

Appunti di Analisi II

Corso di Aldo Pratelli e Carlo Carminati

Pietro Marinari e Simone Sacconi

Università di Pisa
A.A. 2023-2024

Indice

1	Spazi e strutture	1
1.1	Spazi topologici	1
1.2	Spazi normati e spazi metrici	2
1.2.1	Norme su \mathbb{R}^d	6
1.3	Compattezza, completezza, totale limitatezza	7
1.3.1	Insieme di Cantor	10
2	Continuità	13
2.1	Lipschitzianità	18
2.1.1	Equivalenze delle norme su \mathbb{R}^d	20
2.1.2	Distanza di un punto da un insieme	21
2.2	Limiti inferiori e superiori e semicontinuità	22
2.3	Convessità e continuità	25
2.4	Funzioni omogenee	27
3	Calcolo differenziale	29
3.1	Derivate successive e Teorema di Taylor	33
3.2	Massimi e minimi liberi e vincolati	36
3.3	Teorema della funzione implicita	39
3.4	Convessità e differenziabilità	44
3.5	Funzioni omogenee e radiali	45
4	Successioni di funzioni	49
4.1	Convergenza uniforme	49
4.2	Teorema di Ascoli-Arzelà	53
4.3	Serie di funzioni	55
4.3.1	Serie di potenze	56
4.3.2	Convergenza totale	57
4.4	Teorema di Stone-Weierstraß	58
4.4.1	Applicazioni di Stone-Weierstraß	62
5	Equazioni differenziali	67
5.1	Teoremi di esistenza e unicità	68
5.2	Studio qualitativo delle soluzioni	71
5.3	Ordine superiore	75
5.4	Sistemi di equazioni differenziali	76
5.4.1	Esponenziale di matrici	76
5.4.2	Sistemi autonomi	78
5.5	PDE: alcuni esempi	84
6	Integrazione	87
6.1	Teoria della misura	87
6.2	Teoremi di convergenza	93
6.3	Misura prodotto e Teorema di Fubini	97
6.4	Teorema del cambio di variabile	100

6.4.1	Coordinate polari, cilindriche e sferiche	102
6.5	Funzione Gamma e integrale di funzioni radiali	103
6.6	Integrazione su curve	106
6.7	Integrazione su superfici e Teorema della divergenza	108
6.7.1	Integrazione su grafici	112
6.7.2	Integrazione su superfici in \mathbb{R}^3	113
6.7.3	Superfici di rotazione	113
6.7.4	Funzioni armoniche e proprietà della media	114
7	Forme differenziali	115
7.1	Forme chiuse e forme esatte	117
7.2	Lemma di Poincaré e Pullback di forme	122
7.3	Teorema di Stokes	126
7.4	Calcolo di integrali di forme su superfici	128
A	Passaggio di derivata sotto il segno di integrale	131

Capitolo 1

Spazi e strutture

Uno degli argomenti centrali di questo corso sarà l'analisi matematica in più variabili: talvolta sarà possibile estendere nei modi più intuitivi quanto già visto nel corso di Analisi I, talvolta invece la direzione sarà sostanzialmente diversa. Considerato, a questo proposito, il rilievo dell'insieme \mathbb{R} nel corso precedente, iniziamo con la definizione dello spazio \mathbb{R}^d e di alcune sue strutture.

- Lo spazio \mathbb{R}^d visto come insieme è il *prodotto cartesiano* d volte dell'insieme \mathbb{R} , vale a dire

$$\mathbb{R}^d = \mathbb{R} \times \cdots \times \mathbb{R} = \{(x_1, \dots, x_d) \mid \forall i = 1, \dots, d \ x_i \in \mathbb{R}\} ;$$

- L'insieme \mathbb{R}^d è dotato di un'operazione di somma naturale che corrisponde alla somma per componenti:

$$+ : \mathbb{R}^d \times \mathbb{R}^d \rightarrow \mathbb{R}^d \quad (x, y) \mapsto x + y := (x_1 + y_1, \dots, x_d + y_d) ;$$

- L'insieme \mathbb{R}^d è dotato di una moltiplicazione per scalare, che corrisponde a moltiplicare tutte le componenti per tale scalare:

$$\cdot : \mathbb{R} \times \mathbb{R}^d \rightarrow \mathbb{R}^d \quad (\lambda, x) \mapsto \lambda x := (\lambda x_1, \dots, \lambda x_d) .$$

In particolare, la tripla $(\mathbb{R}^d, +, \cdot)$ è un \mathbb{R} -spazio vettoriale.

A questo punto vorremmo definire i limiti di successioni o di funzioni, e a tal proposito vorremmo ricalcare la definizione di limite già data su \mathbb{R} : per quanto riguarda le successioni, ad esempio, avevamo definito il limite di una successione $\{x_n\}_{n \in \mathbb{N}}$ a valori in \mathbb{R} come

$$\lim_{n \rightarrow +\infty} x_n = L \quad \equiv \quad \forall \varepsilon > 0 \ \exists \bar{n} \text{ tale che } n \geq \bar{n} \Rightarrow |x_n - L| < \varepsilon ;$$

se a questo punto $\{x_n\}_{n \in \mathbb{N}}$ fosse una successione in \mathbb{R}^d , la definizione sarebbe ben posta a patto di specificare il significato del "modulo" di un vettore.

1.1 Spazi topologici

Definizione 1.1 (Spazio topologico). Uno *spazio topologico* è una coppia (X, τ) , con X insieme e $\tau \subseteq \mathcal{P}(X)$, tale che:

- $\emptyset, X \in \tau$;
- se $A_1, A_2 \in \tau$, allora $A_1 \cap A_2 \in \tau$;
- se $\{A_i\}_{i \in \mathcal{I}} \subseteq \tau$, allora $\bigcup_{i \in \mathcal{I}} A_i \in \tau$.

Chiamiamo l'insieme τ *topologia*, mentre i suoi elementi sono detti *aperti*. Il complementare di un aperto si dice *chiuso*.

Definizione 1.2 (Chiusura). Dato X spazio topologico e $Z \subseteq X$, la *chiusura* di Z , che denotiamo \bar{Z} , è il più piccolo chiuso che contiene Z , ossia:

$$\bar{Z} = \bigcap_{\substack{C \text{ chiuso} \\ C \supseteq Z}} C .$$

Definizione 1.3 (Parte interna). Dato X spazio topologico e $Z \subseteq X$, la *parte interna* di Z , che denotiamo $\overset{\circ}{Z}$ o $\text{int}(Z)$, è il più grande aperto contenuto in Z , ossia:

$$\overset{\circ}{Z} = \bigcup_{\substack{A \text{ aperto} \\ A \subseteq Z}} A .$$

Osservazione 1.1. Poiché aperti e chiusi sono in bigezione tramite il complementare, vale la seguente relazione:

$$\overset{\circ}{Z} = X \setminus \overline{(X \setminus Z)} .$$

Definizione 1.4 (Frontiera). Dato X spazio topologico e $Z \subseteq X$, la *frontiera* di Z , che denotiamo ∂Z , è la sua chiusura meno la sua parte interna:

$$\partial Z = \bar{Z} \setminus \overset{\circ}{Z} .$$

1.2 Spazi normati e spazi metrici

Definizione 1.5 (Norma). Dato uno spazio vettoriale reale o complesso V , una *norma* su X è una funzione $\|\cdot\| : V \rightarrow \mathbb{R}$ tale che:

- i. $\|v\| \geq 0 \forall v \in V$ e vale $\|v\| = 0$ se e solo se $v = 0$;
- ii. $\|\lambda v\| = |\lambda| \|v\| \forall v \in V, \forall \lambda \in \mathbb{K}$, ossia la norma è *omogenea*;
- iii. $\|v + w\| \leq \|v\| + \|w\| \forall v, w \in V$, vale a dire che vale la *disuguaglianza triangolare*.

Definizione 1.6 (Spazio normato). Uno *spazio normato* è una coppia $(V, \|\cdot\|)$ dove V è uno spazio vettoriale reale o complesso e $\|\cdot\|$ è una norma su V .

Recuperiamo adesso alcuni concetti di Geometria I. Nello spazio vettoriale standard \mathbb{K}^n , con $\text{char} \mathbb{K} \neq 2$, disponiamo di un *prodotto scalare standard* $\langle \cdot, \cdot \rangle$ dato da

$$\langle x, y \rangle = \sum_{i=1}^d x_i y_i .$$

Abbiamo definito *spazio euclideo* (V, ϕ) uno spazio vettoriale V reale (o complesso) con un prodotto scalare (o hermitiano) ϕ definito positivo, vale a dire tale che $\forall v \in V \phi(v, v) \geq 0$ e vale l'uguaglianza se e solo se $v = 0$; abbiamo successivamente osservato che in uno spazio euclideo il prodotto induce una norma: $\|v\| = \sqrt{\phi(v, v)}$. Dato che in particolare il prodotto scalare standard in \mathbb{R}^d è definito positivo, possiamo porre

$$\|\cdot\| : \mathbb{R}^d \rightarrow \mathbb{R} \quad x \mapsto \|x\| := \sqrt{\langle x, x \rangle} = \sqrt{\sum_{i=1}^d x_i^2} .$$

Indichiamo questa norma standard ancora con $|\cdot|$. È immediato osservare che la norma indotta soddisfa le prime due proprietà della definizione. Per dimostrare la terza, introduciamo la disuguaglianza di Cauchy-Schwarz.

Proposizione 1.2: Disuguaglianza di Cauchy-Schwarz

Sia (V, ϕ) uno spazio euclideo, allora vale la seguente disuguaglianza:

$$|\phi(v, w)| \leq \|v\| \|w\| \quad \forall v, w \in V .$$

Dimostrazione. Supponiamo innanzitutto V \mathbb{R} -spazio vettoriale. Sia $t \in \mathbb{R}$ un parametro, allora

$$\begin{aligned} 0 &\leq \phi(v + tw, v + tw) = \phi(v, v) + 2\phi(v, tw) + \phi(tw, tw) = \\ &= \|v\|^2 + 2t\phi(v, w) + t^2\|w\|^2 \quad ; \end{aligned}$$

questo polinomio in t è una parabola non negativa, pertanto ha determinante ridotto

$$\Delta/4 = \phi(v, w)^2 - \|v\|^2 \|w\|^2 \leq 0, \text{ da cui passando alla radice } |\phi(v, w)| \leq \|v\| \|w\| .$$

Consideriamo invece il caso di V \mathbb{C} -spazio vettoriale con ϕ prodotto hermitiano definito positivo. La disuguaglianza è vera per $w = 0$, quindi supponiamo $\phi(w, w) \neq 0$. Similmente al caso reale si ha

$$0 \leq \|v - \lambda w\|^2 = \phi(v - \lambda w, v - \lambda w) = \phi(v, v) - \lambda\phi(v, w) - \bar{\lambda}\phi(w, v) + |\lambda|^2\phi(w, w) .$$

Adesso scegliamo $\lambda = \frac{\phi(w, v)}{\phi(w, w)}$: sviluppando i conti otteniamo

$$\begin{aligned} 0 &\leq \phi(v, v) - \frac{\phi(w, v)}{\phi(w, w)}\phi(v, w) - \left(\frac{\phi(w, v)}{\phi(w, w)}\right)\phi(w, v) + \left(\frac{\phi(w, v)}{\phi(w, w)}\right)\frac{\phi(w, v)}{\phi(w, w)}\phi(w, w) = \\ &= \phi(v, v) - \frac{\phi(w, v)\phi(v, w)}{\phi(w, w)} - \frac{\phi(w, v)\phi(w, v)}{\phi(w, w)} + \frac{\phi(w, v)\phi(w, v)}{\phi(w, w)} = \phi(v, v) - \frac{|\phi(v, w)|^2}{\phi(w, w)} ; \end{aligned}$$

abbiamo dunque trovato che $|\phi(v, w)|^2 \leq \phi(v, v) \cdot \phi(w, w)$, che implica $|\phi(v, w)| \leq \|v\| \|w\|$. □

Osserviamo che $\|a + b\|^2 = \langle a + b, a + b \rangle = \langle a, a \rangle + \langle b, b \rangle + 2\langle a, b \rangle = \|a\|^2 + \|b\|^2 + 2\langle a, b \rangle$. Ora per la disuguaglianza vale $|\langle a, b \rangle| \leq \|a\| \|b\|$, perciò otteniamo

$$\|a + b\|^2 = \|a\|^2 + \|b\|^2 + 2\langle a, b \rangle \leq \|a\|^2 + \|b\|^2 + 2\|a\| \|b\| = (\|a\| + \|b\|)^2 ,$$

da cui, passando alle radici,

$$\|a + b\| \leq \|a\| + \|b\| .$$

La norma è di fatto un'applicazione che attribuisce ad ogni vettore non nullo una grandezza positiva, che possiamo interpretare come *lunghezza* del vettore.

Definizione 1.7 (Distanza). Dato un insieme X , una *distanza* (o *metrica*) su X è una funzione $d : X \times X \rightarrow \mathbb{R}$ tale che:

- i. $d(x, y) \geq 0 \forall x, y \in X$ e vale $d(x, y) = 0$ se e solo se $x = y$;
- ii. $d(x, y) = d(y, x) \forall x, y \in X$, ossia la distanza è *simmetrica*;
- iii. $d(x, z) \leq d(x, y) + d(y, z) \forall x, y, z \in X$, cioè vale la *disuguaglianza triangolare*.

Definizione 1.8 (Spazio metrico). Uno spazio metrico è una coppia (X, d) dove X è un insieme e d è una distanza su X .

In uno spazio normato la norma induce una metrica ponendo $d(x, y) := \|x - y\|$. Le proprietà *i* e *ii* seguono immediatamente. Per la *iii*, osserviamo che ponendo $a = x - y$ e $b = y - z$, otteniamo $d(x, z) = \|x - z\| = \|a + b\| \leq \|a\| + \|b\| = \|x - y\| + \|y - z\| = d(x, y) + d(y, z)$, che è la disuguaglianza triangolare cercata.

Possiamo a questo punto definire il limite di una successione a valori in \mathbb{R}^d , o più in generale in uno spazio metrico.

Definizione 1.9 (Limite di successione). Dato uno spazio metrico (X, d) e una successione $\{x_n\}_{n \in \mathbb{N}}$ a valori in X , diciamo che x_n tende ad un *limite* $\ell \in X$, e scriveremo ancora

$$\lim_{n \rightarrow +\infty} x_n = \ell \quad \text{oppure} \quad x_n \xrightarrow[n \rightarrow +\infty]{} \ell ,$$

se $\forall \varepsilon > 0 \exists \bar{n}$ tale che $\forall n > \bar{n}$ vale $d(x_n, \ell) < \varepsilon$.

Potremmo però pensare di indurre il limite di \mathbb{R} in \mathbb{R}^d , come abbiamo fatto per la somma, definendo un limite per componenti:

Definizione 1.10 (Limite per componenti). Data $\{x_n\}_{n \in \mathbb{N}}$ una successione a valori in \mathbb{R}^d e un valore $\ell \in \mathbb{R}^d$, definiamo il *limite per componenti* di x_n come

$$x_n \xrightarrow[n \rightarrow +\infty]{} \ell \quad \iff \quad \forall i \ (x_n)_i \xrightarrow[n \rightarrow +\infty]{} \ell_i .$$

Proposizione 1.3

In \mathbb{R}^d le definizioni di limite e di limite per componenti sono equivalenti.

Dimostrazione. (\implies) Sia $i \in \{1, \dots, d\}$, dobbiamo dimostrare che $(x_n)_i \rightarrow \ell_i$. Dato $\varepsilon > 0$, sappiamo per ipotesi che $\exists \bar{n}$ tale che $\forall n > \bar{n} \ |x_n - \ell| < \varepsilon$; in particolare

$$|(x_n)_i - \ell_i| = \sqrt{(x_n - \ell)_i^2} \leq \sqrt{(x_n - \ell)_1^2 + \dots + (x_n - \ell)_d^2} = |x_n - \ell| < \varepsilon .$$

(\impliedby) Dato $\varepsilon > 0$, per ipotesi esistono \bar{n}_i per ogni $i \in \{1, \dots, d\}$ tali che per $n > \bar{n}_i$ si ha $|x_{n_i} - \ell_i| < \varepsilon$; allora ponendo $\bar{n} = \max\{\bar{n}_i\}$ troviamo che

$$|x_n - \ell| = \sqrt{(x_n - \ell)_1^2 + \dots + (x_n - \ell)_d^2} < \sqrt{\varepsilon^2 d} = \varepsilon \sqrt{d} ,$$

che tende a 0 per $\varepsilon \rightarrow 0$. □

Definizione 1.11 (Palla aperta). Dato (X, d) spazio metrico, $x \in X$ e $r > 0$, definiamo *palla aperta* di centro x e raggio r l'insieme $B(x, r) = B_r(x) = \{z \in X \mid d(x, z) < r\}$.

Definizione 1.12 (Aperto metrico). Dato (X, d) spazio metrico e $A \subseteq X$, diciamo che A è un *aperto (metrico)* se $\forall x \in A \exists r > 0$ tale che $B(x, r) \subseteq A$.

Definizione 1.13 (Chiuso metrico). Dato (X, d) spazio metrico, un sottoinsieme $C \subseteq X$ si dice *chiuso* se il complementare $X \setminus C$ è aperto.

Proposizione 1.4

- i.* L'unione arbitraria di una famiglia di aperti $\mathcal{F} = \{A_i\}_{i \in \mathcal{I}}$ è aperta.
ii. L'intersezione di una famiglia finita di aperti $\mathcal{F} = \{A_i\}_{i=1, \dots, n}$ è aperta.

Dimostrazione. *i.* Per ogni elemento $x \in \bigcup \mathcal{F}$, $x \in A_i$ per qualche $i \in \mathcal{I}$, ma allora per definizione di aperto $\exists r > 0$ tale che $B(x, r) \subseteq A_i$, quindi necessariamente

$$B(x, r) \subseteq \bigcup_{i \in \mathcal{I}} A_i \quad \text{cioè} \quad \bigcup_{i \in \mathcal{I}} A_i \quad \text{è aperto.}$$

- ii.* Consideriamo il caso di due insiemi $A_1, A_2 \in \mathcal{F}$. Sia $x \in A_1 \cap A_2$: poiché A_1 e A_2 sono aperti, esistono $r_1, r_2 > 0$ tali che $B(x, r_1) \subseteq A_1$ e $B(x, r_2) \subseteq A_2$. Senza perdita di generalità, sia $r_1 \leq r_2$; allora vale $B(x, r_1) \subseteq A_1$ e $B(x, r_1) \subseteq B(x, r_2) \subseteq A_2$, perciò $B(x, r_1) \subseteq A_1 \cap A_2$. Per induzione, si ottiene che l'intersezione finita di aperti è aperta. \square

Osserviamo che in generale non è vero che l'intersezione numerabile di aperti è aperta. Ad esempio, possiamo considerare in \mathbb{R} la famiglia di palle $A_n = (-\frac{1}{n}, \frac{1}{n})$; la loro intersezione è $\bigcap_{n \in \mathbb{N}} A_n = \{0\}$, che non è un aperto.

Per il passaggio al complementare, otteniamo la seguente:

Proposizione 1.5

- i.* L'unione di una famiglia finita di chiusi $\mathcal{F} = \{A_i\}_{i=1, \dots, n}$ è chiusa.
ii. L'intersezione arbitraria di una famiglia di chiusi $\mathcal{F} = \{A_i\}_{i \in \mathcal{I}}$ è chiusa.

Abbiamo quindi dimostrato che uno spazio metrico (X, d) è uno spazio topologico con la topologia indotta dalla distanza; possiamo allora considerare la chiusura, la parte interna e la frontiera di sottinsiemi degli spazi metrici.

Definizione 1.14 (Chiusura). Dato uno spazio metrico (X, d) e un sottoinsieme $D \subseteq S$, si definisce la *chiusura* di D come

$$\overline{D} := \{x \in X \mid \forall \delta > 0 \exists y \in D \ d(x, y) < \delta\} .$$

I punti della chiusura di \overline{D} sono detti *punti di accumulazione* di D ¹.

Lemma 1.6

L'insieme C è chiuso se e solo se $\overline{C} = C$.

Dimostrazione. (\implies) Sia $X \setminus C$ aperto. Poiché $C \subseteq \overline{C}$ è sempre vero, considero $x \in \overline{C}$ e voglio mostrare $x \in C$. Se per assurdo fosse $x \notin C$, allora $x \in X \setminus C$, quindi $\exists r > 0 : B(x, r) \subseteq X \setminus C$. \nexists
 (\impliedby) Se $\overline{C} = C$, devo mostrare che $X \setminus C$ è aperto. Sia $v \in X \setminus C$: se $\exists r > 0$ tale che $B(v, r) \subseteq X \setminus C$, si conclude. Altrimenti, $\forall r > 0 \ B(v, r) \cap C \neq \emptyset$, cioè $v \in \overline{C} = C$. \nexists \square

¹La chiusura di un insieme consiste quindi nei punti *arbitrariamente vicini* all'insieme; altre volte si definiscono *punti di accumulazione* i punti $x \in X$ tali che $\forall \delta > 0 \exists y \in D$ con $y \neq x$ e $d(x, y) < \delta$: in tal caso chiamiamo i punti di \overline{D} *punti di aderenza*, che si dividono in punti di accumulazione e punti isolati.

Lemma 1.7

L'insieme $C \subseteq X$ è chiuso se e solo se $\forall \{x_n\} \subseteq C$ tale che $x_n \rightarrow \bar{x}$ si ha che $\bar{x} \in C$.

Dimostrazione. (\implies) Se $\{x_n\} \subseteq C$ e $x_n \rightarrow \bar{x}$, allora $\bar{x} \in \overline{C} = C$.

(\impliedby) Sia $\bar{x} \in \overline{C}$: allora $\forall \varepsilon B(\bar{x}, \varepsilon) \cap C \neq \emptyset$. Per ogni $n \in \mathbb{N}$, prendo $x_n \in C \cap B(\bar{x}, \frac{1}{n})$. Si ha che $x_n \rightarrow \bar{x}$, dunque per ipotesi $\bar{x} \in C$, cioè $\overline{C} \subseteq C$, e di conseguenza C è chiuso. \square

Gli spazi metrici saranno oggetti centrali in questo corso; la nozione di distanza può però dare origine a spazi metrici molto diversi dagli spazi euclidei o da \mathbb{R}^d : vediamo un esempio.

Esempio 1.8. Consideriamo lo spazio di funzioni $X = \{f : [0, 1] \rightarrow \mathbb{R} \text{ continue}\}$, allora la funzione $d(f, g) = \max\{|f(x) - g(x)|, x \in [0, 1]\}$ è una distanza su X . Infatti

- la distanza è ben definita, in quanto $|f(x) - g(x)|$ è una funzione continua definita su $[0, 1]$ e di conseguenza ammette massimo;
- $|f(x) - g(x)| \geq 0$, e inoltre se $f \neq g \exists x$ tale che $f(x) \neq g(x)$, pertanto $\max |f(x) - g(x)| > 0$;
- $|f(x) - g(x)| = |g(x) - f(x)|$;
- dimostriamo che $d(f, h) \leq d(f, g) + d(g, h)$: $d(f, h) = |f(\bar{x}) - h(\bar{x})|$ per un certo $\bar{x} \in [0, 1]$; inoltre $|f(\bar{x}) - g(\bar{x})| \leq d(f, g)$ e $|g(\bar{x}) - h(\bar{x})| \leq d(g, h)$ per definizione di d , quindi

$$d(f, h) = |f(\bar{x}) - h(\bar{x})| \leq |f(\bar{x}) - g(\bar{x})| + |g(\bar{x}) - h(\bar{x})| \leq d(f, g) + d(g, h) ,$$

dove la prima disuguaglianza vale per disuguaglianza triangolare su \mathbb{R} .

1.2.1 Norme su \mathbb{R}^d

Su \mathbb{R}^d possiamo definire molteplici norme. Le più comuni sono la *norma 1*, la *norma 2* e la *norma ∞* date da

$$\|x\|_1 = \sum_{i=1}^d |x_i| \quad \|x\|_2 = \left(\sum_{i=1}^d x_i^2 \right)^{\frac{1}{2}} \quad \|x\|_\infty = \sup_{1 \leq i \leq d} x_i$$

Le prime due norme rientrano in una categoria di norme più ampia:

Definizione 1.15 (Norma p). Dato $p \in [1, +\infty)$, si definisce la *norma p* come l'applicazione

$$\|\cdot\|_p : \mathbb{R}^d \rightarrow \mathbb{R} \\ x \mapsto \|x\|_p = \left(\sum_{i=1}^d |x_i|^p \right)^{\frac{1}{p}}$$

Osservazione 1.9. Possiamo considerare la norma ∞ una norma p , nel senso che:

$$\|\cdot\|_\infty = \lim_{p \rightarrow +\infty} \|\cdot\|_p .$$

Proposizione 1.10

Dato $x \in \mathbb{R}^d$, vale:

$$\|x\|_1 \leq \sqrt{d} \|x\|_2 \leq d \|x\|_\infty \leq d \|x\|_1$$

Dimostrazione. La prima disuguaglianza segue dal fatto che, dato $x \in \mathbb{R}^d$, scegliendo $y = (\operatorname{sgn}(x_i))_i$, per Cauchy-Schwarz, otteniamo

$$\|x\|_1^2 = \left(\sum_{i=1}^d |x_i| \right)^2 = \langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle = d \|x\|_2^2 ;$$

la seconda si ottiene osservando che

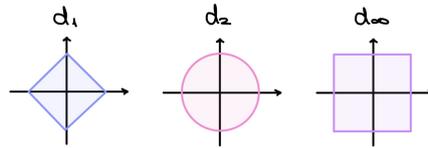
$$\|x\|_2^2 = \sum_{i=1}^d |x_i|^2 \leq \sum_{i=1}^d \|x\|_\infty^2 = d \|x\|_\infty^2 ;$$

l'ultima segue dal fatto

$$\|x\|_\infty = \sup_{1 \leq i \leq d} |x_i| \leq \sum_{i=1}^d |x_i| = \|x\|_1 .$$

□

Ad esempio, le palle unitarie relative alle distanze indotte rispettivamente dalla norma 1, norma 2 e norma ∞ , sono:



In realtà, per gli spazi vettoriali reali (o complessi) di dimensione finita vale un risultato più forte, che mostreremo nel prossimo capitolo: in questi casi tutte le norme sono topologicamente equivalenti.

1.3 Compattezza, completezza, totale limitatezza

Definizione 1.16 (Compattezza per successioni). Dato uno spazio metrico (X, d) , un insieme $D \subseteq X$ si dice *compatto per successioni* o *sequenzialmente compatto* se ogni successione $\{x_n\} \subseteq D$ ammette una sottosuccessione convergente in D .

Osservazione 1.11. Un insieme compatto per successioni è chiuso.

Definizione 1.17 (Ricoprimento). Un *ricoprimento* di X è una famiglia $\{X_i, i \in I\}$ di sottoinsiemi di X tali che

$$\bigcup_{i \in I} X_i = X .$$

Definizione 1.18 (Compattezza). Uno spazio topologico X si dice *compatto* (per *ricoprimenti*) se per ogni ricoprimento di aperti di X esiste un sottoricoprimento finito.

Se (X, d) è uno spazio metrico e $Z \subseteq X$ è un suo sottoinsieme, su Z abbiamo una topologia di sottospazio indotta naturalmente dalla restrizione della metrica d a $Z \times Z$. Per sapere se Z è compatto, dobbiamo sapere come si comportano gli aperti di Z (e non gli aperti di X) rispetto ai sottoricoprimenti. In particolare la topologia di sottospazio è definita come

$$A \subseteq Z \text{ è aperto di } Z \iff \exists B \subseteq X \text{ aperto di } X \text{ tale che } A = B \cap Z .$$

Definizione 1.19 (Totale limitatezza). Uno spazio metrico (X, d) si dice *totalmente limitato* se $\forall r > 0 \exists k \in \mathbb{N}, \exists x_1, \dots, x_k \in X$ tali che

$$\bigcup_{i=1}^k B(x_i, r) = X .$$

La limitatezza è una proprietà prettamente metrica, che non incide in alcun modo sulla topologia di uno spazio: dato infatti (X, d) spazio metrico non limitato, possiamo definire una nuova distanza $d' : X \times X \rightarrow \mathbb{R}$ dove $d'(x, y) := d(x, y) \wedge 1$: risulta che d e d' sono equivalenti, ossia che esprimono gli stessi aperti. Tuttavia (X, d') è chiaramente limitato poiché $\forall x \in X$ vale $B_{d'}(x, 2) = X$.

Definizione 1.20 (Successione di Cauchy). Dato uno spazio metrico (X, d) , una successione $\{x_n\}$ a valori in X si dice *successione di Cauchy* se $\forall \varepsilon > 0 \exists \bar{n}$ tale che $\forall m, n > \bar{n}$ vale $d(x_m, x_n) < \varepsilon$.

Osservazione 1.12. Una successione convergente è sempre di Cauchy.

Infatti, fissato $\varepsilon > 0$, esiste \bar{n} tale che $\forall n \geq \bar{n}$ vale $d(x_n, \bar{x}) < \frac{\varepsilon}{2}$, dove \bar{x} è il limite della successione. Quindi, presi $n, m \geq \bar{n}$, vale $d(x_n, x_m) \leq d(x_n, \bar{x}) + d(\bar{x}, x_m) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$.

Definizione 1.21 (Completezza). Uno spazio metrico (X, d) si dice *completo* se tutte le successioni di Cauchy sono convergenti.

Definizione 1.22 (Spazio di Banach). Uno *spazio di Banach* è uno spazio normato $(E, \|\cdot\|)$ completo rispetto alla distanza indotta dalla norma.

Definizione 1.23 (Proprietà di intersezione finita). Sia \mathcal{H} una famiglia non vuota di insiemi non vuoti, si dice che \mathcal{H} ha la *proprietà dell'intersezione finita* (**PIF**) se per ogni sottoinsieme finito $\{A_1, \dots, A_n\} \subseteq \mathcal{H}$ si ha che $A_1 \cap \dots \cap A_n \neq \emptyset$.

Teorema 1.13

Dato uno spazio topologico X , sono equivalenti:

- i. X è compatto;
- ii. se $\{F_\alpha\}_{\alpha \in \mathcal{A}}$ è una famiglia di chiusi di X con la **PIF**, allora

$$\bigcap_{\alpha \in \mathcal{A}} F_\alpha \neq \emptyset .$$

Dimostrazione.

$i \implies ii$ Supponiamo per assurdo che una famiglia di chiusi di X $\{F_\alpha\}_{\alpha \in \mathcal{A}}$ abbia la **PIF** ma che l'intersezione della famiglia sia vuota, ossia

$$\forall \alpha_1, \dots, \alpha_n \in \mathcal{A} \quad F_{\alpha_1} \cap \dots \cap F_{\alpha_n} \neq \emptyset \quad \text{ma} \quad \bigcap_{\alpha \in \mathcal{A}} F_\alpha = \emptyset ;$$

sia $A_\alpha := F_\alpha^c \forall \alpha$, per cui A_α è aperto. Allora

$$\bigcup_{\alpha \in \mathcal{A}} A_\alpha = \bigcup_{\alpha \in \mathcal{A}} F_\alpha^c = \left(\bigcap_{\alpha \in \mathcal{A}} F_\alpha \right)^c = \emptyset^c = X ,$$

pertanto per compattezza di $X \exists \alpha_1, \dots, \alpha_n \in \mathcal{A}$ tali che

$$X = \bigcup_{i=1}^n A_{\alpha_i} = \left(\bigcap_{i=1}^n F_{\alpha_i} \right)^c \implies \bigcap_{i=1}^n F_{\alpha_i} = \emptyset \quad \text{. } \neq$$

$ii \implies i$ Sia $\{A_\alpha\}_{\alpha \in \mathcal{A}}$ un ricoprimento di aperti di X . Supponiamo per assurdo che $\{A_\alpha\}_{\alpha \in \mathcal{A}}$ non ammetta alcun sottoricoprimento finito, e denotiamo $F_\alpha = A_\alpha^c$: allora $\forall \alpha_1, \dots, \alpha_n$

$$X \supsetneq \bigcup_{i=1}^n A_{\alpha_i} = \bigcup_{i=1}^n F_{\alpha_i}^c = \left(\bigcap_{i=1}^n F_{\alpha_i} \right)^c \implies \bigcap_{i=1}^n F_{\alpha_i} \neq \emptyset ,$$

ossia $\{F_\alpha\}_{\alpha \in \mathcal{A}}$ ha la **PIF**. Per ipotesi allora

$$\emptyset \neq \bigcap_{\alpha \in \mathcal{A}} F_\alpha = \left(\bigcup_{\alpha \in \mathcal{A}} A_\alpha \right)^c \implies \bigcup_{\alpha \in \mathcal{A}} A_\alpha \subsetneq X \quad \text{. } \neq$$

□

Se non riesco a capire che un insieme è compatto mi sparo.

Aldo Pratelli

Teorema 1.14

Dato uno spazio metrico (X, d) , sono equivalenti:

- i.* X è compatto per successioni;
- ii.* X è compatto;
- iii.* X è totalmente limitato e completo;

Dimostrazione. Mostriamo prima l'equivalenza tra *i* e *iii* e poi che *ii* è equivalente a *i* e *iii* assieme (e dunque a ciascuna singolarmente).

$i \implies iii$ Supponiamo per assurdo che X non sia completo, allora $\exists \{x_n\}$ successione di Cauchy che non converge a nessun limite: mostriamo che $\{x_n\}$ non può ammettere sottosuccessioni convergenti. Se esistesse $x_{\alpha(n)}$ sottosuccessione che converge ad un certo \bar{x} , allora dato $\varepsilon > 0 \exists \bar{n}'$ tale che $\forall n > \bar{n}' d(x_{\alpha(n)}, \bar{x}) < \frac{\varepsilon}{2}$; inoltre dal momento che x_n è di Cauchy, $\exists \bar{n}''$ tale che $\forall j, \ell > \bar{n}'' d(x_j, x_\ell) < \frac{\varepsilon}{2}$. A questo punto $\forall j > \max\{\bar{n}', \bar{n}''\}$ si ha per disuguaglianza triangolare $d(x_j, \bar{x}) \leq d(x_j, x_{\alpha(j)}) + d(x_{\alpha(j)}, \bar{x}) < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$, ossia x_n converge a \bar{x} . \neq
Supponiamo invece che X non sia totalmente limitato, allora $\exists r > 0$ tale che non esiste alcun ricoprimento finito di X con palle di raggio r ; costruiamo allora per ricorsione la successione

$$x_0 \in X \text{ qualsiasi, } \quad x_{k+1} \notin \bigcup_{i=0}^k B(x_i, r) ,$$

tuttavia segue dalla costruzione che $\forall i \neq j d(x_i, x_j) \geq r$, e quindi nessuna sottosuccessione può convergere. \neq

$iii \implies i$ Sia $\{x_n\}$ una successione a valori in X . Per ipotesi di totale limitatezza $\forall r > 0$ esiste un ricoprimento finito di palle di raggio r , e dato che il ricoprimento è finito, in almeno una delle palle $\{x_n\}$ deve comparire per infiniti indici, ossia almeno una delle palle contiene una sottosuccessione. Costruiamo per ricorsione:

$$\{x_n^{(1)}\} \subseteq B(\bar{x}_1, 1) \text{ sottosuccessione di } x_n \text{ per } n \geq 1$$

$$\begin{aligned} \{x_n^{(2)}\} &\subseteq B(\bar{x}_2, \frac{1}{2}) \text{ sottosuccessione di } x_n^{(1)} \text{ per } n \geq 2 \\ &\vdots \\ \{x_n^{(k+1)}\} &\subseteq B(\bar{x}_{k+1}, \frac{1}{k+1}) \text{ sottosuccessione di } x_n^{(k)} \text{ per } n \geq k+1 \end{aligned}$$

allora se chiamiamo y_n la successione che ha come k -esimo termine $x_n^{(k)}$, per costruzione sottosuccessione di x_n , y_n è una successione di Cauchy, e per completezza converge.

$ii \implies i$ Sia $\{x_n\}$ una successione a valori in X e sia $C = \{x_n\} \subseteq X$ l'insieme dei valori della successione. Se un elemento di C compare un numero infinito di volte in x_n , possiamo costruire una sottosuccessione definitivamente costante che dunque converge. Supponiamo quindi che ogni valore compaia per un numero finito di indici.

Dato $c \in C$ consideriamo il valore $\inf\{d(c, e), e \in C \text{ e } e \neq c\}$: se per un certo c tale valore è 0, allora in C esistono punti arbitrariamente vicini a c , e dunque troviamo una sottosuccessione che converge a c . Altrimenti $\forall c \in C \exists r(c)$ tale che $B(c, r(c)) \cap C = \{c\}$. Se C non è chiuso, $\exists \ell \in \overline{C} \setminus C$, ma allora possiamo estrarre una sottosuccessione che converge a ℓ ; se invece C fosse chiuso, allora

$$\bigcup_{c \in C} B(c, r(c)) \cup (X \setminus C)$$

sarebbe un ricoprimento di X in aperti che non ammette alcun sottoricoprimento finito. \nexists

$i \wedge iii \implies ii$ Sia $\{A_i\}_{i \in \mathcal{I}}$ un ricoprimento di aperti. Per ogni $x \in X \exists r > 0$ tale che $B(x, r)$ è contenuta in almeno un aperto del ricoprimento. Chiamiamo $r(x) := \sup\{r > 0 \mid \exists i \in \mathcal{I} B(x, r) \subseteq A_i\} \wedge 1$. Supponiamo per assurdo che $\inf\{r(x), x \in X\} = 0$, allora sia $\{x_n\}$ tale che $r(x_n) < \frac{1}{n}$, e per ipotesi possiamo supporre a meno di sottosuccessioni che $x_n \rightarrow \bar{x}$.

Osserviamo che dati $y, z \in X$ vale che $r(y) \geq r(z) - d(y, z)$: infatti $\forall \rho < r(z)$ se $d(x, y) < \rho - d(y, z)$ allora $d(x, z) < \rho$, poiché per la disuguaglianza triangolare $d(x, z) \leq d(x, y) + d(y, z) < \rho$. Abbiamo quindi dimostrato che $\forall \rho < r(z)$ vale $B(y, \rho - d(y, z)) \subseteq B(z, \rho) \subseteq A_i$ per qualche i , ma allora $r(y) \geq \rho - d(y, z)$, e passando al sup, $r(y) \geq r(z) - d(y, z)$. A questo punto $r(x_n) \geq r(\bar{x}) - d(x_n, \bar{x})$, tuttavia $r(x_n) \rightarrow 0$ e $d(x_n, \bar{x}) \rightarrow 0$, pertanto deve valere $r(\bar{x}) = 0$. \nexists

A questo punto $\inf\{r(x), x \in X\} > 0$: poniamo $\delta = \frac{\inf\{r(x), x \in X\}}{2}$, allora $\forall x \in X \exists i \in \mathcal{I}$ tale che $B(x, \delta) \subseteq A_i$. Adesso possiamo concludere sfruttando la totale limitatezza di X : $\exists k \in \mathbb{N}$, $z_1, \dots, z_k \in X$ tali che

$$X \subseteq \bigcup_{i=1}^k B(z_i, \delta)$$

e tuttavia $\exists j(1), \dots, j(k) \in \mathcal{I}$ tali che $B(z_i, \delta) \subseteq A_{j(i)}$, e dunque

$$X \subseteq \bigcup_{i=1}^k A_{j(i)} .$$

□

1.3.1 Insieme di Cantor

Definizione 1.24 (Insieme di Cantor). Consideriamo le mappe $\phi_0, \phi_2 : [0, 1] \rightarrow [0, 1]$ date da $\phi_0(x) = \frac{1}{3}x$ e $\phi_2(x) = \frac{1}{3}x + \frac{2}{3}$. Definiamo per ricorsione

$$C_0 = [0, 1], \quad C_{n+1} = \phi_0(C_n) \cup \phi_2(C_n) \quad \forall n \geq 0 :$$

l'insieme di Cantor è definito come l'insieme

$$C = \bigcap_{n=0}^{+\infty} C_n .$$

Intuitivamente, l'insieme di Cantor è l'insieme che si ottiene a partire dal segmento $[0, 1]$ dividendolo in tre segmenti uguali e scartando il segmento centrale, per poi iterare ricorsivamente il processo ai segmenti ottenuti al passo precedente.

Osserviamo subito che $C \neq \emptyset$, perché ad esempio $0, 1 \in C$.



Si dimostrano per induzione i seguenti fatti:

- C_n è compatto $\forall n$;
- $C_n \subseteq [0, 1] \forall n$;
- $C_{n+1} \subseteq C_n \forall n$.

Da ciò segue in particolare che l'insieme di Cantor è compatto.

Osserviamo che ogni $x \in [0, 1]$ può essere scritto in rappresentazione ternaria come

$$x = \sum_{i=1}^{+\infty} \varepsilon_i 3^{-i} \quad \text{dove } \varepsilon_i \in \{0, 1, 2\}.$$

Ad esempio, se $\varepsilon_i = 2$ per ogni i , otteniamo il valore massimo, infatti:

$$x = \sum_{i=1}^{+\infty} 2 \cdot 3^{-i} = 2 \cdot \left(\frac{1}{1 - \frac{1}{3}} - 1 \right) = 1.$$

Sfruttando la rappresentazione ternaria, possiamo descrivere l'insieme di Cantor come segue.

Proposizione 1.15

Valgono le seguenti espressioni:

$$C_n = \left\{ x = \sum_{i=1}^{+\infty} \varepsilon_i 3^{-i} \mid \varepsilon_i \in \{0, 2\} \forall 1 \leq i \leq n \right\},$$

$$C = \left\{ x = \sum_{i=1}^{+\infty} \varepsilon_i 3^{-i} \mid \varepsilon_i \in \{0, 2\} \forall i \in \mathbb{N} \right\}.$$

Dimostrazione. Per induzione su n .

- La tesi è vera per $n = 1$.

Infatti, se $\varepsilon_1 = 0$,

$$x = \sum_{i=2}^{+\infty} \varepsilon_i 3^{-i} = \sum_{i=1}^{+\infty} \varepsilon_{i+1} 3^{-i-1} = \frac{1}{3} \sum_{i=1}^{+\infty} \varepsilon_{i+1} 3^{-i} \in \left[0, \frac{1}{3} \right];$$

se $\varepsilon_1 = 2$,

$$x = \frac{2}{3} + \sum_{i=2}^{+\infty} \varepsilon_i 3^{-i} = \frac{2}{3} + \frac{1}{3} \sum_{i=1}^{+\infty} \varepsilon_{i+1} 3^{-i} \in \left[\frac{2}{3}, 1 \right].$$

- Supponiamo la tesi valga per n .

Ora $C_{n+1} = \phi_0(C_n) \cup \phi_2(C_n)$. Osserviamo che

$$\begin{aligned} \phi_i(C_n) &= \left\{ x = \frac{i}{3} + \frac{1}{3} \sum_{h=1}^{+\infty} \varepsilon_h 3^{-h} \mid \varepsilon_h \in \{0, 2\} \ 1 \leq h \leq n \right\} = \\ &= \left\{ x = \sum_{h=1}^{+\infty} \varepsilon'_h 3^{-h} \mid \varepsilon'_1 = i, \ \varepsilon'_h = \varepsilon_{h-1} \in \{0, 2\} \text{ per } 2 \leq h \leq n+1 \right\} \end{aligned}$$

e quindi

$$C_{n+1} = \phi_0(C_n) \cup \phi_2(C_n) = \left\{ x = \sum_{h=1}^{+\infty} \varepsilon_h 3^{-h} \mid \varepsilon_h \in \{0, 2\} \text{ per } 1 \leq h \leq n+1 \right\}.$$

□

Osservazione 1.16. C ha la cardinalità del continuo, infatti esiste una funzione surgettiva e monotona

$$F : C \longrightarrow [0, 1]$$

$$x = \sum_{i=1}^{+\infty} \varepsilon_i 3^{-i} \mapsto \sum_{i=1}^{+\infty} \frac{\varepsilon_i}{2} 2^{-i}$$

Inoltre esiste ed è unica \tilde{F} estensione di F a $[0, 1]$ monotona tale che $\tilde{F}|_C = F$, e vale $\tilde{F} = \sup_{t \leq x} F(t)$.

Tale funzione è nota come *scala del diavolo* o *scala di Cantor*.

Per l'insieme di Cantor valgono i seguenti fatti:

- $\text{int}(C) = \emptyset$: infatti C è contenuto nell'unione di 2^n intervalli di ampiezza 3^{-n} . Più precisamente, C è totalmente sconnesso;
- $\partial C = C$: infatti $\partial C \subseteq C$ perché C è chiuso e $C \subseteq \partial C$ perché C non ha punti interni;
- $\partial(\mathbb{R} \setminus C) = C$;
- C non ha punti isolati.

Proposizione 1.17

Definiamo $C + C := \{x + y \mid x, y \in C\}$; allora $C + C = [0, 2]$, ossia per ogni $t \in [0, 2]$ esiste una coppia $(x, y) \in C \times C$ tale che $x + y = t$.

Dimostrazione. Osserviamo che la tesi equivale a

$$C \times C \cap \{(x, y) \in \mathbb{R}^2 \mid x + y = t\} \neq \emptyset \quad \forall t \in [0, 2].$$

Osserviamo anche che

$$C \times C = \bigcap_{n \in \mathbb{N}} C_n \times C_n$$

quindi basta mostrare che

$$F_n := C_n \times C_n \cap \{(x, y) \in \mathbb{R}^2 \mid x + y = t\} \neq \emptyset \quad \forall t \in [0, 2] \quad \forall n \in \mathbb{N} :$$

infatti vale $F_{n+1} \subseteq F_n \quad \forall n$, quindi $F_{n_1} \cap \dots \cap F_{n_k} = F_{\bar{n}}$ dove $\bar{n} = \max\{n_1, \dots, n_k\}$; pertanto, se $F_n \neq \emptyset \quad \forall n$, $\{F_n \mid n \in \mathbb{N}\}$ ha la **PIF** e quindi $\bigcap_{n \in \mathbb{N}} F_n \neq \emptyset$.

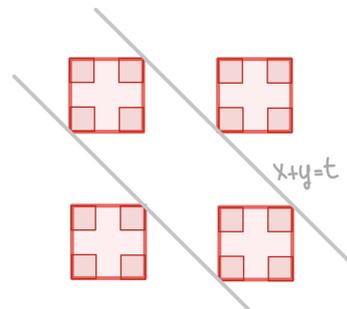
Dimostriamo la tesi per induzione su n .

- La tesi vale per $n = 1$.
- Supponiamo la tesi vera per n .

Sia $t \in [0, 2]$: possiamo scegliere $i, j \in \{0, 2\}$ e $t' \in [0, 2]$ tali che

$$t = \frac{i + j + t'}{3}$$

(idealmente i e j rappresentano la traslazione orizzontale e verticale della copia di C_n in C_{n+1}). Per ipotesi induttiva, esiste $(x', y') \in C_n \times C_n$ tale che $x' + y' = t'$.



Poniamo allora

$$x := \frac{i + x'}{3} \in C_{n+1} \quad \text{e} \quad y := \frac{j + y'}{3} \in C_{n+1}$$

per cui vale $(x, y) \in C_{n+1} \times C_{n+1}$. Ora $x + y = \frac{i + j + x' + y'}{3} = \frac{i + j + t'}{3} = t$.

□

Capitolo 2

Continuità

Occupiamoci adesso di definire il limite di funzioni su spazi metrici.

Definizione 2.1 (Limite). Siano (X, d_X) e (Y, d_Y) due spazi metrici e siano $\bar{x} \in X$, $\bar{y} \in Y$. Data una funzione $f : X \rightarrow Y$, diciamo che \bar{y} è il *limite* di f per $x \rightarrow \bar{x}$ se $\forall \varepsilon > 0 \exists \delta > 0$ tale che $\forall x \in X$ con $0 < d_X(x, \bar{x}) < \delta$ vale $d_Y(f(x), \bar{y}) < \varepsilon$.

Alcune volte si preferisce definire il limite $\lim_{x \rightarrow \bar{x}} f(x) = \bar{y}$ nel seguente modo:

$$\forall \varepsilon > 0 \exists \delta > 0 : \forall x \in X \quad d_X(x, \bar{x}) < \delta \Rightarrow d_X(f(x), \bar{y}) < \varepsilon ,$$

ossia in sostanza si permette il caso $x = \bar{x}$. Questa scelta ha sostanzialmente due effetti collaterali: da un lato il valore assunto dalla funzione in un punto è rilevante rispetto al valore del limite della funzione in quel punto, e in particolare non potrà mai accadere che

$$\lim_{x \rightarrow \bar{x}} f(x) \neq f(\bar{x}) \quad \text{ma soltanto} \quad \lim_{\substack{x \rightarrow \bar{x} \\ x \neq \bar{x}}} f(x) \neq f(\bar{x}) ;$$

dall'altro possiamo definire senza contraddizione il limite di una funzione per $x \rightarrow \bar{x}$ con \bar{x} punto isolato (e di conseguenza sui punti isolati il valore del limite coincide con il valore della funzione).

Definizione 2.2 (Limite infinito). Siano (X, d) uno spazio metrico, $\bar{x} \in X$ e $f : X \rightarrow \mathbb{R}$: diciamo che f tende a $+\infty$ per $x \rightarrow \bar{x}$ se $\forall M > 0 \exists \delta > 0$ tale che $\forall x \in X$ con $0 < d(x, \bar{x}) < \delta$ vale $M < f(x)$; analogamente diciamo che f tende a $-\infty$ per $x \rightarrow \bar{x}$ se $\forall M > 0 \exists \delta > 0$ tale che $\forall x \in X$ con $0 < d(x, \bar{x}) < \delta$ vale $f(x) < -M$.

Definizione 2.3 (Limite della norma). Sia $(V, \|\cdot\|)$ uno spazio normato, (Y, d) uno spazio metrico, $\bar{y} \in Y$. Data una funzione $f : V \rightarrow Y$, diciamo che \bar{y} è il *limite* di f per $\|v\| \rightarrow +\infty$, scriviamo

$$\lim_{\|v\| \rightarrow +\infty} f(v) = \bar{y} ,$$

se $\forall \varepsilon > 0 \exists N > 0$ tale che $\forall v \in V$ con $\|v\| > N$ vale $d(f(v), \bar{y}) < \varepsilon$.

Inoltre se $Y = \mathbb{R}$, diciamo che f tende a $+\infty$ (rispettivamente $-\infty$) per $\|v\| \rightarrow +\infty$ se $\forall M > 0 \exists N > 0$ tale che $\forall v \in V$ con $\|v\| > N$ vale $M < f(x)$ (rispettivamente $f(x) < -M$).

Se abbiamo una funzione definita su \mathbb{R}^d , possiamo fissare una retta passante per l'origine e studiare la funzione ristretta a tale direzione: questa nuova funzione diventa una funzione da \mathbb{R} a \mathbb{R} .

Definizione 2.4 (Limite direzionale). Data $f : \mathbb{R}^d \rightarrow \mathbb{R}$, $\omega \in \mathbb{R}^d \setminus \{0\}$, si chiama *limite direzionale* in direzione ω il limite

$$\lim_{t \rightarrow +\infty} f(t\omega) .$$

Proposizione 2.1

Data $f : \mathbb{R}^d \rightarrow \mathbb{R}$, se esiste $\ell \in \mathbb{R}$ tale che $f(x) \xrightarrow{\|x\| \rightarrow +\infty} \ell$, allora

$$\forall \omega \in \mathbb{R}^d \setminus \{0\} \quad \lim_{t \rightarrow +\infty} f(t\omega) = \ell .$$

Dimostrazione. Sia $\varepsilon > 0$; per ipotesi $\exists K > 0$ tale che per $|x| > K$ vale $|f(x) - \ell| < \varepsilon$, pertanto se $t > K/|\omega|$, $|t\omega| > K$ e dunque $|f(t\omega) - \ell| < \varepsilon$. \square

Osserviamo che il viceversa è falso, ossia è possibile che tutti i limiti direzionali esistano e coincidano, ma che la funzione non abbia limite globale; ad esempio possiamo considerare

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \quad (x, y) \mapsto \begin{cases} 1 & \text{se } y = x^2 \\ 0 & \text{altrimenti} \end{cases} :$$

in questo caso tutti i limiti direzionali sono 0, ma il limite di f per $\|x\| \rightarrow +\infty$ non esiste.

Definizione 2.5 (Continuità). Siano (X, d_X) e (Y, d_Y) due spazi metrici: una funzione $f : X \rightarrow Y$ si dice *continua* in $\bar{x} \in X$ se $\forall \varepsilon > 0 \exists \delta > 0 : \forall x \in X$ con $d_X(x, \bar{x}) < \delta$ vale $d_Y(f(x), f(\bar{x})) < \varepsilon$. Se f è continua in $\bar{x} \forall \bar{x} \in X$, diciamo che f è *continua*.

Osservazione 2.2. Data $f : X \rightarrow Y$ funzione tra due spazi metrici e $\bar{x} \in X$, sono equivalenti:

- i. f è continua in \bar{x} ;
- ii. $\lim_{x \rightarrow \bar{x}} f(x) = f(\bar{x})$.

La dimostrazione è esattamente quella vista nel corso di Analisi I.

Osservazione 2.3. Data una funzione $f : D \rightarrow \mathbb{R}$ con $D \subseteq \mathbb{R}^d$ (o più in generale una funzione tra spazi metrici) possiamo definire il limite di f per ogni $x \in \bar{D}$ non isolato. Se volessimo definire il limite ℓ per $x \rightarrow \bar{x}$ con \bar{x} al di fuori di questo insieme avremmo la seguente situazione: $\forall \varepsilon > 0 \exists \delta > 0$ tale che $0 < |x - \bar{x}| < \delta \Rightarrow |f(x) - \ell| < \varepsilon$; il guaio è che se $\bar{x} \notin \bar{D}$, $\exists \delta' > 0$ tale che $\forall y \in D \quad d(x, y) \geq \delta'$, pertanto scegliendo $\delta = \delta'$, l'antecedente dell'implicazione sarebbe falsa $\forall x$, e di conseguenza l'implicazione sempre vera (a vuoto), indipendentemente dal valore di ℓ .

Definizione 2.6 (Continuità per successioni). Siano (X, d_X) e (Y, d_Y) due spazi metrici: una funzione $f : X \rightarrow Y$ si dice *continua per successioni* se $\forall \{x_n\} \subseteq X$ tale che $x_n \rightarrow \bar{x}$, si ha che $f(x_n) \rightarrow f(\bar{x})$.

Diamo adesso delle caratterizzazioni per le funzioni continue.

Proposizione 2.4

Dati due spazi metrici (X, d_X) e (Y, d_Y) e una funzione $f : X \rightarrow Y$, vale che

$$f \text{ è continua} \iff \forall A \subseteq Y \text{ aperto, } f^{-1}(A) \text{ è aperto.}$$

Dimostrazione. (\implies) Sia $A \subseteq Y$ aperto e sia $x \in f^{-1}(A)$; vogliamo trovare $r > 0$ tale che $B(x, r) \subseteq f^{-1}(A)$. Dato che $f(x) \in A$ e A è aperto, esiste $\varepsilon > 0$ tale che $B(f(x), \varepsilon) \subseteq A$: per la continuità di f esiste $\delta > 0$ tale che $\forall y \in X \ d_X(x, y) < \delta \implies d_Y(f(x), f(y)) < \varepsilon$, cioè $\forall y \in B(x, \delta)$ si ha che $f(y) \in B(f(x), \varepsilon) \subseteq A$, ossia $y \in f^{-1}(A)$.

(\impliedby) Sia f tale che $\forall A \subseteq Y$ aperto, $f^{-1}(A)$ è aperto. Dati $x \in X$ e $\varepsilon > 0$, considero $A = B(f(x), \varepsilon)$. Poiché A è aperto, $f^{-1}(A)$ è aperto e $x \in f^{-1}(A)$, perciò $\exists \delta > 0$ tale che $B(x, \delta) \subseteq f^{-1}(A)$, ossia $\forall y \in X \ d_X(x, y) < \delta \implies d_Y(f(x), f(y)) < \varepsilon$. \square

Proposizione 2.5

Dati due spazi metrici (X, d_X) e (Y, d_Y) e una funzione $f : X \rightarrow Y$, vale che

$$f \text{ è continua} \iff \forall C \subseteq Y \text{ chiuso, } f^{-1}(C) \text{ è chiuso.}$$

Dimostrazione. (\implies) Sia $C \subseteq Y$ un chiuso e sia A' l'aperto complementare, per cui $C = Y \setminus A'$: allora $f^{-1}(C) = f^{-1}(Y \setminus A') = X \setminus f^{-1}(A')$, che è chiuso poiché $f^{-1}(A') = A$ è aperto dato che f è continua.

(\impliedby) Sia f tale che $\forall C \subseteq Y$ chiuso, $f^{-1}(C)$ è chiuso. Ora sia $A' \subseteq Y$ aperto: per ipotesi $X \setminus f^{-1}(A') = f^{-1}(Y \setminus A') = C'$ chiuso, quindi $\forall A \subseteq \mathbb{R}$ aperto, $f^{-1}(A)$ è aperto, ossia f è continua. \square

Proposizione 2.6

Dati due spazi metrici (X, d_X) e (Y, d_Y) e una funzione $f : X \rightarrow Y$, vale che

$$f \text{ è continua} \iff f \text{ è continua per successioni}$$

Dimostrazione. (\implies) Sia $\{x_n\} \subseteq X$ una successione tale che $x_n \rightarrow \bar{x}$; poiché f è continua, $\forall \varepsilon > 0 \exists \delta > 0$ tale che se $d_X(x, \bar{x}) < \delta$ si ha $d_Y(f(x), f(\bar{x})) < \varepsilon$; allora, per definizione di limite, $\exists n_0$ tale che $n \geq n_0 \implies d_X(x_n, \bar{x}) < \delta$ e dunque $d_Y(f(x_n), f(\bar{x})) < \varepsilon$. Quindi, complessivamente, $\forall \varepsilon > 0 \exists n_0$ tale che $n \geq n_0 \implies d_Y(f(x_n), f(\bar{x})) < \varepsilon$, cioè $f(x_n) \rightarrow f(\bar{x})$.

(\impliedby) Sia per assurdo f discontinua in $\bar{x} \in X$, ossia esiste $\varepsilon > 0$ tale che per ogni δ esiste $x \in B_\delta(\bar{x})$ con $d_Y(f(x), f(\bar{x})) \geq \varepsilon$; in particolare troviamo una successione $\{x_n\} \subseteq X$ tale che $x_n \rightarrow \bar{x}$ e per ogni n vale $d_Y(f(x_n), f(\bar{x})) \geq \varepsilon$, pertanto $f(x_n) \not\rightarrow f(\bar{x})$. \nexists \square

Osservazione 2.7. La prima implicazione della proposizione è sempre vera, qualunque siano gli spazi topologici considerati. La seconda implicazione invece vale solo negli *spazi sequenziali*, ed in particolare negli spazi metrici.

Proposizione 2.8

La composizione di funzioni continue è una funzione continua.

Dimostrazione. Siano $f : X \rightarrow Y$, $g : Y \rightarrow Z$ e $g \circ f : X \rightarrow Z$. Dato A aperto, $(g \circ f)^{-1}(A) = f^{-1}(g^{-1}(A))$. Ora $g^{-1}(A) = A'$ aperto perché g è continua; $f^{-1}(A') = A''$ aperto perché f è continua. Quindi $(g \circ f)^{-1}(A) = A''$ aperto, e perciò $g \circ f$ è continua. \square

Osservazione 2.9. Se chiamiamo $f_i : \mathbb{R}^d \rightarrow \mathbb{R}$ la proiezione sulla i -esima coordinata $f_i(x) = x_i$, osserviamo che è una funzione continua.

Infatti $|f_i(x) - f_i(y)| = |x_i - y_i| \leq \sqrt{(x_1 - y_1)^2 + \dots + (x_d - y_d)^2} = |x - y|$.

Perciò lo studio della continuità delle funzioni vettoriali si riconduce allo studio su \mathbb{R} tramite la composizione con le coordinate.

Osservazione 2.10. Data $g : \mathbb{R}^d \rightarrow \mathbb{R}$ continua, l'insieme $D = \{x \in \mathbb{R}^d : g(x) \geq 0\}$ è chiuso e l'insieme $F = \mathbb{R}^d \setminus D = \{x \in \mathbb{R}^d : g(x) < 0\}$ è aperto (analogo per l'altro verso della disuguaglianza).

Infatti, data $\{x_n\} \subseteq D$, cioè $g(x_n) \geq 0 \forall x$, tale che $x_n \rightarrow \bar{x}$, segue che $g(x_n) \rightarrow g(\bar{x})$, quindi $g(\bar{x}) \geq 0$, e perciò $\bar{x} \in D$.

Teorema 2.11: Teorema di Weierstraß

Dato (X, d) spazio metrico, $D \subseteq X$ compatto per successioni e $f : D \rightarrow \mathbb{R}$ continua, allora f ammette massimo e minimo.

Dimostrazione. Sia $\{x_n\} \subseteq D$ tale che $f(x_n) \rightarrow \inf f$. Allora $x_n \rightarrow \bar{x}$ a meno di sottosuccessioni. Quindi, per la continuità di f ,

$$f(\bar{x}) = \lim_{n \rightarrow +\infty} f(x_n) = \inf f ,$$

cioè \bar{x} è un punto di minimo. La dimostrazione è analoga per il punto di massimo. \square

Teorema 2.12: Teorema di Heine-Borel

Un insieme $C \subseteq \mathbb{R}^d$ è compatto per successioni se e solo se C è chiuso e limitato.

Dimostrazione. (\implies) Se C non fosse chiuso, esisterebbe $\{x_n\} \subseteq C$, $x_n \rightarrow \bar{x} \notin C$. Qualunque sottosuccessione converge a \bar{x} , pertanto C non è compatto per successioni, che è un assurdo.

Se C non fosse limitato, esisterebbe $\{x_n\} \subseteq C$ tale che $d(x_n, y) > n$ con $y \in C$. Se \tilde{x}_n è una sottosuccessione che tende a \bar{x} , si ha che per $n \gg 1$ $d(\tilde{x}_n, \bar{x}) < 1$. Ma allora $n < d(\tilde{x}_n, y) \leq d(\tilde{x}_n, \bar{x}) + d(\bar{x}, y) < 1 + d(\bar{x}, y)$, quindi C non è compatto per successioni, che è assurdo.

(\impliedby) Sia $C \subseteq \mathbb{R}^d$ chiuso e limitato e $\{x_n\} \subseteq C$. Ora la successione delle prime componenti $(x_1)_n$ è una successione in \mathbb{R} limitata, quindi esiste una sottosuccessione di (x_n) tale che la prima componente converge. Ora, la successione delle seconde componenti della sottosuccessione appena ottenuta è limitata in \mathbb{R} , quindi si può riapplicare lo stesso ragionamento. In conclusione, in d passaggi otteniamo una sottosuccessione che converge a \bar{x} ; siccome C è chiuso, $\bar{x} \in C$. \square

Osservazione 2.13. Per quanto appena dimostrato, un insieme compatto per successioni è sempre chiuso e limitato; il viceversa vale nel caso in cui lo spazio considerato sia \mathbb{R}^d . Questo è dovuto alla seguente considerazione.

Sia $E \subseteq \mathbb{R}$ chiuso e limitato, allora $E \subseteq [a, b]$. Se si ha $\{e_n\} \subseteq E$, allora $\{e_n\} \subseteq [a, b]$, quindi esiste una sottosuccessione $\{e'_n\}$ tale che $e'_n \rightarrow e \in [a, b]$. Si conclude che $e \in E$ poiché E è chiuso.

Osservazione 2.14. Chiaramente, se $\sup f = +\infty$, f non ammette massimo globale, mentre se $\inf f = -\infty$, f non ammette minimo globale.

Teorema 2.15: Teorema di Weierstraß generalizzato

Data $f : \mathbb{R}^d \rightarrow \mathbb{R}$ continua, se esiste $\ell = \lim_{|x| \rightarrow +\infty} f(x)$ finito, allora vale che:

i. se esiste a tale che $f(a) > \ell$, allora f ammette massimo;

ii. se esiste b tale che $f(b) < \ell$, allora f ammette minimo.

Dimostrazione. Dimostriamo il punto *i* (il punto *ii* è analogo). Poiché $f(x) \rightarrow \ell$ per $|x| \rightarrow +\infty$, scegliendo $\varepsilon = f(a) - \ell > 0$ esiste $R > 0$ tale che se $|x| > R$ vale $|f(x) - \ell| < \varepsilon$ ed in particolare $f(x) < \ell + \varepsilon = f(a)$. Ora consideriamo la palla chiusa $\overline{B}(0, R)$ di centro 0 e raggio R : questo è un insieme chiuso e limitato in \mathbb{R}^d , quindi è compatto, pertanto f ammette massimo M su $\overline{B}(0, R)$. Ora, se $x_M \in \overline{B}(0, R)$ è tale che $f(x_M) = M$, per $|x| > R$ vale che $f(x) < f(a) \leq f(x_M)$, quindi x_M è un massimo globale. \square

In realtà, varianti di questa versione generalizzata del teorema di Weierstraß possono essere utilizzate anche nel caso in cui la funzione f non sia definita su tutto \mathbb{R}^d . Consideriamo i seguenti esempi.

Esempio 2.16. Sia $g : \mathbb{R}^d \setminus \{0\} \rightarrow \mathbb{R}$ tale che $\lim_{|z| \rightarrow +\infty} g(z) = 5$, $\lim_{z \rightarrow 0} g(z) = +\infty$ e $g(\bar{z}) = 2$, allora f ammette minimo.

Infatti, per definizione di limite, $\exists R > 0$ tale che $|z| > R \implies g(z) > 3$, e anche $\exists r > 0$ tale che $0 < z < r \implies g(z) > 3$. Definisco ora $A = \{z \in \mathbb{R}^d \mid r \leq |z| \leq R\}$: per costruzione $A \neq \emptyset$ perché $\bar{z} \in A$; inoltre A è chiuso e limitato, e g è continua, quindi esiste $\hat{z} \in A$ punto di minimo locale di g . Per definizione di A , \hat{z} è minimo globale di g .

Esempio 2.17. Consideriamo la funzione

$$f : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$$

$$f(x, y) = \frac{\sin(x^2 + y^2)}{x^2 + y^4} .$$

Cerchiamo di determinare se f è limitata, e in caso positivo di dire se esistono massimo e minimo. Innanzitutto osserviamo che

$$x^2 + y^4 \geq x^2 + y^2 - 1 = |(x, y)|^2 - 1 \xrightarrow{|(x, y)| \rightarrow +\infty} +\infty \implies \lim_{|(x, y)| \rightarrow +\infty} f(x, y) = 0$$

Calcolando poi il limite nell'origine lungo gli assi cartesiani, troviamo

$$\lim_{x \rightarrow 0} f(x, 0) = 1 \quad \lim_{y \rightarrow 0} f(y, 0) = +\infty$$

Quindi f non ammette limite globale in $(0, 0)$, e perciò non rientriamo nella casistica dell'esempio precedente. Inoltre f non è limitata e non ha massimo. Per determinare se ammette minimo o meno, possiamo percorrere strade diverse.

i. Osserviamo che in $\{0 < x^2 + y^2 < \pi\}$, f è positiva. Inoltre $\forall \delta \exists R > 0$ tale che se $|(x, y)| > R$ allora $-\delta < f(x, y) < \delta$. Scegliamo allora δ tale che

$$f\left(\sqrt{\frac{3}{2}}\pi, 0\right) < -\delta .$$

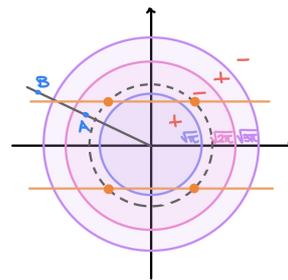
Definiamo ora $A = \{(x, y) : \pi \leq x^2 + y^2 \leq R\}$: $A \neq \emptyset$ perché ad esempio $(\sqrt{\frac{3}{2}}\pi, 0) \in A$. Risulta che A è chiuso e limitato, quindi compatto, e perciò f ammette minimo su A in un certo punto x_m . Segue infine che x_m è punto di minimo globale: infatti, se $x^2 + y^2 < \pi$, $f(x, y) > 0$ e quindi $f(x_m) < f(x, y)$ poiché $f(x_m) < 0$; se invece $x^2 + y^2 > R$, risulta

$$f(x_m) \leq f\left(\sqrt{\frac{3}{2}}\pi, 0\right) < -\delta < f(x, y) .$$

ii. Costruiamo una successione di corone circolari centrate in 0 in cui f è alternativamente positiva e negativa, come in figura. Consideriamo poi la parte negativa di f :

$$f^- = \begin{cases} f & \text{sulle corone negative} \\ 0 & \text{sulle corone positive o nulle} \end{cases}$$

Ora possiamo estendere f^- per continuità in 0, ponendo $f(0) = 0$; quindi, per Weierstraß generalizzato, f ammette minimo.



Questa costruzione ci permette inoltre di calcolare esplicitamente il minimo di f . Innanzitutto il minimo deve essere in una delle corone negative; inoltre, poiché $\sin(x^2 + y^2)$ è continua, assume tutti i valori da 0 a -1 nelle corone negative. Quindi, fissata una direzione, possiamo scegliere un punto A su tale direzione nella prima corona negativa e un punto B sempre su tale direzione ma nella seconda corona negativa, tale che $\sin(A) = \sin(B)$; tuttavia il denominatore è necessariamente cresciuto, quindi risulta $f(A) \leq f(B)$. Concludiamo quindi che il minimo si trova nella prima corona negativa.

Osservando ora che il seno è costante sulle circonferenze (poiché la norma è costante), cerchiamo il minimo di $x^2 + y^4$ sulla circonferenza $x^2 + y^2 = r^2$, con $\sqrt{\pi} < r < \sqrt{2\pi}$:

$$x^2 + y^4 = y^4 - y^2 + r^2 = h(y) \text{ e } h'(y) = 4y^3 - 2y = 0 \iff y = 0 \vee y = \pm\sqrt{\frac{1}{2}}$$

Studiando la funzione h , notiamo che il punto $y = 0$ è di massimo, quindi lo scartiamo. A questo punto bisogna trovare, per simmetria della funzione nei quattro quadranti, il minimo della funzione $g(x, \pm\sqrt{\frac{1}{2}}) = \frac{\sin(x^2 + \frac{1}{2})}{x^2 + \frac{1}{4}}$, che è un problema di Analisi I.

2.1 Lipschitzianità

Definizione 2.7 (Uniforme continuità). Siano (X, d_X) e (Y, d_Y) due spazi metrici: una funzione $f : X \rightarrow Y$ si dice *uniformemente continua* se $\forall \varepsilon > 0 \exists \delta > 0$ tale che $\forall x_1, x_2 \in X$ con $d_X(x_1, x_2) < \delta$ si ha $d_Y(f(x_1), f(x_2)) < \varepsilon$.

Osservazione 2.18. Una funzione uniformemente continua è chiaramente anche continua.

Teorema 2.19: Teorema di Heine-Cantor

Siano (X, d_X) e (Y, d_Y) due spazi metrici e $f : X \rightarrow Y$ una funzione continua. Se X è compatto, allora f è uniformemente continua.

Dimostrazione. Supponiamo per assurdo che $\exists \varepsilon > 0 \forall \delta > 0$ esistono $x, y \in X$ tali che $d_X(x, y) < \delta$ ma $d_Y(f(x), f(y)) \geq \varepsilon$. Prendiamo, in corrispondenza di tale ε , $\delta = \frac{1}{n}$ per $n \geq 1$, e definiamo le successioni $\{x_n\}$ e $\{y_n\}$ date dall'ipotesi assurda. Poiché X è compatto, $x_n \rightarrow \bar{x}$ a meno di sottosuccessioni.

Poiché $d_X(x_n, y_n) < \frac{1}{n} \rightarrow 0$, abbiamo che

$$d_X(y_n, \bar{x}) \leq d_X(y_n, x_n) + d_X(x_n, \bar{x}) \xrightarrow{n \rightarrow +\infty} 0$$

quindi anche $y_n \rightarrow \bar{x}$ a meno di sottosuccessioni. Vale perciò:

$$d_Y(f(x_n), f(y_n)) \leq d_Y(f(x_n), f(\bar{x})) + d_Y(f(\bar{x}), f(y_n))$$

dove il secondo membro tende a 0 per $n \rightarrow +\infty$ per la continuità di f . \nexists

□

Definizione 2.8 (Lipschitzianità). Siano (X, d_X) e (Y, d_Y) due spazi metrici: una funzione $f : X \rightarrow Y$ si dice *lipschitziana di costante L* (oppure *L -Lipschitz*), con $L > 0$, se $\forall x_1, x_2 \in X$ vale $d_Y(f(x_1), f(x_2)) \leq Ld_X(x_1, x_2)$.

Definizione 2.9 (Locale lipschitzianità). Siano (X, d_X) e (Y, d_Y) due spazi metrici e sia $\Omega \subseteq X$ un aperto: una funzione $F : \Omega \rightarrow Y$ si dice *localmente lipschitziana* (su Ω) se per ogni $x_0 \in \Omega$ esistono due costanti $L > 0$ e $r > 0$ tali che

$$d_Y(F(x), F(y)) \leq Ld_X(x, y) \quad \forall x, y \in B(x_0, r) \cap \Omega .$$

Osservazione 2.20. Una funzione lipschitziana è in particolare uniformemente continua: infatti se L è una costante di Lipschitz per la mappa $f : X \rightarrow Y$, dato $\varepsilon > 0$ possiamo scegliere $\delta = \frac{\varepsilon}{L}$ ottenendo che se $d_X(x_1, x_2) < \delta$, allora $d_Y(f(x_1), f(x_2)) \leq Ld_X(x_1, x_2) < L\delta = \varepsilon$. Del tutto analogamente si dimostra che una funzione localmente lipschitziana è in particolare continua.

Osservazione 2.21. Data $F : (X, d) \rightarrow \mathbb{R}$, sono equivalenti:

- i. F è L -Lipschitz;
- ii. $F(x) - F(y) \leq Ld(x, y) \quad \forall x, y \in X$.

Infatti chiaramente $i \implies ii$; viceversa, se $\forall x, y$ vale $F(x) - F(y) \leq Ld(x, y)$, allora in particolare $F(y) - F(x) \leq Ld(y, x) = Ld(x, y)$ e dunque $|F(x) - F(y)| \leq Ld(x, y)$, ossia F è L -Lipschitz.

Proposizione 2.22

Dati (X, d_X) e (Y, d_Y) spazi metrici e $F : X \rightarrow Y$ localmente lipschitziana, se X è compatto allora F è globalmente lipschitziana.

Dimostrazione. Supponiamo per assurdo che f non sia lipschitziana: dunque per ogni $n \in \mathbb{N}$ esistono $x_n, y_n \in X$ tali che

$$d_Y(f(x_n), f(y_n)) > nd_X(x_n, y_n) .$$

Date quindi le successioni $\{x_n\}$ e $\{y_n\}$, poiché X è compatto e quindi compatto per successioni, possiamo estrarre una sottosuccessione x_{n_k} tale che $x_{n_k} \rightarrow \bar{x}$. Ora, poiché f è continua e X è compatto, anche $f(X)$ è compatto, e quindi limitato; dunque $d_Y(f(x_n), f(y_n)) < K$ per una certa costante K . Da questo e dalla disuguaglianza precedente, segue che

$$d_X(x_{n_k}, y_{n_k}) \rightarrow 0 \quad \text{quindi anche} \quad y_{n_k} \rightarrow \bar{x} .$$

Ora, per locale lipschitzianità, esiste una costante L e $\delta > 0$ tali che

$$d_Y(f(x), f(y)) \leq Ld_X(x, y) \quad \forall x, y \in B_\delta(\bar{x}) .$$

Poiché $x_{n_k}, y_{n_k} \in B_\delta(\bar{x})$ definitivamente, vale

$$Ld_X(x_{n_k}, y_{n_k}) > d_Y(f(x_{n_k}), f(y_{n_k})) > n_k d_X(x_{n_k}, y_{n_k}) . \quad \neq$$

□

Definizione 2.10 (Contrazione). Dato uno spazio metrico (X, d) , una funzione $T : X \rightarrow X$ si dice *contrazione* se $\exists L < 1$ tale che $\forall x, y \in X$ vale $d(T(x), T(y)) \leq Ld(x, y)$.

Osservazione 2.23. Segue dalla definizione che una contrazione è in particolare una funzione lipschitziana e dunque continua.

Teorema 2.24: Teorema delle contrazioni (Banach-Caccioppoli)

Dato (X, d) uno spazio metrico completo e $T : X \rightarrow X$ una contrazione, allora T ha un unico punto fisso, ossia:

$$\exists! \tilde{x} \in X \quad \text{tale che} \quad T(\tilde{x}) = \tilde{x} .$$

Dimostrazione. Sia $x_0 \in X$: definiamo per ricorsione la successione $x_{n+1} = T(x_n) = T^n(x_0)$. Mostriamo che la successione x_n è di Cauchy. Per definizione di T e x_n , vale:

$$d(x_n, x_{n+1}) = d(T(x_{n-1}), T(x_n)) \leq Ld(x_{n-1}, x_n) \leq \dots \leq L^n d(x_0, x_1)$$

Ora siano $n, m \in \mathbb{N}$ con $n \geq m$. Per la disuguaglianza appena ottenuta e la disuguaglianza triangolare, abbiamo:

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_{n-1}) + d(x_{n-1}, x_m) \leq \sum_{i=m}^{n-1} d(x_i, x_{i+1}) \leq \\ &\leq d(x_0, x_1) \sum_{i=m}^{n-1} L^i = L^m d(x_0, x_1) \sum_{i=0}^{n-m-1} L^i \end{aligned}$$

Per $n \rightarrow +\infty$, l'ultima è una serie geometrica che converge perché il termine generale è compreso tra 0 e 1, quindi

$$d(x_n, x_m) \leq d(x_0, x_1) \frac{L^m}{1-L} \rightarrow 0 \quad \text{per } m \rightarrow +\infty$$

ossia la successione x_n è di Cauchy. Poiché X è completo, esiste

$$\tilde{x} = \lim_{n \rightarrow +\infty} x_n$$

Poiché T è continua, vale

$$T(\tilde{x}) = \lim_{n \rightarrow +\infty} T(x_n) = \lim_{n \rightarrow +\infty} x_{n+1} = \tilde{x}$$

quindi \tilde{x} è un punto fisso di T .

Sia per assurdo x' un altro punto fisso di T , $T(x') = x'$. Allora

$$d(\tilde{x}, x') = d(T(\tilde{x}), T(x')) \leq Ld(\tilde{x}, x') \stackrel{L < 1}{<} d(\tilde{x}, x') \quad \neq$$

□

2.1.1 Equivalenze delle norme su \mathbb{R}^d

Osservazione 2.25. Ogni norma su \mathbb{R}^d , vista come funzione $(\mathbb{R}^d, \|\cdot\|) \rightarrow (\mathbb{R}^d, \|\cdot\|_1)$ è Lipschitziana, e quindi continua.

Dimostrazione. Sia $M = \max\{\|e_i\|, i = 1, \dots, n\}$. Vale

$$\| \|x\| - \|y\| \| \leq \|x - y\| = \left\| \sum_{i=1}^d (x_i - y_i) e_i \right\| \leq \sum_{i=1}^d |x_i - y_i| \|e_i\| \leq d_1(x, y) M$$

Quindi $\|\cdot\|$ è M -Lipschitziana rispetto a $\|\cdot\|_1$. □

Proposizione 2.26

Tutte le norme su \mathbb{R}^d sono topologicamente equivalenti.

Dimostrazione. Mostriamo che data una norma $\|\cdot\|$, esistono delle costanti α, β tali che

$$\alpha\|x\|_2 \leq \|x\| \leq \beta\|x\|_2 .$$

Consideriamo l'ipersfera unitaria $S^{d-1} = \{x \in \mathbb{R}^d : \|x\|_2 = 1\}$: questo è un insieme chiuso e limitato, quindi per Heine-Borel è compatto.

Allora, per il Teorema di Weierstrass la norma $\|\cdot\|$ ammette minimo e massimo su S^{d-1} : siano rispettivamente α e β . Ora, dato $x \in \mathbb{R}^d \setminus \{0\}$, si ha $\frac{x}{\|x\|_2} \in S^{d-1}$, quindi

$$\alpha \leq \left\| \frac{x}{\|x\|_2} \right\| \leq \beta \quad \text{cioè} \quad \alpha\|x\|_2 \leq \|x\| \leq \beta\|x\|_2 .$$

□

2.1.2 Distanza di un punto da un insieme

Osservazione 2.27. Dato (X, d) spazio metrico e fissato $p \in X$, la funzione

$$x \mapsto d(p, x)$$

è una funzione continua.

Infatti, $\forall x, y \in X$, vale

$$d(x, p) \leq d(x, y) + d(y, p) \iff d(x, p) - d(y, p) \leq d(x, y)$$

Quindi tale funzione è 1-Lipschitz e dunque continua.

Definizione 2.11 (Distanza da un insieme). Dato uno spazio metrico (X, d) , la *distanza dall'insieme* A , con $A \subseteq X$, di un punto $x \in X$ è

$$d(x, A) = \inf_{a \in A} d(x, a)$$

Proposizione 2.28

Dato (X, d) spazio metrico e $A \subseteq X$, valgono le seguenti:

- i. $d(\cdot, A)$ è 1-Lipschitz, e quindi continua;
- ii. $d(x, A) = 0 \iff x \in \bar{A}$.

Dimostrazione. i. Dato $x \in X$, vale

$$d(x, A) \leq d(x, a) \leq d(x, y) + d(y, a) \quad \forall a \in A, y \in X,$$

perciò, passando all'inf su $a \in A$, otteniamo

$$d(x, A) \leq d(x, y) + \inf_{a \in A} d(y, a) = d(x, y) + d(y, A) \implies d(x, A) - d(y, A) \leq d(x, y)$$

ovvero $d(\cdot, A)$ è 1-Lipschitz.

ii. Vale:

$$d(x, A) = \inf_{a \in A} d(x, a) = 0 \iff \forall \delta > 0 \exists y \in A : d(x, y) < \delta \iff x \in \bar{A} .$$

□

Osservazione 2.29. Ogni insieme chiuso è luogo di zeri di una qualche funzione continua: basta infatti prendere la funzione distanza da quel chiuso.

Esercizio 1. Sia $A \subseteq \mathbb{R}^d$ aperto e limitato. Allora A contiene una palla di raggio massimo.

Dimostrazione. Definiamo la funzione

$$\phi : x \mapsto d(x, A^c)$$

che è continua ed è positiva su A .

Ora \bar{A} è chiuso e limitato, perciò compatto, e quindi ϕ ammette massimo su \bar{A} . Poniamo allora $R = \max \phi = \phi(\bar{x})$ per qualche $\bar{x} \in \bar{A}$, e mostriamo che $B_R(\bar{x})$ è una palla di raggio massimo. Infatti osserviamo che, dato $x \in A$, se $r \leq d(x, A^c)$, allora $B_r(x) \subseteq A$, mentre se $r > d(x, A^c)$, si ha $B_r(x) \cap A^c \neq \emptyset$. Infine, osserviamo che $\bar{x} \in A$: altrimenti avremmo $\bar{x} \in \bar{A} \setminus A$ e pertanto $\bar{x} \in A^c$, dunque $\phi(\bar{x}) = d(\bar{x}, A^c) = 0$, mentre A è aperto e certamente contiene una palla di raggio positivo. \square

2.2 Limiti inferiori e superiori e semicontinuità

I limiti sono un vizio di chi se lo può permettere... noi non ce lo possiamo permettere.

Aldo Pratelli

Definizione 2.12 (Limite inferiore e superiore). Data una successione $\{x_n\}_{n \in \mathbb{N}}$ a valori in \mathbb{R} definiamo il *limite inferiore* o *minimo limite* di x_n per $n \rightarrow +\infty$ come

$$\liminf_{n \rightarrow +\infty} x_n := \lim_{j \rightarrow +\infty} \inf_{n \geq j} x_n = \sup_{j \in \mathbb{N}} \inf_{n \geq j} x_n$$

Definiamo il *limite superiore* o *massimo limite* di x_n come

$$\limsup_{n \rightarrow +\infty} x_n := \lim_{j \rightarrow +\infty} \sup_{n \geq j} x_n = \inf_{j \in \mathbb{N}} \sup_{n \geq j} x_n .$$

Indichiamo anche $\underline{\lim}$ il limite inferiore e $\overline{\lim}$ il limite superiore.

Osserviamo che l'uguaglianza tra $\lim_j \inf_n x_n$ e $\sup_j \inf_n x_n$ è dovuta al fatto che la successione $\{\inf_{n \geq j} x_n\}_{j \in \mathbb{N}}$ è monotona crescente, e dunque converge al sup (analoghi ragionamenti valgono per il lim sup). Pertanto questi oggetti hanno la proprietà di esistere sempre, in quanto applicazioni alternate di inf e sup di sottoinsiemi di \mathbb{R} non vuoti.

Proposizione 2.30

Data una successione $\{x_n\}$ a valori in \mathbb{R} , esistono due sottosuccessioni $\{x_{n_k}\}_{k \in \mathbb{N}}$ e $\{x_{n'_k}\}_{k \in \mathbb{N}}$ tali che

$$\lim_{k \rightarrow +\infty} x_{n_k} = \liminf_{n \rightarrow +\infty} x_n \quad \text{e} \quad \lim_{k \rightarrow +\infty} x_{n'_k} = \limsup_{n \rightarrow +\infty} x_n .$$

Dimostrazione. Mostriamo che esiste una sottosuccessione x_{n_k} che converge al lim inf. Sia $\ell = \liminf x_n$, allora per definizione $\forall \varepsilon > 0 \exists \bar{j}$ tale che $\forall j > \bar{j}$ vale $|\inf_{m \geq j} x_m - \ell| < \varepsilon$. Inoltre, fissato j , $\exists n \geq j$ tale che $|x_n - \inf_{m \geq j} x_m| < \varepsilon$, per cui complessivamente $|x_n - \ell| < 2\varepsilon$. Abbiamo dunque dimostrato che $\forall \varepsilon > 0 \exists n$ tale che $|x_n - \ell| < 2\varepsilon$, dunque possiamo costruire la nostra successione per ricorsione:

$$\begin{aligned} y_0 &= x_{n_0} & \text{con} & \quad |x_{n_0} - \ell| < 2 & \quad (\varepsilon = 1) \\ y_1 &= x_{n_1} & \text{con} & \quad |x_{n_1} - \ell| < 1 & \quad (\varepsilon = \frac{1}{2}) \\ & \vdots & & & \\ \forall k & \quad y_k = x_{n_k} & \text{con} & \quad |x_{n_k} - \ell| < 2^{1-k} & \quad (\varepsilon = 2^{-k}) \end{aligned}$$

Come possiamo assicurarci che la costruzione appena data identifichi davvero una sottosuccessione (ossia che la mappa $k \mapsto n_k$ sia strettamente crescente)? In effetti non lo è necessariamente,

ma questo non comporta un vero problema: al passaggio $(k+1)$ -esimo della costruzione appena data sfruttiamo il fatto che $\exists n \geq j$ tale che $|x_n - \inf_{m \geq j} x_m| < \varepsilon$ per poi porre $n_{k+1} = n$; per aggiustare il ragionamento consideriamo $\bar{n} = \max\{j, n_k + 1\}$: a questo punto $\exists n \geq \bar{n}$ tale che $|x_n - \inf_{m \geq \bar{n}} x_m| < \varepsilon$, per cui ponendo $n_{k+1} = n$ otteniamo $n_{k+1} = n \geq \bar{n} > n_k$. \square

Osservazione 2.31. Ripetiamo il primo passaggio della dimostrazione:

$\forall \varepsilon > 0 \exists \bar{j}$ tale che $\forall j > \bar{j} \mid \inf_{n \geq j} x_n - \ell \mid < \varepsilon$ dove $\ell = \liminf x_n$. In particolare, $\inf_{n \geq j} x_n > \ell - \varepsilon$, cioè $\forall \varepsilon > 0 \exists \bar{n}$ tale che $\forall n \geq \bar{n} x_n > \ell - \varepsilon$. Analoghe considerazioni valgono per il \limsup .

Se quindi $\{y_n\}$ è una sottosuccessione di $\{x_n\}$ e $y_n \xrightarrow{n \rightarrow +\infty} \ell$, per quanto appena osservato

$$\liminf_{n \rightarrow +\infty} x_n \leq \ell \leq \limsup_{n \rightarrow +\infty} x_n .$$

Osservazione 2.32. I limiti inferiore e superiore generalizzano il limite; in particolare caratterizzano la sua esistenza

$$\lim_{n \rightarrow +\infty} x_n = \ell \iff \liminf_{n \rightarrow +\infty} x_n = \limsup_{n \rightarrow +\infty} x_n = \ell .$$

Vediamo di definire anche \liminf e \limsup per funzioni.

Definizione 2.13 (Limite inferiore e superiore). Siano (X, d) uno spazio metrico, $\bar{x} \in X$ e $f : X \rightarrow \mathbb{R}$ una funzione. Definiamo il *limite inferiore* di f per $x \rightarrow \bar{x}$ come

$$\liminf_{x \rightarrow \bar{x}} f(x) := \lim_{\delta \rightarrow 0} \inf_{B(\bar{x}, \delta)} f = \sup_{\delta} \inf_{B(\bar{x}, \delta)} f ;$$

analogamente definiamo il *limite superiore* di f per $x \rightarrow \bar{x}$ come

$$\limsup_{x \rightarrow \bar{x}} f := \lim_{\delta \rightarrow 0} \sup_{B(\bar{x}, \delta)} f = \inf_{\delta} \sup_{B(\bar{x}, \delta)} f .$$

Nel caso in cui $(V, \|\cdot\|)$ sia uno spazio normato, definiamo il *limite inferiore* di f per $\|x\| \rightarrow +\infty$ come

$$\liminf_{\|x\| \rightarrow +\infty} f(x) := \lim_{R \rightarrow +\infty} \inf_{{}^c B(y, R)} f = \sup_R \inf_{{}^c B(y, R)} f$$

con $y \in V$ qualsiasi; allo stesso modo

$$\limsup_{\|x\| \rightarrow +\infty} f(x) := \lim_{R \rightarrow +\infty} \sup_{{}^c B(y, R)} f = \inf_R \sup_{{}^c B(y, R)} f .$$

Osservazione 2.33. Sia $f : \mathbb{R}^d \rightarrow \mathbb{R}$ una funzione continua: se

$$\liminf_{|x| \rightarrow +\infty} f(x) = \ell \in \mathbb{R}$$

e $\exists \bar{x}$ tale che $f(\bar{x}) \leq \ell$, allora f ammette un minimo globale.

Dimostrazione. Abbiamo dalle proprietà del \liminf che $\forall \varepsilon > 0 \exists R > 0$ tale che per $|x| > R$ si ha $f(x) > \ell - \varepsilon$.

Se $f(\bar{x}) < \ell$ per quanto detto $\exists R$ tale che $\forall x$ con $|x| > R$ vale $f(x) > f(\bar{x})$; inoltre dato che $\overline{B}(0, R)$ è compatto non vuoto (contiene infatti \bar{x}), $\exists x_{\min}$ punto di minimo di f su $\overline{B}_R(0)$, e dal momento che $\forall x \in \overline{B}(x, r)^c f(x) > f(\bar{x}) \geq f(x_{\min})$, x_{\min} è punto di minimo per f , che quindi ammette minimo globale.

Se invece $f(\bar{x}) = \ell$ si danno due possibilità: o $\forall x \in \mathbb{R}^d f(\bar{x}) \leq f(x)$, e dunque $f(\bar{x})$ è il minimo di f , oppure $\exists \tilde{x} \in \mathbb{R}^d f(\tilde{x}) < f(\bar{x}) = \ell$, e ci riconduciamo al caso precedente. \square

Definizione 2.14 (Semicontinuità). Siano (X, d_X) uno spazio metrico, $x \in X$: una funzione $f : X \rightarrow \mathbb{R}$ si dice *semicontinua inferiormente* (s.c.i.) in x se $\forall \varepsilon > 0 \exists \delta > 0$ tale che $\forall y \in B(x, \delta)$ vale $f(x) - \varepsilon < f(y)$; f si dice *semicontinua superiormente* (s.c.s.) in x se $\forall \varepsilon > 0 \exists \delta > 0$ tale che $\forall y \in B(x, \delta)$ vale $f(y) < f(x) + \varepsilon$.

Proposizione 2.34

Data una funzione $f : X \rightarrow \mathbb{R}$ con X spazio metrico e $\bar{x} \in X$, sono equivalenti:

i. f è semicontinua inferiormente in \bar{x} ;

ii. $\liminf_{x \rightarrow \bar{x}} f(x) \geq f(\bar{x})$.

Allo stesso modo sono equivalenti:

i. f è semicontinua superiormente in \bar{x} ;

ii. $\limsup_{x \rightarrow \bar{x}} f(x) \leq f(\bar{x})$.

Dimostrazione. Dimostriamo la prima equivalenza:

$i \implies ii$ Se f è semicontinua inferiormente, allora per definizione $\forall \varepsilon > 0 \exists \delta > 0$ tale che $\forall x \in B(\bar{x}, \delta)$ vale $f(\bar{x}) - \varepsilon < f(x)$; in particolare, $f(\bar{x}) - \varepsilon$ è un minorante di $\text{Im}f|_{B(\bar{x}, \delta)}$ e di conseguenza $\inf_{B(\bar{x}, \delta)} f \geq f(\bar{x}) - \varepsilon$; concludiamo osservando che

$$\inf_{B(\bar{x}, \delta)} f \geq f(\bar{x}) - \varepsilon \implies \liminf_{x \rightarrow \bar{x}} f(x) \geq f(\bar{x}) - \varepsilon \implies \liminf_{x \rightarrow \bar{x}} f(x) \geq f(\bar{x})$$

dove la prima implicazione vale perché il \liminf è il \sup degli \inf , mentre la seconda per arbitrarietà di ε .

$ii \implies i$ Fissiamo $\varepsilon > 0$, per ipotesi

$$\liminf_{x \rightarrow \bar{x}} f(x) = \sup_{\delta} \inf_{B(\bar{x}, \delta)} f \geq f(\bar{x}) > f(\bar{x}) - \varepsilon ;$$

in particolare $\exists \delta > 0$ tale che

$$\inf_{B(\bar{x}, \delta)} f > f(\bar{x}) - \varepsilon$$

e quindi $\forall x \in B(\bar{x}, \delta)$ vale $f(x) > f(\bar{x}) - \varepsilon$, cioè f è semicontinua inferiormente in \bar{x} . □

Vediamo a questo punto un teorema che generalizza il Teorema di Weierstraß, utilizzando la semicontinuità.

Teorema 2.35

Siano (X, d) spazio metrico, $K \subseteq X$ compatto per successioni e $f : K \rightarrow \mathbb{R}$ una funzione: allora, se f è semicontinua inferiormente ammette minimo globale, mentre se f è semicontinua superiormente ammette massimo globale.

Dimostrazione. Sia f semicontinua inferiormente. Per definizione di estremo inferiore, esiste una successione $\{x_n\} \subseteq K$ tale che $f(x_n) \rightarrow \inf f$. Ora, poiché K è compatto per successioni, esiste una sottosuccessione $\{\tilde{x}_n\}$ tale che $\tilde{x}_n \rightarrow \bar{x}$ per qualche $\bar{x} \in K$. Poiché f è semicontinua inferiormente, vale

$$f(\bar{x}) \leq \liminf_{n \rightarrow +\infty} f(\tilde{x}_n) = \lim_{n \rightarrow +\infty} f(\tilde{x}_n) = \inf f$$

quindi $f(\bar{x}) = \inf f = \min f$.

La dimostrazione è analoga nel caso in cui f sia semicontinua superiormente. □

2.3 Convessità e continuità

Definizione 2.15 (Insieme convesso). Dato uno spazio vettoriale V reale, un insieme $C \subseteq V$ si dice *convesso* se dati $x, y \in C$ vale che $tx + (1-t)y \in C \forall t \in (0, 1)$. Talvolta indichiamo $[x, y] := \{tx + (1-t)y \mid t \in [0, 1]\}$ l'*inviluppo convesso* di x e y .

Definizione 2.16 (Funzione convessa). Dato $(V, \|\cdot\|)$ uno spazio normato, $\Omega \subseteq V$ un insieme convesso. Diciamo che una funzione $f : \Omega \rightarrow \mathbb{R}$ è *convessa* se $\forall x, y \in \Omega, \forall t \in [0, 1]$ vale $f(tx + (1-t)y) \leq tf(x) + (1-t)f(y)$.

Lemma 2.36

Sia $(V, \|\cdot\|)$ uno spazio normato e sia $x_0 \in V$; data $f : \overline{B}_R(x_0) \rightarrow \mathbb{R}$ una funzione convessa, se $\exists M \in \mathbb{R}$ tale che $\forall x \in \overline{B}_R(x_0)$ vale $f(x) - f(x_0) \leq M$, allora

$$\forall x \in \overline{B}_R(x_0) \quad |f(x) - f(x_0)| \leq \frac{M}{R}|x - x_0| .$$

Dimostrazione. Consideriamo la funzione $g : \overline{B}_R(0) \rightarrow \mathbb{R}$ dove $g(y) = f(x_0 + y) - f(x_0)$, per cui $g(0) = 0, g(y) \leq M \forall y$ e g è una funzione convessa. Ponendo $x = x_0 + y$, la tesi diventa $|g(y)| \leq \frac{M}{R}|y|$. Proviamo dunque separatamente che

$$\text{I. } g(y) \leq \frac{M}{R}|y| \quad \text{e} \quad \text{II. } g(y) \geq -\frac{M}{R}|y| .$$

i. Definiamo il punto $\xi = R\frac{y}{|y|}$, per cui $y = t\xi + (1-t) \cdot 0$ con $t = \frac{|y|}{R}$, allora

$$g(y) \leq tg(\xi) + (1-t)g(0) \leq \frac{|y|}{R}M .$$

ii. Analogamente definiamo il punto $\xi = -R\frac{y}{|y|}$, per cui $0 = t\xi + (1-t)y$, da cui ricaviamo

$$t = \frac{|y|}{R+|y|} \text{ e } 1-t = \frac{R}{R+|y|} : \text{ ma allora poich\`e } g(0) = 0 \leq tg(\xi) + (1-t)g(y),$$

$$0 \leq \frac{|y|}{R+|y|}g(\xi) + \frac{R}{R+|y|}g(y) \implies g(y) \geq -\frac{|y|}{R}g(\xi) \geq -\frac{|y|}{R}M .$$

□

Teorema 2.37

Dato $(V, \|\cdot\|)$ spazio normato, siano $\Omega \subseteq V$ un insieme aperto e convesso e $f : \Omega \rightarrow \mathbb{R}$ una funzione convessa e localmente superiormente limitata (**LSL**), allora f è localmente lipschitziana e dunque continua su Ω .

Dimostrazione. Sia $x_0 \in \Omega$: dato che Ω è aperto $\exists r_1 > 0$ tale che $\overline{B}(x_0, r_1) \subseteq \Omega$, e dato che f è **LSL** $\exists r_2 > 0$ tale che $f(x) - f(x_0) \leq M \forall x \in \overline{B}(x_0, r_2)$ per qualche M . In particolare, per il Lemma vale

$$|f(x) - f(x_0)| \leq M \quad \forall x \in \overline{B}(x_0, r_2) .$$

Sia quindi $r = \min\{r_1, r_2\}$. Siano $x, x' \in \overline{B}(x_0, \frac{2}{3}r)$, per cui vale $|x - x'| \leq \frac{2}{3}r$. Ora vale che $\overline{B}(x', \frac{2}{3}r) \subseteq \overline{B}(x_0, r)$, quindi, per la scelta di r ,

$$|f(x') - f(x_0)| \leq M \quad \text{da cui} \quad f(x) - f(x') \leq M - f(x_0) - (f(x_0) - M) = 2M .$$

Ancora per il Lemma precedente, abbiamo $|f(x) - f(x')| \leq \frac{2M}{\frac{2}{3}r} |x - x'| = \frac{3M}{r} |x - x'|$. In conclusione, abbiamo ottenuto

$$|f(x) - f(x')| \leq \frac{3M}{r} |x - x'| \quad \forall x, x' \in B(x_0, \frac{r}{3}),$$

ossia f è localmente lipschitziana. \square

Proposizione 2.38: Disuguaglianza di Jensen

Dato $(V, \|\cdot\|)$, siano $\Omega \subseteq V$ un insieme aperto e convesso e $f : \Omega \rightarrow \mathbb{R}$, allora sono equivalenti:

i. f è convessa;

ii. $\forall n \in \mathbb{N}$ vale C_n : $\forall x_1, \dots, x_n \in \Omega, \forall \lambda_1, \dots, \lambda_n \geq 0$ tali che $\sum_{i=1}^n \lambda_i = 1$ vale

$$f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i) .$$

Dimostrazione. L'implicazione $(ii \implies i)$ è immediata in quanto la definizione di convessità è esattamente la proprietà C_2 . Mostriamo quindi $(i \implies ii)$ per induzione su n . Il passo base è ancora offerto da C_2 . Per il passo induttivo supponiamo vera C_n e siano $x_0, \dots, x_n \in \Omega, \lambda_0, \dots, \lambda_n \geq 0$ tali che $\sum_{i=0}^n \lambda_i = 1$: osserviamo che

$$\sum_{i=0}^n \lambda_i x_i = \lambda_0 x_0 + \sum_{i=1}^n \lambda_i x_i = \lambda_0 x_0 + (1 - \lambda_0) \left(\sum_{i=1}^n \frac{\lambda_i}{1 - \lambda_0} x_i \right) ,$$

$$\text{dove } \sum_{i=1}^n \frac{\lambda_i}{1 - \lambda_0} = 1 \text{ dato che } \sum_{i=1}^n \lambda_i = 1 - \lambda_0;$$

pertanto otteniamo che

$$\begin{aligned} f\left(\sum_{i=0}^n \lambda_i x_i\right) &= f\left(\lambda_0 x_0 + (1 - \lambda_0) \sum_{i=1}^n \left(\frac{\lambda_i}{1 - \lambda_0} x_i\right)\right) \leq \lambda_0 f(x_0) + (1 - \lambda_0) f\left(\sum_{i=1}^n \frac{\lambda_i}{1 - \lambda_0} x_i\right) \leq \\ &\leq \lambda_0 f(x_0) + \cancel{(1 - \lambda_0)} \left(\sum_{i=1}^n \frac{\lambda_i}{\cancel{1 - \lambda_0}} f(x_i)\right) = \sum_{i=0}^n \lambda_i f(x_i) , \end{aligned}$$

dove la prima disuguaglianza è data dal passo base C_2 , mentre la seconda è data dal passo induttivo C_n . \square

Definizione 2.17 (Inviluppo convesso). Dati $x_1, \dots, x_n \in \mathbb{R}^d$, chiamiamo *inviluppo convesso* dei punti x_1, \dots, x_n l'insieme delle loro combinazioni convesse

$$\Delta(x_1, \dots, x_n) = \left\{ \sum_{i=1}^n \lambda_i x_i \mid \lambda_i \geq 0 \ \forall i, \sum_{i=1}^n \lambda_i = 1 \right\} .$$

Osservazione 2.39. Data $f : \Omega \rightarrow \mathbb{R}$ una funzione convessa con Ω aperto convesso, allora dati $x_1, \dots, x_n \in \Omega$ vale che

$$\max_{\Delta(x_1, \dots, x_n)} f = \max \{f(x_1), \dots, f(x_n)\} .$$

Supponiamo infatti per assurdo che $\exists \bar{x} \in \Delta(x_1, \dots, x_n)$ con $\bar{x} \neq x_i \forall i$ tale che $f(\bar{x}) > f(x_i) \forall i$; siano $\lambda_1, \dots, \lambda_n \geq 0$ tali che $\bar{x} = \lambda_1 x_1 + \dots + \lambda_n x_n$, con $\lambda_i \geq 0 \forall i$, $\lambda_1 + \dots + \lambda_n = 1$ e $\lambda_i \neq 1 \forall i$ (altrimenti in corrispondenza di tale i avremmo $\bar{x} = x_i$): allora usando la disuguaglianza di Jensen

$$f(\bar{x}) = f\left(\sum_{i=1}^n \lambda_i x_i\right) \leq \sum_{i=1}^n \lambda_i f(x_i) < \left(\sum_{i=1}^n \lambda_i\right) f(\bar{x}) = f(\bar{x}) . \quad \text{!}$$

Osservazione 2.40. Se Ω è un aperto convesso in \mathbb{R}^d e $f : \Omega \rightarrow \mathbb{R}$ è una funzione convessa, allora f è automaticamente **LSL**, cioè ogni punto P ha un intorno su cui f è limitata superiormente. Di conseguenza per $V = \mathbb{R}^d$ (in generale per V \mathbb{R} -spazio normato di dimensione finita) l'ipotesi di locale superiore limitatezza del Teorema è superflua.

Infatti ogni punto ha come intorno un n -simpleso dato dall'involuppo convesso di $n + 1$ punti in posizione generale, e dunque per l'osservazione precedente f è **LSL**.

Esercizio 2. Sia $\Omega \subseteq \mathbb{R}^d$ aperto convesso: dimostrare che una funzione $f : \Omega \rightarrow \mathbb{R}$ è convessa se e solo se l'*epigrafico* di f , $\text{epi}(f) := \{(x, y) \in \Omega \times \mathbb{R} \mid y \geq f(x)\}$, è un insieme convesso.

Corollario 2.41

Sia $\Omega \subseteq \mathbb{R}^d$ e sia $\{f_\alpha\}_{\alpha \in \mathcal{A}}$ una famiglia di funzioni tale che $f_\alpha : \Omega \rightarrow \mathbb{R}$ è una funzione convessa $\forall \alpha \in \mathcal{A}$. Allora la funzione

$$\bar{f}(x) := \sup_{\alpha \in \mathcal{A}} f_\alpha(x)$$

è convessa (sull'insieme $\{x \in \Omega \mid \bar{f}(x) < +\infty\}$).

2.4 Funzioni omogenee

Definizione 2.18 (Funzione omogenea). Una funzione $f : \mathbb{R}^d \setminus \{0\} \rightarrow \mathbb{R}$ si dice γ -omogenea (o γ -positivamente omogenea) se

$$f(\lambda x) = \lambda^\gamma f(x) \quad \forall \lambda \in \mathbb{R}^+, \forall x \in \mathbb{R}^d \setminus \{0\}.$$

Proposizione 2.42

Se $f : \mathbb{R}^d \setminus \{0\} \rightarrow \mathbb{R}$ è continua e γ -(positivamente) omogenea, allora:

- i. se $\gamma > 0$, f è estendibile per continuità in 0;
- ii. se $\gamma = 0$, f è costante sulle rette oppure non ammette limite per $|x| \rightarrow 0$;
- iii. se $\gamma < 0$, f non è limitata in un intorno di 0, a meno che $f \equiv 0$.

Dimostrazione. Consideriamo la sfera unitaria $S^{d-1} = \{x \in \mathbb{R}^d : |x| = 1\}$: S^{d-1} è un insieme compatto. Inoltre f è continua, e dunque $f|_{S^{d-1}}$ è limitata, ossia $\exists M > 0$ tale che $\forall \xi \in S^{d-1}$ si ha $|f(\xi)| \leq M$.

Dato ora $x \in \mathbb{R}^d \setminus \{0\}$, possiamo scrivere $f(x) = f(|x| \frac{x}{|x|}) = |x|^\gamma f(\frac{x}{|x|})$, pertanto:

- i. se $\gamma > 0$, $|f(x)| \leq |x|^\gamma M \xrightarrow{|x| \rightarrow 0} 0$;
- ii. se $\gamma = 0$ e f non è costante allora non esiste limite perché i limiti direzionali sono diversi (esistono infatti due rette su cui f assume valori distinti);
- iii. se $\gamma < 0$ e $\exists \xi : f(\xi) \neq 0$, $f(t\xi)$ non è limitata in un intorno di 0.

□

Capitolo 3

Calcolo differenziale

Il calcolo differenziale nelle funzioni $f : \mathbb{R} \rightarrow \mathbb{R}$ si concretizzava nello studio delle derivate e consisteva principalmente di due aspetti: da un lato la derivata definita come limite del rapporto incrementale

$$f'(x) = \lim_{h \rightarrow 0} \frac{f(x+h) - f(x)}{h} ,$$

dall'altro, grazie al teorema di Taylor, alla possibilità di approssimare le curve con funzioni lineari:

$$f(x) = f(\bar{x}) + f'(\bar{x})(x - \bar{x}) + o(|x - \bar{x}|).$$

Era in effetti questo doppio aspetto delle derivate a renderle così interessanti: il primo ne permetteva il calcolo, il secondo le applicazioni.

Tentiamo anche in caso multidimensionale di riproporre entrambe le questioni.

Definizione 3.1 (Derivata direzionale). Data $f : \mathbb{R}^d \rightarrow \mathbb{R}$, dati $x_0, v \in \mathbb{R}^d$, la *derivata direzionale* di f in x_0 in direzione v è

$$\frac{\partial f}{\partial v}(x_0) := \lim_{\varepsilon \rightarrow 0} \frac{f(x_0 + \varepsilon v) - f(x_0)}{\varepsilon} ,$$

e la denotiamo anche con $\partial_v f(x_0)$.

Osservazione 3.1. Consideriamo la mappa $v \mapsto \partial_v f(x_0)$. Chiaramente $\partial_0 f(x_0) = 0$, inoltre è un facile calcolo mostrare che $\partial_{\lambda v} f(x_0) = \lambda \partial_v f(x_0)$, ovvero questa mappa è omogenea. In generale però non è additiva come mostra il seguente esempio:

$$f(x, y) = \begin{cases} \frac{x^2 y}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases} \quad \text{per cui}$$

$$\frac{\partial f}{\partial(1,0)}(0,0) = 0 \quad \text{e} \quad \frac{\partial f}{\partial(0,1)}(0,0) = 0 \quad \text{mentre} \quad \frac{\partial f}{\partial(1,1)}(0,0) = \lim_{\varepsilon \rightarrow 0} \frac{f(\varepsilon, \varepsilon) - f(0,0)}{\varepsilon} = \frac{1}{2} .$$

Sappiamo però che $f(x_0 + \varepsilon v) = f(x_0) + \varepsilon \partial_v f(x_0) + o(\varepsilon|v|)$.

Pertanto se la funzione $v \mapsto \partial_v f(x_0)$ fosse lineare, sarebbe possibile rappresentarla con una matrice riga $M \in M(1, n, \mathbb{R})$, tale che $Mv = \partial_v f(x_0)$, o ancora con un vettore $\omega \in \mathbb{R}^d$ tale che $\omega \cdot v = \partial_v f(x_0)$, e in tal caso $\varepsilon \partial_v f(x_0) = \varepsilon \omega \cdot v = \omega \cdot (\varepsilon v)$.

Definizione 3.2 (Differenziabilità). Dati una funzione $f : \mathbb{R}^d \rightarrow \mathbb{R}$ e un punto $x_0 \in \mathbb{R}^d$, si dice che f è *differenziabile* in x_0 se esiste un operatore lineare $Df(x_0) : \mathbb{R}^d \rightarrow \mathbb{R}$ tale che $f(x) = f(x_0) + Df(x_0)(x - x_0) + o(|x - x_0|)$.

Se f è differenziabile, è anche continua in quanto evidentemente le funzioni lineari sono continue; inoltre se f è differenziabile esistono tutte le derivate direzionali e $v \mapsto \partial_v f(x_0)$ è lineare: in particolare $\partial_v f(x_0) = Df(x_0)(v)$ dato che

$$f(x_0 + \varepsilon v) = f(x_0) + Df(x_0)(\varepsilon v) + o(\varepsilon|v|) = f(x_0) + \varepsilon Df(x_0)(v) + o(\varepsilon) .$$

Se esistono tutte le derivate direzionali di f , non è detto che f sia differenziabile, come mostra l'esempio

$$f : \mathbb{R}^2 \rightarrow \mathbb{R} \quad (x, y) \mapsto \begin{cases} 1 & \text{se } y = x^2 \\ 0 & \text{altrimenti} \end{cases} :$$

in questo caso tutte le derivate direzionali in $(0, 0)$ esistono, ma f non è neppure continua.

Proposizione 3.2

Data una funzione $f : \mathbb{R}^d \rightarrow \mathbb{R}$, e v_1, \dots, v_d direzioni linearmente indipendenti, se f ammette le d derivate direzionali in un intorno di x_0 ed esse sono continue in x_0 , allora f è differenziabile in x_0 .

Dimostrazione. Vediamo il caso $d = 2$:

$$\frac{f(x_0 + \varepsilon(\lambda v + \mu w)) - f(x_0)}{\varepsilon} = \frac{f(x_0 + \varepsilon(\lambda v + \mu w)) - f(x_0 + \varepsilon \lambda v) + f(x_0 + \varepsilon \lambda v) - f(x_0)}{\varepsilon} ;$$

per il teorema di Taylor con resto di Lagrange

$$\begin{aligned} f(x_0 + \varepsilon \lambda v) &= f(x_0) + \varepsilon \partial_{\lambda v} f(x_0 + t \lambda v) \quad \text{con } t \in (0, \varepsilon) \\ &= f(x_0) + \varepsilon \lambda \partial_v f(x_0 + t \lambda v) \\ f(x_0 + \varepsilon \lambda v + \varepsilon \mu w) &= f(x_0 + \varepsilon \lambda v) + \varepsilon \partial_{\mu w} f(x_0 + \varepsilon \lambda v + s \mu w) \quad \text{con } s \in (0, \varepsilon) \\ &= f(x_0 + \varepsilon \lambda v) + \varepsilon \mu \partial_w f(x_0 + \varepsilon \lambda v + s \mu w) . \end{aligned}$$

A questo punto

$$\begin{aligned} \frac{f(x_0 + \varepsilon(\lambda v + \mu w)) - f(x_0)}{\varepsilon} &= \frac{\varepsilon \mu \partial_w f(x_0 + \varepsilon \lambda v + s \mu w) + \varepsilon \lambda \partial_v f(x_0 + t \lambda v)}{\varepsilon} = \\ &= \mu \partial_w f(x_0 + \varepsilon \lambda v + s \mu w) + \lambda \partial_v f(x_0 + t \lambda v) \end{aligned}$$

da cui passando al limite per $\varepsilon \rightarrow 0$

$$\frac{\partial f}{\partial(\lambda v + \mu w)}(x_0) = \lambda \frac{\partial f}{\partial v}(x_0) + \mu \frac{\partial f}{\partial w}(x_0) .$$

Nel caso generale si ha

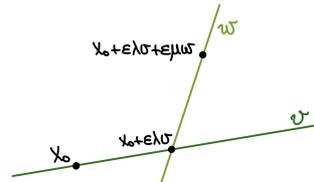
$$\frac{\partial f}{\partial(\lambda_1 v_1 + \dots + \lambda_d v_d)}(x_0) = \lambda_1 \frac{\partial f}{\partial v_1}(x_0) + \dots + \lambda_d \frac{\partial f}{\partial v_d}(x_0)$$

(la dimostrazione è analoga, con qualche addendo in più).

Pertanto l'esistenza di d derivate direzionali localmente continue in direzioni linearmente indipendenti assicura (per linearità) l'esistenza di tutte le derivate direzionali, e di conseguenza l'operatore $Df(x_0) : \mathbb{R}^d \rightarrow \mathbb{R}$ tale che $Df(x_0)(v) = \partial_v f(x_0)$ è ben definito e lineare. Per concludere:

$$\begin{aligned} f(x) &= f(x_0 + (x - x_0)) = f(x_0 + |x - x_0|v) = \quad \text{con } v = \frac{x - x_0}{|x - x_0|} \\ &= f(x_0) + |x - x_0| \frac{\partial f}{\partial v}(x_0) + o(|x - x_0|) = f(x_0) + |x - x_0| Df(x_0)(v) + o(|x - x_0|) = \\ &= f(x_0) + Df(x_0)(x - x_0) + o(|x - x_0|) \end{aligned}$$

che prova la differenziabilità di f . □



Definizione 3.3 (Derivata parziale). Le derivate direzionali nelle direzioni coordinate e_i si dicono *derivate parziali*, e si indicano classicamente con ∂_{x_i} .

Definizione 3.4 (Classe C^k). Sia $\Omega \subseteq \mathbb{R}^d$ un aperto; una funzione $f : \Omega \rightarrow \mathbb{R}$ si dice di *classe C^k* , con $k \in \mathbb{N}$, se f ammette tutte le derivate parziali sino all'ordine k ed esse sono continue in ogni punto di Ω .

Teorema 3.3: Teorema del differenziale totale

Data $f : \mathbb{R}^d \rightarrow \mathbb{R}$, se f è di classe C^1 allora f è differenziabile.

Dimostrazione. Segue direttamente dalla Proposizione precedente. \square

Supponiamo di avere una funzione $f : \mathbb{R}^d \rightarrow \mathbb{R}$ differenziabile in x_0 , e sia $v \in \mathbb{R}^d$ un vettore di coordinate v_1, \dots, v_d (vale a dire $v_i = v \cdot e_i$). Allora sfruttando la linearità del differenziale

$$\begin{aligned} Df(x_0)(v) &= Df(x_0)(v_1 e_1 + \dots + v_d e_d) = \sum_{i=1}^d v_i Df(x_0)(e_i) = \sum_{i=1}^d v_i \frac{\partial f}{\partial x_i}(x_0) = \\ &= v \cdot \begin{pmatrix} \partial_{x_1} f(x_0) \\ \vdots \\ \partial_{x_d} f(x_0) \end{pmatrix}. \end{aligned}$$

Questa semplice osservazione motiva la seguente definizione.

Definizione 3.5 (Gradiente). Data $f : \mathbb{R}^d \rightarrow \mathbb{R}$ differenziabile in x_0 , chiamiamo *gradiente* di f in x_0 il vettore

$$\nabla f(x_0) = \begin{pmatrix} \partial_{x_1} f(x_0) \\ \vdots \\ \partial_{x_d} f(x_0) \end{pmatrix},$$

dunque $Df(x_0)(v) = \nabla f(x_0) \cdot v$, ossia il gradiente è il vettore che rappresenta il differenziale nella base canonica.

Osservazione 3.4. Se consideriamo un vettore di norma unitaria, $|v| = 1$, vale

$$\left| \frac{\partial f}{\partial v}(x_0) \right| = |\nabla f(x_0) \cdot v| \leq |\nabla f(x_0)| |v| = |\nabla f(x_0)|$$

ossia le derivate direzionali valgono al più $|\nabla f(x_0)|$, e l'uguaglianza si ha solo per la direzione di $\nabla f(x_0)$ (se $\nabla f(x_0) \neq 0$).

Quindi il gradiente è la direzione di massima pendenza della funzione nel punto.

Proposizione 3.5

Data $f : \mathbb{R}^d \rightarrow \mathbb{R}$, se \bar{x} è punto di massimo o minimo locale per f e f è differenziabile in \bar{x} , allora $\nabla f(\bar{x}) = 0$.

Dimostrazione. Se fosse per assurdo $\nabla f(\bar{x}) \neq 0$, possiamo scrivere

$$\begin{aligned} f(\bar{x} + \varepsilon \nabla f(\bar{x})) &= f(\bar{x}) + \varepsilon \nabla f(\bar{x}) \cdot \nabla f(\bar{x}) + o(|\varepsilon \nabla f(\bar{x})|) = \\ &= f(\bar{x}) + \varepsilon |\nabla f(\bar{x})|^2 + o(\varepsilon) = \\ &= f(\bar{x}) + \varepsilon (|\nabla f(\bar{x})|^2 + o(1)) \end{aligned}$$

Allora per $\varepsilon > 0$ piccolo, $f(\bar{x} + \varepsilon \nabla f(\bar{x})) > f(\bar{x})$ e per $\varepsilon < 0$ piccolo, $f(\bar{x} + \varepsilon \nabla f(\bar{x})) < f(\bar{x})$, quindi \bar{x} non è punto massimo o minimo locale. ∇

Definizione 3.6 (Punto critico). Data una funzione $f : \mathbb{R}^d \rightarrow \mathbb{R}$, un *punto critico* di f è un punto \bar{x} tale che $\nabla f(\bar{x}) = 0$.

Possiamo generalizzare i discorsi fatti finora considerando funzioni vettoriali, ossia funzioni $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, che scriveremo in coordinate $f = (f_1, \dots, f_m)$, con $f_i : \mathbb{R}^n \rightarrow \mathbb{R}$.

Definizione 3.7 (Differenziabilità). Una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ si dice *differenziabile* in x_0 se esiste un'applicazione lineare $Df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tale che

$$f(x_0 + v) = f(x_0) + Df(x_0)(v) + o(|v|) .$$

Tale applicazione viene detta *differenziale* di f in x_0 .

Definizione 3.8 (Matrice di Jacobi). Data una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ differenziabile in x_0 , si definisce la sua *giacobiana* o *matrice di Jacobi* come la matrice associata al differenziale $Df(x_0)$ nelle basi canoniche di \mathbb{R}^n e \mathbb{R}^m , ossia

$$J_f(x_0) := \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \dots & \frac{\partial f_1}{\partial x_n}(x_0) \\ \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1}(x_0) & \dots & \frac{\partial f_m}{\partial x_n}(x_0) \end{pmatrix} = \begin{pmatrix} \nabla f_1(x_0)^\top \\ \vdots \\ \nabla f_m(x_0)^\top \end{pmatrix} .$$

In effetti, è indifferente considerare il gradiente di una funzione un vettore riga o un vettore colonna, poiché stiamo implicitamente identificando lo spazio \mathbb{R}^n con il suo duale $(\mathbb{R}^n)^*$ (i cui elementi sono effettivamente vettori riga); infatti, per il Teorema di rappresentazione di Riesz, il funzionale $Df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}$ può essere rappresentato con un vettore $v \in \mathbb{R}^n$ tale che $Df(x_0)(x) = \langle v, x \rangle$, e per la definizione di gradiente, tale vettore è proprio $\nabla f(x_0)$.

Teorema 3.6: Differenziazione della funzione composta

Date due funzioni $f : \mathbb{R}^m \rightarrow \mathbb{R}^n$ e $g : \mathbb{R}^n \rightarrow \mathbb{R}^k$ ed un punto $x \in \mathbb{R}^m$ tale che f è differenziabile in x e g è differenziabile in $f(x)$, allora $g \circ f$ è differenziabile in x e vale

$$D(g \circ f)(x) = Dg(f(x)) \circ Df(x) .$$

Dimostrazione. L'idea della dimostrazione è sostanzialmente la medesima del caso unidimensionale: sviluppare al primo ordine prima f e poi g .

$$\begin{aligned} g(f(x+v)) &= g(f_1(x+v), \dots, f_m(x+v)) = \\ &= g(f_1(x) + \nabla f_1(x) \cdot v + o(|v|), \dots, f_m(x) + \nabla f_m(x) \cdot v + o(|v|)) = \\ &= g(f(x) + Df(x)(v) + o(|v|)) = \end{aligned}$$

adesso ponendo $w := Df(x)(v) + o(|v|)$ troviamo

$$\begin{aligned} &= g(f(x) + w) = \\ &= g(f(x)) + Dg(f(x))(w) + o(|w|) = \\ &= g(f(x)) + Dg(f(x))(Df(x)(v) + o(|v|)) + o(|w|) = \\ &= g(f(x)) + Dg(f(x))(Df(x)(v)) + Dg(f(x))(o(|v|)) + o(|w|) = \\ &= g(f(x)) + (Dg(f(x)) \circ Df(x))(v) + Dg(f(x))(o(|v|)) + o(|w|) . \end{aligned}$$

Per concludere mostriamo che $Dg(f(x))(o(|v|)) = o(|v|)$ e che $o(|w|) = o(|v|)$. In generale, possiamo identificare un'applicazione lineare $L : \mathbb{R}^p \rightarrow \mathbb{R}^q$ con una matrice $M \in M(q, p, \mathbb{R})$ tale che $\forall x \in \mathbb{R}^p$ si ha $L(x) = Mx$ (per $Dg(f(x))$ e $Df(x)$ è proprio $J_g(f(x))$ e $J_f(x)$): mostriamo che in tal caso vale $|Mx| \leq k|x|$, con k una certa costante positiva che dipende da M ; ciò conclude in quanto:

- $|Dg(f(x))(o(|v|))| \leq k|o(|v|)| = o(|v|)$;
- $o(|w|) = o(|Df(x)(v) + o(|v|)|) = o(k|v| + o(|v|)) = o(|v|)$.

Per mostrare la disuguaglianza,

$$|Mx| = \left| \begin{pmatrix} M_1x \\ \vdots \\ M_qx \end{pmatrix} \right| = \sqrt{(M_1x)^2 + \dots + (M_qx)^2} \leq \sqrt{(|M_1|^2 + \dots + |M_q|^2)|x|^2} \leq k|x|$$

per $k := \sqrt{n} \max_i |M_i|$. □

Teorema 3.7: Teorema di Lagrange

Sia $f : U \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$ differenziabile e siano $x, y \in U$ tali che $[x, y] \subseteq U$. Allora esiste $\xi \in [x, y]$ tale che

$$f(x) - f(y) = Df(\xi)(x - y) .$$

Dimostrazione. Definiamo la funzione

$$\begin{aligned} \phi : \mathbb{R} &\rightarrow \mathbb{R} \\ t &\mapsto f(x + t(y - x)) \end{aligned}$$

che è differenziabile in quanto composizione di due funzioni differenziabili; allora vale

$$\phi'(t) = Df(x + t(y - x))(y - x) .$$

Per il Teorema di Lagrange in \mathbb{R} , esiste t^* tale che

$$\phi(1) - \phi(0) = \phi'(t^*)(1 - 0) = Df(x + t^*(y - x))(y - x)$$

quindi, ponendo $\xi = x + t^*(y - x) \in [x, y]$, troviamo che

$$f(x) - f(y) = Df(\xi)(x - y) .$$

□

3.1 Derivate successive e Teorema di Taylor

Definizione 3.9 (Matrice di Hesse). La *matrice hessiana* (o *matrice di Hesse*) di una funzione $f : \mathbb{R}^d \rightarrow \mathbb{R}$ è la matrice $d \times d$ delle derivate parziali seconde di f :

$$H_f(x_0) = \begin{pmatrix} \frac{\partial^2 f(x_0)}{\partial x_1^2} & \frac{\partial^2 f(x_0)}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f(x_0)}{\partial x_1 \partial x_d} \\ \frac{\partial^2 f(x_0)}{\partial x_2 \partial x_1} & \frac{\partial^2 f(x_0)}{\partial x_2^2} & \cdots & \frac{\partial^2 f(x_0)}{\partial x_2 \partial x_d} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f(x_0)}{\partial x_d \partial x_1} & \frac{\partial^2 f(x_0)}{\partial x_d \partial x_2} & \cdots & \frac{\partial^2 f(x_0)}{\partial x_d^2} \end{pmatrix} .$$

Osservazione 3.8. Se la funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ è vettoriale, allora il vettore delle derivate parziali seconde non è una matrice, ma un tensore (qualunque cosa voglia dire).

Definizione 3.10 (Hessiano e Laplaciano). Data una funzione $f : U \rightarrow \mathbb{R}$, con $U \subseteq \mathbb{R}^d$ aperto, e un punto $x_0 \in U$ tale che f ammette tutte le derivate seconde in x_0 , definiamo l'hessiano di f nel punto x_0 come il determinante della matrice hessiana $\det H_f(x_0)$; definiamo il laplaciano di f nel punto x_0 come la traccia della matrice hessiana:

$$\Delta f(x_0) := \operatorname{tr}(H_f(x_0)) = \sum_{i=1}^d \frac{\partial^2 f}{\partial x_i^2}(x_0) .$$

Notazione. Un multi-indice è una n -upla $\alpha = (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$.

Dati due multi-indici $\alpha, \beta \in \mathbb{N}^n$ e $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, valgono le seguenti regole:

- i. $\alpha \pm \beta = (\alpha_1 \pm \beta_1, \dots, \alpha_n \pm \beta_n)$;
- ii. $\alpha \leq \beta \iff \alpha_i \leq \beta_i \forall i$;
- iii. $|\alpha| = \alpha_1 + \dots + \alpha_n$;
- iv. $\alpha! = \alpha_1! \dots \alpha_n!$;
- v. $x^\alpha = x_1^{\alpha_1} \dots x_n^{\alpha_n}$;
- vi. $D^\alpha = D_1^{\alpha_1} \dots D_n^{\alpha_n}$, dove $D_i^j := \frac{\partial^j}{\partial x_i^j}$.

Definizione 3.11 (Polinomio di Taylor). Data una funzione $f : U \rightarrow \mathbb{R}$ di classe C^k , con $U \subseteq \mathbb{R}^d$ un aperto, e un punto $x \in U$, il polinomio di Taylor di f di ordine k e centro x valutato in $x + v$ è:

$$T_k(f, x)(x + v) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} D^\alpha f(x) v^\alpha .$$

Teorema 3.9: Teorema di Taylor

Sia $f : U \rightarrow \mathbb{R}$ una funzione di classe C^k , con $U \subseteq \mathbb{R}^n$ aperto, $x \in U$, $v \in \mathbb{R}^n$, allora

$$f(x + v) = T_k(f, x)(x + v) + R_k(f, x)(x + v)$$

dove $T_k(f, x)(x + v)$ è il polinomio di Taylor di f di grado k e centro x valutato in $x + v$, e

$$R_k(f, x)(x + v) = o(|v|^k) . \quad \text{(Resto di Peano)}$$

Se f è anche derivabile $k + 1$ volte in U , allora

$$R_k(f, x)(x + v) = \sum_{|\alpha|=k+1} \frac{1}{\alpha!} D^\alpha f(\xi) v^\alpha \quad \text{(Resto di Lagrange)}$$

dove $\xi \in (x, x + v)$.

Dimostrazione. Dimostriamo innanzitutto la formula del resto di Lagrange. Per farlo, tentiamo di ricondurci al caso unidimensionale definendo una funzione ausiliaria. Definiamo $\varphi : \mathbb{R} \rightarrow \mathbb{R}$ data da $t \mapsto f(x + tv)$, che è di classe C^k in quanto composizione di funzioni C^k . Osserviamo che $\varphi(1) = f(x + v)$.

Ora, calcolando le derivate di φ , otteniamo:

$$\varphi'(t) = \nabla f(x + tv) \cdot v = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x + tv) v_i = \sum_{|\alpha|=1} D^\alpha f(x + tv) v^\alpha ;$$

$$\begin{aligned}\varphi''(t) &= \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i}(x+tv) \right)' v_i = \sum_{i=1}^n \left(\nabla \left(\frac{\partial f}{\partial x_i} \right) (x+tv) \cdot v \right) v_i = \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_i}(x+tv) v_i v_j = \\ &= v^\top H_f(x+tv)v = \sum_{|\alpha|=2} \frac{2!}{\alpha!} D^\alpha f(x+tv) v^\alpha ;\end{aligned}$$

per induzione, otteniamo:

$$\varphi^{(h)}(t) = \sum_{i_1, \dots, i_h=1}^n \frac{\partial^h f}{\partial x_{i_h} \dots \partial x_{i_1}}(x+tv) v_{i_1} \dots v_{i_h} = \sum_{|\alpha|=h} \frac{h!}{\alpha!} D^\alpha f(x+tv) v^\alpha$$

dove il coefficiente numerico $\frac{h!}{\alpha!}$ è dovuto al conteggio ripetuto di $D^\alpha f$ a causa delle permutazioni degli indici che fissano α (e poiché f è di classe C^k). A questo punto, poiché φ è di classe C^k e derivabile $k+1$ volte, possiamo scrivere $\varphi(1)$ come sviluppo di φ in 0:

$$\varphi(1) = \sum_{n=0}^k \frac{1}{n!} \varphi^{(n)}(0) + \frac{1}{(k+1)!} \varphi^{(k+1)}(\eta)$$

dove $\eta \in (0, 1)$. Quindi abbiamo ottenuto

$$f(x+v) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} D^\alpha f(x) v^\alpha + \sum_{|\alpha|=k+1} \frac{1}{\alpha!} D^\alpha f(\xi) v^\alpha$$

dove $\xi = x + \eta v \in (x, x+v)$, che è la formula del resto di Lagrange.

Ora, per dimostrare la formula del resto di Peano, osserviamo che se f è di classe C^k , allora in particolare è di classe C^{k-1} e derivabile k volte, quindi possiamo applicare la formula con il resto di Lagrange e ottenere:

$$f(x+v) = \sum_{|\alpha| \leq k-1} \frac{1}{\alpha!} D^\alpha f(x) v^\alpha + \sum_{|\alpha|=k} \frac{1}{\alpha!} D^\alpha f(\xi) v^\alpha .$$

Poiché f è di classe C^k , le derivate k -esime sono continue, quindi, per $|\alpha| = k$, vale

$$D^\alpha f(\xi) v^\alpha = (D^\alpha f(x) + o(1)) v^\alpha = D^\alpha f(x) v^\alpha + o(|v^\alpha|) = D^\alpha f(x) v^\alpha + o(|v|^k) .$$

Abbiamo quindi ottenuto la formula del resto di Peano:

$$f(x+v) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} D^\alpha f(x) v^\alpha + o(|v|^k) .$$

□

Teorema 3.10: Teorema di Schwarz

Sia $\Omega \subseteq \mathbb{R}^d$ un aperto e sia $f : \Omega \rightarrow \mathbb{R}$ una funzione di classe C^2 . Allora

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(x) = \frac{\partial^2 f}{\partial x_j \partial x_i}(x) \quad \forall 1 \leq i < j \leq d, \quad \forall x \in \Omega .$$

Dimostrazione. Osserviamo innanzitutto che è sufficiente dimostrare l'enunciato per $d = 2$; infatti, se $d > 2$, basta restringersi al piano di x_i e x_j , in cui vale la tesi.

Sia quindi $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ e $\bar{x} \in \Omega$: vogliamo dimostrare che la tesi vale in \bar{x} , e possiamo supporre per semplicità $\bar{x} = (0, 0)$. Poiché Ω è aperto, esiste $\varepsilon > 0$ tale che $A = (0, 0)$, $B = (\varepsilon, 0)$, $C = (0, \varepsilon)$ e $D = (\varepsilon, \varepsilon)$ sono tutti in Ω (in particolare il quadrato di vertici A, B, C, D è in Ω).

Consideriamo la funzione $\phi(t) = f(t, \varepsilon) - f(t, 0)$, differenziabile perché f è differenziabile. Quindi

$$f(D) - f(B) - f(C) + f(A) = \phi(\varepsilon) - \phi(0) = \varepsilon \phi'(\eta) = \quad \text{con } |\eta| < |\varepsilon|$$

$$= \varepsilon \left(\frac{\partial f}{\partial x}(\eta, \varepsilon) - \frac{\partial f}{\partial x}(\eta, 0) \right) = \varepsilon^2 \frac{\partial^2 f}{\partial y \partial x}(\eta, \sigma) \quad \text{con } |\sigma| < |\varepsilon|,$$

dove si è usata la differenziabilità fino al secondo ordine di f .

Analogamente, possiamo definire $\psi(t) = f(\varepsilon, t) - f(0, t)$, nuovamente differenziabile, per cui, ripetendo i conti otteniamo

$$\begin{aligned} f(D) - f(C) - f(B) + f(A) &= \psi(\varepsilon) - \psi(0) = \varepsilon \psi'(\tilde{\sigma}) = \quad \text{con } |\tilde{\sigma}| < |\varepsilon| \\ &= \varepsilon \left(\frac{\partial f}{\partial y}(\varepsilon, \tilde{\sigma}) - \frac{\partial f}{\partial y}(0, \tilde{\sigma}) \right) = \varepsilon^2 \frac{\partial^2 f}{\partial x \partial y}(\tilde{\eta}, \tilde{\sigma}) \quad \text{con } |\tilde{\eta}| < |\varepsilon|. \end{aligned}$$

Uguagliamo le espressioni ottenute, e poiché le derivate seconde di f sono continue, concludiamo:

$$\frac{\partial^2 f}{\partial y \partial x}(\eta, \sigma) = \frac{\partial^2 f}{\partial x \partial y}(\tilde{\eta}, \tilde{\sigma}) \xrightarrow{\varepsilon \rightarrow 0} \frac{\partial^2 f}{\partial y \partial x}(0, 0) = \frac{\partial^2 f}{\partial x \partial y}(0, 0) .$$

□

Corollario 3.11

Data $f : \mathbb{R}^d \rightarrow \mathbb{R}$ una funzione di classe C^2 , allora

$$f(x+v) = f(x) + \nabla f(x) \cdot v + \frac{1}{2} v^\top H_f(x) v + o(|v|^2)$$

e $H_f(x)$ è una matrice simmetrica.

3.2 Massimi e minimi liberi e vincolati

L'Everest è importante, ma anche il monte Bianco è bello. So che c'è la fossa delle Marianne, ma se mi si allaga la cantina voglio saperlo.

Preoccupazioni di Aldo Pratelli

Definizione 3.12 (Insieme di livello). Data una funzione $f : U \rightarrow \mathbb{R}$ e dato $c \in \mathbb{R}$, si definisce *insieme di livello* l'insieme

$$f^{-1}(c) = \{x \in U \mid f(x) = c\}$$

Se $U \subseteq \mathbb{R}^2$, un insieme di livello è una curva ed è detta *curva di livello*.

Teorema 3.12: Moltiplicatori di Lagrange

Se $f, g : \mathbb{R}^d \rightarrow \mathbb{R}$ sono due funzioni di classe C^1 , $\bar{x} \in \Gamma = \{x \in \mathbb{R}^d : g(x) = \text{cost}\}$ e $\nabla g(\bar{x}) \neq 0$, e \bar{x} è punto di massimo o di minimo di $f|_\Gamma$, allora

$$\exists \lambda \in \mathbb{R} : \nabla f(\bar{x}) = \lambda \nabla g(\bar{x})$$

Dimostrazione. Supponiamo per assurdo che $\nexists \lambda$ tale che $\nabla f(\bar{x}) = \lambda \nabla g(\bar{x})$, ossia che $\nabla f(\bar{x})$ e $\nabla g(\bar{x})$ non siano paralleli. Possiamo quindi scomporre $\nabla f(\bar{x})$ come: $\nabla f(\bar{x}) = \lambda \nabla g(\bar{x}) + w$, con $w \cdot \nabla g(\bar{x}) = 0$ (ossia w e $\nabla g(\bar{x})$ perpendicolari).

Sia $y \in \Gamma$, $y \neq \bar{x}$ tale che $|\bar{x} - y| \ll 1$ ¹. Allora $(y - \bar{x}) \cdot \nabla g(\bar{x}) = o(|y - \bar{x}|)$, da cui

$$\begin{aligned} f(y) &= f(\bar{x}) + (y - \bar{x}) \cdot \nabla f(\bar{x}) + o(|y - \bar{x}|) = \\ &= f(\bar{x}) + (y - \bar{x}) \cdot (\lambda \nabla g(\bar{x}) + w) + o(|y - \bar{x}|) = \\ &= f(\bar{x}) + (y - \bar{x}) \cdot w + o(|y - \bar{x}|) . \end{aligned}$$

Ora osserviamo che $y - \bar{x} \neq 0$ e che $(y - \bar{x}) \cdot w = O(|y - \bar{x}|)$, perciò posso scegliere y in modo che risulti $(y - \bar{x}) \cdot w > 0$, che implica $f(y) > f(\bar{x})$ e quindi \bar{x} non è punto di massimo; altrimenti posso scegliere y in modo che risulti $(y - \bar{x}) \cdot w < 0$, che implica $f(y) < f(\bar{x})$ e quindi \bar{x} non è punto di minimo. \neq \square

Vediamo un enunciato più generale del metodo dei moltiplicatori di Lagrange, di cui non diamo la dimostrazione.

Teorema 3.13: Moltiplicatori di Lagrange

Dato $\Omega \subseteq \mathbb{R}^n$ aperto, sia $g = (g_1, \dots, g_m) \in C^1(\Omega, \mathbb{R}^m)$ un vettore di funzioni vincolo, $f \in C^1(\Omega, \mathbb{R})$, $M = \{x \in \Omega \mid g(x) = 0\}$ il luogo di zeri di g e sia $P \in M$ tale che $Dg(P)$ ha rango massimo; se P è punto di estremo locale per $f|_M$, allora

$$\exists \lambda_1, \dots, \lambda_m \text{ tali che } \begin{cases} \nabla f(P) = \sum_{i=1}^m \lambda_i \nabla g_i(P) \\ g(P) = 0 \end{cases} .$$

Osserviamo che la condizione di dipendenza lineare del gradiente di f rispetto ai gradienti dei vincoli g_i corrisponde ad un sistema di n equazioni, mentre la condizione di appartenenza a M corrisponde ad un sistema di m equazioni: il sistema nell'enunciato è dunque un sistema di $n + m$ equazioni in $n + m$ incognite (vale a dire i λ_i e le componenti di P) in genere non lineare. Per queste ragioni, dove possibile si tende fortemente a non usare questo metodo.

Data una funzione $f : \mathbb{R}^d \rightarrow \mathbb{R}$ di classe C^2 per il Teorema di Schwarz l'hessiana $H_f(x)$ è una matrice simmetrica per ogni x . Per il Teorema spettrale, $H_f(x)$ è ortogonalmente diagonalizzabile, ossia esistono $\lambda_1, \dots, \lambda_d$ autovalori e v_1, \dots, v_d autovettori ortonormali per $H_f(x)$, tali che

$$H_f(x)v_i = \lambda_i v_i \text{ e } \langle v_i, v_j \rangle = \delta_{ij} .$$

Ricordiamo che una matrice quadrata simmetrica A si dice:

- **definita positiva** se ha tutti autovalori strettamente positivi, ossia se $x^\top A x > 0 \ \forall x \neq 0$ (in simboli $A > 0$);
- **definita negativa** se ha tutti autovalori strettamente negativi, ossia se $x^\top A x < 0 \ \forall x \neq 0$ (in simboli $A < 0$);
- **semidefinita positiva** se ha tutti autovalori positivi o nulli, ossia se $x^\top A x \geq 0 \ \forall x$ (in simboli $A \geq 0$);
- **semidefinita negativa** se ha tutti autovalori negativi o nulli, ossia se $x^\top A x \leq 0 \ \forall x$ (in simboli $A \leq 0$).

Il prossimo risultato generalizza il calcolo della derivata seconda per lo studio della convessità dei punti critici visto ad Analisi I.

¹L'esistenza di un tale y non è scontata, ma verrà garantita dal teorema della funzione implicita.

Teorema 3.14

Dato $U \subseteq \mathbb{R}^d$, sia $f : U \rightarrow \mathbb{R}$ una funzione di classe C^2 in un intorno di x_0 ; valgono le seguenti:

- i. x_0 minimo locale $\implies H_f(x_0) \geq 0$;
- ii. x_0 massimo locale $\implies H_f(x_0) \leq 0$;
- iii. $H_f(x_0) > 0$ e $\nabla f(x_0) = 0 \implies x_0$ minimo locale;
- iv. $H_f(x_0) < 0$ e $\nabla f(x_0) = 0 \implies x_0$ massimo locale.

Dimostrazione. Dimostriamo i punti i e iii; gli altri sono analoghi.

- i. Per assurdo, esista un autovalore $\lambda < 0$ con relativo autovettore v , ossia tali che $H_f(x_0)v = \lambda v$. A meno di riscalarlo l'autovettore, possiamo supporre $|v|^2 = 1$. Allora, poiché x_0 è un punto di minimo, vale $\nabla f(x_0) = 0$, quindi per Taylor possiamo scrivere:

$$\begin{aligned} f(x_0 + tv) &= f(x_0) + \frac{1}{2}\lambda t^2 \langle v, v \rangle + o(t^2) \quad \text{per } t \rightarrow 0 \\ &= f(x_0) + t^2 \left(\frac{\lambda}{2} + o(1) \right) \quad \text{per } t \rightarrow 0 \end{aligned}$$

Ora, per $t \rightarrow 0$, il termine $\frac{\lambda}{2} + o(1)$ assume valori negativi, che è un assurdo perché $f(x_0)$ è minimo \neq .

- iii. Per Taylor, abbiamo

$$f(x_0 + v) = f(x_0) + \frac{1}{2} \langle H_f(x_0)v, v \rangle + o(|v|^2) \quad \text{per } |v| \rightarrow 0$$

Ora, per il Teorema spettrale, possiamo scrivere $v = \alpha_1 v_1 + \dots + \alpha_d v_d$, dove i v_i sono una base ortonormale di autovettori, ossia tali che $H_f(x_0)v_i = \lambda_i v_i$, $|v_i|^2 = 1$ e $\langle v_i, v_j \rangle = \delta_{ij}$. Sia inoltre $\lambda_{\min} = \min_i \{\lambda_i\}$; poiché $H_f(x_0) > 0$, $\lambda_{\min} > 0$. Quindi abbiamo:

$$\begin{aligned} f(x_0 + v) &= f(x_0) + \frac{1}{2} \left\langle \sum_{i=1}^d \lambda_i \alpha_i v_i, \sum_{i=1}^d \alpha_i v_i \right\rangle + o(|v|^2) \quad \text{per } |v| \rightarrow 0 \\ &= f(x_0) + \frac{1}{2} \sum_{i=1}^d \lambda_i \alpha_i^2 + o(|v|^2) \geq f(x_0) + |v|^2 \left(\frac{\lambda_{\min}}{2} + o(1) \right) \geq f(x_0) \end{aligned}$$

ossia $f(x_0)$ è minimo. □

Esempio 3.15. Osserviamo che le implicazioni non possono essere indebolite. Ad esempio, $\nabla f(x_0) = 0$ e $H_f(x_0) \geq 0$ non implicano che x_0 sia un punto di minimo globale. Basta considerare la funzione $f(x, y) = x^2 - y^4$, che in $(0, 0)$ ha hessiana semidefinita positiva

$$H_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix},$$

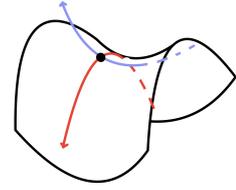
ma chiaramente $(0, 0)$ non è minimo locale.

Nella pratica, per determinare se una matrice H è definita positiva o negativa (o nessuna delle due), si fa ricorso al **criterio dei minori di nord-ovest**, ossia le sottomatrici H_k , ottenute cancellando le ultime $n - k$ righe e colonne di H . Vale quindi:

$$H > 0 \iff \det H_k > 0 \quad \text{per } 1 \leq k \leq d;$$

$$H < 0 \iff \det(-H_k) > 0 \text{ per } 1 \leq k \leq d.$$

Definizione 3.13 (Punto di sella). Un punto critico \bar{x} ($\nabla f(\bar{x}) = 0$) che non è né di massimo né di minimo (ossia $H_f(\bar{x})$ non è definita positiva o negativa) è detto *punto di sella*.



3.3 Teorema della funzione implicita

Sfruttando l'identificazione $\mathbb{R}^{n+m} = \mathbb{R}^n \times \mathbb{R}^m$ nel seguito rappresentiamo gli elementi di \mathbb{R}^{n+m} come coppie $(x, y) \in \mathbb{R}^{n+m}$ dove $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$.

In tal senso, data una funzione $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^k$ di classe C^1 , denoteremo $D_x F(x, y)$ e $D_y F(x, y)$ rispettivamente le prime n e le ultime m colonne di $DF(x, y)$; inoltre data $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$, indicheremo $\partial_{x_i} F(x)$ il vettore $DF(x)(e_i)$.

Teorema 3.16: Teorema di Dini

Sia $g : \mathbb{R}^d \times \mathbb{R} \rightarrow \mathbb{R}$ una funzione C^1 . Se $g(0, 0) = 0$ e $\partial_y g(0, 0) \neq 0$, allora

- i. $\exists r_1, r_2 > 0$ ed esiste una funzione $\varphi : [-r_1, r_1]^d \rightarrow [-r_2, r_2]$ tale che $\forall x \in [-r_1, r_1]^d$ e $\forall y \in [-r_2, r_2]$ si ha $\partial_y g(x, y) \neq 0$ e inoltre $g(x, y) = 0 \iff y = \varphi(x)$;
- ii. φ è di classe C^1 e $\forall x \in [-r_1, r_1]^d$ vale $D\varphi(x) = -\partial_y g(x, \varphi(x))^{-1} D_x g(x, \varphi(x))$.

Dimostrazione. Poiché per ipotesi g è C^1 , la mappa $(x, y) \mapsto \partial_y g(x, y)$ è continua, per cui, dato che $\partial_y g(0, 0) \neq 0$, $\exists r_2 > 0$ tale che $\text{sgn } \partial_y g(x, y) = \text{sgn } \partial_y g(0, 0) \forall (x, y) \in [-r_2, r_2]^{d+1}$. In particolare, la restrizione di g al segmento $[(0, -r_2), (0, r_2)]$ ha derivata (parziale) a segno costante, pertanto su tale segmento g è strettamente monotona; considerando che $g(0, 0) = 0$ troviamo che $g(0, -r_2)g(0, r_2) < 0$. Inoltre per continuità di g $\exists r_1 > 0$, e possiamo richiedere $r_1 \leq r_2$, tale che $\forall x \in [-r_1, r_1]^d$ si ha $g(x, -r_2)g(x, r_2) < 0$; dato che $r_1 \leq r_2$ la derivata parziale $\partial_y g(x, y)$ è ancora a segno costante, pertanto $\forall x \in [-r_1, r_1]^d$ la funzione $y \mapsto g(x, y)$ ristretta all'intervallo $[-r_2, r_2]$ è strettamente monotona e di conseguenza ha un unico zero. Definiamo allora $\varphi : [-r_1, r_1]^d \rightarrow [-r_2, r_2]$ dove

$$\varphi(x) := \text{l'unico zero della funzione } g_x : \begin{matrix} [-r_2, r_2] \rightarrow [-r_2, r_2] \\ y \mapsto g(x, y) \end{matrix} .$$

Consideriamo adesso la funzione $f : [-r_1, r_1]^d \rightarrow [-r_2, r_2]$ con $f(x) = g(x, \varphi(x))$: è chiaro dalla definizione di φ che la funzione f è costantemente nulla; inoltre è data dalla composizione di g e della mappa $x \mapsto (x, \varphi(x))$, che chiamiamo h . Se tentassimo di differenziare f troveremmo che

$$0 = Df(x) = Dg(x, \varphi(x)) \circ Dh(x) = D_x g(x, \varphi(x)) + \frac{\partial g}{\partial y}(x, \varphi(x)) D\varphi(x)$$

che darebbe la formula cercata: il problema di una tale dimostrazione è che non abbiamo alcuna certezza della differenziabilità di φ . Dimostriamo allora che φ è di classe C^1 .

Anzitutto mostriamo che φ è continua tramite la caratterizzazione con la continuità per successioni. Sia $\{x_n\} \subseteq [-r_1, r_1]^d$ una successione tale che $x_n \rightarrow \bar{x}$ per $n \rightarrow +\infty$, e supponiamo per assurdo che $\varphi(x_n) \not\rightarrow \varphi(\bar{x})$: allora $\exists \varepsilon > 0$ tale che $\forall \bar{n} \exists n > \bar{n}$ con $|\varphi(x_n) - \varphi(\bar{x})| \geq \varepsilon$, quindi in particolare $\exists \{x'_n\}$ sottosuccessione tale che $|\varphi(x'_n) - \varphi(\bar{x})| \geq \varepsilon \forall n$. Ora però $\{\varphi(x'_n)\} \subseteq [-r_2, r_2]$ che è compatto, dunque $\exists \{x''_n\}$ sottosottosuccessione tale che $\varphi(x''_n) \rightarrow z$ per qualche $z \in [-r_2, r_2]$; inoltre $z \neq \varphi(\bar{x})$ in quanto $|\varphi(x''_n) - \varphi(\bar{x})| \geq \varepsilon$. Usiamo adesso la continuità di g :

$$0 = g(x''_n, \varphi(x''_n)) \xrightarrow{n \rightarrow +\infty} g(\bar{x}, z) ,$$

pertanto dovremmo avere $g(\bar{x}, z) = 0$, ossia $z = \varphi(\bar{x})$. †

Poiché g è C^1 possiamo svilupparla con Taylor:

$$\begin{aligned} 0 &= g(x + \varepsilon e_i, \varphi(x + \varepsilon e_i)) = \\ &= g(x, \varphi(x + \varepsilon e_i)) + \varepsilon \partial_{x_i} g(x + \eta e_i, \varphi(x + \varepsilon e_i)) = \quad \text{con } |\eta| < |\varepsilon| \\ &= g(x, \varphi(x) + \varphi(x + \varepsilon e_i) - \varphi(x)) + \varepsilon \partial_{x_i} g(x + \eta e_i, \varphi(x + \varepsilon e_i)) = \end{aligned}$$

poniamo $\sigma := \varphi(x + \varepsilon e_i) - \varphi(x)$, allora

$$\begin{aligned} &= g(x, \varphi(x) + \sigma) + \varepsilon \partial_{x_i} g(x + \eta e_i, \varphi(x + \varepsilon e_i)) = \\ &= \cancel{g(x, \varphi(x))} + \sigma \partial_y g(x, \varphi(x) + \xi) + \varepsilon \partial_{x_i} g(x + \eta e_i, \varphi(x + \varepsilon e_i)) \quad \text{con } |\xi| < |\sigma|. \end{aligned}$$

La condizione $(\varphi(x + \varepsilon e_i) - \varphi(x)) \partial_y g(x, \varphi(x) + \xi) + \varepsilon \partial_{x_i} g(x + \eta e_i, \varphi(x + \varepsilon e_i)) = 0$ si riscrive come

$$\frac{\varphi(x + \varepsilon e_i) - \varphi(x)}{\varepsilon} = - \frac{\partial g}{\partial y}(x, \varphi(x) + \xi)^{-1} \frac{\partial g}{\partial x_i}(x + \eta e_i, \varphi(x + \varepsilon e_i))$$

che passando al limite, in virtù della continuità di g , delle sue derivate parziali e di φ , diventa

$$\frac{\partial \varphi}{\partial x_i}(x) = - \frac{\partial g}{\partial y}(x, \varphi(x))^{-1} \frac{\partial g}{\partial x_i}(x, \varphi(x)) \quad :$$

abbiamo quindi dimostrato che φ ammette derivate parziali continue, ovvero φ è C^1 . Dalla relazione precedente segue infine

$$D\varphi(x) = - \frac{\partial g}{\partial y}(x, \varphi(x))^{-1} D_x g(x, \varphi(x)) \quad .$$

□

Commento. Il Teorema di Dini dice che gli insiemi di livello di una mappa $g : \mathbb{R}^n \rightarrow \mathbb{R}$ di classe C^1 sono localmente dei grafici: in particolare possiamo esprimere localmente ogni coordinata con derivata parziale non nulla come funzione C^1 delle altre coordinate.

Vediamo adesso una versione più generale del Teorema di Dini, che ricalca la stessa idea.

Teorema 3.17: Teorema della funzione implicita

Sia $g : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ una funzione di classe C^1 tale che $g(0, 0) = 0$ e $\det D_y g(0, 0) \neq 0$. Allora:

- i. $\exists r_1, r_2 > 0$ ed esiste una funzione $\varphi : [-r_1, r_1]^n \rightarrow [-r_2, r_2]^m$ tale che $\forall x \in [-r_1, r_1]^n$ e $\forall y \in [-r_2, r_2]^m$ si ha $\det D_y g(x, y) \neq 0$ e inoltre $g(x, y) = 0 \iff y = \varphi(x)$;
- ii. φ è di classe C^1 e $\forall x \in [-r_1, r_1]^n$ vale $D\varphi(x) = -D_y g(x, \varphi(x))^{-1} D_x g(x, \varphi(x))$.

Dimostrazione. Mostriamo innanzitutto il secondo punto assumendo che esista φ come nell'enunciato. Per provare la continuità di φ , dimostriamo che φ è continua per successioni (negli spazi metrici le due proprietà sono equivalenti).

Sia $\{x_n\} \subseteq [-r_1, r_1]^n$ una successione, $x_n \rightarrow x$, e supponiamo per assurdo che $\varphi(x_n) \not\rightarrow \varphi(x)$: allora $\exists \varepsilon > 0$ tale che $\forall \bar{n} \exists n > \bar{n}$ con $|\varphi(x_n) - \varphi(x)| \geq \varepsilon$, quindi in particolare $\exists \{x'_n\}$ sottosuccessione tale che $|\varphi(x'_n) - \varphi(x)| \geq \varepsilon \forall n$. Adesso $\{\varphi(x'_n)\} \subseteq [-r_2, r_2]^m$ che è compatto, dunque $\exists \{x''_n\}$ sottosuccessione tale che $\varphi(x''_n) \rightarrow y$ per qualche $y \in [-r_2, r_2]^m$. Usiamo adesso la continuità di g :

$$0 = g(x''_n, \varphi(x''_n)) \xrightarrow{n \rightarrow +\infty} g(x, y) \quad ,$$

pertanto dovremmo avere $g(x, y) = 0$, ossia $y = \varphi(x)$. \neq

Per mostrare che φ è di classe C^1 utilizziamo il teorema di Taylor con resto di Lagrange:

$$\begin{aligned} 0 &= g(x + \varepsilon e_i, \varphi(x + \varepsilon e_i)) = \\ &= g(x, \varphi(x + \varepsilon e_i)) + \varepsilon D_x g(x + \eta e_i, \varphi(x + \varepsilon e_i)) \cdot e_i = \quad \text{con } |\eta| < |\varepsilon| \\ &= g(x, \varphi(x + \varepsilon e_i)) + \varepsilon \partial_{x_i} g(x + \eta e_i, \varphi(x + \varepsilon e_i)) = \\ &= g(x, \varphi(x) + \varphi(x + \varepsilon e_i) - \varphi(x)) + \varepsilon \partial_{x_i} g(x + \eta e_i, \varphi(x + \varepsilon e_i)) = \end{aligned}$$

poniamo $\sigma := \varphi(x + \varepsilon e_i) - \varphi(x)$, allora

$$\begin{aligned} &= g(x, \varphi(x) + \sigma) + \varepsilon \partial_{x_i} g(x + \eta e_i, \varphi(x + \varepsilon e_i)) = \\ &= \underline{g(x, \varphi(x))} + D_y g(x, z)(\sigma) + \varepsilon \partial_{x_i} g(x + \eta e_i, \varphi(x + \varepsilon e_i)) \quad \text{con } z \in [\varphi(x), \varphi(x) + \sigma]. \end{aligned}$$

A questo punto

$$D_y g(x, z)(\sigma) = -\varepsilon \partial_{x_i} g(x + \eta e_i, \varphi(x + \varepsilon e_i))$$

e ricordando che $\sigma := \varphi(x + \varepsilon e_i) - \varphi(x)$

$$\frac{\varphi(x + \varepsilon e_i) - \varphi(x)}{\varepsilon} = -D_y g(x, z)^{-1} \left(\frac{\partial g}{\partial x_i}(x + \eta e_i, \varphi(x + \varepsilon e_i)) \right) .$$

Adesso, il termine di destra è composizione di funzioni continue per ε sufficientemente piccolo: abbiamo dimostrato la continuità di φ ; $\partial_{x_i} g$ è localmente continua dato che per ipotesi g è C^1 ; $D_y g(x, z)$ è continua dato che, vedendola come matrice, dipende in modo continuo dalle sue entrate, che sono continue in quanto derivate parziali di g , e $D_y g(x, z)^{-1}$ dipende a sua volta in modo continuo dalle entrate di $D_y g(x, z)^2$. A questo punto

$$\frac{\partial \varphi}{\partial x_i}(x) = \lim_{\varepsilon \rightarrow 0} \frac{\varphi(x + \varepsilon e_i) - \varphi(x)}{\varepsilon} = -D_y g(x, \varphi(x))^{-1} \left(\frac{\partial g}{\partial x_i}(x, \varphi(x)) \right)$$

ossia, in accordo con la nostra notazione

$$D\varphi(x) \cdot e_i = -D_y g(x, \varphi(x))^{-1} D_x g(x, \varphi(x)) \cdot e_i$$

da cui, applicando il risultato per ogni i ,

$$D\varphi(x) = -D_y g(x, \varphi(x))^{-1} D_x g(x, \varphi(x)) .$$

Dimostriamo adesso che una tale φ esiste.

Per ipotesi g è differenziabile in $(0, 0)$, pertanto per Taylor

$$g(x, y) = \cancel{g(0, 0)} + D_x g(0, 0)x + D_y g(0, 0)y + R(x, y) \quad \text{dove} \quad \frac{|R(x, y)|}{|(x, y)|} \xrightarrow{|(x, y)| \rightarrow 0} 0 ;$$

allora

$$\begin{aligned} g(x, y) = 0 &\iff D_x g(0, 0)x + D_y g(0, 0)y + R(x, y) = 0 \\ &\iff y = -D_y g(0, 0)^{-1} (D_x g(0, 0)x + R(x, y)) . \end{aligned}$$

A questo punto l'idea è scrivervi il termine di destra come funzione di y e cercarne un punto fisso. Facciamo un primo tentativo: definiamo

$$T_x : \mathbb{R}^m \rightarrow \mathbb{R}^m \quad \text{dove} \quad T_x(y) = -D_y g(0, 0)^{-1} (D_x g(0, 0)x + R(x, y)) ,$$

da cui $T_x(y) - T_x(\tilde{y}) = -D_y g(0, 0)^{-1} (R(x, y) - R(x, \tilde{y}))$.

Per lo sviluppo precedente $R(x, y) - R(x, \tilde{y}) = g(x, y) - g(x, \tilde{y}) - D_y g(0, 0)(y - \tilde{y})$, e per Lagrange $\exists z \in [y, \tilde{y}]$ tale che $g(x, y) - g(x, \tilde{y}) = D_y g(x, z)(y - \tilde{y})$, quindi in definitiva

$$R(x, y) - R(x, \tilde{y}) = (D_y g(x, z) - D_y g(0, 0))(y - \tilde{y}) .$$

²La regola della matrice aggiunta classica implica che le entrate dell'inversa di una matrice A dipendono in modo continuo dalle entrate di A .

Adesso $\exists c > 0$ una costante tale che $\forall w \in \mathbb{R}^m$ vale³ $|D_y g(0, 0)^{-1}(w)| \leq c|w|$, e inoltre $\exists r_2 > 0$ tale che $\forall (x, z) \in [-r_2, r_2]^{n+m}$ si ha⁴ $\det D_y g(x, z) \neq 0$ e

$$|(D_y g(x, z) - D_y g(0, 0))(w)| \leq \frac{1}{2c}|w| \quad \forall w \in \mathbb{R}^m.$$

Dai calcoli appena fatti segue che $|R(x, y) - R(x, \tilde{y})| \leq \frac{1}{2c}|y - \tilde{y}|$, e di conseguenza

$$\begin{aligned} |T_x(y) - T_x(\tilde{y})| &= |D_y g(0, 0)^{-1}(R(x, y) - R(x, \tilde{y}))| = \\ &= |D_y g(0, 0)^{-1}((D_y g(x, z) - D_y g(0, 0))(y - \tilde{y}))| \leq \\ &\leq c|(D_y g(x, z) - D_y g(0, 0))(y - \tilde{y})| \leq \ell \frac{1}{2c}|y - \tilde{y}| \leq \\ &\leq \frac{1}{2}|y - \tilde{y}|. \end{aligned}$$

Riassumendo, abbiamo trovato che $\forall x \in [-r_2, r_2]^n$ la funzione $T_x : [-r_2, r_2]^m \rightarrow \mathbb{R}^m$ definita come $y \mapsto -D_y g(0, 0)^{-1}(D_x g(0, 0)x + R(x, y))$ soddisfa la proprietà

$$|T_x(y) - T_x(\tilde{y})| \leq \frac{1}{2}|y - \tilde{y}| \quad \forall y, \tilde{y} \in [-r_2, r_2]^m ;$$

inoltre $[-r_2, r_2]^m$ è compatto e quindi completo, pertanto se valesse $\text{Im}(T_x) \subseteq [-r_2, r_2]^m$, T_x sarebbe una contrazione. Innanzitutto osserviamo che

$$|R(x, y)| - |R(x, 0)| \leq |R(x, y) - R(x, 0)| \leq \frac{1}{2c}|y|$$

pertanto $|R(x, y)| \leq |R(x, 0)| + \frac{1}{2c}|y|$. Stimiamo dunque $\text{Im}(T_x)$:

$$\begin{aligned} |T_x(y)| &= | -D_y g(0, 0)^{-1}D_x g(0, 0)x + (-D_y g(0, 0)^{-1}R(x, y)) | \leq \\ &\leq |D_y g(0, 0)^{-1}D_x g(0, 0)x| + |D_y g(0, 0)^{-1}R(x, y)| \leq \\ &\leq |D_y g(0, 0)^{-1}D_x g(0, 0)x| + c|R(x, y)| \leq \\ &\leq |D_y g(0, 0)^{-1}D_x g(0, 0)x| + c|R(x, 0)| + \ell \frac{1}{2c}|y|, \end{aligned}$$

ossia $|T_x(y)| \leq |D_y g(0, 0)^{-1}D_x g(0, 0)x| + c|R(x, 0)| + \frac{r_2}{2}$.

Adesso però $|D_y g(0, 0)^{-1}D_x g(0, 0)x| + c|R(x, 0)|$ è una funzione continua⁵ della sola x , pertanto $\exists r_1 > 0$ (e imponiamo anche $r_1 \leq r_2$) tale che per $x \in [-r_1, r_1]^n$ vale

$$|D_y g(0, 0)^{-1}D_x g(0, 0)x| + c|R(x, 0)| \leq \frac{r_2}{2}$$

e dunque

$$|T_x(y)| \leq \frac{r_2}{2} + \frac{r_2}{2} = r_2.$$

Abbiamo quindi trovato che $\forall x \in [-r_1, r_1]^n$ la funzione $T_x : [-r_2, r_2]^m \rightarrow [-r_2, r_2]^m$ è una contrazione e dunque ha un unico punto fisso. Definendo $\varphi : [-r_1, r_1]^n \rightarrow [-r_2, r_2]^m$ come la funzione che associa ad x l'unico punto fisso di T_x , e ritornando all'osservazione iniziale, troviamo che

$$\begin{aligned} g(x, y) = 0 &\iff D_x g(0, 0)x + D_y g(0, 0)y + R(x, y) = 0 \\ &\iff y = -D_y g(0, 0)^{-1}(D_x g(0, 0)x + R(x, y)) \\ &\iff y = T_x(y) \\ &\iff y = \varphi(x) \end{aligned}$$

che conclude la dimostrazione. □

³Abbiamo già mostrato questa proprietà nella dimostrazione del Teorema 3.5.

⁴Ciò deriva dal fatto generale che le entrate di $Dg(x, y)$ sono derivate parziali di g , che sono continue poiché g è C^1 : di conseguenza la mappa $(x, y) \mapsto Dg(x, y)$ è continua, ed in particolare lo è anche $(x, y) \mapsto D_y g(x, y)$.

⁵Abbiamo implicitamente definito $R(x, y) := g(x, y) - D_x g(0, 0)x - D_y g(0, 0)y$, che è somma di funzioni continue.

Definizione 3.14 (Punto regolare). Dato $\Omega \subseteq \mathbb{R}^d$ e $f \in C^1(\Omega, \mathbb{R}^n)$, un *punto regolare* di f è un punto di $\{x \in \Omega : f(x) = c\}$ tale che $\text{rk}(Df(x))$ è massimo.

Commento. Dato $\Omega \subseteq \mathbb{R}^{n+m}$, se P è un punto regolare per una funzione $g : \Omega \rightarrow \mathbb{R}^m$ di classe C^1 sull'insieme di livello $M = \{x \in \Omega \mid g(x) = c\}$, allora per il Teorema della funzione implicita M è un grafico C^1 di n variabili in un intorno di P .

Osservazione 3.18. Nel Teorema della funzione implicita, se g è di classe C^k allora anche φ è di classe C^k .

Definizione 3.15 (Tangente affine). Dati $\Omega \subseteq \mathbb{R}^d$ aperto e $f \in C^1(\Omega, \mathbb{R})$, consideriamo $M = \{x \in \Omega : f(x) = 0\}$ con $\nabla f(x) \neq 0 \forall x \in \Omega$. Dato $P \in \Omega$, si definisce il *tangente affine* di M in P

$$T_P M = \{x \in \mathbb{R}^d : \nabla f(P) \cdot (x - P) = 0\} .$$

Definizione 3.16 (Diffeomorfismo). Un applicazione $f : \mathbb{R}^d \rightarrow \mathbb{R}^d$ si dice *diffeomorfismo* di classe C^k se f è di classe C^k , è bigettiva e $f^{-1} : \mathbb{R}^d \rightarrow \mathbb{R}^d$ è ancora di classe C^k .

Teorema 3.19: Teorema di invertibilità locale

Sia $\Omega \subseteq \mathbb{R}^d$ un aperto, $f : \Omega \rightarrow \mathbb{R}^d$ una funzione di classe C^1 , e sia $x_0 \in \Omega$ tale che $\det Df(x_0) \neq 0$; allora esistono U e V intorno di x_0 e $f(x_0)$ tali che $f|_U : U \rightarrow V$ è un diffeomorfismo di classe C^1 . Inoltre $\forall x \in U$ vale la relazione

$$Df^{-1}(f(x)) = (Df(x))^{-1} .$$

Dimostrazione. Poiché $\det Df(x_0) \neq 0$, esiste un intorno sferico I_1 di x_0 tale che $\det Df(x) \neq 0 \forall x \in I_1$. Mostriamo che f è iniettiva, e quindi invertibile, su I_1 . Siano per assurdo $x_1, x_2 \in I_1$, $x_1 \neq x_2$, tali che $f(x_1) = f(x_2)$: per Lagrange, $\exists \xi \in [x_1, x_2] \subseteq I_1$ (il contenimento vale poiché abbiamo scelto I_1 intorno sferico) tale che $0 = f(x_1) - f(x_2) = Df(\xi)(x_1 - x_2)$, dunque il vettore $x_1 - x_2$ appartiene al \ker di $Df(\xi)$, perciò $\det Df(\xi) = 0$, ma $\xi \in I_1$. ∇
Mostriamo ora, che per un opportuno intorno U di x_0 , l'immagine $f(U)$ è intorno di $f(x_0)$ e che su tale insieme f^{-1} è di classe C^1 . Definiamo

$$\begin{aligned} F : \Omega \times \mathbb{R}^d &\longrightarrow \mathbb{R}^d \\ F(x, y) &= f(x) - y \end{aligned}$$

Sia $y_0 = f(x_0)$. Vale $DF(x, y) = (Df(x) \quad -I)$, da cui $\det D_x F(x_0, y_0) = \det Df(x_0) \neq 0$. Quindi, per il Teorema della funzione implicita (a meno di scambiare il ruolo di x e y), esistono un intorno V di y_0 , un intorno I_2 di x_0 e un'unica funzione $g : V \rightarrow I_2$ tale che

$$F(g(y), y) = f(g(y)) - y = 0 \quad \forall y \in V .$$

Possiamo supporre $I_2 \subseteq I_1$, a meno di sostituire I_2 con $I_1 \cap I_2$ e V con $g^{-1}(I_1 \cap I_2)$. Dato che f è iniettiva su I_1 , segue che g è l'inversa di f ristretta a $U = g(V) = f^{-1}(V) \cap I_1$, che è intorno di x_0 per la continuità di f .

Inoltre, dal teorema della funzione implicita, segue

$$Dg(f(x)) = -D_x F(x, f(x))^{-1} D_y F(x, f(x)) ,$$

ma $D_x F(x, f(x)) = Df(x)$, $D_y F(x, f(x)) = -I$, da cui la tesi. \square

3.4 Convessità e differenziabilità

Proposizione 3.20

Sia $\Omega \subseteq \mathbb{R}^d$ aperto e convesso:

i. se $f \in C^1(\Omega, \mathbb{R})$, allora

$$f \text{ convessa} \iff (\nabla f(x) - \nabla f(y)) \cdot (x - y) \geq 0 \quad \forall x, y \in \Omega ;$$

ii. se $f \in C^2(\Omega, \mathbb{R})$, allora

$$f \text{ convessa} \iff H_f(x) \geq 0 \quad \forall x \in \Omega .$$

Dimostrazione. Osserviamo innanzitutto che: f è convessa se e solo se la restrizione di f a ogni retta è convessa.

i. (\implies) Se f è convessa, e $x, y \in \Omega$, allora $\phi(t) = f(y + t(x - y))$ è convessa; inoltre vale $\phi'(t) = \nabla f(y + t(x - y)) \cdot (x - y)$. Poiché ϕ è convessa, la sua derivata prima è crescente, quindi

$$\phi'(1) - \phi'(0) \geq 0 \quad \text{cioè} \quad (\nabla f(x) - \nabla f(y)) \cdot (x - y) \geq 0 .$$

(\impliedby) Considerando la funzione ϕ dell'implicazione precedente, basta mostrare che ϕ è convessa $\forall x, y \in \Omega$, e quindi che f è convessa su ogni retta. Per farlo, mostriamo che ϕ' è debolmente crescente. Dati $t > s$, vale:

$$\phi'(t) - \phi'(s) = (\nabla f(y + t(x - y)) - \nabla f(y + s(x - y))) \cdot (x - y) \frac{t - s}{t - s} .$$

Ponendo $x_t = y + t(x - y)$ e $x_s = y + s(x - y)$, vale

$$\phi'(t) - \phi'(s) = (\nabla f(x_t) - \nabla f(x_s)) \cdot (x_t - x_s) \frac{1}{t - s} \geq 0 .$$

ii. (\implies) Definiamo la funzione $\phi(t) = f(x + tv)$, che è convessa per ogni scelta di $v \in \mathbb{R}^d$ e $x \in \Omega$. Allora si ha che

$$\phi'(t) = \sum_{i=1}^d \frac{\partial f}{\partial x_i}(x + tv)v_i \quad \text{e} \quad \phi''(t) = \sum_{i,j=1}^d \frac{\partial^2 f}{\partial x_j \partial x_i}(x + tv)v_i v_j .$$

Poiché ϕ è convessa, $\phi''(t) \geq 0 \forall t$, e in particolare $\phi''(0) \geq 0$, cioè

$$v^\top H_f(x)v = \sum_{i,j=1}^d \frac{\partial^2 f}{\partial x_j \partial x_i}(x)v_i v_j \geq 0 \quad \forall v ,$$

ossia $H_f(x) \geq 0 \forall x$.

(\impliedby) Se $H_f(x) \geq 0 \quad \forall x \in \Omega$, allora $\phi''_{x,v}(t) \geq 0$, cioè $\phi_{x,v}(t)$ è convessa $\forall x, v$, e quindi anche f è convessa.

□

3.5 Funzioni omogenee e radiali

Proposizione 3.21

Data $f : \mathbb{R}^d \rightarrow \mathbb{R}$ una funzione continua γ -omogenea e $C^1(\mathbb{R}^d \setminus \{0\})$, allora

- i. se $\gamma > 1$, f è differenziabile in 0;
- ii. se $\gamma = 1$, f è differenziabile se e solo se f è lineare;
- iii. se $\gamma < 1$, f non è differenziabile, a meno che $\gamma = 0$ e f sia costante.

Dimostrazione. Poiché f è continua, necessariamente $\gamma > 0$ oppure $\gamma = 0$ e f è costante, nel qual caso f è chiaramente differenziabile. Se invece $\gamma > 0$ allora l'unica estensione per continuità di f in 0 è $f(0) = 0$. Distinguiamo ora i tre casi:

i. ($\gamma > 1$) Per $|x| \rightarrow 0$ si ha che

$$|f(x)| = |x|^\gamma \left| f\left(\frac{x}{|x|}\right) \right| \leq \left(\sup_{|x|=1} |f(x)| \right) |x|^\gamma = o(|x|),$$

quindi

$$\lim_{|x| \rightarrow 0} \frac{f(x) - f(0)}{|x|} = \lim_{|x| \rightarrow 0} \frac{o(|x|)}{|x|} = 0,$$

in particolare f è differenziabile.

ii. ($\gamma = 1$). Verifichiamo che f è differenziabile se e solo se f è lineare:

(\Leftarrow) Chiaro, dato che le lineari sono per definizione 1-omogenee e sono differenziabili (in particolare coincidono con il loro differenziale).

(\Rightarrow) Sviluppiamo in 0 con Taylor

$$f(x) = f(0) + Df(0)(x) + R(x) = \nabla f(0) \cdot x + R(x)$$

con $R(x) = o(|x|)$. Osserviamo che per $t \in \mathbb{R} \setminus \{0\}$

$$f(x) = \frac{f(tx)}{t} = \nabla f(0) \cdot x + \frac{R(tx)}{t|x|} |x|.$$

Dato che $t \rightarrow 0$, e quindi $tx \rightarrow 0$, vale $\lim_{t \rightarrow 0} \frac{R(tx)}{t|x|} = 0$ e quindi passando al limite

$$f(x) = \nabla f(0) \cdot x,$$

ossia f è lineare.

iii. ($0 < \gamma < 1$) Se $f(\partial B(0,1)) = \{0\}$ allora f è costantemente nulla e siamo nel caso $\gamma = 1$ e f lineare. Se $f(\partial B(0,1)) \neq \{0\}$ allora consideriamo $a \in \mathbb{R}^n$ tale che $|a| = 1$ e $f(a) \neq 0$. Si ha che

$$\lim_{t \rightarrow 0^+} \frac{f(ta) - f(0)}{t} = \lim_{t \rightarrow 0^+} \frac{f(ta)}{t} = \lim_{t \rightarrow 0^+} |ta|^{\gamma-1} f\left(\frac{ta}{|ta|}\right) = (|a|^{\gamma-1} f(a)) \lim_{t \rightarrow 0^+} t^{\gamma-1} = \pm\infty,$$

quindi esiste una direzione per cui $\lim_{|x| \rightarrow 0} \frac{f(x) - f(0)}{|x|}$ non è definito, cioè il limite non è definito.

□

Definizione 3.17 (Cono). Un insieme $\Omega \subseteq \mathbb{R}^d$ si dice *cono* se $\forall x \in \Omega, \forall \lambda > 0$, si ha $\lambda x \in \Omega$.

Teorema 3.22: Teorema di Eulero

Sia $\Omega \subseteq \mathbb{R}^d$ un cono aperto e sia $f : \Omega \rightarrow \mathbb{R}$ una funzione differenziabile. Allora f è γ -omogenea se e solo se vale l'**identità di Eulero**, cioè

$$\nabla f(x) \cdot x = \gamma f(x) \quad \forall x \in \Omega.$$

Dimostrazione. Supponiamo che f sia γ -positivamente omogenea e consideriamo la mappa di sostituzione $g : \mathbb{R}^+ \times \Omega \rightarrow \mathbb{R}$ data da $g(\lambda, x) = \lambda x$, allora vale $(f \circ g)(\lambda, x) = f(\lambda x) = \lambda^\gamma f(x)$, pertanto derivando in λ troviamo

$$\begin{aligned} \gamma \lambda^{\gamma-1} f(x) &= \frac{\partial (f \circ g)}{\partial \lambda}(\lambda, x) = (D(f \circ g)(\lambda, x))(e_0) = (Df(g(\lambda, x)) \circ Dg(\lambda, x))(e_0) = \\ &= Df(\lambda x)(x) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\lambda x) x_i = \lambda^{\gamma-1} \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x) x_i = \lambda^{\gamma-1} \nabla f(x) \cdot x ; \end{aligned}$$

in particolare per $\lambda = 1$ troviamo la tesi.

Viceversa supponiamo che per ogni $x \in \Omega$ valga l'identità di Eulero, e fissato x consideriamo la mappa $h : \mathbb{R}^+ \rightarrow \mathbb{R}$ data da $h(\lambda) = \lambda^{-\gamma} f(\lambda x) - f(x)$: chiaramente h è derivabile e vale

$$\begin{aligned} h'(\lambda) &= -\gamma \lambda^{-\gamma-1} f(\lambda x) + \lambda^{-\gamma} \nabla f(\lambda x) \cdot x = -\gamma \lambda^{-\gamma-1} f(\lambda x) + \lambda^{-\gamma-1} (\nabla f(\lambda x) \cdot (\lambda x)) = \\ &= -\gamma \lambda^{-\gamma-1} f(\lambda x) + \lambda^{-\gamma-1} (\gamma f(\lambda x)) = 0 , \end{aligned}$$

pertanto h è costante; dal momento che $h(1) = 0$, h è la mappa nulla, pertanto $f(\lambda x) = \lambda^\gamma f(x)$. \square

Definizione 3.18 (Funzione radiale). Una funzione $f : \mathbb{R}^d \rightarrow \mathbb{R}$ si dice *radiale* se esiste una funzione $\phi : \mathbb{R} \rightarrow \mathbb{R}$ tale $\forall x \in \mathbb{R}^d$ si ha $f(x) = \phi(|x|)$.

Proposizione 3.23

Sia $f : \mathbb{R}^d \rightarrow \mathbb{R}$ una funzione radiale, e sia $\phi : \mathbb{R}^+ \rightarrow \mathbb{R}$ la mappa tale che $f(x) = \phi(|x|)$;

i. se ϕ è C^1 , allora anche f è C^1 e

$$\nabla f(x) = \phi'(|x|) \frac{x}{|x|} ;$$

ii. se ϕ è C^2 , allora anche f è C^2 , vale

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \phi''(|x|) \frac{x_i x_j}{|x|^2} + \frac{\phi'(|x|)}{|x|} (\delta_{ij} - \frac{x_i x_j}{|x|^2})$$

e inoltre

$$\Delta f(x) = \phi''(|x|) + \frac{\phi'(|x|)}{|x|} (d-1) .$$

Dimostrazione. Iniziamo calcolando $\nabla|x|$ e $\nabla \frac{1}{|x|}$:

$$\frac{\partial |x|}{\partial x_k} = \frac{\partial}{\partial x_k} \sqrt{\sum_{i=1}^d x_i^2} = \frac{1}{2} \frac{\frac{\partial}{\partial x_k} \sum_{i=1}^d x_i^2}{\sqrt{\sum_{i=1}^d x_i^2}} = \frac{x_k}{|x|} \quad \text{e dunque} \quad \nabla |x| = \frac{x}{|x|} ;$$

$$\nabla \frac{1}{|x|} = -\frac{1}{|x|^2} \nabla |x| = -\frac{1}{|x|^2} \frac{x}{|x|} = -\frac{x}{|x|^3} .$$

Ora abbiamo:

i. se ϕ è C^1 :

$$\nabla f(x) = \phi'(|x|) \nabla |x| = \phi'(|x|) \frac{x}{|x|} ;$$

ii. se ϕ è C^2 :

$$\begin{aligned} \frac{\partial^2 f}{\partial x_i \partial x_j}(x) &= \frac{\partial}{\partial x_i} \phi'(|x|) \frac{x_j}{|x|} = \phi''(|x|) \frac{x_i x_j}{|x| |x|} + \phi'(|x|) \left(\delta_{ij} \frac{1}{|x|} - \frac{x_i x_j}{|x|^3} \right) = \\ &= \phi''(|x|) \frac{x_i x_j}{|x|^2} + \frac{\phi'(|x|)}{|x|} \left(\delta_{ij} - \frac{x_i x_j}{|x|^2} \right) ; \end{aligned}$$

infine

$$\begin{aligned} \Delta f(x) &= \sum_{i=1}^d \frac{\partial^2 f}{\partial x_i^2}(x) = \phi''(|x|) \left(\sum_{i=1}^d \frac{x_i^2}{|x|^2} \right) + \frac{\phi'(|x|)}{|x|} \left(\sum_{i=1}^d 1 - \sum_{i=1}^d \frac{x_i^2}{|x|^2} \right) = \\ &= \phi''(|x|) + \frac{\phi'(|x|)}{|x|} (d-1) . \end{aligned}$$

□

Definizione 3.19 (Funzione armonica). Una funzione $f : U \rightarrow \mathbb{R}$ differenziabile due volte si dice *funzione armonica* se soddisfa l'equazione di Laplace:

$$\Delta f = 0 .$$

Osservazione 3.24. Le funzioni radiali armoniche sono tali che $f(x) = \phi(|x|)$ e $\Delta f = 0$, quindi, per il punto ii della Proposizione precedente, sono caratterizzate dalle mappe ϕ (e dai d) che risolvono l'equazione differenziale:

$$\phi''(t) + (d-1) \frac{\phi'(t)}{t} = 0 .$$

Essa è un'equazione a variabili separabili, e integrandola otteniamo

$$\frac{\phi''(t)}{\phi'(t)} = \frac{1-d}{t} \implies \log(\phi'(t)) = (1-d) \log(t) + k \implies \phi'(t) = ct^{1-d}$$

e dunque concludiamo

$$\phi(t) = c \log t + b \quad \text{per } d = 2 \text{ e } \quad \phi(t) = \frac{c}{2-d} t^{2-d} + b \quad \text{altrimenti.}$$

Capitolo 4

Successioni di funzioni

Spesso, nella risoluzione di problemi (ad esempio equazioni differenziali), capita di dover determinare una funzione con determinate proprietà; per farlo, un metodo efficace è tentare di approssimare la soluzioni con altre funzioni, con un grado di accuratezza sempre maggiore. Stiamo così implicitamente definendo una successione di funzioni, nella speranza che converga e che il limite sia proprio la soluzione cercata.

4.1 Convergenza uniforme

Perché il capitolo non si chiama “*Convergenza uniforme & friends*” o “*Convergenza puntuale & friends*” oppure “*Convergenza puntuale, uniforme e altri animali*”? Perché la convergenza puntuale fa schifo e non si merita un titolo.

Aldo Pratelli

Definizione 4.1 (Puntuale convergenza). Siano (X, d_X) e (Y, d_Y) due spazi metrici e sia $\{f_n\}_{n \in \mathbb{N}}$ una successione di funzioni $f_n : X \rightarrow Y$. Si dice che f_n converge puntualmente a f se $\forall x \in X \ f_n(x) \rightarrow f(x)$, ossia se

$$\forall x \in X \ \forall \varepsilon > 0 \ \exists \bar{n} : \forall n > \bar{n}, \ d_Y(f_n(x), f(x)) < \varepsilon .$$

In simboli si può indicare: $f_n \xrightarrow{P} f$.

Definizione 4.2 (Uniforme convergenza). Siano (X, d_X) e (Y, d_Y) due spazi metrici e sia $\{f_n\}_{n \in \mathbb{N}}$ una successione di funzioni $f_n : X \rightarrow Y$. Si dice che f_n converge uniformemente a f se

$$\forall \varepsilon > 0 \ \exists \bar{n} : \forall n > \bar{n}, \ \forall x \in X, \ d_Y(f_n(x), f(x)) < \varepsilon .$$

In simboli si può indicare: $f_n \xrightarrow{U} f$.

Osservazione 4.1. La convergenza uniforme implica la convergenza puntuale; la convergenza puntuale invece implica a priori la convergenza uniforme se e solo se X è finito.

Osservazione 4.2. Segue quasi immediatamente che la convergenza uniforme è equivalente alla condizione

$$\lim_{n \rightarrow +\infty} \sup\{d_Y(f_n(x), f(x)), \ x \in X\} = 0$$

Esempio 4.3. La successione $f_n : \mathbb{R} \rightarrow \mathbb{R}$, con

$$f_n(x) = \begin{cases} 1 & \text{se } x \in [n, n+1) \\ 0 & \text{altrimenti} \end{cases}$$

converge puntualmente ma non uniformemente.

Esempio 4.4. La successione $g_n : [0, 1] \rightarrow \mathbb{R}$, con

$$f_n(x) = \begin{cases} 1 & \text{se } x \in [\frac{1}{n+1}, \frac{1}{n}) \\ 0 & \text{altrimenti} \end{cases}$$

converge puntualmente ma non uniformemente.

Esempio 4.5. La successione $f_n : [0, 1] \rightarrow \mathbb{R}$, con $f_n(x) = x^n$ converge puntualmente ma non uniformemente.

Lemma 4.6: Convergenza monotona di Dini

Sia X uno spazio metrico compatto, siano $\{f_n\} \subseteq C(X, \mathbb{R})$ una successione e $f \in C(X, \mathbb{R})$ tali che $f_n \xrightarrow{P} f$ e tale convergenza è monotona ($f_n(x) \nearrow f(x)$ o $f_n(x) \searrow f(x)$). Allora $f_n \xrightarrow{U} f$.

Dimostrazione. Supponiamo, senza perdita di generalità, che $f_n(x)$ sia crescente, $f_n(x) \nearrow f(x)$. Definiamo la successione $g_n(x) = f(x) - f_n(x)$, che è continua perché somma di funzioni continue. Poiché $f_n(x)$ cresce a $f(x)$, vale $g_n(x) \geq 0$ e $g_n(x) \searrow 0$ per $n \rightarrow +\infty$. Fissato $\varepsilon > 0$, consideriamo l'insieme $E_n = \{x \in X \mid g_n(x) < \varepsilon\}$, che sono aperti poiché g_n è continua, e tali che $E_n \subseteq E_{n+1} \forall n$ per costruzione, visto che $g_n(x)$ è decrescente. Poiché $g_n(x)$ decresce a 0, risulta

$$\bigcup_{n=1}^{+\infty} E_n = X$$

e per compattezza di X esistono n_1, \dots, n_k tali che $E_{n_1} \cup E_{n_2} \cup \dots \cup E_{n_k} = X$, da cui segue $X = E_{\bar{n}}$, dove $\bar{n} = \max\{n_1, \dots, n_k\}$.

Quindi, $\forall n \geq \bar{n}$ $E_n = X$, ossia $0 \leq g_n \leq \varepsilon \forall n \geq \bar{n}$: poiché \bar{n} non dipende da x , si ha $f_n \xrightarrow{U} f$. \square

Osservazione 4.7. Le ipotesi non possono essere indebolite: ad esempio, nell'Esempio 4.5, lo spazio $[0, 1]$ è compatto, le f_n sono continue e la convergenza è monotona, tuttavia il limite è discontinuo in 1 e difatti non c'è convergenza uniforme. Analogamente, se restringiamo le f_n allo spazio $[0, 1)$, recuperiamo la continuità del limite, ma perdiamo la compattezza.

Esempio 4.8. Consideriamo la successione

$$f_n(x) = \left(1 + \frac{x}{n}\right)^n \quad \text{con} \quad \lim_{n \rightarrow +\infty} f_n(x) = e^x ;$$

come è noto, se $n > |x|$, $f_n(x)$ è crescente, perciò:

$$\forall a > 0 \quad f_n \xrightarrow{U} \exp \quad \text{in} \quad [-a, a] .$$

Teorema 4.9

Siano (X, d_X) e (Y, d_Y) due spazi metrici e sia $\{f_n\}_{n \in \mathbb{N}}$ una successione di funzioni $f_n : X \rightarrow Y$ continue tali che f_n converge uniformemente a f . Allora f è continua.

Dimostrazione. Sia $\bar{x} \in X$ e sia $\varepsilon > 0$. Dobbiamo trovare $\delta > 0$ tale che $\forall x \in X$ con $d_X(x, \bar{x}) < \delta$ si ha $d_Y(f(x), f(\bar{x})) < \varepsilon$. Sia \bar{n} tale che $d_Y(f(z), f_{\bar{n}}(z)) < \frac{\varepsilon}{3} \forall z \in X$ (che esiste perché f_n converge uniformemente a f), e sia $\delta > 0$ tale che $\forall z \in X$ $d_X(z, \bar{x}) < \delta \implies d_Y(f_{\bar{n}}(z), f_{\bar{n}}(\bar{x})) < \frac{\varepsilon}{3}$. Ora sia x tale che $d_X(x, \bar{x}) < \delta$; applicando la disuguaglianza triangolare, otteniamo:

$$d_Y(f(x), f(\bar{x})) \leq d_Y(f(x), f_{\bar{n}}(x)) + d_Y(f_{\bar{n}}(x), f_{\bar{n}}(\bar{x})) + d_Y(f_{\bar{n}}(\bar{x}), f(\bar{x})) < \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon .$$

\square

Proposizione 4.10

Sia $\{f_n\}_{n \in \mathbb{N}}$ una successione di funzioni continue $f_n : \mathbb{R} \rightarrow \mathbb{R}$ tali che f_n converge uniformemente a f . Allora

$$\int_a^b f_n(x) dx \longrightarrow \int_a^b f(x) dx .$$

Dimostrazione. Per il Teorema f è continua; valgono le seguenti disuguaglianze:

$$\left| \int_a^b f_n(x) - f(x) dx \right| \leq \int_a^b |f_n(x) - f(x)| dx \leq \int_a^b \|f_n - f\|_\infty dx = \|f_n - f\|_\infty (b - a) \xrightarrow{n \rightarrow +\infty} 0$$

da cui segue immediatamente la tesi. \square

Definizione 4.3 (Spazio $C(X, Y)$). Dati (X, d_X) e (Y, d_Y) spazi metrici, definiamo lo spazio metrico $C(X, Y)$ delle funzioni $g : X \rightarrow Y$ continue con la *distanza del sup*

$$d(f, g) = \sup_{x \in X} \{d_Y(f(x), g(x))\} \wedge 1 .$$

Si verifica che d è una distanza, infatti:

i., *ii.* seguono immediatamente;

iii. siano $f, g, h \in Z$: se $d(f, g) = 1$ o $d(g, h) = 1$, segue immediatamente; altrimenti, dato $x \in X$, vale $d_Y(f(x), h(x)) \leq d_Y(f(x), g(x)) + d_Y(g(x), h(x)) \leq d(f, g) + d(g, h)$, per cui, passando al sup del primo membro, $d(f, h) \leq \sup_{x \in X} \{d_Y(f(x), h(x))\} \leq d(f, g) + d(g, h)$.

La convergenza rispetto a d risulta essere proprio la convergenza uniforme.

Osservazione 4.11. Osserviamo che il fatto che $f \in C(X, Y)$ sia continua non ha giocato alcun ruolo nella verifica che d fosse una distanza: infatti questa costruzione può essere fatta con gli spazi di funzioni con una determinata proprietà.

Teorema 4.12

Lo spazio $C(X, Y)$ è completo $\iff Y$ è completo.

Dimostrazione. (\implies) Sia $\{y_n\} \subseteq Y$ una successione di Cauchy: la successione di funzioni costanti data da $f_n(x) \equiv y_n$ è anch'essa di Cauchy in $C(X, Y)$, dato che $d(f_n, f_m) = d_Y(y_n, y_m)$. Poiché $C(X, Y)$ è completo, $\{f_n\}$ converge uniformemente ad una funzione f , ossia $\forall \varepsilon > 0 \exists \bar{n} : \forall n > \bar{n} d(f_n, f) < \varepsilon$. Ora dato $x \in X$ qualsiasi, e detto $y = f(x)$, si ha

$$\lim_{n \rightarrow +\infty} y_n = \lim_{n \rightarrow +\infty} f_n(x) = f(x) = y .$$

(\impliedby) Sia $\{f_n\}$ una successione di Cauchy in $C(X, Y)$.

Per ogni $x \in X$ e ogni $n, m \in \mathbb{N}$, $d_Y(f_n(x), f_m(x)) \wedge 1 \leq d(f_n, f_m)$.

Sia $0 < \varepsilon < 1$ (scegliamo $\varepsilon < 1$ per evitare i casi in cui la distanza è stata effettivamente troncata a 1): esiste $\bar{n} = \bar{n}(\varepsilon)$ tale che $d(f_n, f_m) < \varepsilon \forall n, m \geq \bar{n}$ perché $\{f_n\}$ è di Cauchy.

Quindi $d_Y(f_n(x), f_m(x)) < \varepsilon \forall n, m \geq \bar{n}$, ossia $f_n(x)$ è di Cauchy in Y . Allora $f_n(x) \rightarrow f(x)$ perché Y è completo.

Ma se $n > \bar{n}$, allora

$$d_Y(f_n(x), f(x)) = \lim_{m \rightarrow +\infty} d_Y(f_n(x), f_m(x)) \implies d_Y(f_n(x), f(x)) \leq \varepsilon$$

cioè $f_n \xrightarrow{U} f$ poiché \bar{n} non dipende da x , e, poiché le f_n sono continue, $f \in C(X, Y)$. \square

Definizione 4.4 (Spazio $C^1(\mathbb{R}^n, \mathbb{R}^m)$). Definiamo lo spazio metrico $C^1(\mathbb{R}^n, \mathbb{R}^m)$ delle funzioni $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ di classe C^1 con la distanza

$$d(f, g) = \left(\sup_{x \in \mathbb{R}^n} \{|f(x) - g(x)|\} + \sup_{x \in \mathbb{R}^n} \{|Df(x) - Dg(x)|\} \right) \wedge 1 .$$

Analogamente alla norma del sup in $C(X, Y)$, si verifica che d è effettivamente una distanza.

Osservazione 4.13. Segue facilmente dalla definizione che, se $\{f_n\} \subseteq C^1(\mathbb{R}^n, \mathbb{R}^m)$, $f \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ e $f_n \xrightarrow{d} f$, allora in particolare $f_n \xrightarrow{U} f$ e $Df_n \xrightarrow{U} Df$.

Proposizione 4.14

Lo spazio $C^1(\mathbb{R}^n, \mathbb{R}^m)$ è completo.

Dimostrazione. Sia $\{f_n\} \subseteq C^1(\mathbb{R}^n, \mathbb{R}^m)$ una successione di Cauchy: allora in particolare $\{f_n\}$ è di Cauchy in $C(\mathbb{R}^n, \mathbb{R}^m)$ e $\{Df_n\}$ è di Cauchy in $C(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m))$, quindi $f_n \rightarrow f$ uniformemente e $Df_n \rightarrow g$ uniformemente. Rimane solo da mostrare che $f \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ e che $Df = g$.

Per farlo, scriviamo per Taylor:

$$f_n(x+v) = f_n(x) + Df_n(\xi_n)(v) = f_n(x) + Df_n(x)(v) + (Df_n(\xi_n) - Df_n(x))(v)$$

per un certo $\xi_n \in (x, x+v)$, quindi:

$$\begin{aligned} |f_n(x+v) - f_n(x) - Df_n(x)(v)| &= |(Df_n(\xi_n) - Df_n(x))(v)| \leq |Df_n(\xi_n) - Df_n(x)||v| \leq \\ &\leq (|Df_n(\xi_n) - g(\xi_n)| + |g(\xi_n) - g(x)| + |g(x) - Df_n(x)|)|v| \end{aligned}$$

Ora, poiché $Df_n \rightarrow g$ e g è continua, qualsiasi $\varepsilon > 0$ esistono \bar{n} e $r > 0$ tali che, per $|v| < r$ e per $n > \bar{n}$, ogni addendo del membro di sinistra della disuguaglianza è minore di ε , ossia

$$|f_n(x+v) - f_n(x) - Df_n(x)(v)| \leq 3\varepsilon|v| .$$

Allora, passando al limite per $n \rightarrow +\infty$, abbiamo

$$|f(x+v) - f(x) - g(x)(v)| \leq 3\varepsilon|v|$$

e passando al limite per $\varepsilon \rightarrow 0$ troviamo

$$f(x+v) = f(x) + g(x)(v) + o(|v|) .$$

Quindi $Df = g$ e $f \in C^1(\mathbb{R}^n, \mathbb{R}^m)$. □

In modo completamente analogo, si può generalizzare questo risultato per funzioni più regolari.

Definizione 4.5 (Spazio $C^k(\mathbb{R}^n, \mathbb{R}^m)$). Definiamo lo spazio metrico $C^k(\mathbb{R}^n, \mathbb{R}^m)$ delle funzioni $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ di classe C^k con la distanza

$$d(f, g) = \left(\sup_{x \in \mathbb{R}^n} \{|f(x) - g(x)|\} + \sum_{i=1}^k \sup_{x \in \mathbb{R}^n} \{|D^i f(x) - D^i g(x)|\} \right) \wedge 1 .$$

Si verifica sempre allo stesso modo che d è effettivamente una distanza e che vale il seguente risultato.

Teorema 4.15

Lo spazio $C^k(\mathbb{R}^n, \mathbb{R}^m)$ è completo.

4.2 Teorema di Ascoli-Arzelà

Definizione 4.6 (Equicontinuità). Una famiglia di funzioni $\{f_i\}_{i \in I}$, con $f_i : X \rightarrow Y$, si dice *equicontinua in x_0* se

$$\forall x \in X \quad \forall \varepsilon > 0 \quad \exists \delta > 0 : d_X(x_0, x) < \delta \implies d_Y(f_i(x_0), f_i(x)) < \varepsilon \quad \forall i \in I$$

Definizione 4.7 (Equilimitatezza). Una famiglia di funzioni $\{f_i\}_{i \in I}$, con $f_i : X \rightarrow \mathbb{R}^m$, si dice *equilimitata* se

$$\exists M \quad \forall x \in X : |f_i(x)| < M \quad \forall i \in I$$

Teorema 4.16: Teorema di Ascoli-Arzelà

Sia (X, d) spazio metrico compatto e sia $\{f_n\} \subseteq C(X, \mathbb{R}^m)$ una successione equicontinua ed equilimitata.

Allora, a meno di sottosuccessioni, $f_n \xrightarrow{U} f$.

Dimostrazione. Osserviamo preliminarmente che, dato $x \in X$, la successione $f_n(x)$ è in \mathbb{R}^m ed è limitata (poiché f_n è equilimitata), quindi ammette una sottosuccessione convergente.

Ora, fissato $x \in X$ e dato $\varepsilon > 0$, esiste $\delta > 0$ tale che $d(x, y) < \delta \implies d(f_n(x), f_n(y)) < \varepsilon \quad \forall n$ poiché f_n è equicontinua. Allora, preso \tilde{x} con $d(x, \tilde{x}) < \frac{\delta}{2}$, se $d(\tilde{x}, y) < \frac{\delta}{2}$, vale che $d(x, y) < \delta$, $d(f_n(x), f_n(y)) < \varepsilon$ e quindi $d(f_n(\tilde{x}), f_n(y)) < 2\varepsilon$. Quindi, fissato $\varepsilon > 0$, trovo per ogni $x \in X$ un intorno aperto di x tale che uno stesso δ_x soddisfa la definizione di continuità per ε in ogni punto dell'intorno.

In particolare questi aperti ricoprono X (ogni punto appartiene all'intorno centrato in sé), e siccome X è compatto, esistono finiti intorni che ricoprono X , centrati in $x_1^\varepsilon, x_2^\varepsilon, \dots, x_k^\varepsilon$, dove $k = k(\varepsilon)$. Ripetendo questa operazione per $\varepsilon = 2^{-n}$ per $n = 0, 1, 2, \dots$, otteniamo un insieme numerabile di punti $\{x_j^\varepsilon\}$ (unione numerabile di finiti è numerabile).

A questo punto possiamo costruire una sottosuccessione di f_n tale che $f_n(x) \rightarrow f(x) \quad \forall x \in \{x_j^\varepsilon\}$ (tale fatto è già stato dimostrato in precedenza). Infatti, enumerando i punti di $\{x_j^\varepsilon\}$, possiamo costruire per ricorsione

$\{f_n^{(1)}\}$ sottosuccessione di f_n per $n \geq 1$, tale che $f_n^{(1)}(x_1)$ converge ad un valore, che denominiamo $f(x_1)$;

$\{f_n^{(2)}\}$ sottosuccessione di $f_n^{(1)}$ per $n \geq 2$, tale che $f_n^{(2)}(x_2)$ converge ad un valore, che denominiamo $f(x_2)$;

⋮

$\{f_n^{(h+1)}\}$ sottosuccessione di $f_n^{(h)}$ per $n \geq h+1$, tale che $f_n^{(h+1)}(x_{h+1})$ converge ad un valore, che denominiamo $f(x_{h+1})$.

Quindi la sottosuccessione finale, che ha come elemento k -esimo $f_k^{(k)}$, converge $\forall x \in \{x_j^\varepsilon\}$; per semplicità, chiamiamo ancora f_n questa sottosuccessione. Mostriamo allora che f_n è una successione di Cauchy. Fissato $\tilde{\varepsilon} > 0$, esiste $\varepsilon \in \{2^{-n}, n \in \mathbb{N}\}$ tale che $\varepsilon < \frac{\tilde{\varepsilon}}{8}$; allora, poiché gli x_j^ε sono finiti, esiste \bar{n} tale che $\forall n > \bar{n} \quad d(f_n(x_j^\varepsilon), f(x_j^\varepsilon)) < \varepsilon \quad \forall j = 1, \dots, k(\varepsilon)$. Ma gli intorni centrati in x_j^ε

ricoprono X , quindi, dato $x \in X$, esiste j tale che $x \in B(x_j^\varepsilon, \delta_j^\varepsilon)$, quindi $d(f_n(x), f_n(x_j^\varepsilon)) < \varepsilon \forall n$. Quindi, $\forall n, m > \bar{n}$, vale

$$\begin{aligned} d(f_n(x), f_m(x)) &\leq d(f_n(x), f_n(x_j^\varepsilon)) + d(f_n(x_j^\varepsilon), f(x_j^\varepsilon)) + d(f(x_j^\varepsilon), f_m(x_j^\varepsilon)) + d(f_m(x_j^\varepsilon), f_m(x)) \\ &< 4\varepsilon < \frac{\varepsilon}{2}. \end{aligned}$$

Quindi $d(f_n, f_m) = \sup\{d(f_n(x), f_m(x)) \mid x \in X\} < \tilde{\varepsilon}$.

Quindi, visto che f_n è di Cauchy e \mathbb{R}^m è completo, f_n converge in $C(X, \mathbb{R}^m)$ ad una funzione, che denominiamo $f \in C(X, \mathbb{R}^m)$, dunque $f_n \rightarrow f$ uniformemente. \square

Definizione 4.8 (Relativa compattezza). Dato X spazio topologico, un sottospazio $Y \subseteq X$ si dice *relativamente compatto* se \bar{Y} è compatto.

Corollario 4.17

Sia X compatto e sia $\mathcal{F} \subseteq C(X, \mathbb{R}^m)$ una famiglia di funzioni equicontinua ed equilimitata. Allora \mathcal{F} è relativamente compatto.

Dimostrazione. Mostriamo che $\bar{\mathcal{F}}$ è compatto per successioni. Sia $\{f_n\} \subseteq \bar{\mathcal{F}}$ una successione, e per ogni n sia $g_n \in \mathcal{F}$ tale che $\|f_n - g_n\|_\infty < 2^{-n}$, che esiste per definizione di chiusura. La successione $\{g_n\} \subseteq \mathcal{F}$ è equicontinua ed equilimitata, di conseguenza per il teorema di Ascoli-Arzelà esiste una sottosuccessione $\{g_{n_k}\}$ che converge uniformemente ad una funzione $f \in \bar{\mathcal{F}}$. Mostriamo che la sottosuccessione $\{f_{n_k}\}$ converge uniformemente a f : dato $\varepsilon > 0$, sia \bar{k}_1 tale che $\forall k > \bar{k}_1$ si abbia $2^{-n_k} < \frac{\varepsilon}{2}$ e sia \bar{k}_2 tale che $\forall k > \bar{k}_2$ valga $\|g_{n_k} - f\|_\infty < \frac{\varepsilon}{2}$. Detto allora $\bar{k} = \max\{\bar{k}_1, \bar{k}_2\}$, $\forall k > \bar{k}$ vale

$$\|f_{n_k} - f\|_\infty \leq \|f_{n_k} - g_{n_k}\|_\infty + \|g_{n_k} - f\|_\infty < 2^{-n_k} + \|g_{n_k} - f\|_\infty < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

ossia $f_{n_k} \rightarrow f$ uniformemente, con $f \in \bar{\mathcal{F}}$. \square

Teorema 4.18: Teorema di Ascoli-Arzelà generalizzato

Siano (X, d_X) compatto e (Y, d_Y) completo, e sia $\{f_n\} \subseteq C(X, Y)$ una successione equicontinua e tale che $\forall x \in X \{f_n(x)\}$ sia relativamente compatto.

Allora, a meno di sottosuccessioni, $f_n \xrightarrow{U} f$.

Dimostrazione. Ripercorriamo la stessa dimostrazione del teorema di Ascoli-Arzelà:

- i.* Per ogni x troviamo una sottosuccessione di $\{f_n\}$ tale che $\{f_n(x)\}$ è convergente: nella dimostrazione di Ascoli-Arzelà abbiamo usato l'equilimitatezza delle f_n , ma è sufficiente che l'insieme $\{f_n(x)\}$ sia relativamente compatto, infatti in tal caso $\{f_n(x)\}$ è compatto per successioni e dunque ammette una sottosuccessione convergente.
- ii.* Per ogni ε e per ogni x troviamo un intorno aperto di x in cui tutte le f_n hanno oscillazione minore di ε : per questo passaggio abbiamo usato solo l'equicontinuità.
- iii.* Per ogni $\varepsilon \in \{2^{-n}\}_{n \in \mathbb{N}}$ selezioniamo finiti $x_1^\varepsilon, \dots, x_{k(\varepsilon)}^\varepsilon$, troviamo dunque un insieme numerabile di punti x_j^ε e costruiamo una sottosuccessione di $\{f_n\}$ (che chiamiamo ancora $\{f_n\}$) tale che $f_n(x_j^\varepsilon) \rightarrow f(x_j^\varepsilon)$ per ognuno degli x_j^ε : in questo punto usiamo la compattezza di X .
- iv.* La sottosuccessione delle f_n risulta essere una successione di Cauchy e in quanto tale converge: per giustificare quest'ultimo passaggio è necessaria la completezza di $C(X, Y)$, che abbiamo già dimostrato essere equivalente alla completezza di Y .

\square

Adattiamo nuovamente questa dimostrazione per provare una forma ancora più generale.

Teorema 4.19

Siano X unione numerabile di compatti e Y completo, e sia $\{f_n\} \subseteq C(X, Y)$ una successione equicontinua e tale che $\forall x \in X \{f_n(x)\}$ sia relativamente compatto. Allora esiste una sottosuccessione di $\{f_n\}$ che converge puntualmente e che converge uniformemente sui compatti.

Dimostrazione. Sia $\{X_n\}_{n \in \mathbb{N}}$ l'insieme dei compatti tali che $X = \bigcup X_n$. La successione $\{f_n\}$ è equicontinua, di conseguenza fissato $x \in X$ e dato $\varepsilon > 0$ esiste un intorno U_x tale che per ciascuna delle f_n uno stesso δ soddisfa la definizione di continuità per ε in ogni punto dell'intorno. Fissando ε e lasciando variare $x \in X$ troviamo che

$$\bigcup_{x \in X} U_x = \bigcup_{n \in \mathbb{N}} X_n = X \quad \text{dunque } \forall n \quad X_n \subseteq \bigcup_{x \in X} U_x ;$$

per compattezza di X_n esistono finiti $x_1^{\varepsilon, n}, \dots, x_k^{\varepsilon, n}$ tali che i relativi intorni ricoprono X_n , e dunque esistono numerabili x_m^ε tali che i relativi intorni ricoprono X . L'insieme $\{x_m^\varepsilon \mid \varepsilon = 2^{-n}\}_{m, n \in \mathbb{N}}$ è allora numerabile in quanto unione numerabile di numerabili, pertanto a meno di successioni $\{f_n\}$ converge puntualmente sugli x_m^ε : denotiamo

$$f(x_m^\varepsilon) = \lim_{n \rightarrow +\infty} f_n(x_m^\varepsilon) \quad \forall x_m^\varepsilon.$$

Adesso scelto $x \in X$ e fissato ε , sia x_m^ε tale che $x \in U_{x_m^\varepsilon}$, dunque $d(f_n(x), f_n(x_m^\varepsilon)) < \varepsilon$ e definitivamente $d(f_n(x_m^\varepsilon), f(x_m^\varepsilon)) < \varepsilon$, pertanto $d(f_n(x), f(x_m^\varepsilon)) < 2\varepsilon$. A questo punto si conclude in modo analogo al teorema di Ascoli-Arzelà: fissato x , la successione $\{f_n(x)\} \subseteq Y$ è una successione di Cauchy e per completezza di Y converge ad un punto che denotiamo $f(x)$, ossia $f_n \rightarrow f$ puntualmente; dato invece $K \subseteq X$ compatto e dato ε , esistono finiti $x_{m_1}^\varepsilon, \dots, x_{m_k}^\varepsilon$ tali che gli $U_{x_{m_\ell}^\varepsilon}$ ricoprono K , di conseguenza la successione $\{f_n|_K\}$ è di Cauchy e per completezza di $C(X, Y)$ converge uniformemente. \square

4.3 Serie di funzioni

Proposizione 4.20

Data una successione $\{f_n\} \subseteq C^1(\mathbb{R}^n, \mathbb{R}^m)$, se $\sum_{n=0}^{+\infty} f_n$ e $\sum_{n=0}^{+\infty} Df_n$ convergono uniformemente, allora vale

$$D \left(\sum_{n=0}^{+\infty} f_n \right) = \sum_{n=0}^{+\infty} Df_n .$$

Dimostrazione. Consideriamo i limiti uniformi

$$f = \sum_{n=0}^{+\infty} f_n \in C(\mathbb{R}^n, \mathbb{R}^m) \quad \text{e} \quad g = \sum_{n=0}^{+\infty} Df_n \in C(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)) ;$$

fissiamo $x \in \mathbb{R}^n$ e $\sigma > 0$. Poiché g è continua, esiste ε tale che $|\xi - x| < \varepsilon \implies |g(\xi) - g(x)| < \sigma$. Sia ora $v \in \mathbb{R}^n$ con $|v| < \varepsilon$. Per la convergenza uniforme delle due serie, esiste \bar{n} tale che

$$\left| f(z) - \sum_{n=0}^{\bar{n}} f_n(z) \right| < \sigma |v| \quad \text{e} \quad \left| g(z) - \sum_{n=0}^{\bar{n}} Df_n(z) \right| < \sigma \quad \forall z \in \mathbb{R}^n.$$

Abbiamo allora:

$$\begin{aligned} \frac{f(x+v) - f(x)}{|v|} &= \frac{1}{|v|} \sum_{n=0}^{\bar{n}} [f_n(x+v) - f_n(x)] + \underbrace{\frac{1}{|v|} \sum_{n=\bar{n}+1}^{+\infty} [f_n(x+v) - f_n(x)]}_{E_1} = \\ &= \frac{1}{|v|} \sum_{n=0}^{\bar{n}} Df_n(\xi)(v) + E_1 = \frac{1}{|v|} g(\xi)(v) + \underbrace{\frac{1}{|v|} \left(\sum_{n=0}^{\bar{n}} Df_n(\xi)(v) - g(\xi)(v) \right)}_{E_2} + E_1 = \\ &= \frac{1}{|v|} g(x)(v) + \underbrace{\frac{1}{|v|} (g(\xi)(v) - g(x)(v))}_{E_3} + E_1 + E_2 = \frac{1}{|v|} g(x) + E_1 + E_2 + E_3 . \end{aligned}$$

Ora, E_1 è la coda della serie

$$\frac{1}{|v|} \sum_{n=0}^{+\infty} [f_n(x+v) - f_n(x)]$$

che, per quanto detto, è stimabile in modulo con 2σ . Inoltre, per la scelta di v , valgono $|E_2| < \sigma$ e $|E_3| < \sigma$.

Quindi abbiamo:

$$\frac{|f(x+v) - f(x) - g(x)(v)|}{|v|} < 4\sigma .$$

Perciò, passando al limite per $\sigma \rightarrow 0$, otteniamo che $Df = g$. □

4.3.1 Serie di potenze

Una famiglia fondamentale di serie di funzioni sono le serie di potenze, ossia serie della forma

$$S(x) = \sum_{n=0}^{+\infty} a_n x^n .$$

La regione di definizione di queste serie è caratterizzata dal loro **raggio di convergenza** R , ossia la quantità definita come

$$R = \liminf \frac{1}{\sqrt[n]{|a_n|}} \quad \text{oppure} \quad \frac{1}{R} = \limsup \sqrt[n]{|a_n|} .$$

Riportiamo il risultato principale di Analisi I riguardo le serie di potenze.

Teorema 4.21

Data una serie di potenze

$$S(x) = \sum_{n=0}^{+\infty} a_n x^n$$

con raggio di convergenza $R > 0$, valgono le seguenti implicazioni:

- i.* se $|x| < R$, allora $S(x)$ converge (assolutamente);
- ii.* se $|x| > R$, allora $S(x)$ non converge.

Non possiamo invece stabilire a priori il comportamento dei punti x tali che $|x| = R$, che devono essere studiati separatamente. In generale inoltre la convergenza non è uniforme sull'intera regione $(-R, R)$, tuttavia è uniforme su ogni compatto contenuto in questo intervallo.

Proposizione 4.22

Data una serie di potenze

$$S(x) = \sum_{n=0}^{+\infty} a_n x^n$$

con raggio di convergenza $R > 0$, S converge uniformemente su ogni intervallo $[-\ell, \ell]$, con $0 < \ell < R$.

Dimostrazione. Sia $\ell \in (0, R)$, allora se $|x| \leq \ell$ vale $|a_n x^n| \leq |a_n| \ell^n$.

Se S_n indica il troncamento della serie al termine n -esimo, abbiamo che

$$S(x) - S_n(x) = \sum_{k=n+1}^{+\infty} a_k x^k ;$$

ora per il Teorema precedente la serie converge assolutamente all'interno del raggio di convergenza, pertanto, fissato $x = \ell$, per ogni $\varepsilon > 0$ esiste un \bar{n} tale che per $n > \bar{n}$ si ha

$$|S(\ell) - S_n(\ell)| = \left| \sum_{k=n+1}^{+\infty} a_k \ell^k \right| \leq \sum_{k=n+1}^{+\infty} |a_k| \ell^k < \varepsilon .$$

Ma allora se $n > \bar{n}$ vale

$$|S(x) - S_n(x)| \leq \sum_{k=n+1}^{+\infty} |a_k| x^k \leq \sum_{k=n+1}^{+\infty} |a_k| \ell^k < \varepsilon \quad \forall x \in [-\ell, \ell].$$

Dunque la convergenza di $S_n(x)$ a $S(x)$ è uniforme sull'intervallo detto, poiché \bar{n} dipende solo da ε e non da x . \square

Osservazione 4.23. Una serie di potenze e la sua serie derivata hanno lo stesso raggio di convergenza. Infatti, date le serie di potenze

$$S(x) = \sum_{n=0}^{+\infty} a_n x^n \quad \text{e} \quad S'(x) = \sum_{n=1}^{+\infty} a_n n x^{n-1} = \sum_{n=0}^{+\infty} b_n x^n ,$$

dove $b_n = a_{n+1}(n+1)$, se $R > 0$ è il raggio di convergenza della serie S allora il raggio di convergenza di S' è

$$R' = \liminf_{n \rightarrow +\infty} \frac{1}{\sqrt[n]{a_{n+1}(n+1)}} = \lim_{n \rightarrow +\infty} \frac{1}{\sqrt[n]{n+1}} \liminf_{n \rightarrow +\infty} \frac{1}{\sqrt[n]{a_{n+1}}} = \liminf_{n \rightarrow +\infty} \frac{1}{\sqrt[n]{a_{n+1}}} = R .$$

4.3.2 Convergenza totale

Definizione 4.9 (Totale convergenza). Dati (X, d) spazio metrico, $(E, \|\cdot\|)$ spazio di Banach e una successione $\{u_k\} \subseteq C(X, E)$, si dice che la serie

$$\sum_{k=0}^{+\infty} u_k(x)$$

è *totalmente* (o *normalmente*) *convergente* se $\exists \{M_k\} \subseteq \mathbb{R}$ tale che

$$\|u_k\|_\infty \leq M_k \quad \forall k \quad \text{e} \quad \sum_{k=0}^{+\infty} M_k < +\infty , \quad \text{dove} \quad \|u_k\|_\infty = \sup_{x \in X} \|u_k(x)\| .$$

Proposizione 4.24: Criterio di convergenza totale (Weierstraß)

Dati (X, d) spazio metrico e $(E, \|\cdot\|)$ spazio di Banach, sia $\{u_k\} \subseteq C(X, E)$ una successione: allora se la serie

$$S(x) = \sum_{k=0}^{+\infty} u_k(x)$$

converge totalmente, la convergenza è anche uniforme.

Dimostrazione. Consideriamo la successione delle somme parziali

$$S_n(x) = \sum_{k=0}^n u_k(x)$$

e mostriamo che $\forall x \in X$ la successione $\{S_n(x)\}$ è di Cauchy. Per ipotesi, esiste $\{M_k\} \subseteq \mathbb{R}$ con $\|u_k\|_\infty \leq M_k \forall k$ e tale che la serie delle M_k sia sommabile. Allora,

$$\text{dato } \varepsilon > 0 \exists \bar{n} \text{ tale che } \sum_{k=\bar{n}}^{+\infty} M_k < \varepsilon .$$

Siano dunque $n, m > \bar{n}$ e supponiamo senza perdita di generalità $n \geq m$; allora $\forall x \in X$ vale la seguente stima:

$$\|S_n(x) - S_m(x)\| = \left\| \sum_{k=m+1}^n u_k(x) \right\| \leq \sum_{k=m+1}^n \|u_k(x)\| \leq \sum_{k=m+1}^n M_k \leq \sum_{k=\bar{n}}^{+\infty} M_k < \varepsilon .$$

Dato che $\forall \varepsilon > 0 \exists \bar{n}$ tale che $\forall n, m > \bar{n}$ vale $\|S_n(x) - S_m(x)\| < \varepsilon$, la successione degli $S_n(x)$ è di Cauchy $\forall x \in X$, e per completezza di E la successione S_n converge puntualmente ad una funzione S . Poiché \bar{n} non dipende da x , $\forall n, m > \bar{n}$ vale $\|S_n - S_m\|_\infty \leq \varepsilon$, ossia $\{S_n\}$ è una successione di Cauchy e quindi per completezza converge in $C(X, E)$, ossia $S_n \rightarrow S$ uniformemente. \square

4.4 Teorema di Stone-Weierstraß

Nel 1872 il matematico tedesco Karl Weierstraß pubblica in un articolo le sue scoperte circa una famiglia di funzioni di variabile reale continue in ogni punto ma derivabili in nessuno: prima di allora ogni funzione continua conosciuta era derivabile a meno di un insieme di punti isolati del dominio. Circa dieci anni dopo, Weierstraß si interessa alla possibilità di approssimare qualsiasi funzione continua con mappe derivabili infinite volte, giungendo al seguente risultato.

Teorema 4.25: Teorema di approssimazione di Weierstraß

Sia $f \in C([0, 1], \mathbb{R})$. Allora $\forall \varepsilon > 0 \exists p(x) \in \mathbb{R}[x]$ tale che $d(f, p) < \varepsilon$.

Dimostreremo nelle prossime pagine una versione più generale del Teorema. Si potrebbe pensare di cercare il polinomio p tra i polinomi di Taylor di f , ma, per l'appunto, f potrebbe non essere derivabile, e se anche f fosse liscia, potrebbe comunque succedere che la serie di Taylor di f non converga ad f ; similmente si potrebbe pensare di cercare p tra i polinomi interpolatori di Lagrange, ossia fissando n nodi $(x_i, f(x_i))$ e ponendo

$$p(x) = \sum_{i=1}^n f(x_i) \prod_{j \neq i} \frac{x - x_j}{x_i - x_j} .$$

Entrambi questi metodi risultano piuttosto fallimentari in quanto l'approssimazione a cui danno origine è solo locale e spesso decisamente non uniforme.

Definizione 4.10 (Algebra su campo). Sia $(V, +, \cdot)$ uno spazio vettoriale su un campo \mathbb{K} , e sia $*$: $V \times V \rightarrow V$ un'operazione binaria bilineare, ossia:

- i. $(x + y) * z = x * z + y * z \quad \forall x, y, z \in V$;
- ii. $x * (y + z) = x * y + x * z \quad \forall x, y, z \in V$;
- iii. $(ax) * y = a(x * y) \quad \forall x, y \in V, \forall a \in \mathbb{K}$;
- iv. $x * (ay) = a(x * y) \quad \forall x, y \in V, \forall a \in \mathbb{K}$.

La quadrupla $(V, +, \cdot, *)$ si dice *algebra* sul campo \mathbb{K} o \mathbb{K} -*algebra*.

Nel seguito, chiameremo semplicemente algebre le \mathbb{R} -algebre di funzioni contenute in \mathbb{R}^X con le operazioni di somma, prodotto per scalare e prodotto di funzioni di \mathbb{R}^X .

Definizione 4.11 (Insieme separante). Un insieme di funzioni $\mathcal{F} \subseteq \mathbb{K}^X$ si dice *separante* o *che separa i punti* se $\forall x, x' \in X, x \neq x', \exists f \in \mathcal{F}$ tale che $f(x) \neq f(x')$.

Lemma 4.26

Dato X spazio metrico compatto, sia $\mathcal{F} \subseteq C(X, \mathbb{R})$ un'algebra separante con $1 \in \overline{\mathcal{F}}$ e tale che $\forall f, g \in \mathcal{F}$ vale $f \wedge g$ e $f \vee g \in \overline{\mathcal{F}}$. Allora $\overline{\mathcal{F}} = C(X, \mathbb{R})$.

Dimostrazione. Data $\varphi \in C(X, \mathbb{R})$, sia $\varepsilon > 0$. Poiché l'algebra \mathcal{F} è separante, $\forall z, w \in X$ con $z \neq w \exists f \in \mathcal{F}$ tale che $f(z) \neq f(w)$; allora $\exists f_{zw} \in \overline{\mathcal{F}}$ tale che f_{zw} coincide con φ in z e w : basta considerare la funzione

$$f_{zw}(x) = \varphi(z) + (\varphi(w) - \varphi(z)) \frac{f(x) - f(z)}{f(w) - f(z)} .$$

A tal proposito possiamo considerare l'insieme $A(z, w) = \{x \in X \mid f_{zw}(x) > \varphi(x) - \varepsilon\}$; dato che φ e f_{zw} sono continue, l'insieme $A(z, w)$ è aperto, che chiaramente contiene z e w . Fissando dunque z e facendo variare w in $X \setminus \{z\}$, la famiglia $\{A(z, w)\}_{w \in X \setminus \{z\}}$ è un ricoprimento di X , e pertanto per compattezza di X

$$\exists w_1, \dots, w_n \in X \setminus \{z\} \quad \text{tali che} \quad \bigcup_{i=1}^n A(z, w_i) = X .$$

Definiamo la funzione

$$g_z(x) := \max_{i=1, \dots, n} f_{zw_i}(x)$$

che per ipotesi¹ appartiene a $\overline{\mathcal{F}}$: è chiaro che $\forall x \in X$ vale $g_z(x) > \varphi(x) - \varepsilon$ dato che g_z è massimo di funzioni con questa proprietà. Inoltre $g_z(z) = \varphi(z)$, e ciò vale per ogni scelta di z in X : facendo variare z su tutto X possiamo definire $B(z) = \{x \in X \mid g_z(x) < \varphi(x) + \varepsilon\}$ e dalla continuità di g_z , la famiglia $\{B(z)\}_{z \in X}$ è a sua volta un ricoprimento di X in aperti. Per concludere

$$\exists z_1, \dots, z_k \in X \quad \text{tali che} \quad \bigcup_{i=1}^k B(z_i) = X ,$$

pertanto definendo

$$h(x) := \min_{i=1, \dots, k} g_{z_i}(x),$$

con $h \in \overline{\mathcal{F}}$, si ha la relazione $\forall x \in X \quad \varphi(x) - \varepsilon < h(x) < \varphi(x) + \varepsilon$, ossia $d(\varphi, h) \leq \varepsilon$. \square

¹Si può provare per induzione che approssimare il massimo tra due funzioni equivale ad approssimare il massimo di un numero finito di funzioni; discorso analogo vale per il minimo.

Osserviamo che l'ipotesi $1 \in \overline{\mathcal{F}}$ è equivalente a chiedere che l'algebra approssimi le funzioni costanti.

Lemma 4.27

Dato X spazio metrico compatto, sia $\mathcal{F} \subseteq C(X, \mathbb{R})$ un'algebra separante con $1 \in \overline{\mathcal{F}}$ e tale che $\forall f \in \mathcal{F}$ vale $|f| \in \overline{\mathcal{F}}$. Allora $\overline{\mathcal{F}} = C(X, \mathbb{R})$.

Dimostrazione. Date $f, g \in \mathcal{F}$, si ha

$$f \vee g = \frac{f + g + |f - g|}{2} \quad \text{e} \quad f \wedge g = \frac{f + g - |f - g|}{2}.$$

□

Lemma 4.28

Dato X spazio metrico compatto, sia $\mathcal{F} \subseteq C(X, \mathbb{R})$ un'algebra separante con $1 \in \overline{\mathcal{F}}$ e tale che $\forall f \in \mathcal{F}$ con $f > 0$ vale $\sqrt{f} \in \overline{\mathcal{F}}$. Allora $\overline{\mathcal{F}} = C(X, \mathbb{R})$.

Dimostrazione. Basta osservare che $\forall f \in \mathcal{F}$ vale $|f| = \sqrt{f^2}$.

□

Osservazione 4.29. Data un'algebra $\mathcal{F} \subseteq C(X, \mathbb{R})$, la sua chiusura $\overline{\mathcal{F}}$ è a sua volta un'algebra.

Teorema 4.30: Teorema di Stone-Weierstraß

Dato X spazio metrico compatto, sia $\mathcal{F} \subseteq C(X, \mathbb{R})$ un'algebra separante con $1 \in \overline{\mathcal{F}}$. Allora $\overline{\mathcal{F}} = C(X, \mathbb{R})$.

Dimostrazione. Consideriamo la funzione $\varphi : [0, 1] \rightarrow \mathbb{R}$ dove $\varphi(t) = \sqrt{1-t}$. Essa è di classe C^∞ in un intorno di 0, di conseguenza possiamo prendere in analisi la sua serie di Taylor: abbiamo

$$\begin{aligned} \varphi'(t) &= -\frac{1}{2}(1-t)^{-\frac{1}{2}}, & \varphi''(t) &= -\frac{1}{4}(1-t)^{-\frac{3}{2}}, \\ \varphi^{(n)} &= -\frac{(2n-3)!!}{2^n}(1-t)^{-\frac{2n-1}{2}} \quad \text{per } n \geq 2, \end{aligned}$$

dove l'ultima relazione può essere provata per induzione.

A questo punto possiamo scrivere lo sviluppo di Taylor all'ordine k in 0 della funzione φ :

$$\varphi(t) = P_k(t) + R_k(t) \quad \text{dove}$$

$$P_k(t) = \sum_{n=0}^k \alpha_n t^n \quad \text{e} \quad R_k(t) = \alpha_{k+1} \frac{t^{k+1}}{(1-\xi_k)^{\frac{2k+1}{2}}}$$

per qualche $\xi_k \in (0, 1)$, e abbiamo posto

$$\alpha_0 = 1, \quad \alpha_1 = -\frac{1}{2} \quad \text{e} \quad \alpha_n = -\frac{(2n-3)!!}{2^n n!} \quad \text{per } n \geq 2.$$

L'osservazione alla base della dimostrazione è che se $P_k \rightarrow \varphi$ uniformemente su tutto l'intervallo $[0, 1]$, allora vale la tesi di Stone-Weierstraß: infatti data $f \in \mathcal{F}$, $f \geq 0$, supponiamo $0 \leq f \leq 1$, allora

$$\forall k \quad \sum_{n=0}^k \alpha_n (1-f)^n \in \overline{\mathcal{F}};$$

inoltre $\forall x$ ponendo $t = 1 - f(x)$ troviamo

$$\sum_{n=0}^k \alpha_n (1 - f(x))^n = \sum_{n=0}^k \alpha_n t^n \xrightarrow{k \rightarrow +\infty} \sqrt{1-t} = \sqrt{f(x)}$$

e la convergenza è uniforme (dipende da k ma non da x), dunque \sqrt{f} ha distanza zero da $\overline{\mathcal{F}}$, e pertanto $\sqrt{f} \in \overline{\mathcal{F}}$. Più in generale, data $f \in \mathcal{F}$ con $f \geq 0$, sia $M = \max f$ (che esiste per compattezza di X): dato che $f/M \in \mathcal{F}$ e $0 \leq f/M \leq 1$, per quanto appena osservato

$$\sqrt{\frac{f}{M}} \in \overline{\mathcal{F}} \implies \sqrt{M} \cdot \sqrt{\frac{f}{M}} = \sqrt{f} \in \overline{\mathcal{F}}$$

e si conclude per il Lemma.

Mostriamo quindi che $P_k \rightarrow \varphi$ uniformemente. Innanzitutto, la convergenza è uniforme in $[0, \frac{1}{2}]$, infatti su tale intervallo è semplice mostrare che $R_k(t) \xrightarrow{k \rightarrow +\infty} 0$: per $t \in [0, \frac{1}{2}]$ vale

$$R_k(t) = -\frac{(2k-1)!!}{2^{k+1}(k+1)!} \frac{t^{k+1}}{(1-\xi_k)^{\frac{2k+1}{2}}}$$

con $\xi_k \in [0, \frac{1}{2}]$, pertanto $1 - \xi_k \in [\frac{1}{2}, 1]$ e

$$|R_k(t)| \leq \frac{(2k-1)!!}{2^{k+1}(k+1)!} \frac{(\frac{1}{2})^{k+1}}{(\frac{1}{2})^{k+\frac{1}{2}}} = \frac{(2k-1)!!}{2^{k+1}(k+1)!} \sqrt{\frac{1}{2}} < \frac{(2k-1)!!}{2^{k+1}(k+1)!} = |\alpha_{k+1}|$$

ma $|\alpha_k|$ converge a 0

$$|\alpha_k| = \frac{(2k-3)!!}{2^k k!} = \frac{(2k-3)!!}{(2k-2)!!} \cdot \frac{1}{2k} \leq \frac{1}{2k} \xrightarrow{k \rightarrow +\infty} 0.$$

Inoltre $P_k \rightarrow g$ puntualmente per una certa g sull'intervallo $[0, 1)$, possiamo difatti calcolare il raggio di convergenza con il criterio del rapporto:

$$\frac{\alpha_k}{\alpha_{k+1}} = \frac{(2k-3)!!}{2^k k!} \cdot \frac{2^{k+1}(k+1)!}{(2k-1)!!} = \frac{2(k+1)}{2k-1} \xrightarrow{k \rightarrow +\infty} 1.$$

Dimostriamo che c'è convergenza puntuale anche per $t = 1$, provando per induzione che $|\alpha_n| \leq \frac{1}{n^{3/2}}$.

Per $n = 2$ otteniamo $\frac{1}{8} \leq \frac{1}{\sqrt{8}}$, che è vera; adesso

$$|\alpha_{n+1}| = \frac{|\alpha_n|(2n-1)}{2(n+1)} \stackrel{I.I.}{\leq} \frac{2n-1}{2(n+1)} \cdot \frac{1}{n^{3/2}} \quad \text{e}$$

$$\frac{2n-1}{2(n+1)} \cdot \frac{1}{n^{3/2}} \leq \frac{1}{(n+1)^{3/2}} \iff \left(1 + \frac{1}{n}\right)^{3/2} \leq \frac{2(n+1)}{2n-1} = 1 + \frac{3}{2n-1}$$

che è vera perché stimando il termine di sinistra con il suo sviluppo di Taylor al secondo ordine troviamo

$$\left(1 + \frac{1}{n}\right)^{3/2} - \left[1 + \frac{3}{2n} + \frac{3}{8n^2}\right] = -\frac{1}{16}\zeta_n^3 \quad \text{con } \zeta_n \in [0, \frac{1}{n}], \text{ dunque}$$

$$\left(1 + \frac{1}{n}\right)^{3/2} \leq 1 + \frac{3}{2n} + \frac{3}{8n^2} \quad \text{e} \quad \cancel{\lambda} + \frac{\cancel{\beta}}{2n} + \frac{\cancel{\beta}}{8n^2} \leq \cancel{\lambda} + \frac{\cancel{\beta}}{2n-1} \iff 8n^2 - 2n - 1 \leq 8n^2.$$

A questo punto osserviamo che la convergenza puntuale per $t = 1$ prova automaticamente che $P_k \rightarrow g$ uniformemente su tutto l'intervallo $[0, 1]$, difatti vale la stima

$$\left| \sum_{n=0}^k \alpha_n t^n - g(t) \right| = \left| \sum_{n=k+1}^{+\infty} \alpha_n t^n \right| = \sum_{n=k+1}^{+\infty} |\alpha_n| t^n \leq \sum_{n=k+1}^{+\infty} |\alpha_n| = \left| \sum_{n=k+1}^{+\infty} \alpha_n \right| = \left| \sum_{n=0}^k \alpha_n - g(1) \right|.$$

Infine, abbiamo già calcolato la derivata di φ

$$\varphi'(t) = -\frac{1}{2}(1-t)^{-1/2} = -\frac{1}{2} \frac{\varphi(t)}{1-t}$$

per cui φ risolve il problema di Cauchy

$$\begin{cases} u'(t) = -\frac{1}{2} \frac{u(t)}{1-t} \\ u(0) = \varphi(0) \end{cases}$$

che ha soluzione unica²; inoltre consideriamo la serie di potenze derivata

$$\sum_{n=0}^{+\infty} (\alpha_n t^n)' = \sum_{n=0}^{+\infty} n \alpha_n t^{n-1}$$

che per quanto sappiamo ha ancora raggio di convergenza 1, di conseguenza converge uniformemente a $g'(t)$ su $[0, 1 - \varepsilon]$ $\forall \varepsilon > 0$: mostriamo che su tale intervallo anche g risolve il problema. Chiaramente $g(0) = \varphi(0)$, inoltre

$$\begin{aligned} g'(t) = -\frac{1}{2} \frac{g(t)}{1-t} &\iff (1-t)g'(t) = -\frac{1}{2}g(t) \\ &\iff \sum_{n=0}^{+\infty} [n\alpha_n t^{n-1} - n\alpha_n t^n] = -\frac{1}{2} \sum_{n=0}^{+\infty} \alpha_n t^n \\ &\iff \sum_{n=0}^{+\infty} [(n+1)\alpha_{n+1} t^n - n\alpha_n t^n] = -\frac{1}{2} \sum_{n=0}^{+\infty} \alpha_n t^n \\ &\iff \sum_{n=0}^{+\infty} [(n+1)\alpha_{n+1} - n\alpha_n] t^n = \sum_{n=0}^{+\infty} -\frac{1}{2} \alpha_n t^n \end{aligned}$$

e trattandosi in particolare di serie formali, l'ultima uguaglianza vale se e solo se

$$\begin{aligned} (n+1)\alpha_{n+1} - n\alpha_n &= -\frac{1}{2}\alpha_n \quad \forall n, \quad \text{ossia} \\ -(n+1) \frac{(2n-1)!!}{2^{n+1}(n+1)!} + n \frac{(2n-3)!!}{2^n n!} &= \frac{1}{2} \cdot \frac{(2n-3)!!}{2^n n!} \end{aligned}$$

che è vera per verifica diretta. Allora g ristretta all'intervallo $[0, 1 - \varepsilon]$ risolve il problema di Cauchy, e per unicità su tale intervallo coincide con φ ; per arbitrarietà di ε l'uguaglianza vale su $[0, 1)$ e per continuità anche in 1. Segue che $P_k(t) \rightarrow \varphi(t) = \sqrt{1-t} \forall t \in [0, 1]$ e la convergenza è uniforme. \square

4.4.1 Applicazioni di Stone-Weierstraß

Consideriamo ora due esempi tipici di applicazione del Teorema di Stone-Weierstraß.

Il primo è quello dell'algebra dei **polinomi** $\mathcal{F} = \mathbb{R}[x] \subseteq C(X, \mathbb{R})$, con $X \subseteq \mathbb{R}^n$ compatto. Questa è chiaramente un'algebra tale che $1 \in \mathcal{F}$. Inoltre separa i punti: infatti, dati $x, y \in X$ con $x \neq y$, esiste almeno un indice i tale che $x_i \neq y_i$, quindi basta considerare il polinomio $p(z) = z_i$, che verifica $p(x) \neq p(y)$. Nel caso $X = [0, 1]$, otteniamo il Teorema di approssimazione di Weierstraß, da cui eravamo partiti all'inizio della sezione.

Il secondo è quello dell'algebra dei **polinomi trigonometrici** $\mathcal{F} \subseteq C(X, \mathbb{R})$, con $X = S^1$, generata dall'insieme $\{\sin(nx), \cos(nx) : n \in \mathbb{N}\}$; più precisamente, un elemento generico di \mathcal{F} è del tipo

$$T(x) = a_0 + \sum_{n=1}^k a_n \sin(nx) + \sum_{n=1}^k b_n \cos(nx) .$$

Osserviamo subito che $1 = \cos(0) \in \mathcal{F}$. Inoltre \mathcal{F} è un'algebra, in virtù delle *formule di Werner*:

²Questo fatto seguirà facilmente dal Teorema di Cauchy-Lipschitz, che vedremo nel prossimo capitolo.

$$\sin \alpha \cos \beta = \frac{1}{2} [\sin(\alpha + \beta) + \sin(\alpha - \beta)] \quad \cos \alpha \cos \beta = \frac{1}{2} [\cos(\alpha + \beta) + \cos(\alpha - \beta)]$$

$$\sin \alpha \sin \beta = \frac{1}{2} [\cos(\alpha - \beta) - \cos(\alpha + \beta)] .$$

Infine, \mathcal{F} separa i punti: infatti, dati $x, y \in S^1$ con $x \neq y$, vale almeno una tra $\sin(x) \neq \sin(y)$ e $\cos(x) \neq \cos(y)$.

Teorema 4.31: Teorema di equidistribuzione di Weyl

Data $f \in C(\mathbb{R}, \mathbb{R})$ una funzione 1-periodica, sia $\alpha \in \mathbb{R} \setminus \mathbb{Q}$. Allora

$$\frac{1}{n} \sum_{j=0}^{n-1} f(j\alpha) \xrightarrow{n \rightarrow +\infty} \int_0^1 f(x) dx .$$

Dimostrazione. Dimostriamo innanzitutto la tesi per le funzioni $f_k(x) = e^{2\pi i k x}$, $k \in \mathbb{N}$. Per $k = 0$, f_0 è la funzione costantemente 1; supponiamo dunque $k \neq 0$:

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} f_k(j\alpha) \right| = \left| \frac{1}{n} \sum_{j=0}^{n-1} e^{2\pi i k j \alpha} \right| = \left| \frac{1}{n} \sum_{j=0}^{n-1} (e^{2\pi i k \alpha})^j \right| = \frac{1}{n} \cdot \frac{|1 - e^{2\pi i k n \alpha}|}{|1 - e^{2\pi i k \alpha}|}$$

dove all'ultimo passaggio si è usato $k\alpha \notin \mathbb{N}$; adesso per disuguaglianza triangolare

$$\frac{1}{n} \cdot \frac{|1 - e^{2\pi i k n \alpha}|}{|1 - e^{2\pi i k \alpha}|} \leq \frac{2}{n} \cdot \frac{1}{|1 - e^{2\pi i k \alpha}|} \xrightarrow{n \rightarrow +\infty} 0 .$$

Per quanto detto

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f_k(j\alpha) = \begin{cases} 0 & \text{per } k \neq 0 \\ 1 & \text{per } k = 0 \end{cases} .$$

D'altro canto

$$\int_0^1 f_k(x) dx = \int_0^1 \Re f_k(x) dx + i \int_0^1 \Im f_k(x) dx = \begin{cases} 0 & \text{per } k \neq 0 \\ 1 & \text{per } k = 0 \end{cases}$$

in quanto $\Re f_k(x) = \cos(2\pi k x)$ e $\Im f_k(x) = \sin(2\pi k x)$, che sono funzioni a media nulla per $k > 0$. Abbiamo allora provato che

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f_k(j\alpha) = \int_0^1 f_k(x) dx \quad \text{che vale se e solo se}$$

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \cos(2\pi k j \alpha) = \int_0^1 \cos(2\pi k x) dx \quad \text{e} \quad \lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} \sin(2\pi k j \alpha) = \int_0^1 \sin(2\pi k x) dx .$$

Consideriamo $\mathcal{S} = \text{Span}(\{\cos(2\pi n x), \sin(2\pi n x)\}_{n \in \mathbb{N}})$ l'algebra dei polinomi trigonometrici. Abbiamo dimostrato che per le funzioni generatrici l'enunciato è vero, ma allora vale anche per le loro combinazioni lineari. Sia a questo punto $f \in C(\mathbb{R}, \mathbb{R})$ una funzione 1-periodica, e sia $\varepsilon > 0$: dal Teorema di Stone-Weierstraß sappiamo che $f \in \overline{\mathcal{S}}$, pertanto $\exists p \in \mathcal{S}$ tale che $\|f - p\|_\infty < \frac{\varepsilon}{3}$. Allora per disuguaglianza triangolare

$$\begin{aligned} & \left| \frac{1}{n} \sum_{j=0}^{n-1} f(j\alpha) - \int_0^1 f(x) dx \right| \leq \\ & \leq \left| \frac{1}{n} \sum_{j=0}^{n-1} [f(j\alpha) - p(j\alpha)] \right| + \left| \frac{1}{n} \sum_{j=0}^{n-1} p(j\alpha) - \int_0^1 p(x) dx \right| + \left| \int_0^1 [f(x) - p(x)] dx \right| . \end{aligned}$$

Stimiamo i tre addendi:

I.

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} [f(j\alpha) - p(j\alpha)] \right| \leq \frac{1}{n} \sum_{j=0}^{n-1} |f(j\alpha) - p(j\alpha)| \leq \frac{1}{n} \sum_{j=0}^{n-1} \|f - p\|_{\infty} = \|f - p\|_{\infty} < \frac{\varepsilon}{3}$$

II. $p \in \mathcal{S}$, dunque $\exists \bar{n}$ tale che $\forall n \geq \bar{n}$

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} p(j\alpha) - \int_0^1 p(x) dx \right| < \frac{\varepsilon}{3}$$

III.

$$\left| \int_0^1 [f(x) - p(x)] dx \right| \leq \int_0^1 |f(x) - p(x)| dx \leq \int_0^1 \|f - p\|_{\infty} dx = \|f - p\|_{\infty} < \frac{\varepsilon}{3}$$

Di conseguenza $\forall \varepsilon > 0 \exists \bar{n}$ tale che

$$\left| \frac{1}{n} \sum_{j=0}^{n-1} f(j\alpha) - \int_0^1 f(x) dx \right| < \varepsilon$$

che prova la tesi. □

Tu quando non sai fare una cosa, sai cosa fai? La dividi sempre in tre.

Carlo Cracco

Osservazione 4.32. Il punto chiave del Teorema è che α non è commensurabile con il periodo.

Nel seguito, indichiamo con $\{x\}$ la funzione *parte frazionaria* di x : $\{x\} := x - [x]$.

Corollario 4.33

Sia $\alpha \in \mathbb{R} \setminus \mathbb{Q}$ e sia $x_j = \{j\alpha\}$ al variare di $j \in \mathbb{N}$. Allora dati a, b con $0 \leq a < b \leq 1$ vale che

$$\lim_{n \rightarrow +\infty} \frac{\#\{j \mid 0 \leq j \leq n-1, a \leq x_j \leq b\}}{n} = b - a .$$

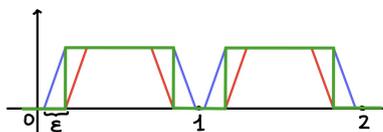
Dimostrazione. Definiamo sull'insieme \mathbb{R} la funzione $f(x) = \chi_{[a,b]}(\{x\})$, dove $\chi_{[a,b]}$ è la funzione indicatrice dell'intervallo $[a, b]$. Valgono allora le identità

$$\#\{j \mid 0 \leq j \leq n-1, a \leq x_j \leq b\} = \sum_{j=0}^{n-1} f(j\alpha) \quad \text{e} \quad b - a = \int_0^1 \chi_{[a,b]}(x) dx .$$

Dato $\varepsilon > 0$, definiamo le seguenti funzioni ausiliarie:

$$\varphi^-(x) = \begin{cases} 0 & \text{se } \{x\} \in [0, a] \cup [b, 1] \\ r_1 & \text{se } \{x\} \in [a, a + \varepsilon] \\ 1 & \text{se } \{x\} \in [a + \varepsilon, b - \varepsilon] \\ r_2 & \text{se } \{x\} \in [b - \varepsilon, b] \end{cases} \quad \text{e} \quad \varphi^+(x) = \begin{cases} 0 & \text{se } \{x\} \in [0, a - \varepsilon] \cup [b + \varepsilon, 1] \\ s_1 & \text{se } \{x\} \in [a - \varepsilon, a] \\ 1 & \text{se } \{x\} \in [a, b] \\ s_2 & \text{se } \{x\} \in [b, b + \varepsilon] \end{cases}$$

dove r_1 e r_2 sono rispettivamente la retta passante per $(a, 0)$ e $(a + \varepsilon, 1)$ e la retta passante per $(b - \varepsilon, 1)$ e $(b, 0)$, mentre s_1 e s_2 sono le rette passanti per le coppie $(a - \varepsilon, 0)$, $(a, 1)$ e $(b, 1)$, $(b + \varepsilon, 0)$.



Dalla definizione segue che φ^-, φ^+ sono funzioni continue e 1-periodiche, dunque soddisfano le ipotesi del Teorema di Weyl; valgono inoltre le seguenti immediate proprietà:

$$\varphi^-(x) \leq f(x) \leq \varphi^+(x) \quad , \quad \int_0^1 \varphi^-(x) dx = b - a - \varepsilon \quad \text{e} \quad \int_0^1 \varphi^+(x) dx = b - a + \varepsilon.$$

Dalla prima disuguaglianza segue che

$$\sum_{j=0}^{n-1} f(j\alpha) \geq \sum_{j=0}^{n-1} \varphi^-(j\alpha) \quad \text{da cui} \quad \liminf_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(j\alpha) \geq \int_0^1 \varphi^-(x) dx = \int_0^1 f(x) dx - \varepsilon;$$

$$\text{analogamente} \quad \limsup_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(j\alpha) \leq \int_0^1 \varphi^+(x) dx = \int_0^1 f(x) dx + \varepsilon .$$

Per arbitrarietà di ε concludiamo che il minimo ed il massimo limite coincidono, e che

$$\lim_{n \rightarrow +\infty} \frac{1}{n} \sum_{j=0}^{n-1} f(j\alpha) = \int_0^1 f(x) dx = b - a .$$

□

Esercizio 3. Data la funzione

$$f : \mathbb{R} \rightarrow \mathbb{R} \quad \text{definita come} \quad f(x) = \sum_{k=0}^{+\infty} \frac{1}{k!} \sin((k!)^2 x)$$

si dimostri che la funzione f è ben definita e continua in \mathbb{R} , 2π -periodica, ma non è derivabile in nessun punto.

Soluzione. Poniamo

$$u_k(x) = \frac{1}{k!} \sin((k!)^2 x) \quad ;$$

tutte le funzioni u_k sono continue e $\|u_k\|_\infty \leq \frac{1}{k!}$, pertanto la serie delle u_k converge uniformemente a f ; inoltre tutte le u_k sono 2π -periodiche. Di conseguenza f è continua e 2π -periodica. Mostriamo che f non è derivabile in nessun punto, e per farlo dimostriamo che $\forall x \in \mathbb{R}$ il rapporto incrementale non è mai limitato in un intorno di x , ossia

$$\limsup_{h \rightarrow 0} \left| \frac{f(x+h) - f(x)}{h} \right| = +\infty .$$

Lemma 1 Per $n \geq 1$ vale

$$\sum_{k=n}^{+\infty} \frac{1}{k!} \leq \frac{2}{n!} .$$

Dimostrazione.

$$\sum_{k=n}^{+\infty} \frac{1}{k!} \leq \frac{1}{n!} + \sum_{h=1}^{+\infty} \frac{1}{(n+h)!} \leq \frac{1}{n!} + \frac{1}{n!} \sum_{h=1}^{+\infty} \frac{1}{(n+1)^h} \leq \frac{1}{n!} + \frac{1}{n!} \sum_{h=1}^{+\infty} \frac{1}{2^h} \leq \frac{2}{n!}$$

□

Lemma 2 Dati $x \in \mathbb{R}$, $N \in \mathbb{N}$, $\exists y \in \left[x, x + \frac{2\pi}{N} \right]$, da cui $|x - y| < \frac{2\pi}{N}$, tale che $|\sin(Nx) - \sin(Ny)| \geq 1$.

Dimostrazione. La funzione $x \mapsto \sin(Nx)$ ha periodo $\frac{2\pi}{N}$ ed ha dunque immagine $[-1, 1]$ ristretta al periodo $\left[x, x + \frac{2\pi}{N} \right]$. □

Dalla definizione di f vale l'identità $f(x) = h_n^-(x) + u_n(x) + h_n^+(x)$, dove

$$h_n^-(x) := \sum_{k=0}^{n-1} u_k(x) \quad \text{e} \quad h_n^+(x) := \sum_{k=n+1}^{+\infty} u_k(x)$$

Per il Lemma 2, scegliamo $x_n \in \left[x, x + \frac{2\pi}{(n!)^2} \right]$ tale che $|\sin((n!)^2 x) - \sin((n!)^2 x_n)| \geq 1$, di conseguenza

$$|u_n(x) - u_n(x_n)| = \frac{1}{n!} |\sin((n!)^2 x) - \sin((n!)^2 x_n)| \geq \frac{1}{n!}.$$

A questo punto $f(x) - f(x_n) = h_n^-(x) - h_n^-(x_n) + u_n(x) - u_n(x_n) + h_n^+(x) - h_n^+(x_n)$: facciamo un po' di contazzi. Poniamo

$$\mathbf{I} = h_n^-(x) - h_n^-(x_n), \quad \mathbf{II} = u_n(x) - u_n(x_n) \quad \text{e} \quad \mathbf{III} = h_n^+(x) - h_n^+(x_n).$$

Stimiamo

$$\begin{aligned} |\mathbf{I}| &= |h_n^-(x) - h_n^-(x_n)| = \left| \sum_{k=0}^{n-1} \frac{1}{k!} [\sin((k!)^2 x) - \sin((k!)^2 x_n)] \right| \leq \\ &\leq \sum_{k=0}^{n-1} \frac{1}{k!} |\sin((k!)^2 x) - \sin((k!)^2 x_n)| \stackrel{*}{\leq} \sum_{k=0}^{n-1} \frac{(k!)^2 |x - x_n|}{k!} \leq 2\pi \sum_{k=0}^{n-1} \frac{k!}{(n!)^2} = \\ &= \frac{2\pi}{n!} \left[\frac{(n-1)!}{n!} + \sum_{k=0}^{n-2} \frac{k!}{n!} \right] \leq \frac{2\pi}{n!} \left[\frac{1}{n} + \sum_{k=0}^{n-2} \frac{(n-2)!}{n!} \right] \leq \frac{2\pi}{n!} \left[\frac{1}{n} + \sum_{k=0}^{n-2} \frac{1}{n(n-1)} \right] = \frac{4\pi}{n \cdot n!} \end{aligned}$$

dove il passaggio (*) segue dalla Ω -lipschitzianità della funzione $\sin(\Omega x)$.

Similmente

$$\begin{aligned} |\mathbf{III}| &= |h_n^+(x) - h_n^+(x_n)| = \left| \sum_{k=n+1}^{+\infty} \frac{1}{k!} [\sin((k!)^2 x) - \sin((k!)^2 x_n)] \right| \leq \\ &\leq \sum_{k=n+1}^{+\infty} \frac{1}{k!} |\sin((k!)^2 x) - \sin((k!)^2 x_n)| \leq \sum_{k=n+1}^{+\infty} \frac{2}{k!} \stackrel{*}{\leq} \frac{4}{(n+1)!} = \frac{4}{(n+1)n!} \end{aligned}$$

e la disuguaglianza (*) segue dal Lemma 1. A questo punto

$$\begin{aligned} |f(x) - f(x_n)| &= |\mathbf{I} + \mathbf{II} + \mathbf{III}| \geq |\mathbf{II}| - \{|\mathbf{I}| + |\mathbf{III}|\} \geq \\ &\geq \frac{1}{n!} \left(1 - \frac{4\pi}{n} - \frac{4}{n+1} \right) = \frac{1}{n!} (1 + o(1)) \geq \frac{1}{2 \cdot n!} \quad \text{definitivamente.} \end{aligned}$$

Di conseguenza

$$\frac{|f(x) - f(x_n)|}{|x - x_n|} \geq \frac{1}{2 \cdot n!} \cdot \frac{(n!)^2}{2\pi} = \frac{n!}{4\pi}$$

che è illimitato.

Capitolo 5

Equazioni differenziali

Definizione 5.1 (Equazione differenziale ordinaria). Un'equazione differenziale ordinaria (ODE) è un'equazione differenziale che coinvolge una funzione in una variabile e le sue derivate, di ordine qualsiasi.

Un'equazione differenziale ordinaria di ordine n è un'equazione del tipo

$$F(t, u(t), \dots, u^{(n)}(t)) = 0$$

dove $u : I \rightarrow \mathbb{R}^m$ e $F : I \times \mathbb{R}^{m(n+1)} \rightarrow \mathbb{R}$, con $I \subseteq \mathbb{R}$.

Definizione 5.2 (Equazione differenziale alle derivate parziali). Un'equazione differenziale alle derivate parziali (PDE) è un'equazione differenziale che coinvolge una funzione in più variabili e le sue derivate parziali, di ordine qualsiasi.

Definizione 5.3 (Problema di Cauchy). Un problema di Cauchy (PC) di primo ordine è un problema ai valori iniziali del tipo:

$$\begin{cases} u'(t) = F(t, u(t)) \quad \forall t \\ u(t_0) = u_0 \end{cases} .$$

Definizione 5.4 (Equazione integrale di Volterra). Un'equazione integrale di Volterra (EV) è un'equazione del tipo

$$u(t) = u_0 + \int_{t_0}^t F(\sigma, u(\sigma)) d\sigma .$$

Osservazione 5.1. Se F è continua, $u \in C^1$ risolve PC $\iff u \in C$ risolve EV.

Dimostrazione. (\implies) La tesi è immediata, infatti da PC segue:

$$u_0 + \int_{t_0}^t F(\sigma, u(\sigma)) d\sigma = u_0 + \int_{t_0}^t u'(\sigma) d\sigma = u_0 + u(t) - u(t_0) = u(t) ;$$

(\impliedby) Essendo u e F continue, anche $F(t, u(t))$ è continua, quindi per il Teorema fondamentale del calcolo integrale

$$u(t) - u_0 = \int_{t_0}^t F(\sigma, u(\sigma)) d\sigma$$

è derivabile. Perciò $u'(t) = (u(t) - u_0)' = F(t, u(t))$, e inoltre $u \in C^1$. \square

5.1 Teoremi di esistenza e unicità

Teorema 5.2: Teorema di Cauchy-Lipschitz-Picard-Lindelöf

Data una funzione $F : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ continua e localmente lipschitziana nel secondo argomento, siano $t_0 \in \mathbb{R}$ e $u_0 \in \mathbb{R}^m$. Allora $\exists S_\infty \in (t_0, +\infty]$ tale che esiste ed è unica la soluzione massimale al **PC** in $[t_0, S_\infty)$.

Dimostrazione. Osserviamo che la soluzione del **PC** sarà una funzione $u \in C^1([t_0, S), \mathbb{R}^m)$: cerchiamo la massima costante S per cui u esiste ed è unica. Per l'osservazione precedente, cerchiamo una funzione u continua che risolve **EV**.

Poiché F è localmente lipschitziana nel secondo argomento, esistono un intorno compatto V di (t_0, u_0) ed una costante $L > 0$ tale che $F|_V$ è uniformemente L -Lipschitz nel secondo argomento, ossia

$$|F(t, u_2) - F(t, u_1)| \leq L|u_2 - u_1| \quad \forall (t, u_1), (t, u_2) \in V .$$

Poiché invece F è continua, la mappa $v \mapsto |F(v)|$ è continua e ammette pertanto un massimo $M > 0$ sul compatto V . Siano allora $\delta > 0$ e $\varepsilon > 0$ tali che $[t_0, t_0 + \delta] \times \overline{B}_\varepsilon(u_0) \subseteq V$, $M\delta \leq \varepsilon$ e $L\delta < 1$. Definiamo lo spazio di funzioni $X = \{u \in C([t_0, t_0 + \delta], \mathbb{R}^m) \mid \|u - u_0\|_\infty \leq \varepsilon\}$ con la distanza del sup: esso è un sottospazio chiuso dello spazio completo $C([t_0, t_0 + \delta], \mathbb{R}^m)$, e dunque è completo. Definiamo allora l'operatore di Volterra

$$T : X \rightarrow X \quad \text{dato da} \quad T(u)(t) = u_0 + \int_{t_0}^t F(\sigma, u(\sigma)) d\sigma .$$

La mappa T è ben definita in quanto $T(u)$ è continua se u lo è¹ e inoltre per ogni $t \in [t_0, t_0 + \delta]$ si ha

$$|T(u)(t) - u_0| = \left| \int_{t_0}^t F(\sigma, u(\sigma)) d\sigma \right| \leq \int_{t_0}^t |F(\sigma, u(\sigma))| d\sigma \leq M\delta \leq \varepsilon$$

e dunque $\|T(u) - u_0\|_\infty \leq \varepsilon$. Osserviamo che u risolve **EV** se e solo se $T(u) = u$; mostriamo allora che T è una contrazione, cosicché il Teorema di Banach garantisca l'esistenza e unicità di una soluzione a **EV** in X in quanto punto fisso di T . Date $u_1, u_2 \in X$ e preso $t \in [t_0, t_0 + \delta]$ si ha

$$\begin{aligned} |T(u_2)(t) - T(u_1)(t)| &= \left| \int_{t_0}^t [F(\sigma, u_2(\sigma)) - F(\sigma, u_1(\sigma))] d\sigma \right| \leq \int_{t_0}^t |F(\sigma, u_2(\sigma)) - F(\sigma, u_1(\sigma))| d\sigma \leq \\ &\leq L \int_{t_0}^t |u_2(\sigma) - u_1(\sigma)| d\sigma \leq L\delta \|u_2 - u_1\|_\infty \end{aligned}$$

pertanto $\|T(u_2) - T(u_1)\|_\infty \leq L\delta \|u_2 - u_1\|_\infty$, e poiché $L\delta < 1$ l'operatore T è una contrazione.

Dall'unicità della soluzione in X possiamo dedurre l'unicità in tutto $C([t_0, t_0 + \delta], \mathbb{R}^m)$. Infatti siano per assurdo $u_1, u_2 : [t_0, t_0 + \delta] \rightarrow \mathbb{R}^m$ due soluzioni continue distinte del **PC** a valori iniziali (t_0, u_0) , e sia $\bar{t} = \inf\{t \text{ tale che } u_1(t) \neq u_2(t)\}$; chiaramente $\bar{u} := u_1(\bar{t}) = u_2(\bar{t})$. Consideriamo allora il **PC'** dato dalla stessa equazione precedente ma con valori iniziali (\bar{t}, \bar{u}) ; siano V' compatto e $L' > 0$ costante dati dalla locale lipschitzianità di F nella seconda variabile per il punto (\bar{t}, \bar{u}) e sia $M' > 0$ il massimo del modulo di F su V' ; siano poi δ' e ε' tali che $[\bar{t}, \bar{t} + \delta'] \times \overline{B}_{\varepsilon'}(\bar{u}) \subseteq V'$, $M'\delta' \leq \varepsilon'$ e $L'\delta' < 1$; a meno di ridurre δ' possiamo richiedere che $\bar{t} + \delta' \leq t_0 + \delta$, e che per $|t - \bar{t}| \leq \delta'$ valgano $|u_1(t) - \bar{u}| \leq \varepsilon'$ e $|u_2(t) - \bar{u}| \leq \varepsilon'$. Per quanto detto allora u_1 e u_2 sono due soluzioni diverse al **PC'** nello spazio

$$X' = \{u \in C([\bar{t}, \bar{t} + \delta'], \mathbb{R}^m) \mid \|u - \bar{u}\|_\infty \leq \varepsilon'\}$$

in cui per quanto provato la soluzione è unica. ζ

Occupiamoci adesso della massimalità. Consideriamo

$$S_\infty = \sup\{S \text{ tale che esiste una soluzione unica su } [t_0, S]\} .$$

¹Questo fatto segue facilmente dal Teorema della media integrale.

Su $[t_0, S_\infty)$ abbiamo unicità della soluzione: infatti per ogni $s \in [t_0, S_\infty)$, esiste un'unica soluzione u definita su $[t_0, s]$; inoltre, per ogni t con $s < t < S_\infty$, la soluzione v definita su $[t_0, t]$ è in particolare una soluzione definita su $[t_0, s]$, cioè $v|_{[t_0, s]} \equiv u$.

Rimane solo da mostrare che l'intervallo $[t_0, S_\infty)$ di definizione della soluzione è massimale: per farlo, mostriamo che non può esistere finito il limite di $u(t)$ per $t \rightarrow S_\infty^-$ a meno che $S_\infty = +\infty$. Infatti, se $S_\infty \neq +\infty$ e per assurdo $\ell \in \mathbb{R}$ tale limite finito, potremmo estendere per continuità $u(t)$ in S_∞ . A questo punto potremmo riapplicare il Teorema di Cauchy-Lipschitz al problema di Cauchy con dato iniziale $u(S_\infty) = \ell$, ottenendo una soluzione unica $v(t)$ definita su $[S_\infty, S'_\infty)$, con $S'_\infty > S_\infty$. Ora, poiché F è continua, la funzione

$$\bar{u}(t) = \begin{cases} u(t) & \text{su } [t_0, S_\infty) \\ v(t) & \text{su } [S_\infty, S'_\infty) \end{cases}$$

è C^1 ed è ancora soluzione del problema di Cauchy iniziale e unica su $[t_0, S'_\infty)$, quindi S_∞ non è il sup dell'insieme $\{S \text{ tale che esiste una soluzione unica su } [t_0, S]\}$. \neq

Dunque l'intervallo $[t_0, S_\infty)$ di definizione della soluzione è massimale. \square

Corollario 5.3

Se la funzione F è uniformemente lipschitziana nel secondo argomento allora la soluzione massimale è globale.

Dimostrazione. Poiché L è uniformemente lipschitziana nel secondo argomento, esiste una costante $L > 0$ tale che

$$|F(t, u_2) - F(t, u_1)| \leq L|u_2 - u_1| \quad \forall (t, u_1), (t, u_2) \in \mathbb{R} \times \mathbb{R}^m .$$

Sia $\delta > 0$ tale che $L\delta < 1$, e definiamo

$$T : C([t_0, t_0 + \delta], \mathbb{R}^m) \rightarrow C([t_0, t_0 + \delta], \mathbb{R}^m) \quad \text{dato da} \quad T(u)(t) = u_0 + \int_{t_0}^t F(\sigma, u(\sigma)) d\sigma ;$$

allora T è ben definita (ossia $T(u)$ è continua per $u \in C([t_0, t_0 + \delta], \mathbb{R}^m)$) e vale

$$\begin{aligned} |T(u_2)(t) - T(u_1)(t)| &= \left| \int_{t_0}^t [F(\sigma, u_2(\sigma)) - F(\sigma, u_1(\sigma))] d\sigma \right| \leq \int_{t_0}^t |F(\sigma, u_2(\sigma)) - F(\sigma, u_1(\sigma))| d\sigma \leq \\ &\leq L \int_{t_0}^t |u_2(\sigma) - u_1(\sigma)| d\sigma \leq L\delta \|u_2 - u_1\|_\infty , \end{aligned}$$

cioè T è una contrazione, e dunque ha un unico punto fisso su $C([t_0, t_0 + \delta], \mathbb{R}^m)$. Come prima concludiamo che esiste ed è unica la soluzione su $[t_0, t_0 + \delta]$; d'altra parte, poiché δ non dipende da t_0 , per lo stesso motivo, esiste ed è unica la soluzione su $[t_0 + \delta, t_0 + 2\delta]$ e così via. Incollando le soluzioni, abbiamo esistenza e unicità globale. \square

Teorema 5.4: Teorema di Cauchy-Peano

Data una funzione $F : \mathbb{R} \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ continua, un punto $t_0 \in \mathbb{R}$ ed un vettore $u_0 \in \mathbb{R}^m$, esiste una soluzione locale al **PC**.

Dimostrazione. Poniamo $k := \max\{|F(t, x)| \mid t \in [t_0 - 1, t_0 + 1], |x - u_0| \leq 1\}$, che certamente esiste per compattezza dell'insieme considerato. Scegliamo invece $I = [t_0 - \frac{1}{k+1}, t_0 + \frac{1}{k+1}]$; ci prefiggiamo di costruire una soluzione definita sull'insieme I . Dato $n \in \mathbb{N}$, poniamo $\ell_n = \frac{1}{(k+1)^n}$ e costruiamo $u_n \in C(I, \mathbb{R})$ come

$$\begin{aligned} u_n(t_0) &= u_0 , \\ u'_n(t) &= F(t_0, u_0) \quad t \in (t_0, t_0 + \ell_n) \\ &= F(t_0 + \ell_n, u_n(t_0 + \ell_n)) \quad t \in (t_0 + \ell_n, t_0 + 2\ell_n) \\ &\vdots \\ &= F(t_0 + (n-1)\ell_n, u_n(t_0 + (n-1)\ell_n)) \quad t \in (t_0 + (n-1)\ell_n, t_0 + \frac{1}{k+1}) \end{aligned}$$

Del tutto analogamente si definisce u_n su $[t_0 - \frac{1}{k+1}, t_0]$. Per costruzione per ogni n vale $|u_n(t) - u_0| < 1 \ \forall t \in I$, infatti altrimenti esisterebbe \bar{t} con $|\bar{t} - t_0|$ minimo tale che $|u_n(\bar{t}) - u_0| \geq 1$, ma allora

$$|u_n(\bar{t}) - u_0| = \left| \int_{t_0}^{\bar{t}} u'_n(\tau) d\tau \right| \leq \int_{t_0}^{\bar{t}} |u'_n(\tau)| d\tau \leq \int_{t_0}^{\bar{t}} k d\tau \leq k \cdot \frac{1}{k+1} < 1 \quad \#$$

Di conseguenza le u_n sono funzioni equilimitate, ed in effetti sono anche equicontinue in quanto k -equilipschitziane: poiché l'insieme I è compatto, per il teorema di Ascoli-Arzelà a meno di sottosuccessioni $u_n \rightarrow u$ uniformemente. Adesso la funzione u è continua in quanto limite uniforme di funzioni continue, e dato che in particolare $u_n \rightarrow u$ puntualmente vale che $u(t_0) = u_0$. Mostriamo dunque che $u'(t) = F(t, u(t))$, o equivalentemente che

$$u(t) = u_0 + \int_{t_0}^t F(\tau, u(\tau)) d\tau \quad (\mathbf{EV})$$

Definiamo $g(t) = F(t, u(t))$ e definiamo $g_n : I \rightarrow \mathbb{R}^m$ come

$$\begin{aligned} g_n(t) &= F(t_0, u_0) \quad t \in [t_0, t_0 + \ell_n] \\ &= F(t_0 + \ell_n, u_n(t_0 + \ell_n)) \quad t \in (t_0 + \ell_n, t_0 + 2\ell_n] \\ &\vdots \\ &= F(t_0 + (n-1)\ell_n, u_n(t_0 + (n-1)\ell_n)) \quad t \in (t_0 + (n-1)\ell_n, t_0 + \frac{1}{k+1}] \end{aligned}$$

e allo stesso modo per $t \in [t_0 - \frac{1}{k+1}, t_0]$.

Mostriamo che $g_n \rightarrow g$ uniformemente.

Innanzitutto se $J = \{x \in \mathbb{R}^m \mid |x - u_0| \leq 1\}$, il cilindro $I \times J$ è compatto, dunque se F è continua, $F|_{I \times J}$ è uniformemente continua. Fissiamo $\varepsilon > 0$ e sia $\delta > 0$ tale che per $|(t_1, x_1) - (t_2, x_2)| < \delta$ con $(t_i, x_i) \in I \times J$ si abbia $|F(t_1, x_1) - F(t_2, x_2)| < \frac{\varepsilon}{2}$.

Sia \bar{n}_1 tale che $\forall n > \bar{n}_1$ valga $\ell_n < \frac{\delta}{2(1+k)}$ e sia \bar{n}_2 tale che $\forall n > \bar{n}_2$ valga $\|u_n - u\|_\infty < \frac{\varepsilon}{2}$; sia poi $\bar{n} = \max\{\bar{n}_1, \bar{n}_2\}$. Scelto $n > \bar{n}$, e preso $t \in I$, esiste un certo m tale che $t \in [t_0 + m\ell_n, t_0 + (m+1)\ell_n]$. Adesso $|g_n(t) - g(t)| = |F(t_0 + m\ell_n, u_n(t_0 + m\ell_n)) - F(t, u(t))|$, pertanto se fosse $|(t_0 + m\ell_n, u_n(t_0 + m\ell_n)) - (t, u(t))| < \delta$, avremmo $|g_n(t) - g(t)| < \frac{\varepsilon}{2}$, e poiché n non dipende da t vale $\|g_n - g\|_\infty < \varepsilon$. Per costruzione

$$\begin{aligned} |(t_0 + m\ell_n, u_n(t_0 + m\ell_n)) - (t, u(t))| &\leq |(t_0 + m\ell_n - t)| + |u_n(t_0 + m\ell_n) - u(t)| \\ &\leq |(t_0 + m\ell_n - t)| + |u_n(t_0 + m\ell_n) - u_n(t)| + |u_n(t) - u(t)| \leq \\ &\ell_n + k\ell_n + \|u_n - u\|_\infty < \frac{\delta}{2(1+k)} + \frac{\delta}{2} = \delta \quad . \end{aligned}$$

A questo punto si conclude osservando che per ogni $t \in I$ vale

$$\begin{aligned} u(t) &= u_n(t) + (u(t) - u_n(t)) = u_0 + \int_{t_0}^t u'_n(\tau) d\tau + (u(t) - u_n(t)) = \\ &= u_0 + \int_{t_0}^t g_n(\tau) d\tau + (u(t) - u_n(t)) = u_0 + \int_{t_0}^t g(\tau) d\tau + \int_{t_0}^t (g_n(\tau) - g(\tau)) d\tau + (u(t) - u_n(t)) \end{aligned}$$

e

$$\left| \int_{t_0}^t (g_n(\tau) - g(\tau)) d\tau + (u(t) - u_n(t)) \right| \leq \frac{1}{k+1} \|g_n - g\|_\infty + \|u_n - u\|_\infty \xrightarrow{n \rightarrow +\infty} 0 \quad ,$$

ossia

$$u(t) = u_0 + \int_{t_0}^t g(\tau) d\tau = u_0 + \int_{t_0}^t F(\tau, u(\tau)) d\tau \quad .$$

□

5.2 Studio qualitativo delle soluzioni

Capita spesso di non riuscire a trovare un'espressione analitica esplicita per una soluzione di un problema ai valori iniziali; l'analisi qualitativa si prefigge allora di ottenere informazioni sulle soluzioni quando esse restano implicite.

Lemma 5.5: Lemma di Grönwall

Sia $I \subseteq \mathbb{R}$ un intervallo della forma $[a, b)$ o $[a, b]$, con $b \in \mathbb{R} \cup \{+\infty\}$. Siano u e α due funzioni continue su I . Se u è derivabile su $\text{int}(I)$ e vale $u'(t) \leq \alpha(t)u(t) \quad \forall t \in \text{int}(I)$, allora

$$u(t) \leq u(a) \exp\left(\int_a^t \alpha(\sigma) d\sigma\right) \quad \forall t \in I.$$

Dimostrazione. Consideriamo la funzione integrale

$$A(t) = \int_a^t \alpha(\sigma) d\sigma .$$

Sia v la soluzione del **PC**:

$$\begin{cases} v'(t) = \alpha(t)v(t) \\ v(a) = 1 \end{cases}$$

cioè $v(t) = \exp(A(t)) = \exp\left(\int_a^t \alpha(\sigma) d\sigma\right)$ e vale $v(t) > 0 \quad \forall t > 0$.

Ora vale:

$$\left(\frac{u}{v}\right)' = \frac{u'v - uv'}{v^2} \leq \frac{\alpha uv - u\alpha v}{v^2} = 0 .$$

Quindi la funzione $\frac{u}{v}$ è decrescente, per cui $\frac{u(t)}{v(t)} \leq \frac{u(a)}{v(a)} = u(a)$ e di conseguenza

$$u(t) \leq u(a)v(t) = u(a) \exp\left(\int_a^t \alpha(\sigma) d\sigma\right) .$$

□

Teorema 5.6: Dipendenza continua dal dato iniziale

Consideriamo il problema di Cauchy

$$\begin{cases} u'(t) = F(t, u(t)) \quad \forall t \\ u(t_0) = u_0 \end{cases} .$$

Se F è continua e localmente lipschitziana nel secondo argomento e $u(t) = \Psi(t_0, u_0)(t)$ è la soluzione del problema di Cauchy, allora Ψ è localmente lipschitziana in u_0 e vale

$$|\Psi(t_0, u_0)(t) - \Psi(t_0, u_1)(t)| \leq |u_0 - u_1| e^{L|t-t_0|}$$

dove L è la costante di locale lipschitzianità.

Dimostrazione. Sia L la costante di locale lipshitzianità di F in u_0 all'istante t_0 , e sia u_1 un punto dell'intorno di u_0 su cui F è L -Lipschitz. Poniamo $u_0(t) = \Psi(t_0, u_0)(t)$ e $u_1(t) = \Psi(t_0, u_1)(t)$, e definiamo la funzione di allontanamento $w(t) = |u_0(t) - u_1(t)|$; allora vale la stima

$$\begin{aligned} w'(t) &= \frac{u_0(t) - u_1(t)}{|u_0(t) - u_1(t)|} (u_0'(t) - u_1'(t)) \leq |u_0'(t) - u_1'(t)| = \\ &= |F(t, u_0(t)) - F(t, u_1(t))| \leq L|u_0(t) - u_1(t)| = Lw(t) . \end{aligned}$$

Osserviamo dunque che se v è la soluzione di

$$\begin{cases} v'(t) = Lv(t) \\ v(t_0) = w(t_0) \end{cases}$$

allora $v(t) \geq w(t)$ per ogni $t \geq t_0$, e di conseguenza

$$w(t) \leq v(t) = |u_0(t_0) - u_1(t_0)|e^{L(t-t_0)} = |u_0 - u_1|e^{L(t-t_0)}$$

Analogamente si trova la stima per $t < t_0$. □

Lemma 5.7: Criterio dell'asintoto

Sia $f : [x_0, +\infty) \rightarrow \mathbb{R}$ derivabile tale che

i. esiste finito il limite $\lim_{x \rightarrow +\infty} f(x)$;

ii. esiste il limite $\lim_{x \rightarrow +\infty} f'(x)$,

Allora $\lim_{x \rightarrow +\infty} f'(x) = 0$.

Dimostrazione. La tesi segue immediatamente dal teorema di De L'Hopital, infatti:

$$\lim_{x \rightarrow +\infty} \frac{f'(x)}{1} = \lim_{x \rightarrow +\infty} \frac{f(x)}{x} = 0 .$$

Alternativamente si poteva applicare il Teorema di Lagrange a f in ogni intervallo del tipo $[n, n+1]$: per Lagrange esiste infatti $\xi_n \in (n, n+1)$ tale che

$$f(n+1) - f(n) = f'(\xi_n)(n+1 - n) = f'(\xi_n) .$$

Passando al limite e sfruttando le ipotesi, si ottiene

$$\lim_{n \rightarrow +\infty} f'(\xi_n) = \lim_{x \rightarrow +\infty} f'(x) = 0 .$$

□

Osservazione 5.8. Una funzione $F : \Omega \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ di classe C^1 è localmente lipschitziana in y . Infatti, fissato $(x_0, y_0) \in \Omega \times \mathbb{R}^n$ e $I \times J \subseteq \text{Dom}(F)$ intorno compatto di (x_0, y_0) , vale

$$|F(x_0, y_1) - F(x_0, y_2)| \leq \left| \frac{\partial F}{\partial y}(x_0, \xi) \right| |y_1 - y_2| \leq L|y_1 - y_2| \quad \forall x \in I, \forall y_1, y_2 \in J$$

con $L = \sup_{I \times J} \left| \frac{\partial F}{\partial y} \right|$.

Esercizio 4. Consideriamo il problema di Cauchy

$$\begin{cases} u'(t) = e^{tu(t)} - u(t)^2 \\ u(0) = u_0 \end{cases}$$

Soluzione. Innanzitutto osserviamo che il problema è del tipo $u'(t) = F(t, u(t))$ con una F di classe C^1 , quindi, per il teorema di Cauchy-Lipschitz, esiste una soluzione massimale unica per qualunque dato iniziale.

Supponiamo adesso che u sia una soluzione, e che per un qualche $\bar{t} \geq 0$ si abbia $u(\bar{t}) > 0$. Ma allora, $u(t) > 0$ per qualunque $t \geq \bar{t}$: infatti una soluzione che passa da 0 ha derivata strettamente positiva in tale punto, e quindi la soluzione era strettamente negativa subito prima di passare da 0; di conseguenza, una soluzione che sia strettamente positiva in qualche punto rimarrà positiva

finché esiste. In realtà lo stesso vale anche per una soluzione per la quale per un qualche $\bar{t} \geq 0$ si abbia $u(\bar{t}) \geq 0$: infatti, se in \bar{t} si ha che $u(\bar{t}) = 0$, allora subito dopo \bar{t} la funzione è strettamente positiva, e quindi ci si riconduce a quanto appena osservato.

Notiamo ora che se $t \geq 2$ e $u(t) \geq 0$ allora si ha

$$u'(t) = e^{tu(t)} - u(t)^2 \geq e^{2u(t)} - u(t)^2 \geq 1 + 2u(t) + 2u(t)^2 - u(t)^2 \geq 1.$$

Si ha allora che una soluzione che in un qualche \bar{t} sia positiva rimarrà sempre positiva, ed avrà derivata maggiore di 1 per tutti gli istanti successivi a \bar{t} e maggiori di 2. Tale soluzione deve quindi esplodere all'infinito. Abbiamo cioè che $u(t) \rightarrow +\infty$ per $t \rightarrow M$ per ogni $u_0 \geq 0$, ma anche per valori di u_0 leggermente negativi.

Consideriamo invece una soluzione u tale che $u(\bar{t}) \leq -1$ per un qualche $\bar{t} \geq 0$. Ma allora

$$u'(\bar{t}) = e^{-\bar{t}|u(\bar{t})|} - u(\bar{t})^2 \leq e^{-\bar{t}} - u(\bar{t})^2 \leq 1 - u(\bar{t})^2 \leq 0,$$

e la disuguaglianza è stretta se $u(\bar{t}) < -1$, così come se $t > 0$. Questo assicura che una soluzione che in un qualunque momento passa sotto a -1 , da quel momento in poi scenderà sempre. Essendo una funzione continua e decrescente, se $t \rightarrow M$ può tendere a $-\infty$ oppure ad un valore $L \in (-\infty, -1)$. Ma questa seconda possibilità può essere esclusa: se infatti $u(t) \rightarrow L$, allora per forza deve essere $M = +\infty$, visto che altrimenti la soluzione non sarebbe massimale. E se $u(t) \rightarrow L$ per $t \rightarrow +\infty$, con $L < -1$, allora di sicuro $u'(t) \rightarrow -L^2$. Una funzione decrescente che tenda ad un limite finito, tuttavia, non può avere derivata che tende ad un numero diverso da 0, e quindi si ha l'assurdo cercato. Ricapitolando, $u(t) \rightarrow -\infty$ per $t \rightarrow M$ per ogni $u_0 \leq -1$, ma anche per valori leggermente superiori a -1 .

Sia adesso \bar{u}_0 un valore tale che la soluzione corrispondente, che chiamiamo \bar{u} , tenda a $+\infty$: allora esiste un qualche $\bar{t} > 0$ tale che $\bar{u}(\bar{t}) > 1$. Per dipendenza continua dal dato iniziale, esiste un intorno di \bar{u}_0 per il quale la soluzione u verifica $|u(\bar{t}) - \bar{u}(\bar{t})| < \frac{1}{2}$. Per quanto detto sopra, per ciascun u_0 in tale intorno la soluzione deve esplodere all'infinito. In altre parole, l'insieme degli u_0 per i quali la soluzione esplosa all'infinito è un aperto. Per unicità della soluzione, inoltre, una soluzione che parta più in alto deve rimanere più in alto, e dunque la soluzione esplosa all'infinito per tutti gli u_0 contenuti in una semiretta aperta, diciamo $(\beta, +\infty)$. Lo stesso identico ragionamento si può fare per soluzioni che tendano a $-\infty$, e quindi l'insieme degli u_0 per i quali la soluzione esplosa a $-\infty$ è un aperto del tipo $(-\infty, \alpha)$. Ovviamente $\alpha \leq \beta$.

Per ogni $\alpha \leq u_0 \leq \beta$, la soluzione non può tendere a $-\infty$ e nemmeno a $+\infty$, e dunque per quanto visto sopra è costretta e restare confinata tra -1 e 0 . Si tratta quindi di una soluzione limitata; d'altra parte la F è Lipschitziana su tutti gli insiemi del tipo $[0, T] \times [-1, 0]$, e quindi una soluzione confinata tra -1 e 0 deve esistere almeno fino all'istante $t = T$; dal momento che T è generico, questo vuol dire che le soluzioni limitate sono tutte globali.

Come appena notato, tutte le soluzioni limitate, dunque quelle corrispondenti a $u_0 \in [\alpha, \beta]$, sono soluzioni globali. D'altra parte per qualunque soluzione non limitata si ha $|u'(t)| \geq \frac{u(t)^2}{2}$ per t abbastanza grande, e visto che le funzioni non nulle per le quali $u'(t) = \pm \frac{u(t)^2}{2}$ esplodono in tempo finito si ottiene per confronto che tutte le soluzioni non limitate esplodono in tempo finito, ossia non sono globali.

Concludiamo mostrando che $\alpha = \beta$, ossia che c'è un'unica soluzione limitata. Per farlo, consideriamo una soluzione u , e supponiamo che per un qualche $\bar{t} > 0$ si abbia $u(\bar{t}) < 0$ e $u'(\bar{t}) < 0$. Allora, per tempi poco superiori a \bar{t} , si ha che la u è diminuita, e quindi sono diminuiti sia il termine $e^{tu(t)}$ che il termine $-u(t)^2$; ossia, per tempi poco superiori a \bar{t} la $u'(t)$ è più piccola di $u'(\bar{t})$, e quindi negativa. In altre parole, una soluzione che sia negativa e decrescente in un qualunque punto resta decrescente per sempre, e quindi deve tendere a $-\infty$ per quanto abbiamo visto. Questo vuol dire che tutte le soluzioni globali e limitate, quindi corrispondenti ad $u_0 \in [\alpha, \beta]$, sono comprese tra -1 e 0 e sono crescenti. Questo assicura che abbiano un limite all'infinito, e come già visto tale limite deve essere 0 . Supponiamo ora che ci siano due diverse soluzioni di questo tipo, chiamiamole u e v , con $v > u$. Si ha allora che

$$(v - u)'(t) = v'(t) - u'(t) = e^{tv(t)} - e^{tu(t)} + (u(t)^2 - v(t)^2) > u(t)^2 - v(t)^2 > 0$$

Due diverse soluzioni globali, quindi, hanno differenza che aumenta sempre. Visto che le soluzioni globali devono tendere a 0 , è impossibile che ce ne siano due diverse, perché la loro differenza

dovrebbe aumentare e al tempo stesso tendere a 0; si ha cioè un'unica soluzione globale e limitata, ossia $\alpha = \beta$.

Ci sono annate in cui gli studenti pensano che fare i conti... sporchi le mani. Se non ti sporchi le mani la verdura non cresce.

Memorie agresti di Aldo Pratelli

Esercizio 5. Consideriamo il problema di Cauchy

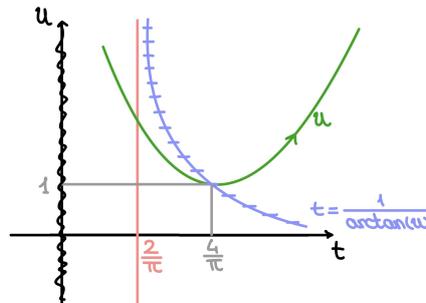
$$\begin{cases} u'(t) = \arctan(u) - \frac{1}{t} \\ u(\frac{4}{\pi}) = 1 \end{cases}$$

Soluzione. Innanzitutto la funzione F che definisce l'equazione è di classe C^∞ , e in particolare è localmente lipschitziana, quindi il problema ammette un'unica soluzione massimale. Osserviamo che la soluzione massimale è definita su $(0, +\infty)$: infatti, per $t \in [c, +\infty)$, con $c > 0$, la funzione $f(u, t)$ rimane limitata, quindi la soluzione u non può scoppiare in tempo finito; poiché c è generico, otteniamo l'esistenza su $(0, +\infty)$.

Dopodiché studiamo gli zeri di F , ossia i punti in cui u passa con derivata orizzontale, e otteniamo:

$$F(t, u(t)) = 0 \iff t = \frac{1}{\arctan(u)} .$$

Da questo otteniamo anche informazioni su quando u è crescente, cioè quando $t > \frac{1}{\arctan(u)}$, e su quando u è decrescente, cioè quando $t < \frac{1}{\arctan(u)}$. Inoltre $F(\frac{4}{\pi}, 1) = 0$.



A questo punto cerchiamo di studiare il comportamento di u agli estremi del dominio. Per $t > 2$, poiché u è crescente, abbiamo che $u'(t) > 1 - \frac{1}{2} = \frac{1}{2}$, da cui otteniamo

$$u(t) = u(2) + u'(\xi)(t - 2) \geq u(2) + \frac{1}{2}(t - 2) \quad \text{da cui} \quad \lim_{t \rightarrow +\infty} u(t) = +\infty .$$

Per studiare il limite in 0, cerchiamo di costruire una *sottosoluzione* per $t < \frac{4}{\pi}$, ossia una funzione y tale che

$$y'(t) \geq F(t, u(t)) \quad \forall t .$$

In questo modo, se $y(\frac{4}{\pi}) \leq u(\frac{4}{\pi})$, avremo $y(t) < u(t) \quad \forall t < \frac{4}{\pi}$; se troviamo y che tende a $+\infty$ per $t \rightarrow 0^+$, potremo dedurre che il comportamento di u è lo stesso. Dopo alcuni tentativi, consideriamo la funzione

$$y_0(t) = -\log t + \frac{\pi}{2}t + c ,$$

che soddisfa l'uguaglianza e che chiaramente tende a $+\infty$ per $t \rightarrow 0^+$; quindi, scelto c tale che $y_0(\frac{4}{\pi}) \leq u(\frac{4}{\pi})$, otteniamo infine che

$$\lim_{t \rightarrow 0^+} u(t) = +\infty .$$

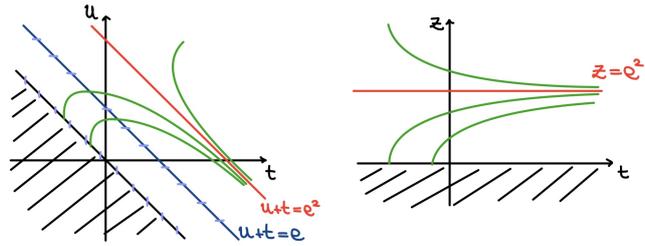
Esercizio 6. Consideriamo l'equazione differenziale

$$u'(t) = 1 - \log(t + u)$$

Studiamo l'intervallo di esistenza delle sue soluzioni e mostriamo che sono tutte asintotiche a una stessa soluzione.

Soluzione. Osserviamo innanzitutto che l'equazione è definita sul semipiano $u + t > 0$ e la funzione F è in particolare C^1 , quindi per ogni dato iniziale esiste unica una soluzione massimale.

Per $u + t \rightarrow 0$ vale $F(u(t), t) \rightarrow +\infty$ e per $u + t = e$ vale $F(u(t), t) = 0$; inoltre osserviamo che $u(t) = e^2 - t$ è soluzione dell'equazione. Quindi tutte le soluzioni con dato iniziale $0 < u_0 < e$ cominciano con derivata verticale sulla retta $u + t = 0$ e poi attraversano la retta $u + t = e$ con derivata orizzontale



Operando il cambio di variabile $z = u + t$, otteniamo che $z' = u' + 1$, quindi l'equazione differenziale diventa $z' = 2 - \log(z)$. Come prima, $z = e^2$ è soluzione dell'equazione, e per unicità le altre soluzioni non possono intersecarla. Quindi per le soluzioni con dato iniziale $z_0 < e^2$ vale sempre $z(t) < e^2$, da cui $z'(t) > 0$: perciò queste soluzioni sono crescenti. Detta u una di queste soluzioni, allora u ammette limite $l \in (0, e^2]$: per il criterio dell'asintoto, abbiamo $\lim_{t \rightarrow +\infty} z'(t) = 2 - \log(l) = 0$, cioè $l = e^2$.

Analogamente per le soluzioni con dato iniziale $z_0 > e^2$ vale sempre $z(t) > e^2$, cioè $z'(t) < 0$, quindi queste soluzioni sono decrescenti. Con lo stesso ragionamento di prima, si conclude che queste soluzioni hanno limite all'infinito e^2 .

Ricordando il cambio di variabile, abbiamo ottenuto che tutte le soluzioni sono asintotiche alla soluzione $u(t) = e^2 - t$.

5.3 Ordine superiore

Consideriamo un problema di Cauchy di ordine n , cioè della forma

$$\begin{cases} u^{(n)}(t) = F(t, u(t), \dots, u^{(n-1)}(t)) \quad \forall t \\ u(t_0) = u_0 \\ u'(t_0) = u'_0 \\ \vdots \\ u^{(n-1)}(t_0) = u_0^{(n-1)} \end{cases}$$

con $u : [0, T] \rightarrow \mathbb{R}^d$, $F : \mathbb{R} \times \mathbb{R}^d \times \dots \times \mathbb{R}^d \rightarrow \mathbb{R}^d$.

In generale, per risolverlo, lo si può ricondurre ad un problema di Cauchy del primo ordine. Infatti, ponendo $v(t) = (u(t), u'(t), \dots, u^{(n-1)}(t))$, otteniamo

$$v'(t) = \begin{pmatrix} u'(t) \\ u''(t) \\ \vdots \\ u^{(n)}(t) \end{pmatrix} = \begin{pmatrix} v_2 \\ v_3 \\ \vdots \\ F(t, v(t)) \end{pmatrix}.$$

Abbiamo cioè formulato il problema di Cauchy equivalente

$$\begin{cases} v(0) = \begin{pmatrix} u_0 \\ u'_0 \\ \vdots \\ u_0^{(n-1)} \end{pmatrix} \\ v'(t) = G(t, v(t)) \end{cases}$$

dove, detto $a = (a_0, a_1, \dots, a_{n-1})$,

$$G(t, a) = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ F(t, a) \end{pmatrix}.$$

Osserviamo infine che G è Lipschitz nella seconda variabile se e solo se F è Lipschitz nelle variabili dalla seconda alla n -esima.

5.4 Sistemi di equazioni differenziali

Definizione 5.5 (Sistema lineare di equazioni differenziali). Siano $I \subseteq \mathbb{R}$, $u : I \rightarrow \mathbb{R}^d$, $A : I \rightarrow M(d, \mathbb{R})$ e $b : I \rightarrow \mathbb{R}^d$. Un sistema lineare di equazioni differenziali è un'equazione differenziale della forma

$$u'(t) = A(t)u(t) + b(t)$$

Se $A(t) \equiv A$, il sistema si dice a coefficienti costanti.

Se $b(t) \equiv 0$, il sistema si dice omogeneo.

5.4.1 Esponenziale di matrici

Consideriamo il sistema omogeneo a coefficienti costanti

$$u' = Au$$

con $A \in M(d, \mathbb{R})$ e cerchiamone le soluzioni $u : I \rightarrow \mathbb{R}^d$.

Definizione 5.6 (Esponenziale di matrice). Data $A \in M(d, \mathbb{R})$, l'esponenziale di A è

$$e^A = \sum_{k=0}^{+\infty} \frac{A^k}{k!} .$$

Osservazione 5.9. Questa è una buona definizione: infatti la serie è convergente in quanto assolutamente convergente (considerando una qualsiasi norma di matrice su $M(d, \mathbb{R})$).

Proposizione 5.10

Valgono le seguenti:

- i. se $AB = BA$, allora $e^{A+B} = e^A e^B = e^B e^A$;
- ii. se M è invertibile, allora $e^{M^{-1}AM} = M^{-1}e^A M$.

Dimostrazione. i. Osserviamo innanzitutto che, poiché A e B commutano, vale la regola del binomio di Newton, ossia

$$(A + B)^k = \sum_{i=0}^k \binom{k}{i} A^i B^{k-i} .$$

Quindi abbiamo:

$$\begin{aligned} e^{A+B} &= \sum_{k=0}^{+\infty} \frac{(A+B)^k}{k!} = \sum_{k=0}^{+\infty} \sum_{i=0}^k \frac{1}{k!} \binom{k}{i} A^i B^{k-i} \\ e^A e^B &= \left(\sum_{i=0}^{+\infty} \frac{A^i}{i!} \right) \left(\sum_{j=0}^{+\infty} \frac{B^j}{j!} \right) = \sum_{k=0}^{+\infty} \sum_{i+j=k} \frac{A^i B^j}{i! j!} = \sum_{k=0}^{+\infty} \sum_{i=0}^k \frac{A^i B^{k-i}}{i!(i-k)!} . \end{aligned}$$

ii. Se M è invertibile, allora

$$e^{M^{-1}AM} = \sum_{k=0}^{+\infty} \frac{(M^{-1}AM)^k}{k!} = \sum_{k=0}^{+\infty} \frac{M^{-1}A^k M}{k!} = M^{-1}e^A M .$$

□

Proposizione 5.11

La funzione $t \mapsto e^{At}$ è differenziabile $\forall t \in \mathbb{R}$ e vale

$$\frac{d}{dt}e^{At} = Ae^{At} .$$

Dimostrazione. Abbiamo:

$$\begin{aligned} \frac{d}{dt}e^{At} &= \lim_{h \rightarrow 0} \frac{e^{A(t+h)} - e^{At}}{h} = e^{At} \lim_{h \rightarrow 0} \frac{e^{Ah} - I}{h} = \\ &= e^{At} \lim_{h \rightarrow 0} \frac{1}{h} \left(\sum_{k=0}^{+\infty} \frac{(Ah)^k}{k!} - I \right) = e^{At} \lim_{h \rightarrow 0} \sum_{k=1}^{+\infty} \frac{A^k h^{k-1}}{k!} = e^{At} \lim_{h \rightarrow 0} \left(A + h \sum_{i=0}^{+\infty} \frac{A^{i+2} h^i}{(i+2)!} \right) = \\ &= Ae^{At} \lim_{h \rightarrow 0} \left(I + h \sum_{i=0}^{+\infty} \frac{A^{i+1} h^i}{(i+2)!} \right) = Ae^{At} \end{aligned}$$

dove l'ultima uguaglianza vale in quanto la serie converge uniformemente, poiché può essere stimata con una geometrica. \square

Corollario 5.12

L'unica soluzione di $\begin{cases} u'(t) = Au(t) \\ u(t_0) = u_0 \end{cases}$ è $u(t) = u_0 e^{A(t-t_0)}$.

Per il calcolo effettivo dell'esponenziale di matrice, consideriamo solo matrici con autovalori reali e distinguiamo due situazioni.

Se la matrice A è diagonalizzabile, con spettro $\sigma(A) = \{\lambda_1, \dots, \lambda_n\}$, allora esiste una matrice di cambio di base M tale che

$$M^{-1}AM = \begin{pmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_n \end{pmatrix} =: D$$

Quindi abbiamo

$$e^A = Me^D M^{-1} = M \begin{pmatrix} e^{\lambda_1} & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & e^{\lambda_n} \end{pmatrix} M^{-1}$$

Altrimenti, consideriamo un singolo blocco di Jordan $J(\lambda) = \lambda I + N$, dove N è la matrice nilpotente con 1 sulla sopradiagonale. Allora $e^{J(\lambda)} = e^\lambda e^N$, quindi basta calcolare l'esponenziale per le nilpotenti (che quindi è una somma finita di termini):

$$e^N = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{3!} & \dots & \frac{1}{(n-1)!} \\ 0 & 1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & \frac{1}{3!} \\ \vdots & & \ddots & 1 & \frac{1}{2} \\ 0 & \dots & \dots & 0 & 1 \end{pmatrix}$$

Consideriamo infine, a titolo di esempio, il caso di una matrice reale 2×2 con autovalori non reali.

Esempio 5.13. Sia A il blocco di Jordan reale

$$A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} ,$$

che ha polinomio caratteristico $p(\lambda) = \det(A - \lambda I) = \lambda^2 + 1$, e quindi ha autovalori $\pm i$. Osserviamo inoltre che valgono le relazioni $A^0 = I, A^1 = A, A^2 = -I, A^3 = -A, A^4 = I$, perciò possiamo scrivere:

$$e^{At} = \sum_{k=0}^{+\infty} \frac{(tA)^k}{k!} = I \sum_{k=0}^{+\infty} \frac{t^{2k}(-1)^k}{(2k)!} + A \sum_{k=0}^{+\infty} \frac{t^{2k+1}(-1)^k}{(2k+1)!} = I \cos(t) + A \sin(t) = \begin{pmatrix} \cos(t) & -\sin(t) \\ \sin(t) & \cos(t) \end{pmatrix} .$$

5.4.2 Sistemi autonomi

Definizione 5.7 (Sistema autonomo). Dato un aperto connesso $\Omega \subseteq \mathbb{R}^d$ e una funzione $F \in C^1(\Omega, \mathbb{R}^d)$, un *sistema autonomo* (\mathbf{S}) è un'equazione differenziale del tipo $u'(t) = F(u(t))$, ossia un'equazione differenziale in cui u' non dipende direttamente dal parametro t .

Osservazione 5.14. Se $u(t)$ è soluzione del sistema, allora $v_c(t) = u(t - c)$ è ancora soluzione, infatti

$$\frac{d}{dt} v_c(t) = \frac{du}{dt}(t - c) = F(u(t - c)) = F(v_c(t)) .$$

Di conseguenza se u e v sono due soluzioni ed esistono t_1, t_2 tali che $u(t_1) = v(t_2)$, allora vale $v(t) = u(t + t_1 - t_2)$, infatti questa è effettivamente soluzione e coincide con v in t_2 .

Definizione 5.8 (Traiettoria). Una *traiettoria* di un sistema autonomo \mathbf{S} è l'immagine di una soluzione massimale, ovvero se $u : (\alpha, \beta) \rightarrow \Omega \subseteq \mathbb{R}^d$ è soluzione massimale di \mathbf{S} , allora l'insieme $\{u(t) \mid t \in (\alpha, \beta)\} \subseteq \Omega$ definisce una *traiettoria* di \mathbf{S} .

Osservazione 5.15. Le diverse traiettorie non si intersecano, e anzi Ω è rivestito da traiettorie (naturalmente orientate). Inoltre se u è una soluzione del sistema ed esistono due punti t_1, t_2 distinti tali che $u(t_1) = u(t_2)$, allora u è necessariamente periodica e definita su tutto \mathbb{R} .

Dato $\Omega \subseteq \mathbb{R}^d$, $F \in C^1(\Omega, \mathbb{R}^d)$, denotiamo \mathbf{SF} il sistema di equazione $u'(t) = F(u(t))$; data una funzione scalare $\lambda \in C^1(\Omega, \mathbb{R})$ sia $G \in C^1(\Omega, \mathbb{R}^d)$ la funzione definita da $G(x) = \lambda(x)F(x) \forall x \in \Omega$, e denotiamo \mathbf{SG} il sistema di equazione $u'(t) = G(u(t))$. Allora \mathbf{SF} e \mathbf{SG} hanno le stesse traiettorie: se v è soluzione di \mathbf{SG} , allora $u(t) := v(\psi(t))$ risolve \mathbf{SF} con la condizione $u(0) = v(t_0)$ dove

$$\varphi(t) = \int_{t_0}^t \lambda(v(\sigma)) d\sigma \quad \text{e} \quad \psi = \varphi^{-1} \text{ inversa insiemistica.}$$

Infatti

$$\begin{aligned} u'(t) &= v'(\psi(t))\psi'(t) = G(v(\psi(t)))\psi'(t) = \lambda(v(\psi(t)))\psi'(t)F(v(\psi(t))) = \\ &= \varphi'(\psi(t))\psi'(t)F(v(\psi(t))) = (\varphi \circ \psi)'(t)F(v(\psi(t))) = (\text{id}_{\mathbb{R}})'(t)F(v(\psi(t))) = \\ &= F(v(\psi(t))) = F(u(t)) \end{aligned}$$

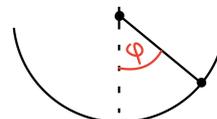
e inoltre $u_0 = v(\psi(0)) = v(\psi(\varphi(t_0))) = v((\psi \circ \varphi)(t_0)) = v(t_0)$.

Definizione 5.9 (Costante del moto). Una funzione $H : \Omega \rightarrow \mathbb{R}$ si dice *costante del moto* (o *integrale del moto*) per il sistema autonomo \mathbf{S} se esiste una costante $c \in \mathbb{R}$ tale che per ogni u soluzione di \mathbf{S} vale $H(u(t)) = c \forall t$.

Osserviamo che H è una costante del moto se e solo se

$$\frac{d}{dt} H(u(t)) = 0 \quad \text{ossia} \quad \nabla H(u(t)) \cdot u'(t) = \nabla H(u(t)) \cdot F(u(t)) = 0 .$$

Studiamo un esempio significativo di sistema autonomo, ossia l'equazione del pendolo di Newton



$$\varphi''(t) = -\sin \varphi(t) .$$

Innanzitutto ponendo $q := \varphi$ e $p := \varphi'$ possiamo riformulare il problema nel sistema equivalente

$$(\mathbf{S}) = \begin{cases} q'(t) = p(t) \\ p'(t) = -\sin q(t) \end{cases}$$

che ammette come integrale del moto la funzione $H(p, q) = \frac{1}{2}p^2 + (1 - \cos q)$: infatti se la funzione $(p(t), q(t))$ soddisfa \mathbf{S} , allora

$$\begin{aligned} \frac{d}{dt}H(p(t), q(t)) &= \frac{\partial H}{\partial p}(p(t), q(t)) \cdot p'(t) + \frac{\partial H}{\partial q}(p(t), q(t)) \cdot q'(t) = \\ &= p(t) \cdot (-\sin q(t)) + \sin q(t) \cdot p(t) = 0 . \end{aligned}$$

In effetti il sistema \mathbf{S} con la funzione H risulta un caso particolare di un'interessante famiglia di sistemi:

Definizione 5.10 (Sistema hamiltoniano). Dato $\Omega \subseteq \mathbb{R}^n \times \mathbb{R}^n$ un aperto connesso, siano $F, G \in C^1(\Omega, \mathbb{R}^n)$ due funzioni: il sistema dinamico

$$(\mathbf{S}) = \begin{cases} q'(t) = F(p(t), q(t)) \\ p'(t) = G(p(t), q(t)) \end{cases}$$

si dice *hamiltoniano* se esiste una funzione $H : \Omega \rightarrow \mathbb{R}^n$ tale che

$$q'(t) = \frac{\partial H}{\partial p}(p(t), q(t)) \quad \text{e} \quad p'(t) = -\frac{\partial H}{\partial q}(p(t), q(t)) ;$$

in tal caso H definisce un integrale del moto.

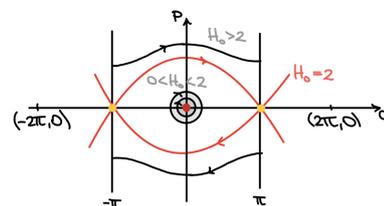
Per quanto detto il sistema \mathbf{S} del pendolo di Newton è un sistema hamiltoniano con la costante del moto

$$H(p, q) = \frac{1}{2}p^2 + (1 - \cos q) .$$

Vediamo come si comportano le soluzioni al variare di $H_0 := H(p(t), q(t))$. La funzione H è non negativa per ogni (p, q) ed ha minimo in uguale a 0, di conseguenza se $H_0 < 0$ si ha $\{H(p, q) = H_0\} = \emptyset$; se $H_0 = 0$ allora $\{H(p, q) = H_0\} = \{(2k\pi, 0)\}$ ed in corrispondenza di questi valori troviamo le soluzioni costanti $\varphi(t) = 2k\pi$; se invece $H_0 > 0$ troviamo

$$p = \pm \sqrt{2(H_0 + \cos q - 1)} \quad \text{da cui} \quad H_0 + \cos q - 1 \geq 0 .$$

Di conseguenza se $H_0 \geq 2$ non ci sono limitazioni su q , e in particolare per $H_0 = 2$ troviamo le soluzioni costanti $\varphi(t) = \pi + 2k\pi$; se invece $0 < H_0 < 2$ troviamo la limitazione $\cos q \geq 1 - H_0$, ovvero $q \in [-\arccos(1 - H_0), \arccos(1 - H_0)]$.



Nel seguito siano $\Omega \subseteq \mathbb{R}^d$ aperto, $F \in C^1(\Omega, \mathbb{R}^d)$, \mathbf{S} il sistema autonomo $u' = F(u)$ e fissato il dato iniziale $u(0) = \xi \in \Omega$ sia $\varphi_\xi(t) = u(t)$ la relativa soluzione.

Definizione 5.11 (Punto di equilibrio). Un punto $\xi_0 \in \Omega$ si dice *punto di equilibrio* (o *punto critico*) del sistema \mathbf{S} se vale $F(\xi_0) = 0$. Se ξ_0 è punto di equilibrio, φ_{ξ_0} è la funzione costantemente ξ_0 .

Definizione 5.12 (Stabilità). Un punto di equilibrio $\xi_0 \in \Omega$ si dice *punto di equilibrio stabile* per il sistema \mathbf{S} se $\forall R > 0 \exists r > 0$ tale che $|\xi - \xi_0| < r \implies |\varphi_\xi(t) - \xi_0| < R \forall t \geq 0$.

Osservazione 5.16. Se ξ_0 è un punto di equilibrio stabile, scelti $R > 0$ e ξ tale che $|\xi - \xi_0| < r$, dove r è la tolleranza di stabilità data dalla definizione, allora la funzione $\varphi_\xi(t)$ è definita $\forall t \geq 0$.

Definizione 5.13 (Asintotica stabilità). Un punto di equilibrio stabile $\xi_0 \in \Omega$ si dice *asintoticamente stabile* se

$$\exists r > 0 \text{ tale che } \forall \xi \in B(\xi_0, r) \text{ vale } \lim_{t \rightarrow +\infty} |\varphi_\xi(t) - \xi_0| = 0 .$$

Introduciamo un attrito nell'equazione del pendolo di Newton, ossia considerando l'equazione $\varphi''(t) = -\sin \varphi(t) - \varepsilon \varphi'(t)$, $\varepsilon > 0$, a cui associamo il sistema

$$(\mathbf{S}_\varepsilon) = \begin{cases} q'(t) = p(t) \\ p'(t) = -\sin q(t) - \varepsilon p(t) \end{cases} ;$$

in questo caso la funzione H non è più un integrale del moto, ma vale

$$\frac{d}{dt} H(p(t), q(t)) = -\varepsilon p^2 ,$$

e dunque, come vedremo attraverso il Lemma di Ljapunov, l'origine diventa un punto di equilibrio asintoticamente stabile.

Definizione 5.14 (Bacino di attrazione). Dato ξ_0 un punto di equilibrio asintoticamente stabile, definiamo il suo *bacino di attrazione* come l'insieme

$$\Omega_{\xi_0} = \{ \xi \in \Omega \mid |\varphi_\xi(t) - \xi_0| \xrightarrow{t \rightarrow +\infty} 0 \} .$$

Osserviamo che il bacino di attrazione di un punto di equilibrio asintoticamente stabile ξ_0 è per definizione un intorno di ξ_0 , ma in realtà è addirittura un aperto: infatti poiché ξ_0 è asintoticamente stabile, $\exists r > 0$ tale che $B_r(\xi_0) \subseteq \Omega_{\xi_0}$; ora se $\xi \in \Omega_{\xi_0}$, in particolare $\exists t_0 > 0$ tale che $|\varphi_\xi(t_0) - \xi_0| < \frac{r}{2}$, e d'altra parte, per dipendenza continua dal dato iniziale, la funzione $\Psi : \Omega \rightarrow \Omega$ data da $\zeta \mapsto \varphi_\zeta(t_0)$ è continua, quindi in particolare $\exists \delta > 0$ tale che se $\zeta \in B_\delta(\xi)$ allora $|\varphi_\zeta(t_0) - \varphi_\xi(t_0)| < \frac{r}{2}$ e dunque complessivamente se $\zeta \in B_\delta(\xi)$ allora $|\varphi_\zeta(t_0) - \xi_0| < r$. Basta allora osservare che la soluzione che ha dato iniziale $\varphi_\zeta(t_0)$ è una traslazione della soluzione che ha per dato iniziale ζ , infatti

$$\varphi_{\varphi_\zeta(t_0)}(t) = \varphi_\zeta(t + t_0)$$

e di conseguenza $\varphi_\zeta(t_0) \in \Omega_{\xi_0}$ implica $\zeta \in \Omega_{\xi_0}$.

Proposizione 5.17

Sia \mathbf{S} un sistema lineare di equazione $F(u) = Au$, con $A \in \text{GL}_n(\mathbb{R})$, ossia 0 è l'unico punto critico del sistema; allora:

- i. 0 è un punto di equilibrio asintoticamente stabile sse $\forall \lambda \in \sigma(A)$ vale $\Re(\lambda) < 0$.
- ii. 0 è un punto di equilibrio stabile sse $\forall \lambda \in \sigma(A)$ vale $\Re(\lambda) \leq 0$ e per $\Re(\lambda) = 0$ tutti i blocchi di Jordan relativi all'autovalore λ hanno dimensione minima.

Definizione 5.15 (Funzione di Ljapunov). Sia $\Omega \subseteq \mathbb{R}^d$, $F \in C^1(\Omega, \mathbb{R}^d)$, sia \mathbf{S} il sistema $u' = F(u)$ e sia $\xi_0 \in \Omega$ un punto di equilibrio per il sistema. Una funzione $V \in C^1(\Omega, \mathbb{R})$ si dice *funzione di Ljapunov* per il sistema \mathbf{S} se

- i. $V(\xi_0) = 0$ ma $V(x) > 0 \forall x \neq \xi_0$;
- ii. $\nabla V(x) \cdot F(x) \leq 0 \forall x \in \Omega$.

Proposizione 5.18: Lemma di Ljapunov

Dato $\xi_0 \in \Omega$ punto di equilibrio del sistema, se esistono Ω' intorno di ξ_0 e $V \in C^1(\Omega', \mathbb{R})$ una funzione di Ljapunov per \mathbf{S} , allora ξ_0 è punto di equilibrio stabile. Se inoltre vale la condizione

$$ii.^+ \quad \nabla V(x) \cdot F(x) < 0 \quad \forall x \in \Omega' \setminus \{\xi_0\},$$

allora ξ_0 è asintoticamente stabile.

Dimostrazione. Per ipotesi Ω' è intorno di ξ_0 , di conseguenza $\exists R > 0$ tale che $\overline{B(\xi_0, R)} \subseteq \Omega'$. Poniamo a questo punto

$$m := \inf_{|x-\xi_0|=R} V(x) = V(x_0) > 0$$

per qualche x_0 . Scegliamo $\varepsilon \in (0, m)$ arbitrario e osserviamo che (per continuità di V) l'insieme $\{x \in \Omega' \mid V(x) < \varepsilon\}$ è un intorno aperto di ξ_0 , e dunque contiene una palla $B(\xi_0, r)$ per qualche $r > 0$: imponiamo $r < R$ e mostriamo che allora $\forall \xi \in B(\xi_0, r)$ vale $|\varphi_\xi(t) - \xi_0| < R \forall t \geq 0$. Dato infatti $\xi \in B(\xi_0, r)$ possiamo definire la funzione ausiliaria $\psi(t) := V(\varphi_\xi(t))$, e derivando

$$\psi'(t) = \nabla V(\varphi_\xi(t)) \cdot \varphi'_\xi(t) = \nabla V(\varphi_\xi(t)) \cdot F(\varphi_\xi(t)) \leq 0$$

ossia ψ è debolmente decrescente: poiché $\psi(0) = V(\xi) < \varepsilon$, $\forall t \geq 0$ vale $\psi(t) < \varepsilon < m$. Allora $\forall t \geq 0$ vale $\varphi_\xi(t) \in B(\xi_0, R)$, altrimenti per continuità esisterebbe \bar{t} tale che $\varphi_\xi(\bar{t}) \in \partial B(\xi_0, R)$ da cui $V(\varphi_\xi(\bar{t})) \geq m$. ∇

Abbiamo allora provato che ξ_0 è punto di equilibrio stabile e $\psi'(t) \leq 0$. Assumiamo a questo punto l'ipotesi $ii.^+$. Osserviamo che $\psi(t) \geq 0$, dunque

$$c_0 := \inf_{t>0} \psi(t) = \lim_{t \rightarrow +\infty} \psi(t) \geq 0 .$$

Supponiamo per assurdo che $c_0 > 0$: allora l'insieme $\{x \in \Omega' \mid V(x) < c_0\}$ è intorno aperto di ξ_0 , pertanto contiene una palla $B(\xi_0, \delta)$, dove scegliamo $\delta < R$; per costruzione $\varphi_\xi(t) \notin B(\xi_0, \delta)$ ($V(\varphi_\xi(t)) = \psi(t) \geq c_0$). Poniamo

$$\sigma := \max_{\delta \leq |x-\xi_0| \leq R} \nabla V(x) \cdot F(x) < 0 ,$$

per cui $\psi'(t) = \nabla V(\varphi_\xi(t)) \cdot F(\varphi_\xi(t)) \leq \sigma$ e dunque $\psi(t) \leq \psi(0) + \sigma t \rightarrow -\infty$ per $t \rightarrow +\infty$. ∇ Dunque necessariamente deve valere $c_0 = 0$. Infine sia $\delta > 0$ e poniamo

$$c_\delta = \min_{\delta \leq |x-\xi_0| \leq R} V(x) > 0 ;$$

poiché $\psi(t) = V(\varphi_\xi(t)) \rightarrow 0$, $\exists t_\delta$ tale che $\forall t > t_\delta$ $\psi(t) < \frac{c_\delta}{2}$, e dunque $|\varphi_\xi(t) - \xi_0| < \delta$, vale a dire $\varphi_\xi(t) \rightarrow \xi_0$. \square

Vediamo un esercizio esemplare.

Esercizio 7. Si studino al variare di $A \in GL_2(\mathbb{R})$ le soluzioni del sistema autonomo $u' = Au$.

Soluzione. Esplicitiamo il sistema: ponendo

$$u = \begin{pmatrix} x \\ y \end{pmatrix} \quad \text{e} \quad A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

dove $\det A \neq 0$, troviamo che l'equazione differenziale è equivalente al sistema

$$\begin{cases} x' = ax + by \\ y' = cx + dy \end{cases} \quad \text{con } a, b, c, d \in \mathbb{R}, \quad ad - bc \neq 0.$$

Dall'algebra lineare sappiamo che $p_A(\lambda) = \lambda^2 - T\lambda + D$, dove indichiamo con p_A il polinomio caratteristico di A e poniamo $T = \text{tr } A$ e $D = \det A$; le radici del polinomio sono dunque

$$\lambda^\pm = \frac{T}{2} \pm \sqrt{\left(\frac{T}{2}\right)^2 - D}.$$

Ricordiamo che $\lambda^- + \lambda^+ = T$ mentre $\lambda^- \lambda^+ = D$.

Se $D < 0$ gli autovalori sono reali e discordi, con $\lambda^- < 0 < \lambda^+$, quindi a meno di un cambio di coordinate il sistema si diagonalizza in

$$\begin{cases} x' = \lambda^+ x \\ y' = \lambda^- y \end{cases} \quad \text{che ha soluzione} \quad \begin{cases} x(t) = x_0 e^{\lambda^+ t} \\ y(t) = y_0 e^{\lambda^- t} \end{cases};$$

lungo la direzione x il moto è espansivo mentre lungo la y è contrattivo; l'origine $O = (0, 0)$ si dice *punto di sella* ed è un punto di equilibrio instabile.

Se $0 < D < (\frac{T}{2})^2$, le soluzioni sono reali, distinte e concordi, pertanto la soluzione del sistema diagonalizzato è ancora

$$\begin{cases} x(t) = x_0 e^{\lambda^+ t} \\ y(t) = y_0 e^{\lambda^- t} \end{cases};$$

se gli autovalori sono positivi il moto è espansivo in entrambe le direzioni, altrimenti è contrattivo in entrambe le direzioni: nel primo caso si dice che l'origine O è un *nodo instabile*, nel secondo che è un *nodo stabile*.

Se $D > (\frac{T}{2})^2$, gli autovalori sono complessi e scriviamo $\lambda^\pm = \alpha \pm i\beta$, dove $\alpha = \frac{T}{2}$ e $\beta = \sqrt{D - \alpha^2}$; il sistema può essere portato in forma normale di Jordan reale, ossia la similitudine

$$A \sim \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix} \quad \text{trasforma il sistema in} \quad \begin{cases} x' = \alpha x - \beta y \\ y' = \beta x + \alpha y \end{cases};$$

passando alle coordinate polari possiamo scrivere

$$\begin{pmatrix} x \\ y \end{pmatrix} = r \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} \quad \text{da cui ricaviamo} \quad \begin{pmatrix} x' \\ y' \end{pmatrix} = r' \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} + r\theta' \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix}$$

e anche

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \alpha r \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} + \beta r \begin{pmatrix} -\sin \theta \\ \cos \theta \end{pmatrix};$$

l'equazione della traiettoria si riscrive come

$$\begin{cases} r'(t) = \alpha r(t) \\ \theta'(t) = \beta \end{cases} \quad \text{che ha soluzione} \quad \begin{cases} r(t) = r_0 e^{\alpha t} \\ \theta(t) = \theta_0 + \beta t \end{cases};$$

poiché $\alpha = \frac{T}{2}$, se $T = 0$ troviamo una rotazione uniforme attorno all'origine, e diciamo che O è un *centro*, se $T > 0$ otteniamo una spirale divergente, e diciamo che O è un *fuoco instabile*, mentre se $T < 0$ otteniamo una spirale convergente a O , e diciamo che O è un *fuoco stabile*.

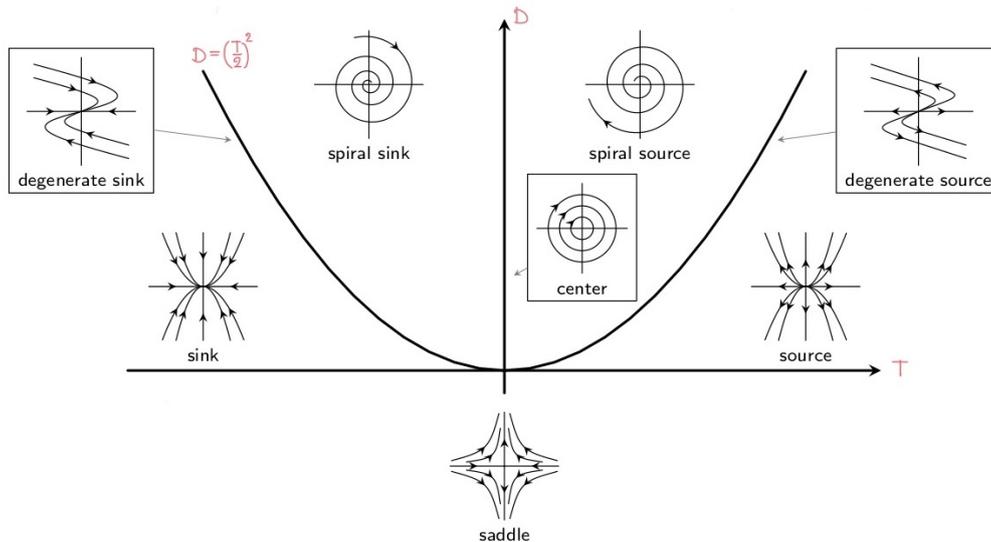
Infine se $D = (\frac{T}{2})^2$, A ha un solo autovalore: dobbiamo allora distinguere due casi a seconda della molteplicità geometrica. Se il sistema è diagonalizzabile, allora lo si può riscrivere come

$$\begin{cases} x' = \lambda x \\ y' = \lambda y \end{cases} \quad \text{che ha soluzione} \quad \begin{cases} x(t) = x_0 e^{\lambda t} \\ y(t) = y_0 e^{\lambda t} \end{cases}$$

e in tal caso l'origine si dice *stella stabile* se $\lambda < 0$ e *stella instabile* se $\lambda > 0$; se invece il sistema non è diagonalizzabile, lo si può trasformare in forma di Jordan trovando

$$\begin{cases} x' = \lambda x + y \\ y' = \lambda y \end{cases} \quad \text{che ha soluzione} \quad \begin{cases} x(t) = (c_1 + c_2 t)e^{\lambda t} \\ y(t) = c_2 e^{\lambda t} \end{cases}$$

e l'origine è detto *nodo degenera*, *stabile* se $\lambda < 0$ e *instabile* se $\lambda > 0$.



Teorema 5.19: Teorema di linearizzazione (Hartman-Grobman)

Siano $\Omega \subseteq \mathbb{R}^d$ aperto, $F \in C^1(\Omega, \mathbb{R}^d)$ e \mathbf{S} il sistema autonomo di equazione $u' = F(u)$; siano inoltre $\xi_0 \in \Omega$ un punto di equilibrio per \mathbf{S} e $A = DF(\xi_0) \in M_d(\mathbb{R})$. Valgono allora le seguenti condizioni sufficienti:

- i. se $\Re(\lambda) < 0 \forall \lambda \in \sigma(A)$, allora ξ_0 è asintoticamente stabile;
- ii. se $\exists \lambda \in \sigma(A)$ tale che $\Re(\lambda) > 0$, allora ξ_0 è instabile.

Dimostrazione. Dimostriamo solo il primo punto con le ipotesi semplificative di A diagonalizzabile e $\sigma(A) = \{\lambda_1, \dots, \lambda_d\} \subseteq \mathbb{R}$. Supponiamo senza perdere generalità $\xi_0 = 0$; per i , esiste $\sigma > 0$ tale che $\lambda_k \leq -\sigma$ per ogni autovalore λ_k . Sia $D = B^{-1}AB = \text{diag}(\lambda_1, \dots, \lambda_d)$. Per il Teorema di Taylor vale $F(u) = Au + \tilde{R}(u)$ dove $\tilde{R}(u) = o(|u|)$; adesso se u è soluzione di \mathbf{S} , poniamo $x(t) = Bu(t)$: derivando otteniamo

$$\begin{aligned} x'(t) &= Bu'(t) = BF(u(t)) = BAu(t) + BR(u(t)) = BAB^{-1}Bu(t) + BR(B^{-1}Bu(t)) = \\ &= Dx(t) + \tilde{R}(x(t)) \quad \text{dove} \quad \tilde{R}(x) = BR(B^{-1}x) = o(|x|), \end{aligned}$$

ossia abbiamo formulato un nuovo sistema \mathbf{N} di equazione $x'(t) = G(x(t))$ dove $G(x) = Dx + \tilde{R}(x)$. Studiamo allora la stabilità di \mathbf{N} , e osserviamo che la mappa $V(x) = \frac{1}{2}|x|^2$ è una funzione di Ljapunov per \mathbf{N} (in un opportuno intorno di $(0, 0)$), infatti chiaramente $V(0) = 0$ e $V(x) > 0$ per $x \neq 0$, e inoltre vale la proprietà ii^+ dato che

$$\nabla V(x) \cdot G(x) = x \cdot (Dx + \tilde{R}(x)) = \sum_{i=1}^d \lambda_i x_i^2 + x \cdot \tilde{R}(x) \leq -\sigma|x|^2 + x \cdot \tilde{R}(x) = |x|^2(-\sigma + o(1))$$

e dunque esiste un intorno Ω' in cui $\nabla V \cdot F$ è negativa. Si conclude allora per il Lemma di Ljapunov. \square

Questa dimostrazione prova anche (sempre sotto ipotesi semplificative) la Proposizione 5.17, infatti chiaramente prova il punto *i*; per quanto riguarda il punto *ii*, osserviamo che in questo caso vale soltanto $\sigma \geq 0$, ma d'altra parte vale $F(u) = Au$ (non a meno di un $o(|u|)$), pertanto ripercorrendo la dimostrazione otteniamo

$$\nabla V(x) \cdot G(x) = x \cdot (Dx) = \sum_{i=1}^d \lambda_i x_i^2 \leq -\sigma |x|^2 \leq 0$$

e concludiamo ancora per il Lemma di Ljapunov.

5.5 PDE: alcuni esempi

Studiamo in questa sezione un esempio di equazione alle derivate parziali, l'*equazione del trasporto*, e alcune sue varianti. Al variare di $a \in \mathbb{R}$, chiamiamo equazione del trasporto il problema differenziale a valori iniziali

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + a \frac{\partial u}{\partial x}(t, x) = 0 \\ u(0, x) = u_0(x) \end{cases} \quad (\mathbf{ET})$$

Definiamo il vettore $\nu = (1 \ a) \in \mathbb{R}^2$ e osserviamo che se u è soluzione del problema, allora

$$\frac{\partial u}{\partial \nu}(t, x) = \frac{\partial u}{\partial t}(t, x) + a \frac{\partial u}{\partial x}(t, x) = 0 \ ,$$

ossia tutte le soluzioni di **ET** sono costanti sulle rette parallele a ν . In particolare deve valere $u(t, x) = u(0, x - at) = u_0(x - at)$, dunque è necessario che $u_0 \in C^1$ (altrimenti certamente non esiste una funzione $u(t, x) \in C^1$ che ristretta all'asse $t = 0$ coincida con u_0); d'altra parte se $u_0 \in C^1$, la funzione definita dalla formula $u(t, x) := u_0(x - at)$ è soluzione: $u(0, x) = u_0(x)$ e

$$\frac{\partial u}{\partial t}(t, x) + a \frac{\partial u}{\partial x}(t, x) = u'_0(x - at) \cdot \frac{\partial(x - at)}{\partial t} + a u'_0(x - at) \cdot \frac{\partial(x - at)}{\partial x} = 0 \ .$$

Se studiamo il problema analogo di ordine superiore, $a \in \mathbb{R}^n$ e

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + a \cdot D_x u(t, x) = 0 \\ u(0, x) = u_0(x) \end{cases} \quad (\mathbf{ET})$$

dove $(t, x) \in \mathbb{R} \times \mathbb{R}^n$ e $D_x u(t, x)$ rappresenta il vettore delle derivate parziali lungo x (vale a dire dalla seconda alla $(n+1)$ -esima), ritroviamo le medesime conclusioni: ponendo $\nu = (1 \ a) \in \mathbb{R}^{n+1}$, se u è soluzione

$$\frac{\partial u}{\partial \nu}(t, x) = \nabla u(t, x) \cdot (1 \ a) = \frac{\partial u}{\partial t}(t, x) + a \cdot D_x u(t, x) = 0 \ ,$$

dunque deve valere $u(t, x) = u_0(x - at)$ e per u_0 di classe C^1 (condizione necessaria e sufficiente) questa descrive effettivamente una soluzione.

Affrontiamo quindi la versione non omogenea del problema: dato $a \in \mathbb{R}^n$

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + a \cdot D_x u(t, x) = f(t, x) \\ u(0, x) = u_0(x) \end{cases} \quad (\mathbf{ET})$$

Tentiamo nuovamente la strada precedente: posto $\nu = (1 \ a) \in \mathbb{R}^{n+1}$, se u è soluzione allora

$$\frac{\partial u}{\partial \nu}(t, x) = \nabla u(t, x) \cdot (1 \ a) = \frac{\partial u}{\partial t}(t, x) + a \cdot D_x u(t, x) = f(t, x) \ ,$$

dunque lungo le rette di direzione ν la funzione u varia con derivata f . Possiamo allora sfruttare lungo tali direzioni il Teorema fondamentale del calcolo: dato un punto $(t, x) \in \mathbb{R}^{n+1}$, definiamo $\omega(\sigma) := u(\sigma, x - at + a\sigma)$ che parametrizza u lungo la retta di direzione ν passante per (t, x) ; valgono dunque $\omega(0) = u(0, x - at)$ e $\omega(t) = (t, x)$. Adesso

$$\begin{aligned}\omega'(\sigma) &= \nabla u(\sigma, x - at + a\sigma) \cdot (1 \ a) = \partial_t u(\sigma, x - at + a\sigma) + a \cdot D_x u(\sigma, x - at + a\sigma) = \\ &= f(\sigma, x - at + a\sigma)\end{aligned}$$

ma allora per il teorema fondamentale del calcolo

$$\omega(t) = \omega(0) + \int_0^t \omega'(\sigma) d\sigma \quad \text{vale a dire} \quad u(t, x) = u_0(x - at) + \int_0^t f(\sigma, x - at + a\sigma) d\sigma .$$

Di conseguenza, per avere soluzione $u \in C^1$ è necessario che f sia continua (in quanto combinazione lineare delle derivate parziali di u) e che u_0 sia C^1 , viceversa se f è continua e $u_0 \in C^1$, la funzione

$$u(t, x) = u_0(x - at) + \int_0^t f(\sigma, x - at + a\sigma) d\sigma$$

risolve effettivamente il problema.

Modifichiamo a questo punto la natura delle condizioni iniziali, e interrogiamoci su quali conclusioni si conservano: dato $\Gamma \subseteq \mathbb{R}^n$ aperto e data una funzione $g \in C^1(\Gamma, \mathbb{R})$ studiamo

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) + a \cdot D_x u(t, x) = f(t, x) \\ u(t_0, x) = g(x) \end{cases} .$$

Il calcolo per trovare la soluzione è analogo al punto precedente, tuttavia in corrispondenza delle rette caratteristiche di direzione $\nu = (1 \ a)$ che non incontrano l'insieme $\{t_0\} \times \Gamma$ non ci sarà unicità di soluzione in quanto viene meno la condizione iniziale; d'altro canto in corrispondenza delle rette di direzione ν che incontrano l'insieme $\{t_0\} \times \Gamma$ più di una volta non è garantita l'esistenza poiché le condizioni iniziali date da g possono non essere compatibili tra loro.

A tal proposito studiamo un ultimo esempio di equazione alle derivate parziali.

Data $u_0 : \mathbb{R} \rightarrow \mathbb{R}$,

$$\begin{cases} \frac{\partial u}{\partial t}(t, x) \sin \theta = \frac{\partial u}{\partial x}(t, x) \cos \theta \\ u(0, x) = u_0(x) \end{cases}$$

dove $\theta = \theta(t, x)$ è l'angolo delle coordinate polari. Definiamo un vettore (funzione) ausiliario $\nu(t, x) = (-\sin \theta, \cos \theta)$ e osserviamo che

$$\frac{\partial u}{\partial \nu}(t, x) = \nabla u(t, x) \cdot \nu = -\frac{\partial u}{\partial t}(t, x) \sin \theta + \frac{\partial u}{\partial x}(t, x) \cos \theta = 0 ;$$

ciò comporta che una soluzione u deve necessariamente essere costante sulle circonferenze di raggio assegnato: parametrizzando la circonferenza di raggio R con la funzione $\varphi(\sigma) = R(\cos \sigma, \sin \sigma)$, otteniamo

$$(u \circ \varphi)'(\sigma) = \nabla u(\varphi(\sigma)) \cdot R \begin{pmatrix} -\sin \sigma \\ \cos \sigma \end{pmatrix} = R \left[-\frac{\partial u}{\partial t}(\varphi(\sigma)) \sin \sigma + \frac{\partial u}{\partial x}(\varphi(\sigma)) \cos \sigma \right] = 0$$

cioè $u \circ \varphi$ è costante.

Pertanto dato un generico punto (t, x) , in tale punto dovrà valere $u(t, x) = u_0(\sqrt{t^2 + x^2})$, e allo stesso tempo anche $u(t, x) = u_0(-\sqrt{t^2 + x^2})$, ossia, per arbitrarietà del punto è necessario che u_0 sia una funzione pari. In queste ipotesi non abbiamo problemi di compatibilità, e per $u_0 \in C^1$ anche u è C^1 su $t > 0$.

Capitolo 6

Integrazione

L'ultimo strumento che vogliamo recuperare dall'analisi in una variabile è l'integrazione. L'integrazione studiata ad Analisi I, alla luce del Teorema fondamentale del calcolo integrale, associava all'integrale di una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$ sia l'insieme delle sue primitive, sia l'area del sottografico della funzione in un determinato intervallo. La teoria dell'integrale indefinito per il calcolo delle primitive era pertanto prima di tutto uno strumento per il calcolo delle aree.

Se devo dipingere casa, trovare la g tale che poi così g' è uguale a f ... non frega niente a nessuno.

Aldo Pratelli

Non ci rimane quindi che estendere il concetto di integrazione come "area" del sottografico, e per farlo bisogna definire cos'è un'"area" multidimensionale: la risposta ci viene fornita dalla teoria della misura.

6.1 Teoria della misura

Definizione 6.1 (σ -algebra). Sia X un insieme. Un insieme $\mathcal{A} \subseteq \mathcal{P}(X)$ si dice σ -algebra se:

- i. $X \in \mathcal{A}$;
- ii. se $E \in \mathcal{A}$, allora $E^c = X \setminus E \in \mathcal{A}$;
- iii. se $E_i \in \mathcal{A} \forall i \in \mathbb{N}$, allora $\bigcup_{i \in \mathbb{N}} E_i \in \mathcal{A}$.

Osservazione 6.1. Segue immediatamente dalla definizione che $\emptyset \in \mathcal{A}$, poiché $\emptyset = X^c$, e che, data $E_i \in \mathcal{A} \forall i \in \mathbb{N}$, allora $\bigcap_{i \in \mathbb{N}} E_i \in \mathcal{A}$, in quanto $\bigcap_{i \in \mathbb{N}} E_i = \left(\bigcup_{i \in \mathbb{N}} E_i^c \right)^c$.

Definizione 6.2 (Misura). Sia X un insieme e $\mathcal{A} \subseteq \mathcal{P}(X)$ una sua σ -algebra. Una funzione $\mu : \mathcal{A} \rightarrow [0, +\infty]$ si dice *misura* se valgono:

- i. $\mu(\emptyset) = 0$;
- ii. se $E_i \in \mathcal{A} \forall i \in \mathbb{N}$, con $E_i \cap E_j = \emptyset \forall i \neq j$, allora $\mu\left(\bigcup_{i \in \mathbb{N}} E_i\right) = \sum_{i \in \mathbb{N}} \mu(E_i)$.

L'insieme \mathcal{A} si dice *insieme dei misurabili*.

Definizione 6.3 (Spazio di misura). Uno *spazio misurabile* è una coppia (X, \mathcal{A}) , con \mathcal{A} una σ -algebra su X . Uno *spazio misurabile* o *spazio di misura* è una terna (X, \mathcal{A}, μ) , dove (X, \mathcal{A}) è uno spazio misurabile e μ è una misura su \mathcal{A} .

Esempio 6.2. La counting measure (misura che conta i punti) è una misura:

$$\mu : \mathcal{O}(X) \rightarrow [0, +\infty], \quad \mu(A) = |A| .$$

Fissato $x_0 \in X$, il delta di Dirac è una misura:

$$\delta(A) = \begin{cases} 1 & \text{se } x_0 \in A \\ 0 & \text{se } x_0 \notin A \end{cases} .$$

Lemma 6.3

Sia (X, \mathcal{A}, μ) uno spazio di misura e sia $\{E_i\}_{i \in \mathbb{N}}$ una famiglia di elementi nella σ -algebra \mathcal{A} . Allora valgono le seguenti:

- i.* se $A, B \in \mathcal{A}$ e $A \subseteq B$, allora $\mu(A) \leq \mu(B)$. Inoltre, se $\mu(A) < +\infty$, vale anche $\mu(B \setminus A) = \mu(B) - \mu(A)$;
- ii.* $\mu\left(\bigcup_{i \in \mathbb{N}} E_i\right) \leq \sum_{i \in \mathbb{N}} \mu(E_i)$;
- iii.* se $E_i \subseteq E_{i+1} \forall i \in \mathbb{N}$, allora $\mu\left(\bigcup_{i \in \mathbb{N}} E_i\right) = \lim_{i \rightarrow +\infty} \mu(E_i) = \sup\{\mu(E_i), i \in \mathbb{N}\}$;
- iv.* se $E_i \supseteq E_{i+1} \forall i \in \mathbb{N}$ e $\exists i \in \mathbb{N}$ tale che $\mu(E_i) < +\infty$, allora $\mu\left(\bigcap_{i \in \mathbb{N}} E_i\right) = \lim_{i \rightarrow +\infty} \mu(E_i) = \inf\{\mu(E_i), i \in \mathbb{N}\}$.

Dimostrazione. *i.* Basta osservare

$$\mu(B) = \mu((B \setminus A) \cup A) = \mu(B \setminus A) + \mu(A) \geq \mu(A) .$$

Se poi A ha misura finita, si ricava dalla seconda uguaglianza

$$\mu(B \setminus A) = \mu(B) - \mu(A) .$$

ii. Similmente, osserviamo che

$$\bigcup_{i \in \mathbb{N}} E_i = E_1 \cup (E_2 \setminus E_1) \cup (E_3 \setminus (E_2 \cup E_1)) \cup \dots$$

da cui, applicando *i*, segue

$$\mu\left(\bigcup_{i \in \mathbb{N}} E_i\right) = \sum_{i \in \mathbb{N}} \mu\left(E_i \setminus \bigcup_{j=1}^{i-1} E_j\right) \leq \sum_{i \in \mathbb{N}} \mu(E_i) .$$

iii. Ripetendo il ragionamento del punto precedente, abbiamo

$$\mu\left(\bigcup_{i \in \mathbb{N}} E_i\right) = \sum_{i \in \mathbb{N}} \mu\left(E_i \setminus \bigcup_{j=1}^{i-1} E_j\right) = \lim_{k \rightarrow +\infty} \sum_{i=1}^k \mu\left(E_i \setminus \bigcup_{j=1}^{i-1} E_j\right) = \lim_{i \rightarrow +\infty} \mu(E_i) .$$

iv. Supponiamo senza perdita di generalità che $\mu(E_1) < +\infty$. Poiché la successione $\{E_i\}_{i \in \mathbb{N}}$ è decrescente, la successione $\{E_1 \setminus E_i\}_{i \in \mathbb{N}}$ risulta crescente, quindi per il punto *iii*:

$$\begin{aligned} \mu(E_1) - \mu\left(\bigcap_{i \in \mathbb{N}} E_i\right) &= \mu\left(E_1 \setminus \bigcap_{i \in \mathbb{N}} E_i\right) = \mu\left(\bigcup_{i \in \mathbb{N}} E_1 \setminus E_i\right) = \lim_{i \rightarrow +\infty} \mu(E_1 \setminus E_i) = \\ &= \mu(E_1) - \lim_{i \rightarrow +\infty} \mu(E_i) \end{aligned}$$

dove l'ultima uguaglianza è lecita poiché $\mu(E_1) < +\infty$.

□

In generale, ci piacerebbe definire una misura su tutti i sottoinsiemi di X , ma ciò non è sempre possibile, come mostra il seguente risultato. Questo motiva la necessità di introdurre le σ -algebre.

Proposizione 6.4: Controesempio di Vitali

Assumendo l'assioma della scelta, non esiste una misura ℓ su $(\mathbb{R}, \mathcal{P}(\mathbb{R}))$ tale che per ogni $a, b \in \mathbb{R}$ con $a < b$ vale $\ell((a, b)) = b - a$ e che sia invariante per traslazioni, ossia tale che $\forall x \in \mathbb{R} \ell(x + A) = \ell(A)$.

Dimostrazione. Definiamo la seguente relazione di equivalenza su \mathbb{R} :

$$x \sim y \iff x - y \in \mathbb{Q} .$$

Da questo otteniamo l'insieme quoziente \mathbb{R}/\sim , che ha infiniti elementi. Ora, grazie all'assioma della scelta, possiamo prendere un insieme $V \subseteq \mathbb{R}$ che contiene un rappresentante per ogni classe di equivalenza: chiaramente vale $\mathbb{R} = \bigsqcup_{q \in \mathbb{Q}} V + q$.

Supponiamo che esista la misura ℓ : se fosse $\ell(V) = 0$, avremmo $\ell(\mathbb{R}) = 0$, e dunque per il Lemma $\ell \equiv 0$; se invece fosse $\ell(V) > 0$, potremmo supporre senza perdere generalità che $\ell(V \cap [0, 1]) > 0$, e dunque avremmo

$$2 \geq \ell \left(\bigcup_{q \in \mathbb{Q} \cap [0, 1]} V \cap [0, 1] + q \right) = +\infty ;$$

in entrambi i casi troviamo un assurdo. □

Definizione 6.4 (Algebra di Borel). Sia (X, τ) uno spazio topologico. I *boreliani*, o *algebra di Borel*, sono la più piccola σ -algebra che contiene gli aperti di X , quindi la più piccola σ -algebra che contiene la topologia τ di X , $\sigma(\tau)$. Indichiamo i boreliani di X con $\mathcal{B}(X)$.

Definizione 6.5 (Misura regolare). Sia (X, \mathcal{A}, μ) uno spazio di misura, con X spazio metrico e $\mathcal{A} = \mathcal{B}(X)$. Diciamo che μ è *regolare* se $\forall E \in \mathcal{A}$ vale

$$\mu(E) = \inf\{\mu(A) : A \supseteq E, A \text{ aperto}\} = \sup\{\mu(K) : K \subseteq E, K \text{ compatto}\} .$$

Cerchiamo ora di dare un'idea intuitiva della costruzione e delle proprietà della *misura di Lebesgue* su \mathbb{R} . I dettagli vengono lasciati a corsi successivi (Giovanni < 3).

Vale il seguente risultato:

Teorema 6.5

Esiste un'unica misura \mathcal{L} su \mathbb{R} tale che $\mathcal{L}((a, b)) = b - a$ e che \mathcal{L} è invariante per traslazione. Tale misura si dice *misura di Lebesgue*.

Dato E misurabile, indicheremo la sua misura di Lebesgue come $\mathcal{L}(E)$ o come $|E|$.

Quindi, dato $I = (a, b)$ intervallo aperto, vale $\mathcal{L}(I) = b - a$.

Dato invece A aperto generico, possiamo scrivere $A = \bigcup_{i \in \mathbb{N}} A_i$, dove ciascun A_i è un intervallo aperto o un punto, e tale unione è disgiunta¹. Poiché l'unione è numerabile, anche $\mathcal{L}(A)$ è determinata (a priori $\mathcal{L}(A)$ dipende dalla decomposizione di A trovata, tuttavia si può dimostrare che

¹Una possibile costruzione consiste nel considerare intervalli del tipo $(n, n + 1), n \in \mathbb{Z}$ che sono contenuti in A , dopodiché considerare quelli del tipo $(\frac{n}{2}, \frac{n+1}{2}), n \in \mathbb{Z}$ che sono contenuti in A ma non negli intervalli della tipologia precedente; a questo punto si itera il processo, in cui ad ogni passo viene aggiunta una quantità al più numerabile di intervalli. Al termine dell'operazione, si dimostra facilmente che mancano solamente punti del tipo $\frac{m}{2^n}$, che sono anch'essi in quantità al più numerabile.

il valore è unico) .

Infine consideriamo K compatto: questo è chiuso e limitato, quindi in particolare esiste un intervallo aperto B tale che $K \subseteq B$. A questo punto la misura di K è immediatamente determinata, poiché vale $\mathcal{L}(K) = \mathcal{L}(B) - \mathcal{L}(B \setminus K)$, e sia B che $B \setminus K$ sono aperti, quindi anche le loro misure sono determinate. Osserviamo in particolare che la misura di Lebesgue è definita su tutti i boreliani di \mathbb{R} .

A questo punto definiamo

$$\mathcal{A} = \{E \subseteq \mathbb{R} : \exists \{K_n\} \text{ compatti, } \exists \{A_n\} \text{ aperti, t.c. } K_n \subseteq E \subseteq A_n, \lim_{n \rightarrow +\infty} \mathcal{L}(K_n) = \lim_{n \rightarrow +\infty} \mathcal{L}(A_n)\}.$$

In particolare, se $E \in \mathcal{A}$ e E è limitato, vale:

$$\forall \varepsilon > 0 \exists K \text{ compatto } \exists A \text{ aperto t.c. } K \subseteq E \subseteq A, \mathcal{L}(A \setminus K) < \varepsilon.$$

Lemma 6.6

\mathcal{A} è una σ -algebra.

Dimostrazione. i. Chiaramente $\emptyset, \mathbb{R} \in \mathcal{A}$;

ii. Sia $E \in \mathcal{A}$. Se E è limitato, allora $E \subseteq I$ intervallo; abbiamo inoltre che $\forall \varepsilon > 0 \exists K$ compatto $\exists A$ aperto t.c. $K \subseteq E \subseteq A \subseteq I$, $\mathcal{L}(A \setminus K) < \varepsilon$, pertanto $I \setminus A \subseteq I \setminus E \subseteq I \setminus K$, dove $I \setminus A$ è chiuso e limitato, quindi compatto, e $I \setminus K$ è aperto, e inoltre $\mathcal{L}((I \setminus K) \setminus (I \setminus A)) = \mathcal{L}(A \setminus K) < \varepsilon$. Nel caso E illimitato possiamo ricondurci al caso precedente osservando che un insieme A è misurabile se e solo se le sue intersezioni con una quantità numerabile di intervalli sono misurabili.

iii. Siano $\{E_n\} \subseteq \mathcal{A}$. Supponiamo che

$$E = \bigcup_{n \in \mathbb{N}} E_n$$

sia limitato, e mostriamo che $E \in \mathcal{A}$. In particolare E_n è limitato $\forall n$, di conseguenza $\forall \varepsilon > 0$ esistono K_n compatto e A_n aperto tali che $K_n \subseteq E_n \subseteq A_n$ e $\mathcal{L}(A_n \setminus K_n) < \frac{\varepsilon}{2^n}$.

Ora abbiamo la catena

$$\bigcup_{i=1}^n K_i \subseteq \bigcup_{i \in \mathbb{N}} K_i \subseteq \bigcup_{i \in \mathbb{N}} E_i \subseteq \bigcup_{i \in \mathbb{N}} A_i \quad \text{e inoltre} \quad \bigcup_{i=1}^n K_i \nearrow \bigcup_{i=1}^{+\infty} K_i$$

che ha misura finita. Pertanto esiste \bar{k} tale che

$$\mathcal{L} \left(\bigcup_{i=1}^{\bar{k}} K_i \right) \geq \mathcal{L} \left(\bigcup_{i \in \mathbb{N}} K_i \right) - \varepsilon,$$

e dunque otteniamo

$$\mathcal{L} \left(\left(\bigcup_{i \in \mathbb{N}} A_n \setminus \bigcup_{i=1}^{\bar{k}} K_i \right) \setminus \left(\bigcup_{i \in \mathbb{N}} A_n \setminus \bigcup_{i \in \mathbb{N}} K_i \right) \right) \leq \varepsilon.$$

Quindi

$$\begin{aligned} \mathcal{L} \left(\bigcup_{i \in \mathbb{N}} A_n \setminus \bigcup_{i=1}^{\bar{k}} K_i \right) &\leq \mathcal{L} \left(\bigcup_{i \in \mathbb{N}} A_n \setminus \bigcup_{i \in \mathbb{N}} K_i \right) + \varepsilon \leq \sum_{n \in \mathbb{N}} \mathcal{L} \left(A_n \setminus \bigcup_{i \in \mathbb{N}} K_i \right) + \varepsilon \leq \\ &\leq \sum_{n \in \mathbb{N}} \mathcal{L}(A_n \setminus K_n) + \varepsilon \leq 2\varepsilon, \end{aligned}$$

da cui la tesi, per l'arbitrarietà di ε . Il caso illimitato è analogo al punto ii. □

Corollario 6.7

Dato $E \subseteq \mathbb{R}$, se esistono $F, G \in \mathcal{A}$ tali che $F \subseteq E \subseteq G$ e $\mathcal{L}(F) = \mathcal{L}(G) < +\infty$, allora $E \in \mathcal{A}$ e vale $\mathcal{L}(E) = \mathcal{L}(F) = \mathcal{L}(G)$.

Definizione 6.6 (Funzione misurabile). Sia (X, \mathcal{A}, μ) uno spazio di misura. Una funzione $f : X \rightarrow \mathbb{R}$ si dice *misurabile* (o *Lebesgue-misurabile*) se

$$\forall B \in \mathcal{B}(\mathbb{R}) \quad f^{-1}(B) \in \mathcal{A} .$$

Proposizione 6.8

Sia (X, \mathcal{A}, μ) uno spazio di misura. Data $f : X \rightarrow \mathbb{R}$, i seguenti fatti sono equivalenti:

- i. f è misurabile;
- ii. $f^{-1}((a, +\infty)) \in \mathcal{A} \quad \forall a \in \mathbb{R}$;
- iii. $f^{-1}([a, +\infty)) \in \mathcal{A} \quad \forall a \in \mathbb{R}$;
- iv. $f^{-1}((-\infty, a)) \in \mathcal{A} \quad \forall a \in \mathbb{R}$;
- v. $f^{-1}((-\infty, a]) \in \mathcal{A} \quad \forall a \in \mathbb{R}$.

Dimostrazione. Chiaramente la condizione *i* implica tutte le altre, inoltre si dimostra facilmente che le condizioni *ii*, *iii*, *iv* e *v* sono equivalenti, ricordando che la preimmagine commuta con unione e intersezione. D'altra parte se vale *ii*, e dunque anche *iv*, per ogni intervallo (a, b) abbiamo $f^{-1}((a, b)) = f^{-1}((-\infty, b)) \cap f^{-1}((a, +\infty)) \in \mathcal{A}$, e si conclude osservando che gli intervalli aperti generano i boreliani in quanto sono una base della topologia euclidea. \square

Proposizione 6.9

Sia (X, \mathcal{A}, μ) uno spazio di misura e siano $f, g : X \rightarrow \mathbb{R}$ funzioni misurabili, con g tale che $\mu(\{x \in X : g(x) = 0\}) = 0$. Allora anche le seguenti sono funzioni misurabili:

$$\lambda f, \quad f + g, \quad f \cdot g, \quad \frac{1}{g}, \quad \sup f, \quad \inf f .$$

Dimostrazione. Utilizzando la caratterizzazione precedente, si tratta di semplici verifiche. Ad esempio

$$\{f + g > a\} = \bigcup_{q \in \mathbb{Q}} (\{f > q\} \cap \{g > a - q\}),$$

che è unione numerabile di misurabili. \square

Definizione 6.7 (Funzioni semplici). Sia (X, \mathcal{A}, μ) uno spazio di misura. Dato $E \in \mathcal{A}$, la *funzione caratteristica*, o *funzione indicatrice*, di E è:

$$\mathbb{1}_E = \chi_E : x \mapsto \begin{cases} 1 & \text{se } x \in E \\ 0 & \text{se } x \notin E \end{cases} .$$

Dati $E_1, \dots, E_n \in \mathcal{A}$ disgiunti e $c_i > 0$ per $i = 1, \dots, n$, una *funzione semplice* o *step-funzion* (*funzione a scala*) è una funzione della forma:

$$f = \sum_{i=1}^n c_i \chi_{E_i} .$$

Definizione 6.8 (Integrale di funzioni semplici). Sia (X, \mathcal{A}, μ) uno spazio di misura e sia $f = \sum_{i=1}^n c_i \chi_{E_i}$ una funzione semplice. Il suo *integrale rispetto a μ* è

$$\int_X f d\mu = \sum_{i=1}^n c_i \mu(E_i) .$$

Definizione 6.9 (Integrale di una funzione misurabile positiva). Sia (X, \mathcal{A}, μ) uno spazio di misura. Se la funzione $f : X \rightarrow \mathbb{R}^+$ è *misurabile*, il suo *integrale rispetto a μ* è

$$\int_X f d\mu = \sup \left\{ \int_X g d\mu : 0 \leq g \leq f, g \text{ funzione a scala} \right\} .$$

Osserviamo che per la misura di Lebesgue la definizione precedente può essere riformulata come

$$\int_{\mathbb{R}} f d\mathcal{L} = \inf \left\{ \int_{\mathbb{R}} g d\mathcal{L} : g \geq f, g \text{ funzione a scala numerabile} \right\} .$$

Infatti, fissato $\varepsilon > 0$, definiamo $E_n = \{x \in X : f(x) \in (n\varepsilon, (n+1)\varepsilon]\} \in \mathcal{A}$. Per definizione, $\exists K_n \subseteq E_n \subseteq A_n$ tali che $\mathcal{L}(A_n \setminus K_n) < \frac{1}{2^n}$.

Ora poniamo

$$f_\varepsilon^- = \sum_{n \in \mathbb{N}} n\varepsilon \chi_{K_n} \quad f_\varepsilon^+ = \sum_{n \in \mathbb{N}} (n+1)\varepsilon \chi_{A_n}$$

per cui vale la disuguaglianza $f_\varepsilon^- \leq f \leq f_\varepsilon^+$. Detto allora $K = \bigcup_{n \in \mathbb{N}} K_n$, possiamo stimare

$$\begin{aligned} \int_{\mathbb{R}} (f_\varepsilon^+ - f_\varepsilon^-) d\mathcal{L} &= \int_K (f_\varepsilon^+ - f_\varepsilon^-) d\mathcal{L} + \int_{\mathbb{R} \setminus K} (f_\varepsilon^+ - f_\varepsilon^-) d\mathcal{L} \leq \varepsilon \mathcal{L}(K) + \int_{\mathbb{R} \setminus K} f_\varepsilon^+ d\mathcal{L} \\ &\leq \varepsilon \mathcal{L}(K) + \sum_{n \in \mathbb{N}} (n+1)\varepsilon \mathcal{L}(A_n \setminus K_n) \leq \varepsilon \mathcal{L}(K) + \varepsilon \sum_{n \in \mathbb{N}} \frac{n+1}{2^n} \xrightarrow{\varepsilon \rightarrow 0} 0. \end{aligned}$$

Osservazione 6.10. Osserviamo che la costruzione appena fatta per definire l'integrale di una funzione misurabile come inf di integrali di funzioni a scala numerabile è stata possibile poiché abbiamo considerato la σ -algebra \mathcal{A} definita in precedenza.

Corollario 6.11

Date $f, g : X \rightarrow \mathbb{R}$ due funzioni misurabili non negative, se $h : X \rightarrow \mathbb{R}$ è una funzione tale che $0 \leq f \leq h \leq g$ e

$$\int_X f d\mu = \int_X g d\mu ,$$

allora h è misurabile e vale

$$\int_X h d\mu = \int_X f d\mu = \int_X g d\mu .$$

Definizione 6.10 (Funzione integrabile). Sia (X, \mathcal{A}, μ) uno spazio di misura. Una funzione misurabile $f : X \rightarrow \mathbb{R}$ si dice *integrabile* se $\int_X f^+ d\mu < +\infty$ o $\int_X f^- d\mu < +\infty$. In tal caso, si pone

$$\int_X f d\mu = \int_X f^+ d\mu - \int_X f^- d\mu .$$

Osservazione 6.12. Segue dalla definizione che l'integrale è lineare, monotono e che per ogni $f : X \rightarrow \mathbb{R}$ integrabile vale

$$\left| \int_X f d\mu \right| \leq \int_X |f| d\mu .$$

6.2 Teoremi di convergenza

In questa sezione ci occupiamo di capire se e sotto quali ipotesi è vero che se $f_n \rightarrow f$ puntualmente allora l'integrale di f_n converge all'integrale di f .

Lemma 6.13

Sia (X, \mathcal{A}, μ) uno spazio di misura e sia $g : X \rightarrow \mathbb{R}$ una funzione misurabile, non negativa e con integrale finito. Allora $\forall \varepsilon > 0 \exists \delta > 0$ tale che $\forall E \in \mathcal{A}$

$$\mu(E) < \delta \implies \int_E g d\mu < \varepsilon.$$

Dimostrazione. La tesi è chiaramente vera per g limitata: difatti se $L = \sup g$, per ogni ε possiamo scegliere $\delta = \frac{\varepsilon}{L}$. Più in generale, poiché g è integrabile per ogni ε esiste una funzione semplice h con $0 \leq h \leq g$ tale che

$$\int_X (g - h) d\mu < \frac{\varepsilon}{2};$$

inoltre h è limitata e di conseguenza soddisfa la tesi, ossia $\exists \delta$ tale che per $E \in \mathcal{A}$ vale l'implicazione

$$\mu(E) < \delta \implies \int_E h d\mu < \frac{\varepsilon}{2}.$$

Pertanto

$$\int_E g d\mu = \int_E (g - h) d\mu + \int_E h d\mu \leq \int_X (g - h) d\mu + \int_E h d\mu < \varepsilon.$$

□

Teorema 6.14: Teorema della convergenza monotona (Beppo Levi)

Sia (X, \mathcal{A}, μ) uno spazio di misura. Sia $\{f_n\}$ una successione crescente di funzioni misurabili non negative, ossia tale che $0 \leq f_1 \leq \dots \leq f_n \leq \dots \leq +\infty$. Allora $f = \lim_{n \rightarrow +\infty} f_n$ è misurabile e vale

$$\lim_{n \rightarrow +\infty} \int_X f_n d\mu = \int_X f d\mu.$$

Dimostrazione 1. Vediamo una dimostrazione per $X = \mathbb{R}$ con la misura di Lebesgue. Chiaramente $\forall n f_n \leq f$, pertanto

$$\int_{\mathbb{R}} f_n d\mathcal{L} \leq \int_{\mathbb{R}} f d\mathcal{L} \quad \text{e dunque} \quad \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f_n d\mathcal{L} \leq \int_{\mathbb{R}} f d\mathcal{L};$$

dimostriamo dunque l'altra disuguaglianza.

Innanzitutto supponiamo f limitata e $\mathcal{L}(\Omega) < +\infty$, dove $\Omega = \{x \mid f(x) > 0\}$. Per ipotesi $f = \sup f_n$, dunque f è misurabile e di conseguenza $f - f_n$ è misurabile per ogni n . Inoltre poiché f è limitata, non negativa e $\mathcal{L}(\Omega)$ è finito, vale

$$\int_{\mathbb{R}} f d\mathcal{L} = \int_{\Omega} f d\mathcal{L} \leq \int_{\Omega} \sup f d\mathcal{L} = \mathcal{L}(\Omega) \sup f$$

che è a sua volta finito. Fissiamo $\varepsilon > 0$ e sia $\delta > 0$ dato dal Lemma tale che $\forall E \in \mathcal{A}$ valga

$$\mathcal{L}(E) < \delta \implies \int_E f d\mathcal{L} < \varepsilon.$$

Per ipotesi $f_n \rightarrow f$ puntualmente, dunque in particolare $\forall x \in \mathbb{R} \exists n(x) \in \mathbb{N}$ minimo tale che per $n > n(x)$ vale $f_n(x) > f(x) - \varepsilon$. Poniamo $E_n := \{x \in \mathbb{R} \mid n(x) > n\}$ e osserviamo che

$E_1 \supseteq E_2 \supseteq \dots$ e $\bigcap_n E_n = \emptyset$, ossia $E_n \searrow \emptyset$: dal momento che $E_n = \{x \mid f(x) - f_n(x) \geq \varepsilon\}$, E_n è misurabile; inoltre $E_1 \subseteq \Omega$ e dunque $\mathcal{L}(E_1) \leq \mathcal{L}(\Omega) < +\infty$: per il Lemma 6.3 troviamo $\mathcal{L}(E_n) \searrow 0$. Sia dunque $\bar{n} \in \mathbb{N}$ tale che $\mathcal{L}(E_{\bar{n}}) < \delta$; possiamo stimare

$$\begin{aligned} \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f_n d\mathcal{L} &\geq \int_{\mathbb{R}} f_{\bar{n}} d\mathcal{L} = \int_{\mathbb{R}} f d\mathcal{L} - \int_{E_{\bar{n}}} (f - f_{\bar{n}}) d\mathcal{L} - \int_{E_{\bar{n}}^c} (f - f_{\bar{n}}) d\mathcal{L} = \\ &= \int_{\mathbb{R}} f d\mathcal{L} - \int_{E_{\bar{n}}} (f - f_{\bar{n}}) d\mathcal{L} - \int_{E_{\bar{n}}^c \cap \Omega} (f - f_{\bar{n}}) d\mathcal{L} \geq \\ &\geq \int_{\mathbb{R}} f d\mathcal{L} - \int_{E_{\bar{n}}} f d\mathcal{L} - \int_{E_{\bar{n}}^c \cap \Omega} \varepsilon d\mathcal{L} \geq \int_{\mathbb{R}} f d\mathcal{L} - \varepsilon(1 + \mathcal{L}(\Omega)) , \end{aligned}$$

dunque per arbitrarietà di ε

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f_n d\mathcal{L} \geq \int_{\mathbb{R}} f d\mathcal{L} .$$

Se Ω ha misura finita ma f è non limitata, per ogni $M > 0$ possiamo considerare la troncatura $f \wedge M$ che è limitata, e chiaramente $f_n \wedge M \nearrow f \wedge M$. Adesso se f ha integrale finito, per ogni $\varepsilon > 0$ esiste h funzione a scala con $0 \leq h \leq f$ tale che

$$\int_{\mathbb{R}} h d\mathcal{L} \geq \int_{\mathbb{R}} f d\mathcal{L} - \varepsilon ,$$

e detto $M = \max h$, pertanto per quanto detto

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f_n d\mathcal{L} \geq \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f_n \wedge M d\mathcal{L} = \int_{\mathbb{R}} f \wedge M d\mathcal{L} \geq \int_{\mathbb{R}} h d\mathcal{L} \geq \int_{\mathbb{R}} f d\mathcal{L} - \varepsilon ;$$

se invece f ha integrale infinito, per ogni $M > 0$ esiste h a scala con supporto finito tale che $0 \leq h \leq f$ e l'integrale di h è almeno M , dunque

$$\lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f_n d\mathcal{L} \geq \lim_{n \rightarrow +\infty} \int_{\mathbb{R}} f_n \wedge h d\mathcal{L} = \int_{\mathbb{R}} h d\mathcal{L} \geq M .$$

Il caso $\mathcal{L}(\Omega) = +\infty$ è analogo. □

Dimostrazione 2. Abbiamo già visto che $f = \lim_n f_n = \sup_n f_n$ è misurabile. Dimostriamo la disuguaglianza non banale.

Dalla definizione di integrale segue che esiste una successione non decrescente g_n di funzioni a scala non negative tali che

$$\lim_{n \rightarrow +\infty} \int_X g_n d\mu = \int_X f d\mu .$$

Perciò basta provare che per ogni $n \in \mathbb{N}$ vale

$$\int_X g_n d\mu \leq \lim_{k \rightarrow +\infty} \int_X f_k d\mu .$$

Si vuole provare che se g è una funzione semplice e $\lim_{k \rightarrow +\infty} f_k(x) \geq g(x)$ quasi ovunque, allora:

$$\lim_{k \rightarrow +\infty} \int_X f_k d\mu \geq \int_X g d\mu .$$

Spezzando la funzione g nelle sue parti a valori costanti, questo si riduce al caso in cui g è la funzione indicatrice di un insieme. Quindi il risultato che si vuole provare è il seguente: si supponga che A sia un insieme misurabile e f_n sia una successione non decrescente di funzioni misurabili su A tali che $\lim_{n \rightarrow +\infty} f_n \geq 1$ per quasi tutti gli $x \in A$. Allora:

$$\lim_{n \rightarrow +\infty} \int_A f_n d\mu \geq \mu(A) .$$

Per provare questo risultato si fissi $\varepsilon > 0$ e si definisca la successione di insiemi misurabili:

$$B_n = \{x \in A : f_n(x) \geq 1 - \varepsilon\} .$$

Per la monotonia dell'integrale, segue che per ogni $n \in \mathbb{N}$ si ha:

$$\mu(B_n)(1 - \varepsilon) = \int_A (1 - \varepsilon)\chi_{B_n} d\mu \leq \int_A f_n d\mu .$$

Dato che per ipotesi

$$\bigcup_{i \in \mathbb{N}} B_i = A$$

a meno di un insieme di misura nulla, per l'addittività della misura μ :

$$\mu(A) = \lim_{n \rightarrow +\infty} \mu(B_n) \leq \lim_{n \rightarrow +\infty} (1 - \varepsilon)^{-1} \int_A f_n d\mu .$$

Poiché questo è vero per ogni ε positivo, segue la tesi. \square

Corollario 6.15

Sia (X, \mathcal{A}, μ) uno spazio di misura. Sia $\{f_n\}$ una successione non crescente di funzioni misurabili non negative, ossia tale che $f_1 \geq \dots \geq f_n \geq \dots$ e $f_n \geq 0 \forall n$, con $\int_X f_1 d\mu < +\infty$. Allora

$$\lim_{n \rightarrow +\infty} \int_X f_n d\mu = \int_X f d\mu .$$

Dimostrazione. Consideriamo la successione $\{f_1 - f_n\}$: questa converge in modo monotono crescente a $f_1 - f$. Ora, per il teorema di convergenza monotona e poiché $\int_X f_1 d\mu < +\infty$, abbiamo

$$\int_X f_1 d\mu - \int_X f_n d\mu = \int_X (f_1 - f_n) d\mu \xrightarrow{n \rightarrow +\infty} \int_X (f_1 - f) d\mu = \int_X f_1 d\mu - \int_X f d\mu .$$

Perciò

$$\int_X f_n d\mu \xrightarrow{n \rightarrow +\infty} \int_X f d\mu .$$

\square

Teorema 6.16: Teorema della convergenza dominata (Lebesgue)

Sia (X, \mathcal{A}, μ) uno spazio di misura e sia $\{f_n\}$ una successione di funzioni misurabili tale che $f_n \rightarrow f$ puntualmente. Se esiste una funzione integrabile g con integrale finito e tale che $|f_n| \leq g \forall n$, allora

$$\lim_{n \rightarrow +\infty} \int_X f_n d\mu = \int_X f d\mu .$$

Dimostrazione. Supponiamo che le funzioni f_n siano non negative. Definiamo per ogni n

$$h_n^- = \inf_{j \geq n} f_j \quad \text{e} \quad h_n^+ = \sup_{j \geq n} f_j$$

dunque per definizione $h_n^- \nearrow \liminf f_n = f$ mentre $h_n^+ \searrow \limsup f_n = f$; inoltre $h_0^+ = \sup_n f_n$ è integrabile in quanto per ipotesi g è un maggiorante della famiglia $\{f_n\}$ per cui

$$h_0^+ \leq g \implies \int_X h_0^+ d\mu \leq \int_X g d\mu < +\infty .$$

Per il Teorema di convergenza monotona ed il successivo Corollario deduciamo che

$$\int_X h_n^- d\mu \xrightarrow{n \rightarrow +\infty} \int_X f d\mu \quad \text{e} \quad \int_X h_n^+ d\mu \xrightarrow{n \rightarrow +\infty} \int_X f d\mu$$

e poiché $h_n^- \leq f_n \leq h_n^+$

$$\int_X f_n d\mu \xrightarrow{n \rightarrow +\infty} \int_X f d\mu .$$

Siano adesso f_n non necessariamente non negative e siano f_n^- e f_n^+ rispettivamente la parte positiva e negativa di f_n , analogamente definiamo f^- e f^+ . A questo punto osserviamo che $f_n^+ \rightarrow f^+$ e $f_n^- \rightarrow f^-$ e che $0 \leq f_n^-, f_n^+ \leq g$, pertanto per quanto provato

$$\int_X f_n^- d\mu \xrightarrow{n \rightarrow +\infty} \int_X f^- d\mu \quad \text{e} \quad \int_X f_n^+ d\mu \xrightarrow{n \rightarrow +\infty} \int_X f^+ d\mu$$

e dunque

$$\int_X f_n d\mu \xrightarrow{n \rightarrow +\infty} \int_X f d\mu .$$

□

Teorema 6.17: Lemma di Fatou

Sia (X, \mathcal{A}, μ) uno spazio di misura. Sia $\{f_n\}$ una successione di funzioni misurabili non negative. Allora

$$\int_X \liminf_{n \rightarrow +\infty} f_n d\mu \leq \liminf_{n \rightarrow +\infty} \int_X f_n d\mu .$$

Dimostrazione. Per ogni n poniamo $g_n = \inf_{j \geq n} f_j$, dunque dalla definizione di limite inferiore $g_n \nearrow \liminf f_n$ e chiaramente g_n è non negativa in quanto ogni f_n lo è: per il Teorema di convergenza monotona troviamo che

$$\int_X \liminf_{n \rightarrow +\infty} f_n d\mu = \lim_{n \rightarrow +\infty} \int_X g_n d\mu = \liminf_{n \rightarrow +\infty} \int_X g_n d\mu ;$$

d'altro canto per ogni n $g_n \leq f_n$, da cui

$$\int_X g_n d\mu \leq \int_X f_n d\mu \implies \liminf_{n \rightarrow +\infty} \int_X g_n d\mu \leq \liminf_{n \rightarrow +\infty} \int_X f_n d\mu .$$

□

Corollario 6.18: Lemma di Fatou inverso

Sia (X, \mathcal{A}, μ) uno spazio di misura. Sia $\{f_n\}$ una successione di funzioni misurabili e sia g una funzione integrabile con integrale finito e tale che $f_n \leq g$ per ogni n ; allora

$$\int_X \limsup_{n \rightarrow +\infty} f_n d\mu \geq \limsup_{n \rightarrow +\infty} \int_X f_n d\mu .$$

Dimostrazione. Consideriamo la successione ausiliaria $h_n = g - f_n$: chiaramente h_n è non negativa e misurabile, pertanto per il Lemma di Fatou

$$\int_X \liminf_{n \rightarrow +\infty} h_n d\mu \leq \liminf_{n \rightarrow +\infty} \int_X h_n d\mu ;$$

adesso osserviamo le seguenti identità

$$\int_X \liminf_{n \rightarrow +\infty} h_n d\mu = \int_X \liminf_{n \rightarrow +\infty} (g - f_n) d\mu = \int_X \left(g - \limsup_{n \rightarrow +\infty} f_n \right) d\mu = \int_X g d\mu - \int_X \limsup_{n \rightarrow +\infty} f_n d\mu$$

e analogamente

$$\liminf_{n \rightarrow +\infty} \int_X h_n d\mu = \liminf_{n \rightarrow +\infty} \int_X (g - f_n) d\mu = \int_X g d\mu + \liminf_{n \rightarrow +\infty} \left(- \int_X f_n d\mu \right) = \int_X g d\mu - \limsup_{n \rightarrow +\infty} \int_X f_n d\mu$$

pertanto

$$\begin{aligned} \int_X g d\mu - \int_X \limsup_{n \rightarrow +\infty} f_n d\mu &\leq \int_X g d\mu - \limsup_{n \rightarrow +\infty} \int_X f_n d\mu \\ \implies \int_X \limsup_{n \rightarrow +\infty} f_n d\mu &\geq \limsup_{n \rightarrow +\infty} \int_X f_n d\mu . \end{aligned}$$

□

Osservazione 6.19. Dati (X, μ) spazio di misura e $\mathcal{I} = (0, \bar{\alpha}) \subseteq \mathbb{R}$, sia $\{f_\alpha\}_{\alpha \in \mathcal{I}} \subseteq \mathbb{R}^X$ una famiglia di funzioni integrabili; supponiamo che esista un limite puntuale f , ossia

$$f(x) = \lim_{\alpha \rightarrow \bar{\alpha}} f_\alpha(x) \quad \forall x \in X.$$

Allora valgono risultati analoghi ai teoremi di convergenza anche per parametri continui. Ad esempio vediamo l'analogo del Teorema della convergenza monotona: supponiamo che $f_\alpha \geq 0$ per ogni α e che per $\alpha < \beta$ si abbia $f_\alpha \leq f_\beta$, ma per assurdo

$$\lim_{\alpha \rightarrow \bar{\alpha}} \int_X f_\alpha d\mu \neq \int_X f d\mu ,$$

allora esiste $\varepsilon_0 > 0$ tale che per ogni $\delta > 0$ esiste $\alpha \in (\bar{\alpha} - \delta, \bar{\alpha})$ tale che

$$\int_X f_\alpha d\mu \leq \int_X f d\mu - \varepsilon ;$$

in particolare scegliendo una successione $\delta_n \searrow 0$, troviamo una successione $\alpha_n \rightarrow \bar{\alpha}$ tale che per ogni n vale

$$\int_X f_{\alpha_n} d\mu \leq \int_X f d\mu - \varepsilon ;$$

allora a meno di sottosuccessioni possiamo supporre $\alpha_n \nearrow \bar{\alpha}$ in modo crescente, ossia la successione delle $f_n := f_{\alpha_n}$, è monotona e converge a f , però

$$\lim_{n \rightarrow +\infty} \int_X f_n d\mu \neq \int_X f d\mu . \quad \nexists$$

6.3 Misura prodotto e Teorema di Fubini

A volte sbaglia uno studente di ingegneria, e pace. Ma se succede a matematica viene voglia di uccidere gli studenti e si finisce in galera.

Aldo Pratelli

Consideriamo due spazi di misura (X_1, μ_1) e (X_2, μ_2) : vogliamo ora definire una misura sullo spazio $X_1 \times X_2$ che “rispetti” le singole misure μ_1 e μ_2 . Chiameremo questa misura la *misura prodotto* $\mu = \mu_1 \times \mu_2$: illustriamo qui l'idea e le caratteristiche principali della sua costruzione.

Dato E_1 misurabile in X_1 e E_2 misurabile in X_2 , definiamo $\mu(E_1 \times E_2) = \mu_1(E_1)\mu_2(E_2)$. Se abbiamo un'unione disgiunta e al più numerabile di rettangoli, definiamo

$$\mu \left(\bigcup_{n \in \mathbb{N}} R_n \right) = \sum_{n \in \mathbb{N}} \mu(R_n) .$$

Possiamo misurare allora anche unioni numerabili generiche di rettangoli, poiché togliere un'unione finita di rettangoli da un rettangolo dà un'unione al più numerabile di rettangoli. Chiameremo questi insiemi *insiemi elementari*.

Se $A \subseteq X$ è un aperto, è anche unione numerabile di rettangoli aperti, quindi è elementare.

Se $K \subseteq X$ è compatto, allora esistono due palle aperte $B_1 \subseteq X_1, B_2 \subseteq X_2$ tali che $K \subseteq B_1 \times B_2$. Allora possiamo scrivere $K = B_1 \times B_2 \setminus (B_1 \times B_2 \setminus K)$, dove $B_1 \times B_2 \setminus K$ è aperto, quindi risulta $\mu(K) = \mu(B_1 \times B_2) - \mu(B_1 \times B_2 \setminus K)$.

Nel seguito approfondiremo queste costruzioni (e di conseguenza anche il Teorema di Fubini) solo nell'ipotesi $X = \mathbb{R}^n$ con la σ -algebra \mathcal{A} definita in precedenza.

Ora, noi sappiamo calcolare l'integrale di funzioni a scala rettangolari. Ci chiediamo se, data f misurabile e non negativa, valga ancora

$$\begin{aligned} \int_X f d\mu &= \sup \left\{ \int_X g d\mu : 0 \leq g \leq f, g \text{ funzione a scala} \right\} = \\ &= \inf \left\{ \int_X g d\mu : g \geq f, g \text{ funzione a scala numerabile} \right\}. \end{aligned}$$

Sia quindi $f \geq 0$ misurabile su X e supponiamo $\{x \in X : f(x) > 0\} \subseteq B_1 \times B_2$. Sia $\varepsilon > 0$ e siano $E_n = \{x \in X : f(x) \in (n\varepsilon, (n+1)\varepsilon]\}$. Ora esistono $K_n \subseteq E_n \subseteq A_n$, con K_n compatto e A_n aperto, tali che $\mu(A_n \setminus K_n) < \frac{1}{2^n}$. Definiamo la funzione a scala rettangolare

$$f_\varepsilon^+ = \sum_{n \in \mathbb{N}} (n+1)\varepsilon \chi_{A_n}$$

tale che $f_\varepsilon^+ \geq f$. Ora vale la stima

$$\int_X (f_\varepsilon^+ - f) d\mu \leq \varepsilon \mu(B_1 \times B_2) + \sum_{n \in \mathbb{N}} (n+1)\varepsilon \frac{1}{2^n} \leq c\varepsilon.$$

Quindi le f_ε^+ sono funzioni a scala tali che $\int_X f_\varepsilon^+ d\mu \xrightarrow{\varepsilon \rightarrow 0} \int_X f d\mu$.

Sia ora $f_M = f \wedge M$: definisco $\varphi = M - f_M$, per cui vale $0 \leq \varphi \leq M$. Per quanto appena visto, esiste una funzione a scala rettangolare g_ε tale che $\int g_\varepsilon \leq \int \varphi + \varepsilon$; posso inoltre scegliere $g_\varepsilon \leq M$. Allora vale $0 \leq M - g_\varepsilon \leq M - \varphi \leq f_M$, e $M - g_\varepsilon$ è a scala rettangolare; inoltre abbiamo che $f_M \nearrow f$ per $M \rightarrow +\infty$, quindi per M abbastanza grande abbiamo che $\int f_M \geq \int f - \varepsilon$. Abbiamo quindi ottenuto (con l'implicita identificazione tra M e $M\chi_{B_1 \times B_2}$)

$$\begin{aligned} \int_X M - g_\varepsilon d\mu &= M\mu(B_1 \times B_2) - \int_X g_\varepsilon d\mu \geq M\mu(B_1 \times B_2) - \int_X \varphi d\mu - \varepsilon = \\ &= \int_X M - \varphi d\mu - \varepsilon = \int_X f_M d\mu - \varepsilon \geq \int_X f d\mu - 2\varepsilon. \end{aligned}$$

Quindi le funzioni $M - g_\varepsilon$ sono funzioni a scala rettangolare tali che $\int_X M - g_\varepsilon d\mu \xrightarrow{\varepsilon \rightarrow 0} \int_X f d\mu$.

Lemma 6.20

Sia f una funzione misurabile non negativa tale che $\int_X f d\mu = 0$. Allora $\mu(\{f > 0\}) = 0$.

Dimostrazione. Supponiamo per assurdo che $\mu(\{f > 0\}) > 0$.

Consideriamo la successione $E_n = \{f > \frac{1}{n}\}$ al variare di $n \in \mathbb{N}$. Osserviamo che $E_n \subseteq E_{n+1}$ e che $\bigcup_{n=1}^{\infty} E_n = \{f > 0\}$. Segue allora che

$$\lim_{n \rightarrow +\infty} \mu(E_n) = \mu(\{f > 0\}) > 0,$$

dunque esiste $n \in \mathbb{N}$ tale che $\mu(E_n) > 0$. Ora valgono le seguenti disuguaglianze:

$$0 < \frac{1}{n} \mu(E_n) = \int_{\{f > \frac{1}{n}\}} \frac{1}{n} d\mu \leq \int_{\{f > \frac{1}{n}\}} f d\mu \leq \int_{\{f > 0\}} f d\mu = 0,$$

che è un assurdo. \nexists

□

Teorema 6.21: Teorema di Fubini

Sia f non negativa e misurabile su $(X, \mu) = (X_1, \mu_1) \times (X_2, \mu_2)$, con integrale finito. Allora

$$\iint_X f d(\mu_1 \times \mu_2) = \int_{X_1} \left(\int_{X_2} f d\mu_2 \right) d\mu_1 = \int_{X_2} \left(\int_{X_1} f d\mu_1 \right) d\mu_2 .$$

Dimostrazione. La tesi è vera se la funzione è del tipo $c\chi_{E \times F}$ o se è a scala rettangolare.

Data f non negativa generica, per definizione di integrale, esiste una successione $g_1 \leq g_2 \leq \dots$, con $g_n \leq f \forall n$, a scala rettangolare tale che

$$\iint_X g_n d\mu \xrightarrow{n \rightarrow +\infty} \iint_X f d\mu .$$

Inoltre la successione g_n converge in modo monotono crescente a una funzione misurabile g tale che $g \leq f$. Ora abbiamo

$$\iint_X f d\mu = \lim_{n \rightarrow +\infty} \iint_X g_n d\mu = \lim_{n \rightarrow +\infty} \int_{X_1} \left(\int_{X_2} g_n(x, y) d\mu_2(y) \right) d\mu_1(x) .$$

Poiché g_n convergono crescendo a g , per il Teorema di convergenza monotona abbiamo che per ogni $x \in X_1$

$$\int_{X_2} g_n(x, \cdot) d\mu_2 \xrightarrow{n \rightarrow +\infty} \int_{X_2} g(x, \cdot) d\mu_2$$

in maniera crescente. Quindi, ancora per convergenza monotona, otteniamo

$$\begin{aligned} \lim_{n \rightarrow +\infty} \int_{X_1} \left(\int_{X_2} g_n(x, y) d\mu_2(y) \right) d\mu_1(x) &= \int_{X_1} \left(\lim_{n \rightarrow +\infty} \int_{X_2} g_n(x, y) d\mu_2(y) \right) d\mu_1(x) = \\ &= \int_{X_1} \left(\int_{X_2} g(x, y) d\mu_2(y) \right) d\mu_1(x) . \end{aligned}$$

Allo stesso modo, esiste una successione $h_1 \geq h_2 \geq \dots$, con $h_n \geq f \forall n$, a scala rettangolare tale che

$$\iint_X h_n d\mu \xrightarrow{n \rightarrow +\infty} \iint_X f d\mu .$$

Come prima, la successione h_n converge in modo monotono decrescente a una funzione misurabile h tale che $h \geq f$. Poiché l'integrale di f è finito, l'integrale delle h_n è definitivamente finito, quindi possiamo applicare il Corollario del Teorema di convergenza monotona per ottenere

$$\begin{aligned} \lim_{n \rightarrow +\infty} \int_{X_1} \left(\int_{X_2} h_n(x, y) d\mu_2(y) \right) d\mu_1(x) &= \int_{X_1} \left(\lim_{n \rightarrow +\infty} \int_{X_2} h_n(x, y) d\mu_2(y) \right) d\mu_1(x) = \\ &= \int_{X_1} \left(\int_{X_2} h(x, y) d\mu_2(y) \right) d\mu_1(x) . \end{aligned}$$

Quindi, in conclusione, otteniamo

$$\iint_X f d\mu = \int_{X_1} \left(\int_{X_2} g(x, y) d\mu_2(y) \right) d\mu_1(x) = \int_{X_1} \left(\int_{X_2} h(x, y) d\mu_2(y) \right) d\mu_1(x) .$$

Ora vale che $0 \leq g \leq f \leq h$, quindi per il Lemma precedente

$$\mu_1 \left(\left\{ x \in X_1 : \int_{X_2} g(x, y) d\mu_2(y) \neq \int_{X_2} h(x, y) d\mu_2(y) \right\} \right) = 0 .$$

Se $x \in X_1$ è tale che

$$\int_{X_2} g(x, y) d\mu_2(y) = \int_{X_2} h(x, y) d\mu_2(y) ,$$

allora, poiché $g(x, \cdot) \leq f(x, \cdot) \leq h(x, \cdot)$, $f(x, \cdot)$ è misurabile per μ_2 e vale

$$\int_{X_2} f(x, \cdot) d\mu_2 = \int_{X_2} g(x, \cdot) d\mu_2 = \int_{X_2} h(x, \cdot) d\mu_2 .$$

Per ottenere l'altro ordine di integrazione, il procedimento è analogo. \square

6.4 Teorema del cambio di variabile

Nel seguito, dati $x \in \mathbb{R}^n$ e $r > 0$, indicheremo con $Q_r(x)$ la palla chiusa di raggio r rispetto alla norma infinito $\bar{B}_\infty(x, r) = \{z \in \mathbb{R}^n \mid \|z - x\|_\infty \leq r\}$.

Lemma 6.22

Sia $R = [a_1, b_1] \times \cdots \times [a_n, b_n]$ un n -parallelepipedo e sia $L \in \mathcal{L}(\mathbb{R}^n)$ un isomorfismo lineare dato da $L(x) = Ax$, con $A \in \text{GL}_n(\mathbb{R})$. Allora

$$|L(R)| = |R| |\det A|$$

Dimostrazione. Dal Teorema di decomposizione polare sappiamo che esistono due matrici $P \in O_n$ e $S \in S_n(\mathbb{R})$ definita positiva tali che $A = PS$; inoltre per il Teorema spettrale S è ortogonalmente diagonalizzabile, dunque complessivamente possiamo scrivere $A = P_1 D P_2$ con D diagonale e $P_1, P_2 \in O_n$. Adesso le trasformazioni ortogonali possono essere decomposte in un numero finito di riflessioni, e dunque non cambiano la misura di Lebesgue; d'altra parte la tesi vale certamente per le matrici diagonali, di conseguenza, se $L_D \in \mathcal{L}(\mathbb{R}^n)$ è la trasformazione data da $L_D(x) = Dx$, vale

$$|L(R)| = |L_D(R)| = |\det D| |R| = |\det A| |R| ,$$

dove abbiamo usato che le trasformazioni ortogonali hanno determinante ± 1 . \square

Lemma 6.23

Dato $\Omega \subseteq \mathbb{R}^n$ un aperto, sia $\Phi : \Omega \rightarrow \mathbb{R}^n$ un diffeomorfismo di classe C^1 con l'immagine; allora per ogni compatto $K \subseteq \Omega$ e per ogni $\varepsilon > 0$ esiste $r_0 > 0$ tale che per ogni $r \in (0, r_0)$ e per ogni $x_0 \in K$ si ha $\Phi(Q_r(x_0)) \subseteq T_{x_0}(Q_{(1+\varepsilon)r}(x_0))$, dove $T_{x_0}(x) = \Phi(x_0) + D\Phi(x_0)(x - x_0)$ è il polinomio di Taylor al primo ordine di Φ centrato in x_0 .

Dimostrazione. Osserviamo preliminarmente, che poiché Φ è C^1 , la mappa $D\Phi : \Omega \rightarrow \mathcal{L}(\mathbb{R}^n)$ data da $x \mapsto D\Phi(x)$ è continua (dove $\mathcal{L}(\mathbb{R}^n) \cong \mathbb{R}^{n^2}$ è topologizzato con una qualsiasi norma), e analogamente $D(\Phi^{-1})$ è continua. Prendiamo allora

$$c := \max_{y \in \Phi(K)} \|D(\Phi^{-1})(y)\|_\infty$$

che esiste per compattezza di $\Phi(K)$, dove $\|D(\Phi^{-1})(y)\|_\infty$ denota la norma ∞ matriciale. Fissato $\varepsilon > 0$ sufficientemente piccolo, consideriamo $M = \{x \in \Omega \mid d_\infty(x, K) \leq \varepsilon\}$: M è chiuso e limitato, dunque è compatto; sia $r_0 > 0$ tale che per ogni $x_1, x_2 \in M$ con $\|x_2 - x_1\|_\infty \leq r_0$ si abbia $\|D\Phi(x_2) - D\Phi(x_1)\|_\infty \leq \varepsilon/c$; senza perdere generalità possiamo supporre $r_0 \leq \varepsilon$. A questo punto, scelto $r < r_0$, sia $y \in \Phi(Q_r(x_0))$, e siano $x \in Q_r(x_0)$ e $x' \in \Omega$ tali che $y = \Phi(x) = T_{x_0}(x')$: la tesi equivale a mostrare che $x' \in Q_{(1+\varepsilon)r}(x_0)$. Stimiamo allora

$$\begin{aligned} \|x' - x\|_\infty &= \|(D(\Phi^{-1})(\Phi(x_0)))(D\Phi(x_0)(x' - x))\|_\infty \leq \|D(\Phi^{-1})(\Phi(x_0))\|_\infty \|D\Phi(x_0)(x' - x)\|_\infty \leq \\ &\leq c \|D\Phi(x_0)(x' - x_0) - D\Phi(x_0)(x - x_0)\|_\infty = \\ &= c \|\Phi(x_0) + D\Phi(x_0)(x' - x_0) - \Phi(x_0) - D\Phi(x_0)(x - x_0)\|_\infty = \\ &= c \|T_{x_0}(x') - \Phi(x_0) - D\Phi(x_0)(x - x_0)\|_\infty = c \|\Phi(x) - \Phi(x_0) - D\Phi(x_0)(x - x_0)\|_\infty = \\ &= c \|D\Phi(\xi_x)(x - x_0) - D\Phi(x_0)(x - x_0)\|_\infty = c \|(D\Phi(\xi_x) - D\Phi(x_0))(x - x_0)\|_\infty = \\ &= c \|D\Phi(\xi_x) - D\Phi(x_0)\|_\infty \|x - x_0\|_\infty \leq \varepsilon \|x - x_0\|_\infty \end{aligned}$$

e dunque $\|x' - x_0\|_\infty \leq \|x - x_0\|_\infty + \|x' - x\|_\infty \leq (1 + \varepsilon)\|x - x_0\|_\infty \leq (1 + \varepsilon)r$. \square

Teorema 6.24: Teorema del cambio di variabile

Dato $\Omega \subseteq \mathbb{R}^n$ un aperto, sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione integrabile con integrale finito e sia $\Phi : \Omega \rightarrow \mathbb{R}^n$ una mappa di classe C^1 , iniettiva e tale che ogni punto è regolare. Allora

$$\int_{\Phi(\Omega)} f(x) dx = \int_{\Omega} (f \circ \Phi)(y) J_{\Phi}(y) dy ,$$

dove $J_{\Phi}(y) = |\det D\Phi(y)|$ è il modulo dello *jacobiano*.

Dimostrazione. Osserviamo che basta dimostrare la disuguaglianza

$$\int_{\Phi(\Omega)} f(x) dx \leq \int_{\Omega} (f \circ \Phi)(y) J_{\Phi}(y) dy$$

per ogni f non negativa; infatti considerando $\Phi^{-1} : \Phi(\Omega) \rightarrow \Omega$ troviamo

$$\begin{aligned} \int_{\Phi(\Omega)} f(x) dx &\leq \int_{\Omega} (f \circ \Phi)(y) J_{\Phi}(y) dy \leq \int_{\Phi(\Omega)} (f \circ \Phi \circ \Phi^{-1})(x) J_{\Phi}(\Phi^{-1}(x)) J_{\Phi^{-1}}(x) dx \leq \\ &\leq \int_{\Phi(\Omega)} f(x) |\det(D\Phi(\Phi^{-1}(x)) \circ D\Phi^{-1}(x))| dx = \int_{\Phi(\Omega)} f(x) dx , \end{aligned}$$

e a quel punto se f è generica basta osservare che l'uguaglianza vale per f^+ e f^- .

Mostriamo a questo punto la disuguaglianza rilassando progressivamente le ipotesi.

Supponiamo che $E \subseteq \Omega$ sia un n -cubo (o unione finita di n -cubi) con lati paralleli agli assi e che $f = \chi_{\Phi(E)}$; sia $\varepsilon > 0$ e sia r_0 dato dal Lemma precedente per ε su E , allora per un certo $r < r_0$ possiamo prendere finiti punti x_1, \dots, x_k tali che $E = Q_r(x_1) \cup \dots \cup Q_r(x_k)$ e l'unione è disgiunta a meno di trascurabili; possiamo inoltre supporre, a meno di ridurre r , che se $\|x - z\|_{\infty} < r$ allora $\|J_{\Phi}(x) - J_{\Phi}(z)\|_{\infty} < \varepsilon$. A questo punto, denotando $r_{\varepsilon} = (1 + \varepsilon)r$, possiamo stimare

$$\begin{aligned} |\Phi(E)| &= \sum_{i=1}^k |\Phi(Q_r(x_i))| \leq \sum_{i=1}^k |T_{x_i}(Q_{r_{\varepsilon}}(x_i))| \stackrel{*}{=} \sum_{i=1}^k \int_{Q_{r_{\varepsilon}}(x_i)} J_{T_{x_i}}(x) dx = \\ &= (1 + \varepsilon)^n \sum_{i=1}^k \int_{Q_r(x_i)} J_{\Phi}(x_i) dx = \\ &= (1 + \varepsilon)^n \sum_{i=1}^k \int_{Q_r(x_i)} J_{\Phi}(x) dx + (1 + \varepsilon)^n \sum_{i=1}^k \int_{Q_r(x_i)} (J_{\Phi}(x_i) - J_{\Phi}(x)) dx \leq \\ &\leq (1 + \varepsilon)^n \int_E J_{\Phi}(x) dx + (1 + \varepsilon)^n \varepsilon |E| \quad \xrightarrow{\varepsilon \rightarrow 0} \int_E J_{\Phi}(x) dx \end{aligned}$$

dove l'uguaglianza (*) vale per il Lemma sul caso lineare.

Se invece $E \subseteq \Omega$ è aperto, esistono numerabili n -cubi $\{Q_m\}$ tali che $E = \bigcup_m Q_m$ ed i Q_m sono disgiunti a meno di trascurabili; allora

$$|\Phi(E)| = \sum_{m \in \mathbb{N}} |\Phi(Q_m)| \leq \sum_{m \in \mathbb{N}} \int_{Q_m} J_{\Phi}(x) dx = \int_{\bigcup_m Q_m} J_{\Phi}(x) dx = \int_E J_{\Phi}(x) dx$$

dove la penultima uguaglianza segue per convergenza monotona.

Se poi $E \subseteq \Omega$ è misurabile, possiamo supporre E limitato a meno di intersecarlo con dei compatti e stimare singolarmente le intersezioni: allora per ogni $\varepsilon > 0$ esiste un aperto limitato A tale che $E \subseteq A$ ma $|A| \leq |E| + \varepsilon$, e di conseguenza

$$|\Phi(E)| \leq |\Phi(A)| \leq \int_A J_{\Phi}(x) dx = \int_E J_{\Phi}(x) dx + \int_{A \setminus E} J_{\Phi}(x) dx \leq \int_E J_{\Phi}(x) dx + \varepsilon \|J_{\Phi}\|_{\infty}$$

e si conclude per arbitrarietà di ε .

Abbiamo allora mostrato che per ogni misurabile $E \subseteq \Omega$ vale

$$\int_{\Phi(\Omega)} \chi_{\Phi(E)}(x) dx = |\Phi(E)| \leq \int_E J_{\Phi}(y) dy = \int_{\Omega} (\chi_{\Phi(E)} \circ \Phi)(y) J_{\Phi}(y) dy$$

ossia la disuguaglianza vale se f è un'indicatrice; per addittività allora vale anche se f è semplice. Per concludere, se f è non negativa e misurabile con integrale finito, esiste una successione h_n crescente di funzioni semplici tali che $0 \leq h_n \leq f$ e tali che

$$\lim_{n \rightarrow +\infty} \int_{\Phi(\Omega)} h_n(x) dx = \int_{\Phi(\Omega)} f(x) dx ,$$

pertanto detto $h = \sup_n h_n$ si ha

$$\begin{aligned} \int_{\Phi(\Omega)} f(x) dx &= \lim_{n \rightarrow +\infty} \int_{\Phi(\Omega)} h_n(x) dx \leq \lim_{n \rightarrow +\infty} \int_{\Omega} (h_n \circ \Phi)(y) J_{\Phi}(y) dy = \\ &= \int_{\Omega} (h \circ \Phi)(y) J_{\Phi}(y) dy \leq \int_{\Omega} (f \circ \Phi)(y) J_{\Phi}(y) dy ; \end{aligned}$$

se f è invece integrabile con integrale finito, ci si riconduce alle parti positiva e negativa f^+ e f^- , per cui abbiamo già la tesi. \square

6.4.1 Coordinate polari, cilindriche e sferiche

Nel calcolo di integrali è spesso necessario fare un cambio di variabile ad hoc; tuttavia alcuni di loro si ripresentano con una certa frequenza.

Il primo consiste nel cambio in **coordinate polari** in \mathbb{R}^2 .

Consideriamo $\Phi : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}^+ \times [0, 2\pi)$ e la sua inversa $\Psi : \mathbb{R}^+ \times [0, 2\pi) \rightarrow \mathbb{R}^2 \setminus \{0\}$ date da

$$(x, y) \xrightarrow{\Phi} (\sqrt{x^2 + y^2}, \arg(x, y)) \quad \text{e} \quad (\rho, \theta) \xrightarrow{\Psi} (\rho \cos \theta, \rho \sin \theta) ,$$

$$\text{allora } \det D\Psi = \det \begin{pmatrix} \cos \theta & \sin \theta \\ -\rho \sin \theta & \rho \cos \theta \end{pmatrix} = \rho .$$

Pertanto, dati $\Omega \subseteq \mathbb{R}^2 \setminus \{0\}$ e $f : \Omega \rightarrow \mathbb{R}$, vale

$$\iint_{\Omega} f(x, y) dx dy = \iint_{\Phi(\Omega)} f(\Psi(\rho, \theta)) \rho d\rho d\theta .$$

Esempio 6.25. Un'applicazione classica è utilizzare le coordinate polari per calcolare l'integrale

$$\int_{\mathbb{R}} e^{-x^2} dx = \sqrt{\pi} ,$$

infatti attraverso il Teorema di Fubini ed un cambio di variabile possiamo calcolare agevolmente il suo quadrato

$$\begin{aligned} \left(\int_{\mathbb{R}} e^{-x^2} dx \right)^2 &= \left(\int_{\mathbb{R}} e^{-x^2} dx \right) \left(\int_{\mathbb{R}} e^{-y^2} dy \right) = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)} dx dy = \\ &= \int_0^{2\pi} \int_0^{+\infty} e^{-\rho^2} \rho d\rho d\theta = 2\pi \int_0^{+\infty} e^{-\rho^2} \rho d\rho = -\pi \left[e^{-\rho^2} \right]_0^{+\infty} = \pi . \end{aligned}$$

A volte la via più breve passa per posti non sospettati.

Carlo Carminati

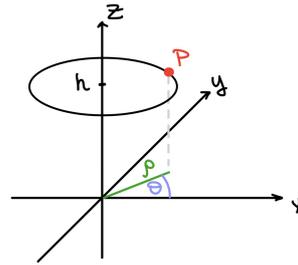
Il secondo consiste nel cambio in **coordinate cilindriche** in \mathbb{R}^3 .

Consideriamo $\Phi : \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}^+ \times [0, 2\pi) \times \mathbb{R}$ e la sua inversa $\Psi : \mathbb{R}^+ \times [0, 2\pi) \times \mathbb{R} \rightarrow \mathbb{R}^3 \setminus \{0\}$ date da

$$(x, y, z) \stackrel{\Phi}{\mapsto} (\rho, \theta, z) \quad \text{e} \quad (\rho, \theta, z) \stackrel{\Psi}{\mapsto} (x, y, z) ,$$

secondo le seguenti relazioni

$$\begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta \\ z = z \end{cases} \quad \begin{cases} \rho = \sqrt{x^2 + y^2} \\ \theta = \arctan(\frac{y}{x}) \\ z = z \end{cases}$$



Risulta così

$$D\Psi = \begin{pmatrix} \cos \theta & \sin \theta & 0 \\ -\rho \sin \theta & \rho \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix} \quad \text{e} \quad \det D\Psi = \rho .$$

Dunque, dati $\Omega \subseteq \mathbb{R}^3 \setminus \{0\}$ e $f : \Omega \rightarrow \mathbb{R}$, vale

$$\iiint_{\Omega} f(x, y, z) dx dy dz = \iiint_{\Phi(\Omega)} f(\Psi(\rho, \theta, z)) \rho d\rho d\theta dz .$$

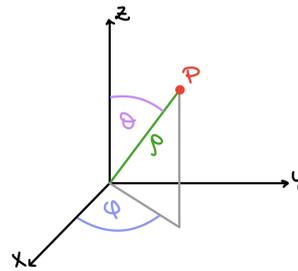
Il terzo consiste nel cambio in **coordinate sferiche** in \mathbb{R}^3 .

Consideriamo $\Phi : \mathbb{R}^3 \setminus \{0\} \rightarrow \mathbb{R}^+ \times [0, \pi) \times [0, 2\pi)$ e la sua inversa $\Psi : \mathbb{R}^+ \times [0, \pi) \times [0, 2\pi) \rightarrow \mathbb{R}^3 \setminus \{0\}$ date da

$$(x, y, z) \stackrel{\Phi}{\mapsto} (\rho, \theta, \varphi) \quad \text{e} \quad (\rho, \theta, \varphi) \stackrel{\Psi}{\mapsto} (x, y, z) ,$$

secondo le seguenti relazioni

$$\begin{cases} x = \rho \sin \theta \cos \varphi \\ y = \rho \sin \theta \sin \varphi \\ z = \rho \cos \theta \end{cases} \quad \begin{cases} \rho = \sqrt{x^2 + y^2 + z^2} \\ \theta = \arccos(\frac{z}{\rho}) \\ \varphi = \arctan(\frac{y}{x}) \end{cases}$$



Risulta così

$$D\Psi = \begin{pmatrix} \sin \theta \cos \varphi & \sin \theta \sin \varphi & \cos \theta \\ \rho \cos \theta \cos \varphi & \rho \cos \theta \sin \varphi & -\rho \sin \theta \\ -\rho \sin \theta \sin \varphi & \rho \sin \theta \cos \varphi & 0 \end{pmatrix} \quad \text{e} \quad \det D\Psi = \rho^2 \sin \theta .$$

Quindi, dati $\Omega \subseteq \mathbb{R}^3 \setminus \{0\}$ e $f : \Omega \rightarrow \mathbb{R}$, vale

$$\iiint_{\Omega} f(x, y, z) dx dy dz = \iiint_{\Phi(\Omega)} f(\Psi(\rho, \theta, \varphi)) \rho^2 \sin \theta d\rho d\theta d\varphi .$$

Esercizio 8. Data una matrice $A \in M_n(\mathbb{R})$ simmetrica e definita positiva, si provi che

$$\int_{\mathbb{R}^n} e^{-(Ax \cdot x)} dx = \sqrt{\frac{\pi^n}{\det A}} .$$

Suggerimento. Si utilizzi il Teorema spettrale per diagonalizzare la matrice A , e si usi il Teorema del cambio di variabile.

6.5 Funzione Gamma e integrale di funzioni radiali

Consideriamo la funzione Gamma, così definita (per $t > 0$):

$$\Gamma(t) = \int_0^{+\infty} x^{t-1} e^{-x} dx .$$

La funzione Gamma risulta fondamentale per calcolare alcuni volumi in più dimensioni; elenchiamo alcune sue proprietà.

Innanzitutto la funzione Gamma è continua. Per dimostrarlo, facciamo vedere che rispetta le ipotesi del Teorema 1.1 dell'Appendice: mostriamo quindi che esiste una funzione integrabile $g : \mathbb{R} \rightarrow \mathbb{R}$ tale $|f(t, x)| \leq g(x) \quad \forall t \in [\frac{1}{n}, n]$.

Distinguendo il caso $0 < x \leq 1$ da quello $x > 1$, otteniamo le seguenti stime:

$$|x^{t-1}e^{-x}| \leq x^{\frac{1}{n}-1} \quad \text{per } 0 < x \leq 1 \quad \text{e} \quad |x^{t-1}e^{-x}| \leq x^{n-1} \quad \text{per } x > 1 .$$

Perciò, definendo la funzione

$$g(x) = \begin{cases} x^{\frac{1}{n}-1} & \text{se } 0 < x \leq 1 \\ x^{n-1} & \text{se } x > 1 \end{cases} ,$$

otteniamo che g è integrabile e che $|f(t, x)| \leq g(x)$.

Ora mostriamo che Γ è derivabile e vale

$$\Gamma'(t) = \int_0^{+\infty} \frac{\partial f}{\partial t}(t, x) dx = \int_0^{+\infty} \log(x) x^{t-1} e^{-x} dx .$$

Per farlo, mostriamo che valgono le ipotesi del Teorema 1.4 dell'Appendice: come per la continuità, mostriamo che esiste una funzione integrabile $g : \mathbb{R} \rightarrow \mathbb{R}$ tale $|\frac{\partial f}{\partial x}(t, x)| \leq g(x) \quad \forall t \in [\frac{1}{n}, n]$.

Distinguendo nuovamente il caso $0 < x \leq 1$ da quello $x > 1$, otteniamo le seguenti stime:

$$|\log(x) x^{t-1} e^{-x}| \leq |\log(x)| x^{\frac{1}{n}-1} \quad \text{per } 0 < x \leq 1 \quad \text{e} \quad |\log(x) x^{t-1} e^{-x}| \leq |\log(x)| x^{n-1} \quad \text{per } x > 1 .$$

Perciò, definendo la funzione

$$g(x) = \begin{cases} |\log(x)| x^{\frac{1}{n}-1} & \text{se } 0 < x \leq 1 \\ |\log(x)| x^{n-1} & \text{se } x > 1 \end{cases} ,$$

otteniamo che g è integrabile.

Inoltre, la funzione Γ soddisfa la seguente proprietà:

$$\Gamma(t+1) = t\Gamma(t) \quad \forall t > 0 .$$

Infatti

$$\Gamma(t+1) = \int_0^{+\infty} x^t e^{-x} dx = [-x^t e^{-x}]_0^{+\infty} + \int_0^{+\infty} t x^{t-1} e^{-x} dx = t\Gamma(t) .$$

Osservando che vale $\Gamma(1) = 1$, abbiamo ottenuto che la funzione Γ interpola il fattoriale sui numeri naturali. Con metodi analoghi si prova che Γ è di classe C^∞ su \mathbb{R}^+ , e inoltre si dimostra che la funzione $\log \Gamma(t)$ è convessa.

Queste condizioni in realtà caratterizzano completamente la funzione Gamma: vale infatti il seguente risultato (che non dimostriamo).

Teorema 6.26: Teorema di Bohr-Mollerup

Sia $f : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ una funzione tale che $f(1) = 1$, $f(x+1) = xf(x)$ e $\log f$ è una funzione convessa: allora $f \equiv \Gamma$.

Nel seguito indicheremo con ω_n la misura della palla unitaria n -dimensionale $B_1^{(n)}(0)$, ometteremo l'apice (n) quando sarà chiaro dal contesto e tutte le palle si intenderanno centrate nell'origine se non specificato altrimenti. Segue immediatamente dal teorema del cambio di variabile che $|B_R^{(n)}| = R^n \omega_n$.

Proposizione 6.27: Integrale di funzioni radiali

Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione radiale e sia $\phi : \mathbb{R} \rightarrow \mathbb{R}$ tale che $f(x) = \phi(|x|)$; se ϕ è integrabile, allora anche f è integrabile e per ogni $R \in [0, +\infty]$ vale

$$\int_{B_R} f(x) dx = \int_0^R \phi(t) n \omega_n t^{n-1} dt .$$

Idea di dimostrazione. Basta dimostrare la tesi per $R = +\infty$, infatti in tal caso possiamo sostituire f con $f \cdot \chi_{B_R}$ e trovare la formula generale. La formula vale se ϕ è una caratteristica $\chi_{[a,b]}$, infatti

$$\begin{aligned} \int_{\mathbb{R}^n} \chi_{[a,b]}(|x|) dx &= \int_{B_b \setminus B_a} 1 dx = |B_b \setminus B_a| = \omega_n (b^n - a^n) = \omega_n \int_a^b n t^{n-1} dt = \\ &= \int_0^{+\infty} \chi_{[a,b]}(t) n \omega_n t^{n-1} dt , \end{aligned}$$

e dunque, per additività, anche se ϕ è una funzione semplice. Si dimostra allora per una ϕ continua e integrabile approssimandola con funzioni semplici e utilizzando il Teorema di convergenza dominata; infine con un argomento di densità delle funzioni continue nelle integrabili si prova il risultato per ϕ integrabile. \square

Applichiamo la Proposizione alla funzione $\phi(t) = e^{-t^2}$:

$$\begin{aligned} \pi^{\frac{n}{2}} &= \int_{\mathbb{R}^n} e^{-|x|^2} dx = \int_0^{+\infty} e^{-t^2} n \omega_n t^{n-1} dt = \frac{n \omega_n}{2} \int_0^{+\infty} t^{n-2} e^{-t^2} 2t dt = \\ &= \frac{n \omega_n}{2} \int_0^{+\infty} s^{\frac{n}{2}-1} e^{-s} ds = \frac{n \omega_n}{2} \Gamma\left(\frac{n}{2}\right) \end{aligned}$$

di conseguenza abbiamo trovato un'espressione di ω_n in termini della funzione Γ :

$$\omega_n = \frac{\pi^{\frac{n}{2}}}{\frac{n}{2} \Gamma\left(\frac{n}{2}\right)} .$$

Esercizio 9. Data una funzione $f \in C^2(\mathbb{R}^n, \mathbb{R})$ e un raggio $r \in (0, +\infty)$, consideriamo l'integrale

$$J(r) = \int_{B_r} f(x) dx ;$$

mostrare che vale lo sviluppo asintotico $J(r) = f(0)\omega_n r^n + c_n \Delta f(0) r^{n+2} + o(r^{n+2})$ per $r \rightarrow 0^+$, dove c_n è una costante che dipende da n .

Soluzione. Chiaramente per la continuità di f , se $x \in B_r$ allora $f(x) = f(0) + o(1)$ per $r \rightarrow 0^+$, e dunque $J(r) = f(0)\omega_n r^n + o(r^n)$; l'esercizio ci dice che per le mappe C^2 si recupera una precisione maggiore sul resto, e in particolare $J(r) = f(0)\omega_n r^n + o(r^{n+1})$. Ora sviluppando f al secondo ordine con Taylor otteniamo

$$f(x) = f(0) + \nabla f(0) \cdot x + \frac{1}{2} (H_f(0)x) \cdot x + o(1)|x|^2 ;$$

chiaramente

$$\int_{B_r} f(0) dx = f(0)\omega_n r^n ,$$

mentre

$$\int_{B_r} \nabla f(0) \cdot x dx = \int_{B_r} \sum_{i=1}^n \frac{\partial f}{\partial x_i}(0) x_i dx = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(0) \int_{B_r} x_i dx = 0$$

per simmetria; ora

$$\int_{B_r} |x|^2 dx = \int_0^r t^2 n \omega_n t^{n-1} dt = \frac{n \omega_n}{n+2} r^{n+2} \quad \text{e quindi} \quad \int_{B_r} x_i^2 dx = \frac{\omega_n}{n+2} r^{n+2} ,$$

invece se $i \neq j$ per simmetria vale anche

$$\int_{B_r} x_i x_j dx = 0$$

e di conseguenza

$$\begin{aligned} \int_{B_r} \frac{1}{2} (H_f(0)x) \cdot x dx &= \frac{1}{2} \int_{B_r} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(0) x_i x_j dx = \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(0) \int_{B_r} x_i x_j dx = \\ &= \frac{1}{2} \sum_{i=1}^n \frac{\partial^2 f}{\partial x_i^2}(0) \int_{B_r} x_i^2 dx = \frac{\omega_n}{2(n+2)} \Delta f(0) r^{n+2} . \end{aligned}$$

6.6 Integrazione su curve

Definizione 6.11 (Divergenza). Dati $\Omega \subseteq \mathbb{R}^n$ un aperto e $F : \Omega \rightarrow \mathbb{R}^n$ una mappa differenziabile, definiamo la *divergenza* di F come la funzione $\operatorname{div} F : \Omega \rightarrow \mathbb{R}$ data da

$$\operatorname{div} F(x) := \frac{\partial F_1}{\partial x_1}(x) + \cdots + \frac{\partial F_n}{\partial x_n}(x) .$$

La divergenza di un campo vettoriale è dunque definita come la traccia della sua matrice jacobiana, e fisicamente misura la tendenza del campo a divergere o a convergere verso un punto dello spazio.

Dopo aver recuperato dal calcolo in una dimensione il Teorema del cambio di variabile, vediamo come si può reinterpretare in più dimensioni il Teorema di integrazione per parti.

Proposizione 6.28

Sia $Q = [0, 1]^2 \subseteq \mathbb{R}^2$ e siano $f : Q \rightarrow \mathbb{R}$ e $g : Q \rightarrow \mathbb{R}^2$ di classe C^1 ; allora

$$\iint_Q f \operatorname{div} g = \int_{\partial Q} f g \cdot \nu - \iint_Q \nabla f \cdot g ,$$

dove ν denota la normale uscente dal quadrato Q .

Dimostrazione. Dal Teorema di integrazione per parti, fissato $y \in [0, 1]$, abbiamo che

$$\int_0^1 f(x, y) \frac{\partial g_1}{\partial x}(x, y) dx = f(1, y) g_1(1, y) - f(0, y) g_1(0, y) - \int_0^1 \frac{\partial f}{\partial x}(x, y) g_1(x, y) dx$$

e integrando questa espressione nella y otteniamo

$$\iint_Q f(x, y) \frac{\partial g_1}{\partial x}(x, y) dx dy = \int_0^1 (f(1, y) g_1(1, y) - f(0, y) g_1(0, y)) dy - \iint_Q \frac{\partial f}{\partial x}(x, y) g_1(x, y) dx dy$$

e analogamente, invertendo i ruoli di x e y e considerando g_2 anziché g_1

$$\iint_Q f(x, y) \frac{\partial g_2}{\partial y}(x, y) dx dy = \int_0^1 (f(x, 1) g_2(x, 1) - f(x, 0) g_2(x, 0)) dx - \iint_Q \frac{\partial f}{\partial y}(x, y) g_2(x, y) dx dy ;$$

sommando le due espressioni termine a termine segue esattamente

$$\iint_Q f \operatorname{div} g = \int_{\partial Q} f g \cdot \nu - \iint_Q \nabla f \cdot g .$$

□

Per definire l'integrale lungo una curva abbiamo bisogno del concetto di lunghezza.

Definizione 6.12 (Lunghezza). Sia $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ una curva C^1 , definiamo la sua lunghezza come

$$\ell(\gamma) := \sup \left\{ \sum_{j=0}^{k-1} |\gamma(t_{j+1}) - \gamma(t_j)| \mid k \in \mathbb{N}, 0 = t_0 \leq t_1 \leq \dots \leq t_k = 1 \right\}.$$

Lemma 6.29: Invarianza per riparametrizzazioni

Date due curve $\gamma, \tilde{\gamma} : [0, 1] \rightarrow \mathbb{R}^n$ di classe C^1 , se esiste una funzione $\phi : [0, 1] \rightarrow [0, 1]$ continua e bigettiva tale che $\tilde{\gamma} = \gamma \circ \phi$ allora $\ell(\gamma) = \ell(\tilde{\gamma})$.

Dimostrazione. Chiaramente ϕ è monotona, e senza perdere generalità supponiamo ϕ crescente; dati $t_0 \leq \dots \leq t_k$ come nella definizione troviamo

$$\ell(\tilde{\gamma}) \geq \sum_{j=0}^{k-1} |\tilde{\gamma}(\phi^{-1}(t_{j+1})) - \tilde{\gamma}(\phi^{-1}(t_j))| = \sum_{j=0}^{k-1} |\gamma(t_{j+1}) - \gamma(t_j)| ;$$

passando al sup troviamo $\ell(\tilde{\gamma}) \geq \ell(\gamma)$. Analogamente si prova la disuguaglianza opposta. \square

Osserviamo che per definire il concetto di lunghezza non abbiamo davvero usato che i cammini fossero C^1 ; la seguente Proposizione però, valida solo sotto questa ipotesi, fornisce un metodo esplicito per il calcolo.

Proposizione 6.30

Se $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ è una curva C^1 (o C^1 a tratti), allora

$$\ell(\gamma) = \int_0^1 |\gamma'(t)| dt .$$

Dimostrazione. Per definizione di lunghezza, per ogni $\varepsilon > 0$ esiste uno spezzamento $0 = t_0 \leq t_1 \leq \dots \leq t_k = 1$ di $[0, 1]$ tale che

$$\sum_{j=0}^{k-1} |\gamma(t_{j+1}) - \gamma(t_j)| \geq \ell(\gamma) - \varepsilon ;$$

d'altra parte

$$\sum_{j=0}^{k-1} |\gamma(t_{j+1}) - \gamma(t_j)| = \sum_{j=0}^{k-1} \left| \int_{t_j}^{t_{j+1}} \gamma'(t) dt \right| \leq \sum_{j=0}^{k-1} \int_{t_j}^{t_{j+1}} |\gamma'(t)| dt = \int_0^1 |\gamma'(t)| dt ,$$

pertanto per arbitrarietà di $\varepsilon > 0$ si ha

$$\ell(\gamma) \leq \int_0^1 |\gamma'(t)| dt .$$

Per l'altra disuguaglianza sia $\varepsilon > 0$ e sia δ dato dall'uniforme continuità di γ' per ε sul compatto $[0, 1]$; sia ora $0 = t_0 \leq t_1 \leq \dots \leq t_k = 1$ uno spezzamento di $[0, 1]$ tale che $|t_{j+1} - t_j| < \delta$ e tale che la relativa somma sinistra di Riemann per $|\gamma'|$ approssima l'integrale con errore al più ε , ossia in particolare

$$\sum_{j=0}^{k-1} |\gamma'(t_j)|(t_{j+1} - t_j) \geq \int_0^1 |\gamma'(t)| dt - \varepsilon ;$$

allora

$$\begin{aligned} \sum_{j=0}^{k-1} |\gamma(t_{j+1}) - \gamma(t_j)| &= \sum_{j=0}^{k-1} \left| \int_{t_j}^{t_{j+1}} \gamma'(t) dt \right| \geq \sum_{j=0}^{k-1} \left| \int_{t_j}^{t_{j+1}} \gamma'(t_j) dt \right| - \sum_{j=0}^{k-1} \left| \int_{t_j}^{t_{j+1}} (\gamma'(t_j) - \gamma'(t)) dt \right| \geq \\ &\geq \sum_{j=0}^{k-1} |\gamma'(t_j)| (t_{j+1} - t_j) - \varepsilon(1 - 0) \geq \int_0^1 |\gamma'(t)| dt - 2\varepsilon, \end{aligned}$$

pertanto

$$\ell(\gamma) \geq \int_0^1 |\gamma'(t)| dt - 2\varepsilon \quad \text{e dunque} \quad \ell(\gamma) \geq \int_0^1 |\gamma'(t)| dt$$

per arbitrarietà di ε . □

Definizione 6.13 (Aperto C^1). Un aperto connesso $\Omega \subseteq \mathbb{R}^2$ si dice *aperto C^1* se esiste un cammino chiuso $\gamma : S^1 \rightarrow \mathbb{R}^2$ iniettivo di classe C^1 tale che $|\gamma'(t)|$ è costante e $\partial\Omega = \gamma(S^1)$.

Definizione 6.14 (Normale uscente). Dato $\Omega \subseteq \mathbb{R}^2$ un aperto C^1 con $\gamma : S^1 \rightarrow \mathbb{R}^2$ dato dalla definizione, chiamiamo *vettore normale uscente* da Ω la funzione $\nu : \partial\Omega \rightarrow \mathbb{R}^2$ data da $\nu(\gamma(t)) \equiv$ unico vettore unitario ortogonale a $\gamma'(t)$ e uscente da Ω .

Definizione 6.15 (Aperto C^1 a tratti). Un aperto $\Omega \subseteq \mathbb{R}^2$ si dice *C^1 a tratti* se $\partial\Omega$ può essere parametrizzato da un numero finito di curve C^1 .

Alla luce di questi nuovi strumenti, enunciamo la versione 2-dimensionale del Teorema della divergenza, che dimostreremo più avanti in forma più generale.

Teorema 6.31

Sia $\Omega \subseteq \mathbb{R}^2$ un aperto C^1 (a tratti) e siano $f : \Omega \rightarrow \mathbb{R}$ di classe C^1 e $g : \bar{\Omega} \rightarrow \mathbb{R}^2$ di classe C^1 . Allora

$$\int_{\Omega} f \operatorname{div} g = \int_{\partial\Omega} fg \cdot \nu - \int_{\Omega} \nabla f \cdot g.$$

6.7 Integrazione su superfici e Teorema della divergenza

Nel seguito denotiamo con B_k la palla unitaria k -dimensionale.

Definizione 6.16 (Varietà C^1). Un insieme $\Gamma \subseteq \mathbb{R}^n$ si dice *varietà k -dimensionale C^1* se per ogni $x \in \Gamma$ esiste un aperto $A \subseteq \mathbb{R}^n$ con $x \in A$ e una carta $\Phi : B_k \rightarrow \mathbb{R}^n$ di classe C^1 e iniettiva tale che $\Phi(B_k) = \Gamma \cap A$.

Definizione 6.17 (Aperto C^1). Un aperto $\Omega \subseteq \mathbb{R}^n$ si dice *aperto C^1* se per ogni $x \in \partial\Omega$ esistono un aperto $A \subseteq \mathbb{R}^n$ con $x \in A$ ed un diffeomorfismo $\Phi : B_n \rightarrow A$ di classe C^1 tale che $\Phi(B_n \cap \{x_n = 0\}) = \partial\Omega \cap A$ e $\Phi(B_n \cap \{x_n < 0\}) = \Omega \cap A$.

Nella definizione precedente non è restrittivo richiedere $\Phi(0) = x$.

Osservazione 6.32. Dalle definizioni appena date si deduce immediatamente che se $\Omega \subseteq \mathbb{R}^n$ è un aperto C^1 , $\partial\Omega$ è una $(n-1)$ -varietà C^1 .

Definizione 6.18 (Piano tangente). Sia Ω un aperto C^1 e $x \in \partial\Omega$, e sia Φ un diffeomorfismo dato dalla definizione di aperto C^1 per il punto x : chiamiamo *piano* (o *iperpiano*) *tangente* a $\partial\Omega$ in x l'insieme $T_x := x + D\Phi(0)(\{x_n = 0\})$.

Possiamo caratterizzare la giacitura del piano tangente a $\partial\Omega$ in x come l'insieme delle direzioni $\gamma'(0)$ al variare di $\gamma : (-\varepsilon, \varepsilon) \rightarrow \partial\Omega$ cammino C^1 tale che $\gamma(0) = x$. Infatti fissati A un intorno di x e $\Phi : B_n \rightarrow A$ diffeomorfismo dati dalla definizione di aperto C^1 , possiamo supporre $\gamma : (-\varepsilon, \varepsilon) \rightarrow \partial\Omega \cap A$ a meno di ridurre ε ; considerando allora il cammino indotto $\gamma_\Phi = \Phi^{-1} \circ \gamma : (-\varepsilon, \varepsilon) \rightarrow B_n \cap \{x_n = 0\}$ troviamo che

$$\gamma'(t) = \frac{\partial}{\partial t} [\Phi \circ \gamma_\Phi(t)] = D\Phi(\gamma_\Phi(t))(\gamma'_\Phi(t))$$

e dunque $\gamma'(0) = D\Phi(0)(\gamma'_\Phi(0))$, dove $\gamma'_\Phi(0) \in \{x_n = 0\}$ in quanto $\gamma_\Phi(t)$ vi appartiene per ogni t . Viceversa, se $w \in \{x_n = 0\}$, possiamo considerare il cammino $\gamma : (-\varepsilon, \varepsilon) \rightarrow \partial\Omega \cap A$ dato da $\gamma(t) = \Phi(tw)$, che è ben definito per ε sufficientemente piccolo, e osservare che γ è C^1 e vale $\gamma'(0) = D\Phi(0)(w)$.

In virtù di tale caratterizzazione, deduciamo che il piano tangente non dipende dalla mappa Φ ed è quindi ben definito.

Definizione 6.19 (Normale uscente). Sia Ω aperto C^1 , $x \in \partial\Omega$ e Φ diffeomorfismo dato dalla definizione per x : definiamo il vettore *normale uscente* da Ω in x , denotato $\nu(x)$, come l'unico vettore unitario ortogonale a T_x che ha prodotto scalare positivo con $D\Phi(0)(e_n)$.

Definizione 6.20 (Integrale di superficie). Data una carta $\Phi : D \rightarrow \mathbb{R}^n$ iniettiva e di classe C^1 , $D = B_n \cap \{x_n = 0\}$, ed una funzione $f : \Phi(D) \rightarrow \mathbb{R}$, definiamo l'*integrale di superficie* di f su $\Phi(D)$ rispetto alla *misura di superficie* dS come

$$\int_{\Phi(D)} f(y) dS(y) := \int_D (f \circ \Phi)(x) J_\Phi(x) dx_1 \dots dx_{n-1}$$

dove $J_\Phi(x) = \sqrt{\det(D\Phi^\top(x)D\Phi(x))}$ è il *fattore di dilatazione*.

Se Ω è un aperto di \mathbb{R}^n , come facciamo a integrare una funzione h su $\partial\Omega$? L'idea è cercare di parametrizzare $\partial\Omega$ con un numero finito di carte $\Phi_i : D \rightarrow \partial\Omega$ tali che $\partial\Omega = \Phi_1(D) \cup \dots \cup \Phi_n(D)$, così da avere

$$\int_{\partial\Omega} h(y) dS(y) = \sum_{i=1}^n \int_{\Phi_i(D) \setminus \bigcup_{j=1}^{i-1} \Phi_j(D)} h(y) dS(y)$$

ed integrare singolarmente sulle varie carte.

Definizione 6.21 (Aperto C^1 a pezzi). Un aperto $\Omega \subseteq \mathbb{R}^n$ si dice *aperto C^1 a pezzi* se per ogni $x \in \partial\Omega$ esistono un aperto $A \subseteq \mathbb{R}^n$ con $x \in A$ ed un omeomorfismo $\Phi : B_n \rightarrow A$ tale che $\Phi(B_n \cap \{x_n = 0\}) = \partial\Omega \cap A$ e $\Phi(B_n \cap \{x_n < 0\}) = \Omega \cap A$ (ossia Ω è un aperto C^0) e inoltre $\partial\Omega$ è unione finita di varietà $(n-1)$ -dimensionali C^1 .

Teorema 6.33: Teorema della divergenza (Ostrogradskij)

Sia $\Omega \subseteq \mathbb{R}^n$ aperto C^1 (a pezzi) e siano $f : \Omega \rightarrow \mathbb{R}$ di classe C^1 e $g : \bar{\Omega} \rightarrow \mathbb{R}^n$ di classe C^1 . Allora

$$\int_{\Omega} f \operatorname{div} g = \int_{\partial\Omega} fg \cdot \nu - \int_{\Omega} \nabla f \cdot g$$

Dimostrazione. Supponiamo che Ω sia un plurirettangolo n -dimensionale, ovvero:

$$\Omega = \prod_{i=1}^n [a_i, b_i],$$

e poniamo $y = (x_2, \dots, x_n)$; allora integrando per parti la funzione $t \mapsto f(t, y) \partial_{x_1} g_1(t, y)$ troviamo

$$\int_{a_1}^{b_1} f(t, y) \frac{\partial g_1}{\partial x_1}(t, y) dt = \left| f(\cdot, y) g_1(\cdot, y) \right|_{a_1}^{b_1} - \int_{a_1}^{b_1} \frac{\partial f}{\partial x_1}(t, y) g_1(t, y) dt.$$

Poniamo dunque

$$F_D^1 = \{b_1\} \times \prod_{i=2}^n [a_i, b_i] \quad \text{e} \quad F_S^1 = \{a_1\} \times \prod_{i=2}^n [a_i, b_i]$$

le facce ortogonali all'asse x_1 di $\partial\Omega$; allora integrando il risultato precedente in dy troviamo

$$\int_{\Omega} f \frac{\partial g_1}{\partial x_1} = \int_{F_D^1} f g_1 dS - \int_{F_S^1} f g_1 dS - \int_{\Omega} \frac{\partial f}{\partial x_1} g = \int_{F_D^1 \cup F_S^1} f g \cdot \nu - \int_{\Omega} \frac{\partial f}{\partial x_1} g_1.$$

Ripetendo il procedimento per ogni i , sommando termine a termine e osservando che

$$\sum_{i=1}^n \int_{F_D^i \cup F_S^i} f g \cdot \nu = \int_{\bigcup_{i=1}^n F_D^i \cup F_S^i} f g \cdot \nu = \int_{\partial\Omega} f g \cdot \nu$$

deduciamo la tesi.

Supponiamo adesso Ω unione finita di plurirettangoli: $\Omega = P_1 \cup \dots \cup P_m$; per ogni plurirettangolo P_i abbiamo dimostrato che

$$\int_{P_i} f \operatorname{div} g = \int_{\partial P_i} f g \cdot \nu_i - \int_{P_i} \nabla f \cdot g$$

e sommando questa espressione termine a termine per ogni i otteniamo che

$$\int_{\Omega} f \operatorname{div} g = \sum_{i=1}^m \int_{\partial P_i} f g \cdot \nu_i - \int_{\Omega} \nabla f \cdot g;$$

d'altra parte fissato i e dato $z \in \partial P_i$, o $z \in \partial\Omega$, e in tal caso $\nu(z) = \nu_i(z)$, oppure esiste un unico $j \neq i$ tale che $z \in \partial P_j$ e in tal caso $\nu_i(z) = -\nu_j(z)$, pertanto

$$\sum_{i=1}^m \int_{\partial P_i} f g \cdot \nu_i = \int_{\partial\Omega} f g \cdot \nu.$$

A questo punto supponiamo che Ω sia un poliedro: mostriamo la tesi per Ω approssimandolo con unioni finite di plurirettangoli. Fissiamo $\varepsilon > 0$: per compattezza di $\bar{\Omega}$ e $\partial\Omega$ troviamo $\delta > 0$ tale che se $x, y \in \Omega$ sono tali che $|x - y| < \delta$ allora $|h(x) - h(y)| < \varepsilon$ dove h è una qualsiasi delle funzioni continue in ballo (in particolare $h = f, g, fg, \nabla f, \operatorname{div} g$). Prendiamo allora Ω_ε un'unione finita di plurirettangoli tale che $\Omega_\varepsilon \subseteq \Omega$, $|\Omega \setminus \Omega_\varepsilon| < \varepsilon$ e inoltre i bordi di Ω e Ω_ε si decompongono in

$$\partial\Omega = \bigcup_{i=1}^m R_i \quad \text{e} \quad \partial\Omega_\varepsilon = \bigcup_{i=1}^m \Gamma_i$$

dove R_i e Γ_i sono facce $(n-1)$ -dimensionali e per ogni i esiste $q_i \in R_i$ tale che $R_i \subseteq B(q_i, \delta)$ e $\Gamma_i \subseteq B(q_i, \delta)$; chiamiamo poi ν_i la normale uscente da Γ_i e ν la normale complessivamente uscente da Ω_ε . A questo punto

$$\begin{aligned}
\int_{\partial\Omega} fg \cdot \nu &= \sum_{i=1}^m \int_{R_i} fg \cdot \nu = \sum_{i=1}^m \int_{R_i} f(q_i)g(q_i) \cdot \nu + \underbrace{\sum_{i=1}^m \int_{R_i} (fg - f(q_i)g(q_i)) \cdot \nu}_{E_1} = \\
&= \sum_{i=1}^m \int_{\Gamma_i} f(q_i)g(q_i) \cdot \nu_i + \underbrace{\sum_{i=1}^m \left(\int_{R_i} f(q_i)g(q_i) \cdot \nu - \int_{\Gamma_i} f(q_i)g(q_i) \cdot \nu_i \right)}_{E_2} + E_1 = \\
&= \sum_{i=1}^m \int_{\Gamma_i} fg \cdot \nu_i + \underbrace{\sum_{i=1}^m \int_{\Gamma_i} (f(q_i)g(q_i) - fg) \cdot \nu_i}_{E_3} + E_1 + E_2 = \\
&= \int_{\partial\Omega_\varepsilon} fg \cdot \nu + E_1 + E_2 + E_3 \stackrel{*}{=} \int_{\Omega_\varepsilon} (f \operatorname{div} g + \nabla f \cdot g) + E_1 + E_2 + E_3 = \\
&= \int_{\Omega} (f \operatorname{div} g + \nabla f \cdot g) - \underbrace{\int_{\Omega \setminus \Omega_\varepsilon} (f \operatorname{div} g + \nabla f \cdot g)}_{E_4} + E_1 + E_2 + E_3
\end{aligned}$$

ossia abbiamo dimostrato la tesi a meno di un errore $E_1 + E_2 + E_3 + E_4$, dove il passaggio (*) segue dal fatto che abbiamo già la tesi per le unioni finite di plurirettangoli. Stimiamo dunque l'errore:

$$\begin{aligned}
|E_1| &= \left| \sum_{i=1}^m \int_{R_i} (fg - f(q_i)g(q_i)) \cdot \nu \right| \leq \sum_{i=1}^m \int_{R_i} |fg - f(q_i)g(q_i)| \cdot |\nu| \leq \sum_{i=1}^m \int_{R_i} |fg - f(q_i)g(q_i)| |\nu| \leq \\
&\leq \varepsilon \sum_{i=1}^m \int_{R_i} 1 = \varepsilon \operatorname{Per}(\Omega) ;
\end{aligned}$$

osserviamo ora che se denotiamo con π_i la proiezione ortogonale sull'iperpiano passante per R_i , vale la relazione

$$\int_{\Gamma_i} f(q_i)g(q_i) \cdot \nu_i = \int_{\pi_i(\Gamma_i)} f(q_i)g(q_i) \cdot \nu$$

pertanto

$$\begin{aligned}
|E_2| &= \left| \sum_{i=1}^m \left(\int_{R_i} f(q_i)g(q_i) \cdot \nu - \int_{\Gamma_i} f(q_i)g(q_i) \cdot \nu_i \right) \right| = \left| \sum_{i=1}^m \int_{R_i \setminus \pi_i(\Gamma_i)} f(q_i)g(q_i) \cdot \nu \right| \leq \\
&= \|fg\|_\infty \sum_{i=1}^m [S(R_i) - S(\pi_i(\Gamma_i))]
\end{aligned}$$

e a meno di raffinare la scelta fatta per R_i e Γ_i , possiamo porre $S(R_i) - S(\pi_i(\Gamma_i)) \leq \varepsilon S(R_i)$, pertanto troviamo $|E_2| \leq \varepsilon \|fg\|_\infty \operatorname{Per}(\Omega)$;

$$|E_3| = \left| \sum_{i=1}^m \int_{\Gamma_i} (f(q_i)g(q_i) - fg) \cdot \nu_i \right| \leq \sum_{i=1}^m \int_{\Gamma_i} |(f(q_i)g(q_i) - fg)| |\nu_i| \leq \varepsilon \operatorname{Per}(\Omega_\varepsilon)$$

mentre

$$|E_4| = \int_{\Omega \setminus \Omega_\varepsilon} (f \operatorname{div} g + \nabla f \cdot g) \leq |\Omega \setminus \Omega_\varepsilon| \|f \operatorname{div} g + \nabla f \cdot g\|_\infty \leq \varepsilon (\|f\|_\infty \|\operatorname{div} g\|_\infty + \|\nabla f\|_\infty \|g\|_\infty) .$$

Osservando allora che $\operatorname{Per}(\Omega_\varepsilon) \leq n \operatorname{Per}(\Omega)$, abbiamo dominato l'errore con una funzione lineare in ε , dunque per arbitrarietà di ε

$$\int_{\partial\Omega} fg \cdot \nu = \int_{\Omega} f \operatorname{div} g + \int_{\Omega} \nabla f \cdot g .$$

Infine supponiamo che Ω sia C^1 (o anche C^1 a pezzi) e ripetiamo i ragionamenti già fatti: fissato $\varepsilon > 0$ scegliamo un poliedro $\Omega_\varepsilon \subseteq \Omega$ tale che $|\Omega \setminus \Omega_\varepsilon| \leq \varepsilon$, così da controllare l'errore E_4 ; scegliamo $\delta > 0$ in modo da far oscillare tutte le funzioni in gioco al più di ε su punti che distano meno di δ , e frazioniamo $\partial\Omega$ e $\partial\Omega_\varepsilon$ in pezzi di diametro minore di δ , dominando così gli errori E_1 ed E_3 ; per concludere, a meno di raffinare il frazionamento scelto, possiamo imporre che la normale ν' uscente da Ω_ε si discosti da ν al più di ε sui singoli pezzi, e ciò sistema l'errore E_2 . Poiché abbiamo già dimostrato la tesi per i poliedri Ω_ε otteniamo che

$$\int_{\partial\Omega} fg \cdot \nu = \int_{\Omega} f \operatorname{div} g + \int_{\Omega} \nabla f \cdot g .$$

□

I buoni trionfano sempre. E sempre nello stesso modo.

Aldo Pratelli

Osservazione 6.34. L'enunciato è equivalente al seguente:

$$\int_{\Omega} \operatorname{div} g = \int_{\partial\Omega} g \cdot \nu .$$

Infatti, dal Teorema, ponendo $f \equiv 1$, si ricava immediatamente questa formula. Viceversa, applicando la formula a fg si riottiene l'enunciato del Teorema, poiché

$$\operatorname{div}(fg) = \sum_{i=1}^n \frac{\partial(fg)_i}{\partial x_i} = \sum_{i=1}^n \left(\frac{\partial f}{\partial x_i} g_i + f \frac{\partial g_i}{\partial x_i} \right) = \nabla f \cdot g + f \operatorname{div} g .$$

Grazie al Teorema della divergenza, possiamo ora calcolare la superficie $(n-1)$ -dimensionale della palla $B_1^{(n)}$. Infatti, considerando le funzioni $f(x) = 1$ e $g(x) = x$, otteniamo

$$n\omega_n = \int_{B_1^{(n)}} n = \int_{\partial B_1^{(n)}} (x \cdot x) dS - \int_{B_1^{(n)}} 0 \quad \text{cioè} \quad \operatorname{Per}(B_1^{(n)}) = n\omega_n .$$

Analogamente al calcolo del volume, per calcolare la superficie della generica palla $B_r^{(n)}$, si opera il cambio di variabile $y = rx$, da cui otteniamo

$$\operatorname{Per}(B_r^{(n)}) = n\omega_n r^{n-1} .$$

6.7.1 Integrazione su grafici

Sia $\Omega \subseteq \mathbb{R}^n$ un aperto e sia $u : \Omega \rightarrow \mathbb{R}$ una mappa C^1 : consideriamo $\Sigma = \{(x, z) \in \mathbb{R}^{n+1} \mid z = u(x)\}$ il grafico di u , che chiaramente può essere parametrizzato con $\Phi : \Omega \rightarrow \Sigma$, $x \mapsto (x, u(x))$. Per definizione, se $f : \Sigma \rightarrow \mathbb{R}$ è misurabile, l'integrale di f su Σ vale

$$\int_{\Sigma} f = \int_{\Omega} f(x, u(x)) J_{\Phi}(x) dx ;$$

cerchiamo allora una formula chiusa per J_{Φ} in termini della sola u . Ora

$$D\Phi = \left(\frac{\partial\Phi}{\partial x_1} \mid \dots \mid \frac{\partial\Phi}{\partial x_n} \right) = \begin{pmatrix} I_n \\ \nabla u^\top \end{pmatrix} \quad \text{pertanto} \quad D\Phi^\top D\Phi = I_n + \nabla u \nabla u^\top ,$$

quindi $D\Phi^\top D\Phi$ ha autovalore 1 con molteplicità $n-1$ data dall'autospazio ∇u^\perp , ed ha autovalore $1 + |\nabla u|^2$ relativo all'autovettore ∇u , di conseguenza $J_{\Phi} = \sqrt{\det D\Phi^\top D\Phi} = \sqrt{1 + |\nabla u|^2}$.

6.7.2 Integrazione su superfici in \mathbb{R}^3

La geometria di \mathbb{R}^3 permette un calcolo particolarmente comodo per gli integrali di funzioni su 2-superfici: se Σ è una 2-superficie in \mathbb{R}^3 parametrizzata da $\Phi : \Omega \rightarrow \Sigma$ diffeomorfismo di classe C^1 , con $\Omega \subseteq \mathbb{R}^2$ aperto, e denotiamo $(u, v) \in \Omega$ le variabili e Φ_u e Φ_v le derivate parziali di Φ , si ha

$$D\Phi^\top D\Phi = \begin{pmatrix} |\Phi_u|^2 & \Phi_u \cdot \Phi_v \\ \Phi_u \cdot \Phi_v & |\Phi_v|^2 \end{pmatrix}$$

e dunque

$$\begin{aligned} J_\Phi &= \sqrt{\det D\Phi^\top D\Phi} = \sqrt{|\Phi_u|^2 |\Phi_v|^2 - (\Phi_u \cdot \Phi_v)^2} = \sqrt{|\Phi_u|^2 |\Phi_v|^2 - \cos^2 \theta |\Phi_u|^2 |\Phi_v|^2} = \\ &= \sin \theta |\Phi_u| |\Phi_v| = |\Phi_u \times \Phi_v| \end{aligned}$$

dove θ è l'angolo compreso tra Φ_u e Φ_v ; di conseguenza in questo caso

$$\iint_\Sigma f(x, y, z) dS(x, y, z) = \iint_\Omega f(\Phi(u, v)) \left| \frac{\partial \Phi}{\partial u} \times \frac{\partial \Phi}{\partial v} \right| dudv .$$

6.7.3 Superfici di rotazione

Consideriamo una curva $\gamma : [a, b] \rightarrow \mathbb{R}^+ \times \mathbb{R}$ in \mathbb{R}^2 , i cui punti hanno coordinate $(\rho(t), z(t))$: vogliamo calcolare la misura della superficie S in \mathbb{R}^3 ottenuta ruotando l'immagine della curva γ attorno all'asse z .

Definiamo quindi la mappa

$$\begin{aligned} \Psi : [a, b] \times [0, 2\pi] &\rightarrow \mathbb{R}^3 \\ (t, \theta) &\mapsto (\rho(t) \cos \theta, \rho(t) \sin \theta, z(t)) . \end{aligned}$$

Il differenziale di Ψ risulta:

$$D\Psi(t, \theta) = \begin{pmatrix} \rho'(t) \cos \theta & -\rho(t) \sin \theta \\ \rho'(t) \sin \theta & \rho(t) \cos \theta \\ z'(t) & 0 \end{pmatrix} ,$$

quindi abbiamo

$$D\Psi^\top D\Psi = \begin{pmatrix} \dot{\rho}^2 + \dot{z}^2 & 0 \\ 0 & \rho^2 \end{pmatrix} \quad \text{e} \quad \sqrt{\det(D\Psi^\top D\Psi)} = \rho \sqrt{\dot{\rho}^2 + \dot{z}^2} .$$

In conclusione, otteniamo:

$$\iint_S dS = \int_a^b \int_0^{2\pi} \rho \sqrt{\dot{\rho}^2 + \dot{z}^2} d\theta dt = 2\pi \int_a^b \rho \sqrt{\dot{\rho}^2 + \dot{z}^2} dt .$$

Esercizio 10. Calcolare la superficie del toro \mathcal{T} , con raggio interno (distanza del centro del tubo dal centro del toro) R e raggio esterno (raggio del tubo) r , con $0 < r < R$.

Soluzione. Per applicare la formula per le superfici di rotazione, consideriamo una sezione del toro lungo il piano $z = \rho$, che ha equazioni

$$\rho(t) = R + r \cos t \quad \text{e} \quad z(t) = r \sin t ,$$

da cui

$$\rho'(t) = -r \sin t \quad \text{e} \quad z'(t) = r \cos t .$$

Quindi, applicando la formula, otteniamo:

$$\text{area}(\mathcal{T}) = 2\pi \int_0^{2\pi} (R + r \cos t) \sqrt{r^2 + \sin^2 t + r^2 \cos^2 t} dt = 2\pi r \int_0^{2\pi} (R + r \cos t) dt = 4\pi^2 r R .$$

6.7.4 Funzioni armoniche e proprietà della media

Proposizione 6.35

Sia $\Omega \subseteq \mathbb{R}^2$ aperto e sia $f \in C^2(\Omega, \mathbb{R})$ una funzione armonica, ossia tale che $\Delta f \equiv 0$, su Ω . Sia inoltre $(x_0, y_0) \in \Omega$. Allora vale la *proprietà di media sulle sfere*:

$$\frac{1}{2\pi r} \int_{\partial B_r(x_0, y_0)} f(x, y) dS = f(x_0, y_0) .$$

Inoltre vale la *proprietà di media sulle palle*:

$$\frac{1}{\pi r^2} \int_{B_r(x_0, y_0)} f(x, y) dx dy = f(x_0, y_0) .$$

Dimostrazione. Dimostriamo solo la proprietà di media sulle sfere. Supponiamo senza perdita di generalità che $(x_0, y_0) = (0, 0)$ e che $f(x_0, y_0) = 0$, e parametrizziamo la superficie $\partial B_r(0, 0)$ con la mappa $\Phi : [0, 2\pi] \rightarrow \partial B_r(0, 0)$, data da

$$\Phi(\theta) = (r \cos \theta, r \sin \theta) \quad \text{da cui} \quad \sqrt{\det(D\Phi^\top D\Phi)} = r .$$

Ora, poiché f è di classe C^2 , valgono le ipotesi del Teorema 1.4 dell'Appendice, per cui possiamo derivare sotto il segno d'integrale:

$$\begin{aligned} \frac{\partial}{\partial r} \frac{1}{2\pi r} \int_{\partial B_r(0,0)} f(x, y) dS &= \frac{\partial}{\partial r} \frac{1}{2\pi} \int_0^{2\pi} f(r \cos \theta, r \sin \theta) d\theta = \\ &= \frac{1}{2\pi} \int_0^{2\pi} \frac{\partial}{\partial r} f(r \cos \theta, r \sin \theta) d\theta = \frac{1}{2\pi r} \int_0^{2\pi} \nabla f(r \cos \theta, r \sin \theta) \cdot \begin{pmatrix} \cos \theta \\ \sin \theta \end{pmatrix} r d\theta = \\ &= \frac{1}{2\pi r} \int_{\partial B_r(0,0)} \nabla f \cdot n \stackrel{*}{=} \frac{1}{2\pi r} \int_{B_r(0,0)} \operatorname{div}(\nabla f) dx dy = \frac{1}{2\pi r} \int_{B_r(0,0)} \Delta f dx dy = 0 , \end{aligned}$$

dove il passaggio (*) segue dal Teorema della divergenza. Dunque la funzione integranda è costante, ed è uguale al suo valore per $r \rightarrow 0$, che è $f(0, 0)$. \square

Capitolo 7

Forme differenziali

Definizione 7.1 (Applicazione multilineare alternante). Sia $k \in \mathbb{N}$. Un'applicazione $F : (\mathbb{R}^n)^k \rightarrow \mathbb{R}$ si dice *multilineare* se F è lineare in ciascuna delle k componenti, ossia

$$F(u_1, \dots, \lambda u'_i + \mu u''_i, \dots, u_k) = \lambda F(u_1, \dots, u'_i, \dots, u_k) + \mu F(u_1, \dots, u''_i, \dots, u_k)$$

per ogni $1 \leq i \leq k$, $u_j, u'_i, u''_i \in \mathbb{R}^n$ e $\lambda, \mu \in \mathbb{R}$.

Un'applicazione multilineare $F : (\mathbb{R}^n)^k \rightarrow \mathbb{R}$ si dice *alternante* se per ogni permutazione $\sigma \in S_k$ vale

$$F(u_{\sigma(1)}, \dots, u_{\sigma(k)}) = \text{sgn}(\sigma) F(u_1, \dots, u_k)$$

per ogni $u_i \in \mathbb{R}^n$.

Scegliamo k indici $i_1, \dots, i_k \in \{1, \dots, n\}$: esiste una funzione F multilineare alternante tale che $F(e_{i_1}, \dots, e_{i_k}) = 1$? Certamente no se due indici sono uguali per alternanza; d'altra parte se tutti gli indici sono diversi possiamo definire

$$F(e_{j_1}, \dots, e_{j_k}) = \begin{cases} 0 & \text{se } \{i_1, \dots, i_k\} \neq \{j_1, \dots, j_k\} \\ \text{sgn}(\sigma) & \text{dove } \sigma(i_1, \dots, i_k) = (j_1, \dots, j_k) \end{cases}$$

ed estendere per multilinearità, ossia

$$F(u_1, \dots, u_k) = F\left(\sum_{\ell=1}^n u_{1\ell} e_\ell, \dots, \sum_{\ell=1}^n u_{k\ell} e_\ell\right) = \sum_{\ell_1, \dots, \ell_k=1}^n u_{1\ell_1} \dots u_{k\ell_k} F(e_{\ell_1}, \dots, e_{\ell_k}) .$$

Denotiamo la mappa F con $dx_{i_1} \wedge \dots \wedge dx_{i_k}$ o con dx_I dove I rappresenta il multi-indice (i_1, \dots, i_k) .

L'insieme delle applicazioni k -multilineari e alternanti su \mathbb{R}^n , che denotiamo $\mathcal{M}_{k,n}$ o anche $\Lambda^k(\mathbb{R}^n)$, formano naturalmente uno spazio vettoriale di dimensione $\binom{n}{k}$, una cui base è offerta dai dx_I al variare di I tra tutti i multi-indici di $\{1, \dots, n\}$: di conseguenza come spazi vettoriali $\Lambda^k(\mathbb{R}^n) \cong \Lambda^{n-k}(\mathbb{R}^n)$.

Definizione 7.2 (Forma differenziale). Dato $\Omega \subseteq \mathbb{R}^n$ aperto, una k -forma differenziale è una funzione $\alpha : \Omega \rightarrow \Lambda^k(\mathbb{R}^n)$. Per ogni $x \in \Omega$ $\alpha(x)$ ammette un'unica scrittura

$$\alpha(x) = \sum_I \alpha_I(x) dx_I ;$$

la forma α si dice di classe C^m se tutte le mappe $\alpha_I : \Omega \rightarrow \mathbb{R}$ sono di classe C^m .

La definizione quasi lascia... un po'... tristi.

Aldo Pratelli

In particolare, le 0-forme sono semplicemente le funzioni scalari, mentre le 1-forme possono essere pensate ($\mathbb{R}^n \cong \Lambda^1(\mathbb{R}^n)$) come funzioni a valori vettoriali $\alpha : \Omega \rightarrow \mathbb{R}^n$ ossia campi vettoriali. Nel seguito sottintenderemo spesso il dominio Ω .

L'insieme $\mathcal{F}_{k,n}$ delle k -forme su \mathbb{R}^n è uno spazio vettoriale con le chiare operazioni di somma e prodotto per scalare. La seguente definizione introduce anche un prodotto tra forme differenziali.

Definizione 7.3 (Prodotto wedge). Il *prodotto wedge* tra k -forme e h -forme è l'operazione $\wedge : \mathcal{F}_{k,n} \times \mathcal{F}_{h,n} \rightarrow \mathcal{F}_{k+h,n}$ che associa a $\alpha \in \mathcal{F}_{k,n}$ e $\beta \in \mathcal{F}_{h,n}$ con

$$\alpha(x) = \sum_I \alpha_I(x) dx_I \quad \text{e} \quad \beta(x) = \sum_J \beta_J(x) dx_J$$

l'elemento $\alpha \wedge \beta \in \mathcal{F}_{k+h,n}$ dato da

$$\alpha \wedge \beta(x) := \sum_{I,J} \alpha_I(x) \beta_J(x) dx_I \wedge dx_J .$$

Osserviamo che il prodotto wedge è costruito ponendo

$$(dx_{i_1} \wedge \cdots \wedge dx_{i_k}) \wedge (dx_{j_1} \wedge \cdots \wedge dx_{j_h}) := dx_{i_1} \wedge \cdots \wedge dx_{i_k} \wedge dx_{j_1} \wedge \cdots \wedge dx_{j_h}$$

ed estendendo per multilinearità in ogni punto. Segue inoltre immediatamente dalla definizione che $\alpha \wedge \beta = (-1)^{hk} \beta \wedge \alpha$.

Sia $\Omega \subseteq \mathbb{R}^n$ un aperto e sia $f \in C^1(\Omega, \mathbb{R})$ una funzione, che interpretiamo come 0-forma. Indichiamo con df la 1-forma data da

$$df(x) := \frac{\partial f}{\partial x_1}(x) dx_1 + \cdots + \frac{\partial f}{\partial x_n}(x) dx_n \quad \text{per ogni } x \in \Omega.$$

Definizione 7.4 (Differenziale esterno). Data $\alpha \in \mathcal{F}_{k,n}$ forma di classe C^m definiamo il suo *differenziale esterno* $d\alpha$ come la $(k+1)$ -forma di classe C^{m-1} data

$$d\alpha(x) := \sum_I d\alpha_I(x) dx_I \quad \text{dove} \quad \alpha(x) = \sum_I \alpha_I(x) dx_I .$$

Proposizione 7.1

Date due forme $\alpha \in \mathcal{F}_{k,n}$ e $\beta \in \mathcal{F}_{h,n}$ di classe C^1 , il differenziale esterno del loro prodotto wedge è

$$d(\alpha \wedge \beta) = d\alpha \wedge \beta + (-1)^k \alpha \wedge d\beta.$$

Dimostrazione. Per le proprietà di additività del differenziale e di distributività del prodotto wedge possiamo considerare il caso

$$\alpha = f dx_{i_1} \wedge \cdots \wedge dx_{i_k} \quad \text{e} \quad \beta = g dx_{j_1} \wedge \cdots \wedge dx_{j_h}.$$

Calcoliamo

$$\begin{aligned}\alpha \wedge \beta &= fg \, dx_{i_1} \wedge \cdots \wedge dx_{i_k} \wedge dx_{j_1} \wedge \cdots \wedge dx_{j_h} \\ d(\alpha \wedge \beta) &= \sum_{\ell=1}^n \frac{\partial fg}{\partial x_\ell} dx_\ell \wedge dx_{i_1} \wedge \cdots \wedge dx_{j_h} = \sum_{\ell=1}^n \left(\frac{\partial f}{\partial x_\ell} g + f \frac{\partial g}{\partial x_\ell} \right) dx_\ell \wedge dx_{i_1} \wedge \cdots \wedge dx_{j_h} = \\ &= \sum_{\ell=1}^n \frac{\partial f}{\partial x_\ell} g \, dx_\ell \wedge dx_{i_1} \wedge \cdots \wedge dx_{j_h} + \sum_{\ell=1}^n f \frac{\partial g}{\partial x_\ell} dx_\ell \wedge dx_{i_1} \wedge \cdots \wedge dx_{j_h}\end{aligned}$$

e osserviamo che il primo termine coincide con

$$\left(\sum_{\ell=1}^n \frac{\partial f}{\partial x_\ell} dx_\ell \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} \right) \wedge (g \, dx_{j_1} \wedge \cdots \wedge dx_{j_h}) = d\alpha \wedge \beta$$

mentre il secondo termine coincide con

$$(-1)^k \left[(f \, dx_{i_1} \wedge \cdots \wedge dx_{i_k}) \wedge \left(\sum_{\ell=1}^n \frac{\partial g}{\partial x_\ell} dx_\ell \wedge dx_{j_1} \wedge \cdots \wedge dx_{j_h} \right) \right] = (-1)^k \alpha \wedge d\beta .$$

□

7.1 Forme chiuse e forme esatte

Esiste una corrispondenza naturale tra 1-forme differenziali e campi vettoriali: data una 1-forma $\alpha = \alpha_1 dx_1 + \cdots + \alpha_n dx_n$ possiamo associarle il campo $H : \Omega \rightarrow \mathbb{R}^n$ dato da $H(x) = (\alpha_1(x), \dots, \alpha_n(x))$ e viceversa.

Se sono un fisico disegno frecce.

Aldo Pratelli

Definizione 7.5 (Rotore). Dato il campo vettoriale $H : \Omega \rightarrow \mathbb{R}^3$ di classe C^1 , definiamo il suo *rotore* come la funzione $\text{Rot}H : \Omega \rightarrow \mathbb{R}^3$ data da

$$\text{Rot}H = \left(\frac{\partial H_3}{\partial x_2} - \frac{\partial H_2}{\partial x_3}, \frac{\partial H_1}{\partial x_3} - \frac{\partial H_3}{\partial x_1}, \frac{\partial H_2}{\partial x_1} - \frac{\partial H_1}{\partial x_2} \right) .$$

Definizione 7.6 (Forma chiusa). Una k -forma α di classe C^1 si dice *chiusa* se $d\alpha = 0$.

Data una 0-forma $f \in C^1$, abbiamo che $df = \partial_{x_1} f dx_1 + \cdots + \partial_{x_n} f dx_n$, pertanto il campo vettoriale associato alla 1-forma df è proprio ∇f ; le 0-forme chiuse sono allora le funzioni a gradiente nullo, ossia localmente costanti.

Sia $\alpha = \alpha_1 dx_1 + \cdots + \alpha_n dx_n$ una 1-forma C^1 : allora α è chiusa, ossia $d\alpha = 0$, se e solo se per ogni $i \neq j$ vale

$$\frac{\partial \alpha_j}{\partial x_i} - \frac{\partial \alpha_i}{\partial x_j} = 0 ;$$

nel caso particolare delle 1-forme su \mathbb{R}^3 segue subito dalla definizione che $\alpha \in \mathcal{F}_{1,3}$ è chiusa se e solo se $\text{Rot}H = 0$.

Similmente se α è una 2-forma C^1 , che scriviamo come

$$\alpha = \sum_{i < j} \alpha_{ij} dx_i \wedge dx_j$$

allora α è chiusa se e solo se per ogni tripla di indici i, j, k distinti vale

$$\frac{\partial \alpha_{jk}}{\partial x_i} - \frac{\partial \alpha_{ik}}{\partial x_j} + \frac{\partial \alpha_{ij}}{\partial x_k} = 0 ;$$

nel caso delle 2-forme su \mathbb{R}^3 , alla forma $\alpha = \alpha_{12}dx_1 \wedge dx_2 + \alpha_{23}dx_2 \wedge dx_3 + \alpha_{31}dx_3 \wedge dx_1$ possiamo associare il campo $H = (\alpha_{23}, \alpha_{31}, \alpha_{12})$: segue allora dall'osservazione precedente che α è chiusa se e solo se $\operatorname{div} H = 0$.

Osservazione 7.2. Per quanto riguarda le forme su \mathbb{R}^3 , abbiamo associato piuttosto naturalmente un campo vettoriale sia ad una 1-forma sia ad una 2-forma: ciò si accorda con l'isomorfismo $\Lambda^k(\mathbb{R}^n) \cong \Lambda^{n-k}(\mathbb{R}^n)$ che chiaramente si estende alle forme.

Definizione 7.7 (Forma esatta). Una k -forma differenziale α si dice *esatta* se esiste una $(k-1)$ -forma β tale che $\alpha = d\beta$.

Lemma 7.3

Se α è una k -forma di classe C^2 allora $d^2\alpha = 0$.

Dimostrazione. Per additività, basta mostrare la tesi per $\alpha = \alpha_I dx_I$: esplicitamente

$$d^2\alpha = d(d\alpha) = \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_i} dx_j \wedge dx_i \wedge dx_I = \sum_{i < j} \left(\frac{\partial^2 f}{\partial x_i \partial x_j} - \frac{\partial^2 f}{\partial x_j \partial x_i} \right) dx_i \wedge dx_j \wedge dx_I = 0$$

dove l'ultima uguaglianza segue dal Teorema di Schwarz. \square

Corollario 7.4

Le forme esatte sono chiuse.

Mostriamo subito che in generale il viceversa è falso. Consideriamo su $\Omega = \mathbb{R}^2 \setminus \{0\}$ la forma

$$\alpha = \frac{1}{x^2 + y^2} (x dy - y dx)$$

che è chiusa in quanto

$$d\alpha = \left(\frac{1}{x^2 + y^2} - \frac{2x^2}{(x^2 + y^2)^2} + \frac{1}{x^2 + y^2} - \frac{2y^2}{(x^2 + y^2)^2} \right) dx \wedge dy = 0 .$$

Se α fosse esatta esisterebbe una funzione $H : \Omega \rightarrow \mathbb{R}$ tale che

$$\nabla H = \left(\frac{-y}{x^2 + y^2}, \frac{x}{x^2 + y^2} \right) ,$$

allora definendo $\phi(\theta) := H(\cos \theta, \sin \theta)$ troveremmo

$$\phi'(\theta) = \frac{d}{d\theta} H(\cos \theta, \sin \theta) = \nabla H(\cos \theta, \sin \theta) \cdot (-\sin \theta, \cos \theta) = 1$$

che è assurdo in quanto

$$H(1, 0) = \phi(2\pi) = \phi(0) + \int_0^{2\pi} \phi'(\theta) d\theta = H(1, 0) + 2\pi .$$

Definizione 7.8 (Integrale di 1-forme su curve). Dato $\Omega \subseteq \mathbb{R}^n$ aperto, sia α una 1-forma su Ω e sia $\gamma : [a, b] \rightarrow \Omega$ un cammino C^1 ; definiamo l'integrale di α sul cammino γ (o sulla curva $\gamma([a, b])$) come

$$\int_{\gamma} \alpha := \int_a^b \alpha(\gamma(t))(\gamma'(t)) dt .$$

Cerchiamo di capire se, sotto opportune ipotesi, può valere l'implicazione opposta del Corollario. Sia Ω connesso per archi: prendiamo α una 1-forma chiusa su Ω e cerchiamo di definire un potenziale H . Osserviamo che se $\tau : [0, 1] \rightarrow \Omega$ è un cammino C^1 e $\alpha = dH$, allora

$$\frac{\partial}{\partial t} H(\tau(t)) = \alpha(\tau(t)) \cdot \tau'(t)$$

e dunque deve essere

$$H(\tau(1)) = H(\tau(0)) + \int_0^1 \alpha(\tau(t)) \cdot \tau'(t) dt = H(\tau(0)) + \int_{\tau} \alpha ,$$

pertanto è necessario che α abbia integrale nullo lungo tutti i cammini chiusi su Ω .

Proposizione 7.5

Una 1-forma α è esatta se e solo se il suo integrale lungo tutti i cammini chiusi è nullo.

Dimostrazione. Abbiamo già osservato il *solo se* della proposizione. Fissiamo ora $P \in \Omega$ e per ogni $Q \in \Omega$ scegliamo $\tau = \tau_Q$ un cammino C^1 tra P e Q : definiamo dunque il potenziale $H : \Omega \rightarrow \mathbb{R}$ di α come

$$H(Q) := \int_{\tau} \alpha ;$$

allora H è ben definito perché se $\tilde{\tau}$ è un secondo cammino C^1 tra P e Q , allora

$$\int_{\tilde{\tau}} \alpha = \int_{\tilde{\tau} * \tau^{-1}} \alpha + \int_{\tau} \alpha = \int_{\tau} \alpha$$

dato che per ipotesi l'integrale di α su $\tilde{\tau} * \tau^{-1}$ è nullo. Mostriamo quindi che $\alpha = dH$. Consideriamo il cammino $\gamma_{\varepsilon} : [0, \varepsilon] \rightarrow \Omega$ dato da $\gamma_{\varepsilon}(t) = Q + te_i$, allora

$$H(Q + \varepsilon e_i) = \int_{\tau * \gamma_{\varepsilon}} \alpha = \int_{\tau} \alpha + \int_{\gamma_{\varepsilon}} \alpha = H(Q) + \int_0^{\varepsilon} \alpha_i(Q + te_i) dt = H(Q) + \varepsilon \alpha_i(Q) + o(\varepsilon)$$

vale a dire

$$\frac{\partial H}{\partial x_i}(Q) = \alpha_i(Q) \quad \text{ossia} \quad \alpha = dH .$$

□

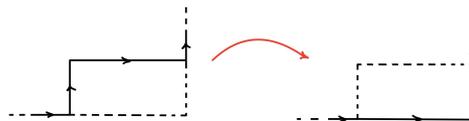
Teorema 7.6

Sia α una 1-forma chiusa su un rettangolo $\Omega \subseteq \mathbb{R}^2$, allora α ha integrale nullo su tutti i cammini chiusi e quindi è esatta.

Dimostrazione. Dimostriamo il Teorema alleggerendo progressivamente le ipotesi sui cammini che consideriamo. Sia innanzitutto γ un cammino C^1 che parametrizza il bordo di un rettangolo $R = [a, b] \times [c, d] \subseteq \Omega$ in senso antiorario, allora

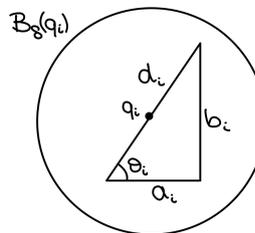
$$\begin{aligned} \int_{\gamma} \alpha &= \int_a^b \alpha_1(t, c) dt + \int_c^d \alpha_2(b, s) ds - \int_a^b \alpha_1(t, d) dt - \int_c^d \alpha_2(a, s) ds = \\ &= \int_a^b (-\alpha_1(t, d) + \alpha_1(t, c)) dt + \int_c^d (\alpha_2(b, s) - \alpha_2(a, s)) ds = \\ &= \int_a^b \int_c^d -\frac{\partial \alpha_1}{\partial y}(t, s) ds dt + \int_c^d \int_a^b \frac{\partial \alpha_2}{\partial x}(t, s) dt ds = \iint_R \left(\frac{\partial \alpha_2}{\partial x} - \frac{\partial \alpha_1}{\partial y} \right) dt ds = 0 . \end{aligned}$$

Se γ parametrizza un'unione finita di tratti orizzontali e verticali, possiamo ricondurci al bordo di un rettangolo attraverso il metodo illustrato in figura.



Se l'immagine di γ è un poligono, basta dimostrare che l'integrale su un tratto obliquo d si riconduce ad integrali su tratti orizzontali e verticali.

Fissato $\varepsilon > 0$, sia $\delta > 0$ tale che per $|v - w| < \delta$ si ha $|\alpha_i(v) - \alpha_i(w)| < \varepsilon$, $i = 1, 2$. Possiamo suddividere il segmento obliquo d in tratti d_1, \dots, d_n e scegliere $q_i \in d_i$ tali che $d_i \subseteq B_\delta(q_i)$: di conseguenza le proiezioni orizzontali a_i e verticali b_i di ogni d_i sono contenuti in $B_\delta(q_i)$.



Allora

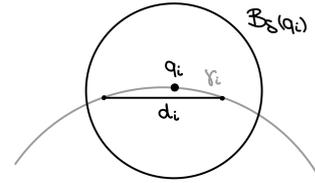
$$\begin{aligned} \int_d \alpha &= \sum_{i=1}^n \int_{d_i} \alpha = \sum_{i=1}^n \int_{d_i} \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix} \cdot \begin{pmatrix} \cos \theta_i \\ \sin \theta_i \end{pmatrix} = \\ &= \sum_{i=1}^n \int_{d_i} \begin{pmatrix} \alpha_1(q_i) \\ \alpha_2(q_i) \end{pmatrix} \cdot \begin{pmatrix} \cos \theta_i \\ \sin \theta_i \end{pmatrix} + \underbrace{\sum_{i=1}^n \int_{d_i} \begin{pmatrix} \alpha_1 - \alpha_1(q_i) \\ \alpha_2 - \alpha_2(q_i) \end{pmatrix} \cdot \begin{pmatrix} \cos \theta_i \\ \sin \theta_i \end{pmatrix}}_{E_1} = \\ &= \sum_{i=1}^n \ell(d_i) [\alpha_1(q_i) \cos \theta_i + \alpha_2(q_i) \sin \theta_i] + E_1 = \sum_{i=1}^n [\alpha_1(q_i) \ell(a_i) + \alpha_2(q_i) \ell(b_i)] + E_1 = \\ &= \sum_{i=1}^n \left(\int_{a_i} \alpha_1(q_i) + \int_{b_i} \alpha_2(q_i) \right) + E_1 = \\ &= \sum_{i=1}^n \left(\int_{a_i} \alpha_1 + \int_{b_i} \alpha_2 \right) + \underbrace{\sum_{i=1}^n \int_{a_i} (\alpha_1(q_i) - \alpha_1)}_{E_2} + \underbrace{\sum_{i=1}^n \int_{b_i} (\alpha_2(q_i) - \alpha_2)}_{E_3} + E_1 ; \end{aligned}$$

stimiamo allora l'errore:

$$\begin{aligned} |E_1| &= \left| \sum_{i=1}^n \int_{d_i} \begin{pmatrix} \alpha_1 - \alpha_1(q_i) \\ \alpha_2 - \alpha_2(q_i) \end{pmatrix} \cdot \begin{pmatrix} \cos \theta_i \\ \sin \theta_i \end{pmatrix} \right| \leq \sum_{i=1}^n \int_{d_i} \left| \begin{pmatrix} \alpha_1 - \alpha_1(q_i) \\ \alpha_2 - \alpha_2(q_i) \end{pmatrix} \right| \left| \begin{pmatrix} \cos \theta_i \\ \sin \theta_i \end{pmatrix} \right| \leq \\ &\leq \sum_{i=1}^n \int_{d_i} \sqrt{2} \varepsilon = \sqrt{2} \varepsilon \ell(d) ; \\ |E_2| &= \left| \sum_{i=1}^n \int_{a_i} (\alpha_1(q_i) - \alpha_1) \right| \leq \sum_{i=1}^n \int_{a_i} |\alpha_1(q_i) - \alpha_1| \leq \sum_{i=1}^n \int_{a_i} \varepsilon = \sum_{i=1}^n \varepsilon \ell(a_i) \leq \\ &\leq \sum_{i=1}^n \varepsilon \ell(d_i) = \varepsilon \ell(d) ; \\ |E_3| &= \left| \sum_{i=1}^n \int_{b_i} (\alpha_2(q_i) - \alpha_2) \right| \leq \sum_{i=1}^n \int_{b_i} |\alpha_2(q_i) - \alpha_2| \leq \sum_{i=1}^n \int_{b_i} \varepsilon = \sum_{i=1}^n \varepsilon \ell(b_i) \leq \\ &\leq \sum_{i=1}^n \varepsilon \ell(d_i) = \varepsilon \ell(d) . \end{aligned}$$

Complessivamente allora l'integrale di α su γ differisce dall'integrale di α su un cammino chiuso di tipo orizzontale-verticale (che dunque è nullo) per un errore dominato in modulo da $\varepsilon(\sqrt{2} + 2)\ell(\gamma)$, e per arbitrarietà di ε si conclude.

Infine se γ è una curva C^1 (o anche C^1 a tratti), e quindi ha derivata continua, fissato ε e scelto δ dato dall'uniforme continuità di α_1 e α_2 , possiamo spezzare γ in archi $\gamma_1, \dots, \gamma_n$ in modo che γ_i' oscilli in modulo al più di ε e che, considerando i cammini d_i che parametrizzano le corde degli archi γ_i a velocità costante, esistano punti q_i giacenti su γ_i tali che $\text{Im}\gamma_i, \text{Im}d_i \subseteq B_\delta(q_i)$.



Stimando i singoli errori come nel caso precedente troviamo che l'integrale di α su γ è dominato da una funzione lineare in ε , e questo conclude la dimostrazione. \square

Teorema 7.7

Dato $\Omega \subseteq \mathbb{R}^n$ semplicemente connesso, le 1-forme chiuse su Ω sono esatte.

Dimostrazione. Mostriamo che le forme chiuse hanno integrali nulli sui cammini chiusi. Sia α una 1-forma su Ω e sia $\gamma : [0, 1] \rightarrow \Omega$ un cammino chiuso C^1 ; poniamo $x = \gamma(0) = \gamma(1)$. Poiché Ω è semplicemente connesso, esiste un'omotopia di cammini $H : [0, 1] \times [0, 1] \rightarrow \Omega$ di classe C^2 tra γ e il cammino costante c_x , ossia $H(s, 0) = \gamma(s)$ e $H(s, 1) = x$; poniamo $\gamma_t : [0, 1] \rightarrow \Omega$ il cammino che segue il tempo t dell'omotopia, ovvero $\gamma_t(s) = H(s, t)$, e mostriamo che la funzione

$$t \mapsto \int_{\gamma_t} \alpha, \text{ che vale } \int_{\gamma} \alpha \text{ per } t = 0 \text{ e } 0 \text{ per } t = 1,$$

è costante: ciò conclude la dimostrazione.

Ora, se $\alpha = \alpha_1 dx_1 + \dots + \alpha_n dx_n$ allora¹

$$\begin{aligned} \frac{d}{dt} \int_{\gamma_t} \alpha &= \frac{d}{dt} \int_a^b \alpha(H(s, t)) \left(\frac{\partial H}{\partial s} \right) ds = \frac{d}{dt} \int_a^b \sum_{i=1}^n \alpha_i(H(s, t)) \frac{\partial H_i}{\partial s} ds = \\ &= \int_a^b \frac{\partial}{\partial t} \left[\sum_{i=1}^n \alpha_i(H(s, t)) \frac{\partial H_i}{\partial s} \right] ds = \int_a^b \sum_{i=1}^n \frac{\partial}{\partial t} \left[\alpha_i(H(s, t)) \frac{\partial H_i}{\partial s} \right] ds = \\ &= \int_a^b \sum_{i=1}^n \left(\frac{\partial}{\partial t} [\alpha_i(H(s, t))] \frac{\partial H_i}{\partial s} + \alpha_i(H(s, t)) \frac{\partial^2 H_i}{\partial t \partial s} \right) ds \end{aligned}$$

quindi integrando per parti il secondo termine troviamo che

$$\int_a^b \alpha_i(H(s, t)) \frac{\partial^2 H_i}{\partial s \partial t} ds = \left[\alpha_i(H(s, t)) \frac{\partial H_i}{\partial t} \right]_{s=a}^b - \int_a^b \frac{\partial}{\partial s} \left[\alpha_i(H(s, t)) \right] \frac{\partial H_i}{\partial t} ds$$

e

$$\left[\alpha_i(H(s, t)) \frac{\partial H_i}{\partial t} \right]_{s=a}^b = \alpha_i(H(b, t)) \frac{\partial H_i}{\partial t}(b, t) - \alpha_i(H(a, t)) \frac{\partial H_i}{\partial t}(a, t) = 0$$

perché l'omotopia è a estremi fissi, quindi

$$\begin{aligned} \frac{d}{dt} \int_{\gamma_t} \alpha &= - \int_a^b \sum_{i=1}^n \left(\frac{\partial}{\partial s} [\alpha_i(H(s, t))] \frac{\partial H_i}{\partial t} - \frac{\partial}{\partial t} [\alpha_i(H(s, t))] \frac{\partial H_i}{\partial s} \right) ds = \\ &= - \int_a^b \sum_{i=1}^n \left[\left(\nabla \alpha_i(H(s, t)) \cdot \frac{\partial H}{\partial s} \right) \frac{\partial H_i}{\partial t} - \left(\nabla \alpha_i(H(s, t)) \cdot \frac{\partial H}{\partial t} \right) \frac{\partial H_i}{\partial s} \right] ds = \\ &= - \int_a^b \sum_{i=1}^n \nabla \alpha_i(H(s, t)) \cdot \left(\frac{\partial H}{\partial s} \frac{\partial H_i}{\partial t} - \frac{\partial H}{\partial t} \frac{\partial H_i}{\partial s} \right) ds = \\ &= - \int_a^b \sum_{i=1}^n \sum_{j=1}^n \frac{\partial \alpha_i}{\partial x_j}(H(s, t)) \left(\frac{\partial H_j}{\partial s} \frac{\partial H_i}{\partial t} - \frac{\partial H_j}{\partial t} \frac{\partial H_i}{\partial s} \right) ds = \\ &= - \int_a^b (d\alpha)(H(s, t)) \left(\frac{\partial H}{\partial s}, \frac{\partial H}{\partial t} \right) ds = 0 \end{aligned}$$

¹Per il passaggio della derivata sotto il segno di integrale si consulti l'Appendice.

perché α è chiusa. \square

Commento. Teoricamente H può essere solo C^0 , ma si può mostrare che possiamo prendere \hat{H} di classe C^∞ tale che $\|H - \hat{H}\|_\infty$ e l'integrale di $\|\partial_s H - \partial_s \hat{H}\|_\infty$ sono piccoli a piacere.

Definizione 7.9 (Integrale di una forma). Sia $S = \Phi(A)$ una k -superficie di classe C^m , dove $A \subseteq \mathbb{R}^k$ è aperto e $\Phi : A \rightarrow \mathbb{R}^n$ è una carta iniettiva di classe C^m tale che $\text{rk } D\Phi = k$; sia α una k -forma su $\Omega \subseteq \mathbb{R}^n$ aperto con $S \subseteq \Omega$:

$$\alpha = \sum_{i_1 < \dots < i_k} \alpha_{i_1, \dots, i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k} .$$

Allora l'integrale della k -forma α sulla k -superficie S è

$$\int_S \alpha := \sum_{i_1 < \dots < i_k} \int_A \alpha_{i_1, \dots, i_k}(\Phi(z)) \frac{\partial(\Phi^{i_1}, \dots, \Phi^{i_k})}{\partial(z_1, \dots, z_k)}(z) dz_1 \dots dz_k$$

dove

$$\frac{\partial(\Phi^{i_1}, \dots, \Phi^{i_k})}{\partial(z_1, \dots, z_k)} := \det \left(\frac{\partial \Phi^{i_\alpha}}{\partial z_\beta} \right)_{\substack{\alpha=1, \dots, k \\ \beta=1, \dots, k}} .$$

L'integrale di una forma è ben definito grazie al Teorema di cambio di variabile: infatti se Ψ è un'altra carta come nella definizione tale che $S = \Psi(A)$, allora per ogni multi-indice $I = (i_1, \dots, i_k)$ si ha

$$\begin{aligned} \int_{\Phi(A)} \alpha_I dx_I &= \int_A \alpha_I(\Phi(z)) \frac{\partial(\Phi^{i_1}, \dots, \Phi^{i_k})}{\partial(z_1, \dots, z_k)}(z) dz = \int_A (\alpha_I \circ \Phi)(z) \det D\Phi_I(z) dz = \\ &\stackrel{*}{=} \int_A (\alpha_I \circ \Phi \circ \Phi^{-1} \circ \Psi)(z) \det D\Phi_I((\Phi^{-1} \circ \Psi)(z)) \det D(\Phi^{-1} \circ \Psi)(z) dz = \\ &= \int_A \alpha_I(\Psi(z)) \det \left(D\Phi_I((\Phi^{-1} \circ \Psi)(z)) D(\Phi^{-1} \circ \Psi)(z) \right) dz = \\ &= \int_A \alpha_I(\Psi(z)) \det \left(D\Phi((\Phi^{-1} \circ \Psi)(z)) D(\Phi^{-1} \circ \Psi)(z) \right)_I dz = \\ &= \int_A \alpha_I(\Psi(z)) \det D(\Phi \circ \Phi^{-1} \circ \Psi)_I(z) dz = \int_A \alpha_I(\Psi(z)) \det D\Psi_I(z) dz = \\ &= \int_A \alpha_I(\Psi(z)) \frac{\partial(\Psi^{i_1}, \dots, \Psi^{i_k})}{\partial(z_1, \dots, z_k)}(z) dz = \int_{\Psi(A)} \alpha_I dx_I , \end{aligned}$$

dove se $M \in M(n, k, \mathbb{R})$ indichiamo con M_I il minore di M ottenuto selezionando le righe i_1, \dots, i_k , ed al passaggio (*) abbiamo cambiato variabile con $z \mapsto \Phi^{-1}(\Psi(z))$. In realtà ciò che abbiamo scritto è vero a meno del segno di $\det D(\Phi^{-1} \circ \Psi)$, che stabilisce se Φ e Ψ forniscono a S la stessa orientazione.

7.2 Lemma di Poincaré e Pullback di forme

Teorema 7.8: Lemma di Poincaré

Sia $\Omega = [0, 1]^n$, allora ogni k -forma chiusa su Ω è esatta.

Dimostrazione. Definiamo \mathcal{A}_ℓ^k , con $1 \leq \ell \leq n$, l'insieme delle k -forme su Ω che possono essere scritte solo in termini di dx_1, \dots, dx_ℓ . Vogliamo mostrare che le forme chiuse in \mathcal{A}_ℓ^k sono esatte: procediamo per induzione su ℓ . Per $\ell = 1$, abbiamo che le forme in \mathcal{A}_1^k sono nulle per $k > 1$, mentre se $k = 1$ abbiamo la tesi per quanto visto sulle 1-forme. Supponiamo adesso la tesi vera per $\ell - 1$,

e sia α una k -forma chiusa in \mathcal{A}_ℓ^k . Possiamo scrivere

$$\alpha = \sum_I \alpha_I dx_I$$

dove il generico I è un multi-indice di lunghezza k in $\{1, \dots, \ell\}$. Adesso per ipotesi $d\alpha = 0$, dunque per ogni I e per ogni $j > \ell$, poiché $\partial_{x_j} \alpha_I$ è l'unico termine che compare come coefficiente di $dx_j \wedge dx_I$, deve essere $\partial_{x_j} \alpha_I = 0$.

Scriviamo adesso $\alpha = \alpha^- + \tilde{\alpha}$ con $\alpha^- \in \mathcal{A}_{\ell-1}^k$, mentre

$$\tilde{\alpha} = \sum_J \alpha_{(J,\ell)} dx_J \wedge dx_\ell$$

dove J è un multi-indice di lunghezza $k-1$ in $\{1, \dots, \ell-1\}$; definiamo le mappe

$$g_I := \int_0^{x_\ell} \alpha_I(x_1, \dots, x_{\ell-1}, t, x_{\ell+1}, \dots, x_n) dt \quad :$$

ogni g_I è di classe almeno C^1 , inoltre

$$\frac{\partial g_I}{\partial x_j} = 0 \quad \text{per } j > \ell, \quad \text{mentre} \quad \frac{\partial g_I}{\partial x_\ell} = \alpha_I \quad .$$

Definiamo allora la forma

$$\eta := (-1)^{k-1} \sum_J g_{(J,\ell)} dx_J$$

chiaramente $\eta \in \mathcal{A}_{\ell-1}^{k-1}$, e inoltre $\beta := \alpha - d\eta$ è una k -forma chiusa. Per concludere osserviamo che

$$\begin{aligned} \beta &= \alpha^- + \tilde{\alpha} - d\eta = \alpha^- + \sum_J \alpha_{(J,\ell)} dx_J \wedge dx_\ell - (-1)^{k-1} \sum_{j=1}^n \sum_J \frac{\partial g_{(J,\ell)}}{\partial x_j} dx_j \wedge dx_J = \\ &= \alpha^- + \sum_J \alpha_{(J,\ell)} dx_J \wedge dx_\ell - (-1)^{k-1} \sum_{j=1}^{\ell-1} \sum_J \frac{\partial g_{(J,\ell)}}{\partial x_j} dx_j \wedge dx_J = \\ &= \alpha^- + \sum_J \alpha_{(J,\ell)} dx_J \wedge dx_\ell - (-1)^{k-1} \left[\sum_{j=1}^{\ell-1} \sum_J \frac{\partial g_{(J,\ell)}}{\partial x_j} dx_j \wedge dx_J + \sum_J \frac{\partial g_{(J,\ell)}}{\partial x_\ell} dx_\ell \wedge dx_J \right] = \end{aligned}$$

e dunque $\beta \in \mathcal{A}_{\ell-1}^k$: per ipotesi induttiva allora β è esatta ed esiste una $(k-1)$ -forma ξ tale che $\beta = d\xi$, pertanto $\alpha = \beta + d\eta = d\xi + d\eta = d(\xi + \eta)$, che prova l'esattezza di α . \square

Osservazione 7.9. Si noti che abbiamo usato che Ω è un n -cubo solo per definire le funzioni g_I , e più in generale la stessa dimostrazione vale se Ω è convesso; d'altra parte vedremo a breve che la tesi vale per tutti i domini diffeomorfi al cubo.

Definizione 7.10 (Pullback). Dati $A \subseteq \mathbb{R}^m$ e $B \subseteq \mathbb{R}^n$, sia $\Phi : A \rightarrow B$ un diffeomorfismo di classe C^1 e sia ω una k -forma su B data da

$$\omega = \sum_{i_1 < \dots < i_k} \omega_{i_1, \dots, i_k} dx_{i_1} \wedge \dots \wedge dx_{i_k} \quad :$$

chiamiamo il *pullback* di ω via Φ la k -forma $\Phi^* \omega$ su A definita da

$$\Phi^* \omega = \sum_{i_1 < \dots < i_k} (\omega_{i_1, \dots, i_k} \circ \Phi) d\Phi_{i_1} \wedge \dots \wedge d\Phi_{i_k} \quad .$$

Definizione 7.11 (Pushforward). Dati $A \subseteq \mathbb{R}^m$ e $B \subseteq \mathbb{R}^n$, sia $\Phi : A \rightarrow B$ un diffeomorfismo di classe C^1 : se ω è una k -forma su A , chiamiamo *pushforward* di ω via Φ la k -forma su B data da $\Phi_*\omega := (\Phi^{-1})^*\omega$.

Per $m = k$, possiamo scrivere

$$\Phi^*\omega = \sum_{i_1 < \dots < i_k} (\omega_{i_1, \dots, i_k} \circ \Phi) \frac{\partial(\Phi^{i_1}, \dots, \Phi^{i_k})}{\partial(y_1, \dots, y_k)} dy_1 \wedge \dots \wedge dy_k,$$

infatti fissato un multi-indice $I = (i_1, \dots, i_k)$, scelti $x \in A$ e $v_1, \dots, v_k \in \mathbb{R}^k$, denotiamo con $D\Phi_I(x)$ il minore di $D\Phi(x)$ dato dalle righe i_1, \dots, i_k e con $V \in M_k(\mathbb{R})$ la matrice data da $V^i = v_i$ e otteniamo che

$$\begin{aligned} d\Phi_{i_1}(x) \wedge \dots \wedge d\Phi_{i_k}(x)(v_1, \dots, v_k) &= dy_1 \wedge \dots \wedge dy_k(D\Phi_I(x)v_1, \dots, D\Phi_I(x)v_k) = \\ &= \det(D\Phi_I(x)V) = \det D\Phi_I(x) \det V = \\ &= \frac{\partial(\Phi^{i_1}, \dots, \Phi^{i_k})}{\partial(y_1, \dots, y_k)}(x) dy_1 \wedge \dots \wedge dy_k(v_1, \dots, v_k) \end{aligned}$$

come voluto. Per m generico invece vale

$$\Phi^*\omega = \sum_{i_1 < \dots < i_k} \sum_{j_1 < \dots < j_k} (\omega_{i_1, \dots, i_k} \circ \Phi) \frac{\partial(\Phi^{i_1}, \dots, \Phi^{i_k})}{\partial(y_{j_1}, \dots, y_{j_k})} dy_{j_1} \wedge \dots \wedge dy_{j_k}$$

dato che

$$\begin{aligned} d\Phi_{i_1} \wedge \dots \wedge d\Phi_{i_k} &= \left(\sum_{j_1=1}^n \frac{\partial\Phi_{i_1}}{\partial y_{j_1}} dy_{j_1} \right) \wedge \dots \wedge \left(\sum_{j_k=1}^n \frac{\partial\Phi_{i_k}}{\partial y_{j_k}} dy_{j_k} \right) = \\ &= \sum_{j_1, \dots, j_k=1}^n \frac{\partial\Phi_{i_1}}{\partial y_{j_1}} \dots \frac{\partial\Phi_{i_k}}{\partial y_{j_k}} dy_{j_1} \wedge \dots \wedge dy_{j_k} = \\ &= \sum_{j_1 < \dots < j_k} \sum_{\sigma \in S_k} \text{sgn}(\sigma) \frac{\partial\Phi_{i_1}}{\partial y_{j_{\sigma(1)}}} \dots \frac{\partial\Phi_{i_k}}{\partial y_{j_{\sigma(k)}}} dy_{j_1} \wedge \dots \wedge dy_{j_k} = \\ &= \sum_{j_1 < \dots < j_k} \frac{\partial(\Phi^{i_1}, \dots, \Phi^{i_k})}{\partial(y_{j_1}, \dots, y_{j_k})} dy_{j_1} \wedge \dots \wedge dy_{j_k}. \end{aligned}$$

Dalla definizione segue immediatamente che l'integrale di una forma differenziale α su una superficie $S = \Phi(A)$ coincide con l'integrale del suo pullback $\Phi^*\alpha$ su A . Vediamo adesso alcune proprietà del pullback.

Proposizione 7.10

Dati $A \subseteq \mathbb{R}^m$ e $B \subseteq \mathbb{R}^n$, sia $\Phi : A \rightarrow B$ un diffeomorfismo di classe C^1 e siano ω una k -forma e η una h -forma su B ; allora $\Phi^*(\omega \wedge \eta) = \Phi^*\omega \wedge \Phi^*\eta$.

Dimostrazione. Scriviamo esplicitamente

$$\omega = \sum_I \omega_I dx_{I_1} \wedge \dots \wedge dx_{I_k} \quad \text{e} \quad \eta = \sum_J \eta_J dx_{J_1} \wedge \dots \wedge dx_{J_h}$$

per cui

$$\begin{aligned}
\Phi^*(\omega \wedge \eta) &= \Phi^* \left(\sum_I \sum_J \omega_I \eta_J dx_{I_1} \wedge \cdots \wedge dx_{I_k} \wedge dx_{J_1} \wedge \cdots \wedge dx_{J_h} \right) = \\
&= \sum_I \sum_J (\omega_I \eta_J \circ \Phi) d\Phi_{I_1} \wedge \cdots \wedge d\Phi_{I_k} \wedge d\Phi_{J_1} \wedge \cdots \wedge d\Phi_{J_h} = \\
&= \left(\sum_I (\omega_I \circ \Phi) d\Phi_{I_1} \wedge \cdots \wedge d\Phi_{I_k} \right) \wedge \left(\sum_J (\eta_J \circ \Phi) d\Phi_{J_1} \wedge \cdots \wedge d\Phi_{J_h} \right) = \\
&= \Phi^*\omega \wedge \Phi^*\eta .
\end{aligned}$$

□

Proposizione 7.11

Dati $A \subseteq \mathbb{R}^m$ e $B \subseteq \mathbb{R}^n$, sia $\Phi : A \rightarrow B$ un diffeomorfismo di classe C^1 e sia ω una k -forma su B ; allora $\Phi^*(d\omega) = d(\Phi^*\omega)$.

Dimostrazione. Scriviamo in coordinate

$$\omega = \sum_{i_1 < \cdots < i_k} \omega_{i_1, \dots, i_k} dx_{i_1} \wedge \cdots \wedge dx_{i_k}$$

da cui

$$d\omega = \sum_{i_1 < \cdots < i_k} \sum_{j=1}^n \frac{\partial \omega_{i_1, \dots, i_k}}{\partial x_j} dx_j \wedge dx_{i_1} \wedge \cdots \wedge dx_{i_k} ;$$

adesso

$$\Phi^*\omega = \sum_{i_1 < \cdots < i_k} (\omega_{i_1, \dots, i_k} \circ \Phi) d\Phi_{i_1} \wedge \cdots \wedge d\Phi_{i_k}$$

e

$$\Phi^*(d\omega) = \sum_{i_1 < \cdots < i_k} \sum_{j=1}^n \left(\frac{\partial \omega_{i_1, \dots, i_k}}{\partial x_j} \circ \Phi \right) d\Phi_j \wedge d\Phi_{i_1} \wedge \cdots \wedge d\Phi_{i_k}$$

mentre

$$\begin{aligned}
d(\Phi^*\omega) &= \sum_{i_1 < \cdots < i_k} \sum_{\ell=1}^m \frac{\partial (\omega_{i_1, \dots, i_k} \circ \Phi)}{\partial y_\ell} dy_\ell \wedge d\Phi_{i_1} \wedge \cdots \wedge d\Phi_{i_k} = \\
&= \sum_{i_1 < \cdots < i_k} \sum_{\ell=1}^m \sum_{j=1}^n \left(\frac{\partial \omega_{i_1, \dots, i_k}}{\partial x_j} \circ \Phi \right) \frac{\partial \Phi_j}{\partial y_\ell} dy_\ell \wedge d\Phi_{i_1} \wedge \cdots \wedge d\Phi_{i_k} = \\
&= \sum_{i_1 < \cdots < i_k} \sum_{j=1}^n \left(\frac{\partial \omega_{i_1, \dots, i_k}}{\partial x_j} \circ \Phi \right) \left(\sum_{\ell=1}^m \frac{\partial \Phi_j}{\partial y_\ell} dy_\ell \right) \wedge d\Phi_{i_1} \wedge \cdots \wedge d\Phi_{i_k} = \\
&= \sum_{i_1 < \cdots < i_k} \sum_{j=1}^n \left(\frac{\partial \omega_{i_1, \dots, i_k}}{\partial x_j} \circ \Phi \right) d\Phi_j \wedge d\Phi_{i_1} \wedge \cdots \wedge d\Phi_{i_k} = \Phi^*(d\omega) .
\end{aligned}$$

□

Osservazione 7.12. Per ogni k -forma ω vale² $\Phi_*\Phi^*\omega = \omega$: infatti per la Proposizione 7.10 basta dimostrare che $\Phi_*\Phi^*(dx_i) = dx_i$, e d'altra parte

$$\begin{aligned}\Phi_*\Phi^*(dx_i) &= (\Phi^{-1})^*(d\Phi_i) = (\Phi^{-1})^*\left(\sum_{j=1}^m \frac{\partial\Phi_i}{\partial y_j} dy_j\right) = \sum_{j=1}^m \left(\frac{\partial\Phi_i}{\partial y_j} \circ \Phi^{-1}\right) d(\Phi^{-1})_j = \\ &= \sum_{j=1}^m \sum_{h=1}^n \left(\frac{\partial\Phi_i}{\partial y_j} \circ \Phi^{-1}\right) \frac{\partial(\Phi^{-1})_j}{\partial x_h} dx_h = \sum_{h=1}^n D\Phi_i(\Phi^{-1}(\cdot)) \left(\frac{\partial\Phi^{-1}}{\partial x_h}\right) dx_h = \\ &= \sum_{h=1}^n \delta_{ih} dx_h = dx_i\end{aligned}$$

come voluto.

Introdurre il pullback ci permette di estendere facilmente il risultato del Lemma di Poincaré ad una classe più ampia di domini. Il seguente Corollario mostra infatti che l'esattezza di una forma può essere controllata su un suo pullback.

Corollario 7.13

Le forme chiuse sono esatte su tutte le immagini dell' n -cubo via un diffeomorfismo C^1 .

Dimostrazione. Se $\Phi : [0, 1]^n \rightarrow \Omega$ è un diffeomorfismo C^1 e ω è una k -forma chiusa su Ω , allora $0 = \Phi^*(d\omega) = d(\Phi^*\omega)$, pertanto $\Phi^*\omega$ è chiusa sul cubo e dunque è esatta. Se allora $\Phi^*\omega = d\xi$, troviamo $\omega = \Phi_*(\Phi^*\omega) = \Phi_*(d\xi) = d(\Phi_*\xi)$. \square

7.3 Teorema di Stokes

Euristicamente, quello che il Teorema della divergenza suggerisce è che la variazione di una quantità in un insieme Ω è data dal flusso della quantità attraverso il bordo $\partial\Omega$. Cerchiamo di riformulare questa idea nel linguaggio delle forme.

Definizione 7.12 (k -catena). Una k -catena singolare è una combinazione lineare formale a coefficienti interi di k -cubi singolari.

Definizione 7.13 (Bordo). Dato $Q_k = [0, 1]^k$ il k -cubo standard, definiamo il suo bordo come $(k-1)$ -catena come

$$\begin{aligned}\partial Q_k &= \bigcup_{i=1}^k (-1)^{i-1} F_{d,i} + (-1)^i F_{s,i} \text{ , dove} \\ F_{d,i} &:= [0, 1]^{i-1} \times \{1\} \times [0, 1]^{k-i} \text{ e } F_{s,i} := [0, 1]^{i-1} \times \{0\} \times [0, 1]^{k-i} \text{ .}\end{aligned}$$

Consideriamo ora una $(k-1)$ -forma α su Q_k ,

$$\alpha = \sum_{i=1}^k \alpha_i dx_1 \wedge \cdots \wedge d\hat{x}_i \wedge \cdots \wedge dx_k$$

dove $d\hat{x}_i$ è il termine omesso; allora il suo differenziale esterno è

$$d\alpha = \sum_{i=1}^k \frac{\partial\alpha_i}{\partial x_i} dx_i \wedge dx_1 \wedge \cdots \wedge d\hat{x}_i \wedge \cdots \wedge dx_k = \left(\sum_{i=1}^k (-1)^{i-1} \frac{\partial\alpha_i}{\partial x_i} \right) dx_1 \wedge \cdots \wedge dx_k$$

²In effetti il pullback è funtoriale, dunque $\Phi_*\Phi^*\omega = (\Phi^{-1})^*\Phi^*\omega = (\Phi \circ \Phi^{-1})^*\omega = \omega$.

e di conseguenza

$$\begin{aligned} \int_{Q_k} d\alpha &= \sum_{i=1}^k (-1)^{i-1} \int_{Q_k} \frac{\partial \alpha_i}{\partial x_i} dx_1 \wedge \cdots \wedge dx_k = \sum_{i=1}^k (-1)^{i-1} \int_{Q_k} \frac{\partial \alpha_i}{\partial x_i}(x) dx_1 \cdots dx_k = \\ &= \sum_{i=1}^k (-1)^{i-1} \int_{x_1=0}^1 \cdots \int_{x_{i-1}=0}^1 \int_{x_{i+1}=0}^1 \cdots \int_{x_k=0}^1 \left(\int_{x_i=0}^1 \frac{\partial \alpha_i}{\partial x_i}(x) dx_i \right) dx_1 \cdots dx_{i-1} \cdots dx_k \end{aligned}$$

ora però

$$\int_{x_i=0}^1 \frac{\partial \alpha_i}{\partial x_i}(x) dx_i = \alpha_i(x_1, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_k) - \alpha_i(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_k)$$

pertanto

$$\begin{aligned} \int_{Q_k} d\alpha &= \sum_{i=1}^k (-1)^{i-1} \left(\int_{F_{d,i}} \alpha_i - \int_{F_{s,i}} \alpha_i \right) = \sum_{i=1}^k \left((-1)^{i-1} \int_{F_{d,i}} \alpha_i + (-1)^i \int_{F_{s,i}} \alpha_i \right) = \\ &= \int_{\partial Q_k} \alpha . \end{aligned}$$

Data adesso $\Phi : Q_k \rightarrow \mathbb{R}^n$ di classe C^2 con differenziale di rango massimo in ogni punto, si ha che $\Phi(Q_k)$ è una k -superficie in \mathbb{R}^n , Φ_* manda h -forme su Q_k in h -forme su $\Phi(Q_k)$ e valgono le relazioni

$$\int_{\Phi(Q_k)} \Phi_* \beta = \int_{Q_k} \beta \quad \text{e} \quad d(\Phi_* \beta) = \Phi_*(d\beta) .$$

Definizione 7.14 (Bordo di superficie). Data una mappa $\Phi : Q_k \rightarrow \mathbb{R}^n$ di classe C^2 con differenziale di rango massimo in ogni punto, definiamo *bordo* della k -superficie $\Phi(Q_k)$ come $(k-1)$ -catena, che denotiamo $\partial(\Phi(Q_k))$, l'insieme $\Phi(\partial Q_k)$ che è unione finita di $(k-1)$ -superfici con segno.

Teorema 7.14: Teorema di Stokes

Siano $S \subseteq \mathbb{R}^n$ una k -superficie di classe C^2 e ∂S il suo bordo come $(k-1)$ -catena, allora per ogni $(k-1)$ -forma α di classe C^2 si ha

$$\int_S d\alpha = \int_{\partial S} \alpha .$$

Dimostrazione. Abbiamo già osservato che la tesi vale per $S = Q_k$; d'altra parte se $S = \Phi(Q_k)$, prendendo β la $(k-1)$ -forma su Q_k data da $\beta = \Phi^* \alpha$ si ha

$$\begin{aligned} \int_S d\alpha &= \int_{\Phi(Q_k)} d(\Phi_* \beta) = \int_{\Phi(Q_k)} \Phi_*(d\beta) = \int_{Q_k} d\beta = \int_{\partial Q_k} \beta = \int_{\Phi(\partial Q_k)} \Phi_* \beta = \\ &= \int_{\partial \Phi(Q_k)} \Phi_* \beta = \int_{\partial S} \alpha . \end{aligned}$$

□

Abbiamo dimostrato come conseguenza del Lemma di Poincaré una condizione sufficiente per l'esattezza delle forme chiuse su un dominio Ω , ossia che Ω sia diffeomorfo ad un qualche cubo n -dimensionale. Per quanto riguarda le 1-forme invece, abbiamo provato che la chiusura equivale all'esattezza se Ω è semplicemente connesso: mostriamo che tale risultato non può essere esteso alle k -forme. Prendiamo $\Omega = \mathbb{R}^3 \setminus \{0\}$, che è semplicemente connesso, e consideriamo la 2-forma

$$\omega = \frac{1}{x^2 + y^2 + z^2} (x dy \wedge dz + y dz \wedge dx + z dx \wedge dy) :$$

un facile conto mostra che ω è chiusa; se per assurdo fosse esatta, ovvero se esistesse una 1-forma ξ tale che $\omega = d\xi$, potremmo considerare la 2-superficie $S^2 \subseteq \Omega$, e per Stokes varrebbe

$$\int_{S^2} \omega = \int_{S^2} d\xi = \int_{\partial S^2} \xi = 0$$

dato che $\partial S^2 = \emptyset$; d'altra parte possiamo parametrizzare S^2 in coordinate sferiche con la mappa $\Phi : [0, 2\pi) \times [0, \pi)$ data da $\Phi(\theta, \phi) = (\sin \phi \cos \theta, \sin \phi \sin \theta, \cos \phi)$, per cui possiamo scrivere

$$D\Phi(\theta, \phi) = \begin{pmatrix} -\sin \phi \sin \theta & \cos \phi \cos \theta \\ \sin \phi \cos \theta & \cos \phi \sin \theta \\ 0 & -\sin \phi \end{pmatrix}$$

e dunque

$$\begin{aligned} \Phi^* \omega &= \sin \phi \cos \theta [-\sin^2 \phi \cos \theta d\theta \wedge d\phi] + \sin \phi \sin \theta [\sin^2 \phi \sin \theta d\phi \wedge d\theta] + \\ &\quad + \cos \phi [-\sin \phi \cos \phi \sin^2 \theta d\theta \wedge d\phi + \sin \phi \cos \phi \cos^2 \theta d\phi \wedge d\theta] = \\ &= \sin \phi [\sin^2 \phi \cos^2 \theta + \sin^2 \phi \sin^2 \theta + \cos^2 \phi \sin^2 \theta + \cos^2 \phi \cos^2 \theta] d\phi \wedge d\theta = \\ &= \sin \phi d\phi \wedge d\theta \end{aligned}$$

pertanto

$$\int_{S^2} \omega = \int_{\Phi^{-1}(S^2)} \Phi^* \omega = \int_0^{2\pi} \int_0^\pi \sin \phi d\phi d\theta = 4\pi \neq 0 \quad \text{!}$$

7.4 Calcolo di integrali di forme su superfici

Siano α, β parametri reali, $0 \leq \alpha < \beta \leq 2\pi$, e sia $f \in C^1([\alpha, \beta], \mathbb{R}^+)$.

Consideriamo il cammino $\gamma : [\alpha, \beta] \rightarrow \mathbb{R}^2$ dato da

$$\gamma : \theta \mapsto \begin{pmatrix} x(\theta) \\ y(\theta) \end{pmatrix} = \begin{pmatrix} f(\theta) \cos \theta \\ f(\theta) \sin \theta \end{pmatrix},$$

allora la mappa γ delimita naturalmente una regione S .

Definiamo la forma ausiliaria $\omega = \frac{1}{2}(xdy - ydx)$, allora

$$\int_\gamma \omega = \text{area}(S),$$

infatti

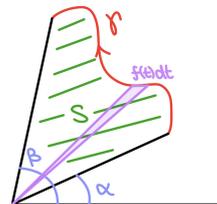
$$\begin{aligned} \int_\gamma \omega &= \frac{1}{2} \int_\alpha^\beta (x(\theta)y'(\theta) - y(\theta)x'(\theta))d\theta = \\ &= \frac{1}{2} \int_\alpha^\beta (f(\theta) \cos \theta (f'(\theta) \sin \theta + f(\theta) \cos \theta) - f(\theta) \sin \theta (f'(\theta) \cos \theta - f(\theta) \sin \theta))d\theta = \\ &= \frac{1}{2} \int_\alpha^\beta (f^2(\theta) \cos^2 \theta + f^2(\theta) \sin^2 \theta)d\theta = \frac{1}{2} \int_\alpha^\beta f^2(\theta)d\theta \end{aligned}$$

e d'altra parte

$$\text{area}(S) = \iint_S dx dy = \int_\alpha^\beta \int_0^{f(\theta)} \rho d\rho d\theta = \frac{1}{2} \int_\alpha^\beta f^2(\theta)d\theta.$$

Consieriamo adesso una curva S (1-superficie) in \mathbb{R}^2 che può essere parametrizzata con un cammino $\gamma : [a, b] \rightarrow \mathbb{R}^2$ di classe C^1 , allora la direzione tangente e la normale uscente a S in punto $\gamma(t)$ possono essere scritte come

$$T_{\gamma(t)} = \frac{\gamma'(t)}{|\gamma'(t)|} \quad \text{e} \quad N_{\gamma(t)} = \frac{1}{|\gamma'(t)|} \begin{pmatrix} \gamma'_2(t) \\ -\gamma'_1(t) \end{pmatrix},$$



dunque se $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ è un campo C^1

$$\int_{\gamma} \vec{F} \cdot \vec{n} = \int_a^b F(\gamma(t)) \cdot N_{\gamma(t)} |\gamma'(t)| dt = \int_a^b (F_1(\gamma(t))\gamma_2'(t) - F_2(\gamma(t))\gamma_1'(t)) dt = \int_{\gamma} (F_1 dy - F_2 dx)$$

ed il calcolo del flusso di F lungo la curva S può essere ricondotto al calcolo dell'integrale di una 1-forma su una 1-superficie.

Consideriamo invece una 2-superficie S in \mathbb{R}^3 ed un campo $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ di classe C^1 ; mostriamo che vale la formula

$$\iint_S \vec{F} \cdot \vec{n} = \iint_S (F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2)$$

e che dunque anche in questo caso il calcolo del flusso di un campo F si riconduce al calcolo dell'integrale della forma $\omega = F_1 dx_2 \wedge dx_3 + F_2 dx_3 \wedge dx_1 + F_3 dx_1 \wedge dx_2$.

Prendiamo $\Phi : D \subseteq \mathbb{R}^2 \rightarrow S$ una parametrizzazione C^1 di S , allora

$$\begin{aligned} \iint_S \vec{F} \cdot \vec{n} &= \iint_D (F \circ \Phi) \cdot (n \circ \Phi) \sqrt{\det D\Phi^T D\Phi} dudv = \\ &= \iint_D (F \circ \Phi) \cdot \left(\frac{\Phi_u \times \Phi_v}{|\Phi_u \times \Phi_v|} \right) |\Phi_u \times \Phi_v| dudv = \iint_D (F \circ \Phi) \cdot (\Phi_u \times \Phi_v) dudv = \\ &= \iint_D \left(F_1 \circ \Phi (\Phi_u^2 \Phi_v^3 - \Phi_u^3 \Phi_v^2) + F_2 \circ \Phi (\Phi_u^3 \Phi_v^1 - \Phi_u^1 \Phi_v^3) + F_3 \circ \Phi (\Phi_u^1 \Phi_v^2 - \Phi_u^2 \Phi_v^1) \right) dudv = \\ &= \iint_D \left(F_1 \circ \Phi \frac{\partial(\Phi^2, \Phi^3)}{\partial(u, v)} + F_2 \circ \Phi \frac{\partial(\Phi^3, \Phi^1)}{\partial(u, v)} + F_3 \circ \Phi \frac{\partial(\Phi^1, \Phi^2)}{\partial(u, v)} \right) dudv = \iint_S \omega . \end{aligned}$$

Osserviamo inoltre che $d\omega = \operatorname{div} F dx_1 \wedge dx_2 \wedge dx_3$, pertanto se $S = \partial\Omega$ con $\Omega \subseteq \mathbb{R}^3$ aperto

$$\iiint_{\Omega} d\omega = \iiint_{\Omega} \operatorname{div} F = \iint_{\partial\Omega} F \cdot n = \iint_{\partial\Omega} \omega$$

ed in questo caso particolare ritroviamo l'enunciato di Stokes attraverso il Teorema della divergenza.

Esercizio 11. Dato un campo differenziabile $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ ed un'ipersuperficie Σ , si trovi un'opportuna $(n-1)$ -forma ω tale che

$$\int_{\Sigma} \vec{F} \cdot \vec{n} = \int_{\Sigma} \omega$$

e si deduca la tesi di Stokes nel caso $k = n$ attraverso il Teorema della divergenza.

Idea di Soluzione. È ragionevole tentare di provare l'uguaglianza per

$$\omega = \sum_{i=1}^n (-1)^{i+1} F_i dx_1 \wedge \cdots \wedge d\hat{x}_i \wedge \cdots \wedge dx_n$$

così che se Ω è un aperto e $\Sigma = \partial\Omega$, denotando con $\operatorname{vol}_{\Omega}$ la n -forma $dx_1 \wedge \cdots \wedge dx_n$, si ha $d\omega = \operatorname{div} F \operatorname{vol}_{\Omega}$ e dunque per il Teorema della divergenza

$$\int_{\Omega} d\omega = \int_{\Omega} \operatorname{div} F = \int_{\partial\Omega} F \cdot n = \int_{\partial\Omega} \omega .$$

Ora se $D \subseteq \mathbb{R}^{n-1}$ e $\Phi : D \rightarrow \Sigma$ è un diffeomorfismo C^1 che parametrizza Σ , si ha che

$$\int_{\Sigma} \omega = \int_D \omega(\Phi(z))(D\Phi(z)) dz = \int_D \det(F(\Phi(z)); D\Phi(z)) dz ;$$

ora sulla superficie Σ il campo F si scrive come $(F \cdot n)n + T$ dove, per ogni $x \in \Sigma$, $T(x)$ è un vettore nel tangente $T_x\Sigma$, di conseguenza $\det(T(\Phi(z)); D\Phi(z)) = 0$ per ogni $z \in D$ e dunque

$$\int_{\Sigma} \omega = \int_D F(\Phi(z)) \cdot n(\Phi(z)) \det(n(\Phi(z)); D\Phi(z)) dz .$$

A questo punto si dimostra che se $v_1, \dots, v_{n-1} \in \mathbb{R}^n$ sono linearmente indipendenti e generano l'iperpiano V , detta ν_V la normale uscente da V e $A \in M(n, n-1, \mathbb{R})$ la matrice data da $A^i = v_i$, allora

$$\det(\nu_V, v_1, \dots, v_{n-1}) = \sqrt{\det A^\top A}$$

in quanto entrambi i termini equivalgono all' $(n-1)$ -volume dell' $(n-1)$ -parallelepipedo generato dai v_i ; dunque $\det(n(\Phi(z)); D\Phi(z)) = \sqrt{\det D\Phi(z)^\top D\Phi(z)}$ e

$$\int_{\Sigma} \omega = \int_D F(\Phi(z)) \cdot n(\Phi(z)) \sqrt{\det D\Phi(z)^\top D\Phi(z)} dz = \int_{\Sigma} F \cdot n \ .$$

Appendice A

Passaggio di derivata sotto il segno di integrale

Consideriamo uno spazio di misura Ω e una funzione $f : (a, b) \times \Omega \rightarrow \mathbb{R}$, e consideriamo la funzione $G : (a, b) \rightarrow \mathbb{R}$ data da

$$G(x) = \int_{\Omega} f(x, y) dy .$$

In questa sezione cerchiamo di capire se e quando è vero che G è continua e che

$$G'(x) = \int_{\Omega} \frac{\partial f}{\partial x}(x, y) dy .$$

Per la continuità, troviamo una risposta immediata:

Teorema 1.1

Sia $f : I \times \mathbb{R}^n \rightarrow \mathbb{R}$, con $I \subseteq \mathbb{R}$ misurabile, una funzione tale che $f(\cdot, y)$ è continua per quasi ogni $y \in \mathbb{R}^n$. Se per ogni $x \in I$ esistono $\delta > 0$ ed una funzione integrabile $g : \mathbb{R}^n \rightarrow \mathbb{R}$ tale che $|f(t, y)| \leq g(y)$ per ogni $t \in [x - \delta, x + \delta]$, allora la funzione

$$G(x) = \int_{\mathbb{R}^n} f(x, y) dy$$

è continua in x .

Dimostrazione. Fissato $x \in I$, vogliamo mostrare

$$\lim_{x \rightarrow 0} G(x + h) - G(x) = 0,$$

che è equivalente alla condizione

$$\forall \{h_n\} \subseteq [-\delta, \delta] \text{ tale che } h_n \xrightarrow{n \rightarrow +\infty} 0 \text{ si ha } \lim_{n \rightarrow +\infty} G(x + h_n) - G(x) = 0 .$$

Vale

$$\left| \int_{\mathbb{R}^n} f(x + h_n, y) - f(x, y) dy \right| \leq \int_{\mathbb{R}^n} |f(x + h_n, y) - f(x, y)| dy .$$

Ora, poiché $f(\cdot, y)$ è continua per quasi ogni $y \in \mathbb{R}^n$

$$f(x + h_n, y) - f(x, y) \xrightarrow{n \rightarrow +\infty} 0 .$$

Inoltre, per ipotesi, vale $|f(x + h_n, y) - f(x, y)| \leq 2g(y)$, quindi, per il Teorema di convergenza dominata, otteniamo

$$\int_{\mathbb{R}^n} |f(x + h_n, y) - f(x, y)| dy \xrightarrow{n \rightarrow +\infty} 0 .$$

□

Riguardo la seconda domanda, iniziamo ad evidenziare alcune condizioni necessarie. Sicuramente per ogni x la mappa $y \mapsto f(x, y)$ deve essere integrabile per la buona definizione G ; inoltre per poter scrivere il termine di destra nella seconda uguaglianza è necessario che f sia derivabile rispetto alla x e che la sua derivata $\partial_x f$ sia integrabile nella y . Supponiamo dunque tutte queste ipotesi.

Fissiamo ε e scriviamo

$$\frac{G(x + \varepsilon) - G(x)}{\varepsilon} = \frac{1}{\varepsilon} \int_{\Omega} (f(x + \varepsilon, y) - f(x, y)) dy = \int_{\Omega} \frac{\partial f}{\partial x}(\xi_y, y) dy$$

con $\xi_y \in (x, x + \varepsilon)$. La dimostrazione adesso potrebbe procedere così: supponiamo che la funzione $t \mapsto \partial_x f(t, y)$ sia continua nella variabile $t = x$ per ogni $y \in \Omega$, allora

$$\int_{\Omega} \frac{\partial f}{\partial x}(x, y) dy = \int_{\Omega} \lim_{\varepsilon \rightarrow 0} \frac{\partial f}{\partial x}(\xi_y, y) dy \stackrel{*}{=} \lim_{\varepsilon \rightarrow 0} \int_{\Omega} \frac{\partial f}{\partial x}(\xi_y, y) dy = \lim_{\varepsilon \rightarrow 0} \frac{G(x + \varepsilon) - G(x)}{\varepsilon} = G'(x) .$$

Resta allora da provare l'uguaglianza (*), non vera in generale, ma a cui possiamo approssiarci con risultati di convergenza. Ad esempio, se $\partial_x f$ è continua su $(a, b) \times \Omega$ e Ω è compatto, dato $\delta > 0$ tale che $[x - \delta, x + \delta] \subseteq (a, b)$, $\partial_x f$ è limitata su $[x - \delta, x + \delta] \times \Omega$, pertanto se Ω ha misura finita, l'uguaglianza vale per convergenza dominata.

Abbiamo così dimostrato il seguente.

Teorema 1.2

Siano Ω spazio di misura compatto e di misura finita e $f : (a, b) \times \Omega \rightarrow \mathbb{R}$ tale che $\forall x \in (a, b)$ la mappa $y \mapsto f(x, y)$ è integrabile, e sia $G : (a, b) \rightarrow \mathbb{R}$ data da

$$G(x) = \int_{\Omega} f(x, y) dy ;$$

se $\exists \delta > 0$ tale che $[x - \delta, x + \delta] \subseteq (a, b)$, $\partial_x f$ esiste su $[x - \delta, x + \delta] \times \Omega$, è limitata e $\forall y \in \Omega$ è continua nella prima variabile allora G è derivabile in x e

$$G'(x) = \int_{\Omega} \frac{\partial f}{\partial x}(x, y) dy .$$

Osserviamo che abbiamo usato la continuità in entrambe le variabili e la compattezza di Ω per ottenere la limitatezza, dunque supponendo $\partial_x f$ localmente limitata possiamo rilassare le altre ipotesi.

Teorema 1.3

Siano Ω spazio di misura di misura finita e $f : (a, b) \times \Omega \rightarrow \mathbb{R}$ tale che $\forall x \in (a, b)$ la mappa $y \mapsto f(x, y)$ è integrabile, e sia $G : (a, b) \rightarrow \mathbb{R}$ data da

$$G(x) = \int_{\Omega} f(x, y) dy ;$$

se $\exists \delta > 0$ tale che $[x - \delta, x + \delta] \subseteq (a, b)$, $\partial_x f$ esiste su $[x - \delta, x + \delta] \times \Omega$ ed è continua allora G è derivabile in x e

$$G'(x) = \int_{\Omega} \frac{\partial f}{\partial x}(x, y) dy .$$

Inoltre abbiamo usato la limitatezza per dominare la derivata $\partial_x f$ con una costante, che è integrabile se Ω ha misura finita, pertanto se disponiamo già di una dominazione integrabile lo spazio di misura Ω può essere generico.

Teorema 1.4

Siano Ω spazio di misura e $f : (a, b) \times \Omega \rightarrow \mathbb{R}$ tale che $\forall x \in (a, b)$ la mappa $y \mapsto f(x, y)$ è integrabile, e sia $G : (a, b) \rightarrow \mathbb{R}$ data da

$$G(x) = \int_{\Omega} f(x, y) dy ;$$

se $\exists \delta > 0$ tale che $[x - \delta, x + \delta] \subseteq (a, b)$, $\partial_x f$ esiste su $[x - \delta, x + \delta] \times \Omega$, è continua nella prima variabile ed esiste $\psi : \Omega \rightarrow \mathbb{R}$ integrabile tale che

$$\left| \frac{\partial f}{\partial x}(t, y) \right| \leq \psi(y) \quad \text{per } t \in [x - \delta, x + \delta]$$

allora G è derivabile in x e

$$G'(x) = \int_{\Omega} \frac{\partial f}{\partial x}(x, y) dy .$$

Altrimenti, possiamo approssimarci all'uguaglianza (*) attraverso il Teorema di convergenza monotona, infatti se per ogni y la funzione $t \mapsto f(t, y)$ è convessa (o concava), allora la mappa $\partial_x f$ è crescente (o decrescente) e dunque si ha convergenza monotona. Possiamo riassumere nel risultato seguente.

Teorema 1.5

Siano Ω spazio di misura e $f : (a, b) \times \Omega \rightarrow \mathbb{R}$ tale che $\forall x \in (a, b)$ la mappa $y \mapsto f(x, y)$ è integrabile, e sia $G : (a, b) \rightarrow \mathbb{R}$ data da

$$G(x) = \int_{\Omega} f(x, y) dy ;$$

supponiamo che $\exists \delta > 0$ tale che $[x - \delta, x + \delta] \subseteq (a, b)$, che $\partial_x f$ esiste su $[x - \delta, x + \delta] \times \Omega$ e che $\forall y \in \Omega$ è continua nella prima variabile; se $\Omega = \Omega^0 + \Omega^+ + \Omega^-$, dove Ω^0 ha misura nulla e la mappa $t \mapsto f(t, y)$ è convessa su Ω^+ e concava su Ω^- , allora G è derivabile in x e

$$G'(x) = \int_{\Omega} \frac{\partial f}{\partial x}(x, y) dy .$$

Supponiamo adesso $\Omega \subseteq \mathbb{R}^n$ e che per ogni $x \in (a, b)$ e per ogni $\delta > 0$ la mappa $\partial_x f$ è integrabile con integrale finito. Allora

$$\begin{aligned} G(t) - G(s) &= \int_{\Omega} f(t, y) dy - \int_{\Omega} f(s, y) dy = \int_{\Omega} (f(t, y) - f(s, y)) dy = \\ &= \int_{\Omega} \left(\int_s^t \frac{\partial f}{\partial x}(\sigma, y) d\sigma \right) dy = \int_s^t \left(\int_{\Omega} \frac{\partial f}{\partial x}(\sigma, y) dy \right) d\sigma \end{aligned}$$

ossia se poniamo

$$H(\sigma) := \int_{\Omega} \frac{\partial f}{\partial x}(\sigma, y) dy$$

troviamo che H è ben definita e integrabile per il Teorema di Fubini, a meno di insiemi di misura finita.

$$G(t) - G(s) = \int_s^t H(\sigma) d\sigma \quad \text{e quindi} \quad G'(x) = H(x) = \int_{\Omega} \frac{\partial f}{\partial x}(x, y) dy$$

almeno per ogni x in cui H è continua¹.

¹Questa condizione è sufficiente: in realtà ciò che importa è che l'integrale medio di $\partial_x f$ converga ad $H(x)$ vicino ad x .

In generale, data una funzione $\phi : \mathbb{R} \rightarrow \mathbb{R}$, abbiamo chiamato *derivata* di ϕ una funzione ϕ' tale che

$$\lim_{\varepsilon \rightarrow 0} \frac{\phi(x + \varepsilon) - \phi(x)}{\varepsilon} \quad \forall x \in \mathbb{R} ,$$

d'altra parte, il Teorema fondamentale del calcolo dice che dati $a < b$ si ha

$$\phi(b) - \phi(a) = \int_a^b \phi'(t) dt ;$$

chiameremo *derivata debole* una qualsiasi mappa $\tau : \mathbb{R} \rightarrow \mathbb{R}$ tale che

$$\phi(b) - \phi(a) = \int_a^b \tau(t) dt ,$$

dunque $H = G'$ in senso debole.