x \geq 0, \\ \varphi_R(-x), & \text{se } x < 0. \end{cases} \quad (\text{A.106})$$

³⁸Per una funzione periodica di periodo T (ed integrabile) $\hat{f}_n = \frac{1}{T} \int_{-T/2}^{T/2} f(x) e^{-i \frac{2\pi}{T} n x} dx$ e la sua serie di Fourier è $\sum_{n \in \mathbb{Z}} \hat{f}_n e^{i \frac{2\pi}{T} n x}$.

³⁹Si ricordi l'osservazione A.42.

Poniamo $f_R := f\psi_R$. Chiaramente $f_R \in C_0^2$ e quindi vale la (A.94) con f_R al posto di f . Poiché f_R coincide con f se $|x| \leq R$, per ogni $R > 0$ e per ogni $|x| \leq R$, si ha

$$f(x) = \int_{-\infty}^{\infty} \hat{f}_R(\xi) e^{ix\xi} d\xi, \quad |x| \leq R. \quad (\text{A.107})$$

Per prendere il limite per $R \rightarrow \infty$ in tal relazione ed ottenere la (A.94), useremo il Lemma A.41. Dalle ipotesi su f e dalla Proposizione A.40, (A.89), segue che

$$|\hat{f}(\xi)| \leq \frac{M_1}{1 + |\xi|^2}, \quad M_1 := 2 \max\{\|f\|_1, \|f^{(2)}\|_1\}. \quad (\text{A.108})$$

Osserviamo che, se

$$c := \max\{1, \sup_{[0,1]} |\varphi'|, \sup_{[0,1]} |\varphi''|\}, \quad (\text{A.109})$$

allora

$$\max\{1, \sup_{\mathbb{R}} |\psi'_R|, \sup_{\mathbb{R}} |\psi''_R|\} \leq c. \quad (\text{A.110})$$

Dunque, essendo $f''_R := (f\psi_R)'' = f''\psi_R + 2f'\psi'_R + f\psi''_R$, si ha che

$$\|f''_R\|_1 \leq \|f''\|_1 + 2\|f'\psi'_R\|_1 + \|f\psi''_R\|_1 \leq c(\|f\|_1 + \|f'\|_1 + \|f''\|_1). \quad (\text{A.111})$$

Quindi, poiché $\|f_R\|_1 \leq \|f\|_1$, per la (A.89), si ha che

$$|\hat{f}_R(\xi)| \leq \frac{M_2}{1 + |\xi|^2}, \quad M_2 := 2c(\|f\|_1 + \|f'\|_1 + \|f''\|_1). \quad (\text{A.112})$$

Inoltre, per ogni $x \in \mathbb{R}$, $\hat{f}_R(\xi) \exp(ix\xi) \rightarrow \hat{f}(\xi) \exp(ix\xi)$ uniformemente per $\xi \in \mathbb{R}$:

$$\begin{aligned} |\hat{f}_R(\xi) - \hat{f}(\xi)| &:= \left| \int_{-\infty}^{\infty} f(x)(\psi_R(x) - 1) e^{-ix\xi} dx \right| \\ &\leq \int_{-\infty}^{\infty} |f(x)| |\psi_R(x) - 1| dx \\ &\leq \int_{\{|x| \geq R\}} |f(x)| dx \end{aligned} \quad (\text{A.113})$$

e quest'ultimo integrale tende a 0 quando $R \rightarrow \infty$ essendo $f \in \mathcal{R}_1(\mathbb{R})$. Ora, se R_j è una qualunque successione tendente a $+\infty$, e se

$$g_j(\xi) := \hat{f}_{R_j}(\xi) \exp(ix\xi), \quad g(\xi) := \hat{f}(\xi) \exp(ix\xi), \quad G(\xi) := M_2/(1 + |\xi|^2),$$

la (A.94) segue dal Lemma A.41.

Ora, assumiamo anche che $f \in \mathcal{R}_2(\mathbb{R})$. Dalla (A.95) per funzioni C_0^2 segue che

$$\int_{-\infty}^{\infty} |\hat{f}_R(\xi)|^2 d\xi = \int_{-\infty}^{\infty} |f_R(x)|^2 dx. \quad (\text{A.114})$$

Poiché

$$\int_{-R}^R |f|^2 \leq \int_{-\infty}^{\infty} |f_R|^2 \leq \int_{-\infty}^{\infty} |f|^2,$$

e poiché $f \in \mathcal{R}_2(\mathbb{R})$, si ha che

$$\lim_{R \rightarrow \infty} \int_{-R}^R |f|^2 = \int_{-\infty}^{\infty} |f|^2. \quad (\text{A.115})$$

Poiché⁴⁰

$$\left| |\hat{f}_R(\xi)|^2 - |\hat{f}(\xi)|^2 \right| \leq |\hat{f}_R(\xi) - \hat{f}(\xi)| |\hat{f}_R(\xi) + \hat{f}(\xi)|,$$

da (A.113) e dalla limitatezza di \hat{f} e \hat{f}_R (vedi (A.108) e (A.112)) segue che $|\hat{f}_R(\xi)|^2$ converge, per $R \rightarrow \infty$, a $|\hat{f}(\xi)|^2$ uniformemente su \mathbb{R} . Queste osservazioni permettono di prendere il limite per $R \rightarrow \infty$ in (A.114) ed ottenere, per il Lemma A.41, la (A.95). ■

Concludiamo questa breve discussione sulla trasformata di Fourier con il ‘Lemma di Riemann-Lebesgue’:

⁴⁰Si ricorda che per ogni coppia di numeri complessi z e w si ha che $||z|^2 - |w|^2| \leq |z - w| |z + w|$.

Proposizione A.46 Sia $f \in \mathcal{R}_1(\mathbb{R})$. Allora $\lim_{|\xi| \rightarrow \infty} \hat{f}(\xi) = 0$.

Dimostrazione Sia $\varepsilon > 0$. Poiché $f \in \mathcal{R}_1(\mathbb{R})$ esiste un aperto $A \subseteq \mathbb{R}$ misurabile secondo Peano-Jordan, su cui f è Riemann integrabile e

$$\int_{A^c} |f| \leq \frac{\varepsilon}{3}. \quad (\text{A.116})$$

Inoltre⁴¹ esiste $\psi \in C_0^\infty(A)$ tale che

$$\int_A |f - \psi| dx < \frac{\varepsilon}{3}. \quad (\text{A.117})$$

Sia, ora, $M > 0$ tale che $|\hat{\psi}(\xi)| \leq \varepsilon/3$ per ogni $|\xi| \geq M$ (tale M esiste per l'osservazione A.42). Allora per $|\xi| \geq M$ si ha che⁴²

$$\begin{aligned} |\hat{f}(\xi)| &= \left| \int_{\mathbb{R}} f(x) e^{-ix\xi} \right| \leq \left| \int_A f(x) e^{-ix\xi} \right| + \int_{A^c} |f| \\ &\leq \frac{\varepsilon}{3} + \left| \int_A f(x) e^{-ix\xi} \right| \\ &\leq \frac{\varepsilon}{3} + \int_A |f - \psi| + \left| \int_A \psi e^{-ix\xi} \right| \\ &= \frac{\varepsilon}{3} + \int_A |f - \psi| + |\hat{\psi}(\xi)| \\ &\leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon. \quad \blacksquare \end{aligned}$$

Esercizi

Esercizio A.1 Dimostrare che se u è una serie di potenze dispari [$u(x) = -u(-x)$], allora tutti i coefficienti pari sono nulli mentre se u è pari [$u(x) = u(-x)$], sono nulli tutti i coefficienti dispari.

Esercizio A.2* Verificare la relazione $\sin x \cos x = \frac{1}{2} \sin 2x$ usando la serie di Taylor; cioè verificare:

$$\left(\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \right) \left(\sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \right) = \sum_{n=0}^{\infty} \frac{(-1)^n}{2} \frac{(2x)^{2n+1}}{(2n+1)!}.$$

Esercizio A.3 Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ periodica e integrabile su $[0, 2\pi]$. Dimostrare che $\int_0^x f(t) dt$ è periodica se e solo se $\int_0^{2\pi} f(t) dt = 0$.

Esercizio A.4 Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ periodica e integrabile su $[0, 2\pi]$ e siano a_n e b_n i coefficienti di Fourier 'reali' definiti in (A.49). Dimostrare le seguenti affermazioni:

- (i) f è pari⁴³ $\implies \hat{f}_n \in \mathbb{R}, \forall n \in \mathbb{Z} \iff b_n = 0, \forall n \geq 1$.
- (ii) f è dispari $\implies i\hat{f}_n \in \mathbb{R}, \forall n \in \mathbb{Z} \iff a_n = 0, \forall n \geq 0$.
- (iii) Se f è C^1 allora (i) e (ii) valgono con ' \iff ' al posto di ' \implies '.

Esercizio A.5 Si calcolino i coefficienti di Fourier delle seguenti funzioni periodiche di periodo 2π i cui valori $f(x)$ per $x \in [-\pi, \pi)$ sono dati da:

$$\begin{aligned} (1) f(x) &= x^k, \quad 1 \leq k \leq 4; & (2) f(x) &= |x|; & (3) f(x) &= \sum_{k=1}^{\infty} \frac{\sin^2 k^2 x}{2^k}, \\ (4) f(x) &= e^x; & (5) f(x) &= e^x, \quad (0 < x \leq \pi) \text{ e } f \text{ dispari}; \\ (6) f(x) &= \sin^3 x; & (7) f(x) &= \cos^6 x; & (8) & \frac{\pi - x}{2}. \end{aligned}$$

⁴¹Si ricorda che se $g \in \mathcal{R}_1(A)$, A misurabile secondo Peano-Jordan, per ogni $\varepsilon > 0$ esiste una funzione $\psi \in C_0^\infty(A)$ tale che $\int_A |g - \psi| < \varepsilon$.

⁴²I seguenti integrali si intendono tutti divisi per $\sqrt{2\pi}$.

⁴³Una funzione g definita su \mathbb{R}^n si dice *pari* se $g(-x) = g(x)$, $\forall x$, e si dice *dispari* se $g(-x) = -g(x)$.

Esercizio A.6 Si calcolino i coefficienti di Fourier delle funzioni periodiche di periodo 2π i cui valori coincidono con i valori $f(x)$ elencati in (1)÷(8) dell'Esercizio A.5 per $x \in [0, 2\pi]$.

Esercizio A.7 (i) Sia $u_n(x) := \int_{-\infty}^{\infty} e^{-t^2} \cos n(t-x) dt$. Si studi la regolarità di $u := \sum_{n \geq 1} u_n$ e se ne calcoli la serie di Fourier.

(ii) Come al punto (i) con $u_n(x) := \int_0^{\infty} e^{-t^2} \cos n(t-x) dt$.

Esercizio A.8 Si dimostri che sotto le ipotesi del Lemma A.22 di Dini si ha che

$$\lim_{N, M \rightarrow \infty} \sum_{n=-M}^N \hat{f}_n e^{inx} = f(x)$$

ossia che $\forall \varepsilon > 0$ esiste N_0 tale che per ogni $N, M \geq N_0$ si ha che $\left| \sum_{n=-M}^N \hat{f}_n e^{inx} - f(x) \right| < \varepsilon$.

Esercizio A.9 Si dimostri che se f è periodica e C^1 a tratti allora soddisfa le ipotesi del Lemma A.28. Dunque nei punti dove f ha un salto (ma esistono i limiti da destra e da sinistra di f e f') la serie di Fourier vale la media del salto.

Esercizio A.10 Sia $k \geq 2$ e sia $u(x) = \sum_{n \geq 1} \frac{\cos nx}{n^k}$. Si dimostri che $u \in C_{\text{per}}^{k-2}$ ma che $u \notin C_{\text{per}}^k$.

Esercizio A.11 Sia f una funzione di periodo $T > 0$ integrabile su $[0, T)$. Si scrivano le formule rilevanti per l'analisi di Fourier.

Appendice B

Equazioni differenziali ordinarie

B.1 Introduzione

B.1.1 Equazioni scalari

Un'equazione differenziale ordinaria di ordine n per una funzione incognita $t \in I \subseteq \mathbb{R} \mapsto u(t) \in \mathbb{R}$, I intervallo, è un'equazione della forma¹

$$F(u(t), u'(t), \dots, u^{(n)}(t), t) = 0, \quad (\text{B.1})$$

dove $F : E \subseteq \mathbb{R}^{n+2} \rightarrow \mathbb{R}$ è una funzione definita su un aperto $E \subseteq \mathbb{R}^{n+2}$; $u^{(j)}(t)$ denota la derivata di ordine j rispetto alla variabile 'indipendente' t .

Una soluzione (classica, locale) di (B.1) con F continua² è una funzione $u \in C^n(I, \mathbb{R})$ che verifica (B.1) per ogni³ $t \in I$.

L'equazione (B.1) si dice in forma normale se esiste una funzione $f : A \subseteq \mathbb{R}^{n+1} \rightarrow \mathbb{R}$ tale che (B.1) abbia la forma $u^{(n)}(t) = f(u(t), u'(t), \dots, u^{(n-1)}(t), t)$ o, più sinteticamente (non indicando la variabile indipendente nell'argomento di u e delle sue derivate):

$$u^{(n)} = f(u, u', \dots, u^{(n-1)}, t). \quad (\text{B.2})$$

Esempi:

- (i) $u' + ku = 0, \quad k > 0$ costante. Circuito RC⁴
- (ii) $u' = ru \cdot \left(1 - \frac{u}{k}\right), \quad r, k > 0$ costanti. Equazione logistica di Verhulst⁵
- (iii) $u' = a(t)u + b(t)u^\alpha, \quad \alpha \in \mathbb{R} \setminus \{0\}$ costante. Equazione di Bernoulli⁶
- (iv) $u = tu' + h(u'), \quad h$ funzione differenziabile. Equazione di Clairaut⁷
- (v) $u'' + \lambda u' + \omega^2 u = 0, \quad \omega > 0, \lambda \geq 0$. Oscillatore armonico (smorzato)⁸

¹Come al solito, $u^{(1)}(t) = u'(t)$. L'aggettivo 'ordinarie' si riferisce al fatto che la funzione incognita dipende da una sola variabile; nel caso la funzione incognita dipenda da più variabili e compaiano nell'equazione le derivate parziali di u , l'equazione si chiamerà equazione differenziale alle derivate parziali; cfr. https://en.wikipedia.org/wiki/Partial_differential_equation.

²In generale, si possono considerare anche funzioni meno regolari chiarendo in maniera opportuna la definizione di 'soluzione'. In questo capitolo considereremo, di norma, funzioni lipschitziane.

³Implicitamente questo significa che $t \rightarrow (u(t), u'(t), \dots, u^{(n)}(t), t) \in A$ per ogni $t \in I$.

⁴https://it.wikipedia.org/wiki/Circuito_RC

⁵https://it.wikipedia.org/wiki/Equazione_logistica

⁶https://it.wikipedia.org/wiki/Equazione_differenziale_di_Bernoulli

⁷https://it.wikipedia.org/wiki/Equazione_di_Clairaut

⁸https://it.wikipedia.org/wiki/Moto_armonico

- (vi) $ku^{(4)} = q(t)$, $k > 0$, q funzione continua. *Equazione statica di Eulero–Bernoulli*⁹
 (vii) $t^n u^{(n)} + a_{n-1} t^{n-1} u^{(n-1)} + \dots + a_0 u = 0$, a_i costanti. *Equazione di Eulero*¹⁰
 (viii)* $u'' + \lambda u' + \sin u = \varepsilon \cos t$, $\lambda, \varepsilon \geq 0$ costanti. *Pendolo smorzato/forzato*¹¹

Il numero naturale n prende il nome di *ordine dell'equazione differenziale*: le equazioni da (i) a (iv) sono del prim'ordine; le equazioni (v) e (viii) del second'ordine; l'equazione (vi) del quart'ordine; l'equazione (vii) di ordine n .

Tutte le equazioni sono in forma normale tranne (in generale) la (iv) e la (vii) (in un intorno di $t = 0$).

Le equazioni (i), (v), (vi) e (vii) sono *equazioni differenziali lineari*, ossia, la funzione incognita u e le sue derivate appaiono linearmente: la forma più generale di un'equazione differenziale scalare, lineare di ordine n è

$$a_n(t)u^{(n)} + a_{n-1}(t)u^{(n-1)} + \dots + a_1(t)u' + a_0(t)u = q(t). \quad (\text{B.3})$$

Tale equazione si dice *omogenea* se $q \equiv 0$, e in tal caso, se u_1 e u_2 sono soluzioni, allora lo sono anche $au_1 + bu_2$ per ogni scelta delle costanti a, b (in particolare, $u \equiv 0$ è soluzione di (B.3) con $q = 0$).

Le equazioni (i), (v) e (vii) sono lineari omogenee.

Soluzione di (i) e (vi):

(i): L'equazione differenziale (i) è lineare, omogenea e di grado uno. Moltiplicando per e^{kt} i due membri dell'equazione si ottiene $e^{kt}u' + e^{kt}u = 0$, ossia

$$(e^{kt}u)' = 0.$$

Fissando \bar{t} ed integrando tra \bar{t} e t tale relazione otteniamo

$$u(t) = e^{-k(t-\bar{t})}\bar{u}, \quad \bar{u} := u(\bar{t}). \quad (\text{B.4})$$

Dunque, per ogni costante \bar{u} , la funzione $u(t) = e^{-k(t-\bar{t})}\bar{u}$ è soluzione di (i) e la costante \bar{u} coincide col valore di u in \bar{t} . Si noti che, per come abbiamo ottenuto la formula (B.4), tutte le soluzioni C^1 di (i) hanno la forma (B.4): se $u \in C^1(I)$, I intervallo, soddisfa (i), allora vale (B.4) per un qualunque $\bar{t} \in I$.

(vi): Questa equazione (del quart'ordine, lineare, a coefficienti costanti) può essere risolta per *integrazione diretta*. Assumiamo $q \in C(I, \mathbb{R})$ con I intervallo. Dividendo per k e ponendo $f(t) := q(t)/k$, si ha $u^{(4)} = f$. Fissando $\bar{t} \in I$ e integrando tale relazione, si ottiene

$$u'''(t) = c_3 + \int_{\bar{t}}^t f(s)ds, \quad c_3 := u'''(\bar{t}).$$

Integrando ancora tre volte ed usando la formula (2.83) per la primitiva di ordine 4 di $f = q/k$, otteniamo:

$$u(t) = c_0 + c_1 t + c_2 \frac{t^2}{2} + c_3 \frac{t^3}{6} + \frac{1}{k} \int_{\bar{t}}^t \frac{(t-s)^3}{6} q(s) ds, \quad c_k = u^{(k)}(\bar{t}). \quad \blacksquare \quad (\text{B.5})$$

Come vedremo più avanti, anche le altre equazioni sopra elencate – tranne la (viii) – possono essere risolte 'esplicitamente' (in termini di integrali di funzioni date).

⁹https://en.wikipedia.org/wiki/Euler-Bernoulli_beam_theory

¹⁰https://it.wikipedia.org/wiki/Equazione_di_Eulero

¹¹John H. Hubbard (1999): *The Forced Damped Pendulum: Chaos, Complication and Control*, The American Mathematical Monthly, 106:8, 741-758, DOI: 10.1080/00029890.1999.12005113. L'asterisco segnala una differenza sostanziale tra questo esempio e i precedenti che verrà chiarito in seguito.

L'equazione (viii) fisicamente rappresenta l'equazione che regola la dinamica di un pendolo che oscilla sotto effetto della forza peso in un piano verticale fissato, soggetto ad una forza di attrito proporzionale alla velocità angolare e ad una forza esterna periodica (per esempio, un campo magnetico generato da un solenoide) di intensità ε . Nel caso $\varepsilon = 0$, le traiettorie possono essere descritte 'esattamente' e tutti i moti classificati (sia nel caso $\lambda = 0$ che $\lambda > 0$), ma nel caso $\lambda, \varepsilon > 0$, questo *non è più possibile* ed, in generale, appaiono orbite 'caotiche' la cui evoluzione è impossibile prevedere per tempi lunghi¹².

B.1.2 Equazioni vettoriali (sistemi di equazioni)

Più in generale si possono considerare *sistemi di equazioni differenziali* della forma

$$F(u', u, t) = 0, \quad (\text{B.6})$$

dove $u : t \in J \mapsto u(t) \in \mathbb{R}^n$ è una funzione incognita *vettoriale* e $F : E \subseteq \mathbb{R}^{n^2+1} \rightarrow \mathbb{R}^n$ (J intervallo e E aperto), o in *forma normale*

$$u' = f(u, t), \quad (\text{B.7})$$

con u come sopra e $f : A \subseteq \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$ (A aperto).

Esempi:

- (ix) Equazioni scalari di ordine superiore al primo possono essere scritte in forma di sistema del prim'ordine: per esempio, l'equazione (v), può essere scritta nella forma (B.7) ponendo

$$v = (v_1, v_2) := (u, u')$$

e

$$v' = f(v) := Av, \quad \text{con} \quad A = \begin{pmatrix} 0 & 1 \\ -\omega^2 & -\lambda \end{pmatrix}.$$

Analogamente, l'equazione scalare, lineare di ordine n (B.3) può essere scritta come sistema di ordine n come segue¹³:

$$v := (u, u', \dots, u^{(n-1)}), \quad \text{e} \quad v' = f(v, t) := A(t)v + g(t)$$

dove

$$A(t) := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & \ddots & \ddots & \\ 0 & 0 & \cdots & 0 & 1 \\ -a_0(t) & -a_1(t) & \cdots & -a_{n-1}(t) & -a_n(t) \end{pmatrix}, \quad g(t) := \begin{pmatrix} 0 \\ \vdots \\ 0 \\ q(t) \end{pmatrix}. \quad (\text{B.8})$$

Più in generale, il sistema di equazioni differenziali

$$u' = A(t)u, \quad (\text{B.9})$$

dove $t \in I \subseteq \mathbb{R} \mapsto A(t) \in \mathbb{R}^{n \times n}$ è una matrice che dipende in modo continuo da t , è un sistema di equazioni differenziali del prim'ordine di dimensione n per la funzione incognita $t \in \mathbb{R} \mapsto u(t) \in \mathbb{R}^n$. Nel caso $A(t) \equiv A$ non dipenda da t parleremo di un sistema di equazioni differenziali di dimensione n a coefficienti costanti.

¹²Si veda il già citato articolo di J.H. Hubbard nella nota 11.

¹³Nell'equazione differenziale il vettore v' va interpretato come *vettore colonna*.

- (x) Naturalmente, anche le equazioni scalari non lineari di ordine n possono risciversi sotto forma di sistema n -dimensionale del prim'ordine:

$$u^{(n)} = g(u, u', \dots, u^{(n-1)}, t), \quad \iff \quad v' = f(v, t) \quad (\text{B.10})$$

dove si è posto

$$v := (u, u', \dots, u^{(n-1)}), \quad f_i(v, t) := \begin{cases} v_{i+1}, & \text{se } 1 \leq i \leq n-1 \\ g(v, t), & \text{se } i = n. \end{cases} \quad (\text{B.11})$$

- (xi) Secondo Newton, il moto $t \in I \mapsto x(t) \in \mathbb{R}^3$ di un punto materiale di massa costante $m > 0$ è descritto dall'equazione $f = m\ddot{x}$, dove t , la variabile indipendente, è il tempo; il punto (anziché l'apice) denota la derivata rispetto a¹⁴ t e $f = (f_1, f_2, f_3)$ la forza esercitata sul punto. Se la forza dipende dalla posizione, dalla velocità del punto e dal tempo, l'equazione di Newton è descritta da un sistema di 3 equazioni differenziali del second'ordine:

$$m\ddot{x} = f(x, \dot{x}, t), \quad \iff \quad m\ddot{x}_i = f_i(x, \dot{x}, t), \quad (i = 1, 2, 3). \quad (\text{B.12})$$

- (xii) Un'altra classe di esempi importante sono i *sistemi hamiltoniani*, che descrivono l'evoluzione di 'sistemi dinamici conservativi'. Data una funzione $H \in C^1$ di $2n+1$ variabili $(p, q, t) = (p_1, \dots, p_n, q_1, \dots, q_n, t)$, il seguente sistema del prim'ordine di $2n$ equazioni descrive l'evoluzione di un sistema hamiltoniano con ' n gradi di libertà', con hamiltoniana H , dove, come d'uso comune, la variabile indipendente t denota il 'tempo':

$$\begin{cases} \dot{p} = -H_q(p, q, t) \\ \dot{q} = H_p(p, q, t) \end{cases} \iff \begin{cases} \dot{p}_i = -H_{q_i}(p, q, t), & 1 \leq i \leq n, \\ \dot{q}_i = H_{p_i}(p, q, t), & 1 \leq i \leq n. \end{cases} \quad (\text{B.13})$$

Per esempio, il 'pendolo matematico', la cui evoluzione è descritta dall'equazione (viii) nel caso $\lambda = \varepsilon = 0$ è un sistema hamiltoniano ad un grado di libertà e con hamiltoniana

$$H(p, q) = \frac{p^2}{2} - \cos q, \quad (\text{B.14})$$

avendo posto, $q = q(t) := u(t)$ e, quindi, $u' := \dot{q} = H_p = p$.

- (xiii) Un ultimo esempio della forma $\dot{u} = f(u)$ con $u(t) := (x(t), y(t), z(t)) \in \mathbb{R}^3$ e f funzione (vettoriale) polinomiale di grado 2, e, quindi, dalla forma matematica assai semplice, ma con una dinamica estremamente ricca e non ancora completamente studiata, è il 'sistema di Lorenz¹⁵':

$$\begin{cases} \dot{x} = \sigma(y - x) \\ \dot{y} = x(\rho - z) - y \\ \dot{z} = xy - \beta z \end{cases} \quad (\text{B.15})$$

dove σ, ρ, β sono costanti fisiche positive; tale sistema è uno dei prototipi alla base della 'teoria del chaos¹⁶'.

B.1.3 Tre esempi risolubili

Alla fine del paragrafo § B.1.1, abbiamo visto come gli esempi (i) e (v) fossero 'esplicitamente' risolubili. Vediamo qui altri tre casi semplici ma importanti di equazioni (o sistemi di equazioni) differenziali risolubili¹⁷.

¹⁴Quindi \dot{x} denota la velocità istantanea del punto x e \ddot{x} (la derivata seconda rispetto a t) la sua accelerazione. Questa notazione è stata introdotta da Newton stesso; cfr. https://en.wikipedia.org/wiki/Notation_for_differentiation

¹⁵https://en.wikipedia.org/wiki/Lorenz_system

¹⁶https://en.wikipedia.org/wiki/Chaos_theory

¹⁷Anche gli esempi (ii), (iii), (iv), (v) e (vii) sono risolubili: cfr. Complemento B.1.

Equazioni scalari lineari del prim'ordine

Consideriamo un'equazione scalare del prim'ordine lineare non omogenea e in forma normale per una funzione $t \mapsto u(t)$, ossia¹⁸

$$u' = a(t)u + b(t), \quad (\text{B.16})$$

dove I è un dato intervallo di \mathbb{R} e $a, b \in C(I)$. Se fissiamo $\bar{t} \in I$, poniamo

$$\alpha(t) := \int_{\bar{t}}^t a(s) ds \quad (\text{B.17})$$

e moltiplichiamo per $e^{-\alpha(t)}$ i termini dell'equazione in (B.16), otteniamo

$$u'e^{-\alpha} - ae^{-\alpha}u = e^{-\alpha}b,$$

ossia¹⁹

$$(ue^{-\alpha})' = e^{-\alpha}b. \quad (\text{B.18})$$

A questo punto, integrando la relazione (B.18) tra \bar{t} e t , osservando che $\alpha(\bar{t}) = 0$, segue dal Teorema fondamentale del calcolo che

$$u(t) = e^{\alpha(t)}\bar{u} + e^{\alpha(t)} \int_{\bar{t}}^t e^{-\alpha(s)}b(s) ds, \quad \text{dove } \bar{u} := u(\bar{t}). \quad (\text{B.19})$$

Queste, sono tutte (e sole²⁰) le soluzioni differenziabili di (B.16).

Si noti che le soluzioni di (B.16) sono definite su tutto l'intervallo I in cui le funzioni a e b sono continue.

Equazioni scalari a variabili separabili

Siano I e J intervalli e $f \in C(J, \mathbb{R})$, $g \in C(I, \mathbb{R})$. Vogliamo determinare le soluzioni C^1 di²¹

$$u' = f(u)g(t). \quad (\text{B.20})$$

Fissiamo $\bar{t} \in I$ e sia $\bar{u} := u(\bar{t})$ il valore della soluzione u (ammesso che esista) in \bar{t} . Se $f(\bar{u}) = 0$, chiaramente, $u(t) \equiv \bar{u}$ è soluzione di (B.20). Supponiamo ora che $f(\bar{u}) \neq 0$. Per il teorema di permanenza del segno, esiste un intervallo aperto $J_0 \subseteq J$ contenente \bar{u} tale che $f(x) \cdot f(\bar{u}) > 0$ per ogni $x \in J_0$. Inoltre se u è una soluzione C^1 di (B.20), per t vicino a \bar{t} , $u(t) \in J_0$ e possiamo, dunque, dividere per $f(u)$ ottenendo

$$\frac{u'}{f(u)} = g(t). \quad (\text{B.21})$$

Integrando tra \bar{t} e t , otteniamo

$$\int_{\bar{t}}^t \frac{u'}{f(u)} dt = G(t), \quad G(t) := \int_{\bar{t}}^t g(s) ds \quad (\text{B.22})$$

Facendo il cambio di variabile $x = u(t)$ nel primo integrale in (B.22), essendo $u(\bar{t}) = \bar{u}$, otteniamo

$$\int_{\bar{u}}^{u(t)} \frac{dx}{f(x)} dx = G(t). \quad (\text{B.23})$$

¹⁸L'esempio (i) è di questa forma con $a(t) \equiv -k$ e $b \equiv 0$.

¹⁹Per il Teorema fondamentale del calcolo si ha $(e^{-\alpha})' = e^{-\alpha}(-\alpha)' = -ae^{-\alpha}$.

²⁰L'unicità deriva dal fatto che abbiamo dedotto la (B.19) direttamente dalla (B.16) (assumendo, naturalmente, che u fosse derivabile).

²¹Naturalmente, l'espressione 'variabili separabili' deriva dal fatto che il membro di destra di (B.21) è il prodotto di una funzione della sola 'variabile indipendente' t e della sola 'variabile dipendente' u .

Sia, ora, $F(x) := F(x; \bar{u})$ la funzione integrale di $1/f$ con punto base \bar{u} , ossia

$$F(x) := \int_{\bar{u}}^x \frac{dy}{f(y)}, \quad (\text{B.24})$$

e si osservi che (per il Teorema fondamentale del calcolo) $F'(x) = 1/f(x) \neq 0$, per $x \in J_0$, e dunque $x \in J_0 \mapsto F(x)$ è una funzione invertibile con inversa C^1 . Se denotiamo con F^{-1} la funzione inversa di F , da (B.23), troviamo $F(u(t)) = G(t)$ e quindi:

$$u(t) = (F^{-1} \circ G)(t). \quad (\text{B.25})$$

Si noti che la funzione $F^{-1} \circ G$ è di classe $C^1(I_0, J_0)$ dove I_0 è il più grande intervallo aperto contenente \bar{t} tale che²² $G(I_0) \subseteq F(J_0)$. Abbiamo dimostrato che l'unica soluzione derivabile $t \in I_0 \mapsto u(t) \in J_0$ tale che $u(\bar{t}) = \bar{u}$ è data da (B.25) e (B.24).

Sistemi lineari a coefficienti costanti

L'esponenziale di matrice²³ è lo strumento giusto per risolvere sistemi di equazioni differenziali lineari a coefficienti costanti: infatti, data una matrice $A \in \mathbb{R}^{n \times n}$ e un punto $\bar{x} \in \mathbb{R}^n$, da (1.89) segue che, $t \mapsto u(t) := e^{At}\bar{x}$ è differenziabile e verifica

$$u' = Au, \quad u(0) = \bar{x}. \quad (\text{B.26})$$

In effetti, si ha

Proposizione B.1 $t \in \mathbb{R} \mapsto \bar{u}(t) := e^{At}\bar{x} \in \mathbb{R}^n$ è l'unica soluzione di (B.26).

Dimostrazione Sia $v \in C^1(E, \mathbb{R}^n)$ una soluzione di (B.26) con E un intervallo aperto di \mathbb{R} contenente 0. Mostriamo che si deve avere $v(t) = \bar{u}(t)$ per ogni $t \in E$. Supponiamo, infatti, per assurdo, che \bar{u} e v non coincidano su E e sia $w := \bar{u} - v$. Sia $I \subseteq E$ il più grande intervallo contenente 0 per cui $w(t) = 0$ per ogni $t \in I$. Secondo la nostra ipotesi I non coincide con E e quindi almeno uno dei suoi estremi è un punto $\tau \in E$. Poichè $w(t)$ è continua, $w(\tau) = 0$. La funzione w soddisfa $w' = Aw$ e $w(\tau) = 0$. Supponendo che τ sia l'estremo destro di I (il ragionamento nell'altro caso è perfettamente analogo), integrando la relazione $w' = Aw$ tra τ e $t > \tau$ con $t \leq \tau + \varepsilon$ ed ε sufficientemente piccolo, si ha²⁴

$$w(t) = w(\tau) + \int_{\tau}^t Aw(s) ds = A \int_{\tau}^t w(s) ds,$$

e, passando alle norme²⁵:

$$|w(t)| = \left| A \int_{\tau}^t w(s) ds \right| \leq \|A\| \left(\sup_{[\tau, \tau + \varepsilon]} |w| \right) \varepsilon \quad (\text{B.27})$$

e, quindi, (prendendo l'estremo superiore su $t \in [\tau, \tau + \varepsilon]$),

$$\sup_{[\tau, \tau + \varepsilon]} |w| \leq \|A\| \cdot \varepsilon \cdot \sup_{[\tau, \tau + \varepsilon]} |w|$$

il che, se scegliamo ε tale che $\|A\|\varepsilon < 1$, è assurdo a meno che $\sup_{[\tau, \tau + \varepsilon]} |w| = 0$; ma in tal caso I non sarebbe il più grande intervallo su cui $w = 0$, e si avrebbe una contraddizione. Dunque I ed E debbono coincidere, ossia, $\bar{u} = v$ su E . ■

La funzione vettoriale $v(t) = e^{A(t-\bar{t})}\bar{x}$ è soluzione di

$$v' = Av, \quad v(\bar{t}) = \bar{x}. \quad (\text{B.28})$$

D'altra parte se v verifica (B.28), allora $u(t) := v(\bar{t} + t)$ verifica (B.26) e quindi, per la Proposizione B.1, $u(t) = e^{At}\bar{x}$ e quindi $v(t) = e^{A(t-\bar{t})}\bar{x}$. In altri termini:

$v(t) = e^{A(t-\bar{t})}\bar{x}$ è l'unica soluzione di (B.28).

²²Si noti che $G(\bar{t}) = 0 = F(\bar{u})$.

²³Cfr. Complemento 1.5.

²⁴Per l'integrazione di funzioni a valori vettoriali si veda il Complemento 1.3 ed in particolare (1.67) e (1.66).

²⁵ $\|\cdot\|$ denota una qualunque norma matriciale rispetto ad una qualunque norma $|\cdot|$ in \mathbb{R}^n .

B.1.4 Osservazioni

Concludiamo questa breve introduzione alle equazioni differenziali ordinarie con alcune osservazioni.

- (i) Nel caso $f(\bar{x}) = 0$ abbiamo osservato che $u \equiv \bar{x}$ è soluzione di (B.20). Ma, in generale, *tale soluzione potrebbe non essere unica*. Per esempio, se $f(x) := x^{2/3}$, $g(t) \equiv 1$, (B.20) diventa

$$u' = u^{2/3} \quad (\text{B.29})$$

e se poniamo $\bar{x} = 0 = \bar{t}$, vediamo che (B.29) con la condizione $u(0) = 0$ ammette le soluzioni $u_1 \equiv 0$ e $u_2 := t^3/27$. Infatti vi sono infinite soluzioni $C^1(\mathbb{R})$ di (B.29) con $u(0) = 0$ tra le quali, per esempio,

$$u_s^+(t) := \begin{cases} 0, & \text{se } t \leq s, \\ (t-s)^3/27, & \text{se } t > s, \end{cases} \quad (\text{B.30})$$

dove s è un parametro non negativo.

- (ii) Un altro punto importante è che la (unica) soluzione (B.25) di (B.20) con $u(\bar{t}) = \bar{x}$ nel caso $f(\bar{x}) \neq 0$ è, in generale, definita solo in un intorno di \bar{t} . In effetti può succedere che l'intervallo massimale di esistenza della soluzione u sia limitato anche se le funzioni f e g sono $C^\infty(\mathbb{R})$, come mostra il seguente esempio: sia $f(x) = x^2$, $g(t) \equiv 1$ e cerchiamo di determinare la soluzione di

$$u' = u^2, \quad u(\bar{t}) = \bar{x}, \quad (\text{B.31})$$

con $\bar{x} \neq 0$. Calcoliamo la funzione integrale $x \mapsto s = F(x)$:

$$F(x) = \int_{\bar{x}}^x \frac{dy}{y^2} = \left[-\frac{1}{y} \right]_{\bar{x}}^x = \frac{1}{\bar{x}} - \frac{1}{x} \quad (\text{B.32})$$

la cui funzione inversa è data da

$$F^{-1}(s) = \frac{1}{\frac{1}{\bar{x}} - s}. \quad (\text{B.33})$$

Essendo, nel presente caso, $G(t) := \int_{\bar{t}}^t ds = t - \bar{t}$, otteniamo

$$u(t) = F^{-1} \circ G(t) = \frac{1}{T_e - t}, \quad T_e := \bar{t} + \frac{1}{\bar{x}}. \quad (\text{B.34})$$

L'intervallo I su cui è definita la soluzione $u(t)$ dipende dal segno di \bar{x} ed è dato da

$$I := I(\bar{t}, \bar{x}) := \begin{cases} (-\infty, T_e), & \text{se } \bar{x} > 0, \\ (T_e, +\infty), & \text{se } \bar{x} < 0. \end{cases} \quad (\text{B.35})$$

Nella singolarità T_e , $u(t)$ ha un asintoto verticale, in altri termini, la soluzione 'esplode in tempi finiti'.

- (iii) Se denotiamo con $u(t; \bar{t}, \bar{x})$ la soluzione $u(t)$ (B.34) dell'equazione (B.31) con $\bar{x} \neq 0$, si vede immediatamente che

$$\lim_{\bar{x} \rightarrow 0} u(t; \bar{t}, \bar{x}) = 0.$$

In effetti, come osservato più volte, $t \in \mathbb{R} \mapsto u(t) \equiv 0$ è soluzione di (B.31) con $\bar{x} = 0$, ma la soluzione nulla non si ottiene col metodo di separazione di variabili (dove abbiamo usato l'ipotesi che $\bar{x} \neq 0$). Rimane il dubbio che ci possa essere un'altra soluzione del problema (B.31) con $\bar{x} = 0$ che non sia identicamente nulla, come accade nell'esempio

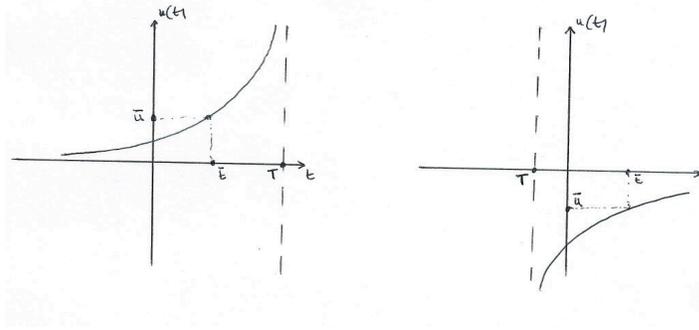


Figura B.1: La soluzione $u(t)$ di (B.31)

del punto (i). Ma questo non è il caso: $u(t) \equiv 0$ è l'unica soluzione di (B.31) con $\bar{x} = 0$. Infatti, sia $t \in I \rightarrow u(t)$ una qualunque soluzione di (B.31) con $\bar{x} = 0$ e $\bar{t} \in I$. Allora, se $v(t) := \int_{\bar{t}}^t u^2(s) ds$, si ha

$$0 = e^{-v}(u' - u^2) = (e^{-v}u)',$$

il che implica che $e^{-v}u$ è costante su I ed essendo $e^{-v(\bar{t})}u(\bar{t}) = 0$ si ha che $u(t)$ è identicamente nulla su I . Quindi $u \equiv 0$ è l'unica soluzione di (B.31) con $\bar{x} = 0$.

- (iv) Negli esempi sinora discussi – a parte il caso dell'esempio (i) di questo paragrafo – abbiamo visto che il problema di determinare la soluzione $u(t)$ di un sistema di equazioni differenziale

$$u'(t) = f(u(t), t) \quad \Longleftrightarrow \quad u'_i(t) = f_i(u(t), t), \quad \forall 1 \leq i \leq n, \quad (\text{B.36})$$

con 'dato iniziale' $u(\bar{t}) = \bar{x}$ assegnato, ammette una ed una sola soluzione, e che tale soluzione, in generale, è definita in un opportuno intorno di \bar{t} . Tale problema prende il nome di *problema di Cauchy* o *problema ai valori iniziali*²⁶.

Un risultato fondamentale che dimostreremo nella prossima sezione è che *il problema di Cauchy ammette una ed una sola soluzione locale in*²⁷ t , assumendo che la funzione f sia *localmente lipschitziana* (in x uniformemente in t).

Si noti che la funzione $u \in \mathbb{R} \mapsto f(x) = x^{2/3}$ (cfr. esempio (i) di questo paragrafo) non è lipschitziana in un intorno di 0.

Vi sono naturalmente molti altri risultati riguardo esistenza e/o unicità di equazioni differenziali²⁸; per esempio, vale il teorema di Peano, secondo il quale, se f in (B.39) è continua, allora esiste sempre almeno una soluzione locale (che, come sappiamo, può non essere unica).

- (v) Dal punto di vista fisico, trovare una soluzione di (B.39) significa trovare una curva (o una 'orbita') $t \in I \mapsto u(t) \in \mathbb{R}^n$, con I intervallo che contiene \bar{t} , che passi per \bar{x} nell'istante \bar{t} e tale che la sua velocità $u'(t)$ (che geometricamente rappresenta un vettore tangente alla curva descritta da u al variare di t) coincida col 'campo vettoriale' $f(u(t), t)$.

²⁶Ancora una volta, si pensa all'interpretazione di t come 'tempo' ed all'equazione differenziale (B.36) come ad una 'equazione di evoluzione' per la funzione incognita u .

²⁷Ossia, definita in un opportuno intorno I di \bar{t} .

²⁸Cfr. [E. Coddington, N. Levinson, "The theory of ordinary differential equations", McGraw-Hill (1955)], uno dei testi di riferimento fondamentali sulla teoria delle equazioni differenziali ordinarie.

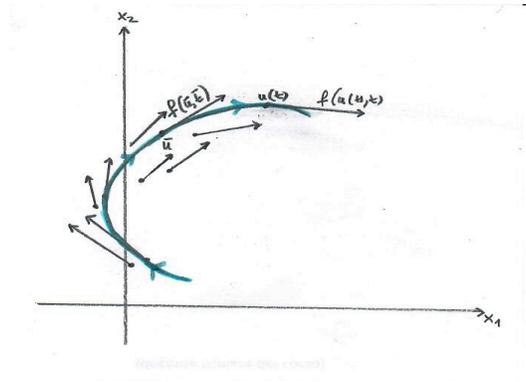


Figura B.2: Campo vettoriale $f(x, t)$ e orbita $t \rightarrow u(t)$

(vi) Come visto, in alcuni (rari) casi, è possibile trovare una ‘formula risolutiva’ di una equazione (o di un sistema di equazioni) differenziale (cfr. § B.1.3). Ma, anche in tali casi, non è sempre ovvio ricavare informazioni *qualitative* sulla soluzione, specialmente per tempi lunghi. Si pensi, per esempio, al caso (assai semplice) delle equazioni scalari a variabili separabili (B.25): la soluzione generale è data in forma implicita ed in termini di integrali, che, di norma, non sono calcolabili esplicitamente; in particolare, può non essere immediato stabilire, per esempio, il comportamento asintotico (per $t \rightarrow \pm\infty$) della soluzione.

D’altra parte, abbiamo già sottolineato come ‘semplici’ equazioni differenziali possano avere soluzioni estremamente complicate, con evoluzione non prevedibile²⁹.

In effetti, il punto di vista moderno sulle equazioni differenziali ordinarie, introdotto, tra gli altri dal grande matematico e fisico teorico Henri Poincaré³⁰ a partire dalla seconda metà del 1800, è proprio quello di determinare il ‘comportamento’ per tempi lunghi delle soluzioni, dando particolare rilevanza a determinate soluzioni (periodiche, quasi-periodiche, omocline, eterocline) ed allo loro stabilità. Questo nucleo di idee è alla base della moderna teoria dei sistemi dinamici³¹.

B.2 Teorema di esistenza e unicità locale

In questo paragrafo discutiamo il problema di Cauchy e il fondamentale teorema di Picard–Lindelöf³² che, sotto opportune ipotesi, garantisce *esistenza, unicità e dipendenza continua dai dati iniziali* di soluzioni locali.

La dimostrazione sarà basata su alcuni lemmi, di notevole interesse per sé stessi.

B.2.1 Definizioni ed enunciato del Teorema

Cominciamo con la definizione precisa di soluzione (locale) di un problema di Cauchy e con un’osservazione preliminare:

Osservazione B.2 (i) Sia $A \subseteq \mathbb{R}^{n+1}$ un insieme aperto, $f \in C(A, \mathbb{R}^n)$ e $(\bar{x}, \bar{t}) \in A$. Ad una funzione $v \in C(I, \mathbb{R}^n)$, I intervallo contenente \bar{t} , tale che $(v(t), t) \in A$ per ogni $t \in I$, possiamo

²⁹Cfr. Esempio (viii) di § B.1.1 e nota 11, o Esempio (xiii) di § B.1.1.

³⁰https://it.wikipedia.org/wiki/Henri_Poincar%C3%A9

³¹<http://www.scholarpedia.org/article/Dynamics>

³²http://en.wikipedia.org/wiki/Picard%E2%80%93Lindel%C3%B6f_theorem e in particolare l’articolo [Lindelöf, E. (1894). “Sur l’application de la méthode des approximations successives aux équations différentielles ordinaires du premier ordre”. Comptes rendus hebdomadaires des séances de l’Académie des sciences. 116: 454–457] in cui Lindelöf discute una generalizzazione del metodo iterativo introdotto da Picard.

associare una nuova funzione $\Phi(v) := \Phi(v; \bar{x}, \bar{t})$ definita come segue:

$$t \in I \mapsto \Phi(v)(t) := \Phi(v; \bar{x}, \bar{t})(t) := \bar{x} + \int_{\bar{t}}^t f(v(s), s) ds \in \mathbb{R}^n. \quad (\text{B.37})$$

Per il Teorema fondamentale del calcolo, $\Phi(v)$ è una funzione $C^1(I, \mathbb{R}^n)$ e³³

$$(\Phi(v; \bar{x}, \bar{t}))'(t) = f(v(t), t), \quad \Phi(v; \bar{x}, \bar{t})(\bar{t}) = \bar{x}. \quad (\text{B.38})$$

L'operatore $v \mapsto \Phi(v)$ prende il nome di **operatore di Picard**.

(ii) Sia $A \subseteq \mathbb{R}^{n+1}$ un insieme aperto, $f \in C(A, \mathbb{R}^n)$ e $(\bar{x}, \bar{t}) \in A$. Una funzione $u \in C^1(I, \mathbb{R}^n)$, I intervallo contenente \bar{t} , tale che $(v(t), t) \in A$ per ogni $t \in I$, soddisfa il problema di Cauchy

$$\begin{cases} u'(t) = f(u(t), t), \\ u(\bar{t}) = \bar{x}, \end{cases} \quad (\text{B.39})$$

se e solo se u è un **punto fisso** dell'operatore di Picard:

$$u = \Phi(u; \bar{x}, \bar{t}) \quad \stackrel{\text{def}}{\iff} \quad u(t) = \bar{x} + \int_{\bar{t}}^t f(u(s), s) ds. \quad (\text{B.40})$$

Infatti (B.40) si ottiene immediatamente da (B.39) integrando ambo i membri dell'equazione differenziale in (B.39) tra \bar{t} e t ed usando il Teorema fondamentale del calcolo, mentre se vale (B.40), da (B.38) segue che u è soluzione di (B.39).

Definizione B.3 Sia $A \subseteq \mathbb{R}^{n+1}$ un insieme aperto. Denotiamo con $\text{Lip}_*(A, \mathbb{R}^n)$ il sottoinsieme di $C(A, \mathbb{R}^n)$ formato dalle funzioni **localmente lipschitziane rispetto a x uniformemente in t** , ossia, tali che:

$$\forall K \subseteq A \text{ compatto } \exists L > 0 \mid |f(x, t) - f(y, t)| \leq L|x - y|, \quad \forall (x, t), (y, t) \in K. \quad (\text{B.41})$$

Osservazione B.4 (i) La (B.41) è equivalente a richiedere che per ogni $(\bar{x}, \bar{t}) \in A$ esistono

$$D := D(\bar{x}, \rho) := \{x \in \mathbb{R}^n \mid |x - \bar{x}| \leq \rho\}, \quad J := J(\bar{t}, \tau) := [\bar{t} - \tau, \bar{t} + \tau], \quad (\text{B.42})$$

con $\rho, \tau > 0$ tali che $D \times J \subseteq A$ e

$$\exists L > 0 \mid |f(x, t) - f(y, t)| \leq L|x - y|, \quad \forall x, y \in D, \forall t \in J. \quad (\text{B.43})$$

Dimostrazione Se vale (B.41), vale (B.43) essendo $D \times J$ in (B.42) compatto. Dimostriamo il viceversa e assumiamo che per ogni $(\bar{x}, \bar{t}) \in A$, esistono $\rho, \tau > 0$, D e J come in (B.42) per cui vale (B.43) con $L(\rho, \tau) := L$. Sia $K \subseteq A$ compatto. Chiaramente $K \subseteq \bigcup_{(\bar{x}, \bar{t}) \in K} B(\rho, \tau)$ con $B(\rho, \tau) := \{x \in \mathbb{R}^n \mid |x - \bar{x}| < \rho\} \times (\bar{t} - \tau, \bar{t} + \tau)$ e per la proprietà di Heine–Borel (cfr. § 5, Cap. 1), esistono N insiemi aperti $B_i := B(\rho_i, \tau_i)$ tali che $K \subseteq \bigcup_{1 \leq i \leq N} B_i$. Se $L = \min\{L_i\}$ con $L_i = L(\rho_i, \tau_i)$, allora vale (B.41). ■

(ii) Se $f \in C^1(A)$ con $A \subseteq \mathbb{R}^{n+1}$ aperto, allora dal teorema del valor medio (Proposizione 2.18–(i)) segue che f è localmente lipschitziana rispetto a x uniformemente in t nell'intorno di ogni punto $(\bar{x}, \bar{t}) \in A$: infatti, se $\rho, \tau > 0$ sono tali che $D \times J := D(\bar{x}, \rho) \times J(\bar{t}, \tau) \subseteq A$, si può prendere $L := \max_{D \times J} |\nabla_x f|$; cfr. (2.39).

Il seguente teorema è il risultato più importante del capitolo.

Teorema B.5 (Picard–Lindelöf) Sia A un aperto di \mathbb{R}^{n+1} e $f \in \text{Lip}_*(A, \mathbb{R}^n)$. Sia $(\bar{x}, \bar{t}) \in A$ e siano D e J come nella Definizione B.3. Definiamo³⁴:

$$M := \max_{D \times J} |f|, \quad T := \min\{\tau, \rho/M\}, \quad I := [\bar{t} - T, \bar{t} + T]. \quad (\text{B.44})$$

Allora, esiste una ed una sola soluzione $u \in C^1(I, D)$ del problema di Cauchy (B.39).

La dimostrazione sarà data alla fine del § B.2.4.

³³Si applichi il Teorema fondamentale del calcolo componente per componente: si noti che, essendo f continua, l'integrando $t \mapsto f_i(u(t), t)$ è una funzione continua dall'intervallo I in \mathbb{R} .

³⁴Se $M = 0$, si ha che $f \equiv 0$ e quindi il relativo problema di Cauchy è banale: l'unica soluzione di (B.39) è data da $u(t) \equiv 0$ (e, in particolare $T = +\infty$).

B.2.2 Lemma di Picard

Il primo dei lemmi su cui si basa la dimostrazione del Teorema B.5 fornisce un *algoritmo – basato sull'iterazione dell'operatore di Picard – per costruire una soluzione di (B.39)*:

Lemma B.6 (Picard) *Siano A , f , D e I come nel Teorema B.5. Si definiscano iterativamente, per $k \geq 0$, le funzioni*

$$u_k := \begin{cases} \bar{x}, & \text{se } k = 0, \\ \Phi(u_{k-1}; \bar{x}, \bar{t}) := \bar{x} + \int_{\bar{t}}^t f(u_{k-1}(s), s) ds, & \text{se } k \geq 1. \end{cases} \quad (\text{B.45})$$

Allora, per ogni $k \geq 1$, $u_k \in C^1(I, D)$ e

$$|u_k(t) - u_{k-1}(t)| \leq \frac{M}{L} \frac{(L|t - \bar{t}|)^k}{k!}, \quad \forall |t - \bar{t}| \leq T. \quad (\text{B.46})$$

La successione $\{u_k\}$ converge uniformemente ad una soluzione $C^1(I, D)$ di (B.39).

Dimostrazione Dimostriamo, innanzitutto, che la successione u_k è ben definita, ossia che $u_k \in C(I, D)$ per ogni k (cosa necessaria per poter definire $\Phi(u_k; \bar{x}, \bar{t})$). Per $k = 0$, $u_0 \equiv \bar{x} \in C(I, D)$ ovviamente. Sia, ora, $k \geq 1$ e assumiamo che $u_{k-1} \in C(I, D)$, allora $\Phi(u_{k-1}; \bar{x}, \bar{t})$ è ben definita e quindi

$$|u_k(t) - \bar{x}| = \left| \int_{\bar{t}}^t f(u_{k-1}(s), s) ds \right| \leq \left| \int_{\bar{t}}^t |f(u_{k-1}(s), s)| ds \right| \stackrel{(\text{B.44})}{\leq} TM \stackrel{(\text{B.44})}{\leq} \rho,$$

il che mostra che $u_k(t) \in D$ per ogni $t \in I$. Dal Teorema fondamentale del calcolo segue poi che $u_k \in C^1(I, D) \subseteq C(I, D)$.

Dimostriamo, ora, (B.46). Sia, per $k \geq 1$, $w_k := u_k - u_{k-1}$ e procediamo per induzione su $k \geq 1$. Per $k = 1$ e per ogni $t \in I$, si ha

$$|w_1(t)| = |\Phi(u_0; \bar{x}, \bar{t})(t) - u_0| = \left| \int_{\bar{t}}^t f(\bar{x}, s) ds \right| \leq \left| \int_{\bar{t}}^t |f(\bar{x}, s)| ds \right| \stackrel{(\text{B.44})}{\leq} M|t - \bar{t}|,$$

e quindi (B.46) vale per $k = 1$. Assumiamo ora vera la (B.46) per un $k \geq 1$ e dimostriamola per $k + 1$. Per ogni $t \in I$ si ha:

$$\begin{aligned} |w_{k+1}(t)| &= |u_{k+1}(t) - u_k(t)| = \left| \int_{\bar{t}}^t (f(u_k(s), s) - f(u_{k-1}(s), s)) ds \right| \\ &\leq \left| \int_{\bar{t}}^t |f(u_k(s), s) - f(u_{k-1}(s), s)| ds \right| \\ &\stackrel{(\text{B.43})}{\leq} L \left| \int_{\bar{t}}^t |u_k(s) - u_{k-1}(s)| ds \right| \\ &= L \left| \int_{\bar{t}}^t |w_k(s)| ds \right| \stackrel{(\text{B.46})}{\leq} ML^k \left| \int_{\bar{t}}^t \frac{|s - \bar{t}|^k}{k!} ds \right| \\ &= \frac{M}{L} \frac{(L|t - \bar{t}|)^{k+1}}{(k+1)!}, \end{aligned}$$

dove l'ultima uguaglianza segue subito distinguendo i due casi $\bar{t} \leq t \leq \bar{t} + T$ e $\bar{t} - T \leq t \leq \bar{t}$.

Ora, $u_k = \bar{x} + \sum_{j=1}^k w_j$ e da (B.46) segue che la serie $\sum w_j$ converge totalmente su I e quindi

converge uniformemente su I ad una funzione $u \in C(I, D)$ (Proposizione 1.56). Si noti che $f(u_k, t) \rightarrow f(u, t)$ uniformemente su I (essendo $|f(u_k, t) - f(u, t)| \leq L|u_k - u|$) e che per (B.38), $u'_k = f(u_{k-1}(t), t) \rightarrow f(u(t), t)$ uniformemente su I . Dunque, dalla Proposizione 1.61 segue che $u'_k \rightarrow u'$ uniformemente su I e che $u' = f(u, t)$. Infine, $u_k(\bar{t}) = \bar{x}$ per ogni k e quindi $u(\bar{t}) = \bar{x}$ cosicché u è soluzione di (B.39). ■

B.2.3 Lemmi di Gronwall

Il prossimo lemma, noto come ‘Lemma di Gronwall³⁵’, è un semplice risultato per funzioni di una variabile reale, ma è di grande utilità nello studio generale delle equazioni differenziali. Qui, da esso dedurremo l’unicità nel problema di Cauchy (B.39) e la dipendenza continua delle soluzioni dal dato iniziale.

Lemma B.7 (Gronwall) *Sia I un intervallo e $\alpha, \delta \geq 0$. Sia $g \in C(I, [0, +\infty))$ tale che, per un $\bar{t} \in I$, si abbia*

$$g(t) \leq \delta + \alpha \left| \int_{\bar{t}}^t g(s) ds \right|, \quad \forall t \in I. \quad (\text{B.47})$$

Allora, per ogni $t \in I$, si ha

$$g(t) \leq \delta e^{\alpha|t-\bar{t}|}. \quad (\text{B.48})$$

Dimostrazione Per $\alpha = 0$ il lemma è una tautologia. Assumiamo $\alpha > 0$. Supponiamo prima che $t \geq \bar{t}$. Sia $G(t) := \int_{\bar{t}}^t g(s) ds$; allora $G \in C^1(I)$ e $G(\bar{t}) = 0$ e, per il Teorema fondamentale del calcolo e (B.47), si ha

$$G' = g \leq \delta + \alpha G. \quad (\text{B.49})$$

Moltiplicando tale relazione per $e^{-\alpha t}$, vediamo che essa è equivalente alla disequaglianza

$$(Ge^{-\alpha t})' \leq \delta e^{-\alpha t}. \quad (\text{B.50})$$

Integrando (B.50) tra \bar{t} e t e ricordando che $G(\bar{t}) = 0$ si ottiene

$$G(t)e^{-\alpha t} \leq \delta \frac{e^{-\alpha \bar{t}} - e^{-\alpha t}}{\alpha}, \quad (\text{B.51})$$

ossia $G \leq \delta \alpha^{-1}(e^{\alpha(t-\bar{t})} - 1)$. Ricordando che $g \leq \delta + \alpha G$, si ottiene l’asserto (quando $t \geq \bar{t}$).

Nel caso $t < \bar{t}$, si ha che $g(t) \leq \delta + \alpha \int_t^{\bar{t}} g(s) ds$, e ponendo, come prima,

$$G(t) := \int_{\bar{t}}^t g(s) ds = - \int_t^{\bar{t}} g(s) ds,$$

si ottiene che $G' = g \leq \delta - \alpha G$ che è equivalente a $(Ge^{\alpha t})' \leq \delta e^{\alpha t}$. Integrando tale relazione tra t e \bar{t} , si ottiene $-G\alpha \leq \delta[e^{\alpha(\bar{t}-t)} - 1]$ da cui segue che $g \leq \delta e^{\alpha(\bar{t}-t)} = \delta e^{\alpha|t-\bar{t}|}$. ■

Anche se non intervengono nella discussione del Teorema di Picard–Lindelöf, includiamo in questo paragrafo altri due semplici risultati di calcolo in una variabile, che utilizzeremo nel § B.3.4 sull’analisi qualitativa di moti unidimensionali: il primo è una versione differenziale del Lemma di Gronwall, il secondo è un ‘teorema di confronto’ per soluzioni di disequazioni differenziali.

Lemma B.8 (Secondo Lemma di Gronwall) *Sia I un intervallo, $\bar{t} \in I$, $L > 0$. Assumiamo che $u \in C^1(I, \mathbb{R})$ sia tale che*

$$u'(t) \leq L|u(t)|, \quad \forall t \in I. \quad (\text{B.52})$$

Allora, se $u(\bar{t}) \leq 0$, si ha che $u(t) \leq 0$, per ogni $t \geq \bar{t}$; se $u(\bar{t}) \geq 0$, si ha che $u(t) \geq 0$, per ogni $t \leq \bar{t}$.

³⁵https://it.wikipedia.org/wiki/Lemma_di_Gronwall#cite_note-gronwall-1

Dimostrazione Innanzitutto, osserviamo che è sufficiente dimostrare il primo caso (il secondo caso segue dal primo sostituendo u con $-u$). Sia $u(\bar{t}) \leq 0$ e sia $\tau = \sup\{t \in I \mid u(t) \leq 0\}$. L'asserto è equivalente a dimostrare che $\tau = b := \sup I$. Supponiamo, per assurdo, che $\tau < b$. Allora $u(\tau) = 0$ e per ogni $s \in (\tau, b)$ si ha che $u(t) > 0$ per ogni $t \in J := (\tau, s]$. Fissiamo un tale s , allora, da (B.52) segue che $u' \leq Lu$ su J , ossia, $u' - Lu \leq 0$, che è equivalente a $(e^{-Lt}u)' \leq 0$. Integrando tale relazione tra τ e s otteniamo $e^{-Ls}u(s) \leq e^{-L\tau}u(\tau) = 0$, che è una contraddizione, essendo $u(s) > 0$. ■

Corollario B.9 Sia $A \subseteq \mathbb{R}^2$, $f, g \in C(A, \mathbb{R})$ con g localmente lipschitziana in x uniformemente in t . Sia I un intervallo, $\bar{t} \in I$ e siano $u, v \in C(I, \mathbb{R})$ tali che $(t, u(t)) \in A$ e $(t, v(t)) \in A$ per ogni $t \in I$ e tali che

$$\begin{cases} u'(t) \leq f(u(t), t) \\ v'(t) \geq g(v(t), t) \end{cases}, \quad (\text{B.53})$$

$$f(u(t), t) \leq g(u(t), t). \quad (\text{B.54})$$

Allora, se $u(\bar{t}) \leq v(\bar{t})$ si ha che $u(t) \leq v(t)$ per ogni $t \in I$, $t \geq \bar{t}$; se $u(\bar{t}) \geq v(\bar{t})$ si ha che $u(t) \geq v(t)$ per ogni $t \in I$, $t \leq \bar{t}$.

Dimostrazione Siano $a < \bar{t} < b$, $a, b \in I$ e sia

$$\Gamma = \{(u(t), t) \mid t \in [a, b]\} \cup \{(v(t), t) \mid t \in [a, b]\},$$

e $K \subseteq A$ un qualunque compatto che contenga Γ . Sia $L > 0$ tale che $|g(x, t) - g(y, t)| \leq L|x - y|$, per ogni $(x, t), (y, t) \in K$. Allora, se $w := u - v$ segue che

$$\begin{aligned} w' &= u' - v' \stackrel{(\text{B.53})}{\leq} f(u, t) - g(v, t) \stackrel{(\text{B.54})}{\leq} g(u, t) - g(v, t) \leq |g(u, t) - g(v, t)| \leq L|u - v| \\ &= L|w|, \end{aligned}$$

e la tesi, per $t \in [a, b]$, segue dal Lemma B.8 con u sostituito da w e I da $[a, b]$. Dall'arbitrarietà di a e b , segue la tesi per $t \in I$. ■

Spesso, il Corollario B.9 si applica alle equazioni (piuttosto che alle disequazioni) differenziali, nella seguente forma (che è un caso speciale del Corollario B.9):

Corollario B.10 (Teorema di confronto) Sia $A \subseteq \mathbb{R}^2$, $f, g \in C(A, \mathbb{R})$, $f \leq g$, e g localmente lipschitziana in x uniformemente in t . Sia I un intervallo, $\bar{t} \in I$ e siano $u, v \in C(I, \mathbb{R})$ tali che $(t, u(t)) \in A$ e $(t, v(t)) \in A$ per ogni $t \in I$ e tali che

$$\begin{aligned} u'(t) &= f(u(t), t), \\ v'(t) &= g(v(t), t). \end{aligned} \quad (\text{B.55})$$

Allora,

$$\begin{aligned} u(\bar{t}) \leq v(\bar{t}) &\implies u(t) \leq v(t), \quad \forall t \geq \bar{t}, \quad t \in I; \\ u(\bar{t}) \geq v(\bar{t}) &\implies u(t) \geq v(t), \quad \forall t \leq \bar{t}, \quad t \in I; \\ u(\bar{t}) = v(\bar{t}) &\implies \begin{cases} u(t) \leq v(t), & \forall t \geq \bar{t}, \quad t \in I, \\ u(t) \geq v(t), & \forall t \leq \bar{t}, \quad t \in I. \end{cases} \end{aligned}$$

B.2.4 Unicità e dipendenza continua dai dati iniziali

Dal Lemma di Gronwall segue facilmente il seguente risultato:

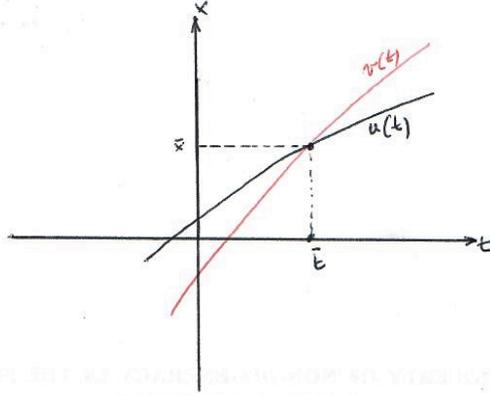


Figura B.3: Due soluzioni di $\dot{u} = f(u, t)$ e $\dot{v} = g(v, t)$ con $f \leq g$

Proposizione B.11 Sia $A \subseteq \mathbb{R}^{n+1}$ un aperto e $f \in \text{Loc}_*(A, \mathbb{R}^n)$. Siano, per $i = 1, 2$, $(\bar{x}_i, \bar{t}) \in A$ e $u_i \in C^1(I, \mathbb{R}^n)$ soluzioni dei problemi di Cauchy:

$$\begin{cases} u_i'(t) = f(u_i(t), t), & \forall t \in I, \\ u_i(\bar{t}) = \bar{x}_i, \end{cases} \quad (\text{B.56})$$

con $I := [a, b]$, $a \leq \bar{t} \leq b$ e

$$\Gamma = \{(u_1(t), t) \mid t \in [a, b]\} \cup \{(u_2(t), t) \mid t \in [a, b]\} \subseteq A.$$

Sia L la costante di Lipschitz in x relativa al compatto Γ . Allora, per ogni $t \in I$

$$|u_1(t) - u_2(t)| \leq |\bar{x}_1 - \bar{x}_2| e^{L|t - \bar{t}|}. \quad (\text{B.57})$$

In particolare, se $\bar{x}_1 = \bar{x}_2$, allora $u_1(t) = u_2(t)$ per ogni $t \in I$.

Dimostrazione Sia $g(t) := |u_1(t) - u_2(t)|$. Poiché le u_i verificano l'equazione integrale (B.40) (con \bar{x}_i al posto di \bar{x} e u_i al posto di u), si ha che

$$\begin{aligned} g(t) &= \left| \bar{x}_1 - \bar{x}_2 + \int_{\bar{t}}^t (f(u_1(s), s) - f(u_2(s), s)) ds \right| \\ &\leq |\bar{x}_1 - \bar{x}_2| + \left| \int_{\bar{t}}^t |f(u_1(s), s) - f(u_2(s), s)| ds \right| \\ &\leq |\bar{x}_1 - \bar{x}_2| + L \left| \int_{\bar{t}}^t |u_1(s) - u_2(s)| ds \right| \\ &= |\bar{x}_1 - \bar{x}_2| + L \left| \int_{\bar{t}}^t g(s) ds \right|. \end{aligned}$$

Dal Lemma di Gronwall con I sostituito da $[a, b]$, $\delta = |u_1 - u_2|$, $\alpha = L$, segue (B.57) per $t \in [a, b]$. L'ultima affermazione segue da (B.57). ■

Osservazione B.12 (Dipendenza continua dai dati iniziali) La Proposizione B.11 asserisce che le soluzioni locali del problema di Cauchy dipendono in modo continuo dal dato iniziale, infatti vi dipendono in modo lipschitziano con costante di Lipschitz data da $e^{L|t - \bar{t}|}$; cfr. (B.57).

Si noti, anche, che due soluzioni inizialmente vicine potrebbero allontanarsi l'una dall'altra, per $t > \bar{t}$ (o $t < \bar{t}$), ad un tasso *esponenziale*. In effetti tale fenomeno può realmente accadere: basti pensare al caso scalare $u' = Lu$ con $L > 0$. In tal caso $u_i(t) = e^{L(t-\bar{t})}\bar{x}_i$ e $|u_1(t) - u_2(t)| = e^{L(t-\bar{t})}|\bar{x}_1 - \bar{x}_2|$, il che mostra che la costante lipschitziana in (B.57) è ottimale.

La dimostrazione del Teorema di Picard–Lindelöf è, a questo punto, immediata:

Dimostrazione (del Teorema di Picard–Lindelöf B.5) Dal Lemma di Picard B.6 segue che $u = \lim \Phi(u_k; \bar{x}, \bar{t})$ (cfr. (B.45)) è una soluzione $C^1(I, D)$ del problema di Cauchy (B.39) e tale soluzione è unica per la Proposizione B.11 applicata con $A = D \times I$. ■

Osservazione B.13 (i) L'enunciato del Teorema B.5 è *quantitativo*, ossia, fornisce una stima del dominio e del codominio della soluzione locale del problema di Cauchy. Una versione sintetica qualitativa del Teorema di estenza ed unicità locale per equazioni differenziali ordinarie che discende immediatamente dal Teorema di Picard–Lindelöf è la seguente:

Teorema di esistenza e unicità locale per EDO³⁶ (enunciato qualitativo)

Sia $f : (x, t) \in A \subseteq \mathbb{R}^{n+1} \mapsto f(x, t) \in \mathbb{R}^n$ una funzione localmente lipschitziana rispetto a x uniformemente in t . Allora, esiste una ed una sola soluzione locale del problema di Cauchy (B.39).

(ii) Si noti che nell'enunciato del Teorema di Picard–Lindelöf, la costante di Lipschitz L non appare; essa compare, come appena osservato, nella dipendenza dai dati iniziali oltre che nella dimostrazione del Lemma di Picard).

(iii) (**Regolarità in t**) Le soluzioni (classiche³⁷) del problema di Cauchy sono, per definizione, differenziabili. In effetti, se il campo f è più regolare, le soluzioni di un'equazione differenziale sono automaticamente più regolari anch'esse. Per esempio, vale la seguente affermazione:

Sia $k \in \mathbb{N}$, A un aperto di \mathbb{R}^{n+1} , $f \in C^k(A, \mathbb{R}^n)$, I un intervallo, $u \in C^1(I, \mathbb{R}^n)$ tale che $(u(t), t) \in A$ per ogni $t \in I$ e $u' = f(u(t), t)$, per ogni $t \in I$. Allora $u \in C^{k+1}(I)$.

Dimostrazione Dimostriamo, per induzione su $1 \leq j \leq k+1$, che $u \in C^j(I)$. Per ipotesi $u \in C^1$. Assumiamo ora $u \in C^j(I)$ per un $1 \leq j \leq k$ e dimostriamo che $u \in C^{j+1}(I)$. Dalla Proposizione 2.33, segue (essendo $k \geq j$) che la funzione composta $t \in I \mapsto f(u(t), t)$ è di classe $C^j(I)$. Ma, per l'equazione differenziale, $f(u(t), t) = u'(t)$ e quindi $t \mapsto u'(t) \in C^j(I)$, che è equivalente a dire che $u \in C^{j+1}(I)$. ■

Questa osservazione permette di calcolare, iterativamente, le derivate successive di u dall'equazione $u' = f(u, t)$. Per esempio, se $f \in C^1(A)$, allora

$$u'' = f_x(u, t)u' + f_t(u, t) = f_x(u, t)f(u, t) + f_t(u, t), \quad (\text{B.58})$$

(dove u e le sue derivate sono calcolate in t), o, in componenti,

$$u_i'' = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(u, t)u_j' + \frac{\partial f_i}{\partial t}(u, t) = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(u, t)f_j(u, t) + \frac{\partial f_i}{\partial t}(u, t). \quad (\text{B.59})$$

B.3 Soluzioni massimali e soluzioni globali

L'unicità locale permette di estendere, in modo massimale, le soluzioni locali del problema di Cauchy. In questa sezione daremo la definizione di soluzione massimale di un problema di Cauchy, discuteremo quando è possibile estendere una soluzione locale. Discuteremo, poi, qual

³⁶ Acronimo per 'Equazioni Differenziali Ordinarie'.

³⁷ Si possono considerare classi di soluzioni più generali ('soluzioni deboli') e non necessariamente differenziabili di equazioni differenziali, cfr. https://en.wikipedia.org/wiki/Weak_solution; in questo testo considereremo solo soluzioni *classiche* ossia differenziabili.

è il comportamento di una soluzione massimale quando uno degli estremi su cui è definita è finito ed, infine, daremo un criterio per poter estendere per $t \rightarrow \pm\infty$ soluzioni locali.

In questa sezione, se non altrimenti specificato, A denota un aperto di \mathbb{R}^{n+1} , U un aperto di \mathbb{R}^n , f una funzione in $\text{Lip}_*(A, \mathbb{R}^n)$.

B.3.1 Definizioni

Definizione B.14 Se $(\bar{x}, \bar{t}) \in A$, denotiamo $\mathcal{C}(\bar{t}, \bar{x}) := \mathcal{C}_f(\bar{t}, \bar{x})$ l'insieme di tutte le soluzioni classiche (locali) del problema di Cauchy (B.39), più precisamente: $u \in \mathcal{C}(\bar{t}, \bar{x})$ se e solo se esiste un intervallo aperto I_u tale che: $\bar{t} \in I_u$; $u \in C^1(I_u, \mathbb{R}^n)$; $(u(t), t) \in A$ per ogni $t \in I_u$; $u(\bar{t}) = \bar{x}$; $u' = f(u(t), t)$ per ogni $t \in I_u$.

L'insieme delle soluzioni classiche $\mathcal{C}_f(\bar{t}, \bar{x})$ è sempre diverso dall'insieme vuoto, grazie al già citato Teorema di Peano (cfr. Osservazione (iv) in § B.1.4); inoltre, nel caso si abbia unicità delle soluzioni locali (cosa garantita, ad esempio, se f è localmente lipschitziana) l'insieme $\mathcal{C}_f(\bar{t}, \bar{x})$ è un insieme parzialmente ordinato:

Lemma B.15 Se $u_1, u_2 \in \mathcal{C}_f(\bar{t}, \bar{x})$, allora $u_1(t) = u_2(t)$ per ogni $t \in I_{u_1} \cap I_{u_2}$.

Dimostrazione Segue immediatamente dalla Proposizione B.11. ■

Definizione B.16 (i) Se $u, v \in \mathcal{C}_f(\bar{t}, \bar{x})$ e $I_u \subseteq I_v$ diremo che v **estende** u (o che v è un'estensione di u) e scriveremo $u \prec v$; se $I_u \subsetneq I_v$, diremo che v **estende** u **propriamente** (o che v è un'estensione propria di u). Se $I_u = (a, b)$ e $I_v = (a, c)$ con $c > b$ [rispettivamente, $I_v = (c, b)$ con $c < a$] diremo che v estende propriamente u a destra [risp., a sinistra] o che v è un'estensione destra [risp. sinistra] propria di u .

(ii) Poniamo

$$I_{\max}(\bar{t}, \bar{x}) := \bigcup_{u \in \mathcal{C}(\bar{t}, \bar{x})} I_u, \quad (\text{B.60})$$

e, per $t \in I_{\max}(\bar{t}, \bar{x})$, definiamo $u_{\max}(t; \bar{t}, \bar{x}) := u(t)$ dove $u \in \mathcal{C}(\bar{t}, \bar{x})$ è tale che³⁸ $t \in I_u$. L'intervallo $I_{\max}(\bar{t}, \bar{x})$, che si denota anche (T_-, T_+) o $(T_-(\bar{t}, \bar{x}), T_+(\bar{t}, \bar{x}))$, prende il nome di **intervallo massimale di esistenza** per (B.39) e la funzione u_{\max} prende il nome di **soluzione massimale del problema di Cauchy** (B.39). Spesso denoteremo la soluzione massimale $u_{\max}(t; \bar{t}, \bar{x})$ con il simbolo $\phi^t(\bar{t}, \bar{x})$.

(iii) Se $I_{\max} = \mathbb{R}$ diremo che u_{\max} è una **soluzione globale** di (B.39).

Osservazione B.17 (i) Nella definizione data di u_{\max} , il valore $u_{\max}(t)$ non dipende dalla particolare scelta di $u \in \mathcal{C}(\bar{t}, \bar{x})$: se $v \in \mathcal{C}(\bar{t}, \bar{x})$ è anche tale che $t \in I_v$, per il Lemma B.15, $v(t) = u(t)$. Dunque la definizione di u_{\max} nella Definizione B.16–(ii) è ben posta. Si noti che dalla definizione data segue che

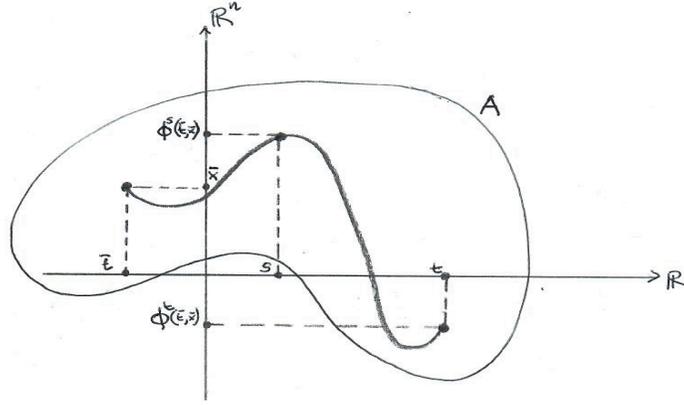
$$\phi^{\bar{t}}(\bar{t}, \bar{x}) = \bar{x}, \quad \forall (\bar{x}, \bar{t}) \in A. \quad (\text{B.61})$$

(ii) Chiaramente, la relazione ' \prec ' è una relazione d'ordine parziale su $\mathcal{C}_f(\bar{t}, \bar{x})$ e u_{\max} è l'elemento massimale in $\mathcal{C}(\bar{t}, \bar{x})$ (ossia $u \prec u_{\max}$ per ogni $u \in \mathcal{C}_f(\bar{t}, \bar{x})$), o, in altri termini, u_{\max} non ammette estensioni proprie. Dunque possiamo identificare $\mathcal{C}_f(\bar{t}, \bar{x})$ con la soluzione massimale $t \in I_{\max}(\bar{t}, \bar{x}) \mapsto \phi^t(\bar{t}, \bar{x})$: ogni elemento di $u \in \mathcal{C}_f(\bar{t}, \bar{x})$ non è altro che una restrizione di $t \rightarrow \phi^t(\bar{t}, \bar{x})$ ad un intervallo aperto $I_u \subseteq I_{\max}(\bar{t}, \bar{x})$ che contenga \bar{t} .

(iii) Per ogni $(\bar{x}, \bar{t}) \in A$, si ha

$$I_{\max}(s, \phi^s(\bar{t}, \bar{x})) = I_{\max}(\bar{t}, \bar{x}), \quad \forall s \in I := I_{\max}(\bar{t}, \bar{x}), \quad (\text{B.62})$$

$$\phi^t(s, \phi^s(\bar{t}, \bar{x})) = \phi^t(\bar{t}, \bar{x}), \quad \forall t, s \in I. \quad (\text{B.63})$$

Figura B.4: Il 'flusso' ϕ^t

Dimostrazione Siano: $u(t) := \phi^t(\bar{t}, \bar{x})$, $I_u := I_{\max}(\bar{t}, \bar{x})$, $y := \phi^s(\bar{t}, \bar{x})$, $v(t) := \phi^t(s, y)$, e $I_v := I_{\max}(s, y)$. La funzione u verifica $u'(t) = f(u(t), t)$ e $u(s) = y$, quindi $u \in \mathcal{C}_f(s, y)$ e dunque, per definizione di $I_{\max}(s, y)$ e $t \rightarrow \phi^t(s, y) = u_{\max}(t; s, y)$, si deve avere $I_u \subseteq I_{\max}(s, y) = I_v$ e $u \prec v = \phi^t(s, y)$. Su $I_u = I_u \cap I_v$, per il Lemma B.15 u e v (essendo due elementi di $\mathcal{C}_f(s, y)$) coincidono; in particolare $\bar{x} = u(\bar{t}) = v(\bar{t})$ e dunque $v \in \mathcal{C}_f(\bar{t}, \bar{x})$. Ma allora, deve essere $I_u = I_v$ (altrimenti v sarebbe un'estensione propria di u in $\mathcal{C}_f(\bar{t}, \bar{x})$, il che è impossibile essendo $u = u_{\max}(\cdot; \bar{t}, \bar{x})$) e quindi $u = v$. ■

(iv) Un esempio in cui tutte le soluzioni del problema di Cauchy sono globali è dato dai sistemi di equazioni lineari a coefficienti costanti (B.26) (cfr. Proposizione B.1).

B.3.2 Criteri di estensione e di globalità

Il seguente risultato fornisce una caratterizzazione semplice dell'estendibilità delle soluzioni del problema di Cauchy:

Lemma B.18 Sia $(\bar{x}, \bar{t}) \in A$. Una soluzione $u \in \mathcal{C}(\bar{t}, \bar{x})$ con $I_u = (a, b)$, $b \in \mathbb{R}$, ammette una estensione destra propria v (ossia, $I_v = (a, b + \varepsilon)$ con $\varepsilon > 0$) se e solo se $\lim_{t \rightarrow b^-} u(t) = x_0 \in \mathbb{R}^n$ e $(x_0, b) \in A$. Analoga affermazione vale per l'estremo sinistro.

Dimostrazione Il 'solo se' è ovvio: se v estende propriamente u in un intorno di b , per il Lemma B.15, $v = u$ su I_u e quindi $\lim_{t \rightarrow b^-} u(t) = \lim_{t \rightarrow b^-} v = v(b)$ e $(v(b), b) \in A$. Dimostriamo il 'se'. Poiché $(x_0, b) \in A$, esiste una soluzione locale $w \in \mathcal{C}(b, x_0)$ definita su $I_w = (b - \varepsilon, b + \varepsilon)$ per un $\varepsilon > 0$. Definiamo,

$$v(t) = \begin{cases} u(t), & \text{se } t \in (a, b), \\ w(t) & \text{se } b \leq t < b + \varepsilon. \end{cases}$$

Per l'ipotesi su u e la definizione di $\mathcal{C}(b, x_0)$, $\lim_{t \rightarrow b^-} u(t) = x_0 = w(x_0)$. Quindi v è continua in b . Inoltre, poiché u e w sono soluzioni della stessa equazione differenziale,

$$\begin{aligned} \lim_{t \rightarrow b^-} v'(t) &= \lim_{t \rightarrow b^-} u'(t) = \lim_{t \rightarrow b^-} f(u(t), t) = f(x_0, b) \\ &= \lim_{t \rightarrow b^+} f(w(t), t) = \lim_{t \rightarrow b^+} w'(t) = \lim_{t \rightarrow b^+} v'(t), \end{aligned}$$

³⁸Per definizione di I_{\max} una tale u esiste sempre.

il che mostra che $v \in C^1((a, b + \varepsilon))$ e, dunque, che $v \in \mathcal{C}(\bar{t}, \bar{x})$.

L'affermazione corrispondente per l'estremo sinistro si dimostra in maniera analoga. ■

Osservazione B.19 (i) Nel problema di Cauchy (B.31) con $(A = \mathbb{R}^2)$ e $\bar{x} \neq 0$, la soluzione u trovata col metodo di separazione delle variabili (B.35) è definita sulla semiretta $I_u = I(\bar{t}, \bar{x})$ definita in (B.35): poiché il $\lim_{t \rightarrow T_c} u(t)$ non è finito, dal Lemma B.18 segue che $I_{\max}(\bar{t}, \bar{x}) = I(\bar{t}, \bar{x})$, intervallo che ha uno dei due estremi finito.

(ii) Se una soluzione u definita su un intervallo $I_u = (a, b)$ è tale che $\exists t_k \nearrow b$ con $u(t_k) \rightarrow x_0$ e $(x_0, b) \in \partial A$ allora u non ammette estensione propria destra.

Dimostrazione Se u avesse una estensione destra $v : (a, b + \varepsilon) \rightarrow \mathbb{R}^n$, per un qualche $\varepsilon > 0$, si dovrebbe avere che $(t, v(t)) \in A$ per ogni $t \in (a, b + \varepsilon)$ il che è falso essendo $(v(b), b) = (x_0, b) \in \partial A \subseteq A^c$ (essendo A aperto). ■

(iii) Dal Lemma B.18 e dalla definizione di soluzione massimale segue immediatamente la seguente affermazione:

Sia $A \subseteq \mathbb{R}^{n+1}$ un aperto, $f \in \text{Lip}_*(A, \mathbb{R}^n)$ e sia $(\bar{x}, \bar{t}) \in A$ e sia $(T_-, T_+) := I_{\max}(\bar{t}, \bar{x})$. Se esiste $\lim_{t \rightarrow T_+} u_{\max}(t; \bar{t}, \bar{x}) =: x_+$, si ha che $(x_+, T_+) \notin A$; analogamente per $t \rightarrow T_-$.

È interessante notare che per soluzioni $u(t)$ di equazioni differenziali, l'esistenza del limite $x_0 = \lim_{t \rightarrow b} u(t)$ è garantita dall'esistenza del limite di una qualunque successione $u(t_k)$ con $t_k \rightarrow b$, purché il punto (x_0, b) appartenga al dominio di f : questo è il contenuto del prossimo lemma, la cui dimostrazione è particolarmente istruttiva.

Lemma B.20 Sia $(\bar{x}, \bar{t}) \in A$, sia $u \in \mathcal{C}(\bar{t}, \bar{x})$ e $(a, b) := I_u(\bar{t}, \bar{x})$. Assumiamo che esista $x_0 \in \mathbb{R}^n$ tale che $(x_0, b) \in A$ e $t_k \in (a, b)$ tali che $t_k \rightarrow b$, $u(t_k) \rightarrow x_0$. Allora, $\lim_{t \rightarrow b} u(t) = x_0$. Analoga affermazione vale per l'estremo sinistro.

Dimostrazione Passando eventualmente ad una sottosuccessione, possiamo assumere che la successione $\{t_k\}$ sia strettamente crescente³⁹.

Supponiamo, per assurdo, che non sia vero che $\lim_{t \rightarrow b} u(t) = x_0$, ossia,

$$\exists \varepsilon > 0 \mid \forall \hat{t} \in (a, b), \quad \exists \hat{s} \in (\hat{t}, b) \quad \text{t.c.} \quad |u(\hat{s}) - x_0| \geq \varepsilon. \quad (\text{B.64})$$

Poiché $(x_0, b) \in A$, esistono $0 < \rho < \varepsilon$ e $0 < \tau$ tali che

$$E := D \times [b - \tau, b] := \{x \in \mathbb{R}^n \mid |x - x_0| \leq \rho\} \times [b - \tau, b] \subseteq A. \quad (\text{B.65})$$

Costruiamo, ora, iterativamente, due successioni s_j e k_j come segue.

Sia k_0 tale che

$$|u(t_k) - x_0| < \rho, \quad \forall k \geq k_0, \quad (\text{B.66})$$

(k_0 esiste poiché $u(t_k) \rightarrow x_0$ per ipotesi) e sia $k_1 \geq k_0$ tale che $t_{k_1} > b - \tau$ (k_1 esiste perché $t_k \nearrow b$). Per (B.64),

$$\exists \hat{s} \in (t_{k_1}, b) \mid |u(\hat{s}) - x_0| \geq \varepsilon > \rho. \quad (\text{B.67})$$

La funzione $t \in [t_{k_1}, \hat{s}] \mapsto \delta(t) := |u(t) - x_0| - \rho$ è continua e, per (B.66) (essendo $k_1 \geq k_0$), $\delta(t_{k_1}) < 0$, mentre, per (B.67), $\delta(\hat{s}) > 0$. Dunque dal teorema di esistenza degli zeri – o meglio ‘del primo zero’ – per funzioni continue⁴⁰ esiste $s_1 \in (t_{k_1}, \hat{s})$ tale che $\delta(s_1) = 0$, ossia $|u(s_1) - x_0| = \rho$. Sia, poi, $k_2 > k_1$, tale che $t_{k_2} > s_1$; si noti che (essendo $k_2 > k_1 \geq k_0$) $|u(t_{k_2}) - x_0| < \rho$ e che, essendo $\{t_k\}$ crescente $t_{k_2} > t_{k_1} > b - \tau$. Come prima, da (B.64) e dalla continuità di u , segue che esiste

³⁹Ad esempio, si può prendere $n_1 := 1$ e ricorsivamente, per $k \geq 2$, $n_k := \min\{k > n_{k-1} \mid t_k > t_{n_{k-1}}\}$: si noti che $t_k < b$ per ogni k e quindi l'insieme appena definito è un sottoinsieme non vuoto di \mathbb{N} e quindi ha minimo. La sottosuccessione t_{n_k} è strettamente crescente e $t_{n_k} \rightarrow b$.

⁴⁰Vedi, per esempio, Teorema 2.51 in [C2019].

$s_2 > t_{k_2}$ tale che $|u(t) - x_0| < \rho$, $\forall t_{k_2} \leq t < s_2$, e $|u(s_2) - x_0| = \rho$. Iterando, troviamo due successioni k_j e s_j tali che

$$k_{j+1} > k_j; \quad b - \tau < t_{k_j} < s_j < t_{k_{j+1}}; \quad \begin{cases} |u(t) - x_0| < \rho, \quad \forall t \in [t_{k_j}, s_j), \\ |u(s_j) - x_0| = \rho. \end{cases} \quad (\text{B.68})$$

Si noti che anche da (B.68), segue che $s_j \rightarrow b$. Allora avremmo⁴¹

$$\begin{aligned} \rho &\stackrel{(\text{B.68})}{=} |u(s_j) - x_0| \leq |u(s_j) - u(t_{k_j})| + |u(t_{k_j}) - x_0| \\ &= \left| \int_{t_{k_j}}^{s_j} u'(s) ds \right| + |u(t_{k_j}) - x_0| \\ &= \left| \int_{t_{k_j}}^{s_j} f(u(s), s) ds \right| + |u(t_{k_j}) - x_0| \\ &\leq \max_E |f| \cdot (s_j - t_{k_j}) + |u(t_{k_j}) - x_0| \rightarrow 0. \end{aligned}$$

ottenendo una contraddizione essendo $\rho > 0$. ■

Osservazione B.21 Mettendo assieme i Lemmi B.18 e B.20 otteniamo il seguente

Criterio di estensione Sia $A \subseteq \mathbb{R}^{n+1}$ un aperto, $f \in \text{Lip}_*(A, \mathbb{R}^n)$, sia $(\bar{x}, \bar{t}) \in A$, sia $u \in \mathcal{C}(\bar{t}, \bar{x})$ e $(a, b) := I_u(\bar{t}, \bar{x})$. Assumiamo che esista $x_0 \in \mathbb{R}^n$ e $t_k \in (a, b)$ tali che $t_k \rightarrow b$, $u(t_k) \rightarrow x_0$ e che $(x_0, b) \in A$. Allora, u ammette una estensione propria destra. Analoga affermazione vale per l'estremo sinistro.

Vediamo due conseguenze notevoli del criterio di estensione.

Il primo risultato descrive il comportamento della soluzione massimale vicino ad un estremo finito dell'intervallo massimale di esistenza:

Proposizione B.22 Sia $(\bar{x}, \bar{t}) \in A$, sia $(a, b) := I_{\max}(\bar{t}, \bar{x})$ con $b \in \mathbb{R}$. Allora:

- (i) per ogni compatto $K \subseteq \mathbb{R}^n$ tale che $K \times \{b\} \subseteq A$, esiste $\hat{t} \in (\bar{t}, b)$ tale che $u_{\max}(\hat{t}) \notin K$;
- (ii) se $A = \mathbb{R}^n \times J$ con J intervallo e $b \in J$, esiste $J \ni t_k \rightarrow b$, tale che $|u_{\max}(t_k)| \rightarrow +\infty$.

Osservazione B.23 (i) Naturalmente, vale un enunciato del tutto analogo nel caso dell'estremo sinistro (che si può desumere dalla Proposizione B.22 con un opportuno cambio di variabile, oppure rifacendo la dimostrazione con gli ovvi cambiamenti).

(ii) A parole, la Proposizione B.22 afferma che, se un estremo c dell'intervallo massimale di esistenza è finito, le soluzioni del problema di Cauchy escono eventualmente da un qualunque compatto di $K \subseteq \mathbb{R}^n$ tale che $K \times \{c\} \subseteq A$.

(iii) L'ipotesi che $K \times \{c\} \subseteq A$ è essenziale, come mostra il seguente esempio banale. Consideriamo il problema di Cauchy $u' = f(u, t)$ con $f := 0$ e $A := \mathbb{R}^n \times J$ con $J = (0, 1)$. Ovviamente, per ogni $(\bar{x}, \bar{t}) \in A$, $I_{\max} = J$ e $u_{\max}(t) \equiv \bar{x}$, che quindi non esplose quando t si avvicina agli estremi di I_{\max} . D'altra parte, in questo caso, l'ipotesi che $K \times \{1\} \subseteq A$ non è verificata per alcun compatto K semplicemente perché $1 \notin J$.

Dimostrazione (della Proposizione B.22) (i): Supponiamo, per assurdo, che esista un compatto K tale che $K \times \{b\} \subseteq A$ e tale che $u_{\max}(t) \in K$ per ogni $t \in (\bar{t}, b)$. Sia $\bar{t} < t_k \nearrow b$ e sia $x_k := u(t_k)$. Poiché K è compatto esiste k_j e $\bar{x} \in K$ tali che $x_{k_j} \rightarrow \bar{x}$. Poiché $(\bar{x}, b) \in K \times \{b\} \subseteq A$, dai Lemmi B.20 e B.18 seguirebbe che u_{\max} potrebbe essere estesa, il che è una contraddizione.

(ii): Sia $k \in \mathbb{N}$. Da (i) con $K = \{x \in \mathbb{R}^n \mid |x| \leq k\}$ segue esiste t_k tale che $|u_{\max}(t_k)| > k$. ■

⁴¹Si noti che, grazie a (B.68), $u(t) \in D$ (cfr. (B.65)) per ogni $t_{k_j} \leq t \leq s_j$, ed essendo D convesso (e $E \subseteq A$ dove f è C^1) possiamo applicare nella seconda uguaglianza il teorema del valor medio scalare applicato componente per componente (o direttamente la Proposizione 2.18–(i)); nella terza uguaglianza abbiamo usato l'equazione differenziale $u'(t) = f(u(t), t)$.

Proposizione B.24 (Criterio di globalità)

Sia I un intervallo aperto, $f \in \text{Lip}_*(\mathbb{R}^n \times I, \mathbb{R}^n)$ e tale che, per ogni $\alpha < \beta$ tali che $[\alpha, \beta] \subseteq I$, esistono $L, M > 0$ per cui

$$|f(x, t)| \leq M + L|x|, \quad \forall t \in [\alpha, \beta], \forall x \in \mathbb{R}^n. \quad (\text{B.69})$$

Allora, per ogni $(\bar{x}, \bar{t}) \in \mathbb{R}^n \times I$, $I_{\max}(\bar{t}, \bar{x}) = I$. In particolare, se $I = \mathbb{R}$, ogni soluzione di $u' = f(u, t)$ è una soluzione globale.

Dimostrazione Fissiamo $(\bar{x}, \bar{t}) \in \mathbb{R}^n \times I$, sia $(a, b) := I_{\max}(\bar{x}, \bar{t}) \subseteq I$ e supponiamo, per assurdo, che $(a, b) \subsetneq I$: supponiamo per esempio che $b < \sup I$ (il caso $\inf I < a$ si discute in maniera del tutto analoga). Poiché $[\alpha, \beta] := [\bar{t}, b] \subseteq I$, segue che esistono $L, M > 0$ tali che vale (B.69). Quindi, per ogni $t \in [\bar{t}, b]$,

$$\begin{aligned} |u(t)| &= \left| \bar{x} + \int_{\bar{t}}^t f(u(s), s) ds \right| \leq |\bar{x}| + \int_{\bar{t}}^t |f(u(s), s)| ds \\ &\stackrel{(\text{B.69})}{\leq} |\bar{x}| + \int_{\bar{t}}^t (M + L|u(s)|) ds \\ &\leq |\bar{x}| + M(b - \bar{t}) + L \int_{\bar{t}}^t |u(s)| ds. \end{aligned}$$

Quindi, dal Lemma di Gronwall B.7 (con I sostituito da $[\bar{t}, b]$ e $g(t) = |u(t)|$, $\delta = |\bar{x}| + M(b - \bar{t})$, $\alpha = L$), segue che $|u(t)| \leq R$ con $R = \delta e^{L(b - \bar{t})}$, ossia che $u(t) \in \{x \in \mathbb{R}^n \mid |x| \leq R\}$ per ogni $t \in [\bar{t}, b]$, ma questo contraddice il punto (ii) della Proposizione B.22. ■

B.3.3 Il caso autonomo: flussi e sistemi dinamici

Un caso particolarmente importante è il caso *autonomo*, ossia, il caso in cui il campo $f : U \subseteq \mathbb{R}^n \mapsto \mathbb{R}^n$ non dipenda esplicitamente dal tempo. In tal caso, l'evoluzione, al variare di t di una soluzione massimale $\phi^t(\bar{t}, \bar{x}) = u_{\max}(t; \bar{t}, \bar{x})$, con $\bar{x} \in U$ descrive un *flusso* in U , che soddisfa la *proprietà di gruppo* descritta nel seguente lemma.

Lemma B.25 (Proprietà di gruppo dei flussi) Sia U un aperto di \mathbb{R}^n e sia $f : U \rightarrow \mathbb{R}^n$ una funzione localmente lipschitziana⁴². Dato $\bar{x} \in U$, poniamo:

$$I_{\max}(\bar{x}) := I_{\max}(0, \bar{x}), \quad \phi^t(\bar{x}) := \phi^t(0, \bar{x}). \quad (\text{B.70})$$

Allora, se $\bar{x} \in U$ e $I_{\max}(\bar{x}) =: (a, b)$, si ha

$$I_{\max}(\bar{t}, \bar{x}) = \bar{t} + I_{\max}(\bar{x}) := (\bar{t} + a, \bar{t} + b), \quad \forall \bar{t} \in \mathbb{R}, \quad (\text{B.71})$$

$$\phi^t(\bar{t}, \bar{x}) = \phi^{t - \bar{t}}(\bar{x}), \quad \forall \bar{t} \in \mathbb{R}, t - \bar{t} \in I_{\max}(\bar{x}), \quad (\text{B.72})$$

$$\phi^{t_1 + t_2}(\bar{x}) = \phi^{t_1} \circ \phi^{t_2}(\bar{x}), \quad \forall t_2, t_1 + t_2 \in I_{\max}(\bar{x}). \quad (\text{B.73})$$

Dimostrazione Sia $\bar{t} \in \mathbb{R}$ e poniamo, per ogni $t \in I_{\bar{t}} := (\bar{t} + a, \bar{t} + b)$, $u(t) := \phi^{t - \bar{t}}(\bar{x})$. Allora, u verifica $u' = f(u)$ su $I_{\bar{t}}$ e $u(\bar{t}) = \bar{x}$. Dunque $u \in \mathcal{C}_f(\bar{t}, \bar{x})$. Inoltre, u non ammette estensioni proprie: se $t \rightarrow v(t)$ fosse un'estensione propria di u , allora $t \rightarrow \tilde{v}(t) := v(\bar{t} + t)$ sarebbe un'estensione propria di u , il che è impossibile essendo $u(t) = \phi^t(\bar{x})$ la soluzione massimale del problema di Cauchy: (B.71) e (B.72) seguono.

Usando (B.72) e la (B.63) (con $t = t_1$, $s = 0$ e $\bar{t} = -t_2$), si ha che

$$\begin{aligned} \phi^{t_1 + t_2}(\bar{x}) &\stackrel{(\text{B.72})}{=} \phi^{t_1}(-t_2, \bar{x}) \stackrel{(\text{B.63})}{=} \phi^{t_1}(0, \phi^0(-t_2, \bar{x})) \stackrel{(\text{B.72})}{=} \phi^{t_1}(\phi^0(-t_2, \bar{x})) \stackrel{(\text{B.72})}{=} \phi^{t_1}(\phi^{t_2}(\bar{x})) \\ &= \phi^{t_1} \circ \phi^{t_2}(\bar{x}). \quad \blacksquare \end{aligned}$$

⁴²Ossia, per ogni compatto $D \subseteq U$ esiste $L > 0$ tale che $|f(x) - f(y)| \leq L|x - y|$, per ogni $x, y \in D$.

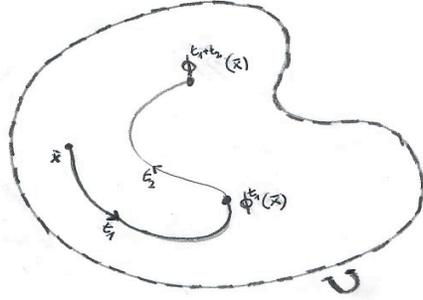


Figura B.5: Proprietà di gruppo del flusso di un'equazione differenziale

Osservazione B.26 (i) Da (B.72) e (B.71) segue che, nel caso autonomo, non è restrittivo prendere $t = 0$ nel problema di Cauchy, poiché il caso generale, segue da questo per traslazione.
(ii) $\{\phi^t\}$ è un 'gruppo (commutativo) ad un parametro di omeomorfismi': $\phi^0 = \text{id}$, l'inverso di $\phi^t(x)$ è ϕ^{-t} .

Concludiamo questa sezione con alcune definizioni relative alla 'dinamica dei flussi', che introducono il linguaggio della 'teoria dei sistemi dinamici'.

Definizione B.27 Sia U un aperto di \mathbb{R}^n e $f : U \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^n$ una funzione localmente lipschitziana.

(i) La coppia $\mathcal{S} = (\phi^t, U)$ è un **sistema dinamico** (su U). L'insieme U si chiama **spazio delle configurazioni** (o 'spazio degli stati') di \mathcal{S} e le funzioni $x(t) = \phi^t(x)$, $x \in U$, sono le **orbite** (o 'traiettorie' o 'moti') di \mathcal{S} .

(ii) Un sottoinsieme $E \subseteq U$ è un **insieme invariante** per \mathcal{S} se $\phi^t : E \rightarrow E$ (o, più precisamente, se $\forall x \in E$, $\phi^t(x) \in E$ per ogni $t \in I_{\max}(x)$).

(iii) Un punto $x_0 \in E$ è un **equilibrio** (o 'punto stazionario') per \mathcal{S} (o, semplicemente, 'per f ') se $\phi^t(x_0) = x_0$ per ogni $t \in \mathbb{R}$.

(iv) Un'orbita $x(t) = \phi^t(x)$ è **periodica** con periodo $T > 0$ se, $I_{\max}(x) = \mathbb{R}$, $x(t+T) = x(t)$ per ogni $t \in \mathbb{R}$, e $T_{\min} := \min\{T > 0 \text{ tale che } x(t+T) = x(t) \text{ per ogni } t \in \mathbb{R}\} > 0$. T_{\min} prende il nome di **periodo minimo** dell'orbita periodica x . Di norma, con 'il periodo di un'orbita periodica' ci si riferisce al periodo minimo.

(v) Una traiettoria non banale (ossia, non un equilibrio) $x(t) = \phi^t(x)$ si dice **asintotica** a $x_0 \in U$ per $t \rightarrow +\infty$ (risp., $t \rightarrow -\infty$) se $\lim_{t \rightarrow +\infty} x(t) = x_0$ (risp., $\lim_{t \rightarrow -\infty} x(t) = x_0$); se $\lim_{t \rightarrow +\infty} x(t) = \lim_{t \rightarrow -\infty} x(t) = x_0$ l'orbita si dice **omoclina**, se $\lim_{t \rightarrow +\infty} x(t) = x_0 \neq x_1 = \lim_{t \rightarrow -\infty} x(t)$ ($x_0, x_1 \in U$) l'orbita si dice **eteroclina**. In generale, analoghe definizioni si possono dare sostituendo x_0 con un insieme invariante E e ' $\lim x(t) = x_0$ ' con ' $\lim \text{dist}(x(t), E) = 0$ '.

Osservazione B.28 Raccogliamo qui alcune osservazioni elementari sulle definizioni appena date.

Nello studio di flussi è d'uso comune usare la notazione di Newton (il punto anziché l'apice per denotare la derivata e la x per la funzione incognita).

(i) Un punto x_0 è un equilibrio per f se e solo se $f(x_0) = 0$.

Dimostrazione Se $x(t) \equiv x_0$ è un'orbita, allora $0 = \dot{x}(t) = f(x(t)) = f(x_0)$. Viceversa, se $f(x_0) = 0$ allora $t \mapsto x(t) = x_0$ è soluzione di $\dot{x} = f(x)$ e dunque $\phi^t(x_0) \equiv x_0$ è un equilibrio. ■

(ii) Un equilibrio $x(t) \equiv x_0$ non è, secondo la definizione in (iv) un'orbita periodica, poiché, in tal caso, ogni $T > 0$ è un periodo e quindi $\min\{T > 0 \mid x(t+T) = x(t)\} = 0$.

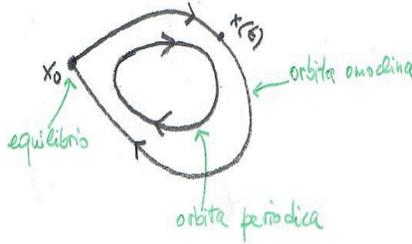


Figura B.6: Orbita omoclina bi-asintotica ad un equilibrio e orbita periodica

(iii) Un'orbita $t \rightarrow x(t)$ è periodica di periodo T se e solo se $x(T) = x(0)$ e $x(t)$ non è identicamente costante.

Dimostrazione Il 'solo se' deriva immediatamente dalla definizione di orbita periodica (prendendo $t = 0$). Assumiamo che $x(t)$ sia un'orbita non identicamente costante tale $x(T) = x(0) =: x_0$. Allora, per ogni t ,

$$x(t+T) = \phi^{t+T}(x_0) \stackrel{\text{(B.73)}}{=} \phi^t(\phi^T(x_0)) = \phi^t(x(T)) = \phi^t(x_0) = x(t). \quad \blacksquare$$

(iv) Sia $x(t) := \phi^t(x_0)$ con $x_0 \in U$ e sia $(T_-, T_+) = I_{\max}(x_0)$. Se esiste $\lim_{t \rightarrow T_+} x(t) = \bar{x} \in U$, allora $T_+ = +\infty$ e \bar{x} è un equilibrio; in altri termini, $x(t)$ è un'orbita asintotica per $t \rightarrow T_+$. Analogamente per T_- .

Dimostrazione Che T_+ sia $+\infty$ segue immediatamente dal Lemma B.18 e dalla definizione di T_+ . Ora, la i -ma componente dell'orbita, $\alpha(t) := x_i(t)$ è tale che $\alpha(t) \rightarrow x_{0i} \in \mathbb{R}$ ed inoltre dall'equazione differenziale segue che $\dot{\alpha} = f_i(x(t)) \rightarrow f_i(x_0) \in \mathbb{R}$ e quindi per il Teorema dell'asintoto⁴³ $f_i(x_0) = 0$. \blacksquare

(v) Ovviamente lo spazio delle configurazioni U è un insieme (banalmente) invariante. Dalla proprietà di gruppo dei flussi segue che la traccia di un'orbita $\gamma := \{x = \phi^t(x_0) \mid t \in I_{\max}(x_0)\}$ è un insieme invariante. Un esempio più interessante di insieme invariante è l'insieme di tutte le orbite periodiche di un sistema dinamico⁴⁴ $\mathcal{O}_{\mathcal{S}} := \{\text{orbite periodiche di } \mathcal{S}\}$.

(vi) Si noti che, dall'invarianza delle tracce delle orbite (cfr. punto precedente) segue:

le orbite di un flusso non possono 'attraversare' la traccia di un'altra orbita: se γ è la traccia di un'orbita e $\bar{x} \notin \gamma$, allora $\phi^t(\bar{x}) \notin \gamma$, per ogni $t \in I_{\max}(\bar{x})$. In particolare, se la traccia γ di un'orbita contiene un equilibrio x_0 , allora $\gamma = \{x_0\}$.

B.3.4 Analisi qualitativa di equazioni scalari del prim'ordine

Le equazioni differenziali scalari costituiscono senz'altro la classe più semplice di equazioni differenziali: anche nei casi in cui non è possibile trovare una soluzione 'esplicita', l'equazione differenziale dà informazioni puntuali sulla derivata della funzione scalare incognita, il che permette di fare una 'analisi qualitativa' del comportamento della soluzione del problema di

⁴³**Teorema dell'asintoto** Sia $\alpha : t \in (a, +\infty) \rightarrow \alpha(t) \in \mathbb{R}$ una funzione differenziabile tale che esistano i limiti $\lim_{t \rightarrow +\infty} \alpha = M$ e $\lim_{t \rightarrow +\infty} \alpha' = L$ con $M \in \mathbb{R}$ (e $L \in \mathbb{R}^*$). Allora, $L = 0$. Analoga affermazione vale per funzioni $\alpha : (-\infty, a) \rightarrow \mathbb{R}$.

Dimostrazione Sia $a < t_n \nearrow +\infty$. Allora, per il teorema del valor medio di Lagrange (cfr. Proposizione 7.26 in [C2019]), esistono $s_n \in (t_n, t_n + 1)$ tali che $0 = M - M = \lim \alpha(t_n + 1) - \alpha(t_n) = \alpha'(s_n) \rightarrow L$, quindi $L = 0$. \blacksquare

⁴⁴Cfr., ad esempio, le seminali ricerche di H. Poincaré in [Les méthodes nouvelles de la mécanique céleste, Tome 1, Ch. III, Gauthier-Villars (Paris), 1892]

Cauchy.

In particolare, le equazioni scalari autonome del prim'ordine, essendo un caso speciale di equazioni a variabili separabili (cfr. § B.1.3), è esplicitamente risolubile in termine di funzioni integrali, ma ciononostante, l'analisi qualitativa, oltre ad essere più semplice, descrive in modo più immediato il comportamento delle soluzioni.

In questa sezione discuteremo brevemente, le equazioni scalari del prim'ordine, anche con l'intento di illustrare alcune delle nozioni introdotte nelle sezioni precedenti.

(i) L'equazione $\dot{x} = x^2$

Come abbiamo visto (col metodo di separazione delle variabili, cfr. (B.31) con $\bar{t} = 0$), il flusso dell'equazione scalare autonoma con $f : x \in U := \mathbb{R} \rightarrow f(x) := x^2 \in \mathbb{R}$

$$\dot{x} = x^2, \quad x(0) = \bar{x} \tag{B.74}$$

è dato da $t \in I_{\max}(\bar{x}) \rightarrow \phi^t(\bar{x})$ con

$$\phi^t(\bar{x}) = \begin{cases} 0, & \text{se } \bar{x} = 0, \\ \frac{1}{T_e - t}, & \text{se } \bar{x} \neq 0, \end{cases} \quad I_{\max}(\bar{x}) := \begin{cases} \mathbb{R}, & \text{se } \bar{x} = 0, \\ (-\infty, T_e), & \text{se } \bar{x} > 0, \\ (T_e, +\infty), & \text{se } \bar{x} < 0, \end{cases} \tag{B.75}$$

con $T_e := 1/\bar{x}$. Dunque, per ogni $\bar{x} > 0$, $\phi^t(\bar{x})$ è asintotica all'unico equilibrio $x_0 = 0$ per $t \rightarrow -\infty$, mentre per $\bar{x} < 0$ è asintotica a \bar{x} per $t \rightarrow +\infty$. Se $\bar{x} \neq 0$, per $t \rightarrow T_e$, $|\phi^t(\bar{x})| \rightarrow +\infty$, in accordo con la Proposizione B.22-(ii).

(ii) Flussi unidimensionali $\dot{x} = f(x)$

La dinamica dell'equazione differenziale scalare $\dot{x} = f(x)$ è legata agli zeri di f :

Proposizione B.29 Sia $f : (a, b) \subseteq \mathbb{R} \rightarrow \mathbb{R}$ ($-\infty < a < b < +\infty$) localmente lipschitziana.

(i) Siano $x_- < x_+$ due zeri consecutivi di $f : f(x_{\pm}) = 0$ con⁴⁵ $f(x) > 0$ per $x \in I := (x_-, x_+)$.

Sia $\bar{x} \in I$. Allora, $I_{\max}(\bar{x}) = \mathbb{R}$, $\phi^t(\bar{x})$ è strettamente crescente e $\lim_{t \rightarrow \pm\infty} \phi^t(\bar{x}) = x_{\pm}$.

(ii) Sia $f(x_0) = 0$ e $f(x) > 0$ per⁴⁶ $x_0 < x < b$, sia $\bar{x} \in (x_0, b)$. Allora, $I_{\max}(\bar{x}) = (-\infty, T_+)$, $\phi^t(\bar{x})$ è strettamente crescente e $\lim_{t \rightarrow -\infty} \phi^t(\bar{x}) = x_0$, $\lim_{t \rightarrow T_+} \phi^t(\bar{x}) = b$; infine,

$$T_+ < +\infty \iff \int_{\bar{x}}^b \frac{dx}{f(x)} < +\infty. \tag{B.76}$$

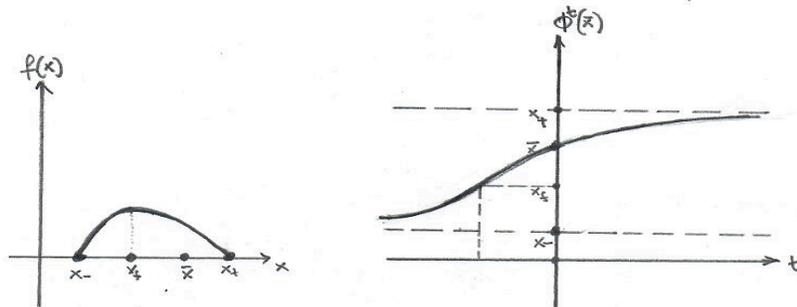


Figura B.7: Campo unidimensionale $f \in C^1$: un'orbita $\phi^t(\bar{x})$ tra due zeri consecutivi

⁴⁵Il caso $f(x) < 0$ si tratta analogamente oppure si deriva da questo sostituendo f con $-f$.
⁴⁶Di nuovo, trattiamo esplicitamente un caso, ma gli altri casi ($f < 0$, o le situazioni simmetriche in $(-\infty, x_0)$ con $f \neq 0$ per $x < 0$) si trattano nella ovvia maniera analoga.

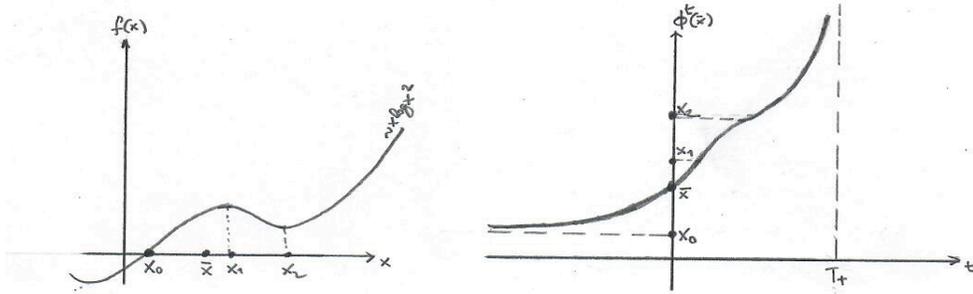


Figura B.8: Campo unidimensionale $f \in C^1$: un'orbita $\phi^t(\bar{x})$ dopo l'ultimo zero

Dimostrazione (i): Sia $(T_-, T_+) := I_{\max}$ e sia $x(t) := \phi^t(\bar{x})$. Poiché $x(t)$ è continua $J := x(I_{\max})$ è un intervallo non banale⁴⁷ che non contiene x_{\pm} (cfr. Osservazione B.28–(vi)). Quindi $J \subseteq I$ (essendo $\bar{x} \in J \cap I$), e $\dot{x}(t) = f(x(t)) > 0$, per ogni $t \in I_{\max}$. Dunque, $x(t)$ è strettamente crescente ed esistono i limiti $\lim_{t \rightarrow T_{\pm}} x(t) =: \hat{x}_{\pm} \in [x_-, x_+]$ e $\lim_{t \rightarrow T_{\pm}} \dot{x}(t) = \lim_{t \rightarrow T_{\pm}} f(x(t))$. Quindi (Osservazione B.28–(iv)) $(T_-, T_+) = \mathbb{R}$ e per il Teorema dell'asintoto (cfr. nota 43), $\lim_{t \rightarrow T_{\pm}} \dot{x} = 0 = f(\hat{x}_{\pm})$ e quindi (non essendoci zeri di f in I) deve essere $\hat{x}_{\pm} = x_{\pm}$. Che $I_{\max} = \mathbb{R}$, nel caso scalare, può anche essere mostrato direttamente come segue. Per separazione di variabili (cfr. § B.1.3) si ha che

$$\int_{\bar{x}}^{x_+} \frac{dx}{f(x)} := \lim_{\tau \rightarrow T_+} \int_{\bar{x}}^{x(\tau)} \frac{dx}{f(x)} = \lim_{\tau \rightarrow T_+} \int_0^{\tau} ds = T_+. \quad (\text{B.77})$$

Essendo f lipschitziana in $[\bar{x}, x_+]$, si ha che, per una opportuna $L > 0$, per ogni $x \in [\bar{x}, x_+]$, $|f(x)| = |f(x) - f(x_+)| \leq L|x - x_+|$ e quindi (si ricordi che $f > 0$ in $[\bar{x}, x_+]$)

$$T_+ = \int_{\bar{x}}^{x_+} \frac{dx}{f(x)} = \int_{\bar{x}}^{x_+} \frac{dx}{|f(x)|} \geq \frac{1}{L} \int_{\bar{x}}^{x_+} \frac{dx}{x_+ - x} = +\infty.$$

(ii): Sia $(T_-, T_+) := I_{\max}$ e sia $x(t) := \phi^t(\bar{x})$. Ragionando come in (i) si ha che $\dot{x} > 0$ su I_{\max} , che $T_- = -\infty$ e che $\lim_{t \rightarrow -\infty} x(t) = x_0$. Dimostriamo, ora, le affermazioni reative a T_+ . Sia $\hat{x} = \lim_{t \rightarrow T_+} x(t)$. Se $T_+ < +\infty$, allora $\hat{x} = b$ per la Proposizione B.22–(ii). Se $T_+ = +\infty$ e, per assurdo, $\hat{x} < b$, allora per il Teorema dell'asintoto (cfr. nota 43) $\dot{x} = f(x) \rightarrow 0$ e quindi $f(\hat{x}) = 0$, il che contraddice l'ipotesi. Dunque, in ogni caso, $\lim_{t \rightarrow T_+} x(t) = b$. Infine, la (B.76) segue per separazione di variabili (essendo T_+ esattamente il valore dell'integrale improprio in (B.76)). ■

Osservazione B.30 (i) Se $f \in C^1$, dall'equazione differenziale, segue che, se $x(t)$ è un'orbita,

$$\ddot{x} = f'(x)\dot{x} = f(x)f'(x), \quad (\text{B.78})$$

relazione che descrive la convessità dell'orbita in termini del campo f . In particolare, se \bar{x} non è un equilibrio, i flessi di $t \rightarrow x(t) = \phi^t(\bar{x})$ corrispondono ai punti critici di f : $\ddot{x}(\tau) = 0$ se e solo se $f'(x(\tau)) = 0$ (cfr. Figura B.8).

(ii) Dalla Proposizione B.29 segue che: gli equilibri corrispondono agli zeri del campo f ; tutte le orbite che iniziano tra due zeri consecutivi sono orbite eterocline; non esistono orbite periodiche o omocline.

(iii) Equazioni non autonome: un esempio

L'analisi qualitativa delle equazioni scalari $\dot{x} = f(x, t)$, con f che dipende esplicitamente dal

⁴⁷Essendo $\dot{x}(0) > 0$.

tempo, è più complessa. Qui ne discuteremo soltanto alcune idee fondamentali tramite un esempio concreto.

Consideriamo il seguente problema di Cauchy

$$\dot{x} = \frac{1}{x - t^2}, \quad x(0) = \bar{x}, \quad \text{con } \bar{x} > 0. \tag{B.79}$$

La funzione $(x, t) \rightarrow 1/(x - t^2)$ ha dominio massimale di esistenza dato da $A_{\max} := \mathbb{R}^2 \setminus S$ dove la ‘curva di singolarità’ S è data dal grafico della parabola $x = t^2$. Su A_{\max} la funzione $1/(x - t^2)$ è C^∞ e dunque la soluzione massimale $x(t) := u_{\max}(t; 0, \bar{x})$ di (B.79) è $C^\infty(I_{\max})$ (Osservazione B.13–(iii)), dove $I_{\max} = I_{\max}(0, \bar{x}) =: (T_-, T_+)$ è l’intervallo massimale di esistenza di (B.79). Il dominio A_{\max} è formato da due componenti connesse⁴⁸ $\Omega_\pm := \{(t, x) \mid \pm(x - t^2) > 0\}$ la cui frontiera comune è la singolarità S . Poiché il dato iniziale $(0, \bar{x}) \in \Omega_+$ si dovrà avere che $(t, x(t)) \in \Omega_+$ per ogni $t \in I_{\max}$. Su Ω_+ , $1/(x - t^2) > 0$ e dunque $\dot{x} > 0$, ossia, $t \in I_{\max} \rightarrow x(t)$ è strettamente crescente. Quindi esistono i limiti laterali $\lim_{t \rightarrow T_\pm} x(t) =: x_\pm \in \mathbb{R}^*$ e, per l’Osservazione B.19–(iii), $(T_\pm, x_\pm) \notin \Omega_+$.

Chiamiamo $\gamma(t) = (t, x(t))$ (per $t \in I_{\max}$) la traiettoria della soluzione massimale nel piano (t, x) . Poiché $\gamma(t) \in \Omega_+$, si ha $x(t) > t^2$ per ogni $t \in I_{\max}$; in particolare per $t < 0$ si ha $0 < t^2 < x(t)$ e, prendendo il limite per $0 > t \rightarrow T_-$, si ha che $0 < T_-^2 \leq x_- < x(0) = \bar{x}$ (si ricordi che $x(t)$ è strettamente decrescente). Quindi, $-\sqrt{\bar{x}} < T_- < 0$, e poiché $(T_-, x_-) \notin \Omega_+$ si deve avere $(T_-, x_-) \in S$, ossia, $x_- = T_-^2$. Dall’equazione differenziale segue poi che $\lim_{t \rightarrow T_-} \dot{x}(t) = +\infty$.

Studiamo, ora, la convessità di $x(t)$. Dall’equazione differenziale segue

$$\ddot{x} = \frac{2t(x - t^2) - 1}{(x - t^2)^3}, \tag{B.80}$$

e quindi studiando il segno del numeratore (il denominatore è sempre positivo su Ω_+) otteniamo che, se poniamo

$$x_f(t) := t^2 + \frac{1}{2t}, \tag{B.81}$$

allora si ha

$$\ddot{x} > 0 \iff x(t) > x_f(t); \quad \ddot{x} < 0 \iff x(t) < x_f(t); \quad \ddot{x} = 0 \iff x(t) = x_f(t).$$

Sia $Q_+ := \{(x, t) \in \Omega_+ \mid x > x_f(t)\}$. Finché $\gamma(t) \in \Omega_+ \setminus \overline{Q_+}$ la soluzione $x(t)$ è strettamente concava ed è concava se $\gamma(t) \in \Omega_+ \setminus Q_+$. Dunque, per $t \in (T_-, 0]$, la soluzione è strettamente concava ed il grafico qualitativo di $x(t)$ sarà come in Figura B.9.

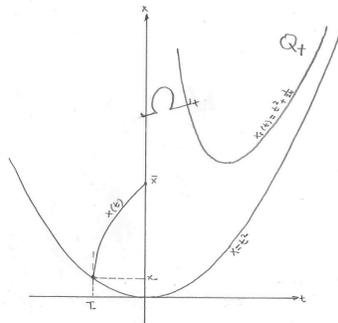


Figura B.9: Una soluzione massimale $x(t)$ di (B.79) per $T \in (T_-, 0)$

⁴⁸In questa discussione, considereremo $\mathbb{R}^2 = \{(t, x) \in \mathbb{R}^2\}$ come piano (t, x) (piuttosto che (x, t)).

Per $t \geq 0$ piccoli, $\gamma(t) \in \Omega_+ \setminus Q_+$ e la soluzione, essendo concava, ha grafico al di sotto di ogni retta tangente al suo grafico⁴⁹; in particolare $\dot{x}(t) \leq \dot{x}(0) = 1/\bar{x}$ per ogni $t > 0$ tale che $\gamma(s) \in \Omega_+ \setminus Q_+$ per ogni $0 \leq s \leq t$. Quindi, se $\gamma(t) \in \Omega_+ \setminus Q_+$ per ogni $t > 0$, $t \in I_{\max}$, si avrebbe $t^2 < x(t) \leq t/\bar{x}$. Questo implica $(T_+, x_+) \in S$ e, quindi, $\lim_{t \rightarrow T_+} \dot{x}(t) = 1/(x(t) - t^2) = +\infty$, il che contraddice $\dot{x}(t) \leq 1/\bar{x}$. Questo significa che esiste un istante $t > 0$ tale che $\gamma(t) \in Q_+$ e, quindi, $t_1 := \inf\{t \mid \gamma(t) \in Q_+\} \in (0, T_+)$. Dalla definizione di t_1 , segue che $\gamma(t_1) \in \partial Q_+$ (ossia, $x(t_1) = x_f(t_1)$) e che esiste $\delta > 0$ tale che $\gamma(t) \in Q_+$ per ogni $t_1 < t < t_1 + \delta$.

Ora, affermiamo che *una volta entrata in Q_+ , la traiettoria $\gamma(t)$ non ne esce più*.

Per dimostrare questa affermazione faremo uso del Teorema di confronto (Corollario B.10), confrontando la soluzione $x(t)$ proprio con ‘la curva dei flessi’ ∂Q_+ che è il grafico della funzione $x_f(t)$. Osserviamo che dalla definizione di $x_f(t)$ ((B.81)) segue che

$$\dot{x}_f(t) = 2t - \frac{1}{2t^2} = \frac{1}{x_f(t) - t^2} - \frac{1}{2t^2}.$$

Dunque possiamo applicare il Corollario B.10 con $\bar{t} = t_1$, $f(x, t) = \frac{1}{x-t^2} - \frac{1}{2t^2}$, $g(x, t) = \frac{1}{x-t^2}$, $u(t) = x_f(t)$ e $v(t) = x(t)$, ottenendo che $x_f(t) < x(t)$ per ogni $t \geq t_1$: quindi, $x(t) \in Q_+$ ed è strettamente convessa per $t > t_1$.

Poiché $(T_+, x_+) \notin \Omega_+$ e $Q_+ \subsetneq \Omega_+$ si deve avere $x_+ = +\infty$ (se fosse $x_+ < +\infty$ si avrebbe $(T_+, x_+) \in S$, il che è impossibile essendo S nel complementare di Q_+).

Infine, $T_+ = +\infty$: se fosse $T_+ < +\infty$, dall’equazione differenziale seguirebbe che $\lim_{t \rightarrow T_+} \dot{x} = 0$, ma $x(t)$ è una funzione convessa con un asintoto verticale in $T_+ \in \mathbb{R}$ e quindi anche la sua derivata deve tendere a⁵⁰ $+\infty$.

Infatti, $x(t)$ è asintotica alla parabola $x = t^2$ nel senso che $x(t) = t^2 + \delta(t)$ con $\delta \rightarrow 0$ per $t \rightarrow +\infty$. Questa affermazione segue dalla convessità di $x(t)$, dall’equazione differenziale e dal fatto che $x(t)$ ha crescita superlineare, essendo $x(t) > t^2$. Infatti, poiché $\ddot{x} > 0$, \dot{x} è strettamente crescente ed esiste $\lim_{t \rightarrow +\infty} \dot{x}(t) = \beta \in \mathbb{R}_+^*$. Se fosse $\beta < +\infty$, si avrebbe, per il Teorema del valor medio di Lagrange, per un qualche $0 < s < t$, $x(t) = \bar{x} + \dot{x}(s)t \leq \beta t$ il che, per t grandi, contraddice $x(t) > t^2$. Quindi, $\beta = +\infty$ e $\dot{x} = 1/(x - t^2) \rightarrow +\infty$ il che equivale a $\delta := x(t) - t^2 \rightarrow 0$. ■

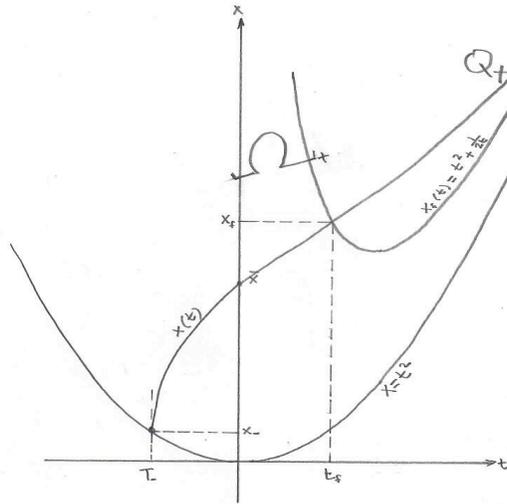


Figura B.10: Una soluzione massimale $x(t)$ di (B.79)

⁴⁹Cfr. [C2019], §7.7.

⁵⁰Poiché x è convessa e regolare \dot{x} è crescente e quindi esiste $\beta = \lim_{t \rightarrow T_+} \dot{x}(t)$ e se fosse $\beta < +\infty$, si avrebbe, per il Teorema del valor medio di Lagrange, che, fissando un qualunque $t_0 < T_+$, per un qualche $t_0 < s < T_+$, $|x(t) - x(t_0)| = |\dot{x}(s)| |t - t_0| \leq \beta(T_+ - t_0)$, ottenendo una contraddizione col fatto che $\lim_{t \rightarrow T_+} x(t) = +\infty$.

B.4 Sistemi lineari

In questa sezione discuteremo sistemi di equazioni differenziali lineari (a coefficienti, in generale, non costanti) su K con $K = \mathbb{R}$ o $K = \mathbb{C}$, ossia, sistemi della forma

$$\dot{x} = A(t)x, \quad (A(t) \in K^{n \times n}, \quad x \in K^n), \quad (\text{B.82})$$

dove gli elementi di matrice $A_{ij} \in C(I, K)$, I intervallo di \mathbb{R} , e $x = x(t) \in K^n$. Nel caso complesso, $x = u + iv$ con $u, v \in \mathbb{R}^n$ e $x \in C(I, K)$ significa $u, v \in C(I, \mathbb{R})$ e la derivata ennesima di x è definita come⁵¹ $x^{(n)} := u^{(n)} + iv^{(n)}$.

Nel caso autonomo (ossia, A indipendente dal tempo) abbiamo visto (Proposizione B.1) che esiste una formula risolutiva, mentre, così non è nel caso generale non autonomo.

Un esempio, il cui studio costituisce un importante capitolo a sé della teoria delle equazioni differenziali è l'equazione (scalare, lineare, omogenea, del secondo ordine) di Sturm–Liouville⁵²

$$(p(t)y')' + q(t)y = \lambda y, \quad (\text{B.83})$$

dove le funzioni $p \in C^1(I, (0, +\infty))$, $q \in C(I, \mathbb{R})$ sono date, $\lambda \in \mathbb{R}$ è un parametro. L'equazione di Sturm–Liouville (B.83) può essere messa nella forma (B.82) ponendo:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y \\ y' \end{pmatrix}, \quad A = \begin{pmatrix} 0 & 1 \\ \frac{\lambda - q}{p} & -\frac{p'}{p} \end{pmatrix}.$$

B.4.1 Spazio delle soluzioni

Sia $K = \mathbb{R}$ o \mathbb{C} , $I \subseteq \mathbb{R}$ un intervallo aperto, $t \in I \mapsto A(t) \in C(I, K^{n \times n})$ e denotiamo

$$\mathcal{S}_A := \{x \in C^1(I, K^n) \mid \dot{x} = A(t)x, \quad \forall t \in I\} \quad (\text{B.84})$$

lo spazio delle soluzioni dell'equazione differenziale lineare $\dot{x} = A(t)x$. Allora, si ha:

Proposizione B.31 (i) \mathcal{S}_A è un sottospazio vettoriale di $C^1(I, K^n)$.

(ii) Se $\bar{t} \in I$ e $\bar{x} \in K^n$, le soluzioni (massimali) del problema di Cauchy $\dot{x} = A(t)x$, $x(\bar{t}) = \bar{x}$ hanno intervallo massimale di esistenza $I_{\max}(\bar{t}, \bar{x}) = I$.

(iii) Siano $x_i \in \mathcal{S}_A$ per $1 \leq i \leq m$ e siano $a_i \in K$. Allora,

$$\exists \bar{t} \in I \mid \sum_{i=1}^m a_i x_i(\bar{t}) = 0 \iff \sum_{i=1}^m a_i x_i(t) = 0, \quad \forall t \in I. \quad (\text{B.85})$$

(iv) \mathcal{S}_A ha dimensione n ed una sua base è data dalle n soluzioni dei problemi di Cauchy $\dot{x} = A(t)x$, $x(\bar{t}) = e^{(i)}$ con $\bar{t} \in I$ arbitrario ($e^{(i)}$ i -esimo versore standard di K^n).

Dimostrazione (i): Se $x, y \in \mathcal{S}_A$ sono soluzioni di (B.82) e $a, b \in K$, chiaramente (per la linearità della derivata e delle matrici identificate con le associate applicazioni lineari⁵³) $ax + by$ è soluzione di (B.82); dunque \mathcal{S}_A è un sottospazio vettoriale di $C^1(I, K^n)$.

(ii) segue immediatamente dal criterio di globalità (Proposizione B.24) con $L = 0$, $M = \max_{[a, b]} \|A(t)\|$.

(iii) L'implicazione ' \Leftarrow ' è ovvia. Dimostriamo ' \Rightarrow '. Assumiamo, per assurdo, che esista $t_0 \neq \bar{t}$ tale che $\sum_{i=1}^m a_i x_i(t_0) = 0$. Allora, per il punto (i), la funzione $x(t) := \sum_{i=1}^m a_i x_i(t) \in \mathcal{S}_A$ e

dunque verifica il problema di Cauchy $\dot{x} = A(t)x$, $x(t_0) = 0$ che ha unica soluzione $x(t) \equiv 0$ e quindi in particolare $x(\bar{t}) = 0$, ottenendo una contraddizione.

Per il punto (ii) le n soluzioni dei problemi di Cauchy $\dot{x} = A(t)x$, $x(\bar{t}) = e^{(i)}$ sono definite su I e per il punto (iii) sono linearmente indipendenti in $C^1(I, \mathbb{R}^n)$. ■

⁵¹Useremo indistintamente le notazioni di Newton e di Lagrange; cfr. nota 14.

⁵²Cfr. Capitoli da 7 a 12 nel già citato [E. Coddington, N. Levinson, "The theory of ordinary differential equations", McGraw–Hill (1955)]; vedi anche https://en.wikipedia.org/wiki/Sturm-Liouville_theory.

⁵³Cfr. Appendice D

Definizione B.32 Una matrice $U(t) := [u^{(1)}, \dots, u^{(n)}]$ non singolare⁵⁴ che soddisfi $U' = AU$ si chiama matrice fondamentale (per il sistema (??)); la funzione $w(t) := \det U$ prende il nome di ‘wronskiano’ (della matrice U o delle soluzioni $u^{(1)}, \dots, u^{(n)}$); una matrice fondamentale per (??) tale che $U(t_0) = I_n$ (dove I_n è la matrice identità) si chiama soluzione fondamentale di (??) (relativa al tempo t_0).

Dunque se U è una matrice fondamentale, la soluzione fondamentale di (??) rispetto al tempo t_0 è data da $V(t) = U(t)U(t_0)^{-1}$ ed in termini della soluzione fondamentale V , la soluzione di (??) con $u(t_0) = u_0$, è data semplicemente da $V(t)u_0$.

Osservazione B.33 Matrici fondamentali, ossia n soluzioni indipendenti di (??), esistono sempre: siano $u^{(j)}(t)$ le soluzioni di

$$u' = A(t)u, \quad u(t_0) = e^{(j)} \quad (\text{B.86})$$

dove $e^{(j)}$ è il versore in \mathbb{R}^n che ha come componenti tutti 0 tranne la j -esima componente che è 1. L'esistenza e l'unicità su tutto I di tali soluzioni è garantito dal Teorema B.5. Allora $V(t) := [u^{(1)}, \dots, u^{(n)}]$ è la soluzione fondamentale di (??).

Osservazione B.34 Dalla Proposizione ?? e dall'osservazione precedente segue che l'insieme di tutte le soluzioni di (??) formano uno spazio vettoriale di dimensione n ed una base di tale spazio vettoriale è data da n soluzioni indipendenti $u^{(1)}, \dots, u^{(n)}$ (per esempio dalle soluzioni di (B.86)).

B.4.2 Il caso non autonomo

L'analisi fatta per i sistemi omogenei può essere usata per discutere i **sistemi lineari non omogenei**

$$u' = A(t)u + b(t) \quad (\text{B.87})$$

con $A(\cdot)$ e $b(\cdot)$ funzioni continue su I a valori, rispettivamente, in $\mathbb{R}^{n \times n}$ ed in \mathbb{R}^n . Sappiamo che, per ogni u_0 esiste una ed una sola soluzione in tutto l'intervallo I di (B.87) con $u(t_0) = u_0$. Per risolvere il problema di Cauchy per (B.87) dobbiamo conoscere la soluzione fondamentale⁵⁵ del problema omogeneo $u' = A(t)u$ ed una qualunque soluzione particolare di (B.87). Supponiamo infatti che V sia la soluzione fondamentale di $u' = Au$ e che $p(t)$ sia una qualunque soluzione di (B.87), allora si verifica immediatamente che

$$u(t) = p(t) + V(t)(u_0 - p(t_0)) \quad (\text{B.88})$$

è la soluzione di (B.87) con $u(t_0) = u_0$.

Nel caso particolare in cui A sia indipendente dal tempo, la soluzione fondamentale (relativa a t_0) è $V(t) = e^{A(t-t_0)}$ e la soluzione di

$$u' = Au + b(t), \quad u(t_0) = u_0$$

è data da

$$u(t) = e^{At} \int_{t_0}^t e^{-A\tau} b(\tau) d\tau + e^{A(t-t_0)} u_0. \quad (\text{B.89})$$

Complementi

Complemento B.1: Altre classi di equazioni differenziali ordinarie risolubili

⁵⁴Cioè con $\det U \neq 0$.

⁵⁵Si ricordi che conoscere la soluzione fondamentale V di (??) è equivalente a conoscere una qualunque matrice fondamentale U essendo queste legate da $V(t) = U(t)U(t_0)^{-1}$.

Equazioni omogenee

Si chiama equazione differenziale omogenea del prim'ordine, un'equazione delle forma

$$x' = f(x, t), \quad (\text{B.90})$$

con f positivamente omogenea di grado 0, ossia, $f(\rho x, \rho t) = f(x, t)$ per ogni $\rho > 0$: in tal caso, per $t > 0$ si ha che

$$f(x, t) = f\left(\frac{1}{t}x, \frac{1}{t}t\right) = g\left(\frac{x}{t}\right), \quad \text{con } g(\xi) := f(\xi, 1).$$

Ponendo, $z(t) := x(t)/t$, si verifica immediatamente che x è soluzione di (B.90) se e solo se z è soluzione di

$$z' = \frac{g(z) - z}{t},$$

che è un'equazione a variabili separabili.

Osservazione B.35 Il motivo legato alla non unicità di questo esempio è che la funzione $x \rightarrow x^{2/3}$ non è sufficientemente regolare in 0. Infatti per avere unicità (almeno locale) in un problema di Cauchy $x' = f(x)$, $x(t_0) = x_0$, è sufficiente che $x \rightarrow f(x)$ sia lipschitziana in un intorno di x_0 .

Equazioni lineari a coefficienti costanti del second'ordine

Infine, consideriamo equazioni differenziali a coefficienti costanti. Sebbene tutti gli argomenti che discuteremo in questa sezione si generalizzino facilmente ad un ordine arbitrario, per concretezza e per il suo particolare interesse, discutiamo solo il caso di ordine 2.

Consideriamo dunque l'equazione differenziale di ordine 2 a coefficienti costanti data da

$$\ddot{x} + a\dot{x} + bx = 0 \quad (\text{B.91})$$

con $a, b \in \mathbb{R}$ e $t \in \mathbb{R} \mapsto x(t) \in \mathbb{R}$ una funzione di classe C^2 . Il relativo problema di Cauchy consiste nel determinare la soluzione di (B.91) avendo assegnati i “dati iniziali” (che in questo caso sono il valore della funzione e della sua derivata in un punto⁵⁶ t_0):

$$\ddot{x} + a\dot{x} + bx = 0, \quad \begin{cases} x(t_0) = x_0, \\ \dot{x}(t_0) = \dot{x}_0, \end{cases} \quad (t, t_0 \in I, t \in I \mapsto x(t) \in C^2(I)), \quad (\text{B.92})$$

dove I è un intervallo.

Osservazione B.36 (i) Se x è soluzione di (B.91) è automaticamente C^∞ ; per esempio, \ddot{x} è derivabile essendo (per (B.91)) uguale a $-a\dot{x} - bx$ e la sua derivata sarà, quindi, data da $-a\ddot{x} - b\dot{x}$ e, iterando, una qualunque derivata di ordine $k \geq 2$ sarà data da $-aD^{k-1}x - bD^{k-2}x$, ($D := d/dt$).

(ii) L'equazione (B.91) è lineare in x e le sue derivate, dunque se x_1 e x_2 sono due soluzioni di (B.91) lo è anche $\alpha x_1 + \beta x_2$ per ogni α, β .

(iii) L'equazione differenziale (B.91) è *autonoma* ossia non dipende esplicitamente dal tempo, o meglio, il tempo appare solo attraverso la funzione x e le sue derivate. Quindi, se $x(t)$ è una soluzione di (B.91) lo è anche $x(t - t_0)$ per ogni t_0 . In particolare, se $x(t)$ è soluzione del problema di Cauchy

$$\ddot{x} + a\dot{x} + bx = 0, \quad \begin{cases} x(0) = x_0, \\ \dot{x}(0) = \dot{x}_0, \end{cases} \quad (\text{B.93})$$

allora $\tilde{x}(t) := x(t - t_0)$ è soluzione del problema di Cauchy (B.92). Dunque basterà considerare il problema di Cauchy (B.93).

Dimostriamo l'unicità del problema di Cauchy (B.93) (da cui segue, per l'osservazione appena fatta, l'unicità del problema di Cauchy (B.92)).

Lemma B.37 *Sia I un intervallo contenente lo 0 e siano x_1 e x_2 due soluzioni di (B.91) per $t \in I$ con $x_1(0) = x_2(0)$, $\dot{x}_1(0) = \dot{x}_2(0)$. Allora, $x_1(t) = x_2(t)$ per ogni $t \in I$.*

⁵⁶Naturalmente per equazioni di ordine n si assegneranno i valori della funzione e delle derivate fino a ordine $(n - 1)$.

Dimostrazione Per linearità, la tesi è equivalente a dimostrare che l'unica soluzione di (B.93) con dati iniziali nulli $x_0 = \dot{x}_0 = 0$ è la soluzione identicamente nulla. Supponiamo dunque, per assurdo, che $t \in I \mapsto x$ sia una soluzione non identicamente nulla di (B.91) con $x(0) = \dot{x}(0) = 0$. Allora, esiste $t_1 \in I$ tale che $x(t_1) \neq 0$. Supponiamo $t_1 > 0$ (l'argomento nel caso $t_1 < 0$ è del tutto analogo). Sia $T = \sup D$ dove $D := \{t \geq 0 \mid x(s) = 0 \forall 0 \leq s \leq t\}$; $0 \in D$ e t_1 è un maggiorante, quindi T è un numero reale maggiore o uguale a 0 e minore di t_1 (dunque $D \subseteq I$). La funzione $y = \dot{x}$ verifica l'equazione lineare del prim'ordine

$$\begin{cases} \dot{y} = -ay - bx(t), & \forall t \in I, \\ y(T) = 0, \end{cases}$$

e dunque, per (B.19), si ha, per ogni $\tau \in I$,

$$\dot{x}(\tau) = -b \int_T^\tau e^{a(s-\tau)} x(s) ds.$$

Se $b = 0$, allora x è costante e quindi $x \equiv 0$ (essendo $x(0) = 0$). Se $b \neq 0$, integrando, tale relazione tra T e $t > T$, otteniamo (essendo $x(T) = 0$)

$$x(t) = -b \int_T^t \left(\int_T^\tau e^{a(s-\tau)} x(s) ds \right) d\tau. \quad (\text{B.94})$$

Sia ora $\varepsilon > 0$ (tale che $T + \varepsilon < t_1$) e denotiamo

$$M_\varepsilon := \sup_{T \leq t \leq T+\varepsilon} |x(t)|.$$

Per definizione di T , $M_\varepsilon > 0$ per ogni $\varepsilon > 0$. Da (B.94) con $t \in [T, T + \varepsilon]$ segue che

$$M_\varepsilon \leq c\varepsilon^2 M_\varepsilon, \quad c = \frac{1}{2} |b| e^{|a|(t_1-T)},$$

che conduce ad una contraddizione se scegliamo $\varepsilon < 1/\sqrt{c}$. ■

Vediamo ora come trovare tutte le soluzioni di (B.91) e la soluzione del problema di Cauchy (B.93). Introduciamo l'“operatore differenziale” $L := D^2 + aD + b$ che agisce sulle funzioni derivabili due volte come

$$Lx = D^2x + aDx + bx = \ddot{x} + a\dot{x} + bx, \quad (\text{B.95})$$

cosicché l'equazione (B.91) si riscrive in forma compatta come $Lx = 0$.

È conveniente considerare in generale soluzioni *complesse* di (B.91), ossia funzioni $z(t) = x(t) + iy(t)$ con $x, y \in \mathbb{R}$. Naturalmente, $\dot{z} = \dot{x} + i\dot{y}$. Quindi, $z(t)$ è soluzione di (B.91) se e solo se lo sono $x(t)$ e $y(t)$.

Poiché la derivata di $e^{\lambda t}$ è $\lambda e^{\lambda t}$ (come è facile verificare anche nel caso⁵⁷ $\lambda \in \mathbb{C}$), è naturale cercare soluzioni di tale forma. Si trova immediatamente

$$Le^{\lambda t} = P(\lambda)e^{\lambda t}, \quad P(\lambda) := \lambda^2 + a\lambda + b \quad (\text{B.96})$$

dove $P(\lambda)$ viene chiamato polinomio caratteristico dell'equazione (B.91). Da (B.96) segue che se λ_0 è una radice del polinomio caratteristico, ossia $P(\lambda_0) = 0$, allora $e^{\lambda_0 t}$ è soluzione di (B.91). Nel caso in esame (equazioni di secondo grado) vi sono dunque tre casi a seconda del segno del discriminante $\Delta := a^2 - 4b$ di $P(\lambda) = 0$:

- (i) $\Delta > 0$: vi sono due soluzioni reali e distinte di $P(\lambda) = 0$ date da $\lambda_\pm := (-a \pm \sqrt{\Delta})/2 \in \mathbb{R}$.
- (ii) $\Delta < 0$: vi sono due soluzioni complesse coniugate e distinte di $P(\lambda) = 0$ date da

$$\lambda_\pm := (-a \pm i\sqrt{-\Delta})/2 \in \mathbb{C}, \quad \bar{\lambda}_+ = \lambda_-.$$

- (iii) $\Delta = 0$: vi è una soluzione reale $\lambda_0 := -a/2$ con molteplicità due, ossia, $P(\lambda) = (\lambda - \lambda_0)^2$.

⁵⁷Si ricordi che dalla formula di Eulero segue che se $\lambda = \alpha + i\beta$, allora $e^{\lambda t} = e^{\alpha t}(\cos \beta t + i \sin \beta t)$.

Nel caso (i) è immediato verificare che la soluzione del problema di Cauchy (B.93) è data da

$$x(t) = A_+ e^{\lambda_+ t} + A_- e^{\lambda_- t}, \quad A_+ := \frac{\dot{x}_0 - \lambda_- x_0}{\lambda_+ - \lambda_-} \quad A_- := \frac{\lambda_+ x_0 - \dot{x}_0}{\lambda_+ - \lambda_-}. \quad (\text{B.97})$$

Anche nel caso (ii) la soluzione del problema di Cauchy (B.93) è data da (B.97) ed osservando che $\bar{A}_+ = A_-$ si ha che $x(t) = 2 \operatorname{Re} (A_+ e^{\lambda_+ t})$, ossia, se chiamiamo $\alpha := -a/2$ e $\omega := \sqrt{-\Delta}/2 > 0$,

$$x(t) = x_0 e^{\alpha t} \cos \omega t + (\dot{x}_0 - \alpha x_0) e^{\alpha t} \operatorname{sen} \omega t, \quad \left(\alpha := -\frac{a}{2}, \quad \omega := \frac{\sqrt{-\Delta}}{2} \right). \quad (\text{B.98})$$

Nel caso (iii), (B.96) diventa

$$L e^{\lambda t} = (\lambda - \lambda_0)^2 e^{\lambda t}, \quad (\text{B.99})$$

per cui $e^{\lambda_0 t} =: e^{\alpha t}$ con $\alpha = -a/2$ è una soluzione di (B.91). Osserviamo, però, che la (B.99) è un'identità che vale per ogni $\lambda \in \mathbb{C}$ e per ogni $t \in \mathbb{R}$; se deriviamo tale relazione rispetto a λ , otteniamo⁵⁸

$$L(t e^{\lambda t}) = (2(\lambda - \lambda_0) + (\lambda - \lambda_0)^2 t) e^{\lambda t}$$

e calcolando tale relazione per $\lambda = \lambda_0$, otteniamo

$$L(t e^{\lambda_0 t}) = 0,$$

che mostra che anche $t e^{\lambda_0 t}$ è soluzione di (B.91). Alternativamente, si può, naturalmente, verificare direttamente che $t e^{\lambda_0 t}$ è soluzione di (B.91). Si verifica, poi, immediatamente che la soluzione di (B.93) è data nel caso (iii) da

$$x(t) = (x_0 + (\dot{x}_0 - \alpha x_0)t) e^{\alpha t}. \quad (\text{B.100})$$

Osservazione B.38 Dall'unicità delle soluzioni del problema di Cauchy segue immediatamente che tutte le soluzioni di (B.91) sono della forma (i), (ii) o (iii) ed in particolare sono definite per tutti i $t \in \mathbb{R}$.

Infatti se $x(t)$ è una qualunque soluzione di (B.91) per $t \in I$ con I intervallo e $t_0 \in I$, possiamo risolvere il problema di Cauchy (B.92) con dati iniziali $x_0 := x(t_0)$ e $\dot{x}_0 := \dot{x}(t_0)$ e se indichiamo con $x(t; x_0, \dot{x}_0)$ tale soluzione, dal Lemma B.37 segue che $x(t) = x(t; x_0, \dot{x}_0)$ per ogni $t \in I$ e quindi $x(t)$ coincide con una delle soluzioni sopra descritte ed inoltre la soluzione è prolungabile per tutti i $t \in \mathbb{R}$ (essendo le soluzioni sopra descritte $C^\infty(\mathbb{R})$).

Come esempio consideriamo l'oscillatore armonico con attrito la cui equazione evolutiva è data da

$$m\ddot{x} + \mu\dot{x} + kx = 0 \quad (\text{B.101})$$

dove $-\mu\dot{x}$ esprime la forza dissipativa di attrito e $\mu > 0$ è una costante fisica legata all'intensità dell'attrito. In questo caso il polinomio caratteristico è dato da $m\lambda^2 \pm \mu\lambda + k = 0$ da cui segue

$$\lambda_{\pm} := \frac{-\mu \pm \sqrt{\mu^2 - 4mk}}{2}.$$

Quindi a seconda della relazione tra μ^2 e $4mk$ si avranno tutti e tre i tipi di soluzione sopra descritti con la caratteristica comune di avere un fattore esponenziale $e^{-\mu t/2}$, la qual cosa implica che tutte le soluzioni avranno limite 0 per $t \rightarrow +\infty$. Questo sistema, al contrario dei modelli conservativi di Newton ($\mu = 0$), ha una dinamica irreversibile.

Per attrito debole (ossia $\mu^2 < 4mk$), si avranno infinite oscillazioni di ampiezza che decade esponenzialmente, mentre per attrito forte ($\mu^2 > 4mk$) non si avrà alcuna oscillazione ("la molla non ha forza sufficiente") ed infine si ha la soluzione speciale nel caso $\mu^2 = 4mk$ con la correzione lineare in t .

⁵⁸Non è difficile verificare che l'ordine delle derivate (rispetto a t e rispetto a λ) può essere scambiato.

Equazioni lineari a coefficienti costanti

Si consideri l'equazione differenziale scalare di ordine n a coefficienti costanti

$$Lx := x^{(n)} + a_{n-1}x^{(n-1)} + \cdots + a_1x' + a_0x = 0 ; \quad (\text{B.102})$$

con $a_i \in \mathbb{R}$ (o anche⁵⁹ $a_i \in \mathbb{C}$).

(i) Se

$$P(\lambda) := \lambda^n + a_{n-1}\lambda^{n-1} + \cdots + a_1\lambda + a_0, \quad (\text{B.103})$$

è facile vedere che $P(\lambda) = \det(\lambda - A)$ dove A è la matrice definita come

$$A := \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \vdots & \vdots & & \ddots & \\ 0 & 0 & 0 & \cdots & 1 \\ -a_0 & -a_1 & -a_2 & \cdots & -a_{n-1} \end{pmatrix}. \quad (\text{B.104})$$

$P(\lambda)$ prende il nome di **polinomio caratteristico** dell'equazione (B.102).

(ii) Poiché $\frac{d^j}{dt^j} e^{\lambda t} = \lambda^j e^{\lambda t}$ si ha che, se λ_0 è una radice (in generale complessa) di $P = 0$ (ossia se $P(\lambda_0) = 0$) allora $e^{\lambda_0 t}$ è soluzione di (B.102). Infatti $L(e^{\lambda_0 t}) = P(\lambda_0)e^{\lambda_0 t} = 0$.

(iii) Si dice che λ_0 è una radice di ordine m di $P = 0$ se esiste un polinomio (a coefficienti complessi) $g(\lambda)$ tale che $P(\lambda) = (\lambda - \lambda_0)^m g(\lambda)$ e $g(\lambda_0) \neq 0$.

Allora, se λ_0 è una radice di ordine m di $P = 0$, $t^j e^{\lambda_0 t}$ è soluzione di (B.102) per ogni $0 \leq j \leq m$. Infatti, se $z(t; \lambda) := e^{\lambda t}$ allora

$$Lz = z P(\lambda) = z (\lambda - \lambda_0)^m g(\lambda),$$

e l'asserto si ottiene derivando rispetto a λ (e calcolando in λ_0) tale relazione.

Da (ii), (iii) e dal Teorema fondamentale dell'algebra⁶⁰ segue che *vi sono n soluzioni distinte della forma $t^j e^{\lambda t}$ di (B.102) dove se λ è una radice semplice di $P(\lambda) = 0$ allora $j = 0$, mentre se λ è una radice di ordine $m \leq n$, allora $0 \leq j \leq m$.*

Se λ è una radice di ordine m di $P = 0$ e se i coefficienti a_i sono reali allora anche $\bar{\lambda}$ è una radice di ordine m . Dunque⁶¹ $L(t^j e^{\lambda t}) = 0$ ossia anche $t^j e^{\bar{\lambda} t}$ è soluzione di (B.102). Poiché l'equazione (B.102) è lineare si che, se $z := t^j e^{\lambda t}$, anche $(z + \bar{z})/2 := \text{Re}(z)$ e $(z - \bar{z})/(2i) := \text{Im}(z)$ sono soluzioni. In altri termini, se $\lambda := \alpha + i\beta$ (con $\alpha, \beta \in \mathbb{R}$) è una radice di ordine m di $P = 0$, sono soluzioni di (B.102) le $2m$ funzioni reali $t^j e^{\alpha t} \cos(\beta t)$ e $t^j e^{\alpha t} \sin(\beta t)$ per $0 \leq j \leq m$.

(iv) Poiché le soluzioni $t^j e^{\lambda t}$ (al variare di j e λ) sono tra loro indipendenti, da § B.4 segue che *tutte le soluzioni di (B.102) sono delle combinazioni lineari delle funzioni $t^j e^{\lambda t}$ (ossia, nel caso $a_i \in \mathbb{R}$, sono delle combinazioni lineari di $t^j e^{\alpha t} \cos(\beta t)$ e $t^j e^{\alpha t} \sin(\beta t)$).*

Sistemi lineari a coefficienti costanti non omogenei

Si dimostra che se A è una matrice (indipendente dal tempo), $b \in C(I, \mathbb{R}^n)$, con I intervallo in \mathbb{R} , allora il vettore $u(t)$ definito come

$$u(t) = e^{At} \int_{t_0}^t e^{-A\tau} b(\tau) d\tau \quad (\text{B.105})$$

soddisfa

$$u' = Au + b, \quad u(t_0) = 0. \quad (\text{B.106})$$

Discutere casi speciali $\ddot{x} + \beta\dot{x} + x = f(t)$

Complemento B.2: Sistemi meccanici unidimensionali. Spazio delle fasi

⁵⁹Si ricordi la nota ??.

⁶⁰Se $P(z)$ è un polinomio complesso di ordine n , l'equazione $P(z) = 0$ ha n soluzioni in \mathbb{C} (contando le molteplicità).

⁶¹Si ricordi che $(e^\alpha) = e^\alpha$.

Uno degli esempi più importanti di equazione differenziale è fornito dalla fisica ed è l'equazione unidimensionale di Newton:

$$f = ma \quad (\text{B.107})$$

dove $m > 0$ denota la massa di un punto materiale vincolato (senza attrito) a muoversi su una linea (o più in generale su una curva), a denota la sua accelerazione e f la forza esterna che agisce sul punto. Se $x \in J$ denota la coordinata del punto⁶² e se la forza è conservativa (ossia, dipende solo dalla coordinata x), l'equazione di Newton (B.107) si scrive più esplicitamente come

$$m\ddot{x} = f(x), \quad (\text{B.108})$$

dove $t \in I \mapsto x(t) \in J$ è la legge oraria (incognita) del punto di classe $C^2(I)$ e $f \in C(J)$ e, seguendo la notazione di Newton, la derivata rispetto a t è denotata con il punto ($\dot{x} = x' = \frac{d}{dt}x$, $\ddot{x} = x''$, etc.).

Questa è una equazione dinamica “conservativa”, ossia, vi è una quantità che verrà chiamata *energia del sistema* che si conserva nel tempo. Moltiplicando, infatti, i membri della equazione (B.108) per la velocità \dot{x} , otteniamo

$$m\ddot{x} \cdot \dot{x} = f(x)\dot{x}. \quad (\text{B.109})$$

Ora, se

$$F(x) := \int_{x_0}^x f(\xi) d\xi \quad (\text{B.110})$$

è una funzione integrale di f (fissato un qualunque $x_0 \in J$), vediamo che (B.109) si può riscrivere come

$$\frac{d}{dt} \frac{m}{2} \dot{x}^2 = \frac{d}{dt} F(x) \quad (\text{B.111})$$

ovvero, come

$$\dot{E} = 0, \quad E := \frac{m}{2} \dot{x}^2 + V(x), \quad V(x) := -F(x), \quad (\text{B.112})$$

dove E denota l'*energia totale del sistema* che si esprime come somma dell'*energia cinetica* $T := \frac{m}{2} \dot{x}^2$ e dell'*energia potenziale* V (che, per definizione, verifica⁶³ $V' = -f$). Dunque essendo la derivata rispetto a t dell'energia identicamente uguale a 0 sull'intervallo I segue che l'*energia totale* E si conserva lungo i moti del sistema⁶⁴:

$$E(t) \equiv E_0 := \frac{1}{2} m \dot{x}_0^2 + V(x_0), \quad \forall t \in I, \quad \text{con} \quad \begin{cases} \dot{x}_0 := \dot{x}(t_0) \\ x_0 := x(t_0) \end{cases} \quad (\text{B.113})$$

(e $t_0 \in I$ denota l'istante “iniziale”). Tale equazione può essere riscritta come

$$\dot{x} = \sigma \sqrt{\frac{2}{m} (E_0 - V(x))}, \quad (\sigma = \text{sgn}(\dot{x})). \quad (\text{B.114})$$

Si noti che, per definizione, $E_0 \geq V(x)$ e $E_0 = V(x)$ se e solo se $\dot{x} = 0$.

Vi sono due casi: o $E_0 = V(x_0)$ e, quindi $\dot{x}_0 = 0$, nel qual caso $x(t) \equiv x_0$ è soluzione di (B.108), oppure $E_0 > V(x_0)$, nel qual caso $\dot{x}_0 \neq 0$ e scegliamo (ai fini di una determinazione della soluzione locale) $\sigma := \text{sgn}(\dot{x}_0)$. In questo secondo caso l'equazione (B.114) è una equazione differenziale del prim'ordine a variabili separabili (§ B.1.3) e dunque l'*unica soluzione locale* di (B.108) con dati iniziali $x(t_0) = x_0$ e $\dot{x}(t_0) = \dot{x}_0$ è data da

$$x(t) = \Phi^{-1}(\sigma(t - t_0)) \quad (\text{B.115})$$

dove Φ^{-1} è la funzione inversa di

$$\Phi(x) := \int_{x_0}^x \frac{d\xi}{\sqrt{\frac{2}{m} (E_0 - V(\xi))}}. \quad (\text{B.116})$$

Osservazione B.39 (i) Dunque, il problema di Cauchy per le equazioni di Newton (o in generale per una equazione differenziale del second'ordine) consiste nel determinare la soluzione dell'equazione differenziale avendo assegnato i dati iniziali $x(t_0)$ e $\dot{x}(t_0)$. Il metodo sopra discusso, mostra che (almeno localmente) è sempre possibile risolvere il problema di Cauchy per le equazioni di Newton della dinamica nel caso unidimensionale (ossia, $x \in \mathbb{R}$ o su una curva). In questo senso si dice che i *sistemi meccanici unidimensionali sono “integrabili”*.

⁶²In generale J sarà un intervallo di \mathbb{R} , ma potrebbe, come detto, essere una curva (per esempio il cerchio unitario S^1).

⁶³La derivata rispetto a x continuiamo a denotarla con l'apice.

⁶⁴I “moti” del sistema non son altro che soluzioni dell'equazione (B.108).

(ii) Come nella discussione generale delle equazioni a variabili separabili, nel caso in cui $E_0 > V(x_0)$, l'unicità delle soluzioni è implicita nel metodo (poiché la soluzione viene esplicitamente determinata); andrebbe, invece, discussa a parte nel caso in cui $E_0 = V(x_0)$. Si noti che in questo secondo caso x_0 è un massimo locale di V e quindi $f(x_0) = -V'(x_0) = 0$ (se x_0 è interno a I) ed, essendo $x(t) \equiv x_0$ una soluzione, diremo che x_0 è un *punto d'equilibrio* del sistema o anche un *punto stazionario*.

(iii) Un metodo utile di analizzare le soluzioni è il *ritratto di fase*: da (B.114) segue che il moto ha luogo sulle curve di energia costante nel piano (*spazio delle fasi*) (x, y) dove x è la posizione al tempo t del punto e $y := m\dot{x}$ è il *momento lineare* che è proporzionale alla velocità \dot{x} al tempo t .

Vediamone un semplice esempio concreto.

L'oscillatore armonico

Un punto di massa $m > 0$ vincolato (senza attrito) su una retta e soggetto alla forza di una molla ideale che lo attrae verso l'origine $x = 0$ della retta subisce, secondo Hook⁶⁵, una forza pari a $-kx$ dove $k > 0$ è la *costante di elasticità della molla*, x la posizione del punto materiale e il segno meno è dovuto al fatto che la molla attrae verso l'origine. Dunque, secondo Newton e Hook, la legge del moto di tale punto è regolata dall'*equazione differenziale*

$$m\ddot{x} = -kx \quad (\text{B.117})$$

che possiamo riscrivere come

$$\ddot{x} + \omega^2 x = 0, \quad (\omega := \sqrt{k/m}). \quad (\text{B.118})$$

Si noti che l'equazione (B.117) è *lineare* e quindi se x_1 e x_2 sono soluzioni di (B.117) lo è anche una qualunque combinazione lineare $\alpha x_1 + \beta x_2$.

Il potenziale del sistema è dato da $V(x) := \frac{1}{2}kx^2$ per cui l'energia totale (che si conserva lungo i moti) sarà data da

$$E(t) := m\frac{\dot{x}^2}{2} + k\frac{x^2}{2}. \quad (\text{B.119})$$

In questo caso (grazie alla linearità) possiamo immediatamente concludere che *la soluzione del problema di Cauchy di (B.117), ossia*⁶⁶,

$$m\ddot{x} = -kx, \quad \begin{cases} x(t_0) = x_0 \\ \dot{x}(t_0) = \dot{x}_0 \end{cases} \quad (\text{B.120})$$

è *unica*.

Infatti, se x_1 e x_2 verificano entrambe (B.120) si ha che $z := x_1 - x_2$ verifica (B.117) e $z(t_0) = \dot{z}(t_0) = 0$, il che implica che la sua energia

$$m\frac{\dot{z}^2}{2} + k\frac{z^2}{2} = E(0) = 0$$

e dunque, necessariamente, $z(t) \equiv 0$ e $\dot{z}(t) \equiv 0$, ossia $x_1 = x_2$.

Per determinare esplicitamente le soluzioni di (B.120) potremmo naturalmente usare il metodo descritto sopra (che sarebbe comunque un utile esercizio sull'integrazione che invitiamo a fare), ma è assai più semplice osservare che ci sono due ovvie funzioni "indipendenti" che verificano $\ddot{x} = -\omega^2 x$ e sono $\sin \omega t$ e $\cos \omega t$ e poiché l'equazione è lineare ci aspettiamo che *tutte le soluzioni possano scriversi come una combinazione lineare di queste due funzioni*. Infatti, è immediato verificare che

$$x(t) = x_0 \cos \omega t + \frac{\dot{x}_0}{\omega} \sin \omega t, \quad (\text{B.121})$$

risolve il problema di Cauchy (B.120) e ne è dunque l'unica soluzione. Si noti che tali soluzioni sono definite per tutti i tempi $t \in \mathbb{R}$.

Anche il ritratto di fase è particolarmente semplice in questo caso. Infatti, nello spazio delle fasi $(x, y) := (x, m\dot{x})$, l'equazione di energia costante è

$$k\frac{x^2}{2} + \frac{y^2}{2m} = E_0 := k\frac{x_0^2}{2} + \frac{m\dot{x}_0^2}{2}, \quad (\text{B.122})$$

⁶⁵Robert Hooke (Freshwater, 18 luglio 1635 – Londra, 3 marzo 1703).

⁶⁶ \dot{x}_0 è un numero assegnato in \mathbb{R} e il punto qui non ha nessun significato matematico.

e l'origine $(0, 0)$ corrisponde all'unico equilibrio del sistema (“punto fermo nell'origine con molla a riposo”) di energia $E_0 = 0$ e se $E_0 > 0$, allora la (B.122) si può scrivere come

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 = 1, \quad \text{con} \quad a := \sqrt{2E_0/k}, \quad b := \sqrt{2E_0m}, \quad (\text{B.123})$$

che è l'equazione di una ellissi di semiassi a e b .

Esercizio B.1 Dimostrare che nello spazio delle fasi dell'oscillatore armonico, durante i moti, le ellissi (B.123) vengono percorse in senso antiorario.

Complemento B.3: Dipendenza C^k da dati iniziali e parametri

Valgono i seguenti risultati sulla dipendenza regolare dai dati iniziali delle soluzioni di equazioni differenziali. Usando le notazioni del Teorema B.5 si ha:

(i) Se f e f_x sono continue su $D \times I$ allora la soluzione $u(t) := \varphi(t; x)$ di (B.39) con $u_0 := x$ è di classe $C^1(\{x, u_0\})$.

(ii) Se tutte le derivate di f rispetto a x_j di ordine $k \geq 1$ sono continue su $D \times I$ allora la soluzione $u(t) := \varphi(t; x)$ di (B.39) con $u_0 := x$ è di classe $C^k(\{x, u_0\})$.

Complemento B.4: Stabilità (cenni)

Consideriamo il sistema (omogeneo) di equazioni differenziali lineari a coefficienti costanti

$$u' = Au \quad (\text{B.124})$$

con $A \in \mathbb{R}^{n \times n}$ matrice indipendente dal tempo.

Definizione B.40 Un punto $x_0 \in \mathbb{R}^n$ si dice punto di equilibrio per (B.124), se la funzione costante $u(t) := x_0$ è soluzione di (B.124).

Quindi x_0 è un punto di equilibrio per (B.124) se e solo se $Ax_0 = 0$; dunque 0 è sempre un punto d'equilibrio per (B.124) e $x_0 \neq 0$ è un equilibrio se e solo se x_0 è un autovettore di A con autovalore nullo. Il problema che ora ci poniamo è come evolvono dati iniziali vicini ad un punto di equilibrio, ossia come si comportano, per $t \gg 1$, le soluzioni $e^{At}x$ di (B.124), quando il dato iniziale x è vicino ad x_0 .

Definizione B.41 Un punto di equilibrio x_0 per (B.124) si dice stabile (secondo Liapunov) se $\forall \varepsilon > 0, \exists \delta > 0$ tale che $|e^{At}x - x_0| \leq \varepsilon$ per ogni $t > 0$ e per ogni $|x - x_0| \leq \delta$. Un punto di equilibrio stabile x_0 si dice asintoticamente stabile se $\exists \delta > 0$ tale che $\lim_{t \rightarrow \infty} e^{At}x = x_0$, per ogni $|x - x_0| \leq \delta$. Un punto di equilibrio asintoticamente stabile si chiama anche ‘attrattore’ per (B.124).

Se x_0 è un autovettore con autovalore nullo, allora tutta la retta $\ell_{x_0} := \{sx_0 : s \in \mathbb{R}\}$ è formata da punti di equilibrio essendo $A(sx_0) = sAx_0 = 0$. Quindi se esiste un autovettore x_0 con autovalore nullo, non ci può essere alcun punto di equilibrio asintoticamente stabile.

Se x_0 è un punto di equilibrio asintoticamente stabile, necessariamente $x_0 = 0$ e A non ha alcun autovettore con autovalore nullo. Inoltre se 0 è asintoticamente stabile, allora $\lim_{t \rightarrow \infty} e^{At}x = 0$, per ogni $x \in \mathbb{R}^n$. Sia infatti $\delta > 0$ quello della definizione di stabilità asintotica, e sia $x \neq 0$ un qualunque vettore di \mathbb{R}^n , allora se poniamo $a = \frac{\delta}{2|x|}$, il vettore ax ha norma uguale a $\frac{\delta}{2}$ ed è quindi all'interno della sfera di raggio δ centrata nell'origine. Allora, $|e^{At}x| = a^{-1}|e^{At}(ax)| \rightarrow 0$, se $t \rightarrow \infty$.

Si noti che, facendo un eventuale cambio di coordinate ($y = x - x_0$), possiamo sempre ridurci al caso in cui il punto di equilibrio sia l'origine.

Se A possiede un autovettore con autovalore avente parte reale strettamente positiva, allora 0 è un punto di equilibrio instabile: infatti è facile vedere che se $Av = \lambda v$, allora $e^{At}v = e^{\lambda t}v$. Dunque se $Av = \lambda v$ e $\text{Re } \lambda > 0$, allora per ogni $\delta > 0$, possiamo trovare, all'interno della sfera di raggio δ attorno all'origine, un vettore w tale che $\lim_{t \rightarrow \infty} |e^{At}w| = \infty$: basta prendere $w = av$ con $a = \frac{\delta}{2|v|}$. In tal caso $|e^{At}w| = a|e^{At}v| = a|e^{\lambda t}v| = a|v|e^{t \text{Re } \lambda} \rightarrow \infty$ quando $t \rightarrow \infty$. E questo dimostra che l'origine è instabile.

Concludiamo questa discussione con un criterio sufficiente affinché 0 sia un punto di equilibrio stabile o asintoticamente stabile.

Proposizione B.42 *Supponiamo che A sia diagonalizzabile e che tutti gli autovalori abbiano parte reale minore o uguale a 0. Allora l'origine $x = 0$ è un punto di equilibrio stabile. Se A è diagonalizzabile e tutti gli autovalori hanno parte reale strettamente negativa, allora $x = 0$ è asintoticamente stabile.*

Dimostrazione Sappiamo che se $T \in \mathbb{C}^{n \times n}$ è una matrice invertibile, allora $T^{-1}e^{AT} = e^{T^{-1}AT}$; inoltre se $D = \text{diag}(d_1, \dots, d_n)$ è la matrice diagonale avente sulla diagonale i numeri complessi d_1, \dots, d_n , allora $\|D\| = \max |d_i|$. Sia, dunque, T la matrice che diagonalizza A : $T^{-1}AT = \Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$. Dalle osservazioni appena fatte segue che

$$\begin{aligned} \|e^{At}v\| &= \|TT^{-1}e^{At}TT^{-1}v\| = \|Te^{T^{-1}AT}T^{-1}v\| = \|Te^{\Lambda t}T^{-1}v\| \\ &\leq \|T\| \|e^{\Lambda t}\| \|T^{-1}\| \|v\| = \|T\| \|T^{-1}\| \|v\| \max_{1 \leq i \leq n} e^{(\text{Re } \lambda_i)t}. \end{aligned}$$

Da tale stima segue subito l'asserto. \blacksquare

In effetti, utilizzando la forma canonica di Jordan, si dimostra in maniera del tutto analoga la stessa affermazione senza l'ipotesi di diagonalizzabilità per A .

Complemento B.5: Equazioni delle onde e del calore (alle derivate parziali)

Corda vibrante con estremi fissi (equazione delle onde)

L'evoluzione dello spostamento verticale di una corda elastica (di lunghezza, a riposo, unitaria) con estremi fissi, che oscilla (senza attrito e senza forze esterne) in un piano, soddisfa, in prima approssimazione, la seguente *equazione differenziale alle derivate parziali*

$$u_{tt} - c^2 u_{xx} = 0, \quad x \in (0, 1), \quad t > 0, \quad (\text{B.125})$$

$$u(0, t) = u(1, t) = 0, \quad \forall t > 0, \quad (\text{B.126})$$

dove $u = u(x, t)$ denota la posizione, all'istante t (in un piano cartesiano $x, y = u$), del punto della corda che a riposo coincide con il punto $(x, y) = (x, 0)$ e c è una costante che dipende dalle caratteristiche fisiche della corda.

Il problema (classico) di Cauchy per (B.125), (B.126) consiste nel trovare

$$u \in C^2((0, 1) \times (0, \infty)), \quad \text{con } u, u_t \in C([0, 1] \times [0, \infty)),$$

che soddisfici (B.125), (B.126) e le seguenti *condizioni iniziali*

$$u(x, 0) = u_0(x), \quad u_t(x, 0) = v_0(x), \quad (\text{B.127})$$

dove u_0 e v_0 sono funzioni assegnate (sufficientemente regolari e tali che $u_0(0) = u_0(1) = v_0(0) = v_0(1) = 0$). Le condizioni 'al contorno' (B.126) si chiamano **condizioni di Dirichlet**.

Definiamo il seguente spazio di funzioni

$$C_{\text{dis}}^k((0, 1)) := \{f \in C^k([0, 1]) : f^{(i)}(0) = f^{(i)}(1) = 0 \forall i \text{ pari e } i \leq k\}. \quad (\text{B.128})$$

Si risolva il problema di Cauchy per (B.125), (B.126) supponendo che i dati iniziali u_0 e v_0 siano di classe $C_{\text{dis}}^\infty((0, 1))$.

Per una funzione $u(x, t)$ di classe C^1 si definisce **l'energia al tempo t** la quantità

$$E(t) := \frac{1}{2} \int_0^1 (u_t^2(x, t) + c^2 u_x^2(x, t)) dx. \quad (\text{B.129})$$

(i) **(Conservazione dell'energia per la corda elastica)** Si dimostri che se

$$u \in C^2((0, 1) \times (0, \infty)) \cap C([0, 1] \times [0, \infty))$$

è soluzione di (B.126) allora

$$\frac{dE}{dt} = 0, \quad \forall t > 0 \quad (\text{B.130})$$

e quindi che $E(t) := E(0)$.

(ii) **(Unicità)** Si usi il punto (i) per dimostrare che la soluzione del problema di Cauchy per (B.125), (B.126) con dati iniziali in C_{dis}^∞ è *unica*.

Equazione del calore

La temperatura di un filo metallico di lunghezza unitaria con gli estremi tenuti a temperatura costante, che (per semplicità) supporremo uguale a 0, soddisfa la seguente equazione alle derivate parziali

$$u_t - k u_{xx} = 0, \quad x \in (0, 1), \quad t > 0, \quad (\text{B.131})$$

$$u(0, t) = u(1, t) = 0, \quad \forall t > 0, \quad (\text{B.132})$$

dove $u(x, t)$ denota la temperatura al tempo t del punto del filo di coordinata uguale ad x e $k > 0$ è la cosiddetta ‘conduttività’.

Il problema (classico) di Cauchy per (B.131), (B.132) consiste nel trovare u tale che

$$u \in C([0, 1] \times [0, \infty)), \quad \text{con } u_t, u_x, u_{xx} \in C((0, 1) \times (0, \infty)),$$

tale che soddisfi (B.131), (B.132) e la *condizione iniziale*

$$u(x, 0) = u_0(x), \quad (\text{B.133})$$

dove u_0 è una funzione assegnata (sufficientemente regolare e tale che $u_0(0) = u_0(1) = 0$). Le condizioni ‘al contorno’ (B.132) si chiamano **condizioni di Dirichlet**.

Si risolva il problema di Cauchy per (B.131), (B.132) supponendo che il dato iniziale u_0 appartenga a $C_{\text{dis}}^1((0, 1))$.

Esercizi

Esercizio B.2 Dimostrare che se $t \in (a, b) \mapsto x(t)$ è una soluzione differenziabile di $x' = x^{2/3}$ e $x(t_0) \neq 0$ per un $t_0 \in (a, b)$, allora esiste un intorno I di t_0 e un polinomio di grado tre, $P(t) = \frac{t^3}{27} + \dots$, tale che $x(t) = P(t)$ per ogni $t \in I$.

Esercizio B.3 Si dimostri l’affermazione fatta nell’Osservazione ??.

Esercizio B.4 Dimostrare che se $f \in C^1(D, \mathbb{R}^m)$ con D sfera chiusa in \mathbb{R}^n , allora f è uniformemente lipschitziana in D e come costante di Lipschitz può prendersi

$$L = \sup_{x \in D} \|f_x(x)\|. \quad (\text{B.134})$$

Esercizio B.5 (i) Sia $\omega \neq 1$. Si trovi una soluzione dell’equazione differenziale (oscillatore armonico forzato)

$$\ddot{x} + x = \text{sen } \omega t.$$

(ii) Si risolva il seguente problema di Cauchy

$$\ddot{x} + x = \text{sen } \omega t, \quad x(0) = 2, \quad \dot{x}(0) = 0.$$

Esercizio B.6 Si trovi una soluzione dell’equazione differenziale (oscillatore armonico forzato risonante)

$$\ddot{x} + x = \text{sen } t.$$

(ii) Si risolva il seguente problema di Cauchy

$$\ddot{x} + x = \text{sen } t, \quad x(0) = 2, \quad \dot{x}(0) = 0.$$

Esercizio B.7 Perché la (B.105) non fornisce, in generale, una soluzione nel caso $A = A(t)$?

Esercizio B.8 (i) Sia $A := \begin{pmatrix} 0 & x \\ y & 0 \end{pmatrix}$ e si calcoli e^A .

(ii) Si verifichi che se $G(t) = \int_0^t \begin{pmatrix} 0 & \text{sen } \tau \\ \tau & 0 \end{pmatrix} d\tau$, allora $GG' \neq G'G$ e $G'e^G \neq (e^G)' \neq e^G G'$.

Esercizio B.9 Sia $t \in I \rightarrow A(t)$ una funzione continua dall'intervallo I a valori matrici ($n \times n$) e sia $G(t) := \int_{t_0}^t A(\tau) d\tau$ con $t_0, t \in I$. Assumendo che G e A commutino, cioè che $AG = GA$, si dimostri che l'unica soluzione $u \in C^1(I)$ di

$$u' = A(t)u + b(t), \quad u(t_0) = u_0 \quad (\text{B.135})$$

dove $b : I \rightarrow \mathbb{R}^n$ è una funzione continua assegnata è data da

$$u(t) = e^{G(t)} \left(\int_{t_0}^t e^{-G(\tau)} b(\tau) d\tau + u_0 \right). \quad (\text{B.136})$$

Esercizio B.10 Si consideri l'equazione differenziale

$$u'' + a(t)u' + b(t)u = f(t), \quad t \in I, \quad (\text{B.137})$$

dove I è un intervallo aperto di \mathbb{R} e a, b e f sono funzioni continue su I .

(i) Si trovi una soluzione di (B.137) assumendo che si conoscano le soluzioni dell'equazione omogenea associata

$$u'' + a(t)u' + b(t)u = 0. \quad (\text{B.138})$$

(ii) Si dimostri che tutte le soluzioni di (B.137) hanno la forma

$$u = c_1 u_1 + c_2 u_2 + p \quad (\text{B.139})$$

dove u_1 e u_2 sono soluzioni indipendenti di (B.138) e $p(t)$ è una qualunque soluzione di (B.137).

(iii) Si scrivano esplicitamente (in termini di f) tutte le soluzioni di (B.137) nel caso in cui i coefficienti $a(t)$ e $b(t)$ siano costanti: $a(t) := a$ e $b(t) := b$.

Esercizio B.11 Sia $x(t)$ la soluzione del seguente problema di Cauchy

$$\ddot{x} + \dot{x} + x = f(t), \quad x(0) = 1, \quad \dot{x}(0) = -1, \quad (\text{B.140})$$

dove f è una funzione continua su \mathbb{R} con $\int_{\mathbb{R}} |f| dt < \infty$ e si trovi, se esiste, il $\lim_{t \rightarrow \infty} x(t)$.

Esercizio B.12 Si consideri l'equazione differenziale

$$\ddot{x} + \dot{x}^2 + \sinh x = 0, \quad \dot{x}(0) = 0, \quad x(0) = \alpha. \quad (\text{B.141})$$

(i) Si trovi $T > 0$ tale che esista un'unica soluzione di (B.141) per $|t| \leq T$.

(ii) Si denoti con $x(t; \alpha)$ la soluzione di (B.141) per $|t| \leq T$ e si calcoli il limite $\lim_{\alpha \rightarrow 0} x(t; \alpha)$ per ogni $|t| \leq T$.

(iii) Il limite definito al punto (ii) è uniforme in $t \in [-T, T]$?

Esercizio B.13 Sia $y(x)$ la soluzione di $(1-x)y' = 1+x-y$, $y(0) = 0$. Usando l'equazione differenziale si calcolino tutte le derivate di $y(x)$ in $x=0$; determinare il raggio di convergenza della serie di Taylor di $y(x)$ e si concluda che in un intorno di $x=0$ la soluzione $y(x)$ è analitica.

Esercizio B.14 Sia $T > 0$ e sia $f \in C^1(A \times \mathbb{R}, \mathbb{R}^n)$, con A aperto di \mathbb{R}^n e f T -periodica: $f(x, t+T) = f(x, t)$ per ogni $(x, t) \in A \times \mathbb{R}$. Dimostrare che se $x(t)$ è soluzione di $\dot{x} = f(x, t)$, $x(0) = x_0$ e se $x(T) = x_0$ allora x è T -periodica ossia $x(t+T) = x(t)$ per ogni t .

Esercizio B.15 Si risolva il problema di Cauchy per (B.125), (B.126) supponendo che i dati iniziali u_0 e v_0 siano:

(i) $u_0 = \sin^3 x$, $v_0 = 0$.

(ii) $u_0 = 0$, $v_0 = \sin^p x$ con $p = 1, 3, 4$.

Esercizio B.16 Discutere il problema di Cauchy per (B.125), (B.126) con $u_0 = \sin^2 x$ e $v_0 = 0$.

Esercizio B.17 Si risolva il problema di Cauchy per (B.131), (B.132) con dato iniziale $u_0 = \sin^p x$ con $p = 1, 2, 3, 4$.

Appendice C

Varietà immerse in \mathbb{R}^n

C.1 Definizioni

Chiameremo **dominio** (o dominio k -dimensionale) un insieme compatto $D = \bar{U}$ di \mathbb{R}^k che sia la chiusura di un insieme $U \subseteq \mathbb{R}^k$ aperto e connesso e tale che $\partial D = \partial U$ sia un insieme trascurabile¹. Per esempio, la sfera unitaria chiusa

$$\bar{B}_1^k := \{u \in \mathbb{R}^k : |u| \leq 1\} \quad (\text{C.1})$$

è un dominio in \mathbb{R}^k (per $k = 1$ è semplicemente l'intervallo chiuso $[-1, 1]$). Sostituendo la norma euclidea $|\cdot|$ con la norma uniforme $|\cdot|_\infty$, si ottiene il dominio dato dal cubo di lato 2 centrato nell'origine; se $k = 2$ (o $k \geq 2$) tali insiemi hanno una ovvia differenza dal punto di vista geometrico: la frontiera del quadrato presenta delle 'singolarità' (i vertici del quadrato) che, invece, la frontiera di \bar{B}_1^2 (e cioè la circonferenza unitaria S^1) non presenta. Nei seguenti paragrafi daremo un significato preciso a queste osservazioni intuitive e studieremo gli elementi del calcolo integrale (di funzioni regolari o regolari a tratti) su tali insiemi.

Definiamo ora una 'inclusione differenziabile'² di un dominio k -dimensionale nello 'spazio ambiente' \mathbb{R}^n , con $n \geq k$:

Definizione C.1 *Sia D un dominio k -dimensionale. Una funzione $\varphi \in C^1(D, \mathbb{R}^n)$ con $n \geq k$ si chiama **inclusione differenziabile** (o semplicemente 'inclusione') di D in \mathbb{R}^n se φ è iniettiva su D e la matrice jacobiana $\frac{\partial \varphi}{\partial u}$ ha rango massimo su ogni punto $u \in D$.*

Una coppia (φ, D) con φ inclusione differenziabile del dominio $D \subseteq \mathbb{R}^k$ verrà anche chiamata 'inclusione differenziabile' o anche 'inclusione k -dimensionale'.

L'immagine³ $\varphi(D)$ di una inclusione è un oggetto geometrico k -dimensionale (essendo descritto dai k parametri $u = (u_1, \dots, u_k) \in D$) immerso in \mathbb{R}^n e che ha le stesse 'proprietà geometriche' del dominio D . In particolare, ad ogni punto $x \in \varphi(D)$ corrisponde un unico $u \in D$ (essendo φ iniettiva): cioè la φ definisce un sistema di coordinate su $\varphi(D)$. Inoltre tale sistema di coordinate sarà 'regolare' nel senso che la corrispondenza $u \in D \rightarrow x \in \varphi(D)$ è descritta da una funzione C^1 e con jacobiano di rango massimo.

Definizione C.2 *Se φ è una inclusione del dominio $D := \bar{U} \subseteq \mathbb{R}^k$ in \mathbb{R}^n (U aperto, connesso e limitato), l'insieme $\mathcal{S} := \varphi(U)$ si chiama **elemento (regolare) di varietà k -dimensionale** (o 'elemento di k -varietà').*

¹Ovvero D è misurabile secondo Peano–Jordan.

²La terminologia non è universale e dipende anche dalla lingua che si sta usando: altri termini che possono indicare oggetti simili (se non lo stesso oggetto) sono 'immersione' oppure (in inglese) 'immersion', 'imbedding', 'embedding' etc. Si faccia dunque attenzione alle definizioni date di volta in volta nei vari testi.

³Anche: 'traccia' o 'supporto'.

Osservazione C.3 (i) Si noti che il ‘bordo’ di \mathcal{S} , $\varphi(\partial U)$, non appartiene a \mathcal{S} .

(ii) Dalla Proposizione 1.30 segue che l’immagine $\varphi(D)$ di una inclusione è un sottoinsieme compatto di \mathbb{R}^n . Inoltre⁴, la funzione inversa $\varphi^{-1} : \varphi(D) \rightarrow D$ è continua su $\varphi(D)$. In particolare φ^{-1} è continua sull’elemento di varietà $\mathcal{S} := \varphi(U)$.

(iii) Per ogni punto $x_0 \in \mathcal{S} = \varphi(U)$ esiste un (unico) $u_0 \in U$ tale che $\varphi(u_0) = x_0$. Poiché U è aperto esiste una sfera aperta B di \mathbb{R}^k interamente contenuta in U e poiché $g = \varphi^{-1}$ è continua su \mathcal{S} , l’insieme $A := \varphi(B) = g^{-1}(B)$ è un insieme aperto in \mathcal{S} (nella topologia relativa⁵) quindi $x_0 \in A = E \cap \mathcal{S}$ per qualche insieme E aperto di \mathbb{R}^n .

(iv) Dati due spazi metrici X ed Y , due sottoinsiemi $A \subseteq X$, $B \subseteq Y$ si dicono **omeomorfi** se esiste una funzione continua $g : A \rightarrow B$ con inversa continua su B ; la funzione biunivoca e bicontinua g si chiama un **omeomorfismo** di A su B . Quindi il punto (iii) può essere riformulato dicendo che: per ogni punto $x_0 \in \mathcal{S}$ esiste un aperto $E \subseteq \mathbb{R}^n$ tale che $E \cap \mathcal{S}$ è omeomorfo ad una sfera B in \mathbb{R}^k .

(v) Dire che la matrice $\frac{\partial \varphi}{\partial u}$ ha rango massimo (e cioè k) equivale a dire che i k vettori $\xi^{(i)} := \frac{\partial \varphi}{\partial u_i} \in \mathbb{R}^n$ sono linearmente indipendenti in \mathbb{R}^n .

(vi) La parola ‘regolare’, che normalmente sottoindenderemo, si riferisce al fatto che l’elemento di varietà è realizzato tramite una funzione di classe C^1 .

(vii) **(Cambiamenti di coordinate)** Siano $\varphi : D = \bar{U} \subseteq \mathbb{R}^k \rightarrow \mathbb{R}^n$ e $\psi : E = \bar{V} \subseteq \mathbb{R}^k \rightarrow \mathbb{R}^n$ due inclusioni differenzibili tali che $\varphi(U) = \psi(V) := \mathcal{S}$. Su \mathcal{S} sono definite e continue [punto (ii)] le funzioni inverse φ^{-1} e ψ^{-1} . Sono dunque definite e continue le seguenti funzioni

$$v := \psi^{-1} \circ \varphi : u \in U \rightarrow v(u) \in V, \quad u := \varphi^{-1} \circ \psi : v \in V \rightarrow u(v) \in U. \quad (\text{C.2})$$

Naturalmente $v(u)$ e $u(v)$ sono iniettive e sono l’una l’inversa dell’altra. In effetti, come ci accingiamo a dimostrare, tali funzioni sono dei **diffeomorfismi** (o, più precisamente dei ‘diffeomorfismi di classe C^1 ’) ossia sono delle funzioni C^1 invertibili con inversa⁶ C^1 . Fissiamo $v_0 \in V$ e sia $u_0 := u(v_0)$ (cosicché $v_0 = v(u_0)$). Poiché $\frac{\partial \psi}{\partial v}(v_0)$ ha rango k esistono $1 \leq i_1 < \dots < i_k \leq n$ tali che, ponendo $\hat{\psi} := (\psi_{i_1}, \dots, \psi_{i_k})$, si ha $\det \frac{\partial \hat{\psi}}{\partial v}(v_0) \neq 0$. Definiamo $F(v, u) := \hat{\psi}(v) - \hat{\varphi}(u)$, con $\hat{\varphi} := (\varphi_{i_1}, \dots, \varphi_{i_k})$. F è di classe $C^1(\{(v_0, u_0)\})$, $\det \frac{\partial F}{\partial v}(v_0, u_0) = \det \frac{\partial \hat{\psi}}{\partial v}(v_0) \neq 0$ e $F(v_0, u_0) = 0$. Dunque dal Teorema delle funzioni implicite segue che esiste un’unica funzione continua \tilde{v} tale che $\tilde{v}(u_0) = v_0$ e $F(\tilde{v}(u), u) = 0$ in un intorno di u_0 ed in più tale funzione è di classe $C^1(\{u_0\})$. Poiché anche $v(u)$ soddisfa $v(u_0) = v_0$ e $F(v(u), u) = 0$ (in un intorno di u_0) per l’unicità della funzione implicita segue che $\tilde{v}(u)$ coincide con $v(u)$ in un intorno di u_0 e questo, in particolare, significa che $v \in C^1(\{u_0\})$. Dal Corollario 3.6 segue che anche lo jacobiano di $v(u)$, $\frac{\partial v}{\partial u}(u_0)$ è invertibile ed ha per inverso lo jacobiano $\frac{\partial u}{\partial v}(v_0)$.

(viii) Sia $u_0 \in \mathbb{R}^k$ e $\varphi \in C^1(\{u_0\}, \mathbb{R}^n)$ con $n \geq k$. Se rango $\frac{\partial \varphi}{\partial u}(u_0) = k$ allora esiste una sfera chiusa $D \subseteq \mathbb{R}^k$ di centro u_0 tale che (φ, D) è un’inclusione differenziabile in \mathbb{R}^n .

Infatti dalle ipotesi segue che esiste un intorno u_0 su cui φ è C^1 e su cui $\det \frac{\partial \hat{\varphi}}{\partial u}(u_0) \neq 0$ con $\hat{\varphi} := (\varphi_{i_1}, \dots, \varphi_{i_k})$ (per opportuni indici $1 \leq i_1 < \dots < i_k \leq n$). Dal Teorema della funzione inversa segue che $u \rightarrow \hat{\varphi}$ è invertibile su di una opportuna sfera chiusa centrata in u_0 . Su tale sfera $\hat{\varphi}$ è iniettiva e quindi lo è anche φ . ■

Diamo ora alcuni esempi di inclusioni (e quindi di elementi di varietà).

(E1) Un insieme $U \subseteq \mathbb{R}^n$ aperto, connesso e limitato è un elemento di n -varietà in \mathbb{R}^n : come inclusione si può prendere l’identità $x = \varphi(u) = u$. Analogamente, se U è un sottoinsieme aperto, connesso e limitato di \mathbb{R}^k e $n > k$, l’insieme $\mathcal{S} := \{x \in \mathbb{R}^n : (x_1, \dots, x_k) \in U, x_{k+1} = \dots = x_n = 0\}$ è un elemento di k -varietà in \mathbb{R}^n e come inclusione si può prendere la mappa $u \in U \rightarrow x = \varphi(u) := (u_1, \dots, u_k, 0, \dots, 0)$.

⁴ D un sottoinsieme compatto di \mathbb{R}^n e $f \in C(\mathbb{R}^n, \mathbb{R}^m)$. Se f è iniettiva, allora la funzione inversa, f^{-1} , appartiene a $C(K, D)$ dove K è l’insieme compatto $f(D)$.

⁵Si ricordi la Proposizione 1.24.

⁶Si ricordi il Corollario 3.6.

(E2) La funzione $t \in [0, 1] \rightarrow z(t) = x + t(y - x) \in \mathbb{R}^n$ è una inclusione dell'intervallo $[0, 1]$ in \mathbb{R}^n che ha come immagine il segmento $P(x, y): z([0, 1]) = P(x, y)$.

(E3) (Caso $k = 1$: elementi di curve regolari in \mathbb{R}^n) Generalizzando l'esempio precedente, definiamo *elemento di curva regolare in \mathbb{R}^n* un elemento di 1-varietà ossia un sottoinsieme di \mathbb{R}^n dato da $\varphi(I)$ con $I = (a, b)$ intervallo aperto e φ una inclusione del dominio unidimensionale $[a, b]$ in \mathbb{R}^n . In questo caso, la matrice jacobiana di φ è semplicemente il vettore⁷ $\varphi'(t) = (\varphi'_1(t), \dots, \varphi'_n(t))$. Dire che tale matrice jacobiana ha rango massimo (e cioè 1) equivale a dire che il vettore $\varphi'(t)$ non si annulla mai:

$$|\varphi'(t)| := \sqrt{\sum_{i=1}^n |\varphi'_i(t)|^2} \neq 0, \quad \forall t \in [a, b]. \tag{C.3}$$

L'iniettività di φ su \bar{I} implica che l'elemento di curva $\varphi(I)$ non si autointerseca mai e che $\varphi(a) \neq \varphi(b)$. Un esempio è dato da 'un elemento di circonferenza' cioè dall'elemento di curva regolare $\varphi(t) = (\cos t, \sin t)$ con $t \in [a, b]$ e $0 < b - a < 2\pi$: φ è iniettiva su $[a, b]$ e $|\varphi'| = |(-\sin t, \cos t)| = 1$. Quindi l'elemento di circonferenza $\varphi((a, b))$ è un elemento di curva regolare sulla circonferenza unitaria

$$S^1 := \{x \in \mathbb{R}^2 : |x| = \sqrt{x_1^2 + x_2^2} = 1\}. \tag{C.4}$$

(E4) (Il caso $k = 2, n = 3$: elementi di superficie in \mathbb{R}^3) Un elemento di 2-varietà in \mathbb{R}^3 (o anche in \mathbb{R}^n) prende il nome di *elemento di superficie*. Se $D = \bar{U}$ è un dominio in \mathbb{R}^2 e $\varphi : (u_1, u_2) \in D \rightarrow (\varphi_1, \varphi_2, \varphi_3) \in \mathbb{R}^3$ è una inclusione di D , dire che la matrice jacobiana $\frac{\partial \varphi}{\partial u}$ ha rango massimo equivale a dire che i due vettori $\frac{\partial \varphi}{\partial u_1}$ e $\frac{\partial \varphi}{\partial u_2}$ sono indipendenti e cioè⁸

$$\left| \frac{\partial \varphi}{\partial u_1}(u) \times \frac{\partial \varphi}{\partial u_2}(u) \right| \neq 0, \quad \forall u \in \bar{U}. \tag{C.5}$$

Un esempio è dato da una 'porzione ellissoidale' realizzata dall'inclusione

$$\varphi(u_1, u_2) = (r_1 \cos u_1 \sin u_2, r_2 \sin u_1 \sin u_2, r_3 \cos u_2)$$

con $D = [a_1, b_1] \times [a_2, b_2]$, $0 < b_1 - a_1 < 2\pi$ e $0 < a_2 < b_2 < \pi$. Allora $\mathcal{S} = \varphi(U)$, ($U := \overset{\circ}{D}$), è un elemento di varietà bidimensionale in \mathbb{R}^3 che descrive una *calotta ellissoidale* ritagliata sull'ellissoide

$$\mathcal{E} := \left\{ x \in \mathbb{R}^3 : \left(\frac{x_1}{r_1}\right)^2 + \left(\frac{x_2}{r_2}\right)^2 + \left(\frac{x_3}{r_3}\right)^2 = 1 \right\}. \tag{C.6}$$

Per $r_1 = r_2 = r_3$ si ha naturalmente un elemento della sfera tridimensionale di raggio $r := r_1$.

(E5) (Grafici in \mathbb{R}^n) Sia $D = \bar{U}$ un dominio in \mathbb{R}^k , sia $f \in C^1(D, \mathbb{R})$ e sia $n := k + 1$. Il grafico

$$G_f := \{x \in \mathbb{R}^n = \mathbb{R}^{k+1} : x_n = f(x_1, \dots, x_k), \forall (x_1, \dots, x_k) \in U\} \tag{C.7}$$

è un elemento di $k = n - 1$ varietà in \mathbb{R}^n realizzato dalla inclusione $\varphi(u) := (u, f(u))$ con $u \in U$. Infatti, la φ è chiaramente iniettiva su D e

$$\frac{\partial \varphi}{\partial u} = \left(\begin{array}{cccc} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & 1 \\ f_{u_1} & f_{u_2} & \dots & f_{u_k} \end{array} \right) \Bigg\} k + 1 \tag{C.8}$$

⁷Di solito, nel caso unidimensionale, denotiamo con I l'aperto connesso e limitato di \mathbb{R} (e cioè l'intervallo) che sopra abbiamo chiamato U e indichiamo con t il punto generico di I (che sopra abbiamo denotato u).

⁸Per informazioni sul prodotto vettoriale 'x' in \mathbb{R}^3 si veda 4.1.

e quindi $\frac{\partial \varphi}{\partial u}$ ha rango uguale a $k = n - 1$ (per ogni $u \in D$).

(E6) (i) Sia $\phi \in C(\mathbb{R}^n, \mathbb{R})$. Per la Proposizione 1.24, l'insieme $A := \{x \in \mathbb{R}^n : \phi(x) < 0\}$ è un insieme aperto; quindi se A è connesso e limitato allora (esempio **(E1)**) è un elemento di n -varietà in \mathbb{R}^n .

(ii) Sia $\phi \in C^1(\mathbb{R}^n, \mathbb{R})$ e sia $E := \{x \in \mathbb{R}^n : \phi(x) = 0\}$ e supponiamo che per un $\bar{x} \in E$ si abbia $\nabla \phi(\bar{x}) \neq 0$ ossia, per un qualche j , $\frac{\partial \phi}{\partial x_j}(\bar{x}) \neq 0$. Allora, per il Teorema delle funzioni implicite (Teorema 3.10) esiste un rettangolo in \mathbb{R}^n ,

$$K_{r,\rho} := \{x \in \mathbb{R}^n : |x_i - \bar{x}_i| \leq r \text{ per } i \neq j \text{ e } |x_j - \bar{x}_j| \leq \rho\}$$

ed una funzione⁹ $g(x_1, \dots, \hat{x}_j, \dots, x_n)$ di classe $C^1(\hat{K}_r, [\bar{x}_j - \rho, \bar{x}_j + \rho])$, dove \hat{K}_r è il cubo in \mathbb{R}^{n-1} di centro $(\bar{x}_1, \dots, \hat{x}_j, \dots, \bar{x}_n)$ e lato $2r$, tale che $E \cap K_{r,\rho}$ coincide con il grafico di g ossia con $\{x \in \mathbb{R}^n : x_j = g(x_1, \dots, \hat{x}_j, \dots, x_n), \forall (x_1, \dots, \hat{x}_j, \dots, x_n) \in \hat{K}_r\}$. Ma questo, per **(E5)**, significa che l'insieme $\mathcal{S} := \{\phi(x) = 0\} \cap K_{r,\rho}^{\circ}$ è un elemento di $(n-1)$ -varietà. Il dominio dell'inclusione è $\hat{K}_r \subseteq \mathbb{R}^{n-1}$ e l'inclusione sarà data da $u \in \hat{K}_r \rightarrow \varphi(u) := (u_1, \dots, u_{j-1}, g(u), u_j, \dots, u_{n-1}) \in K_{r,\rho}$ (avendo posto $x_i = u_i$ per $i \leq j-1$ e $u_i = x_{i+1}$ per $i \geq j$).

(E7) L'esempio precedente (punto (ii)) si generalizza come segue. Supponiamo che siano date $(n-k)$ funzioni $\phi_j \in C^1(\mathbb{R}^n, \mathbb{R})$ e consideriamo l'insieme $E := \{x \in \mathbb{R}^n : \phi_1(x) = \dots = \phi_{n-k}(x) = 0\}$. Assumiamo che per un qualche punto $\bar{x} \in E$, gli $(n-k)$ vettori $\nabla \phi_1(\bar{x}), \dots, \nabla \phi_{n-k}(\bar{x})$ siano linearmente indipendenti (in \mathbb{R}^n). Allora esiste un aperto $A \subseteq \mathbb{R}^n$ tale che $A \cap E$ è un elemento di k -varietà in¹⁰ \mathbb{R}^n .

Concludiamo questo paragrafo introducendo le nozioni di varietà e varietà regolare a tratti:

Definizione C.4 (i) Un insieme connesso $\mathcal{S} \subseteq \mathbb{R}^n$ prende il nome di 'varietà k -dimensionale' se per ogni $x_0 \in \mathcal{S}$ esiste un aperto A di \mathbb{R}^n che contiene x_0 e tale che $A \cap \mathcal{S}$ è un elemento di k -varietà.

(ii) Un insieme connesso $\mathcal{S} \subseteq \mathbb{R}^n$ prende il nome di 'varietà k -dimensionale regolare a tratti' se $\mathcal{S} = \bigcup_{i=1}^N \varphi^{(i)}(E^{(i)})$ dove $(\varphi^{(i)}, E^{(i)})$ godono, per ogni $1 \leq i \leq N$, delle seguenti proprietà: (a) gli $E^{(i)}$ sono insiemi di \mathbb{R}^k con frontiera trascurabile; (b) $\varphi^{(i)} \in C^1(E^{(i)}, \mathbb{R}^n) \cap C(\overline{E^{(i)}}, \mathbb{R}^n)$ e per ogni k -dominio $D^{(i)} \subseteq E^{(i)}$ la coppia $(\varphi^{(i)}, D^{(i)})$ è un'inclusione differenziabile in \mathbb{R}^n ; (c) $\varphi^{(i)}(E^{(i)}) \cap \varphi^{(j)}(E^{(j)}) = \emptyset$ per ogni $i \neq j$.

(iii) Una coppia (φ, E) con φ ed E che soddisfino le proprietà (a) e (b) del punto precedente¹¹ prende il nome di **pseudo-inclusione** o più precisamente di 'pseudo-inclusione differenziabile, k -dimensionale in \mathbb{R}^n '.

Osservazione C.5 (i) Di particolare importanza sono le **varietà compatte**¹²: per esempio le 'superfici sferiche'

$$S^{n-1} := \{x \in \mathbb{R}^n : |x| = 1\}, \quad (\text{C.9})$$

sono delle varietà compatte di dimensione $n-1$.

(ii) Un esempio tipico di varietà regolare a tratti in \mathbb{R}^n è dato dalla frontiera del cubo $[0, 1]^n$ che è una ' $(n-1)$ -varietà chiusa con singolarità' nei vertici del cubo $[0, 1]^n$.

(iii) Una curva regolare a tratti è l'immagine $\varphi(I)$ con I intervallo limitato di \mathbb{R} e con φ continua ed iniettiva su \bar{I} ad eccezione di un numero finito di punti e C^1 a tratti; se $\varphi(a) = \varphi(b)$ la curva (regolare a tratti) si dirà **chiusa**.

⁹ Il simbolo ' $\hat{}$ ' in questo contesto significa che la quantità che vi sta sotto va omessa: $(x_1, \dots, \hat{x}_j, \dots, x_n)$ coincide con $(x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_n)$ se $1 < j < n$, con (x_2, \dots, x_n) se $j = 1$ e con (x_1, \dots, x_{n-1}) se $j = n$.

¹⁰ L'ipotesi che i vettori $\nabla \phi_1, \dots, \nabla \phi_{n-k}$ siano indipendenti è equivalente a dire che la matrice $(n-k) \times n$, $\frac{\partial(\phi_1, \dots, \phi_{n-k})}{\partial(x_{k+1}, \dots, x_n)}$ ha rango massimo e cioè uguale a $n-k$. Si assuma, per fissare le idee, che il minore $\frac{\partial(\phi_1, \dots, \phi_{n-k})}{\partial(x_{k+1}, \dots, x_n)}$ abbia determinante diverso da 0 in \bar{x} . Si definisca $F : (u, y) \in \mathbb{R}^k \times \mathbb{R}^{n-k} \rightarrow F(u, y) = (\phi_1(u, y), \dots, \phi_{n-k}(u, y)) \in \mathbb{R}^{n-k}$ e si noti che $\det \frac{\partial F}{\partial y}(\bar{u}, \bar{y}) \neq 0$, dove $(\bar{u}, \bar{y}) = \bar{x}$. Si applichi ora il Teorema delle funzioni implicite.

¹¹ Con $E^{(i)} = E$ e $\varphi^{(i)} = \varphi$.

¹² A volte le varietà compatte vengono chiamate 'varietà chiuse'.

(iv) Secondo la Definizione C.4 una varietà compatta può essere vista come una varietà regolare a tratti: ad esempio, la circonferenza $\{x \in \mathbb{R}^n : x_1^2 + x_2^2 = 1, x_3 = \dots = x_n = 0\}$ è una 1–varietà regolare a tratti in \mathbb{R}^n poiché $\Gamma = \varphi(E)$ con $E := [0, 2\pi]$ e $\varphi := (\cos t, \sin t, 0, \dots, 0)$.

C.2 Misure e integrali su k –varietà

C.2.1 Definizioni e osservazioni

Sia (φ, E) una pseudo–inclusione differenziabile k –dimensionale in \mathbb{R}^n . Vogliamo definire la ‘misura k –dimensionale’ dell’insieme $\mathcal{S} := \varphi(E)$; la nozione di misura che introdurremo può esser visto come ‘la proiezione ortogonale della misura euclidea n dimensionale sull’oggetto geometrico k –dimensionale rappresentato da \mathcal{S} ’.

Dalla definizione di inclusione segue che almeno uno dei minori $\frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(u_1, \dots, u_k)}(u)$ (dove (i_1, \dots, i_k) per $1 \leq i_1 < \dots < i_k \leq n$ è una qualche scelta di indici) ha determinante diverso da zero; in particolare la somma su tutte le possibili scelte dei minori di rango k dei quadrati dei determinanti di tali minori, ossia il numero

$$\sum_{1 \leq i_1 < \dots < i_k \leq n} \left| \det \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(u_1, \dots, u_k)} \right|^2, \quad (\text{C.10})$$

è diversa da zero su $U := \overset{\circ}{E}$. Definiamo la **misura k –dimensionale di \mathcal{S}** il numero

$$\sigma_k(\mathcal{S}) := \int_U \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \left| \det \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(u_1, \dots, u_k)} \right|^2 \right)^{\frac{1}{2}} du, \quad (\text{C.11})$$

e se $f \in C(\overline{\mathcal{S}}, \mathbb{R})$ definiamo l’**integrale di f su \mathcal{S}** come

$$\int_{\mathcal{S}} f d\sigma_k := \int_U f \circ \varphi(u) \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \left| \det \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(u_1, \dots, u_k)} \right|^2 \right)^{\frac{1}{2}} du. \quad (\text{C.12})$$

Per verificare che tali definizioni sono ben poste osserviamo quanto segue. Innanzitutto, poiché ∂E è un insieme trascurabile, si ha che gli integrali su U possono essere sostituiti da integrali su E senza che il loro valore cambi. Dimostriamo ora che gli integrali in (C.11) e (C.12) non dipendono dalla particolare inclusione che ‘realizza’ \mathcal{S} : sia (ψ, E') un’altra pseudo–inclusione tale che, se $V := \overset{\circ}{E}'$, $\psi(V) = \varphi(U)$. Per il punto (vii) dell’Osservazione C.3 la funzione $v(u) := \psi^{-1} \circ \varphi(u)$ è una funzione di classe $C^1(U, V)$ ed è tale che $\psi(v(u)) = \varphi(u)$, per ogni $u \in U$. Dunque, per ogni scelta di indici $i_1 < \dots < i_k$ si ha $(\psi_{i_1}(v(u)), \dots, \psi_{i_k}(v(u))) = (\varphi_{i_1}(u), \dots, \varphi_{i_k}(u))$ e prendendo lo jacobiano di tale relazione si ottiene

$$\frac{\partial(\psi_{i_1}, \dots, \psi_{i_k})}{\partial(v_1, \dots, v_k)}(v(u)) \frac{\partial v}{\partial u}(u) = \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(u_1, \dots, u_k)},$$

e quindi

$$\det \frac{\partial(\psi_{i_1}, \dots, \psi_{i_k})}{\partial(v_1, \dots, v_k)} \cdot \det \frac{\partial v}{\partial u}(u) = \det \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(u_1, \dots, u_k)}. \quad (\text{C.13})$$

Dunque, tramite il cambio di variabile $v = v(u)$, dal Teorema 4.6, dal fatto che $\psi(v(u)) = \varphi(u)$ e da (C.13) segue che

$$\begin{aligned} & \int_V f \circ \psi(v) \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \left| \det \frac{\partial(\psi_{i_1}, \dots, \psi_{i_k})}{\partial(v_1, \dots, v_k)} \right|^2 \right)^{\frac{1}{2}} dv \\ &= \int_U f \circ \varphi(u) \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \left| \det \frac{\partial(\psi_{i_1}, \dots, \psi_{i_k})}{\partial(v_1, \dots, v_k)}(v(u)) \right|^2 \right)^{\frac{1}{2}} \left| \det \frac{\partial v}{\partial u} \right| du \\ &= \int_U f \circ \varphi(u) \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \left| \det \frac{\partial(\psi_{i_1}, \dots, \psi_{i_k})}{\partial(v_1, \dots, v_k)}(v(u)) \cdot \det \frac{\partial v}{\partial u} \right|^2 \right)^{\frac{1}{2}} du \\ &= \int_U f \circ \varphi(u) \left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \left| \det \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(u_1, \dots, u_k)} \right|^2 \right)^{\frac{1}{2}} du. \quad \blacksquare \end{aligned}$$

Queste definizioni si estendono immediatamente a varietà regolari a tratti: se $\mathcal{S} = \bigcup_{i=1}^N \varphi^{(i)}(E^{(i)}) \subseteq \mathbb{R}^n$ è una varietà k dimensionale regolare a tratti (con $\varphi^{(i)}$ e $E^{(i)}$ come nella Definizione C.4) e se $f \in C(\bar{\mathcal{S}}, \mathbb{R})$ poniamo¹³

$$\sigma_k(\mathcal{S}) := \sum_{i=1}^N \sigma_k(\mathcal{S}^{(i)}), \quad \int_{\mathcal{S}} f d\sigma_k := \sum_{i=1}^N \int_{\mathcal{S}^{(i)}} f d\sigma_k, \quad (\mathcal{S}^{(i)} := \varphi(E^{(i)})). \quad (\text{C.14})$$

Esempio C.6 Applichiamo le definizioni date al calcolo dell'area superficiale di una sfera di raggio r in \mathbb{R}^3 . La sfera $S_r^2 := \{x \in \mathbb{R}^3 : |x| = r\}$ coincide con $\varphi(E)$ se

$$E = [0, 2\pi] \times [0, \pi] \quad \text{e} \quad \varphi(u_1, u_2) := r(\cos u_1 \sin u_2, \sin u_1 \sin u_2, \cos u_2).$$

Dunque

$$\begin{aligned} \sigma_2(S_r^2) &:= \int_E \left(\sum_{1 \leq i < j \leq 2} \left| \det \frac{\partial(\varphi_i, \varphi_j)}{\partial u} \right|^2 \right)^{\frac{1}{2}} du \\ &:= \int_E \left(\left| \det \frac{\partial(\varphi_1, \varphi_2)}{\partial u} \right|^2 + \left| \det \frac{\partial(\varphi_1, \varphi_3)}{\partial u} \right|^2 + \left| \det \frac{\partial(\varphi_2, \varphi_3)}{\partial u} \right|^2 \right)^{\frac{1}{2}} du \\ &= r^2 \int_E \left((\sin u_2 \cos u_2)^2 + (\sin u_1 \sin^2 u_2)^2 + (\cos u_1 \sin^2 u_2)^2 \right)^{\frac{1}{2}} du \\ &= r^2 \int_E \sin u_2 du = 4\pi r^2. \end{aligned}$$

Discutiamo ora alcuni casi speciali cominciando dal caso $k = 1$: se I è un intervallo di estremi $a < b$ e $\Gamma := \varphi(I)$ è un elemento di curva regolare (o una curva regolare a tratti ed eventualmente chiusa) la quantità in (C.10) coincide con la norma euclidea del vettore $\varphi'(t)$ (come al solito $t = u$) e dunque la lunghezza della curva Γ è data da

$$\sigma_1(\Gamma) := L(\Gamma) := \int_a^b |\varphi'(t)| dt, \quad (\text{C.15})$$

mentre l'integrale di una funzione $f \in C(\bar{\Gamma}, \mathbb{R})$ su Γ , che prende il nome di **integrale curvilineo**, è dato da

$$\int_{\Gamma} f d\sigma_1 := \int_a^b f \circ \varphi(t) |\varphi'(t)| dt; \quad (\text{C.16})$$

a volte (nel caso $k = 1$) al simbolo $d\sigma_1$ si preferiscono i simboli ds o dl .

¹³Si noti che la rappresentazione di una varietà regolare a tratti come 'unione di parti regolari' $\varphi^{(i)}(E^{(i)})$ non è ovviamente unica e d'altra parte è altrettanto ovvio che le definizioni in (C.14) non dipendono dalla particolare rappresentazione di \mathcal{S} .

Il caso $k = 2$ e $n = 3$ concerne l'integrazione su superfici in \mathbb{R}^3 . La quantità in (C.10) coincide con il quadrato della norma euclidea del vettore $\frac{\partial\varphi}{\partial u_1} \times \frac{\partial\varphi}{\partial u_2}$ e dunque l'area della superficie \mathcal{S} è data da

$$\sigma_2(\mathcal{S}) := \text{Area}(\mathcal{S}) := \int_U \left| \frac{\partial\varphi}{\partial u_1} \times \frac{\partial\varphi}{\partial u_2} \right| du; \quad (\text{C.17})$$

mentre l'integrale di una funzione $f \in C(\overline{\mathcal{S}}, \mathbb{R})$ su \mathcal{S} , che prende il nome di **integrale superficiale**, è dato da

$$\int_{\mathcal{S}} f d\sigma_2 := \int_U f \circ \varphi \left| \frac{\partial\varphi}{\partial u_1} \times \frac{\partial\varphi}{\partial u_2} \right| du. \quad (\text{C.18})$$

Osservazione C.7 (i) Considerando un'“approssimazione di Riemann” dell'integrale come in (C.15), relativa alla partizione $t_0 = a < t_1 < \dots < t_N = b$, si ha che¹⁴ la lunghezza della poligonale di estremi (in ordine) $\varphi(a), \varphi(t_1), \dots, \varphi(b)$ tende a $L(\Gamma)$, quando l'ampiezza della partizione $\delta := \sup(t_{i+1} - t_i)$ tende a zero, giustificando, in tal modo, il nome ‘lunghezza’ dato a $L(\Gamma)$.

(ii) La quantità $\left| \frac{\partial\varphi}{\partial u_1} \times \frac{\partial\varphi}{\partial u_2} \right|$ rappresenta geometricamente l'area del parallelogramma generato dai vettori $\frac{\partial\varphi}{\partial u_1}$ e $\frac{\partial\varphi}{\partial u_2}$ e quest'interpretazione farebbe pensare alla possibilità di giustificare (tramite opportune approssimazioni di Riemann) la definizione di area in modo analogo a quanto fatto nel punto (i) per la lunghezza. Invece ciò non è possibile: in generale, un elemento di superficie regolare può essere ‘approssimato’ con (ad esempio) triangolini con i vertici sulla superficie in modo tale che la somma delle aree dei triangolini tende ad infinito quando la lunghezza del massimo lato di tali triangolini viene mandata a zero¹⁵.

(iii) La quantità $\left(\sum_{1 \leq i_1 < \dots < i_k \leq n} \left| \det \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(u_1, \dots, u_k)} \right|^2 \right)^{\frac{1}{2}}$ che appare nella definizione di misura k -dimensionale non è altro che la misura euclidea k -dimensionale del parallelepipedo generato dai vettori tangenti (che formano una base dello spazio tangente) $\frac{\partial\varphi}{\partial u_1}, \dots, \frac{\partial\varphi}{\partial u_k}$. Per maggiori informazioni si veda C.3.

C.2.2 Il Teorema della divergenza in \mathbb{R}^n

In questo paragrafo discutiamo la generalizzazione a \mathbb{R}^n del Teorema di Gauss. Di questo fondamentale teorema della matematica e della fisica daremo due dimostrazioni: la prima, nel complemento 4.4, è una dimostrazione ‘diretta’, la seconda, invece, sarà ottenuta come conseguenza della formula di Stokes in \mathbb{R}^n (basata sulla teoria delle k -forme differenziali, di cui parleremo nel prossimo capitolo).

Facciamo un'osservazione preliminare. La sfera unitaria chiusa $\overline{B^n} := \{x \in \mathbb{R}^n : |x| \leq 1\}$ non è una varietà n dimensionale ma è l'unione d'un aperto (cioè la sfera $B^n := \{x \in \mathbb{R}^n : |x| < 1\}$) e d'una varietà compatta di dimensione $n - 1$ (cioè S^{n-1}) che coincide con la sua frontiera insiemistica. L'ovvia importanza di tali insiemi aperti giustifica la seguente definizione

Definizione C.8 Per $n \geq 2$, un insieme $A \subseteq \mathbb{R}^n$ aperto, limitato e connesso si dice **regolare** se esiste una funzione $\phi \in C^1(\mathbb{R}^n, \mathbb{R})$ tale che $A = \{x \in \mathbb{R}^n : \phi(x) < 0\}$, $\partial A = \{x : \phi(x) = 0\}$ e $\nabla\phi(x) \neq 0$ per ogni $x \in \partial A$.

Dunque la sfera aperta B^n è un insieme regolare: $B^n = \{x \in \mathbb{R}^n : \phi(x) < 0\}$ con $\phi(x) = |x|^2 - 1$.

Dall'esempio (E6) di § C.1 e dalla Definizione C.4 segue che la frontiera di un insieme regolare è una $(n - 1)$ -varietà compatta in \mathbb{R}^n . Infatti la frontiera di A divide \mathbb{R}^n in due insiemi connessi

¹⁴Si noti che $L(P(\varphi(t_i), \varphi(t_{i+1}))) = |\varphi(t_{i+1}) - \varphi(t_i)| = |\varphi'(\tilde{t}_i)|(t_{i+1} - t_i)$. Quindi, essendo $|\varphi'|$ continua su $[a, b]$, dalla teoria dell'integrazione di Riemann segue che, se $\delta = \sup_i |t_{i+1} - t_i|$, allora $\lim_{\delta \rightarrow 0} L(P(\varphi(t_i), \varphi(t_{i+1}))) = \int_a^b |\varphi'(t)| dt$.

¹⁵Si veda 4.43.

di cui uno, $\{x : \phi(x) < 0\} = A$, è limitato e l'altro, $\{x : \phi(x) > 0\} = (\bar{A})^c$ è illimitato: in questa situazione l'insieme aperto \bar{A}^c prende il nome di **esterno di A** .

In ogni punto x della frontiera di un insieme regolare $A = \{x' : \phi(x') < 0\}$ definiamo la **normale esterna**¹⁶

$$\nu := \nu(x) := \frac{\nabla\phi(x)}{|\nabla\phi(x)|}. \quad (\text{C.19})$$

Infine se A è un aperto di \mathbb{R}^n e se $F \in C^1(A, \mathbb{R}^n)$ definiamo la **divergenza di F** la funzione continua da A in \mathbb{R} data da

$$\operatorname{div} F := \sum_{i=1}^n \frac{\partial F_i}{\partial x_i}; \quad (\text{C.20})$$

la divergenza di F viene anche denotata con $\nabla \cdot F$.

Teorema C.9 (Teorema della divergenza) *Sia $A := \{x \in \mathbb{R}^n : \phi(x) < 0\}$ un insieme regolare in \mathbb{R}^n (Definizione C.8) e sia $F \in C^1(\bar{A}, \mathbb{R}^n)$. Allora*

$$\int_A \nabla \cdot F \, dx = \int_{\partial A} F \cdot \nu \, d\sigma_{n-1}. \quad (\text{C.21})$$

Osservazione C.10 (i) La quantità $F \cdot \nu$ in (C.21) prende il nome di **flusso esterno** del campo vettoriale¹⁷ F e dunque il Teorema della divergenza può essere formulato dicendo che “l'integrale della divergenza di un campo regolare su di un insieme regolare A coincide con l'integrale del flusso esterno sulla frontiera di A ”.

(ii) Se $F \in C^1([a, b], \mathbb{R})$ e se interpretassimo la quantità $F(b) - F(a)$ come il “flusso totale esterno di F dall'insieme $[a, b]$ ” vedremmo che (nel caso $n = 1$) l'enunciato del Teorema della divergenza dato al punto (i) coincide sostanzialmente con il Teorema fondamentale del calcolo (definendo, se si vuole, gli insiemi regolari in \mathbb{R} come gli intervalli aperti e limitati). In effetti vedremo che la dimostrazione del Teorema della divergenza è *conseguenza* del Teorema fondamentale del calcolo in una variabile.

C.3 k -forme differenziali

Fissiamo due 1-forme elementari in \mathbb{R}^n , diciamo dx_i e dx_j con $1 \leq i, j \leq n$ e definiamo il loro **prodotto esterno** $dx_i \wedge dx_j$ come la seguente forma bilineare ed antisimmetrica su coppie di vettori in \mathbb{R}^n : dati $\xi, \eta \in \mathbb{R}^n$ l'azione di $dx_i \wedge dx_j$ è data da

$$dx_i \wedge dx_j(\xi, \eta) := \det \begin{pmatrix} \xi_i & \eta_i \\ \xi_j & \eta_j \end{pmatrix} = \det \begin{pmatrix} dx_i(\xi) & dx_i(\eta) \\ dx_j(\xi) & dx_j(\eta) \end{pmatrix}. \quad (\text{C.22})$$

È immediato verificare che tale forma è lineare sia in ξ che in η e che

$$\begin{aligned} dx_i \wedge dx_j(\xi, \eta) &= -dx_i \wedge dx_j(\eta, \xi), \\ dx_i \wedge dx_j &= -dx_j \wedge dx_i, \quad \forall i, j = 1, \dots, n \\ dx_i \wedge dx_i &= 0, \quad \forall i = 1, \dots, n. \end{aligned} \quad (\text{C.23})$$

¹⁶Si noti che tale vettore (che è ben definito poiché $\nabla\phi \neq 0$ su ∂A) ha norma unitaria. La parola “esterna” è dovuta al fatto che “ ν punta verso l'esterno” ovvero “esiste $\varepsilon > 0$ tale che $\{x + \nu t \mid 0 < t < \varepsilon\} \subseteq \bar{A}^c$ ”. Infatti: dalla formula di Taylor, se $x \in \partial A$ (ovvero se x tale che $\phi(x) = 0$) e se $t > 0$ è piccolo, $\phi(x + \nu t) = \nabla\phi(x) \cdot \nu t + o(t) = |\nabla\phi(x)|t + o(t) = t \left(|\nabla\phi(x)| + \frac{o(t)}{t} \right)$ che è strettamente positivo se t è sufficientemente piccolo. La parola “normale” si riferisce al fatto che ν è normale (o “ortogonale”) ad un qualunque vettore “tangente a ∂A in x ” [un vettore $\xi \in \mathbb{R}^n$ si dice **tangente** all'elemento di varietà \mathcal{S} in $x \in \mathcal{S}$ se $\xi = z'(0)$ per una qualche applicazione $z \in C^1((-\delta, \delta), \mathbb{R}^n)$ tale che $z(t) \in \mathcal{S}$ per ogni $|t| < \delta$ e $z(0) = x$; per maggiori informazioni si veda C.2]. Infatti se $\xi = z'(0)$ è un vettore tangente a ∂A in x si ha, per definizione, che $z(t) \in \partial A$ per $|t| \leq \delta$ (per un qualche $\delta > 0$), ovvero $\phi(z(t)) = 0$ per $|t| \leq \delta$; derivando tale relazione in $t = 0$ si ottiene $0 = \nabla\phi(z(0)) \cdot z'(0) = \nabla\phi(x) \cdot \xi$ e quindi $\xi \cdot \nu = 0$.

¹⁷Un “campo vettoriale su $A \subseteq \mathbb{R}^n$ non è altro che una funzione F da A in \mathbb{R}^n (ovvero una funzione che ad un vettore $x \in A \subseteq \mathbb{R}^n$ associa un altro vettore, $F(x)$ di \mathbb{R}^n).

Definizione C.11 *Il prodotto esterno di due 1-forme elementari si chiama 2–forma elementare. Una 2–forma differenziale, ω^2 , su A aperto di \mathbb{R}^n , è una combinazione lineare (a coefficienti funzioni) di 2–forme elementari:*

$$\omega^2 = \sum_{i,j=1}^n f_{ij} dx_i \wedge dx_j \quad (\text{C.24})$$

con f_{ij} funzioni regolari su A a valori in \mathbb{R} . Fissato $x \in A$, l'azione di ω^2 su coppie di vettori in \mathbb{R}^n è definita, per linearità, come

$$\omega^2(\xi, \eta) := \sum_{i,j} f_{ij}(x) dx_i \wedge dx_j(\xi, \eta). \quad (\text{C.25})$$

Dunque da (C.23) e (C.25) segue immediatamente che $\omega^2(\xi, \eta)$ è lineare sia in ξ che in η e che

$$\omega^2(\xi, \eta) = -\omega^2(\eta, \xi), \quad \left(\implies \omega^2(\xi, \xi) = 0 \right). \quad (\text{C.26})$$

Si noti che, sempre in virtù di (C.23), gli addendi in (C.24) con $i = j$ sono identicamente nulli, e che i termini con $j > i$ possono esseri messi a fattore (col segno cambiato) di $dx_i \wedge dx_j$ con $i < j$. Quindi, possiamo riscrivere (C.24), come

$$\omega^2 = \sum_{1 \leq i < j \leq n} g_{ij} dx_i \wedge dx_j \quad (\text{C.27})$$

con $g_{ij} = f_{ij} - f_{ji}$ e dove la somma è estesa a tutte le coppie ordinate (i, j) di interi tra 1 ed n con $i < j$.

Poiché il numero di coppie ordinate (i, j) con $1 \leq i < j \leq n$ è $\binom{n}{2}$, vediamo che assegnare una 2–forma in $A \subseteq \mathbb{R}^n$ equivale ad assegnare $\binom{n}{2}$ funzioni regolari su A .

Per esempio, in \mathbb{R}^3 , ogni 2–forma si rappresenta come

$$\omega^2 = g_{12} dx_1 \wedge dx_2 + g_{13} dx_1 \wedge dx_3 + g_{23} dx_2 \wedge dx_3. \quad (\text{C.28})$$

L'estensione al caso $k > 2$ è immediata. Una k –forma elementare in \mathbb{R}^n è data dal prodotto esterno di k 1–forme elementari $dx_{i_1} \wedge \cdots \wedge dx_{i_k}$, con $i_j \in \{1, \dots, n\}$, che agisce sulle k –uple di vettori $(\xi^{(1)}, \dots, \xi^{(k)})$ con $\xi^{(i)} \in \mathbb{R}^n$ secondo la regola

$$\begin{aligned} dx_{i_1} \wedge \cdots \wedge dx_{i_k}(\xi^{(1)}, \dots, \xi^{(k)}) &:= \det \begin{pmatrix} \xi_{i_1}^{(1)} & \cdots & \xi_{i_1}^{(k)} \\ \vdots & \ddots & \vdots \\ \xi_{i_k}^{(1)} & \cdots & \xi_{i_k}^{(k)} \end{pmatrix} \\ &= \det (dx_{i_h}(\xi^{(j)}))_{h,j=1,\dots,k}. \end{aligned} \quad (\text{C.29})$$

Come sopra, da tale definizione, segue immediatamente che

$$dx_{i_1} \wedge \cdots \wedge dx_{i_k}(\xi^{(1)}, \dots, \xi^{(k)})$$

è lineare in ogni $\xi^{(j)}$ e che scambiando di posto due $\xi^{(j)}$ o due dx_{i_h} il valore risultante cambia segno e, quindi, se $i_h = i_{h'}$ o se $\xi^{(j)} = \xi^{(j')}$ per qualcuno degli indici che appaiono in (C.29), il valore alla destra di (C.29) è 0.

Definizione C.12 *Una k –forma, ω^k , in $A \subseteq \mathbb{R}^n$ è una combinazione lineare a coefficienti funzioni delle k –forme elementari definite in (C.29).*

Con ragionamenti identici a quelli fatti più sopra vediamo che ogni k -forma in \mathbb{R}^n può scriversi come

$$\omega^k := \sum_{\substack{i_1 < \dots < i_k \\ i_j = 1, \dots, n}} g_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k} \quad (\text{C.30})$$

e se $\xi^{(j)}$ sono k vettori in \mathbb{R}^n , ($j = 1, \dots, k$), si definisce il numero $\omega^k(\xi^{(1)}, \dots, \xi^{(k)})$ come

$$\omega^k(\xi^{(1)}, \dots, \xi^{(k)}) := \sum_{\substack{i_1 < \dots < i_k \\ i_j = 1, \dots, n}} g_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}(\xi^{(1)}, \dots, \xi^{(k)}). \quad (\text{C.31})$$

Essendo il numero di k -uple di interi ordinati tra 0 ed n uguale a $\binom{n}{k}$, vediamo che *assegnare una k -forma in $A \subseteq \mathbb{R}^n$ equivale ad assegnare $\binom{n}{k}$ funzioni regolari su A* . Una k -forma viene anche chiamata una **forma (differenziale) di grado k** .

Osservazione C.13 (i) Si noti che se $k > n$ ci saranno sempre due dx_{i_j} ripetuti in (C.29) e quindi *ogni k -forma in \mathbb{R}^n con $k > n$ è identicamente nulla*.

(ii) Se $k = n$ vi è una unica k -upla di indici ordinati e cioè $(1, \dots, n)$ e quindi se $k = n$

$$\omega^n = g dx_1 \wedge \dots \wedge dx_n \quad (\text{C.32})$$

per una qualche funzione regolare $g \in C^\infty(A, \mathbb{R})$. In altri termini se definiamo l'**elemento di volume** in \mathbb{R}^n come la n -forma elementare $dx_1 \wedge \dots \wedge dx_n$, vediamo che ω^n è proporzionale all'elemento di volume in \mathbb{R}^n ; vedi (ii) dell'osservazione C.15 più sotto.

(iii) Nel caso $k = n - 1$, $\binom{n}{k} = n$ e l'elenco di tutte (a meno di riordinamenti) le $(n - 1)$ -forme elementari è dato da

$$dx_1 \wedge \dots \wedge \widehat{dx}_i \wedge \dots \wedge dx_n, \quad (i = 1, \dots, n) \quad (\text{C.33})$$

dove il simbolo $\widehat{(\cdot)}$ su dx_i significa che tale forma è omessa¹⁸. Quindi, come nel caso $k = 1$, assegnare una $(n - 1)$ -forma in \mathbb{R}^n equivale ad assegnare un campo vettoriale $F \in C^\infty(A, \mathbb{R}^n)$.

Data $F \in C^\infty(A, \mathbb{R}^n)$ definiamo

$$\alpha_F^{n-1} := \sum_{i=1}^n (-1)^{i-1} F_i dx_1 \wedge \dots \wedge \widehat{dx}_i \wedge \dots \wedge dx_n \quad (\text{C.34})$$

(la presenza del segno $(-1)^{i-1}$ apparirà chiaro in seguito). Si noti che per $n = 2$ una 1-forma è anche una $(n - 1)$ -forma ma che $\omega_F^1 := F_1 dx_1 + F_2 dx_2 \neq \alpha_F^1 := F_1 dx_2 - F_2 dx_1$, essendo, invece,

$$\alpha_F^1 = \omega_{(-F_2, F_1)}^1. \quad (\text{C.35})$$

Questo è il motivo per cui abbiamo usato due simboli diversi (ω e α) per denotare, rispettivamente, le 1-forme e le $(n - 1)$ -forme associate (univocamente) ad un campo vettoriale in $A \subseteq \mathbb{R}^n$.

Definiamo ora, in generale, il prodotto esterno di due forme differenziali di grado rispettivamente k e h .

Definizione C.14 *Siano*

$$\begin{aligned} \omega^k &:= \sum_{i_1 < \dots < i_k} g_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}, \\ \alpha^h &:= \sum_{i_1 < \dots < i_h} f_{i_1 \dots i_h} dx_{i_1} \wedge \dots \wedge dx_{i_h} \end{aligned}$$

¹⁸ Per esempio $\widehat{dx}_1} \wedge dx_2 \wedge dx_3 = dx_2 \wedge dx_3$ oppure $dx_1 \wedge \widehat{dx}_2} \wedge dx_3 = dx_1 \wedge dx_3$.

due forme differenziali di grado, rispettivamente, k e h in $A \subseteq \mathbb{R}^n$. Il prodotto esterno $\omega^k \wedge \alpha^h$ è la $k+h$ -forma in A data da

$$\omega^k \wedge \alpha^h := \sum_{i_1 < \dots < i_k} \sum_{j_1 < \dots < j_h} g_{i_1 \dots i_k} f_{j_1 \dots j_h} dx_{i_1} \wedge \dots \wedge dx_{i_k} \wedge dx_{j_1} \wedge \dots \wedge dx_{j_h}. \quad (\text{C.36})$$

Si noti che molti dei termini nelle sommatorie in (C.36) possono essere nulli (non appena $i_{k'} = j_{h'}$ per qualche $1 \leq k' \leq k$ e $1 \leq h' \leq h$).

Come abbiamo fatto nella sezione precedente, possiamo definire **l'integrale di una k -forma su di un elemento di k -varietà orientato**. Sia $\mathcal{S} = \varphi(U)$ un elemento di k -varietà in $A \subseteq \mathbb{R}^n$ con l'orientamento indotto da φ ($\text{Or} = \text{Or}(\varphi)$) e sia ω^k una k -forma come in (C.30). Si definisce allora

$$\begin{aligned} \int_{\mathcal{S}, \text{Or}} \omega^k &:= \sum_{i_1 < \dots < i_k} \int_U g_{i_1 \dots i_k} \circ \varphi(u) \det \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial(u_1, \dots, u_k)}(u) du \\ &= \int_U \omega^k \left(\frac{\partial \varphi}{\partial u_1}, \dots, \frac{\partial \varphi}{\partial u_k} \right) du. \end{aligned} \quad (\text{C.37})$$

Osservazione C.15 (i) Qualora non ci sia ambiguità (e questo accadrà la maggior parte delle volte) ometteremo l'indicazione esplicita dell'orientamento negli integrali di forme differenziali ma va sempre tenuto a mente che, al contrario degli integrali curvilinei e superficiali, gli integrali di forme differenziali dipendono dall'orientamento fissato e si ha¹⁹:

$$\int_{\mathcal{S}, -\text{Or}} \omega^k = - \int_{\mathcal{S}, \text{Or}} \omega^k. \quad (\text{C.38})$$

(ii) Per chiarire il termine 'elemento di volume' dato alla n -forma $dx_1 \wedge \dots \wedge dx_n$, consideriamo un aperto U connesso e misurabile ed una funzione f continua su \bar{U} . Tale aperto U è un elemento di n -varietà in \mathbb{R}^n (esempio (E1) della sezione C.1). Allora, considerando l'immersione identica $u \in U \rightarrow x = \varphi(u) := u$ con il relativo orientamento Or , ed essendo $\det \frac{\partial \varphi}{\partial u} = 1$ e $f \circ \varphi(u) = f(u)$,

$$\int_U f dx_1 \wedge \dots \wedge dx_n := \int_U f(u) du = \int_U f(x) dx \quad (\text{C.39})$$

dove nell'ultima uguaglianza abbiamo semplicemente cambiato nome alla variabile di integrazione. Quindi *integrare in \mathbb{R}^n la n -forma $f(x)dx_1 \wedge \dots \wedge dx_n$ equivale a calcolare l'integrale di Riemann della funzione f .*

(iii) Uno dei motivi per la definizione data in (C.34) è il seguente. Consideriamo un elemento di superficie orientato in \mathbb{R}^3 , $\mathcal{S} := \varphi(U)$, $\text{Or} = \text{Or}(\varphi)$. Si verifica allora immediatamente che

$$\begin{aligned} \int_{\mathcal{S}} \alpha_F^2 &= \int_U F \cdot \frac{\partial \varphi}{\partial u_1} \times \frac{\partial \varphi}{\partial u_2} du \\ &= \int_{\mathcal{S}} F \cdot \nu d\sigma \end{aligned} \quad (\text{C.40})$$

dove ν è la *normale esterna* definita in (4.202). La quantità nell'ultima riga di (C.40) si chiama **il flusso esterno di F attraverso la superficie orientata \mathcal{S}** .

Consideriamo ora una funzione regolare (almeno C^1)

$$f : y \in B \subseteq \mathbb{R}^m \rightarrow x = f(y) = (f_1(y), \dots, f_n(y)) \in A \subseteq \mathbb{R}^n$$

con A e B insiemi aperti. Data una k -forma ω^k su A definiremo ora una k -forma in B .

¹⁹Esercizio C.39.

Definizione C.16 Siano A, B due insiemi aperti di, rispettivamente, \mathbb{R}^n ed \mathbb{R}^m , sia $f \in C^\infty(B, A)$ e $\omega^k := \sum_{i_1 < \dots < i_k} g_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}$ una k -forma su A . Si chiama ‘pull-back’ della forma ω^k la k -forma su B definita da

$$f^* \omega^k := \sum_{i_1 < \dots < i_k} g_{i_1 \dots i_k} \circ f df_{i_1} \wedge \dots \wedge df_{i_k}. \quad (\text{C.41})$$

Osservazione C.17 (i) Tale definizione è significativa solo se $k \leq \min\{n, m\}$, altrimenti $f^* \omega^k := 0$.

(ii) L’operazione f^* è lineare: se ω^k e α^k sono due k -forme su A , allora²⁰

$$f^*(\omega^k + \alpha^k) = f^* \omega^k + f^* \alpha^k. \quad (\text{C.42})$$

(iii) Se $k = m$, per il punto (ii) dell’osservazione C.13, la forma $f^* \omega^k$ sarà proporzionale all’elemento di volume $dy_1 \wedge \dots \wedge dy_k$ su \mathbb{R}^k . Per (ii) qui sopra, basterà vedere come agisce f^* , in questo caso, su $\omega^k := g dx_{i_1} \wedge \dots \wedge dx_{i_k}$. Calcoliamo prima $f^*(dx_{i_1} \wedge \dots \wedge dx_{i_k})$:

$$\begin{aligned} f^*(dx_{i_1} \wedge \dots \wedge dx_{i_k}) &:= df_{i_1} \wedge \dots \wedge df_{i_k} \\ &= \sum_{\substack{j_1, \dots, j_k \\ j_h = 1, \dots, k}} \frac{\partial f_{i_1}}{\partial y_{j_1}} \dots \frac{\partial f_{i_k}}{\partial y_{j_k}} dy_{j_1} \wedge \dots \wedge dy_{j_k} \\ &= \left(\sum_{(j_1, \dots, j_k) \in \mathcal{P}_k} \varepsilon_{(j_1, \dots, j_k)} \frac{\partial f_{i_1}}{\partial y_{j_1}} \dots \frac{\partial f_{i_k}}{\partial y_{j_k}} \right) dy_1 \wedge \dots \wedge dy_k \\ &= \det \frac{\partial (f_{i_1}, \dots, f_{i_k})}{\partial (y_1, \dots, y_k)} dy_1 \wedge \dots \wedge dy_k, \end{aligned} \quad (\text{C.43})$$

dove \mathcal{P}_k denota l’insieme delle permutazioni dell’insieme $\{1, \dots, k\}$ e, come al solito, $\varepsilon_{(j_1, \dots, j_k)}$ denota il segno della permutazione (j_1, \dots, j_k) . Quindi,

$$f^*(g dx_{i_1} \wedge \dots \wedge dx_{i_k}) = g \circ f \left(\det \frac{\partial (f_{i_1}, \dots, f_{i_k})}{\partial y} \right) dy_1 \wedge \dots \wedge dy_k. \quad (\text{C.44})$$

(iv) Sia $\mathcal{S} = \varphi(U)$ un elemento di k -varietà in A ed ω^k una k -forma su A . Mettendo insieme le osservazioni precedenti, vediamo che l’integrale di ω^k su \mathcal{S} (orientata secondo φ) non è altro che l’integrale su U di $\varphi^* \omega^k$:

$$\int_{\mathcal{S}} \omega^k = \int_U \varphi^* \omega^k \quad (\text{C.45})$$

dove, naturalmente, l’orientamento su U nel secondo integrale è quello indotto dalla mappa identica.

C.4 Derivata esterna di forme differenziali

Una funzione f regolare su A aperto di \mathbb{R}^n è anche chiamata una **0-forma** su A . Il differenziale ‘ d ’ trasforma la 0-forma f in una 1-forma df . Estendiamo tale operazione ‘ d ’, in modo tale che d trasformi k -forme (su \mathbb{R}^n) in $(k+1)$ -forme (sempre su \mathbb{R}^n).

Definizione C.18 Sia $\omega^k := \sum_{i_1 < \dots < i_k} g_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}$ una k -forma su $A \subseteq \mathbb{R}^n$. La **derivata esterna** di ω^k , denotata $d\omega^k$, è la seguente $k+1$ -forma su A :

$$\begin{aligned} d\omega^k &:= \sum_{i_1 < \dots < i_k} dg_{i_1 \dots i_k} \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k} \\ &= \sum_{i_1 < \dots < i_k} \sum_{j=1}^n \frac{\partial g_{i_1 \dots i_k}}{\partial x_j} dx_j \wedge dx_{i_1} \wedge \dots \wedge dx_{i_k}. \end{aligned} \quad (\text{C.46})$$

²⁰Esercizio C.37.

Per esempio, se $\omega^1 = f dx_2$, con $f \in C^1(\mathbb{R}^3)$, si ha:

$$d\omega^1 = df \wedge dx_2 = \sum_{j=1}^3 \frac{\partial f}{\partial x_j} dx_j \wedge dx_2 = f_{x_1} dx_1 \wedge dx_2 - f_{x_3} dx_2 \wedge dx_3.$$

Osservazione C.19 (i) Segue immediatamente dalla definizione che d è lineare:

$$d(\omega^k + \eta^k) = d\omega^k + d\eta^k. \quad (\text{C.47})$$

(ii) Se $k = n$, $d\omega^n$ è una $(n+1)$ forma su \mathbb{R}^n e quindi²¹ è nulla: $d\omega^n = 0$.

(iii) Dalla definizione C.18 segue (ancora immediatamente) che la derivata esterna di una forma elementare è 0:

$$d(dx_{i_1} \wedge \cdots \wedge dx_{i_k}) := d1 \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} = 0 \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} = 0. \quad (\text{C.48})$$

(iv) Sia $F \in C^1(A, \mathbb{R}^n)$ e consideriamo la $(n-1)$ -forma α_F^{n-1} definita in (C.34). Allora

$$\begin{aligned} d\alpha_F^{n-1} &:= \sum_{i=1}^n (-1)^{i-1} dF_i \wedge dx_1 \wedge \cdots \wedge \widehat{dx}_i \wedge \cdots \wedge dx_n \\ &= \sum_{i=1}^n (-1)^{i-1} \sum_{j=1}^n \frac{\partial F_i}{\partial x_j} dx_j \wedge dx_1 \wedge \cdots \wedge \widehat{dx}_i \wedge \cdots \wedge dx_n \\ &= \sum_{i=1}^n (-1)^{i-1} \frac{\partial F_i}{\partial x_i} dx_i \wedge dx_1 \wedge \cdots \wedge \widehat{dx}_i \wedge \cdots \wedge dx_n \\ &= \left(\sum_{i=1}^n \frac{\partial F_i}{\partial x_i} \right) dx_1 \wedge \cdots \wedge dx_n, \end{aligned}$$

Si noti che il segno $(-1)^{i-1}$ conta esattamente il numero di scambi (tra i dx_j) che bisogna fare per mettere il dx_i al posto ‘giusto’: in termini più precisi

$$\varepsilon_{(i,1,2,\dots,i-1,i+1,\dots,n)} = (-1)^{i-1}.$$

(v) In \mathbb{R}^3 ($n = 3$), la derivata esterna di una 1-forma è una 2 = $(n-1)$ forma; quindi, per ogni $F \in C^1(A, \mathbb{R}^3)$, $A \subseteq \mathbb{R}^3$, esiste una unica $G \in C(A, \mathbb{R}^3)$ tale che

$$d\omega_F^1 = \alpha_G^2. \quad (\text{C.49})$$

Osservazione C.20 (i) Ricordando le definizioni di divergenza, vedi (C.20), e quella di rotore, vedi (4.114), e dall’Osservazione C.19 segue che

$$d\alpha_F^{n-1} = (\nabla \cdot F) dx_1 \wedge \cdots \wedge dx_n, \quad (\text{C.50})$$

con $F \in C^1(A, \mathbb{R}^n)$, $A \subseteq \mathbb{R}^n$, e

$$d\omega_F^1 = \alpha_{\nabla \times F}^2, \quad (\text{C.51})$$

con $F \in C^1(A, \mathbb{R}^3)$, e $A \subseteq \mathbb{R}^3$. Tali identità giustificano ulteriormente la presenza del segno $(-1)^{i-1}$ nella definizione (C.34). Si noti che, in virtù del punto (v) dell’osservazione C.46, il rotore di F si può definire direttamente come l’unica funzione $G : A \rightarrow \mathbb{R}^3$ per cui vale (C.49).

(iii) I simboli di prodotto scalare e vettoriale che appaiono in (C.20) e (4.114) suggeriscono la seguente regola mnemonica: $\nabla \cdot F = \sum_{i=1}^n \partial_{x_i} F_i$ (dove il gradiente $\nabla := (\partial_{x_1}, \dots, \partial_{x_n})$) è trattato come fosse un vettore, ed, analogamente, il rotore $\nabla \times F$ è dato dallo sviluppo simbolico²² fatto rispetto alla prima riga del seguente *determinante formale*:

$$\det \begin{pmatrix} e^{(1)} & e^{(2)} & e^{(3)} \\ \partial_{x_1} & \partial_{x_2} & \partial_{x_3} \\ F_1 & F_2 & F_3 \end{pmatrix}. \quad (\text{C.52})$$

²¹ Osservazione C.13, (i).

²² Si ricordi anche (4.175).

Altre proprietà fondamentali dell'operazione d sono elencate nella seguente

Proposizione C.21

- (i) $d(d\omega^k) := d^2\omega^k = 0$.
- (ii) $d(\omega^k \wedge \alpha^h) = d\omega^k \wedge \alpha^h + (-1)^k \omega^k \wedge d\alpha^h$.
- (iii) $d(f^*\omega^k) = f^*d\omega^k$.

Naturalmente, nell'enunciato appena dato, ω^k e α^h sono, rispettivamente, una k -forma ed una h -forma su $A \subseteq \mathbb{R}^n$, e $f \in C^1(B, A)$ dove $B \subseteq \mathbb{R}^m$ (per il primo risultato occorre che $\omega^k \in C^2(A)$ mentre per gli altri risultati basterebbe assumere che le funzioni che intervengono siano $C^1(A)$).

Dimostrazione (i): Vista la linearità di d , basta verificare (i) nel caso $\omega^k = g \bar{\omega}^k$ dove $\bar{\omega}^k$ è la k -forma elementare $\bar{\omega}^k := dx_{i_1} \wedge \cdots \wedge dx_{i_k}$ (e $g \in C^2(A)$):

$$\begin{aligned} d(d\omega^k) &= d(dg \wedge \bar{\omega}^k) = d\left(\sum_{j=1}^n \frac{\partial g}{\partial x_j} dx_j \wedge \bar{\omega}^k\right) = \sum_{j=1}^n d\left(\frac{\partial g}{\partial x_j}\right) \wedge dx_j \wedge \bar{\omega}^k \\ &= \sum_{j,\ell=1}^n \frac{\partial^2 g}{\partial x_j \partial x_\ell} dx_\ell \wedge dx_j \wedge \bar{\omega}^k \\ &= \sum_{\ell < j} \left(\frac{\partial^2 g}{\partial x_j \partial x_\ell} - \frac{\partial^2 g}{\partial x_\ell \partial x_j}\right) dx_\ell \wedge dx_j \wedge \bar{\omega}^k = 0. \end{aligned}$$

(ii): per la linearità di d e del prodotto esterno, basta dimostrare (ii) con $\omega^k = g \bar{\omega}^k$ e $\alpha^h = f \bar{\alpha}^h$ dove $\bar{\omega}^k$ e $\bar{\alpha}^h$ sono le forme elementari

$$\bar{\omega}^k := dx_{i_1} \wedge \cdots \wedge dx_{i_k}, \quad \bar{\alpha}^h := dx_{j_1} \wedge \cdots \wedge dx_{j_h}. \quad (\text{C.53})$$

Usando le proprietà del prodotto esterno (in particolare la linearità e l'associatività) otteniamo:

$$\begin{aligned} d(\omega^k \wedge \alpha^h) &= d(gf \bar{\omega}^k \wedge \bar{\alpha}^h) = d(gf) \wedge \bar{\omega}^k \wedge \bar{\alpha}^h \\ &= (f dg \wedge \bar{\omega}^k \wedge \bar{\alpha}^h) + (g df \wedge \bar{\omega}^k \wedge \bar{\alpha}^h) \\ &= (f dg \wedge \bar{\omega}^k) \wedge \bar{\alpha}^h + (g df \wedge \bar{\omega}^k \wedge \bar{\alpha}^h) \\ &= (f dg \wedge \bar{\omega}^k) \wedge \bar{\alpha}^h + (-1)^k (g \bar{\omega}^k) \wedge (df \wedge \bar{\alpha}^h) \\ &= d\omega^k \wedge \alpha^h + (-1)^k \omega^k \wedge d\alpha^h. \end{aligned}$$

(iii): di nuovo, basta dimostrare (iii) per $\omega^k = g\bar{\omega}^k$ con $\bar{\omega}^k$ come in (C.53); si faccia attenzione al fatto che qui f è una funzione vettoriale $f : y \in B \subseteq \mathbb{R}^m \rightarrow (f_1, \dots, f_n) \in A \subseteq \mathbb{R}^n$ mentre $g : x \in A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}$ è una funzione scalare. Usando (C.121), otteniamo:

$$\begin{aligned} d(f^*\omega^k) &= d(g \circ f df_{i_1} \wedge \cdots \wedge df_{i_k}) = d(g \circ f) \wedge df_{i_1} \wedge \cdots \wedge df_{i_k} \\ &= \sum_{i=1}^n \frac{\partial g}{\partial x_i} \circ f df_i \wedge df_{i_1} \wedge \cdots \wedge df_{i_k} \\ &= f^*\left(\sum_{i=1}^n \frac{\partial g}{\partial x_i} dx_i \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k}\right) \\ &= f^*(d\omega^k). \quad \blacksquare \end{aligned}$$

Dal fatto che $d^2 = 0$ e da (4.99) segue subito che, se $f \in C^2(A)$, $A \subseteq \mathbb{R}^3$,

$$0 = d^2 f = d(df) = d\omega_{\nabla f}^1 = \alpha_{\nabla \times \nabla f}^2$$

ossia²³

$$\nabla \times \nabla f := \operatorname{rot} \nabla f = 0. \quad (\text{C.54})$$

Da (C.51) e da (C.50) segue che, se $F \in C^1(A, \mathbb{R}^3)$ con $A \subseteq \mathbb{R}^3$,

$$0 = d^2 \omega_F^1 = d\alpha_{\nabla \times F}^2 = (\nabla \cdot \nabla \times F) dx_1 \wedge dx_2 \wedge dx_3,$$

ossia

$$\nabla \cdot \nabla \times F := \operatorname{div} \operatorname{rot} F = 0. \quad (\text{C.55})$$

C.5 Forme chiuse. Forme esatte. Lemma di Poincaré

Consideriamo un insieme $A \subseteq \mathbb{R}^n$ e consideriamo l'insieme delle k -forme su A .

Definizione C.22 Una k -forma, ω , su A si dice **chiusa** se $d\omega = 0$. Una k -forma, ω , su A si dice **esatta** se esiste una $(k-1)$ -forma, α , su A tale che $\omega = d\alpha$.

Osservazione C.23 (i) Ricordando che le 0-forme sono le funzioni, vediamo che una 1-forma ω è esatta se e solo se è il differenziale di una funzione; equivalentemente $\omega := \omega_F^1$ è esatta se e solo se $F = \nabla f$ per una qualche funzione f . La proposizione 4.13 dà quindi una condizione necessaria e sufficiente (in termini dell'integrale di ω^1 su curve chiuse) affinché una 1-forma sia esatta.

(ii) Dal punto (i) della proposizione C.21 segue che ogni forma esatta è chiusa (su qualunque regione A). Il viceversa non è, in generale, vero. Infatti, se ω^1 è la 1-forma definita in (4.111) su $A = \mathbb{R}^2 \setminus \{0\}$, si vede subito che²⁴ $d\omega^1 = 0$ e cioè che ω^1 è chiusa. D'altra parte, per 4.52 e per l'osservazione (i) precedente, ω^1 non è esatta in A .

Il prossimo risultato mostra che su una larga classe di regioni A le forme chiuse coincidono con quelle esatte. Per formulare tale risultato, diamo la seguente

Definizione C.24 Un insieme aperto $A \subseteq \mathbb{R}^n$ si dice **stellato** rispetto al punto $x_0 \in A$ se, per ogni $x \in A$, il segmento $\ell(x_0, x)$ di estremi x e x_0 è interamente contenuto in A .

L'insieme $A = \mathbb{R}^2 \setminus \{0\}$ non è stellato. Gli insiemi convessi sono stellati.

Teorema C.25 (Lemma di Poincaré) Sia $A \subseteq \mathbb{R}^n$ un aperto stellato. Una k -forma ω^k su A è chiusa se e solo se è esatta.

La dimostrazione di questo teorema è basata sulla esistenza della seguente operazione \mathcal{I} che associa ad ogni k -forma su un insieme stellato A una $(k-1)$ -forma su A . Definiamo, dapprima, l'azione di \mathcal{I} su k -forme del tipo $g\bar{\omega}^k$, con $\bar{\omega}^k := dx_{i_1} \wedge \cdots \wedge dx_{i_k}$ una k -forma elementare fissata; estenderemo poi, per linearità, la definizione di \mathcal{I} ad una arbitraria k -forma su A . Assumiamo, per semplicità, che l' x_0 nella definizione C.24 coincida con l'origine: $x_0 = 0$ (essendo il caso generale riconducibile a questo mediante una semplice traslazione).

Nel caso $k = 1$, poniamo

$$\mathcal{I}(g dx_j) := x_j \int_0^1 g(tx) dt, \quad (\text{C.56})$$

nel caso $k > 1$, poniamo

$$\begin{aligned} \mathcal{I}(g dx_{i_1} \wedge \cdots \wedge dx_{i_k}) := & \\ \left\{ \int_0^1 t^{k-1} g(tx) dt \right\} \sum_{j=1}^k (-1)^{j-1} x_{i_j} dx_{i_1} \wedge \cdots \wedge \widehat{dx_{i_j}} \wedge \cdots \wedge dx_{i_k}. & \end{aligned} \quad (\text{C.57})$$

²³Si ricordi C.37 e C.38.

²⁴Esercizio C.41.

Infine per una generica k -forma poniamo

$$\mathcal{F}\left(\sum_{i_1 < \dots < i_k} g_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}\right) := \sum_{i_1 < \dots < i_k} \mathcal{F}(g_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}). \quad (\text{C.58})$$

Vale allora il seguente risultato che implica immediatamente il Lemma di Poincaré.

Proposizione C.26 *Sia A stellato e sia*

$$\omega := \sum_{i_1 < \dots < i_k} g_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k} \quad (\text{C.59})$$

una arbitraria²⁵ k forma su A . Allora

$$d(\mathcal{F}\omega) + \mathcal{F}(d\omega) = \omega. \quad (\text{C.60})$$

Quindi se ω è chiusa (cioè $d\omega = 0$) segue che $\omega = d(\mathcal{F}\omega)$ ossia ω è anche esatta e quindi vale il Lemma di Poincaré.

Dimostrazione Dimostriamo la proposizione nel caso $n = 3$ e $k = 2$. La dimostrazione del caso generale è assai simile e viene lasciata per esercizio.

Sia quindi

$$\omega := \sum_{i_1 < i_2} g_{i_1 i_2} dx_{i_1} \wedge dx_{i_2} \quad (\text{C.61})$$

con $i_1 < i_2$ ($i_j = 1, 2, 3$). Innanzitutto

$$d\omega = \sum_{i_1 < i_2} \sum_{j=1}^3 \frac{\partial g_{i_1 i_2}}{\partial x_j} dx_j \wedge dx_{i_1} \wedge dx_{i_2}. \quad (\text{C.62})$$

Dalla definizione di \mathcal{F} (essendo $k = 2$):

$$\mathcal{F}(\omega) = \sum_{i_1 < i_2} \left\{ \int_0^1 t g_{i_1 i_2}(tx) dt \right\} (x_{i_1} dx_{i_2} - x_{i_2} dx_{i_1}), \quad (\text{C.63})$$

$$\mathcal{F}(d\omega) = \sum_{i_1 < i_2} \sum_{j=1}^3 \left\{ \int_0^1 t^2 \frac{\partial g_{i_1 i_2}}{\partial x_j}(tx) dt \right\} [x_j dx_{i_1} \wedge dx_{i_2} - (x_{i_1} dx_j \wedge dx_{i_2} - x_{i_2} dx_j \wedge dx_{i_1})].$$

Calcoliamo, infine, $d(\mathcal{F}\omega)$: poiché

$$d\left(\int_0^1 t g_{i_1 i_2}(tx) dt\right) = \sum_{j=1}^3 \left(\int_0^1 t^2 \frac{\partial g_{i_1 i_2}}{\partial x_j}(tx) dt\right) dx_j \quad (\text{C.64})$$

e

$$d(x_{i_1} dx_{i_2} - x_{i_2} dx_{i_1}) = 2 dx_{i_1} \wedge dx_{i_2}$$

otteniamo (ricordando anche che $d(fg) = fdg + gdf$)

$$\begin{aligned} d(\mathcal{F}\omega) &= \sum_{i_1 < i_2} \sum_{j=1}^3 \left(\int_0^1 t^2 \frac{\partial g_{i_1 i_2}}{\partial x_j}(tx) dt\right) (x_{i_1} dx_j \wedge dx_{i_2} - x_{i_2} dx_j \wedge dx_{i_1}) \\ &\quad + 2 \sum_{i_1 < i_2} \left\{ \int_0^1 t g_{i_1 i_2}(tx) dt \right\} dx_{i_1} \wedge dx_{i_2}. \end{aligned} \quad (\text{C.65})$$

²⁵E quindi non necessariamente chiusa.

Quindi

$$\begin{aligned}
\mathcal{I}(d\omega) + d(\mathcal{I}\omega) &= \sum_{i_1 < i_2} \left\{ \int_0^1 \left(\sum_{j=1}^3 x_j t^2 \frac{\partial g_{i_1 i_2}}{\partial x_j}(tx) + 2t g_{i_1 i_2}(tx) \right) dt \right\} dx_{i_1} \wedge dx_{i_2} \\
&= \sum_{i_1 < i_2} \left\{ \int_0^1 \frac{d}{dt} (t^2 g_{i_1 i_2}(tx)) dt \right\} dx_{i_1} \wedge dx_{i_2} \\
&= \sum_{i_1 < i_2} g_{i_1 i_2}(x) dx_{i_1} \wedge dx_{i_2} = \omega. \quad \blacksquare
\end{aligned} \tag{C.66}$$

Dal Lemma di Poincaré e dalla discussione fatta alla fine della sezione C.4 segue immediatamente il seguente

Corollario C.27 (i) Sia A un aperto stellato di \mathbb{R}^n e $F \in C^1(A, \mathbb{R}^n)$. Condizione necessaria e sufficiente affinché $F = \nabla f$ per una qualche $f \in C^2(A, \mathbb{R})$ è che

$$\frac{\partial F_j}{\partial x_i} = \frac{\partial F_i}{\partial x_j}. \tag{C.67}$$

(ii) Sia A un aperto stellato di \mathbb{R}^3 ed $F \in C^1(A, \mathbb{R}^3)$. Condizione necessaria e sufficiente affinché $F = \nabla \times H$ per una qualche $H \in C^2(A, \mathbb{R}^3)$ è che

$$\operatorname{div} F := \nabla \cdot F = 0. \tag{C.68}$$

Dimostrazione La ‘necessarietà’ di tali condizioni è già stata verificata alla fine della sezione precedente (si ricordi in particolare (C.54) e (C.55) e C.38). Basta poi osservare²⁶ che (C.67) è equivalente a dire che $d\omega_F^1 = 0$ e quindi, per il lemma di Poincaré, esiste una 0-forma (cioè una funzione) f tale che $df = \omega_F^1$ che equivale a dire $\nabla f = F$.

Analogamente, da (C.50) segue che (C.68) è equivalente a dire che $d\alpha_F^2 = 0$ e quindi, per il Lemma di Poincaré, si ha che $\alpha_F^2 = d\omega_H^1$ per una qualche $H \in C^2(A, \mathbb{R}^3)$, il che, in base a (C.51), significa che $F = \operatorname{rot} H$. \blacksquare

C.6 Varietà con bordo

Sia H^k il semipiano (chiuso) k -dimensionale:

$$H^k := \{u \in \mathbb{R}^k : u_k \geq 0\}. \tag{C.69}$$

Definizione C.28 Un sottoinsieme di \mathbb{R}^n connesso e chiuso $\mathcal{S} \subseteq \mathbb{R}^n$ è una k -varietà con bordo se per ogni $x_0 \in \mathcal{S}$ esiste una inclusione $\varphi : \bar{U} \subseteq \mathbb{R}^k \rightarrow \mathcal{S}$ tale che U è un aperto (nella topologia relativa) di H^k , $\varphi(U)$ è un aperto (nella topologia relativa) di \mathcal{S} e $x_0 \in \varphi(U)$; assumiamo inoltre²⁷ che l’insieme $\hat{U} := \{\hat{u} \in \mathbb{R}^{k-1} : (\hat{u}, 0) \in U\}$ sia un sottoinsieme connesso e misurabile di \mathbb{R}^{k-1} .

Come nella sezione precedente, chiameremo **un atlante di \mathcal{S}** una collezione di inclusioni $\{(\varphi^{(i)}, U^{(i)})\}_{i \in J}$ che verifichino le proprietà della definizione appena data e tali che $\varphi^{(i)}(U^{(i)})$ ricoprano, al variare di i in J , la varietà \mathcal{S} .

Definizione C.29 Sia \mathcal{S} una k -varietà con bordo e $\{(\varphi^{(i)}, U^{(i)})\}_{i \in J}$ un suo atlante. Sia \hat{J} l’insieme degli indici i tali che $U^{(i)} \cap \{u_k = 0\} \neq \emptyset$

$$\hat{J} := \{i \in J : U^{(i)} \cap \{u_k = 0\} \neq \emptyset\}. \tag{C.70}$$

²⁶Esercizio C.45.

²⁷Per evitare patologie inutili.

Si chiama **bordo di \mathcal{S}** , e si denota $\partial\mathcal{S}$, il sottoinsieme di \mathcal{S} costituito dai punti x le cui ‘coordinate’ u appartengono alla frontiera di H^k (cioè abbiano $u_k = 0$):

$$\partial\mathcal{S} = \{x = \varphi^{(i)}(u) : u \in U^{(i)} \cap \{u_k = 0\}, i \in \hat{J}\}. \quad (\text{C.71})$$

Nel caso $\hat{J} = \emptyset$ ²⁸ poniamo $\partial\mathcal{S} = \emptyset$.

Osservazione C.30 Per denotare il bordo di \mathcal{S} si è usato simbolo $\partial\mathcal{S}$: si osservi, però, che, in generale, il bordo di \mathcal{S} non coincide con la frontiera di \mathcal{S} né nella topologia di \mathbb{R}^n , né nella topologia relativa su \mathcal{S} . Si consideri, per esempio, un intervallo chiuso immerso in \mathbb{R}^2 :

$$I := \{x \in \mathbb{R}^2 : x_1 \in [0, 1], x_2 = 0\}. \quad (\text{C.72})$$

Come sottoinsieme di \mathbb{R}^2 , I è chiuso ma non ha interno, cioè $\bar{I} = I$ ma $\overset{\circ}{I} = \emptyset$; dunque la frontiera di I (nella topologia di \mathbb{R}^2) coincide con I stesso: $\partial I = I$. Nella topologia relativa, ogni sottoinsieme $E \subseteq \mathbb{R}^n$ è sia aperto che chiuso²⁹; dunque, la frontiera, nella topologia relativa, è vuota: $\partial I = \emptyset$. D'altra parte, I è una 1-varietà con bordo in \mathbb{R}^2 ed il suo bordo è dato dai due punti $x := (0, 0)$ e $y := (1, 0)$: $\partial I = \{x, y\}$. Vi è dunque un'ambiguità di notazione nel simbolo ∂I ; però l'unica interpretazione non banale (nel caso di \mathcal{S} varietà con bordo) è quella di bordo. Quindi, d'ora in avanti, sarà quest'ultima, l'unica interpretazione che daremo del simbolo $\partial\mathcal{S}$ nel caso in cui \mathcal{S} sia una varietà con bordo in \mathbb{R}^n . Notiamo, infine, che se \mathcal{S} è una n -varietà con bordo in \mathbb{R}^n , allora il bordo di \mathcal{S} coincide con la frontiera di \mathcal{S} (nella topologia di \mathbb{R}^n)³⁰

Osservazione C.31 (i) Si noti che se $U \cap \{u \in \mathbb{R}^k : u_k = 0\} \neq \emptyset$ l'insieme $\varphi(U)$ non è un elemento di varietà nel senso delle definizioni date precedentemente³¹. In generale, chiameremo **elemento di varietà con bordo** l'insieme $\varphi(U)$ dove (φ, U) è una carta di un atlante di una varietà con bordo \mathcal{S} . Il bordo di tale elemento di varietà è, per definizione, l'insieme

$$\varphi(U) \cap \partial\mathcal{S}. \quad (\text{C.73})$$

Essenzialmente, tutte le definizioni relative alla integrazione (di funzioni e di forme differenziali) su elementi di varietà si estendono in maniera ovvia ad elementi di varietà con bordo. Per esempio rimangono invariate le definizioni date in (C.16) e (C.18) di, rispettivamente, integrale curvilineo ed integrale superficiale su di un elemento di, rispettivamente, 1 e 2-varietà con bordo. Analogamente rimane invariata la definizione di integrale di una k -forma su di un elemento $\varphi(U)$ di k -varietà con bordo; si veda (C.40). In particolare si ha

$$\int_{\varphi(U)} \omega = \int_U \varphi^* \omega. \quad (\text{C.74})$$

(ii) Se $U \subseteq \{u_k > 0\}$, dire che U è un aperto nella topologia relativa di H^k è equivalente a dire che U è un aperto di \mathbb{R}^k . Dunque se tutte le carte di un atlante di \mathcal{S} sono tali che $U \subseteq \{u_k > 0\}$, allora \mathcal{S} è una varietà nel senso della sezione precedente ossia, più precisamente, è una varietà chiusa (come sottoinsieme di \mathbb{R}^n) con bordo vuoto. sia $\hat{f} := \hat{f}_{ij}$ la corrispondente funzione di transizione relativa all'atlante $\hat{\mathcal{A}}$, e sia $u := (\hat{u}, u_k)$ con $\hat{u} := (u_1, \dots, u_{k-1})$. Si noti che

$$f := (f_1, \dots, f_k) : U_j^{(i)} \subseteq \mathbb{R}^k \rightarrow U_i^{(j)} \subseteq \mathbb{R}^k$$

mentre

$$\hat{f} := (\hat{f}_1, \dots, \hat{f}_{k-1}) : \hat{U}_j^{(i)} \subseteq \mathbb{R}^{k-1} \rightarrow \hat{U}_i^{(j)} \subseteq \mathbb{R}^{k-1}.$$

²⁸come, per esempio, nel caso di S^1 .

²⁹ E è aperto perché $E = E \cap \mathbb{R}^n$ e \mathbb{R}^n è un insieme aperto; E è anche chiuso poiché $E = E \setminus \emptyset$ e l'insieme vuoto è un sottoinsieme aperto di E ($\emptyset = E \cap \emptyset$ e \emptyset è un insieme aperto di \mathbb{R}^n).

³⁰Esercizio C.5.

³¹Si veda anche l'esempio (S5) dato più in basso.

Se $1 \leq \ell, s \leq k-1$, si ha

$$\frac{\partial \hat{f}_\ell}{\partial u_s}(\hat{u}) := \frac{\partial f_\ell}{\partial u_s}(\hat{u}, 0), \quad (\text{C.75})$$

mentre, essendo $f_k(\hat{u}, 0) := 0$ per ogni $\hat{u} \in \hat{U}_j^{(i)}$, si ha, per ogni $1 \leq s \leq k-1$,

$$\frac{\partial f_k}{\partial u_s}(\hat{u}, 0) = 0. \quad (\text{C.76})$$

Inoltre, dalla definizione di varietà con bordo, segue che se $h > 0$ è tale che $(\hat{u}, h) \in U_j^{(i)}$ si ha $f_k(\hat{u}, h) > 0$ e quindi

$$\frac{\partial f_k}{\partial u_k}(\hat{u}, 0) = \lim_{h \searrow 0} \frac{f_k(\hat{u}, h)}{h} \geq 0. \quad (\text{C.77})$$

Mettendo queste relazioni assieme si ottiene

$$\begin{aligned} \det \frac{\partial f}{\partial u}(\hat{u}, 0) &= \det \begin{pmatrix} \frac{\partial \hat{f}_1}{\partial u_1}(\hat{u}) & \dots & \frac{\partial \hat{f}_1}{\partial u_{k-1}}(\hat{u}) & \frac{\partial f_1}{\partial u_k}(\hat{u}, 0) \\ \vdots & \ddots & \vdots & \vdots \\ \frac{\partial \hat{f}_{k-1}}{\partial u_1}(\hat{u}) & \dots & \frac{\partial \hat{f}_{k-1}}{\partial u_{k-1}}(\hat{u}) & \frac{\partial f_{k-1}}{\partial u_k}(\hat{u}, 0) \\ 0 & \dots & 0 & \frac{\partial f_k}{\partial u_k}(\hat{u}, 0) \end{pmatrix} \\ &= \frac{\partial f_k}{\partial u_k}(\hat{u}, 0) \det \frac{\partial \hat{f}}{\partial \hat{u}}(\hat{u}). \end{aligned} \quad (\text{C.78})$$

Da tale relazione, essendo l'atlante \mathcal{A} orientato (e quindi $\det \frac{\partial f}{\partial u}(\hat{u}, 0) > 0$), si evince che $\frac{\partial f_k}{\partial u_k}(\hat{u}, 0) > 0$, e quindi che $\det \frac{\partial \hat{f}}{\partial \hat{u}} > 0$ ossia che l'atlante $\hat{\mathcal{A}}$ è orientato. ■

C.7 Integrazione di forme differenziali su varietà

Per definire l'integrale di una k -forma in $A \subseteq \mathbb{R}^n$ su di una k -varietà $\mathcal{S} \subseteq A$ con bordo e limitata, abbiamo bisogno del seguente risultato. Denotiamo $C_0^\infty(\mathbb{R}^n)$ l'insieme delle funzioni $C^\infty(\mathbb{R}^n)$ con supporto, $\text{supp}(f)$, compatto³². Si ricordi il Lemma 4.64.

Osservazione C.32 Una conseguenza immediata del lemma della partizione dell'unità è il seguente

Corollario C.33 (“Lemma di Urysohn”) *Siano $D \subseteq U \subseteq \mathbb{R}^n$ con D compatto e U aperto. Allora esiste una funzione $f \in C_0^\infty(\mathbb{R}^n)$ a valori in $[0, 1]$ con supporto contenuto in U tale che $f(x) = 1$ per ogni $x \in D$.*

Dimostrazione Per ogni $x \in D$, sia $B(x)$ una sfera centrata in x e la cui chiusura è contenuta in U . Allora $\{B(x) : x \in D\}$ è un ricoprimento di D e, se $\{f_j : 1 \leq j \leq N\}$ è una partizione dell'unità di D subordinata a tale ricoprimento, possiamo prendere $f := \sum_{j=1}^N f_j$. ■

Possiamo ora definire l'integrale di una k -forma su una k -varietà compatta con bordo. Cominciamo col definire l'integrale di una k -forma

$$\omega := \sum_{i_1 < \dots < i_k} g_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}$$

³² Si ricorda che $\text{supp}(f)$ è la chiusura di $\{x \in \mathbb{R}^n : f(x) \neq 0\}$. Per $0 \leq p \leq \infty$, una funzione $f \in C_0^p(\mathbb{R}^n)$ è una funzione $C^p(\mathbb{R}^n)$ tale che esista un $r > 0$ per cui $f(x) = 0$ se $|x| \geq r$. In tal caso $\text{supp}(f) \subseteq B_r(0)$.

che abbia supporto contenuto³³ in un aperto A tale che $A \cap \mathcal{S} = \varphi(U)$ per una qualche carta (φ, U) di un fissato atlante orientato $\mathcal{A} := \{\varphi^{(i)}, U^{(i)}\}_{i \in J}$. Sia Or l'orientamento indotto dall'atlante \mathcal{A} ossia $\text{Or} := \text{Or}(\varphi^{(i)})$ con $(\varphi^{(i)}, U^{(i)}) \in \mathcal{A}$. In tal caso si definisce

$$\int_{\mathcal{S}, \text{Or}} \omega := \int_U \varphi^*(\omega) := \sum_{i_1 < \dots < i_k} \int_U g_{i_1 \dots i_k} \circ \varphi \det \frac{\partial(\varphi_{i_1}, \dots, \varphi_{i_k})}{\partial u} du. \quad (\text{C.79})$$

Sia ora ω una arbitraria k -forma su un aperto A ; sia $\mathcal{S} \subseteq A$ una k -varietà con bordo e orientabile e Or un orientamento fissato. Fissiamo una partizione³⁴ dell'unità $\{f_j\}$ su \mathcal{S} subordinata a $\{A_i\}_{i \in J}$ dove gli A_i sono aperti tali che $A_i \cap \mathcal{S} = \varphi^{(i)}(U^{(i)})$. Allora, per ogni j esiste un $i_j \in J$ tale che $\text{supp}(f_j) \subseteq A_{i_j}$. Possiamo allora definire le seguenti N forme a supporto compatto

$$\omega_j := f_j \omega. \quad (\text{C.80})$$

Dalle proprietà della partizione dell'unità segue che

$$\text{supp}(\omega_j) \subseteq A_{i_j} \quad (\text{C.81})$$

e che

$$\sum_{j=1}^N \omega_j = \omega \quad \text{su } \mathcal{S}. \quad (\text{C.82})$$

Si definisce allora

$$\int_{\mathcal{S}, \text{Or}} \omega := \sum_{j=1}^N \int_{\mathcal{S}, \text{Or}(\varphi^{(i_j)})} \omega_j. \quad (\text{C.83})$$

Come già fatto ometteremo il simbolo Or nella notazione tutte le volte che tale omissione non produca ambiguità. Non è difficile verificare che *tale definizione non dipende né dall'atlante orientato fissato (purché l'orientamento su \mathcal{S} sia tenuto fisso) né dalla partizione dell'unità scelta*. Da tale definizione segue anche immediatamente che se α e ω sono k -forme su A e se a e b sono due costanti

$$\int_{\mathcal{S}, \text{Or}} (a\alpha + b\omega) = a \int_{\mathcal{S}, \text{Or}} \alpha + b \int_{\mathcal{S}, \text{Or}} \omega, \quad (\text{C.84})$$

ed infine, se $\overline{\text{Or}}$ è l'orientamento opposto a Or

$$\int_{\mathcal{S}, \overline{\text{Or}}} \omega = - \int_{\mathcal{S}, \text{Or}} \omega. \quad (\text{C.85})$$

C.8 Formula di Stokes

Il seguente teorema dà una generalizzazione (multidimensionale) del Teorema fondamentale del calcolo che lega l'integrale della derivata di una funzione ai valori di tale funzione sugli estremi (cioè sul 'bordo') del dominio di integrazione.

Teorema C.34 (Formula di Stokes) *Sia \mathcal{S} una k -varietà compatta, orientabile e con bordo e sia ω una $(k-1)$ -forma su $A \supseteq \mathcal{S}$. Allora*

$$\int_{\mathcal{S}} d\omega = (-1)^k \int_{\partial \mathcal{S}} \omega, \quad (\text{C.86})$$

dove gli orientamenti su \mathcal{S} e su $\partial \mathcal{S}$ sono compatibili³⁵.

³³Una k -forma $\omega = \sum_{i_1 < \dots < i_k} g_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}$ si dice **a supporto compatto** se tutte le $g_{i_1 \dots i_k}$ sono funzioni a supporto compatto; il supporto di ω , $\text{supp}(\omega)$, è, allora, dato dall'unione dei $\text{supp}(g_{i_1 \dots i_k})$ al variare di $i_1 < \dots < i_k$. Dunque dire che $\text{supp}(\omega) \subseteq A$ con $A \subseteq \mathbb{R}^n$ significa dire che $g_{i_1 \dots i_k}(x) = 0$ per ogni $x \notin A$ e per ogni $i_1 < \dots < i_k$.

³⁴Si ricordi il Lemma 4.64.

³⁵Ovvero l'orientamento della mappa f , definita per \mathcal{S} , ma ristretta a $\partial \mathcal{S}$, è lo stesso della mappa g definita per $\partial \mathcal{S}$.

In particolare, se $\partial\mathcal{S} = \emptyset$, segue da tale teorema che $\int_{\mathcal{S}} d\omega = 0$ per ogni $(k-1)$ -forma ω .

Dimostrazione Fissiamo un atlante orientato $\mathcal{A} := \{(\varphi^{(i)}, U^{(i)})\}_{i \in J}$ ed osserviamo che possiamo assumere che ω abbia supporto compatto e che $\text{supp}(\omega) \subseteq A$ dove A è un aperto tale che $A \cap \mathcal{S} = \varphi(U)$ per una qualche carta $(\varphi, U) \in \mathcal{S}$. Infatti, come è stato fatto per definire l'integrale di una forma su di una varietà con bordo, è possibile decomporre ω in

$$\omega = \sum_{j=1}^N \omega_j \quad (\text{C.87})$$

con $\text{supp}(\omega_j) \subseteq A_{i_j}$ per una qualche i_j e quindi, se il teorema vale per le suddetti forme, otteniamo

$$\begin{aligned} \int_{\mathcal{S}} d\omega &= \int_{\mathcal{S}} \sum_{j=1}^N d\omega_j = \sum_{j=1}^N \int_{\mathcal{S}} d\omega_j \\ &= (-1)^k \sum_{j=1}^N \int_{\partial\mathcal{S}} \omega_j = (-1)^k \int_{\partial\mathcal{S}} \sum_{j=1}^N \omega_j \\ &= (-1)^k \int_{\partial\mathcal{S}} \omega. \end{aligned} \quad (\text{C.88})$$

Assumiamo dunque che $\text{supp}(\omega) \subseteq A$ con $\varphi(U) = A \cap \mathcal{S}$ e $(\varphi, U) \in \mathcal{A}$. Vi sono due casi:

$$(i) \quad \varphi(U) \cap \partial\mathcal{S} = \emptyset, \quad (ii) \quad \varphi(U) \cap \partial\mathcal{S} \neq \emptyset. \quad (\text{C.89})$$

Consideriamo (i). Chiaramente, in tal caso,

$$\int_{\partial\mathcal{S}} \omega = 0. \quad (\text{C.90})$$

Ricordando le proprietà del ‘pull-back’, si ottiene

$$\int_{\mathcal{S}} d\omega = \int_U \varphi^*(d\omega) = \int_U d(\varphi^*\omega). \quad (\text{C.91})$$

La forma $\varphi^*\omega$ è una $(k-1)$ -forma su $U \subseteq \mathbb{R}^k$, quindi $\varphi^*\omega = \alpha_f^{k-1}$ per un qualche campo vettoriale $f \in C^1(U, \mathbb{R}^k)$ ossia

$$\varphi^*\omega = \sum_{i=1}^k (-1)^{i-1} f_i(u) du_1 \wedge \cdots \wedge \widehat{du}_i \wedge \cdots \wedge du_k \quad (\text{C.92})$$

e, quindi,

$$d(\varphi^*\omega) = \nabla \cdot f du_1 \wedge \cdots \wedge du_k. \quad (\text{C.93})$$

Da (C.91) e (C.92) segue allora che

$$\int_{\mathcal{S}} d\omega = \sum_{i=1}^k \int_U \frac{\partial f_i}{\partial u_i} du. \quad (\text{C.94})$$

Ma essendo ω a supporto compatto anche le funzioni f_i hanno supporto compatto e se $R := [a_1, b_1] \times \cdots \times [a_k, b_k]$ è un rettangolo che contiene al suo interno tutti i supporti delle f_i , dal Teorema fondamentale del calcolo segue che, per ogni $i = 1, \dots, k$,

$$\int_{a_i}^{b_i} \frac{\partial f_i}{\partial u_i}(u) du_i = f_i(u_1, \dots, b_i, \dots, u_k) - f_i(u_1, \dots, a_i, \dots, u_k) = 0. \quad (\text{C.95})$$

Questo dimostra la validità dell’asserto nel caso (i).

Consideriamo (ii). Anche in questo caso vale (C.94) e, ragionando come sopra, si vede che anche (C.95) vale se $1 \leq i \leq k-1$. L'integrale in u_k va trattato in modo diverso, poiché sebbene anche in questo caso f_k sia nulla quando u_k è esterno all'intervallo $[a_k, b_k]$, il dominio U ha la forma $V \cap H^k$ con V aperto di \mathbb{R}^k con intersezione non nulla con l'iperpiano $\{u_k = 0\}$ quindi integrando in u_k si ottiene

$$\int_0^{b_k} \frac{\partial f_k}{\partial u_k} du_k = -f_k(u_1, \dots, u_{k-1}, 0) := -f_k(\hat{u}, 0). \quad (\text{C.96})$$

Quindi, ricordando le notazioni usate nella sezione § C.1, si ha

$$\int_{\mathcal{S}} d\omega = - \int_{\hat{U}} f_k(\hat{u}, 0) d\hat{u}, \quad (\text{C.97})$$

dove $\hat{U} := U \cap \{u_k = 0\}$. D'altra parte, essendo l'orientamento su $\partial\mathcal{S}$ l'orientamento compatibile,

$$\begin{aligned} \int_{\partial\mathcal{S}} \omega &= \int_{\hat{U}} \hat{\varphi}^* \omega \\ &= \sum_{i=1}^k (-1)^{i-1} \int_{\hat{U}} f_i(\hat{u}, 0) \cdot du_1 \wedge \dots \wedge \widehat{du_i} \wedge \dots \wedge du_k \end{aligned} \quad (\text{C.98})$$

Si osservi ora che, per ogni $i < k$, la matrice

$$\frac{\partial(u_1, \dots, u_{i-1}, u_{i+1}, \dots, u_k)}{\partial(u_1, \dots, u_{k-1})}$$

ha l'ultima riga nulla e dunque ha determinante nullo. Allora

$$\int_{\hat{U}} f_i(\hat{u}, 0) du_1 \wedge \dots \wedge \widehat{du_i} \wedge \dots \wedge du_k = 0, \quad (1 \leq i \leq k-1), \quad (\text{C.99})$$

il che implica

$$\int_{\partial\mathcal{S}} \omega = (-1)^{k-1} \int_{\hat{U}} f_k(\hat{u}, 0) d\hat{u}. \quad (\text{C.100})$$

Tale relazione, assieme a (C.97), implica l'asserto anche nel caso (ii). \blacksquare

Complementi

Complemento C.1: Retta tangente a una curva in \mathbb{R}^n

Sia $\Gamma := \varphi(I)$ un elemento di curva regolare in \mathbb{R}^n ($I = (a, b)$ e φ è una inclusione di $[a, b]$ in \mathbb{R}^n). Ricordiamo che una qualunque retta in \mathbb{R}^n passante per un punto x_0 è data da

$$\ell := \{x \in \mathbb{R}^n : x = x_0 + t\xi \text{ con } t \in \mathbb{R}\}, \quad (\text{C.101})$$

dove ξ è un vettore non nullo di \mathbb{R}^n . Non è restrittivo assumere che³⁶ $|\xi| = 1$ cosa che d'ora in avanti faremo. Ricordiamo anche che dato un punto $x \in \mathbb{R}^n$ ed un sottoinsieme $X \subseteq \mathbb{R}^n$ si definisce la *distanza di x da X* (rispetto alla norma euclidea $|\cdot|$) come

$$\text{dist}(x, X) := \inf_{y \in X} |x - y|. \quad (\text{C.102})$$

Si ha allora la seguente semplice

³⁶Basta infatti cambiare il parametro t con $s = |\xi|t$.

Proposizione C.35 Dato un punto $x \in \mathbb{R}^n$ ed una retta passante per $x_0 \in \mathbb{R}^n$, $\ell := \{x_0 + t\xi\}$ (con $|\xi| = 1$), si consideri il seguente punto su ℓ

$$p_\ell(x) := x_0 + (x - x_0) \cdot \xi \xi. \quad (\text{C.103})$$

Allora $p_\ell(x)$ verifica

$$\text{dist}(x, \ell) = |x - p_\ell(x)| \quad (\text{C.104})$$

ed è l'unico punto di ℓ per cui (C.104) valga.

Il punto $p_\ell(x)$ si chiama **proiezione di x su ℓ** .

Dimostrazione Per trovare la distanza di x da ℓ bisogna minimizzare la funzione $t \rightarrow |x - x_0 - t\xi|$ o, equivalentemente, la funzione $|x - x_0 - t\xi|^2$. Tale quantità è uguale a

$$|x - x_0|^2 + t^2 - 2t(x - x_0) \cdot \xi \quad (\text{C.105})$$

che raggiunge il suo minimo per l'unico valore $t = (x - x_0) \cdot \xi$. ■

Definizione C.36 Sia $x_0 \in \Gamma$ elemento di curva regolare in \mathbb{R}^n . Una retta ℓ passante per x_0 si dice *tangente a Γ* (nel punto x_0) se per ogni successione di punti $x^{(j)} \in \Gamma$, $x^{(j)} \neq x_0$, tali che $\lim_{j \rightarrow \infty} x^{(j)} = x_0$ si ha

$$\lim_{j \rightarrow \infty} \frac{\text{dist}(x^{(j)}, \ell)}{|x^{(j)} - x_0|} = 0. \quad (\text{C.106})$$

Si noti che la quantità $\frac{\text{dist}(x^{(j)}, \ell)}{|x^{(j)} - x_0|}$ non è altro che il seno dell'angolo tra la retta ℓ e la retta passante per i punti x_0 e $x^{(j)}$.

Proposizione C.37 Sia $\Gamma := \varphi((a, b))$ un elemento di curva regolare³⁷ in \mathbb{R}^n e sia $x_0 \in \Gamma$. Allora esiste un'unica retta tangente a Γ in x_0 data da

$$\ell := \{x = \varphi(t_0) + s\tau : s \in \mathbb{R}\}, \quad \text{dove} \quad \tau := \frac{\varphi'(t_0)}{|\varphi'(t_0)|}. \quad (\text{C.107})$$

Il vettore τ definito in (C.107) si chiama *versore tangente* alla curva Γ nel punto x_0 mentre $\varphi'(t_0)$ è un **vettore tangente** a Γ in x_0 .

Osservazione C.38 Il versore tangente τ definito in (C.107) è definito a meno del segno³⁸. Il verso scelto in (C.107) fissa un verso (cioè un 'orientamento') sulla retta tangente che corrisponde al verso in cui $\varphi(t)$ 'percorre' l'elemento Γ .

Dimostrazione (della Proposizione C.37) Assumiamo, per semplicità di notazione che $t_0 = 0$ e che $x_0 = 0$ (situazione che può sempre essere raggiunta tramite traslazioni in t e x). Dimostriamo prima che ℓ è tangente a Γ in $x_0 = 0$. Sia $x^{(j)} \in \Gamma$ una successione convergente a 0 (con $x^{(j)} \neq 0$); questo equivale a dire che $x^{(j)} = \varphi(t_j)$ con $t_j \rightarrow 0$, ($t_j \neq 0$), e $\varphi(0) = 0$. Dalla Proposizione C.35 si ha che

$$\frac{\text{dist}(x^{(j)}, \ell)}{|x^{(j)}|} = \frac{|\varphi(t_j) - \varphi(t_j) \cdot \tau \tau|}{|\varphi(t_j)|}. \quad (\text{C.108})$$

D'altra parte per la formula di Taylor al prim'ordine si ha che esiste una funzione (vettoriale) $\delta(t)$ tale che

$$\varphi(t) = \varphi'(0)t + \delta(t), \quad \text{con} \quad \lim_{t \rightarrow 0} \frac{|\delta(t)|}{t} = 0. \quad (\text{C.109})$$

Inoltre, dalla definizione di τ , segue che

$$\varphi'(0) \cdot \tau \tau = \varphi'(0). \quad (\text{C.110})$$

Sostituendo (C.109) in (C.108) ed usando (C.110) otteniamo

$$\frac{\text{dist}(x^{(j)}, \ell)}{|x^{(j)}|} = \frac{|\delta(t_j) - \delta(t_j) \cdot \tau \tau|}{|t_j| |\varphi'(0) + \frac{\delta(t_j)}{t_j}|} \quad (\text{C.111})$$

³⁷ Come al solito φ è una inclusione dell'intervallo $[a, b]$ in \mathbb{R}^n .

³⁸ In effetti, la Proposizione C.37 asserisce l'unicità della retta tangente e ad una retta corrispondono *due* versori.

e tale quantità tende a 0, quando $j \rightarrow \infty$, in virtù di (C.109)(ed essendo $|\varphi'(0)| \neq 0$).

Dimostriamo ora l'unicità della retta tangente. Supponiamo che esista un'altra retta tangente $\ell' := \{x_0 + s\tau' : s \in \mathbb{R}\}$ con $|\tau'| = 1$ e $\tau' \neq \pm\tau$ (il che equivale a dire che le rette ℓ e ℓ' sono distinte). Sia $x^{(j)} \neq 0$ una successione di punti in Γ che tendono a 0. Essendo le due rette tangenti a Γ si ha che

$$\lim_{j \rightarrow \infty} \frac{|x^{(j)} - x^{(j)} \cdot \tau \tau|}{|x^{(j)}|} = \lim_{j \rightarrow \infty} \frac{|x^{(j)} - x^{(j)} \cdot \tau' \tau'|}{|x^{(j)}|} = 0. \quad (\text{C.112})$$

Sia $y^{(j)} := \frac{x^{(j)}}{|x^{(j)}|}$. Tale successione appartiene a $S^{n-1} := \{x \in \mathbb{R}^n : |x| = 1\}$. Tale insieme è compatto, quindi esiste una sottosuccessione $y^{(j_k)}$ convergente ad un punto \bar{y} di S^{n-1} . Inserendo tale sottosuccessione in (C.112) e prendendo il limite per $k \rightarrow \infty$, otteniamo che

$$\bar{y} - \bar{y} \cdot \tau \tau = 0 = \bar{y} - \bar{y} \cdot \tau' \tau' \quad (\text{C.113})$$

ossia $\bar{y} = \bar{y} \cdot \tau \tau = \bar{y} \cdot \tau' \tau'$. Prendendo la norma di tali relazioni e ricordando che $|\bar{y}| = 1$, vediamo che $\bar{y} \cdot \tau = \pm 1$ e $\bar{y} \cdot \tau' = \pm 1$, il che implica $\tau = \pm\tau'$ (essendo $\bar{y} \cdot \tau \tau = \bar{y} \cdot \tau' \tau'$) producendo una contraddizione. ■

Complemento C.2: Spazio tangente ad una varietà

Sia $\mathcal{S} = \varphi(U)$ un elemento di k -varietà in \mathbb{R}^n e $x_0 := \varphi(u_0) \in \mathcal{S}$.

Definizione C.39 Si chiama spazio tangente a \mathcal{S} in x_0 , e si denota $T\mathcal{S}_{x_0}$, l'insieme dei vettori $\xi \in \mathbb{R}^n$ tali che $\xi = z'(t_0)$ per una qualche $z : (a, b) \rightarrow \mathcal{S}$ di classe C^1 tale che $z(t_0) = x_0$ con $a < t_0 < b$.

Si noti che se $z'(t_0) \neq 0$, l'immagine di un intorno sufficientemente piccolo di t_0 è un elemento di curva regolare in \mathcal{S} passante per x_0 ; d'altra parte prendendo $z : t \in (0, 1) \rightarrow z(t) := x_0 \in \mathcal{S}$ si vede che $0 \in T\mathcal{S}_{x_0}$. Per ogni $x_0 = \varphi(u_0) \in \mathcal{S} = \varphi(U)$ passano k -curve speciali, dette **curve coordinate**, date da

$$\gamma^{(i)}(t) := \varphi(u_0 + te^{(i)}) := \varphi(u_{01}, \dots, u_{0i} + t, \dots, u_{0k}) \quad (\text{C.114})$$

dove $\{e^{(i)}\}$ è la base standard in \mathbb{R}^k e t varia in $[-\delta, \delta]$ per δ sufficientemente piccolo³⁹. Quindi i k vettori

$$\xi^{(i)} := (\gamma^{(i)})'(0) = \frac{\partial \varphi}{\partial u_i}(u_0) \quad (\text{C.115})$$

appartengono a $T\mathcal{S}_{x_0}$. Infatti, vale la seguente

Proposizione C.40 Lo spazio $T\mathcal{S}_{x_0}$ è un sottospazio vettoriale di dimensione k di \mathbb{R}^n ed una base di tale spazio vettoriale è data dai k vettori $\xi^{(i)}$, $i = 1, \dots, k$ definiti in (C.115).

Dimostrazione Tramite una traslazione ($t = t_0 + s$) possiamo assumere che gli elementi di $T\mathcal{S}_{x_0}$ siano realizzati da applicazioni differenziabili z definite in un intorno di 0 tali che $z(0) = x_0$. Dato $\xi \in T\mathcal{S}_{x_0}$, facciamo vedere che ξ può scriversi come combinazione lineare dei vettori $\xi^{(i)}$. Sia $\xi = z'(0)$. Definiamo $v(t) := \varphi^{-1} \circ z(t)$. Tale funzione (per il punto (vii) dell'Osservazione C.3) appartiene a $C^1(\{0\}, \mathbb{R}^k)$ ed inoltre $\varphi(v(t)) = z(t)$ e $u_0 := v(0)$. Derivando $\varphi(v(t)) = z(t)$ si ottiene che, per ogni $j = 1, \dots, n$,

$$\sum_{i=1}^k \frac{\partial \varphi_j}{\partial u_i}(u_0) v_i'(0) = z_j'(0), \quad (\text{C.116})$$

cioè $\xi = \sum_{i=1}^k a_i \xi^{(i)}$ con $a_i := v_i'(0)$. Mostriamo che $T\mathcal{S}_{x_0}$ è uno spazio vettoriale. Se $\xi := z'(0) \in T\mathcal{S}_{x_0}$ e $a \in \mathbb{R}$ allora $a\xi = w'(0)$ dove $w(t) := z(at)$ e quindi $a\xi \in T\mathcal{S}_{x_0}$. Se $\xi, \eta \in T\mathcal{S}_{x_0}$ per quanto visto precedentemente si ha che $\xi = \sum_{i=1}^k a_i \xi^{(i)}$ e $\eta = \sum_{i=1}^k b_i \xi^{(i)}$ e quindi $\xi + \eta = \frac{d}{dt} \Big|_{t=0} \varphi(u_{01} + (a_1 + b_1)t, \dots, u_{0k} + (a_k + b_k)t)$ ossia $\xi + \eta \in T\mathcal{S}_{x_0}$. Infine che i vettori $\xi^{(i)}$ siano linearmente indipendenti deriva immediatamente dalla definizione di inclusione (in particolare dal fatto che lo jacobiano $\frac{\partial \varphi}{\partial u}(u_0)$ ha rango massimo). ■

Si noti che se $k = 1$, $T\mathcal{S}_{x_0}$ non è la retta ℓ tangente definita in C.1 ma è la retta parallela a ℓ passante per l'origine.

³⁹ 'Sufficientemente piccolo' significa che $\gamma^{(i)}([-\delta, \delta])$ è contenuto in \mathcal{S} .

Complemento C.3: Volumi di parallelepipedi k -dimensionali in \mathbb{R}^n ($k \leq n$)

Siano $v^{(1)}, \dots, v^{(k)}$ k vettori in \mathbb{R}^n con $n \geq k$; sia $\Pi_0 := \Pi(v^{(1)}, \dots, v^{(k)})$ il parallelepipedo generato da $v^{(1)}, \dots, v^{(k)}$. Lo scopo di questo paragrafo è di definire il volume (la misura) k -dimensionale di Π_0 (per le notazioni si veda § 4). Tale definizione è basata sui due seguenti lemmi (la cui dimostrazione viene lasciata per esercizio). Sia $V := V(v^{(1)}, \dots, v^{(k)})$ lo spazio vettoriale generato da $v^{(1)}, \dots, v^{(k)}$.

Lemma C.41 Siano $v^{(1)}, \dots, v^{(k)}$ vettori indipendenti in \mathbb{R}^n con $n \geq k$; sia $B := \{w^{(1)}, \dots, w^{(k)}\}$ una base ortonormale in V . Denotiamo con $u^{(i)} \in \mathbb{R}^k$ il vettore delle coordinate di $v^{(i)}$ rispetto a B (cioè $u^{(i)} := (u_1^{(i)}, \dots, u_k^{(i)})$ e $v^{(i)} = \sum_{j=1}^k u_j^{(i)} w^{(j)}$), e denotiamo con U e A le matrici $U := [u^{(1)}, \dots, u^{(k)}]$ e $A := [v^{(1)}, \dots, v^{(k)}]$; $U \in \mathbb{R}^{k \times k}$ e $A \in \mathbb{R}^{n \times k}$. Allora $U^T U = A^T A$.

Lemma C.42 Siano $k, n, v^{(i)}$ ed A come nel Lemma precedente e sia $\varphi : V \rightarrow \mathbb{R}^d$ un isomorfismo lineare che conservi il prodotto scalare. Allora

$$\text{mis}_k \left(\varphi(\Pi[v^{(1)}, \dots, v^{(k)}]) \right) = \text{mis}_k \left(\Pi[\varphi(v^{(1)}), \dots, \varphi(v^{(k)})] \right) = \sqrt{A^T A}. \quad (\text{C.117})$$

Definizione C.43 Se $v^{(1)}, \dots, v^{(k)}$ sono vettori in \mathbb{R}^n con $n \geq k$ e se $\Pi_0 := \Pi(v^{(1)}, \dots, v^{(k)})$ denota il parallelepipedo generato in \mathbb{R}^n dai vettori $v^{(1)}, \dots, v^{(k)}$, si definisce la misura euclidea k -dimensionale di Π_0 come

$$\text{mis}_k(\Pi_0) := \text{mis}_k \left(\varphi(\Pi[v^{(1)}, \dots, v^{(k)}]) \right), \quad (\text{C.118})$$

dove $\varphi : V \rightarrow \mathbb{R}^d$ è un qualunque isomorfismo lineare che conservi il prodotto scalare.

Osservazione C.44 (i) Un isomorfismo che conservi il prodotto scalare conserva anche gli angoli e le lunghezze e quindi conserva la ‘geometria euclidea’.

(ii) In particolare da (C.117) mostra che la misura k -dimensionale di $\varphi(\Pi_0)$ non dipende dal particolare isomorfismo φ (purchè, naturalmente φ conservi il prodotto scalare). Dunque la Definizione C.43 è ben posta.

(iii) Per una nota formula di algebra lineare⁴⁰ si ha che se A è una matrice $n \times k$ con $k \leq n$ allora il determinante di $A^T A$ è uguale alla somma dei quadrati dei determinanti di tutti i minori di rango massimo (ossia k). Dunque l’oggetto formale ‘ $d\sigma_k$ ’ che appare in (C.12) (e cioè nella definizione della misura di una varietà k -dimensionale in \mathbb{R}^n) può essere interpretata come la misura k -dimensionale del parallelepipedo ‘infinitesimo’ generato dai vettori $\frac{\partial \varphi}{\partial u_i} du_i$ (e si ricordi che i vettori $\frac{\partial \varphi}{\partial u_1}, \dots, \frac{\partial \varphi}{\partial u_k}$ generano lo spazio tangente alla varietà).

Esercizi

Esercizio C.1 Se qualcuna delle affermazioni fatte in (E1)÷(E7) di § C.1 non fosse completamente chiara, la si dimostri con cura!

Esercizio C.2 Dire se le seguenti funzioni definiscono delle inclusioni differenziabili e, in caso affermativo, si disegni la traccia dei relativi elementi di curva e se ne calcoli la lunghezza:

- (i) $\varphi(t) = r(t - \sin t, 1 - \cos t)$, $t \in [\varepsilon, 1 - \varepsilon]$, ($r > 0, 1 > \varepsilon > 0$),
(ii) $\varphi(t) = t(\cos t, \sin t)$, $t \in [a, b]$, ($0 < a < b$),

(iii) φ come in (i) ma con $t \in [0, 1]$.

[L’elemento di curva in (i) prende il nome di ‘elemento di cicloide’; quello in (ii) ‘elemento di spirale’.]

Esercizio C.3 Trovare un’inclusione, nell’intorno di $x_0 = (1, 10, 2, 0)$, che realizzi $\{\phi = 0\}$ con

$$\phi(x) = x_1^2 + x_2^5(x_3 - 2) + \cos x_4 - 2.$$

Esercizio C.4 Si diano i dettagli della dimostrazione relativa all’affermazione fatta nel punto (i) dell’Osservazione C.7.

⁴⁰Tale formula è una generalizzazione della formula per il determinante del prodotto di due matrici quadrate ed in alcuni testi è nota come ‘Teorema di Binet’; si veda, per esempio, ‘E. Sernesi Geometria I, ed. Boringhieri.

Esercizio C.5 Si dimostri che se \mathcal{S} è una n -varietà con bordo in \mathbb{R}^n allora il suo bordo coincide con la sua frontiera (nella topologia usuale di \mathbb{R}^n).

Esercizio C.6 Si dimostri che le varietà descritte in (S3) e (S4) hanno bordo vuoto.

Esercizio C.7 Sia $\mathcal{S} := \varphi(U)$ un elemento di k -varietà in \mathbb{R}^n . Si trovi una inclusione (ψ, V) tale che $\mathcal{S} = \psi(V)$ e tale che V sia un aperto di \mathbb{R}^k contenuto in $\{u_k > 0\}$.

[Risposta nel caso in cui $U \cap \{u_k \leq 0\} \neq \emptyset$: Sia $m := \inf_{u \in U} u_k$ e sia $u_0 := (0, \dots, 0, 1 - m)$; si può allora prendere $V := u_0 + U$ e $\psi(v) := \varphi(v - u_0)$.]

Esercizio C.8 Sia $\mathcal{S} = \{\phi \geq 0\}$ un insieme regolare in \mathbb{R}^n (si veda (S7)) e sia ν la normale esterna in $x_0 \in \partial\mathcal{S}$. Si dimostri che ‘muovendosi’ da x_0 nella direzione ν si ‘esce’ da \mathcal{S} , mentre ‘muovendosi’ nella direzione opposta si ‘entra’ in \mathcal{S} , più precisamente, si dimostri che esiste $\varepsilon > 0$ tale che

$$\{x_0 + \nu t : 0 < t < \varepsilon\} \subseteq \mathbb{R}^n \setminus \mathcal{S}, \quad \{x_0 - \nu t : 0 < t < \varepsilon\} \subseteq \mathcal{S}. \quad (\text{C.119})$$

Esercizio C.9 (La funzione Γ di Eulero) Sia $\Gamma(z) := \int_0^\infty t^{z-1} e^{-t} dt$. Si dimostrino le seguenti proprietà:

$$\begin{aligned} \Gamma(z+1) &= z\Gamma(z), \quad (\forall z); & \Gamma(1/2) &= \sqrt{\pi}; \\ \Gamma(n+1) &= n!, \quad (n \in \mathbb{N}); & \Gamma(n + \frac{1}{2}) &= \frac{(2n-1)!!}{2^n} \sqrt{\pi}, \quad (n \in \mathbb{N}). \end{aligned}$$

Da queste relazioni e da (4.200) segue che il volume della sfera unitaria in \mathbb{R}^n è dato da

$$\Omega_n = \begin{cases} \frac{\pi^{\frac{n}{2}}}{(n/2)!}, & \text{se } n \text{ è pari,} \\ \frac{2(2\pi)^{\frac{n-1}{2}}}{n!}, & \text{se } n \text{ è dispari.} \end{cases}$$

Esercizio C.10 Sia $\Gamma \subseteq \mathbb{R}^2$ l’unione dei due segmenti $\{x \in \mathbb{R}^2 : x_2 = 0, 0 \leq x_1 < 1\}$ e $\{x \in \mathbb{R}^2 : x_1 = 0, 0 \leq x_2 < 1\}$. Si dimostri che Γ è una 1-varietà regolare a tratti ma non è un elemento di 1-varietà regolare in \mathbb{R}^2 .

Esercizio C.11 Verificare che i vettori tangenti alla cicloide (C.2) nei punti $(x, y) = \varphi(n\pi)$ sono paralleli all’asse delle x .

Esercizio C.12 (Orientamento su superfici) Fissare una particolare inclusione φ che realizza \mathcal{S} , elemento di superficie, equivale a fissare un **orientamento su \mathcal{S}** : si definisce $\text{Or}(\varphi)$ la classe di equivalenza costituita da tutte le inclusioni $\psi : \bar{V} \rightarrow \mathbb{R}^3$ con $\psi(V) = \mathcal{S}$ tali che $\det \frac{\partial \psi}{\partial u} > 0$, dove $v(u) := \psi^{-1} \circ \varphi(u)$. Dimostrare che l’orientamento $\text{Or} = \text{Or}(\varphi)$ è una classe di equivalenza.

Esercizio C.13 Sia $\text{Or} = \text{Or}(\varphi)$ un orientamento fissato su $\mathcal{S} = \varphi(U)$. Trovare l’orientamento opposto.

Esercizio C.14 Dimostrare che su ogni elemento di superficie $\mathcal{S} = \varphi(U)$ si hanno due soli orientamenti, ossia ogni inclusione ψ che realizza \mathcal{S} o appartiene a $\text{Or}(\varphi)$ o appartiene a $\text{Or}(\bar{\varphi}) = -\text{Or}(\varphi)$.

Esercizio C.15 (Orientamento su elementi di k -varietà) Le definizioni date nel caso di curve e superfici in \mathbb{R}^3 si estendono immediatamente al caso generale. Sia $\mathcal{S} = \varphi(U)$ un elemento di k -superficie in \mathbb{R}^n . Si definisce l’**orientamento** indotto da φ su \mathcal{S} la classe di equivalenza costituita da tutte le inclusioni $\psi : \bar{V} : \mathbb{R}^n$ con $\psi(V) = \mathcal{S}$ tali che $\det \frac{\partial \psi}{\partial v} > 0$ dove $v(u) := \psi^{-1} \circ \varphi(u)$. Si generalizzino gli esercizi C.12, C.13 e C.14.

Esercizio C.16 (i) Si trovi una inclusione $\varphi : U \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}^3$ tale che $\varphi(U) = S^2 \cap \{x_3 > \varepsilon\} := \{x \in \mathbb{R}^3 : |x|^2 = 1 \text{ e } x_3 > \varepsilon\}$ ($0 < \varepsilon < 1$)

(ii) Si determinino gli spazi tangenti a $\varphi(U)$ (φ come in (i) con $\varepsilon = \frac{1}{3}$) nei punti $x_0 = (0, 0, 1)$ e $(\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}})$.

Esercizio C.17 (i) Se \mathcal{S} è un elemento di $(n-1)$ -superficie in \mathbb{R}^n dato da $\{\phi = 0\} \cap B$ con $\phi \in C^1(\bar{B}, \mathbb{R})$ e B sfera aperta in \mathbb{R}^n , si dimostri che

$$\nabla\phi(x) \perp T\mathcal{S}_x \quad (\text{C.120})$$

per ogni $x \in \mathcal{S}$.

(ii) Si trovino i versori (cioè ‘i vettori di norma 1’) normali a \mathcal{S} in x .

Esercizio C.18 Siano

$$g : x \in A \subseteq \mathbb{R}^n \rightarrow g(x) \in \mathbb{R}, \quad f : y \in B \subseteq \mathbb{R}^m \rightarrow (f_1(y), \dots, f_n(y)) \in A,$$

funzioni C^1 sui rispettivi domini. Dimostrare che

$$d(g \circ f) = \sum_{i=1}^n \left(\frac{\partial g}{\partial x_i} \circ f \right) df_i. \quad (\text{C.121})$$

Esercizio C.19 (i) Si trovino tutti i vettori $\xi \in \mathbb{R}^2$ tali che $(dx_1 + dx_2)(\xi) = 0$.

(ii) Sia $\omega^1 = \omega_F^1$ una 1-forma in $A \subseteq \mathbb{R}^n$ e si fissi $x_0 \in A$. Si dimostri che $\{\xi \in \mathbb{R}^n : \omega_{F, x_0}^1(\xi) = 0\}$ è uno spazio vettoriale. Qual è la dimensione di tale spazio vettoriale?

(iii) Sia $F(x, y) := (x, y)$. Si trovi una curva $\Gamma = \varphi((0, 1))$ in \mathbb{R}^2 tale che $\omega_{F, \varphi}^1(\varphi') = 0$ per ogni $t \in (0, 1)$.

Esercizio C.20 Si dimostri (4.104).

Esercizio C.21 Sia $A \subseteq \mathbb{R}^n$ un dominio stellato che non contenga l'origine e sia f una funzione continua su $[0, \infty)$. Si dica se $\omega^1 := f(|x|) \sum_{i=1}^n x_i dx_i$ è esatta su A ed in caso affermativo se ne calcoli una primitiva.

Esercizio C.22 Si dimostri che una k -varietà regolare a tratti in \mathbb{R}^n con $k < n$ è un insieme trascurabile.

Esercizio C.23 Si dimostri l'affermazione fatta dopo (4.106).

Esercizio C.24 Siamo a_j , $j = 1, \dots, N$ numeri complessi. Si dimostri (per induzione su N) che vale

$$a_1 + (1 - a_1)a_2 + \dots + (1 - a_1) \dots (1 - a_{N-1})a_N = 1 - \prod_{j=1}^N (1 - a_j). \quad (\text{C.122})$$

Esercizio C.25 (Integrali curvilinei e superficiali su varietà) Analogamente a come si è fatto per definire l'integrale di una k -forma su di una k -varietà con bordo, è possibile definire gli integrali curvilinei e superficiali su, rispettivamente, 1 e 2-varietà. Se $F \in C(\mathcal{S}, \mathbb{R}^n)$ con $\mathcal{S} \subseteq \mathbb{R}^n$ k -varietà con bordo ($k = 1, 2$) e se poniamo $F_j := f_j F$ dove $\{f_j\}$ è una partizione dell'unità come sopra (e cioè subordinata ad un atlante di \mathcal{S}) definiamo

$$\begin{aligned} \int_{\mathcal{S}} F(s) ds &:= \sum_{j=1}^N \int_{\varphi(U^{i_j})} F_j(s) ds, & (k=1) \\ \int_{\mathcal{S}} F(\sigma) d\sigma &:= \sum_{j=1}^N \int_{\varphi(U^{i_j})} F_j(\sigma) d\sigma, & (k=2) \end{aligned} \quad (\text{C.123})$$

(si ricordi il punto (i) dell'Osservazione 4.12).

Esercizio C.26 Sia $f := 0$ su $[0, 1] \setminus \mathbb{Q}$ e $f(x) := 1/n$ se $x = m/n$ con $0 \leq m \leq n \leq 1$ (m ed n relativamente primi). (i) Dimostrare che l'insieme di discontinuità di f è $\mathbb{Q} \cap [0, 1]$. (ii) Dimostrare direttamente (senza usare il Teorema di Vitali–Lebesgue) che $f \in \mathcal{R}([0, 1])$.

Esercizio C.27 Sia $f \in S(E)$ e sia C il suo supporto. Si dimostri che per ogni $\varepsilon > 0$ esiste $g \in S(E)$ con $\text{supp}(g) \subseteq \overset{\circ}{C}$ tale che $\int_E |f - g| \leq \varepsilon$.

Esercizio C.28 Sia $B_r(x_0)$ la sfera (euclidea) aperta in \mathbb{R}^n centrata in x_0 e di raggio r . Dimostrare che esiste $\bar{x}_0 \in \mathbb{Q}^n$ e $0 < \bar{r} \in \mathbb{Q}$ tali che il cubo aperto di centro \bar{x}_0 e lato \bar{r} sia contenuto in $B_r(x_0)$ e contenga x_0 .

Esercizio C.29 Dare un esempio di insieme $Q \subseteq [0, 1]^2$ trascurabile (in \mathbb{R}^2) tale che per ogni $x, y \in \mathbb{Q} \cap [0, 1]$, Q_x e Q_y non sono trascurabile in \mathbb{R} .

Esercizio C.30 (tratto da M. Reed, B. Simon *Functional Analysis*) Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ così definita: $f = 1$ su $\{x > 0, y > 0, 0 \leq x - y \leq 1\}$, $f = -1$ su $\{x > 0, y > 0, 0 < y - x < 1\}$ e $f = 0$ altrimenti. Si calcolino gli integrali iterati di f e si spieghi il risultato.

Esercizio C.31 Per ogni $k \in \mathbb{N}$, sia $a_k := \sum_{j=1}^k \frac{1}{j}$. Se $\{x\}$ denota la funzione ‘parte frazionaria di x ’ [ossia $\{x\}$ è l’unico numero in $[0, 1)$ tale che $x = n + \{x\}$ con $n \in \mathbb{Z}$], sia $A_k := \{x = \{y\} : a_k \leq y \leq a_{k+1}\}$. Dimostrare che A_k è un insieme elementare e che

$$\int_0^1 \chi_{A_k} := m(A_k) = \frac{1}{k+1}, \quad \limsup_{k \rightarrow \infty} \chi_{A_k}(x) = 1, \quad \forall x \in [0, 1].$$

Esercizio C.32 Sia A_k come nell’esercizio precedente e $f_k := \chi_{A_k}$. Trovare (esplicitamente) una sottosuccessione f_{k_j} tale che $f_{k_j} \rightarrow 0$ q.o. in $[0, 1]$.

Esercizio C.33 Trovare una successione di funzioni a scalini e non negative in $[0, 1]$ tale che $\int_0^1 f_k = 1$ per ogni k e $\lim f_k(x) = 0$ per ogni $x \in [0, 1]$.

Esercizio C.34 (i) Si calcoli il valore della 2-forma $\cos(x_1 x_2^2) dx_1 \wedge dx_3 - dx_2 \wedge dx_3$ su $(\xi^{(1)}, \xi^{(2)})$ con $\xi^{(1)} = (1, 0, 2)$ e $\xi^{(2)} = (1, 1, 1)$ per $x \in \{x \in \mathbb{R}^3 : x_1 = \pi, x_2 = 1\}$.

(ii) Si trovino tutte le coppie di vettori ξ ed η in \mathbb{R}^3 tali che $(dx_1 \wedge dx_2)(\xi, \eta) = 0$.

Esercizio C.35 (i) Si dimostri che il prodotto esterno è associativo:

$$(\omega^k \wedge \alpha^h) \wedge \beta^\ell = \omega^k \wedge (\alpha^h \wedge \beta^\ell). \quad (\text{C.124})$$

(ii) Si dimostri che se f_1 e f_2 sono due funzioni

$$(f_1 \omega_1^k + f_2 \omega_2^k) \wedge \alpha^h = f_1 \omega_1^k \wedge \alpha^h + f_2 \omega_2^k \wedge \alpha^h. \quad (\text{C.125})$$

(iii) Si dimostri che

$$\omega^k \wedge \alpha^h = (-1)^{kh} \alpha^h \wedge \omega^k. \quad (\text{C.126})$$

Esercizio C.36 Si verifichi la seconda uguaglianza in (C.40).

Esercizio C.37 Sia $\omega^2 = f dx_1 \wedge dx_2 + g dx_1 \wedge dx_3$ una 2-forma su $A \subseteq \mathbb{R}^3$, si dimostri che $\omega^2 = 0$ se e solo se $f(x) = g(x) = 0$ per ogni $x \in A$.

[Suggerimento: $\omega^2 = 0$ su A significa che $\omega_x^2(\xi, \eta) = 0$ per ogni $x \in A$ e per ogni coppia di vettori ξ, η in \mathbb{R}^3 . Il ‘se’ è banale. Per dimostrare il ‘solo se’ si calcoli $\omega^2(\xi, \eta)$ con $\xi = (1, 0, 0)$ e $\eta = (0, 1, 0)$. Si calcoli, successivamente, $\omega^2(\xi, \eta)$ con $\xi = (1, 0, 0)$ e $\eta = (0, 0, 1)$.]

Esercizio C.38 Si generalizzi l’esercizio precedente. Si dimostri, cioè, che se

$$\omega^k = \sum_{i_1 < \dots < i_k} g_{i_1 \dots i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k}$$

è una k -forma su $A \subseteq \mathbb{R}^n$, $\omega^k = 0$ se e solo se $g_{i_1 \dots i_k} := 0$ su A per ogni scelta degli indici (i_1, \dots, i_k) .

[Suggerimento: si fissi una scelta di indici (i'_1, \dots, i'_k) (con $i'_1 < \dots < i'_k$) e si calcoli il valore di $\omega^k(\xi^{(1)}, \dots, \xi^{(k)})$ con $\xi^{(j)} := e^{(i'_j)}$, essendo come al solito $e^{(i)}$ la base standard in \mathbb{R}^n .]

Esercizio C.39 Si dimostri (C.38).

Esercizio C.40 Verificare (C.47), (C.50) e (C.51).

Esercizio C.41 Verificare che la 1-forma definita in (4.111) è chiusa (sul suo dominio di definizione).

Esercizio C.42 Si dimostri il Lemma di Poincaré nel caso $k = 1$, n arbitrario.

[Suggerimento: $\omega := \omega_F^1$, $\mathcal{I}\omega_F^1 = \sum_{i=1}^n \left(\int_0^1 F_i(tx) dt \right) x_i$, etc.]

Esercizio C.43 Si dimostri il Lemma di Poincaré nel caso generale $1 \leq k \leq n$.

Esercizio C.44 Si calcoli $\mathcal{I}\omega$ con $\omega := (x_2 + \cos x_1)dx_1 + x_1 dx_2$.

[Risposta: $x_1 x_2 + \sin x_1$.]

Esercizio C.45 Si verifichi che $d\omega_F^1 = 0$ se e solo se vale (C.67).

Appendice D

Algebra lineare

D.1 Spazi vettoriali

(i) Uno **spazio vettoriale** (o, anche, ‘spazio lineare’) V su un campo di scalari K è una coppia di insiemi dotati di due operazioni: $(x, y) \in V \times V \mapsto x + y \in V$ e $(a, x) \in K \times V \mapsto ax \in V$ tali che $(V, +, 0)$ è un *gruppo*¹ *abeliano*, $(K, +, 0, \cdot, 1)$ è un campo², e valgono le seguenti relazioni di compatibilità, $\forall a, b \in K, x, y \in V$:

$$a(x + y) = ax + by; \quad (a + b)x = ax + bx; \quad (a \cdot b)x = a(bx); \quad 1x = x.$$

Normalmente l’operazione di prodotto nel campo K si denota semplicemente ab al posto di $a \cdot b$. In questo testo si considereranno, di norma, campi vettoriali su \mathbb{R} (ossia, con campo di scalari $K = \mathbb{R}$) e campi vettoriali su \mathbb{C} (ossia, con campo di scalari $K = \mathbb{C}$). Gli elementi di uno spazio vettoriale vengono chiamati **vettori**, e gli elementi del campo K **scalari**.

(ii) Un sottoinsieme U di uno spazio vettoriale V , si dice **linearmente indipendente** se

$$\forall n \in \mathbb{N}, \forall v_1, \dots, v_n \in U, \forall a_1, \dots, a_n \in K : \quad \sum_{i=1}^n a_i v_i = 0 \quad \implies \quad a_i = 0, \quad \forall i; \quad (\text{D.1})$$

i vettori di un insieme linearmente indipendente si dicono **linearmente indipendenti**.

Si noti che se $U \subseteq V$ è linearmente indipendente, allora $u \neq 0, \forall u \in U$ (altrimenti la (D.1) sarebbe violata con $n = 1$ e $v_1 = 0$).

(iii) Un sottoinsieme U di uno spazio vettoriale V (su un campo K) è un **sottospazio vettoriale di V** se U è uno spazio vettoriale (su K).

Se U è un sottoinsieme non vuoto di uno spazio vettoriale V , l’insieme

$$\langle U \rangle := \{v \in V \mid \exists v_1, \dots, v_n \in U \text{ e } a_1, \dots, a_n \in K \mid v = a_1 v_1 + \dots + a_n v_n\}, \quad (\text{D.2})$$

è (evidentemente) un sottospazio vettoriale di V e si chiama **sottospazio generato da U** .

(iv) Uno spazio vettoriale V si dice **finito dimensionale** se esiste un suo sottoinsieme finito, linearmente indipendente β tale che $\langle \beta \rangle = V$; tale insieme β è una **base** (finita) per V ; se un tale insieme non esiste (e $V \neq \emptyset$), V si dice **infinito dimensionale**.

In questo *Complemento* considereremo solo spazi vettoriali finito dimensionali (e quindi il termine ‘base’ sarà, qui, sinonimo di ‘base finita’³).

Se $\beta = \{v_1, \dots, v_n\}$ è una base per V , allora:

$$\forall v \in V, \quad \exists! a_1, \dots, a_n \in K \mid \quad v = \sum_{i=1}^n a_i v_i \quad (\text{D.3})$$

¹Per la definizione di gruppo, vedi [https://it.wikipedia.org/wiki/Gruppo_\(matematica\)](https://it.wikipedia.org/wiki/Gruppo_(matematica)).

²Per la definizione di campo, si veda [https://it.wikipedia.org/wiki/Campo_\(matematica\)](https://it.wikipedia.org/wiki/Campo_(matematica)).

³Nel caso infinito dimensionale, la nozione di base è meno naturale e si possono introdurre varie definizioni, non equivalenti; una nozione di base particolarmente naturale per l’analisi in ambiente infinito dimensionale è quello di base di uno spazio Hilbert separabile; cfr. Sez 1.6.

dove l'unicità segue dalla lineare indipendenza dei vettori⁴ v_i ; gli scalari a_i in tale rappresentazione di v prendono il nome di **componenti di v nella base β** .

Esempio D.1 (i) \mathbb{R}^n e \mathbb{C}^n sono spazi vettoriali⁵ (finito dimensionale) con campo di scalari, rispettivamente, \mathbb{R} e \mathbb{C} , e una loro base (detta ‘base standard’) è data da $e^{(i)}$ (detto ‘versore nell’ i -ma direzione’) le cui componenti sono date da:

$$(e^{(j)})_i = \delta_{ij} := \begin{cases} 1 & \text{se } i = j, \\ 0 & \text{se } i \neq j. \end{cases} \quad (\text{D.4})$$

(ii) L'insieme dei polinomi su un campo K di grado al più n ,

$$K_n[x] := \{P(x) = a_0 + a_1x + \dots + a_nx^n \mid a_i \in K, x \in K\},$$

è uno spazio vettoriale (finito dimensionale) ed una sua base è data da $\{x^i \mid 0 \leq i \leq n\}$.

(iii) L'insieme di tutti i polinomi reali $\mathbb{R}[x] := \bigcup_{n \geq 0} \mathbb{R}_n[x]$ è uno spazio vettoriale infinito dimensionale: supponiamo, per assurdo, che $\mathbb{R}[x]$ sia finitamente generato, $\mathbb{R}[x] = \langle \{P_1, \dots, P_n\} \rangle$ con $P_i \in \mathbb{R}[x]$. Sia $d_i = \deg(P_i)$ il grado di P_i e sia $d = 1 + \max_{1 \leq i \leq n} d_i$. Chiaramente $x^d \neq \sum_{i=1}^n a_i P_i(x)$ per qualunque scelta di a_i e avremmo quindi una contraddizione⁶.

(v) Il prossimo lemma permette di definire univocamente **la dimensione di uno spazio vettoriale finito dimensionale** come la cardinalità di una qualunque sua base:

Lemma D.2 (di Steinitz⁷) *Sia V uno spazio vettoriale finito dimensionale e sia $\{v_1, \dots, v_n\}$ una sua base. Se $\{w_1, \dots, w_m\}$ è un sottoinsieme di V linearmente indipendente, allora $m \leq n$. Inoltre, se $m < n$, esistono $n - m$ vettori v_{j_i} tali che*

$$\langle \{w_1, \dots, w_m, v_{j_{m+1}}, \dots, v_{j_n}\} \rangle = V. \quad (\text{D.5})$$

Dimostrazione Il caso $n = 1$ è ovvio: se fosse $m > 1$, (poiché $\langle \{v_1\} \rangle = V$) si avrebbe $w_1 = \alpha v_1$ e $w_2 = \beta v_1$ con $\alpha, \beta \neq 0$, e, quindi, $w_1 = \frac{\alpha}{\beta} w_2$ contrariamente all'ipotesi che w_1 e w_2 sono linearmente indipendenti.

Assumiamo ora $n \geq 2$ e supponiamo, per assurdo che $m > n \geq 2$. Dimostriamo la seguente affermazione:

$$\exists 1 \leq i_1 \leq n \mid \langle \{w_1\} \cup \{v_i \mid 1 \leq i \leq n, i \neq i_1\} \rangle = V. \quad (\text{D.6})$$

Infatti, esistono a_i tali che $w_1 = \sum_{i=1}^n a_i v_i$ e (poiché $w_1 \neq 0$) non tutti i a_i sono nulli: sia i_1 tale che $a_{i_1} \neq 0$. Allora

$$v_{i_1} = \frac{1}{a_{i_1}} \left(w_1 - \sum_{i \neq i_1} a_i v_i \right),$$

relazione dalla quale segue immediatamente (D.6).

Ora, sia $1 \leq k < n$ e assumiamo che esistano k indici (diversi) i_1, \dots, i_k con $1 \leq i_j \leq n$ tali che

$$\langle \{w_1, \dots, w_k\} \cup \{v_i \mid i \neq i_j, \forall 1 \leq j \leq k\} \rangle = V. \quad (\text{D.7})$$

Allora, esistono scalari α_j e β_j tali che

$$w_{k+1} = \sum_{j=1}^k \alpha_j w_j + \sum_{i \neq i_1, \dots, i_k} \beta_j v_j,$$

⁴Se $v = \sum_{i=1}^n a_i v_i = \sum_{i=1}^n a'_i v_i$, allora $\sum_i (a_i - a'_i) v_i = 0$ e quindi $a_i = a'_i$ per ogni $1 \leq i \leq n$.

⁵ $x + y := (x_1, \dots, x_n) + (y_1, \dots, y_n) := (x_1 + y_1, \dots, x_n + y_n)$ e $ax := a(x_1, \dots, x_n) = (ax_1, \dots, ax_n)$.

⁶In questo caso c'è un candidato naturale per una ‘base’ per $\mathbb{R}[x]$, ossia l'insieme numerabile $\beta := \{x^n \mid n \in \mathbb{N}_0\}$: per ogni $P \in \mathbb{R}[x]$, $\exists!$ a_0, \dots, a_n tali che $P = \sum_{j=0}^n a_j x^j$ (tali a_j non sono altro che i coefficienti del polinomio P). Questo è un esempio di ‘base di Hamel’; cfr. [https://en.wikipedia.org/wiki/Basis_\(linear_algebra\)](https://en.wikipedia.org/wiki/Basis_(linear_algebra)).

⁷https://de.wikipedia.org/wiki/Ernst_Steinitz

e, poiché i w_j sono linearmente indipendenti) non tutti i β_j sono nulli. Quindi esiste $j_0 \neq i_1, \dots, i_k$ tale che $\beta_{j_0} \neq 0$, per cui, ponendo $i_{k+1} := j_0$, da (D.7) segue che

$$v_{i_{k+1}} = \frac{1}{\beta_{i_{k+1}}} \left(w_{k+1} - \sum_{j=1}^k \alpha_j w_j - \sum_{j \neq i_1, \dots, i_{k+1}} \beta_j v_j \right),$$

relazione che, a sua volta, implica la (D.7) con k sostituito da $k+1$. Per induzione (finita) su $k \leq n-1$ otteniamo dunque che $\langle \{w_1, \dots, w_n\} \rangle = V$. Ma, allora, $w_{k+1} \in \langle \{w_1, \dots, w_n\} \rangle$ contraddicendo la linear indipendenza di $\{w_1, \dots, w_{k+1}\}$. Dunque $m \leq n$. Inoltre, se $k < m$, iterando m volte la costruzione sopra descritta, si ottiene la (D.5). ■

Elenchiamo alcune conseguenze immediate del Lemma di Steinitz:

(a) **Definizione di dimensione di uno spazio vettoriale finito dimensionale**

Se $V \neq \{0\}$ è uno spazio vettoriale finito dimensionale e $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_m\}$ sono due basi per V , dal Lemma di Steinitz segue che $m = n$ e quindi la **dimensione di V** , $\dim V$, è, per definizione, *la cardinalità di una sua base qualunque ed è univocamente determinata*. Se V è lo spazio vettoriale banale $V = \{0\}$ poniamo, per definizione, $\dim V = 0$.

- (b) Se V è generato da una base $\beta := \{v_1, \dots, v_n\}$ e $U := \{w_1, \dots, w_m\}$ è un sottoinsieme di V linearmente indipendente, o $m = n$ nel qual caso anche U è una base per V , o è sempre possibile **completare** U in una base $\beta' := \{w_1, \dots, w_m, v_{j_{m+1}}, \dots, v_{j_n}\}$ (con opportuni j_i).
- (c) Un sottospazio vettoriale W di uno spazio vettoriale finito dimensionale, è uno spazio finito dimensionale con $\dim(W) \leq \dim(V)$.
- (d) Se V è uno spazio vettoriale che contiene infiniti vettori linearmente indipendenti, allora V è infinito dimensionale.

Somma e somma diretta di sottospazi vettoriali

Siano V_i , per $1 \leq i \leq m$ sottospazi vettoriali di uno spazio vettoriale V . Chiamiamo **somma dei V_i** l'insieme

$$\sum_{i=1}^m V_i := V_1 + \dots + V_m := \{v = v_1 + \dots + v_m \mid v_i \in V_i\}. \quad (\text{D.8})$$

Chiaramente tale insieme è un sottospazio vettoriale di V .

(vi) Diremo che V è la **somma diretta di sottospazi V_i** se valgono le seguenti due condizioni:

$$\left\{ \begin{array}{l} V = V_1 + \dots + V_m, \\ \sum_{i=1}^m v_i = 0 \text{ con } v_i \in V_i \implies v_i = 0, \forall i. \end{array} \right. \quad (\text{D.9})$$

In tal caso, scriveremo

$$V = \bigoplus_{i=1}^m V_i = V_1 \oplus \dots \oplus V_m. \quad (\text{D.10})$$

Per esempio, se $V_1 = \mathbb{R}^2 \times \{0\}$ e $v = (v_1, v_2, v_3) \in \mathbb{R}^3$ con $v_3 \neq 0$, allora $\mathbb{R}^3 = V_1 \oplus \langle \{v\} \rangle$.

Osservazione D.3 (i) È immediato verificare che la seconda condizione in (D.9) è equivalente a richiedere che se $v_1 + \dots + v_m = v'_1 + \dots + v'_m$, con $v_i, v'_i \in V_i$, allora $v_i = v'_i$ per ogni i . Dunque,

per ogni $v \in V_1 \oplus \cdots \oplus V_m$, esistono e sono unici $v_i \in V_i$ tali che $v = v_1 + \cdots + v_m$.

Si noti anche che se $v_i \in V_i \setminus \{0\}$ allora $\{v_i\}$ è un insieme indipendente: se $\sum a_i v_i = 0$, poiché $a_i v_i \in V_i$ dalla condizione di somma diretta (D.9) segue che $a_i v_i = 0$ ed essendo $v_i \neq 0$ si deve avere $a_i = 0$ per ogni i .

(ii) Nel caso $m = 2$, (D.9) è equivalente a

$$V = V_1 + V_2 \quad \text{e} \quad V_1 \cap V_2 = \{0\}. \quad (\text{D.11})$$

Dimostrazione Assumiamo (D.11). Allora, se $v_i \in V_i$ e $v_1 + v_2 = 0$, si ha che $V_1 \ni v_1 = -v_2 \in V_2$, e quindi $v_i \in V_1 \cap V_2$; ma allora, $v_1 = 0 = v_2$. Viceversa, se assumiamo (D.11) e se $v_1 + v_2 = 0$, come prima $v_i \in V_1 \cap V_2$ e quindi $v_1 = 0 = v_2$. ■

Nel caso generale, è facile verificare che la condizione (D.11) si generalizza come segue:

$$V = \sum_{i=1}^m V_i \quad \text{e} \quad V_i \cap \sum_{j \neq i} V_j = \{0\} \quad \forall i. \quad (\text{D.12})$$

Ma, in generale, se $m \geq 3$, $V = \sum V_i$ e $V_i \cap V_j = \{0\}$ se $i \neq j$ non è sufficiente a garantire che $V = \oplus V_i$: per esempio se $V_1 = \langle \{e^{(3)}\} \rangle$, $V_2 = \langle \{e^{(1)} + e^{(3)}\} \rangle$ e $V_3 = \langle \{e^{(1)}, e^{(2)}\} \rangle$, allora $\mathbb{R}^3 = V_1 + V_2 + V_3$ ma non è vero che⁸ $\mathbb{R}^3 = V_1 \oplus V_2 \oplus V_3$.

(iii) $\beta = \{v_1, \dots, v_n\}$ è una base per V se e solo se $V = \langle v_1 \rangle \oplus \cdots \oplus \langle v_n \rangle$.

(iv) Se $V = V_1 \oplus \cdots \oplus V_m$ è uno spazio n -dimensionale e se $n_i = \dim V_i$, allora $n = \sum_{i=1}^m n_i$.

Infatti, vettori $v_i \in V_i$ sono tra loro indipendenti e quindi l'unione di basi di V_i (la cui cardinalità è $\sum n_i$) formano un insieme indipendente e generano V dunque $\sum n_i = n$.

Si può dimostrare che vale anche il viceversa, ossia: se $V = V_1 + \cdots + V_m$ e $n = \sum_{i=1}^m n_i$, dove $n = \dim V$ e $n_i = \dim V_i$, allora $V = V_1 \oplus \cdots \oplus V_m$.

Prodotto scalare e ortogonalità

In \mathbb{R}^n è definito il prodotto scalare (1.4) che conduce alla nozione di ortogonalità (cfr. Osservazione 1.3-(iv)).

(vii) Una **base ortonormale** di uno spazio vettoriale $V \subseteq \mathbb{R}^n$ è una base $\{u_1, \dots, u_m\}$ di V tale che $u_i \cdot u_j = \delta_{ij}$.

Se $\beta = \{u_1, \dots, u_m\}$ è una base ortonormale di uno spazio vettoriale $V \subseteq \mathbb{R}^n$, ogni vettore $v \in V$ si rappresenta come

$$v = \sum_{j=1}^m (v \cdot u_j) u_j, \quad (\text{D.13})$$

ed inoltre

$$|v|^2 = v \cdot v = \sum_{j=1}^m |v \cdot u_j|^2. \quad (\text{D.14})$$

Dimostrazione Essendo β una base di V si ha che esistono e sono unici $a_j \in \mathbb{R}$ tali che $v = \sum_{j=1}^m a_j u_j$ e calcolando il prodotto scalare di v con u_i e usando la proprietà di ortonormalità di β si ottiene:

$$v \cdot u_i = \left(\sum_{j=1}^m a_j u_j \right) \cdot u_i = \sum_{j=1}^m a_j u_j \cdot u_i = \sum_{j=1}^m a_j \delta_{ij} = a_i,$$

⁸Se $v_1 = -e^{(3)}$, $v_2 = e^{(1)} + e^{(3)}$ e $v_3 = -e^{(1)}$, allora $v_i \in V_i$ e $0 = v_1 + v_2 + v_3$, ma $v_i \neq 0$.

da cui segue (D.13). La (D.14) segue sviluppando il prodotto scalare e usando la relazione $u_j \cdot u_i = \delta_{ij}$: in generale, per ogni $a_j \in \mathbb{R}$,

$$\left| \sum_{j=1}^m a_j u_j \right|^2 = \left(\sum_{j=1}^m a_j u_j \right) \cdot \left(\sum_{j=1}^m a_j u_j \right) = \sum_{i,j=1}^m a_i a_j u_i \cdot u_j = \sum_{i,j=1}^m a_i a_j \delta_{ij} = \sum_{i=1}^m |a_i|^2. \quad \blacksquare$$

In generale se V è un sottospazio di dimensione m di \mathbb{R}^n è sempre possibile trovare una base ortonormale di⁹ V :

(viii) (Lemma di Gram–Schmidt) Sia V è un sottospazio di dimensione $1 \leq m \leq n$ di \mathbb{R}^n e sia $\{v_1, \dots, v_m\}$ una base di V . Si definiscano, ricorsivamente, i seguenti vettori¹⁰:

$$\begin{cases} w_1 := v_1, & u_1 := \frac{w_1}{|w_1|}, \\ w_j := v_j - \sum_{h=1}^{j-1} (v_j \cdot u_h) u_h, & u_j := \frac{w_j}{|w_j|}, \quad \forall 2 \leq j \leq m. \end{cases} \quad (\text{D.15})$$

Allora, per ogni $1 \leq i, j \leq m$, si ha

$$u_i \cdot u_j = \delta_{ij}, \quad \text{e} \quad \langle \{u_1, \dots, u_j\} \rangle = \langle \{v_1, \dots, v_j\} \rangle. \quad (\text{D.16})$$

In particolare $\{u_1, \dots, u_m\}$ è una base ortonormale di V .

Dimostrazione Dimostriamo (D.16) per induzione su $i, j \leq p$ con $1 \leq p \leq m$. Se $p = 1$ e $i, j = 1$ la (D.16) è ovvia. Assumiamo (D.16) per $1 \leq i, j \leq p < m$ e dimostriamola per $p + 1$. Da (D.16) con $i, j \leq p$ segue che il vettore $\tilde{w}_p := \sum_{h=1}^p (v_{p+1} \cdot u_h) u_h \in \langle \{u_1, \dots, u_p\} \rangle$ e quindi $w_{p+1} := v_{p+1} - \tilde{w}_p$ è linearmente indipendente da $\{u_1, \dots, u_p\}$ e $\langle \{u_1, \dots, v_{p+1}\} \rangle = \langle \{u_1, \dots, u_{p+1}\} \rangle$. Per ogni $h \neq k \leq p$, $u_h \cdot u_k = \delta_{hk}$ (per ipotesi induttiva), e quindi, per ogni $k \leq p$, si ha:

$$\begin{aligned} w_{p+1} \cdot u_k &= (v_{p+1} - \tilde{w}_p) \cdot u_k = v_{p+1} \cdot u_k - \sum_{h=1}^p (v_{p+1} \cdot u_h) u_h \cdot u_k \\ &= v_{p+1} \cdot u_k - v_{p+1} \cdot u_k = 0. \end{aligned}$$

Quindi $u_{p+1} \cdot u_k = 0$ per ogni $k \leq p$ e poiché $|u_{p+1}| = 1$, si ha che la (D.16) vale con p sostituito da $p + 1$. \blacksquare

Esempio Siano $v_1 = (1, 1, 0)$, $v_2 = (1, 0, 1)$, $v_3 = (0, 0, 1)$. Si verifica facilmente che tali vettori sono indipendenti su \mathbb{R}^3 e, da (D.15), troviamo una base ortonormale $\{u_1, u_2, u_3\}$ con:

$$u_1 = \frac{1}{\sqrt{2}}(1, 1, 0), \quad u_2 = \frac{1}{\sqrt{6}}(1, -1, 2), \quad u_3 = \frac{1}{\sqrt{3}}(-1, 1, 1). \quad (\text{D.17})$$

(ix) Se V è un sottospazio di \mathbb{R}^n , si definisce il **complemento ortogonale** di V l'insieme

$$V^\perp := \{u \in \mathbb{R}^n \mid u \cdot v = 0, \forall v \in V\}. \quad (\text{D.18})$$

Dalla linearità del prodotto scalare nella prima componente segue immediatamente che V^\perp è uno spazio vettoriale.

Se V e W sono spazi vettoriali in \mathbb{R}^n poniamo

$$V \perp W \stackrel{\text{def}}{\iff} v \cdot w = 0 \quad \forall v \in V, \forall w \in W. \quad (\text{D.19})$$

⁹Il Lemma viii, a volte, è chiamato 'algoritmo di ortonormalizzazione di Gram–Schmidt'; Cfr. https://it.wikipedia.org/wiki/Ortogonalizzazione_di_Gram-Schmidt.

¹⁰Nel caso $m = 1$, la seconda riga non va considerata.

Si noti che se $\dim(V) = n$, allora $V = \mathbb{R}^n$ e $V^\perp = \{0\}$.

Assumiamo, ora, che $\dim(V) = m < n$. Sia $\beta = \{u_1, \dots, u_m\}$ una base ortonormale di V (Lemma viii). Possiamo completare¹¹ β in una base $\beta' = \{u_1, \dots, u_m, \dots, u_n\}$ di \mathbb{R}^n , che, sempre per il Lemma viii, possiamo assumere ortonormale. Sia $W = \langle \{u_{m+1}, \dots, u_n\} \rangle$. Dimostriamo che $W = V^\perp$. Chiaramente, essendo la base β' ortonormale, $W \subseteq V^\perp$. Sia ora $v \in V^\perp$, allora, in particolare, $v \cdot u_j = 0$ per ogni $j \leq m$, quindi (cfr. (D.13) che vale essendo β' una

base ortonormale di \mathbb{R}^n) $v = \sum_{j=m+1}^n (v \cdot u_j) u_j \in W$. Dunque $V^\perp \subseteq W$ e quindi $V^\perp = W$.

Poiché chiaramente $\mathbb{R}^n = \langle \{u_1, \dots, u_m\} \rangle \oplus \langle \{u_{m+1}, \dots, u_n\} \rangle$, segue che $\mathbb{R}^n = V \oplus V^\perp$ e che $\dim V^\perp = n - m$. Si noti poi che, dalla simmetria tra V e W (ossia che gli elementi della base dell'uno sono il complementare degli elementi della base β'), segue immediatamente che $W^\perp = V$, ossia che $(V^\perp)^\perp = V$. Abbiamo dunque dimostrato:

(x) Sia V un sottospazio di \mathbb{R}^n di dimensione $m < n$. Allora, $\dim(V^\perp) = n - m$ e $\mathbb{R}^n = V \oplus V^\perp$. Inoltre $(V^\perp)^\perp = V$.

Sia $u \in \mathbb{R}^n$ e V un sottospazio di \mathbb{R}^n . Da **(x)** segue che esiste un unico $v \in V$ tale che $u = v + w$ con $w \in V^\perp$: tale v si chiama **proiezione ortogonale di u su V** e si denota con $\pi_V(u)$. Se $\{u_1, \dots, u_m\}$ è una base ortonormale di V , per quanto visto sopra, si ha

$$\pi_V(u) = \sum_{j=1}^m (u \cdot u_j) u_j, \quad \forall u \in \mathbb{R}^n. \quad (\text{D.20})$$

Ovviamente¹²,

$$\pi_V|_V = \text{id}, \quad \pi_V^2 := \pi_V \circ \pi_V = \pi_V. \quad (\text{D.21})$$

La proiezione ortogonale ha anche la seguente caratterizzazione ‘variazionale’:

(xi) Sia V un sottospazio di dimensione m di \mathbb{R}^n e sia $v \in \mathbb{R}^n$. Allora, esiste, ed è unico, $v_0 \in V$ tale che

$$|u - v_0| = \min\{|u - v| \mid v \in V\}, \quad (\text{D.22})$$

e tale v_0 coincide con $\pi_V(u)$.

Dimostrazione Sia $\{u_1, \dots, u_m\}$ è una base ortonormale di V e $\{u_1, \dots, u_n\}$ una base ortonormale di \mathbb{R}^n . Un qualunque $u \in \mathbb{R}^n$ si sviluppa come $u = \sum_{j=1}^n (u \cdot u_j) u_j$ e un qualunque

$v \in V$ come $v = \sum_{j=1}^m a_j u_j$, con $a_j \in \mathbb{R}$. Allora, usando (D.14),

$$\begin{aligned} |u - v|^2 &= \left| \sum_{j=1}^n (u \cdot u_j) u_j - \sum_{j=1}^m a_j u_j \right|^2 = \left| \sum_{j=1}^m ((u \cdot u_j) - a_j) u_j + \sum_{j=m+1}^n (u \cdot u_j) u_j \right|^2 \\ &= \sum_{j=1}^m |(u \cdot u_j) - a_j|^2 + \sum_{j=m+1}^n |u \cdot u_j|^2, \end{aligned}$$

il che mostra che il minimo è raggiunto quando $a_j = u \cdot u_j$ per ogni $1 \leq j \leq m$, ossia quando $v = \pi_V(u)$. ■

(xii) Su \mathbb{C}^n si definisce il seguente **prodotto hermitiano**:

$$\langle u, v \rangle := u \cdot \bar{v} := \sum_{j=1}^n u_j \bar{v}_j, \quad (\text{D.23})$$

dove la barra denota complesso coniugato.

¹¹Cfr. (v).

¹²Se $f : A \rightarrow B$ e $V \subseteq A$, $f|_V : x \in V \mapsto f(x) \in B$ denota la restrizione di f a V .

Tale forma quadratica è *sesquilineare*¹³, ossia, lineare nella prima variabile e antilineare nella seconda:

$$\langle au, v \rangle = a \langle u, v \rangle, \quad \langle u, av \rangle = \bar{a} \langle u, v \rangle, \quad \forall a \in \mathbb{C}, \forall u, v \in \mathbb{C}^n. \quad (\text{D.24})$$

Come nel caso reale, il prodotto hermitiano è legato alla norma dalla relazione

$$|u|^2 := \sum_{j=1}^n |u_j|^2 = \langle u, u \rangle, \quad \forall u \in \mathbb{C}^n. \quad (\text{D.25})$$

Si noti anche che

$$\overline{\langle u, v \rangle} = \langle v, u \rangle, \quad \forall u, v \in \mathbb{C}^n. \quad (\text{D.26})$$

Una base ortonormale di un sottospazio vettoriale V di \mathbb{C}^n è una base $\{u_1, \dots, u_m\}$ di V tale che $\langle u_i, u_j \rangle = \delta_{ij}$.

A questo punto è facile verificare che i risultati descritti da (vii) a (xi) si estendono al caso di \mathbb{C}^n , sostituendo il prodotto scalare $u \cdot v$ con il prodotto hermitiano $\langle u, v \rangle$.

D.2 Applicazioni lineari

(xiii) Dati due spazi vettoriali V e W (sullo stesso campo K), una funzione¹⁴ $L : u \in V \mapsto Lu \in W$ è una **applicazione lineare** se $L(au + bv) = aLu + bLv$ per ogni $u, v \in V$ e $a, b \in K$.

Si noti che, se L è lineare, $L0 = L(0v) = 0Lv = 0$.

Un'applicazione lineare tra spazi vettoriali V e W che sia iniettiva e suriettiva si chiama **isomorfismo** (di spazi lineari) e V e W si dicono **isomorfi**.

Un'applicazione lineare da V in se stesso prende il nome di **endomorfismo lineare**.

Chiaramente, come in generale per funzioni, l'insieme delle applicazioni lineari da uno spazio vettoriale V in uno spazio vettoriale W (con lo stesso campo di scalari) forma esso stesso uno spazio vettoriale che si denota con $\mathcal{L}(V, W)$ e nel caso $V = W$ con¹⁵ $\mathcal{L}(V)$.

Dal lemma di Steinitz segue:

(xiv) Due spazi vettoriali (finito dimensionali) sono isomorfi se e solo se hanno la stessa dimensione.

Dimostrazione Supponiamo che V e W abbiano la stessa dimensione e siano $\{v_1, \dots, v_n\}$ e $\{w_1, \dots, w_n\}$ due, rispettive, basi. Definiamo

$$j : v = \sum_{j=1}^n \lambda_j v_j \in V \mapsto j(v) := \sum_{j=1}^n \lambda_j w_j \in W. \quad (\text{D.27})$$

Dall'unicità dei coefficienti di un dato vettore rispetto ad una data base, segue che j è un isomorfismo lineare.

Viceversa, sia $j : V \rightarrow W$ un isomorfismo e $\beta := \{v_1, \dots, v_n\}$ e $\beta' := \{w_1, \dots, w_m\}$ rispettive basi. Essendo j un isomorfismo, segue che $j(\beta)$ è linearmente indipendente in W e, analogamente, $j^{-1}(\beta')$ è linearmente indipendente in V . Quindi, dal Lemma di Steinitz, segue che $n = m$.

■

In particolare:

Qualunque spazio vettoriale su K di dimensione n è isomorfo a K^n e, fissata una base $\beta = \{v_1, \dots, v_n\}$ di V , chiamiamo l'isomorfismo

$$j_\beta : v = \sum_{i=1}^n x_i v_i \in V \mapsto j_\beta(v) = x = (x_1, \dots, x_n) \in K^n, \quad (\text{D.28})$$

¹³Cfr. https://it.wikipedia.org/wiki/Forma_sesquilineare.

¹⁴In genere, per mappe lineari, si omettono le parentesi tonde, se ciò non crea confusione.

¹⁵Come al solito, se $a, b \in K$ e $L, M \in \mathcal{L}(V, W)$, $(aL + bM)(u) := aLu + bMu$, per ogni $u \in V$.

isomorfismo canonico tra V e K^n subordinato alla base β .

(xv) Data un'applicazione lineare $L \in \mathcal{L}(V, W)$, definiamo, rispettivamente, il **nucleo** (o kernel) di L e l'**immagine** di L gli insiemi

$$\ker L := \{u \in V \mid Lu = 0\}, \quad \text{im } L := \{w = Lu \in W \mid u \in V\}. \quad (\text{D.29})$$

Dalla linearità di L segue immediatamente che questi insiemi sono spazi vettoriali. Un'ulteriore conseguenza del Lemma di Steinitz è il seguente fondamentale

Teorema D.4 (nullità–rango¹⁶) Sia $L : V \rightarrow W$ un'applicazione lineare tra spazi vettoriali (finito dimensionali). Allora,

$$\dim(\ker L) + \dim(\text{im } L) = \dim V. \quad (\text{D.30})$$

Infatti, se $L \neq 0$, esistono m vettori $u_j \in V$ tali che gli m vettori $Lu_j \in W$ sono indipendenti e

$$V = \ker L \oplus \langle u_1, \dots, u_m \rangle \quad \text{e} \quad \langle Lu_1, \dots, Lu_m \rangle = \text{im } L. \quad (\text{D.31})$$

Dimostrazione Se $L = 0$, allora $\ker L = V$ e $\dim(\text{im } L) = 0$ e il teorema è banalmente vero. Assumiamo dunque¹⁷ $L \neq 0$.

Se vettori Lu_j sono indipendenti, lo sono anche i vettori¹⁸ u_j . Dunque, da (D.31) segue che $\langle u_1, \dots, u_m \rangle$ e $\text{im } L$ hanno la stessa dimensione e, quindi, dal punto (iv) dell'Osservazione D.3 e (D.31) segue immediatamente (D.30). Dimostriamo direttamente (D.31).

Sia $n = \dim V$ e siano¹⁹ $\{v_1, \dots, v_k\}$ e $\{w_1, \dots, w_m\}$ basi di, rispettivamente, $\ker L$ e $\text{im } L$, dove $m \geq 1$ e $k \geq 0$, con la convenzione che se $k = 0$, $\ker L = \{0\}$ e i vettori v_j non esistono. Siano $u_j \in V$, tali che $Lu_j = w_j$, per ogni $1 \leq j \leq m$. I vettori $\{u_j\}$ sono linearmente indipendenti (altrimenti, sarebbero linearmente dipendenti anche i $\{w_j\}$). Mostriamo che l'insieme $\beta := \{v_1, \dots, v_k, u_1, \dots, u_m\}$ è linearmente indipendente. Se

$$\sum_{i=1}^k \alpha_i v_i + \sum_{j=1}^m \beta_j u_j = 0,$$

allora, $u := \sum \beta_j u_j = -\sum \alpha_i v_i \in \ker(L)$, e quindi $0 = Lu = \sum \beta_j Lu_j = \sum \beta_j w_j$, il che implica che $\beta_j = 0$, per ogni j (essendo i vettori w_j indipendenti); dunque si avrebbe anche $\sum \alpha_i v_i = 0$, che implica anche $\alpha_i = 0$ per ogni i : in conclusione β è un insieme indipendente. Mostriamo, infine, che $\langle \beta \rangle = V$. Sia $v \in V$. Se $v \in \ker(L)$, allora $v \in \langle \beta \rangle$. Se $v \notin \ker L$, allora $Lv \neq 0$ e quindi $Lv = \sum \beta_i w_i$ per opportuni β_i non tutti nulli. Ora, $L(v - \sum \alpha_i u_i) = 0$, ossia, $v - \sum \alpha_i u_i \in \ker(L)$ e quindi $v - \sum \alpha_i u_i = \sum \beta_i v_i$, ossia, $v = \sum \alpha_i u_i + \sum \beta_i v_i$ e quindi, anche in questo caso, $v \in \langle \beta \rangle$. ■

Osservazione D.5 (i) Dalla formula (D.31) segue facilmente che, se $V_1 := \langle u_1, \dots, u_m \rangle$, allora la restrizione di L su V_1 , $L|_{V_1} : V_1 \rightarrow \text{im}(L)$, è un isomorfismo lineare tra V_1 e $\text{im}(L)$.

(ii) La dimensione di W non gioca nessun ruolo nella dimostrazione e infatti il teorema e la sua dimostrazione continuano a valere anche se W non ha dimensione finita.

(xvi) (**Endomorfismi nilpotenti**) Un endomorfismo lineare $N \in \mathcal{L}(V)$ si dice **nilpotente** se esiste $m \in \mathbb{N}$ tale che $N^m = 0$. Se N è nilpotente si definisce il suo **indice di nilpotenza** il numero naturale $\nu := \nu(N) := \min\{p \in \mathbb{N} \mid N^p = 0\}$.

Per esempio, la mappa $N : K^2 := K \times K \rightarrow K^2$ tale che $N(x, y) := (y, 0)$ è (chiaramente) un endomorfismo lineare di K^2 e $N^2(x, y) = N(y, 0) = (0, 0)$ per ogni $(x, y) \in K^2$ e quindi N è

¹⁶La 'nullità di L ' è la dimensione del nucleo, mentre il 'rango' è la dimensione dell'immagine.

¹⁷ $L \neq 0$ significa che esiste $u \in V$ tale che $Lu \neq 0$ e quindi $\ker L \neq V$ e $\dim(\text{im } L) \geq 1$.

¹⁸Se $\sum a_j u_j = 0 \implies 0 = L \sum a_j u_j = \sum a_j Lu_j \implies a_j = 0$ per $L \neq 0$ e ogni j .

¹⁹Qui usiamo il Lemma di Steinitz.

nilpotente con indice $\nu(N) = 2$. Ovviamente, se N è nilpotente, $\nu(N) = 1$ se e solo se $N = 0$; inoltre, $N^p = 0$ per ogni $p \geq \nu$.

Sia V uno spazio vettoriale di dimensione n . Se $N \in \mathcal{L}(V)$ è nilpotente, allora $\nu = \nu(N) \leq n$, infatti esiste $u \in V$ tale che l'insieme $J := \{u, Nu, \dots, N^{\nu-1}u\}$ è linearmente indipendente.

Dimostrazione Per definizione di indice di nilpotenza, $N^\nu = 0$ e $N^{\nu-1} \neq 0$, dunque esiste $u \in V$ tale che $N^{\nu-1}u \neq 0$ e $N^\nu u = 0$. Se $\nu = 1$, la tesi è ovvia con $J = \{u\}$, con u un qualunque vettore di V diverso da 0. Sia $\nu > 1$ e sia $w = \sum_{k=0}^{\nu-1} a_k N^k u$ una combinazione lineare di elementi di J e assumiamo che $w = 0$. Allora, $0 = N^{\nu-1}w = a_0 N^{\nu-1}u$ e quindi (essendo $N^{\nu-1}u \neq 0$) si ha che $a_0 = 0$. Ma ora, $0 = N^{\nu-2}w = a_1 N^{\nu-1}u$ e quindi anche $a_1 = 0$. Iterando (se $\nu > 2$) si ha che $a_k = 0$ per ogni k . La tesi segue ora dal Lemma di Steinitz. ■

Un insieme $J := \{u, Nu, \dots, N^{j-1}u\}$ tale che $N^{j-1}u \neq 0$ e $N^j u = 0$ prende il nome di **catena di Jordan**²⁰. Le catene di Jordan saranno centrali nella discussione della forma canonica di Jordan più avanti (cfr., in particolare, (lii)).

(xvii) (Spettro, autovalori ed autovettori) Sia V uno spazio vettoriale n -dimensionale su K e $T \in \mathcal{L}(V)$ un endomorfismo lineare. Un elemento $\lambda \in K$ si dice **autovalore** di T e $u \in V \setminus \{0\}$ un **autovettore** relativo a λ se $Tu = \lambda u$. Il sottospazio $V_\lambda := \ker(T - \lambda I) \subseteq V$ di tutti gli autovettori relativi ad un autovalore λ (assieme al vettore nullo) prende il nome di **autospatio** relativo a λ ; la dimensione di V_λ prende il nome di **molteplicità geometrica** di λ e si denota $g(\lambda)$. L'insieme $\sigma(T) := \sigma_K(T) \subseteq K$ di tutti gli autovalori di T prende il nome di **spettro di T** (su K).

(xviii) Ad autovalori distinti corrispondono autovettori indipendenti. Più precisamente, siano $T \in \mathcal{L}(V)$, $\lambda_1, \dots, \lambda_m \in \sigma(T)$ con $\lambda_i \neq \lambda_j$ se $i \neq j$ e $Tu_i = \lambda_i u_i$ con $u_i \in V \setminus \{0\}$. Allora $\{u_1, \dots, u_m\}$ è un insieme indipendente.

Dimostrazione Supponiamo che $\sum_{j=1}^m a_j u_j = 0$ con $a_j \in K$. Fissiamo $1 \leq i \leq m$. Definiamo²¹

$S = \prod_{j \neq i} (T - \lambda_j I)$. Si noti che gli endomorfismi $(T - \lambda_k I)$ e $(T - \lambda_h I)$ commutano²² e quindi l'ordine nel prodotto che definisce S è irrilevante. Ora,

$$0 = S \left(\sum_{j=1}^m a_j u_j \right) = \sum_{j=1}^m a_j S u_j = a_i S u_i = a_i \prod_{j \neq i} (\lambda_i - \lambda_j) \implies a_i = 0. \quad \blacksquare$$

Dunque, la dimensione geometrica $g(\lambda)$ di un autovalore $\lambda \in \sigma(T)$ è un intero tra 0 e n .

Osservazione D.6 $\dim V_\lambda = 0$ equivale a dire che $\sigma(T) = \emptyset$. Un esempio di endomorfismo con spettro vuoto è il seguente. Sia $V = \mathbb{R}^2$ ($K = \mathbb{R}$) e $T(u_1, u_2) := (-u_2, u_1)$. Allora $\lambda \in \mathbb{R}$ è autovalore se e solo se $\exists u \neq 0$ tale che $Tu = \lambda u$, ossia $(-u_2, u_1) = \lambda(u_1, u_2)$, il che implica $u_1 = -\lambda^2 u_1$ e $u_2 = -\lambda^2 u_2$ e, dunque, $\lambda^2 = -1$, equazione che non ha soluzione in \mathbb{R} ; quindi, deve essere $\sigma(T) = \emptyset$.

Vedremo in seguito che se il campo degli scalari è il campo complesso ($K = \mathbb{C}$), allora $\sigma(T) \neq \emptyset$.

D.3 Matrici

(xix) Sia $\mathcal{F}_n := \{i \in \mathbb{N} \mid i \leq n\}$. Una **matrice $n \times m$ a coefficienti nel campo K** è una funzione

$$A : (i, j) \in \mathcal{F}_n \times \mathcal{F}_m \mapsto A_{ij} \in K.$$

²⁰https://it.wikipedia.org/wiki/Camille_Jordan

²¹ I denota l'endomorfismo identità: $Iu = u$ per ogni u .

²²Ossia, $[T, S] := TS - ST = 0$.

I valori ²³ A_{ij} di una matrice A vengono detti ‘elementi di matrice’. L’insieme delle matrici $n \times m$ a coefficienti in K si denota con $K^{n \times m}$.

È uso comune rappresentare una matrice $A \in K^{n \times m}$ come una ‘tabella rettangolare’ con n righe e m colonne

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1j} & \cdots & A_{1m} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ A_{i1} & \cdots & A_{ij} & \cdots & A_{im} \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ A_{n1} & \cdots & A_{nj} & \cdots & A_{nm} \end{pmatrix} \quad (\text{D.32})$$

L’insieme delle matrici $K^{n \times m}$ ha una struttura naturale di spazio vettoriale su K : se $A = (A_{ij}), B = (B_{ij}) \in K^{n \times m}$, $a \in K$, $(A+B)_{ij} := A_{ij} + B_{ij}$ e $(aA)_{ij} = aA_{ij}$. Tale spazio vettoriale ha dimensione nm (una base è data dalle nm matrici (δ_{ij}) (con $1 \leq i \leq n$ e $1 \leq j \leq m$), dove $\delta_{ij} = 1$, se $i = j$ e $\delta_{ij} = 0$ se $i \neq j$) ed è dunque isomorfo (come spazio vettoriale) a K^{nm} .

Le **righe** di una matrice $A = (A_{ij}) \in K^{n \times m}$ sono gli n vettori in K dati da (A_{i1}, \dots, A_{im}) con $1 \leq i \leq n$; le **colonne** di una matrice $A = (A_{ij}) \in K^{n \times m}$ sono gli m vettori in K^n dati da (A_{1j}, \dots, A_{nj}) con $1 \leq j \leq m$.

Dati m vettori in $x^{(j)} \in K^n$, indichiamo con $[x^{(1)}, \dots, x^{(m)}]$ la matrice $n \times m$ che ha come colonne i vettori $x^{(j)}$, ossia

$$([x^{(1)}, \dots, x^{(m)}])_{ij} := x_i^{(j)}; \quad (\text{D.33})$$

nel caso $m = 1$, la matrice $(n \times 1)$, $[x]$, si identifica, di solito, col vettore x .

(xx) Data una matrice $A \in K^{n \times m}$, la sua **matrice trasposta**, A^T è una matrice $(m \times n)$ in cui le righe e le colonne sono scambiate tra loro, ossia, $A_{ij}^T := A_{ji}$. Una matrice $A \in K^{n \times n}$ si dice **simmetrica** se $A^T = A$.

Se $K = \mathbb{C}$, $A^* = \bar{A}^T \in \mathbb{C}^{m \times n}$ denota la matrice **aggiunta** (o ‘complesso-coniugata’) data da $(A^*)_{ij} := \bar{A}_{ji}$ (dove la barra, come al solito, denota complesso-coniugato). Una matrice $A \in \mathbb{C}^{n \times n}$ si dice **autoaggiunta** se $A^* = A$.

(xxi) Date due matrici $A \in K^{n \times m}$ e $B \in K^{m \times p}$ si definisce il **prodotto righe per colonne** di A per B , e si denota AB , la matrice $K^{n \times p}$, i cui elementi di matrice sono dati da:

$$(AB)_{ij} = \sum_{k=1}^m A_{ik} B_{kj}, \quad 1 \leq i \leq n, \quad 1 \leq j \leq p. \quad (\text{D.34})$$

Tale prodotto è (come si verifica immediatamente) associativo: se $A \in K^{n \times m}$, $B \in K^{m \times p}$, $C \in K^{p \times q}$, si ha $(AB)C = A(BC) \in K^{n \times q}$.

Si noti che se $A \in K^{n \times m}$ e $[v^{(1)}, \dots, v^{(p)}] \in K^{m \times p}$ (ossia, $v^{(j)} \in K^m$), allora

$$A[v^{(1)}, \dots, v^{(p)}] = [Av^{(1)}, \dots, Av^{(p)}] \in K^{n \times p}, \quad (A \in K^{n \times m}, \quad v^{(j)} \in \mathbb{R}^m). \quad (\text{D.35})$$

Infine, si osservi che se $A := [v^{(1)}, \dots, v^{(m)}] \in K^{n \times m}$ (ossia, $v^{(j)} \in K^n$), allora

$$v^{(j)} = Ae^{(j)}, \quad (\text{D.36})$$

dove $e^{(j)}$ è il j -mo versore in K^m .

Particolarmente importanti sono le matrici ‘quadrate’:

(xxii) Una **matrice quadrata** è una matrice $(n \times n)$; il numero n prende il nome di **ordine della matrice** A .

È importante osservare che, se $n \geq 2$, il prodotto tra matrici in $K^{n \times n}$ non è commutativo. Per esempio:

$$\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} = \begin{pmatrix} 4 & 2 \\ 1 & 1 \end{pmatrix} \neq \begin{pmatrix} 4 & 1 \\ 2 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}. \quad (\text{D.37})$$

²³A volte, i valori di una matrice A si denota con la lettera minuscola a_{ij} ; noi, di norma, useremo la notazione $A = (A_{ij})$ per indicare una matrice e i suoi valori.

La **matrice identità** in $K^{n \times n}$ è definita come:

$$I_n := \begin{pmatrix} 1 & 0 & 0 \cdots & 0 \\ 0 & 1 & 0 \cdots & 0 \\ \vdots & & \ddots & \\ 0 & & & 1 \end{pmatrix}, \quad (I_n)_{ij} = \delta_{ij}, \quad (\text{D.38})$$

e soddisfa $AI_n = I_nA$ per ogni $A \in K^{n \times n}$.

(xxiii) Una matrice quadrata $A \in K^{n \times n}$ si dice **invertibile** se esiste una matrice $B \in K^{n \times n}$ tale che $AB = BA = I_n$; tale matrice se esiste è unica²⁴ e si denota con A^{-1} .

L'insieme delle matrici invertibili ($n \times n$) in $K^{n \times n}$ forma un gruppo che si denota con $GL(n, K)$ o con²⁵ $GL_n(K)$. Se $n \geq 2$, $GL(n, K)$ è un gruppo *non commutativo*.

D.4 Determinanti

Nella teoria delle matrici gioca un ruolo fondamentale la ‘funzione determinante’ (o, semplicemente, ‘il determinante’) $\det : K^{n \times n} \mapsto K$. Per definire e studiare le proprietà del determinante di una matrice, discutiamo brevemente il cosiddetto *gruppo simmetrico* delle permutazioni di n oggetti.

(xxiv) Il **gruppo simmetrico su n elementi (o, n oggetti)** è l'insieme delle funzioni biunivoche²⁶

$$\sigma : i \in \mathcal{F}_n := \{1, \dots, n\} \mapsto \sigma_i \in \mathcal{F}_n.$$

Spesso denoteremo un elemento $\sigma \in S_n$ come una n -pla $(\sigma_1, \dots, \sigma_n) \in \mathcal{F}_n^n$. Gli elementi di S_n si chiamano *permutazioni* (di n oggetti). Per esempio,

$$S_1 = \{1\}; S_2 = \{(1, 2), (2, 1)\}, S_3 = \{(1, 2, 3), (1, 3, 2), (2, 1, 3), (2, 3, 1), (3, 1, 2), (3, 2, 1)\}.$$

Per induzione, si vede facilmente che $\#S_n = n!$.

Poiché componendo funzioni biunivoche si ottengono funzioni biunivoche, se $\sigma, \tau \in S_n$, allora $\sigma \cdot \tau := \sigma \circ \tau \in S_n$. Rispetto a tale prodotto, l'unità è data dalla identità $\sigma_0 := \text{id} = (1, 2, \dots, n)$. Inoltre, poiché l'inversa di una funzione biunivoca è biunivoca, $\sigma^{-1} \in S_n$ e $\sigma^{-1} \cdot \sigma = \sigma \cdot \sigma^{-1} = \sigma_0$. Dunque S_n *forma un gruppo* (rispetto alla composizione, con unità $\sigma_0 = \text{id}$; l'associatività è ovvia).

Per $n \geq 2$, esempi di permutazione sono le *trasposizioni*: $\tau \in S_n$ è una **trasposizione** se scambia tra loro due indici, lasciando gli altri invariati: ossia, $\exists 1 \leq k < \ell \leq n$, tali che $\tau_k = \ell$, $\tau_\ell = k$ e $\tau_i = i$ per ogni $i \neq k, \ell$. Per esempio, $S_2 = \{\sigma_0, \tau\}$ dove $\sigma_0 = (1, 2)$ è l'unità e $\tau = (2, 1)$ è una trasposizione. In generale, se τ è una trasposizione $\tau^{-1} = \tau$ e dunque $\tau^2 := \tau \cdot \tau = \sigma_0$.

Lemma D.7 *Sia $n \geq 2$. Per ogni $\sigma \in S_n$, esistono N trasposizioni $\tau^{(1)}, \dots, \tau^{(N)}$ tali che $\sigma = \tau^{(1)} \dots \tau^{(N)}$.*

Dimostrazione Se $\sigma = \sigma_0$, la tesi è vera con $N = 2$; infatti, se τ_1 è una qualunque trasposizione e $\tau_2 = \tau_1$: per quanto osservato sopra, $\tau_1 \cdot \tau_2 = \tau_1^2 = \text{id} = \sigma$.

Assumiamo ora che $\sigma \neq \sigma_0$. Sia i il più piccolo naturale tale $\sigma_i \neq i$. In tal caso, esiste $i < j \leq N$ tale che $\sigma_j = i$; si noti in particolare che, necessariamente, $i < N$. Scegliamo $\tau^{(1)}$ come la trasposizione che scambia i con j : allora, $\sigma^{(1)} := \tau^{(1)} \cdot \sigma$ è tale che $\sigma_\ell = \ell$ per ogni $\ell \leq i$. Se $\sigma^{(1)} = \sigma_0$, si ha $\tau^{(1)} = \sigma$ e la tesi è vera con $N = 1$. Se $\sigma^{(1)} \neq \sigma_0$, possiamo ripetere quanto fatto prima: sia k il più piccolo naturale tale che $\sigma_k \neq k$: chiaramente, $k > i$ e possiamo ripetere

²⁴Se anche $C \in K^{n \times n}$ verifica $CA = I_n$, allora, $C = CI_n = C(AB) = (CA)B = I_nB = B$. Le matrici nell'esempio (D.37) sono invertibili, infatti: $\begin{pmatrix} 2 & 0 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & 0 \\ 0 & 1 \end{pmatrix}$ e $\begin{pmatrix} 2 & 1 \\ 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}$.

²⁵Cfr. https://en.wikipedia.org/wiki/General_linear_group.

²⁶In tal caso gli ‘oggetti’ sono i primi n numeri naturali \mathcal{F}_n , ma, ovviamente, al posto di $\mathcal{F}(n)$ si possono considerare n elementi distinti di un qualunque insieme.

la costruzione fatta prima, definendo $\tau^{(2)}$ in modo che $\sigma^{(2)} := \tau^{(2)} \cdot \sigma^{(1)}$ sia tale che $\sigma_\ell^{(2)} = \ell$ per ogni $\ell \leq k$. Per induzione finita, esiste $N \geq 2$, tale che $\sigma^{(N)} := \tau^{(N)} \cdots \tau^{(1)} \cdot \sigma = \sigma_0$, il che equivale a $\sigma = \tau^{(1)} \cdots \tau^{(N)}$, che è la tesi. ■

Sia $\sigma \in S_n$. Si definisce il **segno di σ** il numero²⁷

$$\varepsilon_\sigma := \prod_{i < j} \operatorname{sgn}(\sigma_j - \sigma_i), \quad (\text{D.39})$$

dove $x \in \mathbb{R} \rightarrow \operatorname{sgn}(x)$ è l'ordinaria funzione 'segno di x ', ossia $\operatorname{sgn}(x) = 1$ se $x > 0$, $\operatorname{sgn}(x) = -1$ se $x < 0$, $\operatorname{sgn}(0) = 0$. Si noti che, essendo σ iniettiva $\sigma_i \neq \sigma_j$ se $i < j$ e quindi $\varepsilon_\sigma \in \{-1, 1\}$. Dalla definizione segue immediatamente che $\varepsilon_{\sigma_0} = 1$, mentre se τ è una trasposizione $\varepsilon_\tau = -1$. Più in generale, si vede subito che se τ è una trasposizione, allora $\varepsilon_{\tau \cdot \sigma} = -\varepsilon_\sigma$, per ogni $\sigma \in S_n$. Dunque dal Lemma D.7 segue che

la mappa $\sigma \in S_n \rightarrow \varepsilon_\sigma \in \{-1, 1\}$ è un omeomorfismo²⁸ dal gruppo simmetrico S_n sul gruppo moltiplicativo di ordine due $(\{-1, 1\}, \cdot)$.

Da questa proprietà e dal Lemma D.7 segue che se $\sigma = \tau^{(1)} \cdots \tau^{(N)}$, allora $\varepsilon_\sigma = (-1)^N$ e dunque la parità di N nella rappresentazione di σ come prodotto di trasposizioni dipende solo da σ (e non dalla scelta delle trasposizioni, che ovviamente non è unica).

(xxv) Data una matrice $A = [v^{(1)}, \dots, v^{(n)}] \in K^{n \times n}$ si definisce **determinante di A** il numero in K dato da:

$$\det A := \sum_{\sigma \in S_n} \varepsilon_\sigma A_{1\sigma_1} \cdots A_{n\sigma_n} = \sum_{\sigma \in S_n} \varepsilon_\sigma v_1^{(\sigma_1)} \cdots v_n^{(\sigma_n)}. \quad (\text{D.40})$$

Esempio D.8 (1) Se $n = 1$, $K^{1 \times 1} = K$ e $\det A = A$, per ogni $A \in K^{1 \times 1}$.

(2) Se $n = 2$,

$$\det \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix} = \sum_{\sigma \in S_2} \varepsilon_\sigma A_{1\sigma_1} A_{2\sigma_2} = \varepsilon_{(1,2)} A_{11} A_{22} + \varepsilon_{(2,1)} A_{12} A_{21} = A_{11} A_{22} - A_{12} A_{21}.$$

(3) Se $n = 3$,

$$\begin{aligned} \det \begin{pmatrix} A_{11} & A_{12} & A_{13} \\ A_{21} & A_{22} & A_{23} \\ A_{31} & A_{32} & A_{33} \end{pmatrix} &= \sum_{\sigma \in S_3} \varepsilon_\sigma A_{1\sigma_1} A_{2\sigma_2} A_{3\sigma_3} \\ &= \varepsilon_{(1,2,3)} A_{11} A_{22} A_{33} + \varepsilon_{(1,3,2)} A_{11} A_{23} A_{32} + \varepsilon_{(2,1,3)} A_{12} A_{21} A_{33} \\ &\quad + \varepsilon_{(2,3,1)} A_{12} A_{23} A_{31} + \varepsilon_{(3,1,2)} A_{13} A_{21} A_{32} + \varepsilon_{(3,2,1)} A_{13} A_{22} A_{31} \\ &= A_{11} A_{22} A_{33} - A_{11} A_{23} A_{32} - A_{12} A_{21} A_{33} + A_{12} A_{23} A_{31} + A_{13} A_{21} A_{32} - A_{13} A_{22} A_{31}. \end{aligned}$$

(4) Se A è una **matrice triangolare**, ossia, $A_{ij} = 0$, $\forall i < j$ (in tal caso, si dice 'sopra-triangolare'), oppure, $A_{ij} = 0$, $\forall i > j$ (in tal caso, si dice 'sotto-triangolare'), allora

$$\det A = A_{11} \cdots A_{nn}. \quad (\text{D.41})$$

Infatti, se A è triangolare, $\forall \sigma \neq \sigma_0$ se $i = \min\{k \mid \sigma_k \neq k\}$, allora esiste $j > i$ tale che $\sigma_j = i$ e dunque nel prodotto $A_{1\sigma_1} \cdots A_{n\sigma_n}$ compare il prodotto $A_{i\sigma_i} A_{j\sigma_j} = A_{i\sigma_i} A_{ji}$ in cui $i < \sigma_i$ e $j > i$ quindi o l'uno o l'altro sono 0. Dunque, l'unico termine non nullo è quello corrispondente a $\sigma = \sigma_0$ ed essendo $\varepsilon_{\sigma_0} = 1$ si ottiene l'asserto.

In particolare, se $A = \operatorname{diag}(\lambda_1, \dots, \lambda_n)$ è **diagonale**, ossia $A_{ij} = 0$ se $i \neq j$ e $A_{ii} = \lambda_i$, allora $\det A = \lambda_1 \cdots \lambda_n$.

²⁷In altri testi, ε_σ si denota $\operatorname{sgn}(\sigma)$.

²⁸Ossia $\varepsilon_{\sigma \cdot \tau} = \varepsilon_\sigma \varepsilon_\tau$ e $\varepsilon_{\sigma_0} = 1$.

- (5) Sia $M \in K^{(n+m) \times (n+m)}$ una matrice con un ‘blocco angolare di $(m \times n)$ zeri’: per esempio, $M_{ij} = 0$ se $i > n \geq j$. Una tale matrice M ha la forma

$$M = \begin{pmatrix} A & B \\ 0 & C \end{pmatrix}$$

con $A \in K^{n \times n}$, $C \in K^{m \times m}$, $B \in K^{n \times m}$ e lo 0 è una matrice $K^{m \times n}$ composta da tutti zeri. Allora,

$$\det M = \det A \cdot \det C. \quad (\text{D.42})$$

Infatti, se $\sigma_i > n$ per un qualche $i \leq n$, allora deve esistere un $1 \leq j \leq m$ tale che²⁹ $\sigma_{n+j} \leq n$, il che implica $M_{(n+j)\sigma_{n+j}} = 0$. Dunque, nella somma su $\sigma \in S_{n+m}$ (che appare nella definizione di determinante di M) si possono escludere tutte le permutazioni con $\sigma_i > n$ per un qualche $i \leq n$, il che significa che la somma è fatta sulle permutazioni $\sigma \in S_{n+m}$ della forma $\sigma = (\bar{\sigma}, n + \hat{\sigma})$ con $\bar{\sigma} \in S_n$ e $\hat{\sigma} \in S_m$. Quindi, osservando che³⁰ $\varepsilon_\sigma = \varepsilon_{\bar{\sigma}} \cdot \varepsilon_{\hat{\sigma}}$, si ha:

$$\begin{aligned} \det M &= \sum_{\substack{\sigma = (\bar{\sigma}, n + \hat{\sigma}) \\ \bar{\sigma} \in S_n, \hat{\sigma} \in S_m}} \varepsilon_\sigma M_{1\bar{\sigma}_1} \cdots M_{n\bar{\sigma}_n} \cdot M_{(n+1)(n+\hat{\sigma}_1)} \cdots M_{(n+m)(n+\hat{\sigma}_m)} \\ &= \sum_{\substack{\sigma = (\bar{\sigma}, n + \hat{\sigma}) \\ \bar{\sigma} \in S_n, \hat{\sigma} \in S_m}} \varepsilon_{\bar{\sigma}} \cdot \varepsilon_{\hat{\sigma}} M_{1\bar{\sigma}_1} \cdots M_{n\bar{\sigma}_n} \cdot M_{(n+1)(n+\hat{\sigma}_1)} \cdots M_{(n+m)(n+\hat{\sigma}_m)} \\ &= \sum_{\bar{\sigma} \in S_n} \varepsilon_{\bar{\sigma}} M_{1\bar{\sigma}_1} \cdots M_{n\bar{\sigma}_n} \cdot \sum_{\hat{\sigma} \in S_m} \varepsilon_{\hat{\sigma}} M_{(n+1)(n+\hat{\sigma}_1)} \cdots M_{(n+m)(n+\hat{\sigma}_m)} \\ &= \det A \cdot \det C. \end{aligned}$$

(xxvi) Dalla definizione di determinante seguono le seguenti fondamentali tre proprietà:

- (d₁) se $\tau \in S_n$ è una trasposizione, allora $\det[v^{(\tau_1)}, \dots, v^{(\tau_n)}] = -\det[v^{(1)}, \dots, v^{(n)}]$; (scambiando due colonne il segno del determinante cambia);
- (d₂) $\det[av + bw, v^{(2)}, \dots, v^{(n)}] = a \det[v, v^{(2)}, \dots, v^{(n)}] + b \det[w, v^{(2)}, \dots, v^{(n)}]$; (il determinante è una funzione lineare nella prima variabile, ossia, nella prima colonna);
- (d₃) $\det[e^{(1)}, \dots, e^{(n)}] = \det(I_n) = 1$.

Le proprietà (d₁) e (d₂) implicano immediatamente anche le seguenti proprietà del determinante:

- (d₄) il determinante è una funzione lineare di una sua qualunque colonna;
- (d₅) se una matrice A ha due colonne uguali, allora $\det A = 0$;
- (d₆) se in una matrice A una colonna è combinazione lineare delle altre colonne, allora $\det A = 0$;
- (d₇) Se $v^{(1)}, \dots, v^{(n)}$ sono linearmente dipendenti, allora³¹ $\det[v^{(1)}, \dots, v^{(n)}] = 0$;
- (d₈) se in una matrice A ad una colonna si aggiunge una qualunque combinazione lineare di tutte le colonne, il valore del determinante non cambia

In effetti le prime tre proprietà *caratterizzano* il determinante:

²⁹ σ_{n+j} per j che varia in $\{1, \dots, m\}$ non può assumere tutti i valori $\{n+1, \dots, n+m\}$ essendo uno di questi valori assunto da σ_i .

³⁰ Cfr. (xxiv).

³¹ Se $\sum \alpha_i v^{(i)} = 0$ e $\alpha_j \neq 0$ per un qualche j , allora $v^{(j)} = \frac{1}{\alpha_j} \sum_{i \neq j} \alpha_i v^{(i)}$ e (d₇) segue da (d₆).

Lemma D.9 Sia $\delta : K^n \rightarrow K$ tale che

(i) se $\tau \in S_n$ è una trasposizione, allora $\delta(v^{(\tau_1)}, \dots, v^{(\tau_n)}) = -\delta(v^{(1)}, \dots, v^{(n)})$;

(ii) $\delta(av + bw, v^{(2)}, \dots, v^{(n)}) = a\delta(v, v^{(2)}, \dots, v^{(n)}) + b\delta(w, v^{(2)}, \dots, v^{(n)})$.

Allora, $\delta = c \det$ con $c := \delta(e^{(1)}, \dots, e^{(n)})$. In particolare, se $\delta(e^{(1)}, \dots, e^{(n)}) = 1$, allora $\delta = \det$.

Dimostrazione Le proprietà (i) e (ii) (come già osservato nel caso del determinante), implicano che δ è lineare in ogni argomento. Dunque, espandendo $v^{(j)}$ nella base canonica, $v^{(j)} = \sum_i v_i^{(j)} e^{(i)}$, si ha, per ogni $(v^{(1)}, \dots, v^{(n)}) \in K^n$:

$$\begin{aligned} \delta(v^{(1)}, \dots, v^{(n)}) &= \delta\left(\sum_{i_1} v_{i_1}^{(1)} e^{(i_1)}, \dots, \sum_{i_n} v_{i_n}^{(n)} e^{(i_n)}\right) \\ &= \sum_{i_1, \dots, i_n} v_{i_1}^{(1)} \dots v_{i_n}^{(n)} \delta(e^{(i_1)}, \dots, e^{(i_n)}) \\ &\stackrel{(*)}{=} \sum_{\sigma \in S_n} v_{\sigma_1}^{(1)} \dots v_{\sigma_n}^{(n)} \delta(e^{(\sigma_1)}, \dots, e^{(\sigma_n)}) \\ &\stackrel{(**)}{=} \sum_{\sigma \in S_n} \varepsilon_\sigma v_{\sigma_1}^{(1)} \dots v_{\sigma_n}^{(n)} \delta(e^{(1)}, \dots, e^{(n)}), \end{aligned}$$

dove nell'uguaglianza (*), abbiamo usato il fatto che se due indici i_j nella somma nella seconda riga, allora da (i) segue che $\delta(e^{(i_1)}, \dots, e^{(i_n)}) = 0$ e dunque la somma su tutti gli indici si restringe ad una somma in cui ogni indice è diverso dall'altro, il che è equivalente a sommare su tutte le permutazioni di S_n ; nell'uguaglianza (**), abbiamo usato il fatto che ogni permutazione σ è prodotto di N trasposizioni e che ogni volta che scambiamo due argomenti di δ il suo valore cambia segno e che $\varepsilon_\sigma = (-1)^N$. ■

Un corollario immediato di questo lemma è

$$(d_9) \det A = \det A^T, \text{ per ogni } A \in K^{n \times n}.$$

Dimostrazione Sia

$$\delta(v^{(1)}, \dots, v^{(n)}) := \det[v^{(1)}, \dots, v^{(n)}]^T = \sum_{\sigma \in S_n} \varepsilon_\sigma v_{\sigma_1}^{(1)} \dots v_{\sigma_n}^{(n)}.$$

Si verifica subito che δ verifica (d_1) , (d_2) e (d_3) e dunque dal Lemma D.9 segue che

$$\delta(v^{(1)}, \dots, v^{(n)}) = \det[v^{(1)}, \dots, v^{(n)}],$$

ossia $\det A^T = \det A$. ■

Osservazione D.10 Poiché la matrice trasposta di A è la matrice che ha come righe le colonne di A , da (d_9) segue immediatamente che le proprietà $(d_1) \div (d_8)$ valgono sostituendo le parole 'colonna/e' con 'riga/he' e il simbolo $[v^{(1)}, \dots, v^{(n)}]$ con il simbolo

$$\begin{bmatrix} v^{(1)} \\ \vdots \\ v^{(n)} \end{bmatrix} =: A, \quad \text{dove } A_{ij} := v_j^{(i)}. \quad (\text{D.43})$$

Un'altra conseguenza importante del Lemma D.9 è:

(d_{10}) il determinante è una funzione moltiplicativa:

$$\det(AB) = \det A \cdot \det B, \quad \forall A, B \in K^{n \times n}, \quad (\text{D.44})$$

e, quindi:

$$\det(AB) = \det(BA), \quad \forall A, B \in K^{n \times n} \quad (\text{D.45})$$

Dimostrazione Fissiamo A , e, per ogni $(v^{(1)}, \dots, v^{(n)})$, poniamo

$$\delta(v^{(1)}, \dots, v^{(n)}) := \det[Av^{(1)}, \dots, Av^{(n)}] \stackrel{\text{(D.35)}}{=} \det(A[v^{(1)}, \dots, v^{(n)}]).$$

Chiaramente, δ verifica (i) e (ii) del Lemma D.9. Dunque $\delta = c \det$ con $c = \delta(I_n)$, ossia $c = \det(A)$, ma questo implica la (D.44). La (D.45) segue immediatamente dalla (D.44). ■

Da (d_{10}) segue la seguente importante caratterizzazione.

(xxvii) È possibile rappresentare una qualunque matrice in $GL(n, K)$ come prodotto di ‘matrici elementari’

Chiameremo **matrice elementare** una matrice del seguente tipo:

- (a) *riscaldamento di un fattore λ nella i -esima direzione:* $R_\lambda^i := [e^{(1)}, \dots, \lambda e^{(i)}, \dots, e^{(n)}]$, dove $1 \leq i \leq n$ e $0 \neq \lambda \in K$;
- (b) *scambio di versori:* $T := [e^{(\tau_1)}, \dots, e^{(\tau_n)}]$ dove $\tau \in S_n$ è una trasposizione; più specificamente, indicheremo con T^{ij} , per $i \neq j$, la matrice associata alla trasposizione $\tau_i = j$ $\tau_j = i$; è conveniente anche porre $T^{ii} := I_n$;
- (c) *somma di versori:*

$$S^{ij} = [e^{(1)}, \dots, \overbrace{e^{(i)} + e^{(j)}}^{\text{colonna } i}, \dots, e^{(n)}], \quad i \neq j.$$

Osservazione D.11 (i) Gli elementi delle matrici elementari sono date da:

$$\begin{aligned} \text{(a)} \quad (R_\lambda^i)_{hk} &= \begin{cases} \delta_{hk}, & \text{se } k \neq i, \\ \lambda \delta_{ik}, & \text{se } k = i; \end{cases} \\ \text{(b)} \quad T_{hk} &= \delta_{h\tau_k}; \\ \text{(c)} \quad (S^{ij})_{hk} &= \delta_{hk} + \delta_{ki} \delta_{hj}. \end{aligned} \tag{D.46}$$

(ii) Dalla definizione di determinante segue che, per ogni i e per ogni $j \neq i$, si ha

$$\det R_\lambda^i = \lambda; \quad \det T^{ij} = -1; \quad \det S^{ij} = 1. \tag{D.47}$$

(iii) Da (D.46) segue che, moltiplicando una matrice arbitraria $[v^{(1)}, \dots, v^{(n)}]$ a destra per le matrici elementari si ottiene:

$$\begin{aligned} [v^{(1)}, \dots, v^{(n)}] \cdot R_\lambda^i &= [v^{(1)}, \dots, \lambda v^{(i)}, \dots, v^{(n)}], \\ [v^{(1)}, \dots, v^{(n)}] \cdot [e^{(\tau_1)}, \dots, e^{(\tau_n)}] &= [v^{(\tau_1)}, \dots, v^{(\tau_n)}], \\ [v^{(1)}, \dots, v^{(n)}] \cdot S^{ij} &= [v^{(1)}, \dots, v^{(i)} + v^{(j)}, \dots, v^{(n)}]. \end{aligned} \tag{D.48}$$

(iv) Sia E una matrice elementare e $A = [v^{(1)}, \dots, v^{(n)}]$ con $v^{(i)}$ linearmente indipendenti. Allora, le colonne di AE sono linearmente indipendenti.

Dimostrazione Nel caso E è di tipo (a) o (b) la tesi è ovvia. Supponiamo $E = S^{ij}$. Da (D.48) segue che le colonne di AE sono date da $\{v^{(h)} \mid h \neq i\} \cup \{v^{(i)} + v^{(j)}\}$. Ora, se questi vettori fossero linearmente dipendenti si avrebbe $0 = \sum_{k \neq i} \alpha_k v^{(k)} + \alpha(v^{(i)} + v^{(j)})$ e $\alpha \neq 0$ (essendo $\{v^{(h)} \mid h \neq i\}$ linearmente indipendente), ma questo significherebbe che $v^{(i)} = \sum_{k \neq i} \alpha^{-1} \alpha_k v^{(k)} - v^{(j)}$, ma (si ricordi che $i \neq j$) questo contraddice il fatto che $\{v^{(k)}\}$ è un insieme linearmente indipendente. ■

Proposizione D.12 Data $A \in GL(n, K)$ esistono $N \geq 1$ matrici elementari $E^{(i)}$ tali che

$$A = E^{(1)} \dots E^{(N)}. \tag{D.49}$$

Dimostrazione Chiaramente, per $n = 1$ l'enunciato è banale, essendo, in tal caso A scalare e $A = R_A^1$. Sia ora $n \geq 2$ e sia $B := A^{-1} =: [v^{(1)}, \dots, v^{(n)}]$. Dimosteremo che esistono $E^{(i)}$ elementari tali che $B \cdot E^{(1)} \cdots E^{(N)} = I_n$ il che è equivalente a (D.49).

Primo passo: Poiché i vettori $\{v^{(i)}\}$ sono indipendenti³², formano una base; quindi esistono: $1 \leq m \leq n$, $1 \leq j_i \leq m$ strettamente crescenti in i e $\alpha_i \neq 0$ tali che

$$e^{(1)} = \sum_{i=1}^m \alpha_i v^{(j_i)}. \quad (\text{D.50})$$

Dunque, da (D.48) segue che

$$\begin{aligned} B^1 &:= [v^{(1)}, \dots, v^{(n)}] \cdot R_{\alpha_1}^{j_1} \cdots R_{\alpha_m}^{j_m} = [\dots, \alpha_{j_1} v^{(j_1)}, \dots, \alpha_{j_m} v^{(j_m)}, \dots], \\ B^2 &:= B^1 \cdot S^{j_1 j_2} \cdots S^{j_1 j_m} = [\dots, e^{(1)}, \dots, \alpha_{j_m} v^{(j_m)}, \dots], \\ B^3 &:= B^2 T^{1 j_1} =: [e^{(1)}, u^{(2)}, \dots, u^{(n)}]. \end{aligned} \quad (\text{D.51})$$

dove, in vista dell'Osservazione D.11–(iv), i vettori $u^{(i)}$ definiti nell'ultima uguaglianza, formano assieme al versore $e^{(1)}$, una base; quindi, in particolare, $B^3 \in GL(n, K)$.

Secondo passo (iterativo): Sia $1 \leq i \leq n - 1$ e assumiamo $B^4 \in GL(n, K)$ sia della forma

$$B^4 := [e^{(1)}, \dots, e^{(i-1)}, u^{(i)}, \dots, u^{(n)}]. \quad (\text{D.52})$$

Allora, poiché $\{e^{(1)}, \dots, e^{(i-1)}, u^{(i)}, \dots, u^{(n)}\}$ è una base, esistono:

$$0 \leq p \leq i - 1, \text{ e se } p > 0, 1 \leq \ell_k \leq p \text{ strettamente crescenti in } k, \beta_k \neq 0,$$

$$1 \leq m \leq n - 1 + i, i \leq j_k \leq n \text{ strettamente crescenti in } k, \gamma_k \neq 0,$$

tali che:

$$e^{(i)} = \sum_{k=1}^p \beta_k e^{(\ell_k)} + \sum_{k=1}^m \gamma_k u^{(j_k)}, \quad (\text{D.53})$$

dove, se $p = 0$, la prima somma è assente (ma che $m \geq 1$ segue dal fatto che $e^{(i)}$ è indipendente da $\{e^{(1)}, \dots, e^{(i-1)}\}$). Con passaggi del tutto analoghi a quelli fatti in (D.51), si ottiene

$$B^5 := B^4 \cdot \tilde{R}^1 \cdots \tilde{R}^q \cdot \tilde{S}^1 \cdots \tilde{S}^{q-1} \cdot T = [e^{(1)}, \dots, e^{(i)}, w^{(i+1)}, \dots, w^{(n)}],$$

dove, $q = m + p$, \tilde{R}^k sono della forma $R_{\beta_k}^{\ell_k}$ o $R_{\gamma_k}^{j_k}$ (se $p > 0$) oppure (se $p = 0$) $\tilde{R}_{\gamma_k}^{j_k}$, \tilde{S}^k sono della forma $S^{\ell_1 j}$ se $p > 0$ oppure $S^{j_1 j}$ se $p = 0$ e $T = T^{ij}$. Se $i = n - 1$, i vettori $w^{(j)}$ sono assenti, $B^5 = I_n$ e il lemma è dimostrato, altrimenti (essendo $B^5 \in GL(n, K)$) si itera il passo due (un numero finito di volte) fino a ottenere la matrice I_n . ■

(xxviii) Formula di Laplace per il calcolo del determinante e formula per la matrice inversa

Data una matrice $A \in K^{n \times n}$, con $n \geq 2$ e dati $1 \leq i, j \leq n$, definiamo $A^{(i,j)} \in K^{(n-1) \times (n-1)}$ la matrice ottenuta 'cancellando da A la i -esima riga e la j -esima colonna', ossia, se $A = [v^{(1)}, \dots, v^{(n)}]$ e se³³

$$\bar{v}^{(k)} := (v_1^{(k)}, \dots, v_{i-1}^{(k)}, v_{i+1}^{(k)}, \dots, v^{(n)}),$$

allora,

$$A^{(i,j)} := [\bar{v}^{(1)}, \dots, \bar{v}^{(j-1)}, \bar{v}^{(j+1)}, \dots, \bar{v}^{(n)}].$$

Definiamo poi il **minore** (i, j) di A come il determinante di $A^{(i,j)}$:

$$m_{ij} := m_{ij}(A) := \det A^{(i,j)}. \quad (\text{D.54})$$

³²Si ricordi che $[v^{(1)}, \dots, v^{(n)}] \in GL(n, K) \iff$ i vettori $v^{(i)}$ sono linearmente indipendenti: cfr. (xxiii).

³³Naturalmente, se $i = 1$, $\bar{v}^{(1)} := [v^{(2)}, \dots, v^{(n)}]$ e se $i = n$, $\bar{v}^{(n)} = [v^{(1)}, \dots, v^{(n-1)}]$.

Allora, per ogni $A \in K^{n \times n}$, per ogni $1 \leq i \leq n$ si ha

$$\det A = \sum_{j=1}^n (-1)^{i+j} A_{ij} m_{ij}. \quad (\text{D.55})$$

La dimostrazione della formula (D.55) si basa sul seguente lemma combinatorio. Chiamiamo $S_n^{ij} := \{\sigma \in S_n \mid \sigma_i = j\}$ e osserviamo che S_n è l'unione disgiunta per $1 \leq j \leq n$ di S_n^{ij} .

Lemma D.13 Se $\sigma \in S_n^{ij}$, allora

$$\varepsilon_\sigma = (-1)^{i+j} \varepsilon_{\hat{\sigma}} \quad (\text{D.56})$$

dove $\hat{\sigma}$ è la permutazione di $n-1$ oggetti $\hat{\sigma} : k \in J \mapsto \hat{\sigma}_k := \sigma_k \in J$, dove $J := \{k \in \mathbb{N} \mid k \leq n \text{ e } k \neq j\}$.

Dimostrazione Dalle definizioni di ε_σ , S_n^{ij} e $\hat{\sigma}$, segue che

$$\begin{aligned} \varepsilon_\sigma &:= \prod_{h < k} \text{sgn}(\sigma_k - \sigma_h) = \prod_{\substack{h < k \\ h, k \neq i}} \text{sgn}(\sigma_k - \sigma_h) \cdot \prod_{\substack{h < k \\ h=i}} \text{sgn}(\sigma_k - \sigma_h) \cdot \prod_{\substack{h < k \\ k=i}} \text{sgn}(\sigma_k - \sigma_h) \\ &= \prod_{\substack{h < k \\ h, k \neq i}} \text{sgn}(\sigma_k - \sigma_h) \cdot \prod_{h > i} \text{sgn}(\sigma_h - \sigma_i) \cdot \prod_{h < i} \text{sgn}(\sigma_i - \sigma_h) \\ &= \varepsilon_{\hat{\sigma}} \cdot \prod_{h > i} \text{sgn}(\sigma_h - j) \cdot \prod_{h < i} \text{sgn}(j - \sigma_h) \end{aligned} \quad (\text{D.57})$$

Ora, sia $p = \#\{h < i \mid \sigma_h < j\}$ e $q := \#\{h > i \mid \sigma_h < j\}$. Chiaramente,

$$\prod_{h > i} \text{sgn}(\sigma_h - j) = (-1)^q, \quad \prod_{h < i} \text{sgn}(j - \sigma_h) = (-1)^{i-1-p}.$$

Essendo, $p + q = j - 1$, da (D.57) segue la tesi. \blacksquare

Dimostrazione della formula di Laplace (D.55):

$$\begin{aligned} \det A &:= \sum_{\sigma \in S_n} \varepsilon_\sigma \prod_{k=1}^n A_{k\sigma_k} = \sum_{j=1}^n \sum_{\sigma \in S_n^{ij}} \varepsilon_\sigma \prod_{k=1}^n A_{k\sigma_k} \\ &= \sum_{j=1}^n A_{ij} \sum_{\sigma \in S_n^{ij}} \varepsilon_\sigma \prod_{k \neq j} A_{k\sigma_k} \\ &\stackrel{(\text{D.56})}{=} \sum_{j=1}^n (-1)^{i+j} A_{ij} \sum_{\sigma \in S_n^{ij}} \varepsilon_{\hat{\sigma}} \prod_{k \neq j} A_{k\sigma_k} \\ &= \sum_{j=1}^n (-1)^{i+j} A_{ij} m_{ij}. \quad \blacksquare \end{aligned}$$

Osservazione D.14 Dalla definizione di S_n^{ij} e di minore (i, j) segue immediatamente che $m_{ij}(A^T) = m_{ji}(A)$, dunque essendo $\det A^T = \det A$ otteniamo la seguente ‘versione trasposta’ della formula di Laplace³⁴

$$\begin{aligned} \det A &= \det A^T = \sum_{j=1}^n (-1)^{i+j} A_{ij}^T m_{ij}(A^T) \\ &= \sum_{j=1}^n (-1)^{i+j} A_{ji} m_{ji}. \end{aligned} \quad (\text{D.58})$$

³⁴Tale formula viene a volte chiamata ‘espansione di Laplace per colonne’, mentre la (D.55) viene chiamata ‘espansione di Laplace per righe’.

Vediamo, ora, come dalla formula di Laplace segue una formula analitica per il calcolo dell'inversa di una matrice con determinante non nullo.

Definiamo la **matrice dei cofattori di A** come la matrice $C = C(A)$ con elementi di matrice $C_{ij} := (-1)^{i+j} m_{ij}$ e la **matrice aggiuntiva**³⁵ di A , $\text{adj}(A)$, come la trasposta della matrice dei cofattori di A :

$$\text{adj}(A) := C^T. \quad (\text{D.59})$$

Proposizione D.15 Per ogni $A \in K^{n \times n}$, si ha

$$A \text{adj}(A) = \text{adj}(A) A = (\det A) I_n. \quad (\text{D.60})$$

Dimostrazione Per ogni i si ha

$$(A \text{adj}(A))_{ii} = \sum_{j=1}^n A_{ij} \text{adj}(A)_{ji} := \sum_{j=1}^n (-1)^{i+j} A_{ij} m_{ij} \stackrel{(\text{D.55})}{=} \det A.$$

Analogamente,

$$(\text{adj}(A) A)_{ii} = \sum_{j=1}^n \text{adj}(A)_{ij} A_{ji} := \sum_{j=1}^n (-1)^{i+j} A_{ji} m_{ji} \stackrel{(\text{D.58})}{=} \det A.$$

Se $i \neq k$,

$$(A \text{adj}(A))_{ik} = \sum_{j=1}^n A_{ij} \text{adj}(A)_{jk} := \sum_{j=1}^n (-1)^{i+j} A_{ij} m_{kj} \stackrel{(\text{D.55})}{=} \det \hat{A},$$

dove \hat{A} è la matrice ottenuta da A sostituendo la k -esima riga con la i -esima riga. Poiché \hat{A} ha due righe uguali (la i -esima e la k -esima), per l'Osservazione D.10, $\det \hat{A} = 0$. ■

Da (D.60) segue immediatamente che

$$\det A \neq 0 \quad \implies \quad A \in GL(n, K) \quad \text{e} \quad A^{-1} = \frac{1}{\det A} \text{adj}(A). \quad (\text{D.61})$$

Dunque si ha la seguente importante caratterizzazione:

Proposizione D.16 $A \in GL(n, K)$ (ossia, A è invertibile) $\iff \det A \neq 0$, o, equivalentemente,

$$GL(n, K) = \{A \in K^{n \times n} \mid \det A \neq 0\}. \quad (\text{D.62})$$

Se $A \in GL(n, K)$, $\det(A^{-1}) = (\det A)^{-1}$.

Dimostrazione Se A è invertibile allora esiste A^{-1} tale che $AA^{-1} = I_n$ e da (d₁₀) segue, che, essendo $\det I_n = 1$, $1 = \det I_n = \det(AA^{-1}) = \det A \cdot \det(A^{-1})$ e quindi $\det A \neq 0$ e $\det A^{-1} = 1/\det A$.

Viceversa, se $\det A \neq 0$ allora da (D.61) segue che A è invertibile e, come prima, per (d₁₀), si ha che $\det A^{-1} = 1/\det A$. ■

(xxix) Due matrici $A, B \in K^{n \times n}$ si dicono **simili** se esiste $U \in GL(n, K)$ tale che $A = U^{-1}BU$; in tal caso, diremo anche che U coniuga B a A . Da (d₁₀) e dalla Proposizione D.16 segue che due matrici simili hanno lo stesso determinante.

³⁵Adjugate matrix in inglese.

D.5 Matrici e applicazioni lineari

(xxx) Ad una matrice $A \in K^{n \times m}$ si può associare *in modo canonico* una applicazione lineare $L_A \in \mathcal{L}(K^m, K^n)$ come segue³⁶

$$L_A : x = (x_1, \dots, x_m) \in K^m \mapsto y = (y_1, \dots, y_n) := L_A x \in K^n, \quad (\text{D.63})$$

$$\text{dove } y_i := \sum_{j=1}^m A_{ij} x_j, \quad 1 \leq i \leq n.$$

È uso comune identificare la matrice A con la mappa lineare L_A : scriveremo, quindi, Ax al posto di $L_A x$.

Osservazione D.17 (i) Se $x \in K^m$ si identifica con il vettore colonna $[x] \in K^{m \times 1}$, allora $L_A x$ non è altro che il vettore colonna $(n \times 1)$ ottenuto dal prodotto righe per colonne delle matrici, $A \in K^{n \times m}$ con $x = [x] \in K^{m \times 1}$, ossia la matrice Ax .

(ii) La parola ‘canonico’ usata in (xxx) si riferisce al fatto che nella corrispondenza $L_A \longleftrightarrow A$ abbiamo fissato a priori spazi vettoriali canonici, ossia, K^m e K^n , insieme alle loro basi canoniche $\{e^{(j)}\}$.

(xxxii) Più in generale, è possibile far corrispondere ad una applicazione lineare $T \in \mathcal{L}(V, W)$ tra due spazi vettoriali V e W di dimensione m e n con campo di scalari K e due basi $\beta \subseteq V$ e $\beta' \subseteq W$ fissate una matrice $A \in K^{n \times m}$:

Siano, infatti, $\beta := \{v_1, \dots, v_m\}$ e $\beta' := \{w_1, \dots, w_n\}$ due basi di, rispettivamente, V e W e $T \in \mathcal{L}(V, W)$. Per ogni $1 \leq j \leq m$, il vettore $Tv_j \in W$ ha un’unica espansione nella base β' ,

$$Tv_j = \sum_{i=1}^n a_i^{(j)} w_i =: \sum_{i=1}^n A_{ij} w_i, \quad (\text{D.64})$$

e i numeri A_{ij} formano una matrice $(n \times m)$ che rende commutativo il seguente diagramma:

$$\begin{array}{ccc} V & \xrightarrow{T} & W \\ \downarrow j_\beta & & \downarrow j_{\beta'} \\ K^m & \xrightarrow{A} & K^n \end{array} \quad (\text{D.65})$$

ossia:

$$j_{\beta'} \circ T = A \circ j_\beta \quad \iff \quad A = j_{\beta'} \circ T \circ j_\beta^{-1}. \quad (\text{D.66})$$

Infatti, ricordando la definizione di isomorfismo canonico j_β in (D.28), se $v = \sum_{j=1}^m x_j v_j$, si ha

$$j_\beta(v) = x = (x_1, \dots, x_m) \text{ e}$$

$$Tv = \sum_{j=1}^m x_j Tv_j \stackrel{(\text{D.64})}{=} \sum_{j=1}^m x_j \sum_{i=1}^n A_{ij} w_i = \sum_{i=1}^n \left(\sum_{j=1}^m A_{ij} x_j \right) w_i$$

e quindi $j_{\beta'}(Tv) = Ax = A j_\beta(v)$, ossia (D.66)

Denotiamo la matrice A così ottenuta con $\text{mat}_{\beta, \beta'}(T)$. Chiaramente, la mappa

$$T \in \mathcal{L}(V, W) \mapsto A = \text{mat}_{\beta, \beta'}(T) \in K^{n \times m} \quad (\text{D.67})$$

è un isomorfismo di spazi vettoriali³⁷. Nel caso $n = m$, sceglieremo $\beta = \beta'$ e denoteremo $\text{mat}_{\beta, \beta'}$ con mat_β .

³⁶La linearità della mappa definita in (D.63) è ovvia.

³⁷In particolare, $\dim \mathcal{L}(V, W) = \dim K^{n \times m} = nm$.

(xxxii) **(Cambi di base)** Sia V uno spazio vettoriale di dimensione n su K . Se $T \in \mathcal{L}(V)$ e se β e β' sono basi di V , allora $\text{mat}_\beta(T)$ e $\text{mat}_{\beta'}(T)$ sono matrici simili.

Dimostrazione Siano $\beta = \{u_1, \dots, u_n\}$ e $\beta' = \{u'_1, \dots, u'_n\}$. Per ogni j , espandendo, rispettivamente, u_j nella base β' e u'_j nella base β si ha, per opportuni $U'_{ij}, U_{ij} \in K$,

$$u_j = \sum_{i=1}^n U'_{ij} u'_i, \quad u'_j = \sum_{i=1}^n U_{ij} u_i. \quad (\text{D.68})$$

I numeri U'_{ij} e U_{ij} costituiscono gli elementi di due matrici $U', U \in K^{n \times n}$ che sono una l'inversa dell'altra: infatti, per ogni j si ha

$$u_j = \sum_{i=1}^n U'_{ij} u'_i = \sum_{i,k=1}^n U'_{ij} U_{ki} u_k = \sum_{k=1}^n \left(\sum_{i=1}^n U_{ki} U'_{ij} \right) u_k,$$

relazione che implica $\sum_{i=1}^n U_{ki} U'_{ij} = \delta_{kj}$, ossia $UU' = I_n$. Sia, ora, $A = \text{mat}_\beta(T)$ e $B = \text{mat}_{\beta'}(T)$. Da (D.64) segue che

$$Tu_j = \sum_{i=1}^n A_{ij} u_i, \quad Tu'_j = \sum_{i=1}^n B_{ij} u'_i. \quad (\text{D.69})$$

Dunque dalla definizione di U' e U e da (D.69) segue che

$$\begin{aligned} \sum_{h=1}^n B_{hj} u'_h &\stackrel{(\text{D.69})}{=} Tu'_j \stackrel{(\text{D.68})}{=} \sum_{i=1}^n U_{ij} Tu_i \stackrel{(\text{D.69})}{=} \sum_{i,k=1}^n U_{ij} A_{ki} u_k \stackrel{(\text{D.68})}{=} \sum_{i,k,h=1}^n U_{ij} A_{ki} U'_{hk} u'_h \\ &= \sum_{h=1}^n \left(\sum_{k,i=1}^n U'_{hk} A_{ki} U_{ij} \right) u'_h = \sum_{h=1}^n (U'AU)_{hj} u'_h. \end{aligned}$$

da cui segue $B = U'AU$, ossia, $B = U^{-1}AU$. ■

Osservazione D.18 Osservando che la seconda uguaglianza in (D.68) può essere scritta formalmente come³⁸ $\hat{\beta}' = U^T \hat{\beta}$, dove $\hat{\beta}$ è la n -pla di vettori (u_1, \dots, u_n) e $\hat{\beta}'$ è la n -pla di vettori (u'_1, \dots, u'_n) la **formula del cambio di basi** può essere scritta come:

$$\text{mat}_{\beta'}(T) = U^{-1} \text{mat}_\beta(T) U, \quad \hat{\beta}' = U^T \hat{\beta}. \quad (\text{D.70})$$

Esempio Sia $V = \mathbb{R}_2[x]$ lo spazio vettoriale dei polinomi su \mathbb{R} di grado al più 2 e sia T il ‘troncamento’ $T : p(x) = a_0 + a_1x + a_2x^2 \mapsto a_0 + a_1x$. Chiaramente, $T \in \mathcal{L}(V)$. Fissiamo in V le seguenti due basi: $\beta := \{1, x, x^2\}$ e $\beta' := \{1, 2x - 1, x + x^2\}$. L’isomorfismo canonico $j_\beta : V \rightarrow \mathbb{R}^3$ è dato da $j_\beta(a_0 + a_1x + a_2x^2) = (a_0, a_1, a_2)$, mentre, l’isomorfismo canonico $j_{\beta'}$ è dato da

$$j_{\beta'}(a_0 + a_1x + a_2x^2) = \left(a_0 + \frac{a_1 - a_2}{2}, \frac{a_1 - a_2}{2}, a_2 \right).$$

Da (D.66) e dalla relazione $\hat{\beta}' = U^T \hat{\beta}$ segue che

$$A := \text{mat}_\beta(T) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B := \text{mat}_{\beta'}(T) = \begin{pmatrix} 1 & 0 & \frac{1}{2} \\ 0 & 1 & \frac{1}{2} \\ 0 & 0 & 0 \end{pmatrix}, \quad U = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 2 & 1 \\ 0 & 1 & 1 \end{pmatrix},$$

e, come si può anche verificare direttamente, U coniuga A a B cioè, $B = U^{-1}AU$.

³⁸ $\hat{\beta}$ è una n -pla di vettori di V , mentre β è un insieme di vettori di V (di cardinalità n); ‘formalmente’ significa che la relazione $\hat{\beta}' = U^T \hat{\beta}$ è definita dalla seconda uguaglianza in (D.68).

(xxxiii) Se $A = [v^{(1)}, \dots, v^{(m)}] \in K^{n \times m}$, ($v^{(j)} \in K^n$), e $x \in K^m$, allora

$$Ax = \sum_{j=1}^m x_j v^{(j)}. \quad (\text{D.71})$$

Dunque, $\text{im}(A) := \text{im}(L_A) := \{Ax \mid x \in K^m\}$ è il sottospazio vettoriale di K^n generato dai vettori colonna (dalle colonne) di A e il **rango** di L_A , $\text{rank}(L_A) := \text{rank}(A)$, per definizione, ne è la dimensione, che quindi coincide con il massimo numero di colonne indipendenti di A .

Dal Teorema nullità–rango D.4 segue che: una matrice $A \in K^{n \times n}$ è invertibile se e solo $\text{rank}(A) = n$.

Possiamo, ora, raccogliere, varie caratterizzazioni dell'invertibilità di una matrice:

(xxxiv) (**condizioni di invertibilità**) Sia $A = [v^{(1)}, \dots, v^{(n)}] \in K^{n \times n}$. Allora:

$\det A \neq 0 \stackrel{(\text{D.62})}{\iff} A \in GL(n, K) \stackrel{(\text{xxx})}{\iff} L_A \text{ è un isomorfismo su } K^n \stackrel{(\text{xxxiii})}{\iff} \text{rank}(A) = n$

$\stackrel{\text{def}}{\iff} \langle v^{(1)}, \dots, v^{(n)} \rangle = K^n \stackrel{(\text{D.71})}{\iff} \text{ l'unica soluzione } x \in K^n \text{ dell'equazione } Ax = 0 \text{ è } x = 0.$

(xxxv) Se $A \in K^{n \times n}$ e $\{e^{(i)}\}$ è la base ortonormale standard di K^n , allora è immediato verificare che

$$Ae^{(i)} \cdot e^{(j)} = A_{ij}, \quad (K = \mathbb{R}); \quad \langle Ae^{(i)}, e^{(j)} \rangle = A_{ij}, \quad (K = \mathbb{C}). \quad (\text{D.72})$$

(xxxvi) Se $A \in \mathbb{R}^{n \times n}$ o $A \in \mathbb{C}^{n \times n}$, allora (come si verifica immediatamente) si ha, rispettivamente, che³⁹

$$Au \cdot v = u \cdot A^T v, \quad (\forall u, v \in \mathbb{R}^n); \quad \langle Au, v \rangle = \langle u, A^* v \rangle, \quad (\forall u, v \in \mathbb{C}^n). \quad (\text{D.73})$$

(xxxvii) $A \in \mathbb{R}^{n \times n}$ è simmetrica⁴⁰ se e solo se $Au \cdot v = u \cdot Av$ per ogni $u, v \in \mathbb{R}^n$. Analogamente, $A \in \mathbb{C}^{n \times n}$ è autoaggiunta se e solo se $\langle Au, v \rangle = \langle u, Av \rangle$ per ogni $u, v \in \mathbb{C}^n$.

Dimostrazione Verifichiamo il caso complesso (il caso reale è del tutto analogo). Se $A^* = A$, da (D.73) segue immediatamente che $\langle Au, v \rangle = \langle u, Av \rangle$ per ogni $u, v \in \mathbb{C}^n$. Viceversa se $\langle Au, v \rangle = \langle u, Av \rangle$ per ogni $u, v \in \mathbb{C}^n$, prendendo $u = e^{(i)}$ e $v = e^{(j)}$, per (xxxv), si ha che

$$A_{ij} = \langle Ae^{(i)}, e^{(j)} \rangle = \langle e^{(i)}, Ae^{(j)} \rangle = \overline{\langle Ae^{(j)}, e^{(i)} \rangle} = \overline{A_{ji}} = A_{ij}^*. \quad \blacksquare$$

(xxxviii) Una matrice $U \in GL(n, \mathbb{R})$ si dice **ortogonale** se $U^{-1} = U^T$ o, equivalentemente, se $U^T U = I_n$. Se $U = [u^{(1)}, \dots, u^{(n)}]$, la relazione $U^T U = I_n$ è equivalente a $u^{(i)} \cdot u^{(j)} = \delta_{ij}$. Dunque:

una matrice è ortogonale se e solo se le sue colonne formano una base ortonormale di \mathbb{R}^n .

Una matrice $U \in GL(n, \mathbb{C})$ si dice **unitaria** se $U^{-1} = U^*$ o, equivalentemente, se $U^* U = I_n$. Se $U = [u^{(1)}, \dots, u^{(n)}]$, la relazione $U^* U = I_n$ è equivalente a $\langle u^{(j)}, u^{(i)} \rangle = \delta_{ij}$. Dunque:

una matrice è unitaria se e solo se le sue colonne formano una base ortonormale di \mathbb{C}^n .

L'insieme delle matrici ortogonali in $GL(n, \mathbb{R})$ forma (come è immediato verificare) un sottogruppo di $GL(n, \mathbb{R})$ che si denota con $O(n)$.

L'insieme delle matrici unitarie in $GL(n, \mathbb{C})$ forma (come è immediato verificare) un sottogruppo di $GL(n, \mathbb{C})$ che si denota con $U(n)$.

Le matrici ortogonali/unitarie sono le matrici che conservano il prodotto scalare/hermitiano:

- (i) $U \in O(n)$ se e solo se $Uu \cdot Uv = u \cdot v$, $\forall u, v \in \mathbb{R}^n$; ossia le matrici ortogonali sono le matrici che conservano il prodotto scalare in \mathbb{R}^n .

³⁹Si ricordi (xx).

⁴⁰Si ricordi (xx).

- (ii) $U \in U(n)$ se e solo se \mathbb{C}^n : $\langle Uu, Uv \rangle = \langle u, v \rangle$, $\forall u, v \in \mathbb{C}^n$; ossia, le matrici unitarie sono le matrici che conservano il prodotto hermitiano in \mathbb{C}^n .

Dimostrazione Consideriamo il caso complesso (il caso reale è analogo): se $U \in U(n)$, $\langle Uu, Uv \rangle = \langle u, U^*Uv \rangle = \langle u, v \rangle$, per ogni $u, v \in \mathbb{C}^n$. Viceversa se $\langle Uu, Uv \rangle = \langle u, v \rangle$ per ogni $u, v \in \mathbb{C}^n$, si ha che $\langle u, U^*Uv \rangle = \langle u, v \rangle$ per ogni $u, v \in \mathbb{C}^n$, e prendendo $u = e^{(i)}$, $v = e^{(j)}$ si ha che $(U^*U)_{ij} = \delta_{ij}$, ossia, $U^*U = I_n$ e quindi $U \in U(n)$. ■

(xxxix) (**Polinomio caratteristico**) Data una matrice $A \in K^{n \times n}$ il polinomio in $z \in K$ definito da

$$p_{\text{car}}(z; A) := \det(A - zI_n) = 0, \quad (\text{D.74})$$

prende il nome di **polinomio caratteristico di A** (su K); se non è necessario esprimere la dipendenza da A , scriveremo $p_{\text{car}}(z)$ al posto di $p_{\text{car}}(z; A)$.

Ricordando la definizione di⁴¹ spettro (data per una applicazione lineare generale), e ricordando che $\det(A - \lambda I_n) = 0$ se e solo se l'equazione $(A - \lambda I_n)u = 0$ ha una soluzione non banale⁴² $u \in K^n \setminus \{0\}$, segue che *le radici del polinomio caratteristico sono gli autovalori di A* :

$$\lambda \in \sigma(A) \iff p_{\text{car}}(\lambda) = 0. \quad (\text{D.75})$$

Se $\lambda \in \sigma(A)$, allora⁴³ $(z - \lambda) | p_{\text{car}}(z)$ e il numero

$$a(\lambda) := \max\{1 \leq j \leq n \text{ tale che } (\lambda - z)^j | p_{\text{car}}(z)\} \quad (\text{D.76})$$

prende il nome di **molteplicità algebrica** (su K) dell'autovalore $\lambda \in \sigma(A)$.

Vediamo alcune proprietà del polinomio caratteristico.

- (i) *Lo spettro è un invariante per matrici simili*: se A e B sono due matrici quadrate simili, allora $p_{\text{car}}(z; A) = p_{\text{car}}(z; B)$ e quindi $\sigma(A) = \sigma(B)$.

Dimostrazione Se $B = U^{-1}AU$, allora

$$\begin{aligned} p_{\text{car}}(z; A) &= \det(A - zI_n) \stackrel{(\text{D.44})}{=} \det(U^{-1}(A - zI_n)U) = \det(U^{-1}AU - zI_n) \\ &= \det(B - zI_n) = p_{\text{car}}(z; B). \quad \blacksquare \end{aligned}$$

- (ii) Dal punto precedente e dalla formula di cambio di base (D.70) segue che possiamo definire il **polinomio caratteristico $p_{\text{car}}(z; T)$ di un qualunque endomorfismo $T \in \mathcal{L}(V)$** , V spazio vettoriale n dimensionale su K , come $\det(A - zI_n)$ dove $A = \text{mat}_\beta(T)$ è la matrice rispetto ad una *qualunque base β di V* .

Questo permette di estendere tutte le considerazioni sullo spettro di una matrice al caso più generale di endomorfismi lineari.

- (iii) Il polinomio caratteristico di $A \in K^{n \times n}$ è un polinomio di grado n della forma

$$p_{\text{car}}(z) = \alpha_n z^n + \alpha_{n-1} z^{n-1} + \dots + \alpha_0. \quad (\text{D.77})$$

Dal punto precedente segue che *i coefficienti $\alpha_k := \alpha_k(A) \in K$ sono degli invarianti di A* , ossia, non cambiano per coniugazioni a matrici simili. In particolare, *gli invarianti α_n, α_{n-1} e α_0 sono dati da*:

$$\alpha_n = (-1)^n, \quad \alpha_{n-1} = (-1)^{n-1} \sum_{i=1}^n A_{ii}, \quad \alpha_0 = \det A. \quad (\text{D.78})$$

⁴¹Cfr. (xvii).

⁴²Cfr (xxxiv).

⁴³ $(z - \lambda)$ divide $p_{\text{car}}(z)$, ossia, $p_{\text{car}}(z) = (z - \lambda)Q(z)$ con Q polinomio in z di grado $n - 1$.

Dimostrazione $p_{\text{car}}(0) = \det A$ e quindi $\alpha_0 = \det A$. Dimostriamo le prime due relazioni per induzione su n . Per $n = 1$, si ha $A = A_{11} \in K$ e $p_{\text{car}}(z) = A - z$ e dunque $\alpha_1 = -1 = (-1)^n$ e $\alpha_0 = A = \det A$ (si noti che $\alpha_{n-1}z^{n-1} = \alpha_0$). Assumiamo ora che (D.77) valga per $n \geq 1$ e dimostriamola con $n + 1$ al posto di n . Sia dunque $A \in K^{(n+1) \times (n+1)}$. Usando lo sviluppo di Laplace del determinante (D.55) con $i = 1$ (‘sviluppo rispetto alla prima riga’) e denotando con $O(z^k)$ un polinomio in z di ordine al più k , otteniamo

$$p_{\text{car}}(A - \lambda I_{n+1}) = (A_{11} - z) \cdot m_{11}(A) + O(z^{n-1}),$$

dove $m_{11}(A) = \det(A^{(1,1)} - \lambda I_n)$ con $A^{(1,1)} \in K^{n \times n}$ ottenuta da A cancellando la prima colonna e la prima riga. Per ipotesi induttiva,

$$m_{11}(A) = (-1)^n z^n + (-1)^{n-1} z^{n-1} \sum_{i=2}^{n+1} A_{ii} + O(z^{n-2}),$$

dove se $n = 1$, il termine $O(z^{n-2})$ è assente. Svolgendo l'algebra, si ottiene

$$p_{\text{car}}(A - \lambda I_{n+1}) = (-1)^{n+1} z^{n+1} + (-1)^n \sum_{i=1}^{n+1} A_{ii} + O(z^{n-1}). \quad \blacksquare$$

(iv) Il numero $\text{tr}(A) := \sum_{i=1}^n A_{ii}$, che appare in (D.78), prende il nome di **traccia** di $A \in K^{n \times n}$.

È immediato verificare direttamente che la traccia verifica le seguenti tre proprietà⁴⁴:

$$\text{tr}(A + B) = \text{tr}(A) + \text{tr}(B), \quad \text{tr}(aA) = a \text{tr}(A), \quad \text{tr}(AB) = \text{tr}(BA). \quad (\text{D.79})$$

(v) Il polinomio caratteristico di una matrice 2×2 , come è immediato verificare, è dato da

$$p_{\text{car}}(z) = z^2 - \text{tr}A z + \det A, \quad (A \in K^{2 \times 2}), \quad (\text{D.80})$$

e quindi gli autovalori dipendono solo dai due invarianti $\text{tr}(A)$ e $\det A$.

(vi) Se K non è algebricamente chiuso, lo spettro di una matrice può essere vuoto: per esempio, se $A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \in \mathbb{R}^{2 \times 2}$, $p_{\text{car}}(z) = z^2 + 1$ che non ha radici reali e quindi $\sigma_{\mathbb{R}}(A) = \emptyset$. D'altra parte, $A \in \mathbb{C}^{2 \times 2}$ e, su \mathbb{C} , p_{car} ha le due radici $\pm i$ e dunque $\sigma_{\mathbb{C}}(A) = \{i, -i\}$. Si noti che $A^2 = -I_2$.

In generale, se $K = \mathbb{C}$, dal teorema fondamentale dell'algebra, segue che esistono (e sono unici) $1 \leq m \leq n$, $a_j \in \mathbb{N}$ e $\lambda_j \in \mathbb{C}$ tali che

$$p_{\text{car}}(z) = (-1)^n (z - \lambda_1)^{a_1} \cdots (z - \lambda_m)^{a_m}, \quad a_1 + \cdots + a_m = n, \quad \lambda_i \neq \lambda_j \text{ se } i \neq j, \quad (\text{D.81})$$

cosicché $\sigma(A) = \{\lambda_1, \dots, \lambda_m\}$ e $a_i = a(\lambda_i)$ è la molteplicità algebrica di λ_i . Inoltre, da (D.77), (D.78) e da (D.81) segue che

$$\det A = \prod_{j=1}^m \lambda_j^{a_j}, \quad \sum_{i=1}^n A_{ii} = \sum_{i=1}^m \lambda_i. \quad (\text{D.82})$$

In particolare, un endomorfismo lineare su uno spazio vettoriale su \mathbb{C} ha sempre almeno un autovalore.

⁴⁴Si noti che l'ultima proprietà dà una dimostrazione alternativa che la traccia è un invariante per matrici simili.

D.6 Sottospazi invarianti e matrici a blocchi

(xl) Date m matrici quadrate $A^{(i)}$ di ordine n_i , $\text{block}(A^{(1)}, \dots, A^{(m)})$ denota la **matrice a blocchi** B di ordine $\bar{n} = \sum_{j=1}^m n_j$, con le matrici $A^{(i)}$ (in ordine) sui blocchi centrali: più precisamente gli elementi di $B := \text{block}(A^{(1)}, \dots, A^{(m)}) \in K^{\bar{n} \times \bar{n}}$ sono definiti come segue: sia $\bar{n}_0 := 0$ e per $1 \leq k \leq m$, $\bar{n}_k := \sum_{h=1}^k n_s$, ($\bar{n}_m = \bar{n}$), allora, per ogni $1 \leq k \leq m$, se poniamo, per ogni $1 \leq i, j \leq n_k$, $i' = \bar{n}_{k-1} + i$ e $j' = \bar{n}_{k-1} + j$, si ha $B_{i'j'} := A_{ij}^{(k)}$ e $B_{ij} = 0$ per ogni altra coppia di i, j .

Per esempio,

$$B = \begin{pmatrix} 2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 & 2 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & -1 & \sqrt{2} \end{pmatrix} = \text{block}(A^{(1)}, A^{(2)}, A^{(3)}) \quad \text{con} \quad \begin{cases} A^{(1)} = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix} \\ A^{(2)} = 4 \\ A^{(3)} = \begin{pmatrix} 3 & 0 & 2 \\ 1 & 0 & 0 \\ 3 & -1 & \sqrt{2} \end{pmatrix} \end{cases}$$

Si osservi che se V_i sono gli ‘spazi coordinati’ definiti da

$$V_1 = \mathbb{C}^2 \times \{0\} \times \{0\} \times \{0\} \times \{0\}, \quad V_2 = \{0\} \times \{0\} \times \mathbb{C} \times \{0\} \times \{0\} \times \{0\}, \quad V_3 = \{0\} \times \{0\} \times \{0\} \times \mathbb{C}^3,$$

allora $A : V_i \rightarrow V_i$, ossia, gli spazi V_i sono **sottospazi invarianti per A** . Più in generale, se B è una matrice a blocchi come in (xl), allora $K^{\bar{n}} = K^{n_1} \oplus \dots \oplus K^{n_m}$ e $A : V_i \mapsto V_i$ con

$$V_i := \underbrace{\{0\} \times \dots \times \{0\}}_{(n_1 + \dots + n_{i-1} \text{ volte})} \times K^{n_i} \times \underbrace{\{0\} \times \dots \times \{0\}}_{(n_{i+1} + \dots + n_m \text{ volte})}.$$

Su tali sottospazi invarianti la matrice A si ‘riduce’ alla matrice $A^{(i)}$. Questo ‘processo di riduzione’ è assai importante da un punto di vista teorico e pratico. A tal proposito si ha la seguente caratterizzazione:

(xli) (**Lemma di riduzione a blocchi**) Sia $A \in K^{n \times n}$ con $n = \sum_{j=1}^m n_j$, ($n_j \in \mathbb{N}$). Esiste

una matrice $U \in GL(n, K)$ e m matrici $A^{(j)} \in K^{n_j \times n_j}$ tali che

$$U^{-1}AU = \text{block}(A^{(1)}, \dots, A^{(m)}), \quad (\text{D.83})$$

se e solo se $K^n = V_1 \oplus \dots \oplus V_m$ con $AV_i \subseteq V_i$ e $U = [u_1^{(1)}, \dots, u_{n_1}^{(1)}, \dots, u_1^{(m)}, \dots, u_{n_m}^{(m)}]$ con $\{u_j^{(i)} \mid 1 \leq j \leq n_i\}$ base di V_i .

Dimostrazione Facciamo la dimostrazione nel caso $m = 2$; il caso generale presenta differenze solo formali. Poniamo, dunque,

$$m = 2, \quad n_1 = p, \quad n_2 = q, \quad C := A^{(1)} := \begin{bmatrix} c^{(1)} \\ \vdots \\ c^{(p)} \end{bmatrix}, \quad D := A^{(2)} := \begin{bmatrix} d^{(1)} \\ \vdots \\ d^{(q)} \end{bmatrix}, \quad B = \text{block}(C, D).$$

con $c^{(j)} \in K^p$ e $d^{(j)} \in K^q$. Allora,

$$\begin{aligned} \exists U \in GL(n, K) \mid U^{-1}AU = B &\iff \exists U \in GL(n, K) \mid AU = UB \\ &\iff \begin{cases} U = [u^{(1)}, \dots, u^{(p)}, v^{(1)}, \dots, v^{(q)}] \\ \langle u^{(1)}, \dots, u^{(p)}, v^{(1)}, \dots, v^{(q)} \rangle = K^n \end{cases} \\ &\& \quad [Au^{(1)}, \dots, Au^{(p)}, Av^{(1)}, \dots, Av^{(q)}] = U \begin{pmatrix} C & 0 \\ 0 & D \end{pmatrix}. \end{aligned}$$

Esplicitando l'ultima uguaglianza si ha che

$$Au^{(k)} = \sum_{j=1}^p c_j^{(k)} u^{(j)} \in \langle u^{(1)}, \dots, u^{(p)} \rangle = V_1, \quad Av^{(k)} = \sum_{j=1}^q d_j^{(k)} v^{(j)} \in \langle v^{(1)}, \dots, v^{(q)} \rangle = V_2,$$

che equivale a dire che gli spazi V_i sono invarianti per A . ■

D.7 Diagonalizzazione

(xlii) Un caso particolarmente interessante del Lemma di riduzione a blocchi è quando $m = n$ e $n_j = 1$ per ogni j . In tal caso, B è una **matrice diagonale**, ossia

$$B = \text{diag}(\lambda_1, \dots, \lambda_n) \iff B_{ij} = \lambda_i \delta_{ij}, \quad (\text{D.84})$$

e la matrice A , agisce lungo gli 'assi invarianti' $V_i = \{tu^{(i)} \mid t \in K\}$ associati ad una base $\{u^{(1)}, \dots, u^{(n)}\}$ di K^n , semplicemente come moltiplicazione per lo scalare λ_i .

Una matrice $A \in K^{n \times n}$ si dice **diagonalizzabile** (sul campo K) se è simile ad una matrice diagonale.

Da (xli) segue immediatamente:

(xliii) $A \in K^{n \times n}$ è diagonalizzabile (su K) $\iff K^n$ ha una base di autovettori di $A \iff a(\lambda) = g(\lambda)$ per ogni $\lambda \in \sigma(A)$.

Da (xviii) e (xliii) segue:

(xliv) Se una matrice di ordine n ha n autovalori distinti, allora è diagonalizzabile.

Naturalmente, non è vero il viceversa.

Esempi (i) Sia $A = \begin{pmatrix} 2 & 0 & 3 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$, considerata come mappa lineare su \mathbb{R}^3 ($K = \mathbb{R}$). Allora,

$p_{\text{car}}(\lambda) = (1 - \lambda)^2(2 - \lambda)$ e quindi $\sigma(A) = \{1, 2\}$ (e le dimensioni algebriche sono $a(1) = 2$ e $a(2) = 1$). Cerchiamo gli autovettori:

$$Au = u \iff u_1 = -3u_3, \quad Au = 2u \iff u_2 = 0 = u_3$$

e quindi, $u^{(1)} := (0, 1, 0)$ e $u^{(2)} := (-3, 0, 1)$ sono due autovettori indipendenti relativi all'autovalore 1 e $u^{(3)} := (1, 0, 0)$ è un autovettore relativo all'autovalore 2 (le dimensioni geometriche sono $g(1) = 2$ e $g(2) = 1$). Da (xli) segue che

$$U^{-1}AU = \text{diag}(1, 1, 2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 2 \end{pmatrix}, \quad \text{con } U := [u^{(1)}, u^{(2)}, u^{(3)}] = \begin{pmatrix} 0 & -3 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix};$$

e quindi abbiamo diagonalizzato A su \mathbb{R} .

(ii) Sia $A = \begin{pmatrix} 1 & -2 \\ 1 & -1 \end{pmatrix}$. Allora, $p_{\text{car}}(\lambda) = \lambda^2 + 1$ e A non ha autovalori reali e dunque A non è diagonalizzabile su \mathbb{R} . D'altra parte, A ha due autovalori complessi distinti $\lambda_1 = i$ e $\lambda_2 = -i$ ed è dunque diagonalizzabile su \mathbb{C} . Infatti, $u^{(1)} := (1 + i, 1)$ è un autovettore relativo a $\lambda_1 = i$ e $u^{(2)} := (1 - i, 1)$ è un autovettore relativo a $\lambda_2 = -i$. Dunque,

$$U^{-1}AU = \text{diag}(i, -i) = \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix}, \quad \text{con } U := [u^{(1)}, u^{(2)}] = \begin{pmatrix} 1+i & 1-i \\ 1 & 1 \end{pmatrix}.$$

(iii) Se $g(\lambda) < a(\lambda)$ per un autovalore $\lambda \in \sigma(A)$, allora A non può essere diagonalizzata: l'esempio più semplice è $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$. Gli autovalori sono $\lambda_1 = \lambda_2 = 0$ (molteplicità algebrica

2), ma l'autospazio corrispondente $V_0 = \{(t, 0) \mid t \in K\}$ (K può essere sia \mathbb{R} che \mathbb{C}) e quindi $\dim V_0 = 1$ e A non può essere diagonalizzata (né in \mathbb{R} , né in \mathbb{C}). Si noti che $A^2 = 0$.

Una classe importante di matrici diagonalizzabili su \mathbb{C} o su \mathbb{R} sono, rispettivamente, le matrici autoaggiunte o le matrici reali simmetriche. Cominciamo con alcuni risultati preliminari.

(xlv) Se $A \in \mathbb{C}^{n \times n}$ è autoaggiunta, allora $\sigma_{\mathbb{C}}(A) \subseteq \mathbb{R}$.

Dimostrazione Sia $\lambda \in \sigma_{\mathbb{C}}(A)$ e $Au = \lambda u$ con $u \in \mathbb{C}^n \setminus \{0\}$. Allora,

$$\lambda |u|^2 = \lambda \langle u, u \rangle = \langle \lambda u, u \rangle = \langle Au, u \rangle = \langle u, Au \rangle = \langle u, \lambda u \rangle = \bar{\lambda} \langle u, u \rangle = \bar{\lambda} |u|^2,$$

il che implica $\lambda = \bar{\lambda}$, ossia, $\lambda \in \mathbb{R}$. ■

Osservazione D.19 il risultato è ancora vero nel caso speciale in cui A ha elementi in \mathbb{R} , nel qual caso l'ipotesi $A^* = A$ si riduce a $A^T = A$ (A è una matrice reale simmetrica). In generale, un autovettore corrispondente a λ è un elemento di⁴⁵ \mathbb{C}^n . Cionondimeno, possiamo sempre trovare un autovettore reale corrispondente all'autovalore λ . Infatti, se $Au = \lambda u$ con $A \in \mathbb{R}^{n \times n}$, $\lambda \in \mathbb{R}$ e $u \in \mathbb{C}^n$, allora si ha anche⁴⁶ $A\bar{u} = \lambda \bar{u}$. Per cui, $Av = \lambda v$ con $v := u + \bar{u} \in \mathbb{R}^n$.

(xlvi) Se $A \in \mathbb{C}^{n \times n}$ è autoaggiunta, $Au = \lambda u$ e $Aw = \mu w$ con $\lambda \neq \mu$, allora $\langle u, w \rangle = 0$.

Dimostrazione Essendo, per (xlv), $\mu \in \mathbb{R}$, si ha

$$\lambda \langle u, w \rangle = \langle \lambda u, w \rangle = \langle Au, w \rangle = \langle u, Aw \rangle = \langle u, \mu w \rangle = \mu \langle u, w \rangle,$$

per cui, essendo $\lambda \neq \mu$, si deve avere $\langle u, w \rangle = 0$. ■

È immediato verificare che, nel caso in cui A abbia elementi reali, vale l'affermazione analoga, ossia:

Se $A \in \mathbb{R}^{n \times n}$ è simmetrica, $Au = \lambda u$ e $Aw = \mu w$ con $u, w \in \mathbb{R}^n$ e $\lambda \neq \mu$, allora $u \cdot w = 0$.

(xlvii) (Teorema spettrale in \mathbb{R}^n) Se $A \in \mathbb{R}^{n \times n}$ è una matrice simmetrica, allora $\sigma(A) \subseteq \mathbb{R}$ e A è diagonalizzabile tramite una matrice ortogonale⁴⁷ U , ossia:

$$\exists U \in O(n) \quad \text{tale che} \quad U^T A U = \text{diag}(\alpha_1, \dots, \alpha_n) \quad \text{con} \quad \alpha_i \in \mathbb{R}. \quad (\text{D.85})$$

Dimostrazione Da (xlv) sappiamo che $\sigma_{\mathbb{C}}(A) \subseteq \mathbb{R}$ e da (xlvi) che autovettori reali corrispondenti ad autovalori diversi sono ortogonali in \mathbb{R}^n . Sia $\sigma_{\mathbb{C}}(A) = \{\lambda_1, \dots, \lambda_m\}$ con $1 \leq m \leq n$ e $\lambda_1 < \dots < \lambda_m$ (se $m > 1$). Siano V_i gli autospazi reali di λ_i , ossia, $V_i := \{u \in \mathbb{R}^n \mid Au = \lambda_i u\}$. Per l'Osservazione D.19, si ha che $\dim V_i \geq 1$ per ogni i . Per ogni i , fissiamo una base ortonormale di V_i (che esiste sempre per il Lemma di Gram-Schmidt (viii)) e sia $\{u_1, \dots, u_p\}$ l'unione (ordinata) di tali basi. Sia $V := V_1 + \dots + V_m$ e sia $W = V^\perp \cap \mathbb{R}^n$ l'ortogonale reale di V . Se $W = \{0\}$, il teorema è dimostrato, poiché, in tal caso, $V_1 \oplus \dots \oplus V_m = \mathbb{R}^n$, $p = n$ e, se $a_i = a(\lambda_i)$ denota la molteplicità algebrica (su \mathbb{R}) di λ_i , $\alpha_1 = \dots = \alpha_{a_1} = \lambda_1$, $\alpha_{a_1+1} = \dots = \alpha_{a_1+a_2} = \lambda_2$, etc. Chiaramente, in tal caso, si ha che $a_1 + \dots + a_m = n$ (e quindi le molteplicità algebriche degli autovalori su \mathbb{R} coincidono con quelle su \mathbb{C}^n).

Assumiamo, ora, che $\dim W \geq 1$. In tal caso, $V \oplus W = \mathbb{R}^n$ con $V \perp W$. W è un sottospazio invariante per A : sia, infatti, $w \in W$ e $u \in V$ (per cui $u \cdot w = 0$), allora esiste λ_i tale che $Au = \lambda_i u$ e quindi

$$0 = (Au - \lambda_i u) \cdot w = Au \cdot w - \lambda_i u \cdot w = Au \cdot w = u \cdot Aw,$$

e quindi $Aw \in W$. Sia $\{u_{p+1}, \dots, u_n\}$ una base ortonormale di W (di nuovo Gram-Schmidt) cosicché $\{u_1, \dots, u_n\}$ forma una base ortonormale di \mathbb{R}^n . Sia $U = [u_1, \dots, u_n]$; per (xxxviii),

⁴⁵Si faccia attenzione: qui stiamo considerando la mappa lineare $L_A : \mathbb{C}^n \rightarrow \mathbb{C}^n$ con $A \in \mathbb{R}^{n \times n}$, cosa diversa da considerare la mappa $L_A : \mathbb{R}^n \rightarrow \mathbb{R}^n$.

⁴⁶Ovviamente, $\bar{u} := (\bar{u}_1, \dots, \bar{u}_n)$.

⁴⁷Cfr. (xxxviii)

$U \in O(n)$. Per (xli) si ha che $U^T A U = \text{block}(A_1, A_2)$ con $A_1 \in \mathbb{R}^{p \times p}$ e $A_2 \in \mathbb{R}^{q \times q}$ con $q := n - p \geq 1$. Essendo $U^T A U$ una matrice simmetrica, lo sono anche A_1 e A_2 . Quindi A_2 (vista come mappa da $\mathbb{R}^q \rightarrow \mathbb{R}^q$) ha un autovalore reale λ ed un autovettore $\hat{w} \in \mathbb{R}^q \setminus \{0\}$: $A_2 \hat{w} = \lambda \hat{w}$. E se $w := (0, \hat{w}) \in \mathbb{R}^n$, Allora $\text{block}(A_1, A_2)w = (0, A_2 \hat{w}) = \lambda(0, \hat{w}) = \lambda w$, il che implica $AUw = \lambda Uw$. Ma

$$Uw = [u_1, \dots, u_p, u_{p+1}, \dots, u_n] \begin{pmatrix} 0 \\ \hat{w} \end{pmatrix} = \hat{w}_1 u_{p+1} + \dots + \hat{w}_q u_n \in W,$$

il che vuol dire che Uw è un autovettore (reale) di A (corrispondente all'autovalore λ) che appartiene a W , ma questa è una contraddizione perché tutti gli autovalori reali di A appartengono per definizione a $V = W^\perp$. Dunque questa seconda alternativa non esiste. ■

Osservazione D.20 (i) Come già osservato nel corso della dimostrazione, *le dimensioni degli autospazi reali e degli autospazi complessi di una matrice simmetrica coincidono.*

(ii) Il teorema spettrale dà una ‘rappresentazione’ particolarmente semplice delle matrici simmetriche: siano, infatti, come sopra, V_1, \dots, V_m gli autospazi distinti di $A = A^T \in \mathbb{R}^{n \times n}$ corrispondenti allo spettro $\{\lambda_1, \dots, \lambda_m\}$ di A ($\lambda_1 < \dots < \lambda_m$, se $m > 1$) e denotiamo con π_i la proiezione ortogonale su V_i (cfr. (x)). Allora, A è una *combinazione lineare di proiezioni ortogonali*:

$$A = \sum_{i=1}^m \lambda_i \pi_i, \quad (\text{D.86})$$

infatti, ogni vettore $u \in \mathbb{R}^n$ si decompone in modo unico come somma di vettori negli autospazi

V_i , ossia, $u = \sum_{i=1}^m \pi_i u$, e

$$Au = A \left(\sum_{i=1}^m \pi_i u \right) = \sum_{i=1}^m A(\pi_i u) = \sum_{i=1}^m \lambda_i \pi_i u, \quad \forall u \in \mathbb{R}^n. \quad (\text{D.87})$$

In altre parole, A ‘agisce separatamente su ogni proiezione $\pi_i u$ come una moltiplicazione per il numero reale λ_i ’.

(iii) Nel caso complesso, si ha un enunciato del teorema spettrale del tutto analogo a quello reale:

Teorema spettrale in \mathbb{C}^n *Se $A \in \mathbb{C}^{n \times n}$ è una matrice autoaggiunta, allora $\sigma(A) \subseteq \mathbb{R}$ e A è diagonalizzabile tramite una matrice unitaria U , ossia:*

$$\exists U \in U(n) \quad \text{tale che} \quad U^* A U = \text{diag}(\lambda_1, \dots, \lambda_n) \quad \text{con } \lambda_i \in \mathbb{R}. \quad (\text{D.88})$$

La dimostrazione è simile (infatti, più semplice) alla dimostrazione del caso reale⁴⁸.

Esempio Sia

$$A = \begin{pmatrix} 4 & -1 & -1 \\ -1 & 4 & 1 \\ -1 & 1 & 4 \end{pmatrix}. \quad (\text{D.89})$$

A è simmetrica e il suo polinomio caratteristico è dato da $p_{\text{car}}(z) = -z^3 + 12z^2 - 45z + 54 = -(z-3)^2(z-6)$, quindi $\sigma(A) = \{3, 6\}$. Si verifica facilmente che $w_1 = (1, 1, 0)$ e $w_2 = (1, 0, 1)$ sono autovettori indipendenti dell'autovalore 3 e che $w_3 = (-1, 1, 1)$ è un autovettore di 6. Tali vettori sono perpendicolari tra loro. Normalizzando, otteniamo

$$u_1 := \frac{w_1}{|w_1|} = \frac{1}{\sqrt{2}}(1, 1, 0), \quad u_2 := \frac{w_2}{|w_2|} = \frac{1}{\sqrt{2}}(1, 0, 1) \quad u_3 := \frac{w_3}{|w_3|} = \frac{1}{\sqrt{3}}(-1, 1, 1).$$

⁴⁸Sostituendo sistematicamente \mathbb{R} con \mathbb{C} e il prodotto scalare col prodotto hermitiano; naturalmente i V_i saranno definiti come gli autospazi complessi in \mathbb{C}^n di A corrispondenti agli autovalori (sempre reali).

Quindi, se $U := [u_1, u_2, u_3]$ e $\Lambda = \text{diag}(3, 3, 6)$, si ha che $U^T A U = \Lambda$.

Una semplice conseguenza del teorema spettrale è la seguente caratterizzazione della norma di una matrice autoaggiunta o simmetrica:

(xlviii) Se $A \in \mathbb{C}^{n \times n}$ è autoaggiunta allora

$$\|A\| := \sup_{x \in \mathbb{C}^n: |x|=1} |Ax| = \sup_{x \in \mathbb{C}^n: |x|=1} |\langle Ax, x \rangle| = \max\{|\lambda| : \lambda \in \sigma(A)\}, \quad (\text{D.90})$$

dove $|x| := \sqrt{\langle x, x \rangle} = \sqrt{\sum_{j=1}^n |x_j|^2}$. Analogamente, se $A \in \mathbb{R}^{n \times n}$ è simmetrica allora

$$\|A\| := \sup_{x \in \mathbb{R}^n: |x|=1} |Ax| = \sup_{x \in \mathbb{R}^n: |x|=1} |Ax \cdot x| = \max\{|\lambda| : \lambda \in \sigma(A)\}, \quad (\text{D.91})$$

dove $|x| := \sqrt{x \cdot x} = \sqrt{\sum_{j=1}^n |x_j|^2}$.

Dimostrazione Diamo la dimostrazione nel caso complesso, essendo quella nel caso reale completamente analoga. Dal teorema spettrale segue che $A = U^* \Lambda U$ con $U \in U(n)$ e $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$ diagonale e reale. Allora, poiché le matrici unitarie conservano il prodotto hermitiano e la norma euclidea,

$$\begin{aligned} \|A\|^2 &= \sup_{|x|=1} \langle Ax, Ax \rangle = \sup_{|x|=1} \langle U \Lambda U^* x, U \Lambda U^* x \rangle = \sup_{|x|=1} \langle \Lambda U^* x, \Lambda U^* x \rangle = \sup_{|Uy|=1} \langle \Lambda y, \Lambda y \rangle \\ &= \sup_{|y|=1} \langle \Lambda y, \Lambda y \rangle = \sup_{|y|=1} \sum_{j=1}^n \lambda_j^2 |y_j|^2 = \max_j \lambda_j^2. \end{aligned}$$

Allo stesso modo,

$$\begin{aligned} \sup_{|x|=1} |\langle Ax, x \rangle| &= \sup_{|x|=1} |\langle U \Lambda U^* x, x \rangle| = \sup_{|x|=1} \langle \Lambda U^* x, U^* x \rangle = \sup_{|Uy|=1} |\langle \Lambda y, y \rangle| = \sup_{|y|=1} |\langle \Lambda y, y \rangle| \\ &= \sup_{|y|=1} \sum_{j=1}^n |\lambda_j| |y_j|^2 = \max_j |\lambda_j|. \quad \blacksquare \end{aligned}$$

D.8 Matrici definite positive/negative

Una matrice simmetrica $A \in \mathbb{R}^{n \times n}$ si dice **semi definita positiva**, e si scriverà $A \geq 0$, se

$$Ax \cdot x \geq 0, \quad \forall x \in \mathbb{R}^n; \quad (\text{D.92})$$

si dice **definita positiva**, e si scriverà $A > 0$, se in (D.92) vale la disuguaglianza stretta per ogni $x \neq 0$; A si dice **(semi) definita negativa** se $-A$ è (semi) definita positiva (e si scriverà, rispettivamente $A \leq 0$ o $A < 0$).

Analoghe definizioni si danno nel caso complesso:

Una matrice autoaggiunta $A \in \mathbb{C}^{n \times n}$ si dice **semi definita positiva** se

$$\langle Ax, x \rangle \geq 0, \quad \forall x \in \mathbb{C}^n; \quad (\text{D.93})$$

si dice **definita positiva** se in (D.93) vale la disuguaglianza stretta per ogni $x \neq 0$; A si dice **(semi) definita negativa** se $-A$ è (semi) definita positiva.

(xlix) Sia $A \in K^{n \times n}$ una matrice simmetrica se $K = \mathbb{R}$ o autoaggiunta se $K = \mathbb{C}$. Allora, A è semidefinita positiva se e solo se $\sigma(A) \subseteq [0, +\infty)$.

Dimostrazione Facciamo la dimostrazione nel caso reale, il caso complesso essendo del tutto analogo. Se A è semidefinita positiva e se $\lambda \in \sigma(A)$ con $Au = \lambda u$ con $u \in \mathbb{R}^n \setminus \{0\}$, allora da (D.92), segue che $0 \leq Au \cdot u = \lambda |u|^2$, e quindi $\lambda \geq 0$. Viceversa, se $\sigma(A) \subseteq [0, +\infty)$, dal

Teorema spettrale segue che $U^T A U = \text{diag}(\lambda_1, \dots, \lambda_n) =: \Lambda$ con $\lambda_i \geq 0$ e quindi se $x \in \mathbb{R}^n$ e $y := U^T x$ si ha

$$Ax \cdot x = U \Lambda U^T x \cdot x = \Lambda U^T x \cdot U^T x = \Lambda y \cdot y = \sum_{i=1}^n \lambda_i y_i^2 \geq 0. \quad \blacksquare$$

Analoghe affermazioni (con le ovvie modifiche) si hanno negli altri casi.

Dunque se A è una matrice reale definita positiva, la forma quadratica $(x, y) \mapsto Ax \cdot y$ definisce un prodotto scalare su \mathbb{R}^n , ossia, una forma bilineare, simmetrica e tale che $Ax \cdot x > 0$ per ogni $x \neq 0$. L'immagine di A della sfera euclidea unitaria di raggio 1, ossia l'insieme $\{Ax \mid |x| \leq 1\}$ è un *ellissoide* di semiassi principali $0 < \lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$ dove λ_i sono gli autovalori (contati con molteplicità) di A .

Osservazione D.21 (i) La matrice $A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}$ è simmetrica e $Ax \cdot x = (x_1 + x_2)^2 \geq 0$ e quindi è semi-definita positiva ($Ax \cdot x = 0$ sulla retta $\{x \in \mathbb{R}^2 \mid x_1 + x_2 = 0\}$ e gli autovalori sono 0 e 2).

(ii) La matrice A in (D.89) è simmetrica reale e, come abbiamo visto, ha autovalori 3 e 6 e dunque è una matrice definita positiva. Infatti,

$$Ax \cdot x = 2|x|^2 + (x_1 - x_3)^2 + (x_2 - x_1)^2 + (x_2 + x_3)^2 > 0, \quad \forall x \in \mathbb{R}^3 \setminus \{0\}.$$

(iii) Se $A \in \mathbb{C}^{n \times n}$ non è difficile verificare che la condizione $\langle Au, u \rangle \geq 0$ per ogni $u \in \mathbb{C}^n$ implica che A sia *autoaggiunta*⁴⁹ e quindi nelle definizioni date sopra per matrici complesse la condizione che A sia autoaggiunta può essere omessa. Questo *non è vero nel caso reale*.

Per esempio, se $A = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix}$, allora, per ogni $x \in \mathbb{R}^2 \setminus \{0\}$ si ha $Ax \cdot x = x_1^2 + x_2^2 > 0$, ma $\sigma(A) = \{1 \pm i\}$ e, se $u = (1, i)$, $\langle Au, u \rangle = 2(1 + i) \notin \mathbb{R}$ (e quindi tale matrice A , non è una matrice definita positiva vista come matrice complessa e senza la condizione di simmetria si avrebbe una inconsistenza tra la definizione per matrici reali e quella per matrici complesse).

D.9 Forma normale (o canonica) di Jordan

In questa sezione finale discutiamo un risultato fondamentale – dovuto al matematico francese Camille Jordan⁵⁰ – sulle trasformazioni lineari di spazi finito-dimensionali (su \mathbb{C}) che generalizza il teorema spettrale e permette, in particolare, di trasformare, tramite un opportuno cambio di base, una qualunque matrice quadrata complessa in una forma particolarmente semplice.

Un **blocco di Jordan** di ordine $n \in \mathbb{N}$ e autovalore $\lambda \in \mathbb{C}$ è una matrice quadrata $J := J_n(\lambda) \in \mathbb{C}^{n \times n}$ con tutti λ sulla diagonale principale e tutti 1 sulla sopradiagonale principale, ossia, se $n \geq 2$,

$$J_n(\lambda) := \begin{pmatrix} \lambda & 1 & 0 & 0 & \dots \\ 0 & \lambda & 1 & 0 & \dots \\ \vdots & & \ddots & \ddots & \\ & & & \lambda & 1 \\ \dots & & & 0 & \lambda \end{pmatrix} \iff (J_n(\lambda))_{ij} = \begin{cases} 1, & \text{se } j = i + 1 \\ 0, & \text{altrimenti} \end{cases}, \quad (\text{D.94})$$

⁴⁹Infatti, sia $A = U + iV$ con $U := (A + A^*)/2$ e $V := (A - A^*)/2$. Le matrici U e V sono autoaggiunte e, per ogni $x \in \mathbb{C}^n$, $\langle Ax, x \rangle = \langle Ux, x \rangle + i\langle Vx, x \rangle =: \alpha + i\beta$. I numeri α e β sono reali (per esempio, $\alpha = \langle Ux, x \rangle = \overline{\langle x, Ux \rangle} = \overline{\langle Ux, x \rangle} = \alpha$) e quindi dovendo essere $\langle Ax, x \rangle$ reale si deve avere $\beta = 0$ ossia $\langle Vx, x \rangle = 0$ per ogni $x \in \mathbb{C}^n$. Ma questo, per (xlix), significa che $\|V\| = 0$, ossia che $V = 0$, il che implica che $A = U$, ossia che A è autoaggiunta.

⁵⁰https://it.wikipedia.org/wiki/Camille_Jordan

mentre, nel caso $n = 1$, si pone $J_1(\lambda) := \lambda$. Si noti che⁵¹

$$J_n(\lambda) = \lambda I_n + J_n(0) = \lambda I_n + [0, e^{(1)}, \dots, e^{(n-1)}]. \quad (\text{D.95})$$

(I) Teorema della forma normale di Jordan Sia $A \in \mathbb{C}^{n \times n}$. Esiste una matrice $U \in GL(n, \mathbb{C})$ tale che

$$U^{-1}AU = \text{block}(J_1, \dots, J_k), \quad (\text{D.96})$$

dove J_i sono blocchi di Jordan. La ‘forma normale di Jordan’ $\text{block}(J_1, \dots, J_k)$ è unica a meno dell’ordine dei blocchi.

La dimostrazione (cfr. (liv)) di questo risultato fondamentale si basa su due risultati, interessanti di per sé, sugli endomorfismi di spazi vettoriali complessi: un lemma di ‘splitting’ di V in somma diretta di spazi invarianti per T (‘autospazi generalizzati’) su cui $(T - \lambda_i I)$ è nilpotente e sull’esistenza di opportune basi formate da catene di Jordan per matrici nilpotenti⁵².

(ii) Splitting spettrale Cominciamo con alcune definizioni preliminari. Sia V uno spazio vettoriale complesso di dimensione n , sia $T \in \mathcal{L}(V)$ e sia $\sigma(T) = \{\lambda_1, \dots, \lambda_m\}$ il suo spettro ($\lambda_i \in \mathbb{C}$ e $\lambda_i \neq \lambda_j$ se $i \neq j$). Poniamo⁵³, per ogni $1 \leq i \leq m$, $N_i := T - \lambda_i I$ e consideriamo gli spazi vettoriali $\ker N_i^p$ con $p \geq 1$. Chiaramente, $\ker N_i^p \subseteq \ker N_i^{p+1}$ per ogni p e se u è un autovettore di T con autovalore⁵⁴ λ_i , allora $N_i u = 0$ e quindi $u \in \ker N_i$ e dunque $\dim N_i^p \geq 1$ per ogni $p \geq 1$.

D’altra parte se $\ker N_i^p \subsetneq \ker N_i^{p+1}$ allora $\dim \ker N_i^p < \dim \ker N_i^{p+1}$ e quindi, poiché, per ogni p , $\ker N_i^p \subseteq V$ (che ha dimensione finita n) deve esistere un primo numero naturale p tale che $\ker N_i^p = \ker N_i^{p+1}$. Chiamiamo **indice di nilpotenza relativo all’autovalore λ_i** il numero naturale $\nu_i := \nu(\lambda_i) := \min\{p \in \mathbb{N} \mid \ker(T - \lambda_i I)^p = \ker(T - \lambda_i I)^{p+1}\}$:

$$\begin{cases} \ker N_i = \ker N_i^2, & \text{se } \nu_i = 1, \\ \ker N_i \subsetneq \dots \subsetneq \ker N_i^{\nu_i} = \ker N_i^{\nu_i+1}, & \text{se } 1 < \nu_i \leq n. \end{cases} \quad (\text{D.97})$$

Dimostriamo, ora, che, a partire da $p = \nu_i$ in poi, gli spazi $\ker N_i^p$ si ‘stabilizzano’:

$$\ker N_i^{\nu_i+p} = \ker N_i^{\nu_i}, \quad \forall p \geq 0. \quad (\text{D.98})$$

Dimostrazione di (D.98) Per $p = 0$ (D.98) è una tautologia e per $p = 1$ è vera per definizione di ν_i . Assumiamo (D.98) vera per $p \geq 1$ e dimostriamola per $p + 1$. Se (tralasciando per semplicità di notazione l’indice i) $\ker N^{\nu+p} = \ker N^{\nu+p+1}$ la tesi è vera. Se $\ker N^{\nu+p} \subsetneq \ker N^{\nu+p+1}$ allora esiste u tale che $N^{\nu+p+1}u = 0$ e $N^{\nu+p}u \neq 0$. Sia $v = Nu$. Allora, $0 = N^{\nu+p}v \implies v \in \ker N^{\nu+p} = \ker N^{\nu} \implies N^{\nu}v = 0 \implies N^{\nu+p-1}v = N^{\nu+p}u = 0$, ma questa è una contraddizione, quindi solo la prima alternativa è vera. ■

Da (D.98) segue che

$$V = \ker N_i^{\nu_i} \oplus \text{im } N_i^{\nu_i}, \quad \forall 1 \leq i \leq m. \quad (\text{D.99})$$

Infatti, $v \in \ker N_i^{\nu_i} \cap \text{im } N_i^{\nu_i}$ se e solo se $N_i^{\nu_i}v = 0$ ed esiste u tale che $N_i^{\nu_i}u = v$ ed allora $v = N_i^{\nu_i}u \stackrel{(\text{D.98})}{=} N_i^{2\nu_i}u = N_i^{\nu_i}v = 0$. Quindi $\ker N_i^{\nu_i} \cap \text{im } N_i^{\nu_i} = \{0\}$, e (D.99) segue dal Teorema nullità–rango.

Lo spazio vettoriale $K_i := \ker N_i^{\nu_i} \subseteq V$ prende il nome di **autospazio generalizzato** relativo all’autovalore λ_i e i suoi elementi si chiamano **autovettori generalizzati**. Si noti che, dalla definizione di ν_i , segue che esiste $v \in K_i$ tale che $u := N_i^{\nu_i-1}v \neq 0$ (essendo $\ker N_i^{\nu_i-1} \subsetneq \ker N_i^{\nu_i}$) e poiché $N_i u = N_i^{\nu_i}v = 0$, tale u è un autovettore con autovalore λ_i e quindi ogni **autospazio generalizzato contiene almeno un autovettore di T** . In particolare $\dim K_i \geq g(\lambda_i) \geq$

⁵¹ $J_1(0) = 0$.

⁵² Per la definizione di applicazione nilpotente e di catena di Jordan, si veda (xvi).

⁵³ I denota, come al solito, la mappa identità $I : v \mapsto Iv = v$.

⁵⁴ L’ipotesi che lo spazio vettoriale sia complesso è essenziale e garantisce, per esempio, che $\ker N_i \neq \{0\}$.

1. Si noti, anche, che, poiché T commuta con $N_i^{\nu_i}$, l'autospazio generalizzato K_i è invariante per T , ossia $TK_i \subseteq K_i$. Inoltre, la restrizione di N_i su K_i è nilpotente con indice di nilpotenza⁵⁵ ν_i .

Lemma D.22 (Splitting spettrale) Sia V uno spazio vettoriale complesso⁵⁶ di dimensione n , sia $T \in \mathcal{L}(V)$, sia $\sigma(T) = \{\lambda_1, \dots, \lambda_m\}$ il suo spettro ($\lambda_i \in \mathbb{C}$ e $\lambda_i \neq \lambda_j$ se $i \neq j$). Allora,

$$V = K_1 \oplus \dots \oplus K_m, \quad (\text{D.100})$$

dove K_i è l'autospazio generalizzato relativo a λ_i .

Dimostrazione (Induzione finita su m) Se $\sigma(T) = \{\lambda_1\}$, da (D.99) si ha che $V = \ker N_1^{\nu_1} \oplus \text{im } N_1^{\nu_1}$ e se fosse $\text{im } N_1^{\nu_1} \neq \{0\}$, $T|_{\text{im } N_1^{\nu_1}}$ avrebbe un autovalore che necessariamente (per somma diretta) dovrebbe essere diverso da λ_1 e si avrebbe una contraddizione: dunque se $m = 1$, $V = K_1$ e (D.100) è vera. Assumiamo, ora, il risultato vero per $1 \leq m-1 \leq n-1$ e dimostriamolo per m . Di nuovo per (D.99) con $i = m$, si ha che

$$V = \ker N_m^{\nu_m} \oplus \text{im } N_m^{\nu_m}. \quad (\text{D.101})$$

Se $W := \text{im } N_m^{\nu_m} = \{0\}$, si avrebbe che $V = \ker N_m^{\nu_m}$, ma questo implicherebbe che T non ha altri autovalori eccetto λ_m ottenendo una contraddizione (stiamo assumendo che $m \geq 2$): infatti, se esistesse $\lambda_1 \neq \lambda_m$ e $u \in V \setminus \{0\}$ tale che $(T - \lambda_1 I)u = N_1 u = 0$, allora si avrebbe $0 = N_m^{\nu_m} u = (\lambda_1 - \lambda_m)^{\nu_m} u$, che è una contraddizione. Dunque, $\dim W \geq 1$. Poiché T e N_m commutano, W è uno spazio invariante per T . Poniamo $\bar{T} := T|_W : u \in W \mapsto Tu \in W$. Poiché $\dim \ker N_m^{\nu_m} \geq g(\lambda_m) \geq 1$, si ha che $1 \leq \dim W \leq n-1$. Poiché, come visto sopra, $T|_{K_m}$ (T ristretto a K_m) ha il solo autovalore λ_m , si deve avere $\sigma(\bar{T}) = \{\lambda_1, \dots, \lambda_{m-1}\}$ e

$$\bar{\nu}_i := \min\{p \in \mathbb{N} \mid \ker(\bar{T} - \lambda_i I)^p = \ker(\bar{T} - \lambda_i I)^{p+1}\} = \nu_i, \quad \forall 1 \leq i \leq m-1.$$

Dunque, per l'ipotesi induttiva,

$$W = \ker(\bar{T} - \lambda_1 I)^{\bar{\nu}_1} \oplus \dots \oplus \ker(\bar{T} - \lambda_{m-1} I)^{\bar{\nu}_{m-1}} = \ker(T - \lambda_1 I)^{\nu_1} \oplus \dots \oplus \ker(T - \lambda_{m-1} I)^{\nu_{m-1}},$$

ed essendo, per (D.101), $V = W \oplus \ker(T - \lambda_m I)^{\nu_m}$ si ha la tesi. ■

Osservazione D.23 (Polinomio minimo) Siano V , T , λ_i e ν_i come nel Lemma di *splitting* D.22. Si definisce **polinomio minimo di T** il polinomio

$$p_{\min}(z) := (z - \lambda_1)^{\nu_1} \dots (z - \lambda_m)^{\nu_m}, \quad (z \in \mathbb{C}). \quad (\text{D.102})$$

Tale polinomio è il *polinomio monico di grado minimo che annulla T* , ossia tale che⁵⁷ $p_{\min}(T) = 0$. Infatti, dal Lemma di *splitting* segue che $p_{\min}(T)u = 0$, per ogni $u \in V$. Assumiamo ora che $p(z)$ sia un altro polinomio monico che annulla T . Se u_i è un autovettore con autovalore λ_i , allora $0 = p(T)u = p(\lambda_i)u$, che implica che $p(\lambda_i) = 0$, ossia che λ_i è una radice di p . Dunque $p(z) = q(z)(z - \lambda_i)^r$ per un qualche $r \in \mathbb{N}$ e q polinomio monico tale che $q(\lambda_i) \neq 0$. Supponiamo (per assurdo) che $r < \nu_i$. Sia $u \in K_i$ un vettore tale che $N_i^{\nu_i-1}u \neq 0$ e poniamo $v := N_i^{\nu_i-1-r}u$. Allora si avrebbe $0 = p(T)v = q(T)(T - \lambda_i I)^r N_i^{\nu_i-1-r}u = q(T)N_i^{\nu_i-1}u = q(\lambda_i)N_i^{\nu_i-1}u \neq 0$, contraddizione. Dunque $r \geq \nu_i$ e $p(z) = \tilde{q}(z)p_{\min}(z)$ per un qualche polinomio monico \tilde{q} che non ha per radici λ_i . In conclusione o $p = p_{\min}$ oppure $\deg p > \deg p_{\min}$.

(lii) (Catene di Jordan) Sia V uno spazio vettoriale di dimensione n e $N \in \mathcal{L}(V)$ un'applicazione lineare nilpotente di indice ν . Ricordiamo ((xvi)) che una **catena di Jordan** per N è un insieme della forma $\gamma = \{u, Nu, \dots, N^{j-1}u\}$ con $1 \leq j \leq \nu$ tale che $N^{j-1}u \neq 0$ e $N^j u = 0$; la **lunghezza** (o 'ordine') della catena γ è $j := \#\gamma$. Chiameremo i vettori u e $N^{j-1}u$, rispettivamente, **base** e **vertice** della catena γ . Una catena γ si dice **massimale** se la sua base $u \notin \text{im}(N)$.

⁵⁵Cfr. (xvi).

⁵⁶L'ipotesi che V sia uno spazio vettoriale su \mathbb{C} è essenziale e serve a garantire che $\sigma(T) \neq \emptyset$ (cosa in generale non vera, per esempio, per spazi vettoriali su \mathbb{R}).

⁵⁷ $p_{\min}(T) = 0$ significa che $((T - \lambda_1 I)^{\nu_1} \dots (T - \lambda_m I)^{\nu_m})u = 0$ per ogni $u \in V$. Osserviamo che i 'monomi' $(T - \lambda_i I)^{\nu_i}$ commutano tra loro e quindi l'ordine nella definizione di p_{\min} è irrilevante.

Osservazione D.24 (i) Dalla definizione segue immediatamente che il vertice di una catena di Jordan è un elemento del nucleo di N (ossia, un autovettore con autovalore 0) e che una catena massimale di ordine 1 è formata da un vettore $u \in \ker N$ tale che $u \notin \text{im } N$.

(ii) Se $\gamma = \{u, \dots, N^{j-1}u\}$ è una catena di Jordan di ordine $j \geq 2$, allora

$$N\gamma = \{Nu, \dots, N^{j-1}u, N^j u\} = \{Nu, \dots, N^{j-1}u, 0\} = \bar{\gamma} \cup \{0\}, \quad (\text{D.103})$$

con $\bar{\gamma} := \{Nu, \dots, N^{j-1}u\}$ catena di Jordan di ordine $j - 1$ e con lo stesso vertice di γ .

(iii) *Catene di Jordan con vertici indipendenti formano insiemi indipendenti.* Più precisamente:

Siano γ_i catene di Jordan di uno spazio vettoriale V . Se i vertici delle catene sono indipendenti, allora l'insieme $\cup \gamma_i$ forma un insieme indipendente; in particolare, gli elementi di una catena di Jordan sono tra loro indipendenti.

Dimostriamo questa affermazione per induzione su $k := \# \cup \gamma_i$. Se $k = 1$ il risultato è banalmente vero. Assumiamo il risultato vero per $k \geq 1$ e dimostriamolo per $k + 1$. Siano dunque γ_i p catene di Jordan con $\# \cup \gamma_i = k + 1$ ($1 \leq p \leq k + 1$) e vertici indipendenti w_i . Sia $w = \sum a_i w_i + \tilde{w}$ una combinazione lineare di elementi di $\cup \gamma_i$ (\tilde{w} combinazione lineare di vettori in $\cup(\gamma_i \setminus \{w_i\})$ o il vettore nullo). Se $\tilde{w} = 0$, allora, essendo i w_i indipendenti si deve avere $a_i = 0$ per ogni i e abbiamo finito. Assumiamo ora $\tilde{w} \neq 0$ e osserviamo (essendo $Nw_i = 0$) che $Nw = N\tilde{w}$. Ora, \tilde{w} è una combinazione lineare di catene di Jordan (cfr. osservazione (ii) qui sopra) la cui cardinalità è $\bar{k} = k + 1 - p$ con $1 \leq \bar{k} < k$ e quindi, per ipotesi induttiva, formata da vettori indipendenti: dunque \tilde{w} deve essere la combinazione lineare nulla, ossia, $\tilde{w} = 0$. Ma allora, per l'indipendenza dei vertici w_i , anche w deve essere la combinazione lineare nulla. ■

(iv) *Ad una catena di Jordan corrisponde un blocco di Jordan con autovalore 0.* Più precisamente:

Se $N \in \mathbb{C}^{k \times k}$ e $\mathbb{C}^k = \langle \{u, Nu, \dots, N^{k-1}u\} \rangle$, con $N^{k-1}u \neq 0$, allora⁵⁸

$$U^{-1}NU = J_k(0), \quad \text{con } U := [N^{k-1}u, \dots, Nu, u]. \quad (\text{D.104})$$

Infatti, $NU = [N^k u, N^{k-1}u, \dots, Nu] = [0, N^{k-1}u, \dots, Nu]$ e

$$UJ_k(0) \stackrel{(\text{D.95})}{=} U[0, e^{(1)}, \dots, e^{(k-1)}] = [0, Ue^{(1)}, \dots, Ue^{(k-1)}] \stackrel{(\text{D.36})}{=} [0, N^{k-1}u, \dots, Nu],$$

e dunque vale (D.104). Inoltre, se $A = N + \lambda I_k$, da (D.104) segue che

$$U^{-1}AU = U^{-1}NU + \lambda I_k = J_k(0) + \lambda I_k = J_k(\lambda).$$

(v) *Lo spazio vettoriale generato da una catena di Jordan $\gamma = \{u, \dots, N^{k-1}u\}$ di N è uno spazio invariante per N .* Infatti, da (D.103) segue che, se $k = 1$, $N : \langle \gamma \rangle \rightarrow \{0\}$ e se $k > 1$, $N : \langle \gamma \rangle \rightarrow \langle \{Nu, \dots, N^{k-1}u\} \rangle$.

(liii) Lemma delle basi di catene di Jordan *Sia V uno spazio vettoriale di dimensione n e $N \in \mathcal{L}(V)$ un endomorfismo nilpotente. Allora, esiste una base $\beta = \cup_i \gamma_i$ di V formata da k catene di Jordan γ_i con $1 \leq k \leq n$ e*

$$V = \langle \gamma_1 \rangle \oplus \dots \oplus \langle \gamma_k \rangle, \quad (\gamma_i \text{ catena di Jordan}). \quad (\text{D.105})$$

Dimostrazione Se $N = 0$, il lemma è banalmente vero con $\beta = \{u_1, \dots, u_n\}$ base qualunque di V ($k = n$). Se $n = 1$, l'unico endomorfismo nilpotente è $N = 0$ e quindi il lemma è vero. Assumiamo, quindi, $N \neq 0$ e $n \geq 2$ e sia $1 \leq \nu \leq n$ l'indice di nilpotenza di N (ossia $N^\nu = 0$, $N^{\nu-1} \neq 0$). Per $0 \leq j \leq \nu - 1$, siano W_j i seguenti insiemi indipendenti: $W_{\nu-1}$ è una base

⁵⁸Si ricordi (D.95).

di $N^{\nu-1}V = N^{\nu-1}V \cap \ker N$ e (ricorsivamente per $\nu - 2 \geq j \geq 0$) W_j è un completamento di $W_{\nu-1} \cup \dots \cup W_{j+1}$ in una base di $N^jV \cap \ker N$. Poniamo⁵⁹ $\kappa_j = \#W_j$. Per costruzione⁶⁰,

$$N^jV \cap \ker N = W_j \oplus \dots \oplus W_{\nu-1}, \quad \dim(N^jV \cap \ker N) = \sum_{i=j}^{\nu-1} \kappa_i, \quad \forall 0 \leq j \leq \nu-1. \quad (\text{D.106})$$

In particolare, $W := \cup W_i$ è una base di $\ker N$. Si noti anche, se $W_j \neq \emptyset$, ad ogni elemento $w \in W_j$ corrisponde una catena di Jordan $\gamma(w)$ di ordine $j+1$: infatti, se $w \in W_j$ si ha $w = N^j u$ per un qualche $u \in V$ e $N^{j+1}u = 0$, e $\gamma(w) = \{u, Nu, \dots, N^j u\}$ con $N^j u = w$. Chiamiamo $\Gamma = \{\gamma_1, \dots, \gamma_k\}$ l'insieme di queste catene di Jordan: se $W = \{w_i\}$ dove $1 \leq i \leq k := \sum \kappa_j$, allora $\gamma_i := \gamma(w_i)$. Si noti che

$$k := \#\Gamma = \sum_{j=0}^{\nu-1} \kappa_j = \#\ker N. \quad (\text{D.107})$$

Sia, ora, $\hat{\Gamma} = \cup \gamma_i$ l'insieme dei vettori che formano tutte le catene di Jordan in Γ : dall'Osservazione D.24–(iii) segue che $\hat{\Gamma}$ è un insieme indipendente con cardinalità⁶¹

$$\#\hat{\Gamma} = \sum_{j=0}^{\nu-1} (j+1)\kappa_j. \quad (\text{D.108})$$

Osserviamo, ora, che

$$\dim N^jV = \dim N^{j+1}V + \dim(\ker N \cap N^jV), \quad \forall 0 \leq j \leq \nu-1; \quad (\text{D.109})$$

infatti, se chiamiamo $\hat{N} = N|_{N^jV}$ la restrizione di N allo spazio N^jV , si ha che $\hat{N}(N^jV) = N^{j+1}V$ e che $\ker \hat{N} = (\ker N \cap N^jV)$ e la (D.109) segue dal teorema di nullità-rango. Se chiamiamo, per⁶² $0 \leq j \leq \nu-1$, $d_j = \dim N^jV$, da (D.109) e (D.106) segue che

$$d_j = d_{j+1} + \sum_{i=j}^{\nu-1} \kappa_i, \quad \forall 0 \leq j \leq \nu-1, \quad (d_j := \dim N^jV), \quad (\text{D.110})$$

relazione che, iterata, implica

$$\dim V = d_0 = \sum_{j=0}^{\nu-1} \sum_{i=j}^{\nu-1} \kappa_i = \sum_{i=0}^{\nu-1} \sum_{j=0}^i \kappa_i = \sum_{i=0}^{\nu-1} (i+1)\kappa_i \stackrel{(\text{D.108})}{=} \#\hat{\Gamma},$$

e dunque (essendo $\hat{\Gamma}$ un insieme indipendente) $\langle \hat{\Gamma} \rangle = V$. Inoltre, essendo gli spazi $\langle \gamma_i \rangle$ invarianti per N (Osservazione D.24–(v)) si ha (D.105). ■

Osservazione D.25 (i) La dimostrazione appena fatta è *costruttiva* e contiene più informazioni dell'enunciato: in particolare, fornisce un *algoritmo che permette di trovare le catene γ_i che generano V* , algoritmo riassunto qui di seguito:

1. determinare l'indice di nilpotenza ν di N ($\nu = \min\{p \geq 1 \mid N^p = 0\}$)
2. determinare una base $W_{\nu-1}$ di $N^{\nu-1}V$
3. ricorsivamente determinare (per $\nu-2 \geq j \geq 0$) un insieme di vettori W_j che completino $W_{\nu-1} \cup \dots \cup W_{j+1}$ in una base di $N^jV \cap \ker N$; definire $\kappa_j := \#W_j$

⁵⁹ $\kappa_j = 0 \iff W_j = \emptyset$. Si noti che $\kappa_{\nu-1} \geq 1$, ma altri κ_j potrebbero essere nulli.

⁶⁰ $N^0 = I$ e quindi $N^0V \cap \ker N = \ker N$; $N^{\nu-1}V \subseteq \ker N$ e quindi $N^{\nu-1}V \cap \ker N = N^{\nu-1}V$.

⁶¹ Ci sono κ_j catene di Jordan di ordine (e quindi di cardinalità) $j+1$.

⁶² $d_0 = \dim V$, $d_{\nu+1} = 0$.

4. per ogni $\kappa_j > 0$ ($\kappa_{\nu-1} > 1$ sempre) e per ogni $w_i \in W_j$ si determini una catena di Jordan $\gamma_i = \gamma(w_i)$, determinando, in modo ricorsivo, j vettori u_i^k tali che $u_i^j := w_i$ e $Nu_i^{k-1} = u_i^k$.

(ii) Applicando il Lemma delle basi di catene di Jordan ad una matrice $N \in \mathbb{C}^{n \times n}$ nilpotente (identificata, come al solito con l'endomorfismo L_N), da (D.105), dal Lemma di riduzione a blocchi (xli) e dai punti (iv) e (v) dell'Osservazione D.24 segue

$$U^{-1}NU = \text{block}(J_{n_1}(0), \dots, J_{n_k}(0)), \quad \text{con} \quad U := \underbrace{[N^{n_1-1}u_1, \dots, u_1]}_{\gamma_1}, \dots, \underbrace{[N^{n_k-1}u_k, \dots, u_k]}_{\gamma_k}. \quad (\text{D.111})$$

(iii) Abbiamo dimostrato, in particolare, che:

Se N è una trasformazione nilpotente come in (liii) con $\gamma_i = \{u_i, \dots, N^{n_i-1}u_i\}$ per $1 \leq i \leq k$, allora⁶³: $\sigma(N) = \{0\}$, $a(0) = n$, $\dim(\ker N) = g(0) = k$, $\nu(0) = \max n_i$. Una base di $\ker N$ (ossia dell'autospazio dell'unico autovalore 0) è data dai k autovettori $w_i := N^{n_i-1}u_i$.

(iv) Le catene di Jordan che generano V sono tutte massimali: se una di esse non lo fosse, diciamo γ_i con base u_i , allora esisterebbe $u \in V$ tale che $Nu = u_i$, ma allora (Osservazione D.24–(iii)) il vettore u sarebbe indipendente da $\{\cup \gamma_i\}$ il che è impossibile poichè $\langle \cup \gamma_i \rangle = V$.

A questo punto il teorema di Jordan (I) segue facilmente:

(liv) Dimostrazione del Teorema della forma normale di Jordan (I)

Sia $\sigma(A) = \{\lambda_1, \dots, \lambda_m\}$ lo spettro di A ($\lambda_i \in \mathbb{C}$ e $\lambda_i \neq \lambda_j$ se $i \neq j$). Per il Lemma di *splitting* spettrale D.22 con $V = \mathbb{C}^n$ e $T = A$ (identificando, come al solito, la matrice A con la mappa lineare L_A), si ha che

$$\mathbb{C}^n = K_1 \oplus \dots \oplus K_m, \quad K_i := \ker(A - \lambda_i I_n)^{\nu_i}, \quad (\text{D.112})$$

dove K_i è l'autospazio generalizzato associato a λ_i e ν_i l'indice di nilpotenza associato a λ_i . Gli spazi K_i sono invarianti per A . Denotiamo con $A_i : K_i \rightarrow K_i$ la restrizione di A a K_i e con $N_i := A_i - \lambda_i I$. Allora, $N_i : K_i \rightarrow K_i$ è un'applicazione lineare nilpotente di indice ν_i , quindi, per il Lemma delle basi di catene di Jordan (liii), esistono k_i catene di Jordan $\gamma_1^{(i)}, \dots, \gamma_{k_i}^{(i)}$ tali che (cfr. (D.105)) $K_i = \langle \gamma_1^{(i)} \rangle \oplus \dots \oplus \langle \gamma_{k_i}^{(i)} \rangle$ e (D.96) segue dal Lemma di riduzione a blocchi (xli), dall'Osservazione D.25–(ii) e dall'Osservazione D.24–(iv).

La forma normale di Jordan (a meno dell'ordine) è determinata da: gli autovalori λ_i e i relativi indici di nilpotenza ν_i (che determinano lo *splitting* (D.100)); gli autospazi generalizzati $K_i = \ker N_i^{\nu_i}$ (relativi a λ_i) che hanno dimensione a_i ; e poi, per ogni $1 \leq i \leq m$ dal numero di blocchi di Jordan $\kappa_j(\lambda_i)$ con autovalore λ_i di ordine j che è dato da (cfr. (D.106)) da $\kappa_j(\lambda_i) := q_j - q_{j+1}$ con $q_j := \dim(N_i^j K_i \cap \ker N_i)$. Tutti gli oggetti coinvolti sono geometrici (ossia, non dipendono da scelte di basi) e dunque la forma canonica di Jordan è unica a meno dell'ordine. ■

Esempi

(1) L'esempio non banale più semplice è ovviamente il caso di $A \in \mathbb{C}^{2 \times 2}$. In questo caso lo spettro è formato o da due elementi distinti $\lambda_1 \neq \lambda_2$ o da un solo elemento λ (di molteplicità algebrica 2). Nel primo caso la molteplicità geometrica è $g = 2$ e la matrice è diagonalizzabile tramite una matrice $U = [u^1, u^2]$ con u^i autovettore relativo a λ_i . Nel caso di unico autovalore λ si definisce $N = A - \lambda I_2$ e si cerca il polinomio minimo che è o $(z - \lambda)$ oppure $(z - \lambda)^2$. Nel caso $p_{\min} = (z - \lambda)$, si ha $\nu = 1$ e per il Lemma di *splitting* $\mathbb{C}^2 = \ker(A - \lambda I_2)$, e quindi $g = 2$ e la matrice è di nuovo diagonalizzabile tramite $U = [u^1, u^2]$ con u^i autovettori indipendenti di

⁶³Come al solito, $a(\lambda) =$ dimensione algebrica di λ ; $g(\lambda) =$ dimensione geometrica di λ ; $\nu(\lambda)$ indice di nilpotenza di λ .

A . Nel caso $p_{\min} = (z - \lambda)^2$, si ha $\nu = 2$ ($N \neq 0$, $N^2 = 0$), e $\mathbb{C}^2 = \ker(A - \lambda I_2)^2$. In questo caso la forma canonica di Jordan è $J_2(\lambda)$ e per trovare U che coniughi A a $J_2(\lambda)$ bisogna trovare una catena di Jordan (di ordine 2), ossia, bisogna trovare un vettore u tale che $Nu \neq 0$ e si prenderà $U = [Nu, u]$.

Vediamo un esempio concreto. Sia $A = \begin{pmatrix} 1 & 1 \\ -1 & 3 \end{pmatrix}$. Il polinomio caratteristico di A è $z^2 - 4z + 4 = (z - 2)^2$. Dunque, $\sigma(A) = \{2\}$ e $a(2) = 2$. La matrice $N = A - \lambda I_2 = \begin{pmatrix} -1 & 1 \\ -1 & 1 \end{pmatrix}$. Quindi $N \neq 0$ e $N^2 = 0$, dunque $\nu = 2$. La forma canonica di Jordan è quindi data da $J_2(2)$. Determiniamo una catena di Jordan di ordine 2. Poiché $Nu = (u_2 - u_1, u_2 - u_1)$, possiamo prendere $u = (0, 1)$ e $Nu = (1, 1)$. In conclusione, se $U := \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix}$, si ha che $U^{-1}AU = J_2(2)$.

(2) Sia $A = \begin{pmatrix} 3 & 1 & -1 \\ 0 & 2 & 0 \\ 1 & 1 & 1 \end{pmatrix}$. Allora, $p_{\text{car}}(z) = -(z - 2)^3$. Dunque, $\sigma(A) = \{2\}$ e $a(2) = 3$.

Sia $N := A - 2I_3 = \begin{pmatrix} 1 & 1 & -1 \\ 0 & 0 & 0 \\ 1 & 1 & -1 \end{pmatrix}$. Dal Lemma di *splitting* D.22 segue che $\mathbb{C}^3 = \ker N^\nu$ con

ν indice di nilpotenza di N (relativo all'unico autovalore $\lambda = 2$). Poiché $N^2 = 0$ si ha che $\nu = \nu(2) = 2$. Da questo già possiamo concludere che la forma normale di Jordan di A è $J := \text{block}(J_2(2), J_1(2)) = \text{block}(J_2(2), 2)$. Per trovare esplicitamente una matrice $U \in GL(3)$ tale che $U^{-1}AU = J$ implementiamo l'algoritmo descritto nell'Osservazione D.25-(i).

Determiniamo $N^{\nu-1}\mathbb{C}^3 = N\mathbb{C}^3$. Poiché $Nu = (u_1 + u_2 - u_3, 0, u_1 + u_2 - u_3)$, si ha che $N\mathbb{C}^3$ è la retta complessa $\{z(1, 0, 1) \mid z \in \mathbb{C}\}$ e una sua base è $W_1 = \{w\}$ con $w := (1, 0, 1)$. Determiniamo ora W_0 in modo che $W_1 \cup W_0$ sia una base di $\ker N$, che coincide col piano complesso $\{u \in \mathbb{C}^3 \mid u_1 + u_2 - u_3 = 0\}$. Possiamo, prendere $W_0 = \{w^0\}$, con $w^0 = (0, 1, 1)$. Infine per calcolare la catena di Jordan di vertice w , determiniamo u tale che $Nu = w$. Possiamo prendere $u := (1, 0, 0)$. In conclusione:

$$U := [Nu, u, w^0] = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 1 \end{pmatrix} : \quad U^{-1}AU = \begin{pmatrix} 2 & 1 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$