Analisi 2

Stefano Scrobogna

26 febbraio 2024

Indice

1 Integrali generalizzati

Definizione 1.1 (Funzione integrabile secondo Riemann)**.** Sia $f : I := [a,b] \subset \mathbb{R} \to \mathbb{R}$ limitata, $J \in \mathbb{N}$ e consideriamo la *partizione* di I

$$I = \bigsqcup_{j=1}^{J} I_j, \qquad\qquad I_j := \left[a + \frac{b-a}{J}(j-1)\ ,\ a + \frac{b-a}{J}\, j\right),$$

consideriamo le *somme parziali superiori ed inferiori* definite come

$$\Sigma_J := \sum_{j=1}^{J} \left|I_j\right| \sup_{x \in I_j} f(x) \qquad\qquad \sigma_J := \sum_{j=1}^{J} \left|I_j\right| \inf_{x \in I_j} f(x)$$

Se esiste, finito, il limite

$$\lim_{J\to\infty} \sigma_J = \lim_{J\to\infty} \Sigma_J = \mathcal{I},$$

allora f è R-integrabile in $[a,b)$ e sia ha che

$$\int_a^b f(x)\,\mathrm{d}x := \mathcal{I}.$$

Definizione 1.2 (Funzioni Riemann integrabili)**.** Ricordiamo che dato un insieme chiuso e limitato $K \subset \mathbb{R}$ (che possiamo assumere essere un intervallo senza perdita di generalità) definiamo con

$$\mathcal{R}(K;\mathbb{R}) := \left\{ f : K \to \mathbb{R} \;\middle|\; f \text{ Riemann-integrabile in } K \right\}.$$

Osservazione 1.3.
- Ricordiamo che nella teoria dell'integrale secondo Riemann una funzione integrabile è automaticamente limitata;
- È immediato che $\mathcal{C}(K;\mathbb{R}) \subset \mathcal{R}(K;\mathbb{R})$;
- L'esempio canonico di funzione non integrabile è la funzione di Dirichelet $\mathbb{1}_{\mathbb{Q}}$.

Definizione 1.4 (Funzioni integrabili in senso generalizzato in domini illimitati)**.** Consideriamo una funzione $f : [a,+\infty) \to \mathbb{R}$ appartenente allo spazio

$$f \in \bigcap_{c>a} \mathcal{R}([a,c];\mathbb{R}),$$

Se esiste finito il limite

$$\lim_{c\to\infty} \int_a^c f(x)\,\mathrm{d}x$$

allora si dice che f è *integrabile in senso improprio* (o *in senso generalizzato*) in $[a,+\infty)$ o che f è a integrale convergente, e si definisce

$$\int_a^{+\infty} f(x)\,\mathrm{d}x := \lim_{c\to\infty} \int_a^c f(x)\,\mathrm{d}x.$$

In tal caso scriviamo

$$f \in \tilde{\mathcal{R}}([a,+\infty);\mathbb{R}).$$

Data $g : (-\infty,b] \to \mathbb{R}$, tale che

$$g \in \bigcap_{c<b} \mathcal{R}([c,b];\mathbb{R}),$$

se esiste finito il limite

$$\lim_{c\to-\infty} \int_c^b f(x)\,\mathrm{d}x$$

allora si dice che g è integrabile in senso improprio (o in senso generalizzato) in $(-\infty,b]$ o che f è a integrale convergente, e si definisce

$$\int_{-\infty}^b g(x)\,\mathrm{d}x := \lim_{c\to-\infty} \int_c^b g(x)\,\mathrm{d}x,$$

e scriviamo

$$g \in \tilde{\mathcal{R}}((-\infty, b];\mathbb{R}).$$

Data una funzione $h : \mathbb{R} \to \mathbb{R}$ integrabile in ogni intervallo, si dice che h è integrabile in senso improprio (o in senso generalizzato) in $\mathbb{R}$ se esiste $a \in \mathbb{R}$ tale che

$$h \in \tilde{\mathcal{R}}((-\infty, a];\mathbb{R}) \cap \tilde{\mathcal{R}}([a, +\infty);\mathbb{R}) =: \tilde{\mathcal{R}}(\mathbb{R};\mathbb{R}).$$

Osservazione 1.5. Quest'ultima definizione non dipende dalla scelta di $a \in \mathbb{R}$, come si verifica facilmente utilizzando la proprietà di additività dell'integrale. Inoltre, se f è integrabile in $\mathbb{R}$, allora

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = \lim_{c\to+\infty} \int_{-c}^{c} f(x)\,\mathrm{d}x. \tag{1.1}$$

Osservazione 1.6. Non è vero che la condizione nel lato destro di (1.1) implica che f sia integrabile in $\mathbb{R}$; basta considerare la funzione $f(x) = x$ o $f(x) = \sin x$.

Esempio 1.7. Dato $\alpha > 0$, la funzione $f(x) = x^{-\alpha}$ è integrabile in $[1, +\infty)$ se e solo se $\alpha > 1$. Infatti

$$\int_1^c x^{-\alpha}\mathrm{d}x = \begin{cases} \frac{c^{1-\alpha}-1}{1-\alpha}, & \text{se } \alpha \neq 1 \\ \log c, & \text{se } \alpha = 1 \end{cases}$$

da cui

$$\lim_{c\to+\infty} \int_1^c x^{-\alpha}\mathrm{d}x = \begin{cases} +\infty, & \text{se } \alpha \leq 1 \\ \frac{1}{\alpha-1}, & \text{se } \alpha > 1 \end{cases}$$

Definizione 1.8 (Funzioni illimitate integrabili in senso generalizzato)**.** Si consideri una funzione $f : [a, b) \to \mathbb{R}$, $f \in \bigcap_{c\in(a,b)} \mathcal{R}([a, c];\mathbb{R})$. Se esiste finito il limite

$$\lim_{c\to b^-} \int_a^c f(x)\,\mathrm{d}x$$

allora si dice che f è integrabile in senso improprio (o in senso generalizzato) in $[a, b)$ o che $f \in \tilde{\mathcal{R}}([a, b);\mathbb{R})$, e si definisce

$$\int_a^b f(x)\,\mathrm{d}x := \lim_{c\to b^-} \int_a^c f(x)\,\mathrm{d}x$$

Analogamente possiamo definire funzioni appartenenti allo spazio $\tilde{\mathcal{R}}((a, b];\mathbb{R})$. Dato $x_0 \in \mathring{I}$ definiamo lo spazio $\tilde{\mathcal{R}}([a, b] \setminus \{x_0\};\mathbb{R}) := \tilde{\mathcal{R}}([a, x_0);\mathbb{R}) \cap \tilde{\mathcal{R}}((x_0, b];\mathbb{R})$.

Osservazione 1.9. Le definizioni precedenti sono utili quando f non sia estendibile con continuità all'intervallo chiuso $[a, b]$, dato che in tal caso l'integrale $\int_a^b f(x)\,\mathrm{d}x$ è già definito dalla teoria dell'integrazione; osserviamo che, per la continuità della funzione integrale, in questo caso integrale improprio e integrale usuale coincidono. La definizione si applica principalmente al caso di funzioni illimitate in un intorno di uno degli estremi dell'intervallo.

Esempio 1.10. Dato $\alpha > 0$, la funzione $f(x) = x^{-\alpha}$ è integrabile in $(0, 1]$ se e solo se $\alpha < 1$. Infatti

$$\int_c^1 x^{-\alpha}\mathrm{d}x = \begin{cases} \frac{1-c^{1-\alpha}}{1-\alpha}, & \text{se } \alpha \neq 1 \\ -\log c, & \text{se } \alpha = 1 \end{cases}$$

da cui

$$\lim_{c\to0^+} \int_c^1 x^{-\alpha}\mathrm{d}x = \begin{cases} +\infty, & \text{se } \alpha \geq 1 \\ \frac{1}{1-\alpha}, & \text{se } \alpha < 1 \end{cases}$$

Osservazione 1.11. Notiamo che la funzione $x \mapsto x^{-1}$ *non* è integrabile in $[-1, 0) \cap (0, 1]$, tuttavia

$$\int_{-1}^{-c} \frac{\mathrm{d}x}{x} + \int_c^1 \frac{\mathrm{d}x}{x} = 0 \qquad \forall c \in (0, 1).$$

Osservazione 1.12. Dalle definizioni date seguono facilmente le proprietà di linearità e di monotonia per gli integrali impropri. Non è invece detto che il prodotto di funzioni integrabili in senso improprio sia una funzione integrabile in senso improprio: $f(x) = \frac{1}{\sqrt{x}}$ è integrabile in senso improprio in $(0,1]$ma $(f \cdot f)(x) = \frac{1}{x}$ non lo è.

Quando i limiti introdotti in tutte le definizioni precedenti esistono e sono infiniti, si parla di integrali divergenti. Ad esempio, data $f \in \bigcap_{c>a} \mathcal{R}([a,c];\mathbb{R})$, se

$$\lim_{c\to\infty} \int_a^c f(x)\,\mathrm{d}x = +\infty \quad \text{o} \quad \lim_{c\to\infty} \int_a^c f(x)\,\mathrm{d}x = -\infty$$

si dice che f è a *integrale divergente*, e si definisce rispettivamente

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = +\infty \quad \text{o} \quad \int_a^{+\infty} f(x)\,\mathrm{d}x = -\infty$$

Da quanto visto, si ha allora che la funzione $f(x) = x^{-\alpha}$ in $[1,+\infty)$ è a integrale convergente se $\alpha > 1$ e a integrale divergente se $\alpha \leq 1$; $f(x) = x^{-\alpha}$ in $(0,1]$ è a integrale convergente se $\alpha < 1$ e a integrale divergente se $\alpha \geq 1$.

Proposizione 1.13 (Criterio del confronto)**.** *Sia $a \in \mathbb{R}$ e siano $f,g \in \bigcap_{b>a} \mathcal{R}([a,b];\mathbb{R})$, e supponiamo esiste $x_0 \geq a$ tale che $0 \leq f(x) \leq g(x)$, per ogni $x \geq x_0$. Se $g \in \tilde{\mathcal{R}}([a,\infty);\mathbb{R})$ allora $f \in \tilde{\mathcal{R}}([a,\infty);\mathbb{R})$. Viceversa, se f è a integrale divergente in $[a,+\infty)$, allora lo è anche g.*

Dimostrazione. Per l'additività dell'integrale è sufficiente provare la tesi sulla semiretta $[x_0,+\infty)$. La funzione $F(c) = \int_{x_0}^c f(x)\,\mathrm{d}x$ è crescente, per la positività di f e la monotonia dell'integrale. Dunque esiste $\lim_{c\to+\infty} \int_{x_0}^c f(x)\,\mathrm{d}x = \sup_{c\geq x_0} \int_{x_0}^c f(x)\,\mathrm{d}x$, quindi la funzione f è a integrale convergente o divergente a seconda che il limite sia finito o $+\infty$. Lo stesso vale per $G(c) = \int_{x_0}^c g(x)\,\mathrm{d}x$.

Se g è integrabile in $[a,+\infty)$, allora per ogni $c \geq x_0$, si ha

$$\int_{x_0}^c f(x)\,\mathrm{d}x \leq \int_{x_0}^c g(x)\,\mathrm{d}x \leq \int_{x_0}^{+\infty} g(x)\,\mathrm{d}x \in \mathbb{R}$$

Allora $F(c)$ è superiormente limitata e quindi $\lim_{c\to+\infty} \int_{x_0}^c f(x)\,\mathrm{d}x$ è finito. Sia ora f a integrale divergente. Si ha che $\lim_{c\to+\infty} \int_{x_0}^c f(x)\,\mathrm{d}x = +\infty$. Dato che $\int_{x_0}^c g(x)\,\mathrm{d}x \geq \int_{x_0}^c f(x)\,\mathrm{d}x$, si ha $\lim_{c\to+\infty} \int_{x_0}^c g(x)\,\mathrm{d}x = +\infty$. $\square$

Osservazione 1.14. Dagli argomenti usati nella dimostrazione della precedente proposizione segue che se una funzione f ha segno costante, allora esiste il $\lim_{c\to\infty} \int_a^c f(x)\,\mathrm{d}x$ e quindi f è a integrale convergente o a integrale divergente.

Proposizione 1.15. *Sia $f : [a,+\infty) \to \mathbb{R}$ integrabile in $[a,b]$ per ogni $b > a$. Se $|f|$ è integrabile in $[a,+\infty)$, allora lo è anche f e*

$$\left| \int_a^{+\infty} f(x)\ \mathrm{d}x \right| \leq \int_a^{+\infty} |f(x)|\,\mathrm{d}x$$

Dimostrazione. Si possono rappresentare le funzioni f e $|f|$ in funzione delle funzioni $f_+ = \max\{f,0\}$ e $f_- = -\min\{f,0\}$. Precisamente, si ha $f = f_+ - f_-$ e $|f| = f_+ + f_-$. Dato che $0 \leq f_+ \leq |f|$ e $0 \leq f_- \leq |f|$, dalla Proposizione 1.13 segue che f_+ e f_- sono integrabili in $[a,+\infty)$, dunque lo è anche f che si esprime come loro combinazione lineare. La disuguaglianza segue al solito dalla monotonia dell'integrale. $\square$

Alla luce della Proposizione 1.15, la Proposizione 1.13 si rivela utile anche per funzioni f che cambiano segno; lo si applica a $|f|$ e se si ottiene che $|f|$ è a integrale convergente allora lo è anche f. In particolare dal confronto con le potenze negative si deduce il seguente risultato:

Corollario 1.16. *Sia $a > 0$;*

i *Sia $f \in \bigcap_{b>a} \mathcal{R}([a,b];\mathbb{R})$. Se esistono $\alpha > 1$ e $M > 0$ tali che $|f(x)| \leq Mx^{-\alpha}$ per ogni $x \geq a$ allora $f \in \tilde{\mathcal{R}}([a,\infty);\mathbb{R})$.*

ii *Se esiste $m > 0$ tale che $f(x) \geq mx^{-1}$ per ogni $x \geq a$ allora f è a integrale divergente in $[a,+\infty)$.*

Dimostrazione. Dalla Proposizione 1.13 e Esempio 1.7, si ha che nel primo caso $|f|$ è integrabile in $[a,+\infty)$ e quindi pure f, grazie alla Proposizione 1.15, mentre nel secondo caso f è a integrale divergente in $[a,+\infty)$. □

Osservazione 1.17. La Proposizione 1.15 stabilisce che se il valore assoluto di una funzione è integrabile in senso improprio, allora lo è la funzione data. Vediamo con un controesempio che tale enunciato non si può invertire, ovvero esistono funzioni f integrabili in senso improprio tali che $|f|$ non lo è.

Consideriamo la funzione $f(x) = \frac{\sin x}{x}$, vogliamo dimostrare che

$$f \in \tilde{\mathcal{R}}([\pi,\infty);\mathbb{R}), \qquad |f| \notin \tilde{\mathcal{R}}([\pi,\infty);\mathbb{R}).$$

Integrando per parti si ha

$$\int_\pi^c \frac{\sin x}{x}\mathrm{d}x = \left[-\frac{\cos x}{x}\right]_\pi^c - \int_\pi^c \frac{\cos x}{x^2}\mathrm{d}x$$

Dato che

$$\lim_{c\to+\infty}\left[-\frac{\cos x}{x}\right]_\pi^c = -\frac{1}{\pi}, \qquad \left|\frac{\cos x}{x^2}\right| \le \frac{1}{x^2}, \quad \forall x \in [\pi,+\infty)$$

quindi il Corollario 1.16 ci garantisce che $f \in \tilde{\mathcal{R}}([\pi,\infty);\mathbb{R})$.

Vediamo ora che $|f| \notin \tilde{\mathcal{R}}([\pi,\infty);\mathbb{R})$,

$$\begin{aligned}\int_\pi^{n\pi}\left|\frac{\sin x}{x}\right|\mathrm{d}x &= \sum_{k=1}^{n-1}\int_{k\pi}^{(k+1)\pi}\frac{|\sin x|}{x}\mathrm{d}x \ge \sum_{k=1}^{n-1}\frac{1}{(k+1)\pi}\int_{k\pi}^{(k+1)\pi}|\sin x|\mathrm{d}x = \\ &= \sum_{k=1}^{n-1}\frac{1}{(k+1)\pi}\int_0^\pi \sin x\mathrm{d}x = \frac{2}{\pi}\sum_{k=1}^{n-1}\frac{1}{k+1} \ge \epsilon\log n \xrightarrow{n\to\infty} \infty,\end{aligned}$$

pertanto applicando la Proposizione 1.13 concludiamo.

2 Serie numeriche

Definizione 2.1 (successione)**.** Sia $\mathbb{I} \subset \mathbb{Z}$ un sottoinsieme dei numeri interi ed $(X, \|\bullet\|_X)$ uno spazio normato, una *successione* è una funzione

$$k \in \mathbb{I} \mapsto a_k \in X \tag{2.1}$$

se $X = \mathbb{R}$ allora si ha una *successione reale*, se $X = \mathbb{C}$ allora si ha una *successione complessa.* Per denotare la successione in (2.1) si utilizzeranno le più compatte notazioni $(a_k)_{k\in\mathbb{I}}, (a_k)_k$ e a invece della più corretta $k \mapsto a_k$. Denotiamo con

$$\ell^\infty(\mathbb{I};X) := \left\{k \in \mathbb{I} \mapsto a_k \in X \;\middle|\; \sup_{k\in\mathbb{I}}\|a_k\|_X < \infty\right\},$$

l'insieme delle *successioni limitate* in X.

Osservazione 2.2. Dal punto di vista applicativo considereremo sempre e solo i casi in cui $\mathbb{I} = \mathbb{N},\mathbb{Z}$ e $X = \mathbb{R},\mathbb{C}$ nella Definizione 2.1.

Notazione 2.3. Utilizzeremo la notazione $\ell^\infty := \ell^\infty(\mathbb{N};\mathbb{R})$. In questo corso ci ridurremo allo studio di successioni reali

Definizione 2.4 (Ridotta n-esima)**.** Sia $a \in \ell^\infty(\mathbb{N};X)$ e $n \in \mathbb{N}$, definiamo come *ridotta n-esima di a* la quantità

$$s_n = s_n[a] := \sum_{k=1}^n a_k. \tag{2.2}$$

Definizione 2.5 (Serie)**.** Con le stesse ipotesi come in Definizioni 2.4 denotiamo con

$$\sum_{k=1}^\infty a_k := \{n \mapsto s_n[a]\}, \tag{2.3}$$

la *serie di termine generico a_k*.

Definizione 2.6. La serie (2.3):

1. si dice convergente se converge la successione s_n delle sue somme parziali (2.2). In tal caso, il numero reale $s_\infty = \lim_{n\to\infty} s_n$ è detto somma della serie e si scrive $\sum_{k=1}^{\infty} a_k = s_\infty$;

2. si dice divergente positivamente/negativamente se la successione delle sue somme parziali (2.4) diverge positivamente/negativamente e si scrive $\sum_{k=1}^{\infty} a_k = +\infty/-\infty$;

3. si dice indeterminata se la successione delle sue somme parziali (2.4) non ammette limite.

Esempio 2.7 (Serie geometrica)**.** Dato $q \in \mathbb{R}$, la serie

$$\sum_{k=0}^{\infty} q^k,$$

è detta *serie geometrica di ragione q*. È di verifica immediatamente che $(1+s_n)\big(1-q\big) = 1-q^{n+1}$. Se $q \neq 1$ si ricava allora

$$1+s_n = \frac{1-q^{n+1}}{1-q}, \tag{2.4}$$

mentre, se $q = 1, s_n = n$. Dalla formula (2.4), si ha che

1. se $|q| < 1$ esiste $\lim_{n\to\infty} s_n = \frac{1}{1-q} - 1$;

2. se $q \geq 1$ esiste $\lim_{n\to\infty} s_n = +\infty$;

3. se $q = -1, s_n = 1, 0, 1, 0, \ldots$ e la serie è indeterminata;

4. se $q < -1, \lim_{n\to\infty} |q|^{n+1} = +\infty$, quindi $\lim_{n\to\infty} |s_n| = +\infty$, ma la successione s_n alterna valori positivi e negativi, dunque la serie è indeterminata. Si osservi ad esempio che la sottosuccessione di indice dispari s_{2n+1} diverge a $-\infty$, mentre la sottosuccessione di indice pari s_{2n} diverge a $+\infty$

Riassumendo, la serie geometrica converge (con somma $\frac{1}{1-q}$) se e soltanto se $|q| < 1$.

Esempio 2.8 (Serie armonica)**.** Si tratta della serie

$$\sum_{k=1}^{\infty} \frac{1}{k} = 1 + \frac{1}{2} + \ldots + \frac{1}{k} + \ldots \tag{2.5}$$

Dato che gli addendi sono positivi, la successione delle ridotte s_n è crescente, quindi esiste $\lim_{n\to\infty} s_n = \sup_{n\in\mathbb{N}} s_n$. Consideriamo il caso in cui $n = 2^p$

$$s_{2^p} = 1 + \frac{1}{2} + \underbrace{\left(\frac{1}{3} + \frac{1}{4}\right)}_{\geq 2} + \underbrace{\left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right)}_{\geq 2} + \ldots + \underbrace{\left(\frac{1}{2^{p-1}+1} + \ldots + \frac{1}{2^p}\right)}_{\geq 2} > 1 + \underbrace{\frac{1}{2} + \frac{1}{2} + \ldots + \frac{1}{2}}_{p} = 1 + \frac{p}{2}, \tag{2.6}$$

da cui

$$\lim_{p\to+\infty} s_{2^p} = +\infty$$

Dunque la successione s_n è superiormente illimitata e quindi divergente a $+\infty$. Più in generale, definendo i *numeri armonici*

$$\mathsf{H}_n := \sum_{k=1}^{n} \frac{1}{k},$$

è possibile dimostrare che

$$\lim_{n\to\infty} \big(\mathsf{H}_n - \log n\big) = \gamma_{\mathrm{EM}},$$

dove γ_{EM} è la *costante di Eulero-Mascheroni.*

Esempio 2.9 (Serie di Mengoli)**.** Si consideri la serie

$$\sum_{k=1}^{\infty} \frac{1}{k(k+1)} = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} \ldots + \frac{1}{k\cdot(k+1)} + \ldots$$

Si osservi che $\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1}$, da cui

$$s_n = \left(1 - \frac{1}{2}\right) + \left(\frac{1}{2} - \frac{1}{3}\right) + \ldots + \left(\frac{1}{n} - \frac{1}{n+1}\right) = 1 - \frac{1}{n+1} \xrightarrow[n\to+\infty]{} 1$$

La serie di Mengoli è una *serie telescopica.*

Teorema 2.10 (Criterio di convergenza di Cauchy)**.** *La serie* (2.3) *converge se e solo se per ogni* $\epsilon > 0$ *esiste* $\bar{n} \in \mathbb{N}$ *tale che per ogni* $n \geq \bar{n}$ *e per ogni* $p \in \mathbb{N}, p \geq 1$*, si ha*

$$\left| \sum_{k=n+1}^{n+p} a_k \right| < \epsilon. \tag{2.7}$$

Dimostrazione. Dal criterio di Cauchy per le successioni di numeri reali, si ha che s_n converge se e solo se per ogni $\epsilon > 0$ esiste $\bar{n} \in \mathbb{N}$ tale che per ogni $n, m \geq \bar{n}$ si ha

$$|s_m - s_n| < \epsilon \tag{2.8}$$

Essendo tale disuguaglianza banalmente vera per $m = n$, possiamo supporre ad esempio che sia $m > n$, quindi $m = n + p$, al variare di $p \in \mathbb{N}, p \geq 1$. La tesi segue allora dalle seguente riscrittura di $|s_m - s_n|$

$$|s_m - s_n| = \left|s_{n+p} - s_n\right| = \left| \sum_{k=1}^{n+p} a_k - \sum_{k=1}^{n} a_k \right| = \left| \sum_{k=n+1}^{n+p} a_k \right|. \tag{2.9}$$

Le Equazioni (2.8) e (2.9) danno dunque (2.7). □

Osservazione 2.11. Il Teorema 2.10 ci dice che la successione delle ridotte n-esime è di Cauchy.

Corollario 2.12 (Criterio necessario di convergenza)**.** *Se la serie* (2.3) *converge allora il suo termine generale è infinitesimo, ovvero* $\lim_{k\to\infty} a_k = 0$.

Dimostrazione. Se (2.3) converge, allora vale il Teorema 2.10. Si scelga in esso in particolare $p = 1$ e si ottiene che per ogni $\epsilon > 0$ esiste $\bar{n} \in \mathbb{N}$ tale che per ogni $n \geq \bar{n}$ si ha

$$\left| \sum_{k=n+1}^{n+1} a_k \right| = |a_{n+1}| < \epsilon.$$

□

Osservazione 2.13. La condizione $\lim_{k\to\infty} a_k = 0$ non è sufficiente per la convergenza: un controesempio è dato dalla serie armonica in (2.5).

Definizione 2.14 (Serie a termini positivi)**.** Una serie come in (2.3) si dice a termini positivi se $a_k \geq 0$, per ogni k.

Osservazione 2.15. La successione delle ridotte di una serie a termini positivi è crescente, quindi esiste

$$\lim_{n\to\infty} s_n \in \mathbb{R} \cup \{+\infty\}.$$

Vale perciò il principio noto come:

> *AUT-AUT per le serie a termini positivi:* una serie a termini positivi o converge o diverge positivamente.

Definizione 2.16 (Assoluta convergenza)**.** Una serie si dice *assolutamente convergente* se è convergente la serie $\sum_{k=1}^{\infty}|a_k|$, che ha per termine generale il valore assoluto del suo termine generale; si dice *assolutamente divergente* se la serie $\sum_{k=1}^{\infty}|a_k|$ è divergente.

Osservazione 2.17. La serie dei valori assoluti di una serie data è a termini positivi, quindi per essa vale l'aut-aut. Segue che una serie è o assolutamente convergente o assolutamente divergente.

Proposizione 2.18. *Una serie assolutamente convergente è convergente.*

Dimostrazione. Fissiamo $\epsilon > 0$, siccome la serie è assolutamente convergente applichiamo il Teorema 2.10: esiste $\bar{n} = \bar{n}(\epsilon) \in \mathbb{N}$ tale che per ogni $n \geq \bar{n}$ e per ogni $p \in \mathbb{N}$, $p \geq 1$, si ha

$$\sum_{k=n+1}^{n+p} |a_k| < \epsilon.$$

Applicando la disuguagianza triangolare otteniamo che

$$\left|\sum_{k=n+1}^{n+p} a_k\right| \leq \sum_{k=n+1}^{n+p} |a_k| < \epsilon,$$

e pertanto possiamo applicare il Teorema 2.10 alla serie (2.3). □

Osservazione 2.19. Il viceversa non vale, cioè la convergenza di una serie non implica l'assoluta convergenza: un controesempio è dato dalla seguente serie

$$1 - 1 + \frac{1}{2} - \frac{1}{2} + \frac{1}{3} - \frac{1}{3} + \ldots \frac{1}{n} - \frac{1}{n} + \ldots$$

Indicata con s_n la successione delle ridotte della serie data e con $\tilde{s}_n$ la successione delle ridotte della serie dei valori assoluti, si ha

$$s_n = 1, 0, \frac{1}{2}, 0, \frac{1}{3}, 0, \ldots \frac{1}{n}, 0 \ldots, \qquad \lim_{n\to\infty} s_n = 0,$$
$$\tilde{s}_{2n} = 2\left(1 + \frac{1}{2} + \frac{1}{3} + \ldots \frac{1}{n}\right), \qquad \lim_{n\to\infty} \tilde{s}_{2n} = +\infty.$$

Definizione 2.20 (Serie di resto)**.** Sia $N \in \mathbb{N}$. Si dice *serie resto* di indice N della serie (2.3) la serie

$$\sum_{k=1}^{\infty} a_{N+k} = a_{N+1} + a_{N+2} + \ldots + a_{N+k} + \ldots$$

che si ottiene trascurando i primi N termini di (2.3).

Proposizione 2.21. *Ogni serie resto ha lo stesso carattere della serie data. Se la serie* (2.3) *è convergente con somma s_∞, la somma della serie resto di indice N converge a*

$$r^{(N)} := s_\infty - s_N, \qquad \lim_{N\to\infty} r^{(N)} = 0.$$

Dimostrazione. La ridotta ennesima della serie resto è $\hat{s}_n = s_{N+n} - s_N$, che ha lo stesso carattere di s_n. Se poi $\lim_{n\to\infty} s_n = s_\infty$, allora $r^{(N)} = \lim_{n\to\infty} \hat{s}_n = \lim_{n\to\infty} s_{N+n} - s_N = s_\infty - s_N$. Inoltre $\lim_{N\to\infty} r^{(N)} = \lim_{N\to\infty} s_\infty - s_N = 0$. □

Osservazione 2.22. Segue dalla precedente proposizione che il carattere di una serie non cambia modificando, aggiungendo o togliendo un numero finito di suoi termini.

Definizione 2.23. Date due serie

$$\sum_{k=1}^{\infty} a_k \tag{2.10}$$

$$\sum_{k=1}^{\infty} b_k \tag{2.11}$$

si definisce serie somma di (2.10) e (2.11) la serie

$$\sum_{k=1}^{\infty} (a_k + b_k) \tag{2.12}$$

che ha per termine generale la somma dei termini generali delle serie date.

Proposizione 2.24. *Se le serie* (2.10) *e* (2.11) *convergono con somme s' e s'' rispettivamente, allora* (2.12) *converge con somma $s' + s''$. Se* (2.10) *e* (2.11) *divergono positivamente, allora* (2.12) *diverge positivamente. Se* (2.10) *diverge positivamente e la successione delle ridotte di* (2.11) *è inferiormente limitata, allora* (2.12) *diverge positivamente. Similmente per la divergenza a $-\infty$.*

Dimostrazione. La dimostrazione segue dalle proprietà del limite della somma di successioni. □

Definizione 2.25 (Serie prodotto)**.** Data una serie $\sum_{k=1}^{\infty} a_k$ e una costante $\alpha \in \mathbb{R} \setminus \{0\}$, si definisce serie prodotto della serie data per la costante α la serie $\sum_{k=1}^{\infty} \alpha a_k$. E' evidente che se la serie data ha somma $\sum_{k=1}^{\infty} a_k = s \in \mathbb{R} \setminus \{0\}$, allora $\sum_{k=1}^{\infty} \alpha a_k = \alpha s$ (dove $\alpha \cdot (\pm\infty) = \text{sgn}(\alpha) \cdot (\pm\infty)$); se la serie data è indeterminata, la serie prodotto è indeterminata.

Definizione 2.26. Si consideri una serie come in Eq. (2.3), sia $n \mapsto k_n$ una successione crescente di numeri naturali. Poniamo

$$b_1 = \sum_{k=1}^{k_1} a_k, \qquad b_2 = \sum_{k=k_1+1}^{k_2} a_k, \qquad \dots \qquad b_n = \sum_{k=k_{n-1}+1}^{k_n} a_k, \qquad \dots$$

e consideriamo la serie

$$\sum_{k=1}^{\infty} b_k \tag{2.13}$$

è detta *serie associata* di (2.3).

Proposizione 2.27. *Se* (2.3) *converge ogni sua serie associata converge con la stessa somma; se* (2.3) *diverge a $+\infty / -\infty$, ogni sua serie associata diverge a $+\infty / -\infty$.*

Dimostrazione. La dimostrazione segue dal fatto che la successione $\tilde{s}_n$ delle ridotte della serie (2.13) è una sottosuccessione della successione s_n delle ridotte della serie (2.3) $\left(\tilde{s}_1 = s_{k_1}, \tilde{s}_2 = s_{k_2}, \dots, \tilde{s}_n = s_{k_n}, \dots\right)$. □

Osservazione 2.28.

- Non vale il viceversa, ad esempio la serie $a - a + a - a \dots$, con $a \neq 0$, è indeterminata, dato che $s_n = a, 0, a, 0, \dots$, mentre la serie ottenuta associandone i termini a due a due $(a - a) + (a - a) \dots$ ovviamente converge con somma 0.
- Se la serie data è a termini positivi, allora sappiamo che s_n è regolare, quindi se una qualunque sua serie associata converge o diverge, allora la serie data converge (con la stessa somma) o diverge dato che ogni sottosuccessione tende allo stesso limite della successione.
- Si dice pure che le serie a termini positivi godono della proprietà associativa, dato che se ne può valutare carattere e somma associandone i termini a piacere.
- Se una qualunque serie associata della serie data è indeterminata, tale è anche la serie data (ragionare per assurdo).

2.1 Serie a termini positivi

Il fatto che le serie a termini positivi abbiano la successione delle ridotte crescente, quindi regolare, è di grande utilità per stabilire criteri di convergenza o di divergenza.

Definizione 2.29. Date due serie come nelle Equazioni (2.10) e (2.11) si dice che la serie (2.10) è minorante della serie (2.11) o, equivalentemente, che (2.11) è maggiorante di (2.10), se $a_k \leq b_k$ per ogni $k \in \mathbb{N}$.

Teorema 2.30 (Criterio del confronto)**.** *Date le serie* (2.10), (2.11), *tali che* $0 \leq a_k \leq b_k$ *per ogni* $k \in \mathbb{N}$

i *Se* (2.11) *converge, allora* (2.10) *converge;*

ii *Se* (2.10) *diverge, allora allora* (2.11) *diverge.*

Dimostrazione. Siano $s_n, \tilde{s}_n$ le successioni delle ridotte di (2.10), (2.11) rispettivamente. Vale ovviamente $s_n \leq \tilde{s}_n$ per ogni $n \in \mathbb{N}$. Se $\tilde{s}_n$ converge, è superiormente limitata, quindi anche s_n è superiormente limitata, ed essendo crescente è convergente. Se s_n diverge a $+\infty$, ovviamente anche $\tilde{s}_n$ diverge a $+\infty$. □

Teorema 2.31 (Criterio della radice)**.** *Si consideri una serie come in Equazione* (2.3) *a termini positivi*

1. *Se esiste* $\delta \in (0,1)$ *e* $k_0 = k_0(\delta)$, *tale che* $\sqrt[k]{a_k} \leq \delta$ *per ogni* $k > k_0$, *allora la serie converge;*

2. *Se* $\sqrt[k]{a_k} \geq 1$ *per infiniti indici* k, *allora la serie diverge.*

Dimostrazione. Nel primo caso esiste $k_0 \in \mathbb{N}$ tale che $\sqrt[k]{a_k} \leq \delta$ per ogni $k \geq k_0$. Dunque $a_k \leq \delta^k$ per ogni $k \geq k_0$. Quindi la serie resto di indice k_0 è minorante della serie resto dello stesso indice della serie geometrica di ragione δ, che è convergente essendo $0 < \delta < 1$, e quindi converge per il Teorema 2.30.

Nel secondo caso, $a_k \geq 1$ per infiniti indici e quindi a_k non è infinitesima, violando il Corollario 2.12. Quindi la serie, essendo a termini positivi, diverge. □

Osservazione 2.32. La condizione 1 è equivalente a $\limsup_{k\to\infty} \sqrt[k]{a_k} < 1$. Infatti se vale 1, allora esiste ν tale che $\Lambda_\nu = \sup_{k>\nu} \sqrt[k]{a_k} \leq \delta$, quindi $\limsup_{k\to\infty} \sqrt[k]{a_k} = \inf_{n\in\mathbb{N}} \Lambda_n \leq \delta < 1$. Viceversa, se $\lim_{\sup_{k\to\infty}} \sqrt[k]{a_k} < 1$, scelto $\delta \in \left(\limsup_{k\to\infty} \sqrt[k]{a_k}, 1\right)$, per la prima proprietà caratteristica del limite superiore si ha che definitivamente $\sqrt[k]{a_k} < \delta$. Se $\limsup_{k\to\infty} \sqrt[k]{a_k} > 1$, per la seconda proprietà caratteristica del limite superiore si ha che $\sqrt[k]{a_k} > 1$ per infiniti indici, quindi vale 2.

Da quanto osservato, si ha il seguente corollario, molto utile nello svolgimento degli esercizi.

Corollario 2.33. *Si consideri una serie come in Equazione* (2.3) *a termini positivi:*

i *Se esiste* $\lim_{k\to\infty} \sqrt[k]{a_k} < 1$, *allora la serie converge;*

ii *Se esiste* $\lim_{k\to\infty} \sqrt[k]{a_k} > 1$, *allora la serie diverge.*

Osservazione 2.34. Se esiste $\lim_{k\to\infty} \sqrt[k]{a_k} = 1$, il criterio della radice non permette di stabilire il carattere della serie, come testimonia il fatto che la serie armonica e la serie di Mengoli, la prima divergente e la seconda convergente, verificano entrambe la condizione $\lim_{k\to\infty} \sqrt[k]{a_k} = 1$.

Teorema 2.35 (Criterio del rapporto)**.** *Si consideri una serie come in Equazione* (2.3) *a termini positivi:*

1. *Se esiste* $\delta \in (0,1)$ *e* $\bar{k} \in \mathbb{N}$, *tale che* $\frac{a_{k+1}}{a_k} \leq \delta$ *per ogni* $k > \bar{k}$, *allora la serie converge;*

2. *Se* $\frac{a_{k+1}}{a_k} \geq 1$ *per ogni* $k > \bar{k} \in \mathbb{N}$, *allora la serie diverge.*

Dimostrazione. Nel primo caso esiste $\bar{k} \in \mathbb{N}$ tale che $a_{k+1} \leq \delta a_k$ per ogni $k \geq \bar{k}$. Dunque $a_{\bar{k}+p} \leq \delta^p a_{\bar{k}}$ per ogni $p \in \mathbb{N}$. Quindi la serie resto di indice $\bar{k}$ è minorante della serie prodotto di $a_{\bar{k}}$ per la serie geometrica di ragione δ, con $0 < \delta < 1$, e quindi converge grazie al Teorema 2.30. Nel secondo caso, esiste $\bar{k}$ tale che $a_{k+1} \geq a_k \geq a_{\bar{k}} > 0$ per ogni $k \geq \bar{k}$ e quindi a_k non è infinitesima, violando il Corollario 2.12. Quindi la serie, essendo a termini positivi, diverge. □

Osservazione 2.36. La condizione 1 è equivalente a $\limsup_{k\to\infty} \frac{a_{k+1}}{a_k} < 1$. Se $\liminf_{k\to\infty} \frac{a_{k+1}}{a_k} > 1$, per la prima proprietà caratteristica del limite inferiore si ha che $\frac{a_{k+1}}{a_k} > 1$ definitivamente, quindi vale 2.

Da quanto osservato, si ha il seguente corollario, molto utile nello svolgimento degli esercizi:

Corollario 2.37. *Si consideri una serie come in Equazione* (2.3) *a termini positivi:*

- **i** *Se esiste* $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} < 1$*, allora la serie converge;*
- **ii** *Se esiste* $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} > 1$*, allora la serie diverge.*

Osservazione 2.38. Se esiste $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = 1$, il criterio del rapporto non permette di stabilire il carattere della serie, come testimonia il fatto che la serie armonica e la serie di Mengoli, la prima divergente e la seconda convergente, verificano entrambe la condizione $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = 1$.

Osservazione 2.39. Per una successione $(a_k)_{k\in\mathbb{N}}$, vale la seguente catena di disuguaglianze

$$\liminf_{k\to\infty} \frac{a_{k+1}}{a_k} \le \liminf_{k\to\infty} \sqrt[k]{a_k} \le \limsup_{k\to\infty} \sqrt[k]{a_k} \le \limsup_{k\to\infty} \frac{a_{k+1}}{a_k}$$

da cui segue che il criterio della radice è in generale più forte di quello del rapporto, dato che può accadere che $\limsup_{k\to\infty} \frac{a_{k+1}}{a_k} \ge 1$, il che non permette di concludere che la serie converge (e nemmeno che diverge), mentre $\limsup_{k\to\infty} \sqrt[k]{a_k} < 1$, che garantisce la convergenza della serie; similmente può accadere che $\liminf_{k\to\infty} \frac{a_{k+1}}{a_k} \le 1$, il che non permette di concludere che la serie diverga, mentre $\limsup_{k\to\infty} \sqrt[k]{a_k} > 1$, che garantisce la divergenza della serie.
Si consideri ad esempio la serie

$$\frac{1}{2} + \frac{1}{3} + \frac{1}{2^2} + \frac{1}{3^2} + \ldots + \frac{1}{2^k} + \frac{1}{3^k} + \ldots$$

Si ha $\liminf_{k\to\infty} \frac{a_{k+1}}{a_k} = 0$, $\limsup_{k\to\infty} \frac{a_{k+1}}{a_k} = +\infty$, da cui non si può stabilire il carattere della serie, mentre si ha invece $\limsup_{k\to\infty} \sqrt[k]{a_k} = \frac{1}{\sqrt{2}}$, che garantisce la convergenza della serie. Se però esiste il limite $\lim_{k\to\infty} \frac{a_{k+1}}{a_k}$, allora esiste anche $\lim_{k\to\infty} \sqrt[k]{a_k}$ e i due limiti sono uguali. Dunque se vale $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = 1$, è inutile fare ricorso al criterio della radice.

Teorema 2.40 (Criterio della serie condensata di Cantor)**.** *Si consideri una serie come in* (2.3) *tale che*

- $a_k > 0$ *per ogni* k*;*
- $k \mapsto a_k$ *decrescente.*

La serie (2.3) *converge se e solo se converge la serie condensata*

$$\sum_{k=1}^{\infty} 2^k a_{2^k} = a_1 + 2a_2 + 4a_4 + \ldots + 2^k a_{2^k} + \ldots \tag{2.14}$$

La serie (2.3) *diverge se e solo se diverge* (2.14).

Dimostrazione. Consideriamo la serie ottenuta da (2.3) associando così i suoi termini

$$a_1 + (a_2 + a_3) + (a_4 + a_5 + a_6 + a_7) + \ldots + \left(a_{2^{k-1}} + \ldots + a_{2^k - 1}\right) + \ldots \tag{2.15}$$

che, essendo la successione a_k decrescente, è minorante di (2.14). Quindi se (2.14) converge, converge la sua minorante (2.15) per il Teorema 2.30. Essendo la serie data (2.3) a termini positivi, per essa vale la proprietà associativa e quindi anch'essa converge.

Consideriamo ora la serie ottenuta associando i termini della serie resto (cf. Definizione 2.20) di indice 1 di (2.3) al seguente modo

$$a_2 + (a_3 + a_4) + (a_5 + a_6 + a_7 + a_8) + \ldots + \left(a_{2^{k-1}+1} + \ldots + a_{2^k}\right) + \ldots \tag{2.16}$$

che, essendo a_k decrescente, è maggiorante della serie

$$a_2 + 2a_4 + 4a_8 + \ldots + 2^{k-1} a_{2^k} + \ldots \tag{2.17}$$

Se la serie data (2.3) converge, allora converge la serie (2.16), che è ottenuta associando i termini della sua serie resto di indice 1 , quindi converge la sua minorante (2.17). Infine converge la serie prodotto $2 \cdot (2.17)$ che è la serie resto di indice 1 della serie condensata (2.14), da cui la tesi. □

Esempio 2.41. La serie armonica generalizzata

$$\zeta(s) := \sum_{k=1}^{\infty} \frac{1}{k^s}, \qquad s>0, \tag{2.18}$$

converge se $s>1$, diverge se $0<s\le 1$.
Osserviamo che se $s=1$ si ritrova la serie armonica, che sappiamo essere divergente. Il termine generale è decrescente, essendo $s>0$, e si può quindi applicare il Teorema 2.40.
La serie condensata di (2.18) è

$$\sum_{k=1}^{\infty} \frac{2^k}{\left(2^k\right)^s} = \sum_{k=1}^{\infty} \left(2^{1-s}\right)^k$$

che è la serie geometrica di ragione 2^{1-s}. Dato che $2^{1-s}<1$ se e solo se $s>1$, si ha la tesi.

Osservazione 2.42. Si noti che per (2.18) i criteri della radice e del rapporto non danno informazioni utili, dato che $\lim_{k\to\infty} \frac{a_{k+1}}{a_k} = 1$.

Teorema 2.43 (Criterio del confronto asintotico o dell'ordine di infinitesimo)**.** *Si considerino due serie come nelle Equazioni* (2.10) *e* (2.11)*, a termini positivi, tali che*

$$\lim_{k\to\infty} a_k = \lim_{k\to\infty} b_k = 0.$$

1. *Se* (2.10) *converge e se b_k è infinitesimo di ordine superiore o uguale ad a_k, allora anche* (2.11) *converge.*
2. *Se* (2.10) *diverge e se b_k è infinitesimo di ordine inferiore o uguale ad a_k, allora anche* (2.11) *diverge.*

Dimostrazione. Se b_k è infinitesimo di ordine superiore o uguale ad a_k, esistono $M>0$ e $\bar{k}\in\mathbb{N}$, tale che $b_k \le M a_k$ per ogni $k\ge\bar{k}$. Applicando il Teorema 2.30 alle serie resto di indice $\bar{k}$, si ha che la convergenza della prima serie implica la convergenza della seconda.
La seconda parte dell'enunciato segue dalla prima parte ragionando per assurdo. □

Corollario 2.44. *Si consideri una serie come in Equazione* (2.3)*, a termini positivi ed infinitesimi. Se esiste $s>1$ tale che a_k è infinitesimo di ordine superiore o uguale a $\frac{1}{k^s}$, allora* (2.3) *converge.*
Se a_k è infinitesimo di ordine inferiore o uguale a $\frac{1}{k}$, allora (2.3) *diverge.*

Dimostrazione. La tesi segue dal Teorema 2.43 e Esempio 2.41. □

Osservazione 2.45. Non è sufficiente che a_k sia infinitesimo di ordine superiore a $\frac{1}{k}$ per la convergenza della serie (2.3), come illustra il seguente esempio:

$$\sum_{k=2}^{\infty} \frac{1}{k\log k}$$

in cui a_k è infinitesimo di ordine superiore a $\frac{1}{k}$, eppure la serie diverge. Infatti il termine generale è decrescente e la condensata di Cantor associata è data da

$$\sum_{k=1}^{\infty} 2^k a_{2^k} = \sum_{k=1}^{\infty} \frac{1}{\log 2^k} = \frac{1}{\log 2} \sum_{k=1}^{\infty} \frac{1}{k}$$

pertanto diverge grazie al Teorema 2.40.

Esercizio 2.46. Usando il Teorema 2.40 dimostrare che la serie

$$\sum_{k=2}^{\infty} \frac{1}{k\left(\log k\right)^s}$$

converge per $s>1$ e diverge per $s\le 1$.

Teorema 2.47 (Criterio dell'integrale)**.** *Sia* $f : [1,+\infty) \to \mathbb{R}$ *una funzione continua, positiva e decrescente. Sia* $a_k = f(k)$*. La serie* (2.3) *converge se e solo se* f *è a integrale convergente in* $[1,+\infty)$ *e, in tal caso, si ha*

$$\int_1^{+\infty} f(x)\,\mathrm{d}x \le \sum_{k=1}^{\infty} a_k \le a_1 + \int_1^{+\infty} f(x)\,\mathrm{d}x$$

La serie (2.3) *diverge se e solo se* f *è a integrale divergente in* $[1,+\infty)$.

Dimostrazione. Dalla decrescenza di f si ha che per ogni $x \in [k, k+1]$

$$a_{k+1} = f(k+1) \le f(x) \le f(k) = a_k.$$

Integrando in $[k, k+1]$, si ha

$$a_{k+1} \le \int_k^{k+1} f(x)\,\mathrm{d}x \le a_k,$$

e sommando per k che varia da 1 a n,

$$\sum_{k=1}^{n} a_{k+1} \le \int_1^{n+1} f(x)\,\mathrm{d}x \le \sum_{k=1}^{n} a_k \qquad \Leftrightarrow \qquad s_{n+1} - a_1 \le \int_1^{n+1} f(x)\,\mathrm{d}x \le s_n.$$

Dalla positività di f e dalla monotonia dell'integrale, si ha che esiste il limite $\lim_{c\to+\infty} \int_1^c f(x)\,\mathrm{d}x \in \mathbb{R}^+ \cup \{+\infty\}$, similmente, essendo la serie a termini positivi, la successione delle ridotte s_n o converge o diverge a $+\infty$. Possiamo quindi passare al limite per $n \to \infty$ ottenendo il risultato cercato. □

Osservazione 2.48. Naturalmente il criterio si applica a funzioni definite su semirette superiormente illimitate $[m,+\infty)$.

Esercizio 2.49. Usando il Teorema 2.47 determinare il carattere della serie armonica generalizzata e della serie $\sum_{k=2}^{\infty} \frac{1}{k(\log k)^s}$.

2.2 Serie a termini di segno misto

Definiamo allora serie a termini di segno misto quelle in cui compaiono infiniti termini (strettamente) positivi e infiniti termini (strettamente) negativi. Si osservi che i criteri per le serie a termini positivi possono essere utili anche per stabilire il carattere delle serie a termini di segno misto. Infatti possono essere applicati alla serie dei valori assoluti: la sua eventuale convergenza garantisce la convergenza della serie data grazie alla Proposizione 2.18. La divergenza della serie dei valori assoluti invece non implica la divergenza della serie data (vedi Osservazione 2.19), ma se si prova che $\limsup_{k\to\infty} \sqrt[k]{|a_k|} > 1$ o che $\liminf_{k\to\infty} \left|\frac{a_{k+1}}{a_k}\right| > 1$ allora il termine generale a_k non è infinitesimo (si vedano le rispettive dimostrazioni) e dunque si può affermare che la serie data non è convergente, potendo essere divergente o indeterminata.

Definizione 2.50 (Semplice convergenza)**.** Una serie si dice semplicemente convergente se è convergente ma non è assolutamente convergente.

Esempio 2.51 (Esempio di serie semplicemente convergente)**.** Consideriamo la serie

$$\sum_{k=1}^{\infty} \frac{(-1)^k}{k}, \tag{2.19}$$

ovviamente essa non è assolutamente convergente in quanto la serie armonica diverge, tuttavia si noti che la somma di due termini consecutivi ha segno definito ed inoltre

$$\frac{1}{k} - \frac{1}{k+1} = \frac{1}{k(k+1)} < \frac{1}{k^2}.$$

La definitezza del segno fa sì che (2.19) o converge o diverge (i.e. non può essere indeterminata), pertanto usando il Teorema 2.47 sulla serie $\sum_{k=1}^{\infty} \frac{1}{k^2}$ ottengo

$$\left|\sum_{k=1}^{\infty} \frac{(-1)^k}{k}\right| < \sum_{k=1}^{\infty} \frac{1}{k^2} < \infty,$$

ossia la serie è semplicemente convergente.

L'esempio 2.51 è un caso particolare di una configurazione più generale che presentiamo adesso.

Definizione 2.52 (Serie di segno alternato)**.** Le *serie a termini di segno alternato* sono serie della forma

$$\sum_{k=1}^{\infty}(-1)^{k+1}a_k = a_1 - a_2 + a_3 - a_4 + \ldots + (-1)^{k+1}a_k + \ldots \qquad a_k \geq 0 \tag{2.20}$$

Teorema 2.53 (Criterio di Leibniz)**.** *Sia* (2.20) *una serie a termini di segno alternato. Se a_k è decrescente e infinitesima, allora* (2.20) *converge ad un valore $s_\infty \in \mathbb{R}$. Inoltre $|s_n - s_\infty| \leq a_n$, per ogni $n \in \mathbb{N}$.*

Dimostrazione. Consideriamo la sottosuccessione delle ridotte di indice pari

$$s_{2n} = (a_1 - a_2) + (a_3 - a_4) + \ldots + (a_{2n-1} - a_{2n})\,.$$

Raccogliendo gli addendi come indicato sopra e tenuto conto della monotonia della successione a_k, si ha che la successione s_{2n} è crescente e non negativa, ovvero $0 \leq s_{2n} \leq s_{2n+2}$ per ogni $n \in \mathbb{N}$. Consideriamo ora la sottosuccessione delle ridotte di indice dispari

$$s_{2n+1} = a_1 - (a_2 - a_3) - (a_4 - a_5) - \ldots - (a_{2n} - a_{2n+1})\,.$$

Raccogliendo gli addendi come indicato sopra e tenuto conto della monotonia della successione a_k, si ha che la successione s_{2n+1} è decrescente. Inoltre

$$s_{2n+1} - s_{2n} = a_{2n+1} \geq 0, \tag{2.21}$$

da cui si ha che $s_{2n+1} \geq 0$ per ogni $n \in \mathbb{N}$. Abbiamo quindi ottenuto che

$$0 \leq s_{2n} \leq s_{2n+1} \leq a_1, \qquad \forall n \in \mathbb{N}, \tag{2.22}$$

inoltre per monotonia delle successioni $n \mapsto (s_{2n}, s_{2n+1})$ ottengo $(s_{2n}, s_{2n+1}) \xrightarrow{n\to\infty} \left(\underline{s_\infty}, \overline{s_\infty}\right) \in [0, a_1]^2$ inoltre siccome $a_{2n+1} \xrightarrow{n\to\infty} 0$ passando al limite per $n \to \infty$ in (2.21) otteniamo che $\underline{s_\infty} = \overline{s_\infty}$ e quindi

$$\lim_n s_{2n} = \lim_n s_{2n+1} = s_\infty \in [0, a_1]\,.$$

Applichiamo quindi il teorema dei due carabinieri ed otteniamo la convergenza di (2.20).
Per valutare l'errore fissiamo $n \in \mathbb{N}$ e sia $m := \lfloor n/2 \rfloor$ in modo che $n \in [2m, 2m+1]$, (2.22) e la monotonia della successione $k \mapsto a_k$ ci dà che

$$|s_n - s_\infty| \leq s_{2m+1} - s_{2m} = a_{2m+1} \leq a_n$$

□

3 Limiti superiori ed inferiori

Definizione 3.1. Definiamo $\bar{\mathbb{R}}$ *esteso* l'insieme

$$\bar{\mathbb{R}} := \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$$

In $\bar{\mathbb{R}}$ si estende l'ordinamento tra numeri reali ponendo

$$-\infty < a < +\infty, \quad \forall a \in \mathbb{R}$$

Osservazione 3.2. In base a tale definizione, $\bar{\mathbb{R}}$ ammette massimo $+\infty$ e minimo $-\infty$.

Definizione 3.3 (Successione regolare)**.** Una successione di numeri reali $(a_n)_{n\in\mathbb{N}}$ si dice *regolare* se esiste $\lim_{n\to\infty} a_n \in \bar{\mathbb{R}}$.

Definizione 3.4 (Limite superiore)**.** Data una successione di numeri reali $(a_n)_{n\in\mathbb{N}}$ si definisce *limite superiore* di $(a_n)_{n\in\mathbb{N}}$

$$\limsup_{n\to+\infty} a_n := \inf_{n\in\mathbb{N}} \sup_{k\geq n} a_k. \tag{3.1}$$

Osservazione 3.5. 1. Posto

$$\Lambda_n = \sup_{k\geq n} a_k = \sup\{a_n, a_{n+1}, a_{n+2}, \dots, a_{n+k}, \dots\}$$

la (3.1) si può riscrivere equivalentemente nella forma

$$\limsup_{n\to+\infty} a_n = \inf_{n\in\mathbb{N}} \Lambda_n.$$

Dato che la successione Λ_n è decrescente, esiste $\lim_{n\to+\infty} \Lambda_n = \inf_{n\in\mathbb{N}} \Lambda_n$ e quindi possiamo anche scrivere

$$\limsup_{n\to+\infty} a_n = \lim_{n\to+\infty} \Lambda_n.$$

Si usano equivalentemente le seguenti notazioni

$$\limsup_{n\to+\infty} a_n = \overline{\lim_{n\to+\infty}}\, a_n$$

2. Osserviamo che se a_n è superiormente illimitata, allora $\Lambda_n = +\infty$ per ogni n, e pertanto $\limsup_{n\to+\infty} a_n = +\infty$. Altrimenti, $\limsup_{n\to+\infty} a_n$ sarà un numero reale oppure $-\infty$ a seconda che la successione Λ_n sia convergente o tenda a $-\infty$. D'altra parte, $\Lambda_n \to -\infty$ se e solo se per ogni $m \in \mathbb{R}$, esiste $\bar{n}$ tale che $\Lambda_n < m$, per ogni $n \geq \bar{n}$, ovvero $a_n < m$, per ogni $n \geq \bar{n}$ (dalla definizione di Λ_n). Quindi $\limsup_{n\to+\infty} a_n = -\infty$ se e solo se $\lim_{n\to+\infty} a_n = -\infty$.

Definizione 3.6 (Limite inferiore)**.** Data una successione di numeri reali $(a_n)_{n\in\mathbb{N}}$ si definisce *limite inferiore*

$$\liminf_{n\to+\infty} a_n := \sup_{n\in\mathbb{N}} \inf_{k\geq n} a_k. \tag{3.2}$$

Osservazione 3.7. 1. Posto

$$\lambda_n = \inf_{k\geq n} a_k = \inf\{a_n, a_{n+1}, a_{n+2}, \dots, a_{n+k}, \dots\}$$

la (3.2) si può riscrivere equivalentemente nella forma

$$\liminf_{n\to+\infty} a_n = \sup_{n\in\mathbb{N}} \lambda_n$$

Dato che la successione λ_n è crescente, esiste $\lim_{n\to+\infty} \lambda_n = \sup_{n\in\mathbb{N}} \lambda_n$e quindi possiamo anche scrivere

$$\liminf_{n\to+\infty} a_n = \lim_{n\to+\infty} \lambda_n$$

Si usano equivalentemente le seguenti notazioni

$$\liminf_{n\to+\infty} a_n = \min_{n\to+\infty} \lim_{n} a_n = \underline{\lim}_{n\to+\infty} a_n = \lim_{n\to+\infty}{}' a_n.$$

2. Osserviamo che se a_n è inferiormente illimitata, allora $\lambda_n = -\infty$ per ogni n, e pertanto $\liminf_{n\to+\infty} a_n = -\infty$. Altrimenti, $\liminf_{n\to+\infty} a_n$ sarà un numero reale oppure $+\infty$ a seconda che la successione λ_n sia convergente o tenda a $+\infty$. D'altra parte, $\lambda_n \to +\infty$, se e solo se per ogni $M \in \mathbb{R}$, esiste $\bar{n}$ tale che $\lambda_n > M$, per ogni $n \geq \bar{n}$, ovvero $a_n > M$, per ogni $n \geq \bar{n}$ (dalla definizione di λ_n). Quindi $\liminf_{n\to+\infty} a_n = +\infty$ se e solo se $\lim_{n\to+\infty} a_n = +\infty$.

Esempio 3.8. 1. Sia $a_n = (-1)^n : \limsup_{n\to+\infty} a_n = 1, \liminf_{n\to+\infty} a_n = -1$.

2. Sia $a_n = (-1)^n n : \limsup_{n\to+\infty} a_n = +\infty, \liminf_{n\to+\infty} a_n = -\infty$.

3. Sia $a_n = \frac{(-1)^n}{n} : \limsup_{n\to+\infty} a_n = \liminf_{n\to+\infty} a_n = 0$.

Lemma 3.9. *Sia $(a_n)_n$ una cuccessione a valori in $\bar{\mathbb{R}}$, si ha che*

$$\liminf_{n\to+\infty} a_n \leq \limsup_{n\to+\infty} a_n.$$

Dimostrazione. Dalla loro definizione, segue che $\lambda_n \leq \Lambda_n$, per ogni $n \in \mathbb{N}$. Passando al limite per $n \to \infty$, si ha che $\liminf_{n\to+\infty} a_n = \lim_{n\to+\infty} \lambda_n \leq \lim_{n\to+\infty} \Lambda_n = \limsup_{n\to+\infty} a_n$. □

Lemma 3.10. *1. Il numero reale L è il limite superiore di $(a_n)_{n\in\mathbb{N}}$ se e solo se valgono le due proprietà seguenti:*

(a) $\forall \epsilon > 0, \exists \nu \in \mathbb{N}$ tale che $a_n < L + \epsilon, \forall n \geq \nu$;

(b) $\forall \epsilon > 0, a_n > L - \epsilon$, per infiniti indici n.

2. Il numero reale l è il limite inferiore di $(a_n)_{n\in\mathbb{N}}$ se e solo se valgono le due proprietà seguenti:

(a) $\forall \epsilon > 0, \exists \nu \in \mathbb{N}$ tale che $a_n > l - \epsilon, \forall n \geq \nu$;

(b) $\forall \epsilon > 0, a_n < l + \epsilon$, per infiniti indici n.

Dimostrazione. Vediamo la dimostrazione per il limite superiore. Supponiamo che L sia il limite superiore di a_n, ovvero $L = \inf_{n\in\mathbb{N}} \Lambda_n$. Allora, per la seconda proprietà caratteristica dell'estremo inferiore, per ogni $\epsilon > 0$ esiste $\nu \in N$ tale che $\Lambda_\nu < L + \epsilon$, per cui $L + \epsilon > \Lambda_\nu \geq a_n$ per ogni $n \geq \nu$, cioè vale la proprietà 1a.

Se, per assurdo, non valesse la proprietà 1b, esisterebbero $\epsilon > 0$ e $\nu \in \mathbb{N}$ tale che $a_n \leq L - \epsilon$ per ogni $n \geq \nu$. Quindi si avrebbe $L - \epsilon \geq \Lambda_\nu \geq \inf_{n\in\mathbb{N}} \Lambda_n = L$, che è una contraddizione.

Viceversa, sia $L \in \mathbb{R}$ verificante le proprietà 1a e 1b. La proprietà 1a afferma che per ogni $\epsilon > 0$, esiste $\nu \in \mathbb{N}$ tale che $L + \epsilon \geq \Lambda_\nu \geq \inf_{n\in\mathbb{N}} \Lambda_n = \limsup_{n\to+\infty} a_n$. Dall'arbitrarietà di ϵ, si ha che $L \geq \limsup_{n\to+\infty} a_n$.

La proprietà 1b implica che per ogni $\epsilon > 0 L - \epsilon < \Lambda_\nu$ per ogni ν; infatti, altrimenti, esisterebbe $\nu \in \mathbb{N}$ tale che $a_n \leq L - \epsilon$ per ogni $n \geq \nu$. Passando all'estremo inferiore, si ha allora che $L - \epsilon \leq \limsup_{n\to+\infty} a_n$. Dall'arbitrarietà di ϵ, si ha che $L \leq \limsup_{n\to+\infty} a_n$. In conclusione, $L = \limsup_{n\to+\infty} a_n$. □

Proposizione 3.11. *La successione $(a_n)_{n\in\mathbb{N}}$ ha limite $\ell \in \bar{\mathbb{R}}$ se e solo se*

$$\limsup_{n\to+\infty} a_n = \liminf_{n\to+\infty} a_n = \ell$$

Dimostrazione. Supponiamo che esista $\lim_{n\to+\infty} a_n = \ell$. Si hanno tre casi: a) $\ell = +\infty$, b) $\ell = -\infty$, c) $\ell \in \mathbb{R}$.

a La succcessione diverge a $+\infty$, quindi $\liminf_{n\to+\infty} a_n = +\infty$. Poichè $\limsup_{n\to+\infty} a_n \geq \liminf_{n\to+\infty} a_n$, si ha che $\limsup_{n\to+\infty} a_n = +\infty$ quindi limite superiore, limite inferiore e limite coincidono.

b La successione diverge a $-\infty$, quindi $\limsup_{n\to+\infty} a_n = -\infty$. Poichè $\liminf_{n\to+\infty} a_n \leq \limsup_{n\to+\infty} a_n$, si ha che $\liminf_{n\to+\infty} a_n = -\infty$ quindi limite superiore, limite inferiore e limite coincidono.

c Per ogni $\epsilon > 0$ esiste $\nu \in \mathbb{N}$ tale che $\ell - \epsilon < a_n < \ell + \epsilon$ per ogni $n \geq \nu$. Allora ℓ soddisfa le proprietà caratteristiche sia del limite superiore che del limite inferiore, quindi $\limsup_{n\to+\infty} a_n = \liminf_{n\to+\infty} a_n = \ell$.

Viceversa, supponiamo ora che $\limsup_{n\to+\infty} a_n = \liminf_{n\to+\infty} a_n = \ell$. Distinguiamo i tre casi: a) $\ell = +\infty$, b) $\ell = -\infty$, c) $\ell \in \mathbb{R}$.

a Dato che $\liminf_{n\to+\infty} a_n = +\infty$, la successione a_n diverge a $+\infty$.

b Dato che $\limsup_{n\to+\infty} a_n = -\infty$, la successione a_n diverge a $-\infty$.

c Dalle proprietà a) del limite superiore e a') del limite inferiore segue che la successione converge a ℓ.

□

4 Topologia in spazi di funzioni

4.1 Spazi metrici

Definizione 4.1 (Spazio metrico)**.** Dato un insieme non vuoto X, si dice *distanza* una funzione $\mathrm{d} : X \times X \to \mathbb{R}$ tale che per ogni $x, y, x \in X$ le seguenti condizioni sono verificate:

1. **Positività:** $\mathrm{d}\left(x, y\right) \geq 0$ e $\mathrm{d}\left(x, y\right) = 0 \Leftrightarrow x = y$;
2. **Simmetria:** $\mathrm{d}\left(x, y\right) = \mathrm{d}\left(y, x\right)$;
3. **Disuguaglianza triangolare:** $\mathrm{d}\left(x, z\right) \leq \mathrm{d}\left(x, y\right) + \mathrm{d}\left(y, z\right)$.

Un insieme X dotato di una distanza d è detto *spazio metrico* ed indicato con (X, d). I suoi elementi sono detti *punti*.

Esempio 4.2. Si consideri $\mathrm{d} \in \mathbb{N}$, $\mathrm{d} \geq 1$ e lo spazio vettoriale

$$\mathbb{R}^{\mathrm{d}} := \left\{ x = (x_1, \ldots, x_{\mathrm{d}}) \;\middle|\; x_i \in \mathbb{R}, i = 1, \ldots, \mathrm{d} \right\},$$

identificabile, per $\mathrm{d} = 1, \mathrm{d} = 2, \mathrm{d} = 3$, con la retta reale, il piano cartesiano, lo spazio cartesiano, rispettivamente. Sia d_2 definita in $\mathbb{R}^{\mathrm{d}}$ da

$$\mathrm{d}_2\left(x, y\right) := \sqrt{\sum_{i=1}^{\mathrm{d}} \left(x_i - y_i\right)^2} \tag{4.1}$$

Le prime due proprietà di distanza si verificano banalmente, mentre la disuguaglianza triangolare segue dal significato geometrico della distanza euclidea nelle dimensioni $\mathrm{d} = 1, \mathrm{d} = 2, \mathrm{d} = 3$; vediamo la sua dimostrazione nel caso generale. Posto $a_i = x_i - y_i, b_i = y_i - z_i$, per $i = 1, .., \mathrm{d}$, si tratta di verificare che

$$\sqrt{\sum_{i=1}^{\mathrm{d}} (a_i + b_i)^2} \leq \sqrt{\sum_{i=1}^{\mathrm{d}} a_i^2} + \sqrt{\sum_{i=1}^{\mathrm{d}} b_i^2} \tag{4.2}$$

Se $a_i = 0$ per ogni $i = 1, \ldots, \mathrm{d}$, o se $b_i = 0$ per ogni $i = 1, \ldots, \mathrm{d}$, (4.2) è ovvia, supponiamo quindi che $\sum_{i=1}^{\mathrm{d}} a_i^2 > 0$ e $\sum_{i=1}^{\mathrm{d}} b_i^2 > 0$. Dalla seguente disuguaglianza elementare conosciuta come *disuguaglianza di Young*

$$2ab \leq a^2 + b^2$$

si ha che, per ogni $\lambda > 0$,

$$2ab = 2(\sqrt{\lambda} a)\left(\frac{b}{\sqrt{\lambda}}\right) \leq \lambda a^2 + \frac{1}{\lambda} b^2,$$

quindi

$$2\left|\sum_{i=1}^{\mathrm{d}} a_i b_i\right| \leq 2\sum_{i=1}^{\mathrm{d}} |a_i||b_i| \leq \lambda \sum_{i=1}^{\mathrm{d}} a_i^2 + \frac{1}{\lambda}\sum_{i=1}^{\mathrm{d}} b_i^2.$$

Scegliendo il valore di λ che rende uguali i due addendi a secondo membro della precedente disuguaglianza, ovvero

$$\lambda = \frac{\sqrt{\sum_{i=1}^{\mathrm{d}} b_i^2}}{\sqrt{\sum_{i=1}^{\mathrm{d}} a_i^2}}$$

si ha

$$\left|\sum_{i=1}^{\mathrm{d}} a_i b_i\right| \leq \sqrt{\sum_{i=1}^{\mathrm{d}} a_i^2} \cdot \sqrt{\sum_{i=1}^{\mathrm{d}} b_i^2} \tag{4.3}$$

detta *disuguaglianza di Schwarz*. L'Equazione (4.3) in Eq. (4.2) da

$$\sum_{i=1}^{\mathrm{d}} (a_i + b_i)^2 = \sum_{i=1}^{\mathrm{d}} a_i^2 + \sum_{i=1}^{\mathrm{d}} b_i^2 + 2\sum_{i=1}^{\mathrm{d}} a_i b_i \leq \sum_{i=1}^{\mathrm{d}} a_i^2 + \sum_{i=1}^{\mathrm{d}} b_i^2 + 2\sqrt{\sum_{i=1}^{\mathrm{d}} a_i^2} \cdot \sqrt{\sum_{i=1}^{\mathrm{d}} b_i^2} = \left(\sqrt{\sum_{i=1}^{\mathrm{d}} a_i^2} + \sqrt{\sum_{i=1}^{\mathrm{d}} b_i^2}\right)^2,$$

che implica (4.2).
La coppia $\left(\mathbb{R}^{\mathsf{d}}, \mathsf{d}_2\right)$ è detto spazio euclideo di dimensione d e d_2 è detta distanza euclidea. Se non si specificherà diversamente, $\mathbb{R}^{\mathsf{d}}$ si considererà sempre munito della distanza euclidea.

Esempio 4.3. Dato un qualunque insieme X non vuoto, si definisce *distanza discreta* la seguente distanza

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

Esempio 4.4. $(\mathbb{R}, \mathsf{d}_Y)$, con d_Y definita da $\mathsf{d}_Y(x,y) = \left|\arctan x - \arctan y\right|$.

Osservazione 4.5. Dato un sottinsieme E di uno spazio metrico (X, d), la restrizione di d a $E \times E$ è ovviamente una distanza, detta *distanza indotta*, che rende (E, d) spazio metrico. Ad esempio $(\mathbb{Q}, \mathsf{d}_2)$ è spazio metrico con la distanza indotta dalla distanza euclidea in $\mathbb{R}$.

Definizione 4.6. Dati un punto $x_0 \in X$, e un numero reale $r > 0$, si definisce *palla (aperta) di centro x_0 e raggio r* l'insieme

$$B(x_0, r) := \{x \in X \mid \mathsf{d}(x, x_0) < r\}$$

Esercizio 4.7. Le palle in $\mathbb{R}^2$ sono cerchi. Come sono le palle rispetto alla distanza discreta?

Definizione 4.8 (Sfera metrica)**.** Sia (X, d) uno spazio metrico. L'insieme $\mathbb{S}_X(x_0, r) := \{x \in X \mid \mathsf{d}(x, x_0) = r\}$ è detto *sfera* di centro x_0 e raggio r. Se $(X, \mathsf{d}) = \left(\mathbb{R}^{\mathsf{d}}, \mathsf{d}_2\right)$ allora denotiamo $\mathbb{S}_{\mathbb{R}^{\mathsf{d}}}(x_0, r) := \mathbb{S}^{\mathsf{d}-1}(x_0, r)$. Se $(x_0, r) = (0, 1)$ scriviamo semplicemente $\mathbb{S}^{\mathsf{d}-1} := \mathbb{S}^{\mathsf{d}-1}(0,1)$ e $\mathbb{S}_X := \mathbb{S}_X(0,1)$.

Definizione 4.9. Un sottinsieme E di uno spazio metrico X si dice *limitato* se esistono $x_0 \in X$ e $r > 0$ tali che $E \subset B(x_0, r)$.

Esempio 4.10. Lo spazio metrico dell'esempio 4.4 è limitato.

Definizione 4.11. Dato $x_0 \in X$, un sottinsieme U di X si dice *intorno* di x_0 se esiste $r > 0$ tale che $B(x_0, r) \subset U$. Si noti che se U è un intorno di x_0 e V è un soprainsieme di U, allora V è intorno di x_0.

Definizione 4.12 (Distanze equivalenti)**.** Due distanze d e d' su uno stesso insieme X si dicono *equivalenti* se esiste $C \geq 1$ tale che per ogni $x, y \in X$ si ha che

$$\frac{1}{C}\mathsf{d}_Y(x,y) \leq \mathsf{d}(x,y) \leq C\mathsf{d}_Y(x,y).$$

Osservazione 4.13. Notiamo che la costante C nella Definizione 4.12 è indipendente da $x, y \in X$.

Definizione 4.14 (Interno di un insieme)**.** Dato $E \subset X$, un punto $x_0 \in X$ si dice interno a E se esiste $r > 0$ tale che $B(x_0, r) \subset E$, ovvero se E è intorno di x_0. Si definisce interno di E l'insieme

$$\mathring{E} = \{x \in X \mid x \text{ è interno a } E\}$$

Ovviamente, si ha $\mathring{E} \subset E$.

Definizione 4.15 (Insieme aperto)**.** Un sottinsieme E di X si dice aperto se ogni suo punto è interno a E, ovvero se E è intorno di ogni suo punto. Ovviamente E è aperto se e solo se $\mathring{E} = E$.

Esercizio 4.16. Si verifichi che

a l'unione di una famiglia di insiemi aperti è un insieme aperto;

b l'intersezione di una famiglia finita di insiemi aperti è un insieme aperto.

Definizione 4.17 (Punto di aderenza)**.** Dato $E \subset X$, un punto $x_0 \in X$ si dice *aderente* a E se ogni suo intorno contiene almeno un punto appartenente a E.

Definizione 4.18 (chiusura di un insieme)**.** Si definisce *chiusura* di E l'insieme

$$\bar{E} = \{x \in X \mid x \text{ è aderente a } E\}$$

Ovviamente, si ha $E \subset \bar{E}$.

Definizione 4.19 (Insieme chiuso)**.** Un sottinsieme E di X si dice chiuso se contiene tutti i punti aderenti a E. E è chiuso se e solo se $\bar{E} = E$.

Esercizio 4.20. Si verifichi che

a l'intersezione di una famiglia di insiemi chiusi è un insieme chiuso;

b l'unione di una famiglia finita di insiemi chiusi è un insieme chiuso.

Definizione 4.21 (Punto di frontiera)**.** Dato $E \subset X$, un punto $x_0 \in X$ si dice punto di frontiera per E se ogni suo intorno contiene almeno un punto di E e almeno un punto di $\complement E$. Si definisce frontiera di E l'insieme

$$\partial E = \{x \in X \mid x \text{ è di frontiera per } E\}$$

Ovviamente, si ha

$$\partial E = \bar{E} \cap \overline{E^{\complement}}, \qquad \partial\left(E^{\complement}\right) = \partial E.$$

Esempio 4.22. Sia $E = \left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1, y \geq 0\right\}$. Si ha

$$\begin{aligned}
\mathring{E} &= \left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 < 1, y > 0\right\} \\
\bar{E} &= \left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 \leq 1, y \geq 0\right\} \\
\partial E &= \left\{(x, y) \in \mathbb{R}^2 \mid x^2 + y^2 = 1, y > 0\right\} \cup \left\{(x, y) \in \mathbb{R}^2 \mid -1 \leq x \leq 1 y = 0\right\}
\end{aligned}$$

Esercizio 4.23. Dato un insieme $E \subset X$, si verifichi che:

i E è aperto se e solo se non contiene alcun punto della sua frontiera ($\partial E \cap E = \emptyset$);

ii E è chiuso se e solo se contiene tutti i punti della sua frontiera ($\partial E \subset E$).

Definizione 4.24 (Insieme derivato)**.** Dato un sottinsieme E di uno spazio metrico (X, d), un punto $x_0 \in X$ si dice *punto di accumulazione* per E se ogni suo intorno contiene almeno un punto di E diverso da x_0. Si definisce *derivato di* E l'insieme

$$\mathcal{D}(E) = \{x \in X \mid x \text{ è di accumulazione per } E\}$$

Osservazione 4.25. Si ha che $\bar{E} = E \cup \mathcal{D}(E)$.

Definizione 4.26 (Punto isolato)**.** Dato un sottinsieme E di uno spazio metrico (X, d), un punto $x_0 \in E$ si dice *punto isolato* di E se esiste $r > 0$ tale che $B(x_0, r) \cap E = \{x_0\}$.

Esempio 4.27. Sia $E = \left\{\frac{1}{n}\right\}_{n \in \mathbb{N}} \subset \mathbb{R}$. L'unico punto di accumulazione di E è 0 , mentre ogni punto di E è un punto isolato di E.

Lemma 4.28. *E è aperto se e solo se $E^{\complement}$ è chiuso*

Lemma 4.29. *X è aperto e chiuso; $\emptyset$ è aperto e chiuso.*

Lemma 4.30. *Per ogni $x_0 \in X$ e per ogni $r > 0, B(x_0, r)$ è un insieme aperto.*

Lemma 4.31. *Per ogni $E \subset X$, si ha che $\mathring{E}$ è aperto e $\bar{E}$ è chiuso.*

Esercizio 4.32. Dimostrare i Lemmi da 4.28 a 4.31.

Esercizio 4.33. Verificare le seguenti inclusioni

$$E_1 \subset E_2 \Rightarrow \mathring{E}_1 \subset \mathring{E}_2 \qquad\qquad E_1 \subset E_2 \Rightarrow \overline{E_1} \subset \overline{E_2}. \tag{4.4}$$

Osservazione 4.34. Dalla prima formula in (4.4) si ha che se $A \subset E$ e A è aperto, allora $A = \mathring{A} \subset \mathring{E}$, ovvero ogni aperto contenuto in E è contenuto in $\mathring{E}$. Dal Lemma 4.31 , segue allora che " $\mathring{E}$ è il più grande aperto contenuto in E ". Similmente, si dimostra che " $\bar{E}$ è il più piccolo chiuso contenente E ". In formule matematiche

$$\mathring{E} := \bigcup_{\substack{A \subset E \\ A \text{ aperto}}} A, \qquad\qquad \bar{E} := \bigcap_{\substack{C \supset E \\ C \text{ chiuso}}} C.$$

Lemma 4.35. *Per ogni $x_0 \in X$ e per ogni $r > 0$, l'insieme $F := \{x \in X \mid \mathsf{d}(x, x_0) \le r\}$ è chiuso.*

Osservazione 4.36. Da quanto sopra, segue che l'insieme $\overline{B(x_0, r)} \subset F$, ma non è detto che i due insiemi coincidano. Ad esempio, se si considera la distanza discreta introdotta nell'esempio 4.3, per ogni $x_0 \in X$ si ha

$$B_1(x_0) = \{x_0\}, \qquad\qquad \overline{B_1(x_0)} = \{x_0\} \subset \left\{x \in X \mid \mathsf{d}(x, x_0) \le 1\right\} = X,$$

quindi se X contiene almeno due punti $\overline{B_1(x_0)} \subsetneqq \{x \in X \mid \mathsf{d}(x, x_0) \le 1\}$.

Definizione 4.37 (Successione convergente)**.** Sia (X, d) uno spazio metrico. Si dice che una successione $(x_n)_n$ a valori in X *converge* a un punto $x_0 \in X$, e si scrive $\lim_{n\to\infty} x_n = x_0$, se $\forall \epsilon > 0, \exists \bar{n} \in \mathbb{N} : \forall n \ge \bar{n}, \mathsf{d}(x_n, x_0) < \epsilon$.

Definizione 4.38 (Successione di Cauchy)**.** Sia (X, d) uno spazio metrico. Una successione $(x_n)_n$ a valori in X è detta successione di Cauchy se $\forall \epsilon > 0, \exists \bar{n} \in \mathbb{N} : \forall n, m \ge \bar{n}, \mathsf{d}(x_n, x_m) < \epsilon$.

Lemma 4.39. *Ogni successione convergente è di Cauchy. Ogni successione di Cauchy è limitata.*

Dimostrazione. Se $(x_n)_n$ converge a x_0, allora

$$\forall \epsilon > 0, \exists \bar{n} \in \mathbb{N} : \forall n \ge \bar{n}, \mathsf{d}(x_n, x_0) < \frac{\epsilon}{2}$$

da cui

$$\forall n, m \ge \bar{n}, \mathsf{d}(x_n, x_m) \le \mathsf{d}(x_n, x_0) + \mathsf{d}(x_m, x_0) < \epsilon.$$

Se $(x_n)_n$ è di Cauchy, scegliendo $\epsilon = 1$, si ha che esiste $\bar{n} \in \mathbb{N}$ tale che per ogni $n \ge \bar{n}, \mathsf{d}(x_n, x_{\bar{n}}) < 1$ Sia $M = \max\{\mathsf{d}(x_1, x_{\bar{n}}), \dots, \mathsf{d}(x_{\bar{n}-1}, x_{\bar{n}}), 1\}$. Per ogni $n \in \mathbb{N}, \mathsf{d}(x_n, x_{\bar{n}}) \le M$, da cui $\{x_n\}_{n\in\mathbb{N}} \subset B_{2M}(x_{\bar{n}})$. □

Definizione 4.40 (Spazio metrico completo)**.** Uno spazio metrico (X, d) si dice *completo* se ogni successione di Cauchy in X è convergente a un punto $x_0 \in X$. Un sottinsieme E di X si dice completo se lo è come spazio metrico con la metrica indotta da quella di X.

Lemma 4.41. *Ogni sottinsieme chiuso E di uno spazio metrico completo (X, d) è completo.*

Dimostrazione. Sia $(x_n)_n$ una successione di Cauchy in E. Sappiamo che esiste $x_0 \in X$ tale che x_n converge a x_0. Quindi x_0 è aderente all'insieme $S = \{x_n\}_{n\in\mathbb{N}} \subset E$, da cui $x_0 \in \bar{E} = E$. □

Teorema 4.42 (Completezza di $\mathbb{R}^\mathsf{d}$)**.** *Lo spazio euclideo $\left(\mathbb{R}^\mathsf{d}, \mathsf{d}_2\right)$ è completo.*

Dimostrazione. Sia $\left(x^{(n)}\right)_n$ una successione di Cauchy in $\left(\mathbb{R}^\mathsf{d}, \mathsf{d}_2\right)$. Dunque

$$\forall \epsilon > 0, \exists \bar{n} \in \mathbb{N} : \forall n, m \ge \bar{n}, \mathsf{d}\left(x^{(n)}, x^{(m)}\right) < \epsilon$$

Dalla definizione di distanza euclidea, si ha che

$$\left|x_i^{(n)} - x_i^{(m)}\right| \le \sqrt{\sum_{i=1}^{\mathsf{d}} \left(x_i^{(n)} - x_i^{(m)}\right)^2} = \mathsf{d}\left(x^{(n)}, x^{(m)}\right), \text{ per ogni } i = 1, \dots, \mathsf{d}$$

da cui segue che per ogni $i = 1, \dots, \mathsf{d}$, la successione $x_i^{(n)}$ è di Cauchy in $\mathbb{R}$ e dunque, per la completezza di $\mathbb{R}$, converge a un numero reale x_i. Posto $x = (x_1, \dots, x_{\mathsf{d}}) \in \mathbb{R}^{\mathsf{d}}$, si ha allora che

$$\mathsf{d}\left(x^{(n)}, x\right) = \sqrt{\sum_{i=1}^{\mathsf{d}} \left(x_i^{(n)} - x_i\right)^2} \xrightarrow{n \to \infty} 0,$$

cioè $x^{(n)}$ converge a x in $\mathbb{R}^{\mathsf{d}}$. $\square$

Osservazione 4.43. Gli spazi metrici $\mathbb{Q}$ e $(0,1)$ con la distanza indotta dalla distanza euclidea in $\mathbb{R}$ non sono completi $x_n = \left(1 + \frac{1}{n}\right)^n$ converge al numero irrazionale e quindi è di Cauchy in $\mathbb{Q}$ ma non converge in $\mathbb{Q}$, analogamente $x_n = \frac{1}{n}$ converge a $0 \notin (0,1)$.

4.2 Spazi normati e spazi dotati di prodotto scalare

Definizione 4.44 (Spazi normati)**.** Sia X uno spazio vettoriale reale sul campo $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Si dice *norma* un'applicazione $\|\cdot\| : X \to \mathbb{R}_+$ verificante per ogni $x, y, z \in X$ e $\lambda \in \mathbb{K}$:

1. **Positività:** $\|x\| \geq 0$, $\|x\| = 0 \Leftrightarrow x = 0$,
2. **Positiva omogeneità:** $\|\lambda x\| = |\lambda| \|x\|$,
3. **Subadditività:** $\|x + y\| \leq \|x\| + \|y\|$.

Uno spazio vettoriale X dotato di una norma $\|\cdot\|$ è detto *spazio normato* ed indicato con $(X, \|\cdot\|)$.

Definizione 4.45 (Seminorma)**.** Un'applicazione $|\cdot| : X \to \mathbb{R}_+$ come in Definizioni 4.44 nella quale la condizione 1 è sostituita con la più semplice $|x| \geq 0$, $\forall x \in X$ si dice *seminorma*.

Osservazione 4.46. Notiamo che dalla proprietà di subaddittività nella Definizione 4.44 abbiamo immediatamente che la norma è una funzione continua da X a $\mathbb{R}_+$, infatti

$$\left\| x + \epsilon y \right\| \leq \|x\| + |\epsilon| \left\| y \right\| \xrightarrow{\epsilon \to 0} \|x\|.$$

Osservazione 4.47. La norma induce canonicamente una distanza, definita come

$$\mathsf{d}\left(x, y\right) := \|x - y\|. \tag{4.5}$$

Le proprietà di distanza sono di immediata verifica. Dunque *uno spazio normato è anche spazio metrico.*

Definizione 4.48 (Spazio di Banach)**.** Uno spazio normato è completo rispetto alla distanza indotta dalla norma, è detto spazio di Banach.

Lemma 4.49. *Sia* $(X, \|\cdot\|)$ *uno spazio normato, allora per ogni* $x_0 \in X$ *e* $r > 0$

$$\overline{B\left(x_0, r\right)} = \{x \in X \mid \|x - x_0\| \leq r\}.$$

Dimostrazione. Dal Lemma 4.35 sappiamo che $\overline{B\left(x_0, r\right)} \subset \{x \in X \mid \mathsf{d}\left(x, x_0\right) \leq r\}$. Proviamo ora l'inclusione inversa. Sia x_1 un qualunque punto appartenente a $\{x \in X \mid \mathsf{d}\left(x, x_0\right) \leq r\}$. Se $\mathsf{d}\left(x_1, x_0\right) < r$, allora $x_1 \in B\left(x_0, r\right) \subset \overline{B\left(x_0, r\right)}$. Sia $\mathsf{d}\left(x_1, x_0\right) = r$. Si tratta di provare che x_1 è aderente a $B\left(x_0, r\right)$, ovvero che per ogni $\epsilon > 0$ si ha che $B\left(x_1, \epsilon\right)$ contiene almeno un punto di $B\left(x_0, r\right)$. Non è restrittivo supporre $\epsilon < 2r$. Consideriamo il punto $x_2 = x_0 + \left(1 - \frac{\epsilon}{2r}\right)\left(x_1 - x_0\right)$. Si ha che $\mathsf{d}\left(x_2, x_1\right) = \|x_2 - x_1\| = \left\|\frac{\epsilon}{2r}\left(x_0 - x_1\right)\right\| = \frac{\epsilon}{2} < \epsilon$, cioè $x_2 \in B\left(x_1, \epsilon\right)$, $\mathsf{d}\left(x_2, x_0\right) = \|x_2 - x_0\| = \left\|\left(1 - \frac{\epsilon}{2r}\right)\left(x_0 - x_1\right)\right\| = r - \frac{\epsilon}{2} < r$, cioè $x_2 \in B\left(x_0, r\right)$. $\square$

Osservazione 4.50. Negli spazi normati si ha

$$\mathbb{S}_X\left(x_0, r\right) = \partial B\left(x_0, r\right).$$

Infatti, dato che $B\left(x_0, r\right)$ è aperto, il suo complementare è chiuso e quindi si ha

$$\partial B\left(x_0, r\right) = \overline{B\left(x_0, r\right)} \cap \overline{B\left(x_0, r\right)^{\complement}} = \{x \in X \mid \mathsf{d}\left(x, x_0\right) \leq r\} \cap \{x \in X \mid \mathsf{d}\left(x, x_0\right) \geq r\} = \{x \in X \mid \mathsf{d}\left(x, x_0\right) = r\}$$

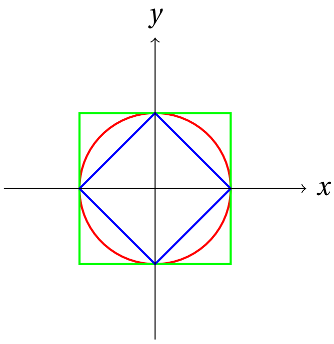

Figura 1: Rappresentazione delle palle unitarie nel piano 2D secondo diverse norme: norma euclidea (cerchio rosso), norma infinito (quadrato verde) e norma uno (diamante blu).

Definizione 4.51 (Equivalenza di norme)**.** Due norme definite in uno stesso spazio vettoriale X si dicono equivalenti se le distanze indotte (cf. Eq. (4.5)) sono equivalenti come in Definizione 4.12.

Esempio 4.52. In $\mathbb{R}^{\mathtt{d}}$, consideriamo le seguenti norme

$$\|x\|_2 = \sqrt{\sum_{i=1}^{\mathtt{d}} x_i^2}, \qquad \|x\|_\infty = \max_{i=1,\dots,\mathtt{d}} \{|x_i|\}, \qquad \|x\|_1 = \sum_{i=1}^{\mathtt{d}} |x_i|. \tag{4.6}$$

La verifica delle proprietà di norma sono immediate. Per $\mathtt{d} = 2$, le palle associate dalle precedenti norme sono rappresentate in Figura 1.

Si verifica in modo elementare che

$$\|x\|_\infty \le \|x\|_2 \le \|x\|_1 \qquad \|x\|_2 \le \sqrt{\mathtt{d}}\|x\|_\infty \qquad \|x\|_1 \le \sqrt{\mathtt{d}}\|x\|_2, \tag{4.7}$$

tali norme sono pertanto equivalenti.

Definizione 4.53 (Prodotto scalare)**.** Sia X uno spazio vettoriale su $\mathbb{K} = \mathbb{R}, \mathbb{C}$. Si dice *prodotto scalare* un'applicazione $\langle \cdot \mid \cdot \rangle \; X \times X \to \mathbb{K}$ tale che per ogni $x, y, z \in X$ e $\lambda \in \mathbb{K}$:

1. **Positività:** $\langle x \mid x \rangle \ge 0$ e $\langle x \mid x \rangle = 0 \Leftrightarrow x = 0$;
2. **Hermitiana:** $\left\langle x \,\middle|\, y \right\rangle = \overline{\left\langle y \,\middle|\, x \right\rangle}$;
3. **Lineare:** $\left\langle x + \lambda y \,\middle|\, z \right\rangle = \langle x \mid z \rangle + \lambda \left\langle y \,\middle|\, z \right\rangle$.

Uno spazio vettoriale X dotato di un prodotto scalare $\langle \cdot \mid \cdot \rangle$ è indicato con $(X, \langle \cdot \mid \cdot \rangle)$.

Osservazione 4.54. Notiamo che si ha che $\left\langle x \,\middle|\, y + \lambda z \right\rangle = \left\langle x \,\middle|\, y \right\rangle + \bar{\lambda} \langle x \mid z \rangle$.

Osservazione 4.55. Notiamo che nel caso si abbia un prodotto scalare *reale* la condizione 2 nella Definizioni 4.53 diventa una condizione di simmetria, ossia $\left\langle x \,\middle|\, y \right\rangle = \left\langle y \,\middle|\, x \right\rangle$.

Osservazione 4.56. Il prodotto scalare induce canonicamente una norma, definita come

$$\|x\| := \sqrt{\langle x \mid x \rangle}.$$

Teorema 4.57 (Disuguaglianza di Cauchy-Schwarz)**.** *Sia X uno spazio vettoriale dotato di prodotto scalare. Allora per ogni $x, y \in X$ si ha che*

$$\left|\left\langle x \,\middle|\, y \right\rangle\right| \le \|x\| \left\|y\right\|.$$

Dimostrazione. In base alla precedente osservazione, la disuguaglianza è ovvia se $y = 0$. Sia allora $y \neq 0$, così che $\|y\| > 0$. Per ogni $\lambda \in \mathbb{R}$, si ha

$$0 \leq \|x+\lambda y\|^2 = \langle x+\lambda y \mid x+\lambda y\rangle = \|x\|^2 + \bar{\lambda}\langle x \mid y\rangle + \lambda\langle y \mid x\rangle + |\lambda|^2 \left\|y\right\|^2 = = \|x\|^2 + 2\bar{\lambda}\langle x \mid y\rangle + |\lambda|^2 \|y\|^2$$

Scegliamo $\lambda = -\frac{\langle x \mid y\rangle}{\|y\|^2}$, ottenendo

$$\|x\|^2 - \frac{\left|\langle x \mid y\rangle\right|^2}{\left\|y\right\|^2} \geq 0,$$

da cui segue la tesi. □

Definizione 4.58 (Spazio di Hilbert)**.** Uno spazio dotato di prodotto scalare che sia completo rispetto alla metrica indotta dal prodotto scalare, è detto *spazio di Hilbert.*

Esempio 4.59. In $\mathbb{R}^{\mathrm{d}}$, si definisce prodotto scalare euclideo

$$x \cdot y := \sum_{i=1}^{\mathrm{d}} x_i y_i,$$

Esso induce la norma euclidea (4.6) e la distanza euclidea (4.1). $\mathbb{R}^{\mathrm{d}}$, munito del prodotto scalare euclideo, è dunque uno spazio di Hilbert. Se non si specificherà diversamente, $\mathbb{R}^{\mathrm{d}}$ si considererà sempre munito del prodotto scalare euclideo.

Definizione 4.60. Sia X uno spazio vettoriale dotato di prodotto scalare, $x, y \in X$ sono *ortogonali* se $\langle x \mid y\rangle = 0$.

Esempio 4.61. Ricordiamo che il toro unidimensionale $\mathbb{T}$ è definito come la classe di equivalenza $\mathbb{T} := \mathbb{R}/(2\pi\mathbb{Z})$. La famiglia di funzioni $\left(x \mapsto e^{\mathrm{i}jx}\right)_{j\in\mathbb{Z}}$ è composta da elementi ortogonali rispetto al prodotto scalare

$$\langle f \mid g\rangle_{L^2(\mathbb{T};\mathbb{C})} := \int_{-\pi}^{\pi} f(x)\,\overline{g(x)}\,\mathrm{d}x.$$

4.3 Compattezza

Definizione 4.62 (Totale limitatezza)**.** Un sottinsieme E di uno spazio metrico (X, d) si dice *totalmente limitato* se per ogni $\epsilon > 0$ esiste un' ϵ-rete per E, cioè un $k = k(\epsilon) \in \mathbb{N}$ ed un insieme finito $\{x_1, \dots, x_k\}$ di punti di X tale che $E \subset \bigcup_{i=1}^{k} B(x_i, \epsilon)$.

Osservazione 4.63. Un insieme totalmente limitato è limitato; infatti, considerata una 1-rete $\{x_1, \dots, x_k\}$ di E, si ha che $E \subset B(x_1, \mathrm{d}+1)$, dove $\mathrm{d} := \max_{i=2,\dots,k} \mathrm{d}(x_i, x_1)$.

Definizione 4.64 (Spazio compatto)**.** Un sottinsieme K di uno spazio metrico (X, d) si dice *sequenzialmente compatto* se per ogni $(x_n)_{n\in\mathbb{N}} \subset K$ esiste una sottosuccessione $\left(x_{n_k}\right)_{k\in\mathbb{N}} \subset K$ ed un punto $x_0 \in K$ tale che $\mathrm{d}\left(x_{n_k}, x_0\right) \xrightarrow{k\to\infty} 0$.

Proposizione 4.65. *Un sottinsieme K di uno spazio metrico (X, d) è sequenzialmente compatto se e soltanto se è completo e totalmente limitato.*

Dimostrazione. L'insieme vuoto è banalmente sequenzialmente compatto, completo e totalmente limitato. Sia quindi $K \neq \emptyset$. Supponiamo che K sia sequenzialmente compatto. Sia $(x_n)_n$ una successione di Cauchy in K. Per ipotesi, $(x_n)_n$ ammette una sottosuccessione $\left(x_{n_k}\right)_k$ convergente a un punto $x_0 \in K$. Resta da provare che l'intera successione converge a x_0. Sappiamo che

$$\forall\epsilon > 0, \exists\bar{n} \in \mathbb{N} : \forall n, m \geq \bar{n}, \mathrm{d}(x_n, x_m) < \frac{\epsilon}{2}, \qquad \forall\epsilon > 0, \exists\bar{k} \in \mathbb{N}, n_{\bar{k}} \geq \bar{n} : \forall k \geq \bar{k}, \mathrm{d}\left(x_{n_k}, x_0\right) < \frac{\epsilon}{2}.$$

Allora per ogni $n \geq \bar{n}, \mathrm{d}(x_n, x_0) \leq \mathrm{d}\left(x_n, x_{n_{\bar{k}}}\right) + \mathrm{d}\left(x_{n_{\bar{k}}}, x_0\right) < \epsilon$. Dunque K è completo. Per provare che K è totalmente limitato, ragioniamo per assurdo. Se non lo fosse, esisterebbe $\epsilon > 0$ tale che K è privo di ϵ-rete.

Scelto un qualunque punto $x_1 \in K$, l'insieme finito $\{x_1\}$ non è una ϵ-rete, e dunque esiste $x_2 \in K$ tale che $x_2 \notin B(x_1,\epsilon)$, cioè $\mathsf{d}(x_2,x_1) \geq \epsilon$. Ora, l'insieme finito $\{x_1,x_2\}$ non può costituire una ϵ-rete, e perciò esiste $x_3 \in K$ tale che $x_3 \notin \cup_{i=1}^{2} B_\epsilon(x_i)$, cioè $\mathsf{d}(x_3,x_i) \geq \epsilon, i = 1,2$. Così proseguendo si costruisce una successione $(x_n)_n$ a valori in K tale che $\mathsf{d}(x_n,x_m) \geq \epsilon$ per ogni $n,m \in \mathbb{N}, n \neq m$. Essendo K sequenzialmente compatto, $(x_n)_n$ ammette una sottosuccessione x_{n_k} convergente a un punto di K, e perciò di Cauchy, in contraddizione con la condizione $\mathsf{d}(x_n,x_m) \geq \epsilon$ per ogni $n,m \in \mathbb{N}, n \neq m$.

Supponiamo ora che K sia completo e totalmente limitato. Sia $(x_n)_n$ una successione a valori in K. Esiste una 1-rete per K, cioè esiste un sottinsieme finito F_1 di X tale che $K \subset \bigcup_{x\in F_1} B_1(x)$. Almeno una di queste palle, che indichiamo per semplicità con B_1 senza specificarne il centro, deve contenere x_n per infiniti valori dell'indice. Esiste una $\frac{1}{2}$-rete per K, ovvero esiste un sottinsieme finito F_2 di X tale che $K \subset \cup_{x\in F_2} B_{\frac{1}{2}}(x)$. Si ha che $B_1 \cap K \subset B_1 \cap \left(\bigcup_{x\in F_2} B_{\frac{1}{2}}(x)\right) = \bigcup_{x\in F_2}\left(B_1 \cap B_{\frac{1}{2}}(x)\right)$. Esiste almeno una di queste palle di raggio $\frac{1}{2}$, che indichiamo con $B_{\frac{1}{2}}$, tale che $B_1 \cap B_{\frac{1}{2}}$ contiene x_n per infiniti indici. Proseguendo, per ogni $k \in \mathbb{N}$ esiste una palla di raggio $B_{\frac{1}{k}}$ di raggio $\frac{1}{k}$ tale che $B_1 \cap B_{\frac{1}{2}} \cap \ldots \cap B_{\frac{1}{k}}$ contiene x_n per infiniti indici. Si può dunque scegliere $x_{n_1} \in B_1, x_{n_2} \in B_1 \cap B_{\frac{1}{2}}$, con $n_2 > n_1$, eccetera, estraendo una sottosuccessione x_{n_k} tale che $x_{n_k} \in \bigcap_{i=1}^{k} B_{\frac{1}{i}}$. Se $k,l \geq j, x_{n_k},\ x_{n_l} \in B_{\frac{1}{j}}$, e quindi $\mathsf{d}\left(x_{n_k},x_{n_l}\right) < \frac{2}{j}$. Scelto $\epsilon > 0$, sia $\bar{k} > \frac{2}{\epsilon}$ si ha che per ogni $k,l \geq \bar{k}, \mathsf{d}\left(x_{n_k},x_{n_l}\right) < \frac{2}{\bar{k}} < \epsilon$. Dunque x_{n_k} è di Cauchy ed essendo K completo, converge a un punto $x_0 \in K$. □

Teorema 4.66. *Un sottinsieme K di $\mathbb{R}^{\mathrm{d}}$ è sequenzialmente compatto se e solo se è chiuso e limitato.*

Dimostrazione. Un chiuso in uno spazio completo è completo (cf. Lemma 4.41). Dimostriamo che ogni sottoinsieme limitato di $\mathbb{R}^{\mathrm{d}}$ è totalmente limitato. Grazie all'equivalenza della norma Eulclidea con la norma ∞ dimostrata in (4.7) è sufficiente ragionare con ipercubi (ossia palle per norma del sup). Se K è limitato allora esiste $R > 0$ tale che $K \subset B_\infty(0,R)$, possiamo ovviamente suddividere $B_\infty(0,R)$ in un reticolo finito di taglia $\epsilon > 0$, questo conclude la dimostrazione. □

Corollario 4.67 (Teorema di Bolzano-Weierstrass)**.** *Ogni successione limitata $(x_n)_n$ in $\mathbb{R}^{\mathrm{d}}$ ammette una sottosuccessione convergente.*

Dimostrazione. Sia $F = \{x_n, n \in \mathbb{N}\}$. La sua chiusura $\bar{F}$ è chiuso e limitato. Dal Teorema 4.66, $\bar{F}$ è sequenzialmente compatto, da cui la tesi. □

5 Funzioni continue

Definizione 5.1. Un'applicazione tra spazi metrici $f : (X,\mathsf{d}_X) \to (Y,\mathsf{d}_Y)$ si dice *continua* in $x_0 \in X$ se

$$\forall \epsilon > 0, \exists \delta > 0 \,\big|\, \forall x \in X, \quad \mathsf{d}_X(x,x_0) < \delta \;\Rightarrow\; \mathsf{d}_Y\left(f(x),f(x_0)\right) < \epsilon. \tag{5.1}$$

La funzione f si dice continua in $E \subset X$ se lo è in ogni punto di E.

Definizione 5.2. Dati $f : E \subset (X,\mathsf{d}_X) \to (Y,\mathsf{d}_Y), x_0 \in X$ punto di accumulazione per E e $y_0 \in Y$, si dice che

$$\lim_{x\to x_0} f(x) = y_0,$$

se

$$\forall \epsilon > 0, \exists \delta > 0 \,\big|\, \forall x \in E, \quad 0 < \mathsf{d}_X(x,x_0) < \delta \Rightarrow \mathsf{d}_Y\left(f(x),y_0\right) < \epsilon\,.$$

Osservazione 5.3. Segue dalla Definizione 5.2 che $f : E \subset (X,\mathsf{d}_X) \to (Y,\mathsf{d}_Y)$ è continua in un punto $x_0 \in E, x_0$ di accumulazione per E, se e solo se

$$\exists \lim_{x\to x_0} f(x) = f(x_0)\,.$$

Teorema 5.4 (Teorema di compattezza)**.** *Sia $f : (X,\mathsf{d}_X) \to (Y,\mathsf{d}_Y)$ continua. Se$K \subset X$ è sequenzialmente compatto allora $f(K)$ è sequenzialmente compatto.*

Dimostrazione. Sia $\left(y_n\right)_n$ una successione a valori in $f(K)$. Per ogni $n \in \mathbb{N}$, esiste $x_n \in K$ tale che $y_n = f(x_n)$. Essendo K sequenzialmente compatto, esiste $\left(x_{n_k}\right)_k$ convergente a un punto $x_0 \in K$. Dalla continuità di f in x_0 segue che $\lim_{k\to\infty} f\left(x_{n_k}\right) = f(x_0) \in f(K)$. Quindi y_{n_k} converge a un punto di $f(K)$. □

Teorema 5.5 (di Weierstrass)**.** *Sia $f : (X, \mathrm{d}) \to \mathbb{R}$ continua. Se $K \subset X, K \neq \emptyset$, è sequenzialmente compatto allora f assume massimo e minimo su K.*

Dimostrazione. $f(K)$ è sequenzialmente compatto per il Teorema 5.4, dunque grazie al Teorema 4.66 è chiuso e limitato in $\mathbb{R}$. Ne segue che $\sup_{x\in K} f(x) = M \in \mathbb{R}$. Dalle proprietà dell'estremo superiore, segue che M è aderente a $f(K)$, che è chiuso, quindi $M \in f(K)$, ovvero esiste $x_M \in K$ tale che $f(x_M) = M$. Similmente, si prova che f assume minimo su K. □

Osservazione 5.6. Per le funzioni continue a valori reali valgono, con dimostrazione del tutto analoga a quella vista per le funzioni reali di variabile reale, le seguenti proprietà:

a ogni combinazione lineare di funzioni continue è continua;

b il prodotto di funzioni continue è continuo;

c il quoziente di due funzioni continue è continuo nei punti in cui non si annulla il denominatore;

d la funzione composta di due funzioni continue è continua.

È di particolare interesse il caso in cui $X = \mathbb{R}^{\mathrm{d}}$ e $Y = \mathbb{R}$.

Vediamo un esempio in cui la continuità di una funzione in un punto del suo dominio non è banale.

Esempio 5.7. 1. Sia $f : \mathbb{R}^2 \to \mathbb{R}$, definita da

$$f(x,y) := \begin{cases} \dfrac{xy}{x^2+y^2} & \text{se } (x,y) \neq (0,0) \\ 0 & \text{se } (x,y) = (0,0) \end{cases}$$

Le restrizioni di f alle rette di equazione $x = 0$ e $y = 0$ sono continue in $(0,0)$, ma non lo sono le restrizioni alle rette di equazione $y = mx$, per $m \neq 0$, dato che, per $x \neq 0, f(x, mx) = \frac{m}{m^2+1} \neq 0$. Dunque f non è continua in $(0,0)$.

Osserviamo che $\lim_{x\to 0} f(x, mx) = \frac{m}{m^2+1}$ varia con m.

2. Definiamo ora la funzione $g : \mathbb{R}^2 \to \mathbb{R}$, definita da

$$g(x,y) := \begin{cases} \dfrac{x^2 y}{x^4+y^2} & \text{se } (x,y) \neq (0,0) \\ 0 & \text{se } (x,y) = (0,0) \end{cases}$$

Avviciniamoci ora al punto $(0,0)$ seguendo la curva $y = mx$, $m \in [0,\infty]$, si ha che

$$\lim_{x\to 0} g(x, mx) = 0.$$

Avviciniamoci ora allo zero seguendo la curva $y = qx^2$, $q \in (0,\infty)$, otteniamo che

$$\lim_{x\to 0} g\left(x, qx^2\right) = \frac{q}{1+q^2} \neq 0.$$

Il precedente esempio suggerisce l'utilità di considerare le restrizioni a determinate curve per provare che non esiste il limite di una funzione in un punto e quindi che la funzione non è continua in quel punto.

Definizione 5.8 (Funzione vettoriale di variabili reali)**.** Una funzione

$$\begin{aligned} f : \quad & E \subset \mathbb{R}^{\mathrm{d}} & \to \quad & \mathbb{R}^{\mathrm{M}} \\ & x = (x_1, \dots, x_{\mathrm{d}}) & \mapsto \quad & \left(f_1(x_1,\dots,x_{\mathrm{d}}), f_2(x_1,\dots,x_{\mathrm{d}}), \dots, f_{\mathrm{M}}(x_1,\dots,x_{\mathrm{d}})\right) \end{aligned}$$

è detta funzione vettoriale delle d variabili reali $x_1, \dots, x_{\mathrm{d}}$. Le funzioni f_i sono dette funzioni componenti di f. Si verifica immediatamente che f è continua in un punto $x \in E$ se e solo se lo sono le sue componenti f_i.

Osservazione 5.9. Siccome distanze equivalenti determinano le stesse famiglie di intorni, date due distanze d_1 e d_2 in uno stesso insieme X, una successione x_n a valori in X converge a un punto $x_0 \in X$ rispetto alla distanza d_1 se e solo se converge a x_0 rispetto alla distanza d_2. Similmente, data una coppia di distanze equivalenti d_1 e d_2 in X e una coppia di distanze equivalenti d'_1 e d'_2 in Y, una funzione $f : (X, \mathsf{d}_1) \to \left(Y, \mathsf{d}'_1\right)$ è continua in un punto $x \in X$ se e solo se lo è $f : (X, \mathsf{d}_2) \to \left(Y, \mathsf{d}'_2\right)$.

Definizione 5.10 (Funzioni limitate)**.** Dati uno spazio metrico (X, d) e uno spazio normato $(Y, \|\cdot\|)$, si consideri lo spazio vettoriale $\mathcal{B}(X, Y)$ costituito dalle funzioni limitate da X in Y e il suo sottospazio $\mathcal{C} \cap \mathcal{B}(X, Y)$ costituito dalle funzioni continue e limitate da X in Y. Definiamo in $\mathcal{B}(X, Y)$ la norma infinito, o norma del sup, o norma della convergenza uniforme,

$$\|f\|_\infty := \sup_{x \in X} |f(x)|. \tag{5.2}$$

Teorema 5.11. *Se Y è completo allora*

1. $\left(\mathcal{B}(X, Y), \|\cdot\|_\infty\right)$ *è completo;*
2. $\left(\mathcal{C} \cap \mathcal{B}(X, Y), \|\cdot\|_\infty\right)$ *è completo.*

Dimostrazione. 1. Sia f_n una successione di Cauchy in $\mathcal{B}(X, Y)$. Allora

$$\forall \epsilon > 0, \exists \bar{n} \in \mathbb{N} : \forall n, m \geq \bar{n}, \left\|f_n - f_m\right\|_\infty < \epsilon,$$

da cui

$$\forall \epsilon > 0, \exists \bar{n} \in \mathbb{N} : \forall n, m \geq \bar{n}, \forall x \in X, \left\|f_n(x) - f_m(x)\right\| \leq \left\|f_n - f_m\right\|_\infty < \epsilon$$

Dunque, per ogni $x \in X$, la successione $f_n(x)$ è di Cauchy in Y e dunque convergente, essendo Y completo. Definiamo, per ogni $x \in X, f(x) = \lim_{n \to \infty} f_n(x)$. Passando al limite per $m \to \infty$ nella formula precedente, si ha

$$\forall \epsilon > 0, \exists \bar{n} \in \mathbb{N} : \forall n \geq \bar{n}, \forall x \in X, \left\|f_n(x) - f(x)\right\| \leq \epsilon$$

ovvero

$$\forall \epsilon > 0, \exists \bar{n} \in \mathbb{N} : \forall n \geq \bar{n}, \left\|f_n - f\right\|_\infty \leq \epsilon$$

da cui, poichè $f_{\bar{n}}$ è limitata, segue che f è limitata, ovvero che $f \in \mathcal{B}(X, Y)$.

2. Sia f_n una successione di Cauchy in $\mathcal{C} \cap \mathcal{B}(X, Y)$. Alla luce del teorema precedente, sappiamo che esiste $f \in \mathcal{B}(X, Y)$ tale che f_n converge a f in $\|\cdot\|_\infty$, cioè

$$\forall \epsilon > 0, \exists \bar{n} \in \mathbb{N} : \forall n \geq \bar{n}, \left\|f_n - f\right\|_\infty < \epsilon/3$$

Resta da provare che f è continua. Sia $x_0 \in X$. Esprimiamo la continuità di $f_{\bar{n}}$ in x_0 :

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x \in X, \quad \mathsf{d}(x, x_0) < \delta \Rightarrow \left\|f_{\bar{n}}(x) - f_{\bar{n}}(x_0)\right\| < \epsilon/3$$

Allora, per ogni $x \in X$ tale che $\mathsf{d}(x, x_0) < \delta$, si ha $\left\|f(x) - f(x_0)\right\| \leq \left\|f(x) - f_{\bar{n}}(x)\right\| + \left\|f_{\bar{n}}(x) - f_{\bar{n}}(x_0)\right\| + \left\|f_{\bar{n}}(x_0) - f(x_0)\right\| < \epsilon$, quindi f è continua in x_0.

□

Osservazione 5.12. 1. Consideriamo in particolare il caso in cui $X = [a, b], Y = \mathbb{R}$. Allora $\mathcal{C} \cap \mathcal{B}([a, b], \mathbb{R}) = \mathcal{C}([a, b], \mathbb{R})$ è lo spazio vettoriale costituito dalle funzioni reali continue sull'intervallo chiuso $[a, b]$. Per semplificare la notazione lo indicheremo semplicemente con $\mathcal{C}([a, b])$. Osserviamo che l'insieme delle potenze

$$\mathcal{P} := \left\{x^n,\ n \in \mathbb{N}\right\} \subset \mathcal{C}([a, b]), \tag{5.3}$$

è una famiglia linearmente indipendente, quindi $C([a, b])$ è uno spazio vettoriale di dimensione infinita.

2. Definiamo in $\mathcal{C}([a,b])$ il seguente prodotto scalare

$$\langle f \mid g\rangle_2 := \langle f \mid g\rangle_{L^2([a,b];\mathbb{C})} := \int_a^b f(x)\overline{g(x)}\mathrm{d}x$$

con norma indotta (cf. Osservazione 4.56)

$$\|f\|_2 := \sqrt{\int_a^b |f(x)|^2 \mathrm{d}x}$$

Verifichiamo che tale norma non è equivalente alla norma del sup definita in Equazione (5.2). Per ogni $n \geq (b-a)^{-1}$, sia f_n la funzione continua definita da $f_n(a) = 1$, $f_n(x) = 0$ per ogni $x \in \left[a + \frac{1}{n}, b\right]$, f_n lineare nell'intervallo $\left[a, a + \frac{1}{n}\right]$. Si ha che $\|f_n\|_\infty = 1$, per ogni $n \geq (b-a)^{-1}$, mentre

$$\|f_n\|_2 = \left(\int_a^{a+\frac{1}{n}} |f_n(x)|^2 \mathrm{d}x\right)^{\frac{1}{2}} \leq \left(\int_a^{a+\frac{1}{n}} 1 \mathrm{d}x\right)^{\frac{1}{2}} \leq \frac{1}{\sqrt{n}}.$$

Allora f_n converge alla funzione identicamente nulla in $\|\cdot\|_2$, ma non in $\|\cdot\|_\infty$, dunque le due norme non sono equivalenti.

5.1 Funzioni uniformemente continue

Definizione 5.13 (Uniforme continuità)**.** Un'applicazione tra spazi metrici $f : (X, \mathrm{d}_X) \to (Y, \mathrm{d}_Y)$ si dice *uniformemente continua* se

$$\forall \epsilon > 0, \exists \delta > 0 \mid \forall (x, y) \in X^2, \quad \mathrm{d}_X(x, y) < \delta \Rightarrow \mathrm{d}_Y(f(x), f(y)) < \epsilon. \tag{5.4}$$

Osservazione 5.14. Comparando le definizioni nelle Equazioni (5.1) e (5.4) notiamo che l'ϵ in Definizione 5.13 *non dipende dai punti* $x, y \in X$, mentre l'ϵ in Definizione 5.1 è una funzione dipendente da x_0.

Esempio 5.15 (Funzione continua e non uniformemente continua nel suo dominio)**.** Sia $f : A = (0,1) \subset \mathbb{R} \to \mathbb{R}$ definita da $f(x) = \frac{1}{x}$. f è continua in A, strettamente decrescente, quindi invertibile con inversa $f^{-1} : f(A) = (1, +\infty) \to (0,1)$, $f^{-1}(y) = \frac{1}{y}$. Dati $x \in A$ e $\epsilon > 0$, si ha $f^{-1}(f(x) - \epsilon, f(x) + \epsilon) = (x_1, x_2)$, con

$$x_1 = f^{-1}(f(x) + \epsilon) = \frac{x}{1 + \epsilon x},$$

e

$$x_2 = \begin{cases} f^{-1}(f(x) - \epsilon) = \frac{x}{1 - \epsilon x}, & \text{se } \frac{1}{x} - \epsilon > 1 \\ 1, & \text{se } \frac{1}{x} - \epsilon \leq 1 \end{cases}$$

ovvero $x_2 = \min\left\{1, \frac{x}{1-\epsilon x}\right\}$. Il massimo valore di δ che si può trovare affinché valga

$$\forall y \in A, |x - y| < \delta \Rightarrow |f(x) - f(y)| < \epsilon$$

è $\delta = \min\{x - x_1, x_2 - x\}$. In particolare, $\delta(\epsilon, x) \leq x - x_1 = \frac{\epsilon x^2}{1+\epsilon x} \leq \epsilon x^2$ che tende a zero per x tendente a zero. Dunque non esiste $\delta > 0$ indipendente da $x \in A$ per cui valga (5.4). In altre termini, servirebbe $\inf_{x\in A} \delta(\epsilon, x) = \delta(\epsilon) > 0$, mentre si ha $\inf_{x\in A} \delta(\epsilon, x) \leq \epsilon \inf_{x\in A} x^2 = 0$.

Proposizione 5.16. *Sia $f : (X, \mathrm{d}_X) \to (Y, \mathrm{d}_Y)$ uniformemente continua. Se $(x_n)_n$ una successione di Cauchy in X, allora $(f(x_n))_n$ è una successione di Cauchy in Y.*

Dimostrazione. Vogliamo dimostrare che

$$\forall \epsilon > 0, \exists \bar{n} \in \mathbb{N} : \forall n, m \geq \bar{n}, \mathrm{d}_Y(f(x_n), f(x_m)) < \epsilon$$

Fissato $\epsilon > 0$, sia $\delta > 0$ determinato da (5.4). Siccome $(x_n)_n$ è di Cauchy, si ha che

$$\exists \bar{n} \in \mathbb{N} : \forall n, m \geq \bar{n}, \mathrm{d}_X(x_n, x_m) < \delta,$$

e da (5.4) segue subito la tesi.

□

Osservazione 5.17. Se una funzione è solamente continua la Proposizione 5.16 non è più vera, basti pensare alla successione $(1/n)_{n\geq 1}$ e alla funzione $f(x) := 1/x$ in $(0,1]$.

Proposizione 5.18. *Sia $f : E \subset \mathbb{R}^{\mathrm{d}} \to \mathbb{R}$ uniformemente continua. Sia $B \subset E$, B limitato. Allora $f(B)$ è limitato.*

Dimostrazione. Supponiamo per assurdo che $f(B)$ sia illimitato. Allora esiste una successione $(y_n)_n \subset f(B)$ tale che $\lim_{n\to\infty} |y_n| = +\infty$. Sia $x_n \in B$ tale che $y_n = f(x_n)$. La successione $(x_n)_n$ è limitata e quindi ammette una sottosuccesione x_{n_k} convergente in $\mathbb{R}^{\mathrm{d}}$ e dunque di Cauchy. Dal teorema 5.3, segue che $y_{n_k} = f(x_{n_k})$ è di Cauchy, quindi limitata, da cui si ha una contraddizione. □

Teorema 5.19 (di Heine-Cantor)**.** *Sia $f : K \subset (X, \mathsf{d}_X) \to (Y, \mathsf{d}_Y)$ continua in K, con K sequenzialmente compatto. Allora f è uniformemente continua in K.*

Dimostrazione. Procediamo per assurdo, negando la tesi (5.4):

$$\exists \epsilon > 0 : \forall \delta > 0, \exists x, y \in K \text{ per cui } \mathsf{d}_X(x, y) < \delta \text{ e } \mathsf{d}_Y(f(x), f(y)) \geq \epsilon$$

Scegliendo in particolare $\delta = 1/n$ per $n \in \mathbb{N}$, si ha che esistono due successioni $(x_n)_n, (y_n)_n \subset K$ tali che $\mathsf{d}_X(x_n, y_n) < 1/n$ e $\mathsf{d}_Y(f(x_n), f(y_n)) \geq \epsilon$. Essendo K sequenzialmente compatto, esistono una sottosuccessione x_{n_k} di x_n e un punto $x_0 \in K$ tali che $\lim_{k\to\infty} x_{n_k} = x_0$. Dalla continuità di f, segue che $\lim_{k\to\infty} f(x_{n_k}) = f(x_0)$. D'altra parte $\mathsf{d}_X(y_{n_k}, x_0) \leq \mathsf{d}_X(x_{n_k}, y_{n_k}) + \mathsf{d}_X(x_{n_k}, x_0) < 1/n_k + \mathsf{d}_X(x_{n_k}, x_0)$, da cui $\lim_{k\to\infty} y_{n_k} = x_0$. Allora, sempre per la continuità di f, si ha $\lim_{k\to\infty} f(y_{n_k}) = f(x_0)$, e perciò $\lim_{k\to\infty} \mathsf{d}_Y(f(y_{n_k}), f(x_{n_k})) = 0$, in contraddizione con $\mathsf{d}_Y(f(y_{n_k}), f(x_{n_k})) \geq \epsilon$. □

Definizione 5.20 (Funzioni Hölderiane)**.** Sia $\alpha \in (0,1]$. Una funzione $f : (X, \mathsf{d}_X) \to (Y, \mathsf{d}_Y)$ si dice *Hölderiana* di ordine α (su X) se esiste una costante positiva L (detta costante di Hölder) tale che

$$\mathsf{d}_Y(f(x), f(y)) \leq L(\mathsf{d}_X(x, y))^{\alpha}, \quad \forall x, y \in X.$$

Denotiamo con $\mathcal{C}^{0,\alpha}(X; Y)$ lo spazio delle funzioni Hölderiane da X a Y, e su di esso definiamo la seminorma (cf. Definizione 4.45)

$$|f|_{\mathcal{C}^{0,\alpha}(X;Y)} := \sup_{x,y\in X,\ x\neq y} \frac{\mathsf{d}_Y(f(x), f(y))}{(\mathsf{d}_X(x, y))^{\alpha}},$$

e la norma

$$\|f\|_{\mathcal{C}^{0,\alpha}(X;Y)} := \|f\|_{\infty} + |f|_{\mathcal{C}^{0,\alpha}(X;Y)}.$$

Si ha che $\left(\mathcal{C}^{0,\alpha}(X; Y), \|\cdot\|_{\mathcal{C}^{0,\alpha}(X;Y)}\right)$ è uno spazio di Banach.
Se $\alpha = 1$ si dice che la funzione è *Lipschitziana*

Esercizio 5.21.

1. Si fissi $\alpha \in (0,1]$, si dimostri che la funzione $x \mapsto |x|^{\alpha}$ appartiene allo spazio $\mathcal{C}^{0,\alpha'}([-1,1]; \mathbb{R})$ per ogni $\alpha' \in (0, \alpha]$.

2. Si dimostri che una funzione Hölderiana è sia continua che uniformemente continua.

Definizione 5.22 (Applicazione contrattiva)**.** Un'applicazione da uno spazio metrico in sé $f : (X, \mathsf{d}) \to (X, \mathsf{d})$ è detta *contrazione* se esiste una costante positiva $q \in (0,1)$ tale che

$$\mathsf{d}(f(x), f(y)) \leq q\, \mathsf{d}(x, y), \quad \forall x, y \in X.$$

Definizione 5.23 (Punto fisso)**.** Definizione 5.9. Sia X un insieme non vuoto e $f : X \to X$. Un elemento $x \in X$ si dice punto fisso di f se $f(x) = x$.

Teorema 5.24 (Teorema delle contrazioni di Banach-Caccioppoli)**.** *Sia (X, d) uno spazio metrico completo e sia $T : (X, \mathsf{d}) \to (X, \mathsf{d})$ una contrazione. Allora esiste un unico punto fisso di T.*

Dimostrazione. Sappiamo che esiste $q<1$ tale che $\mathsf{d}\big(Tx,Ty\big)\le q\ \mathsf{d}\big(x,y\big)$, per ogni $x,y\in X$. Sia $x_0\in X$ un qualunque punto di X e definiamo, ricorsivamente, la successione

$$x_{n+1}:=Tx_n=T^{n+1}x_0, \qquad\qquad n\in\mathbb{N},$$

Proviamo, per induzione, che

$$\mathsf{d}\left(x_{n+1},x_n\right)\le q^n\mathsf{d}\left(x_1,x_0\right). \tag{5.5}$$

Per $n=1$, si ha $\mathsf{d}\left(x_2,x_1\right)=\mathsf{d}\left(Tx_1,Tx_0\right)\le q\mathsf{d}\left(x_1,x_0\right)$. Proviamo ora il passo induttivo, $\mathsf{d}\left(x_{n+1},x_n\right)=\mathsf{d}\left(Tx_n,Tx_{n-1}\right)\le q\mathsf{d}\left(x_n,x_{n-1}\right)\le q^n\mathsf{d}\left(x_1,x_0\right)$, provando (5.5).

Sia ora $m>n$. Si ha che

$$\begin{aligned}\mathsf{d}\left(x_m,x_n\right)\le\sum_{i=n}^{m-1}\mathsf{d}\left(x_{i+1},x_i\right)\le\sum_{i=n}^{m-1}q^i\mathsf{d}\left(x_1,x_0\right)&=\mathsf{d}\left(x_1,x_0\right)q^n\sum_{j=0}^{m-1-n}q^j\\&=\mathsf{d}\left(x_1,x_0\right)q^n\frac{1-q^{m-n}}{1-q}\le\mathsf{d}\left(x_1,x_0\right)\frac{q^n}{1-q}\xrightarrow{n\to\infty}0.\end{aligned}$$

Quindi per ogni $\epsilon>0$ esiste $n_\epsilon\in\mathbb{N}$ tale che

$$\mathsf{d}\left(x_m,x_n\right)<\epsilon,\quad\forall n,m\ge n_\epsilon,$$

cioè la successione x_n è di Cauchy. Dalla completezza di X, segue che essa converge a un punto $x\in X$. Essendo una contrazione, T è continua e quindi Tx_n converge a Tx, ma $Tx_n=x_{n+1}$ e quindi converge a x. Per l'unicità del limite in spazi metrici, $Tx=x$ e perciò x è punto fisso di T. Se, per assurdo, esistesse un altro punto fisso $y\in X, y\neq x$, allora si avrebbe

$$\mathsf{d}\big(x,y\big)=\mathsf{d}\big(Tx,Ty\big)\le q\ \mathsf{d}\big(x,y\big) \tag{5.6}$$

Essendo $q<1$ e $\mathsf{d}\big(x,y\big)>0$, (5.6) non ha soluzione, pertanto è assurdo assumere $x\neq y$. □

6 Calcolo differenziale

6.1 Derivate direzionali e differenziabilità

Notazione 6.1. In questa sezione ogni qual volta appaia un concetto metrico lo si considera relativamente allo spazio metrico Euclideo $\big(\mathbb{R}^{\mathrm{d}},\mathsf{d}_2\big)$, salvo eccezioni esplicitamente menzionate.

Notazione 6.2. Sia $(X,\|\cdot\|_X)$ uno spazio di Banach. In questa sezione denoteremo con Ω un sottoinsieme aperto, rispetto alla topologia indotta da $\|\cdot\|_X$, di X.

Definizione 6.3 (Differenziabilità secondo Gateaux)**.** Siano $(X,\|\cdot\|_X)$ e $(Y,\|\cdot\|_Y)$ due spazi di Banach (cf. Definizione 4.48). Sia $f:\Omega\subset X\to Y$ e sia $x^0\in\Omega$. Diremo che f è differenziabile in x^0 nella direzione $h\in X$ se esiste finito in Y il limite

$$\lim_{\varepsilon\to0}\frac{f\left(x^0+\varepsilon h\right)-f\left(x^0\right)}{\varepsilon}. \tag{6.1}$$

Tale limite è detto *derivata direzionale di f nella direzione h*, è indicato con

$$\frac{\partial f}{\partial h}\left(x^0\right),\qquad\qquad \partial_h f\left(x^0\right),\qquad\qquad \mathrm{d_G}f\left(x^0;h\right). \tag{6.2}$$

La funzione f si dice *Gateaux differenziabile o G-differenziabile in x^0* se il limite in (6.1) esiste finito per ogni $h\in X$.

Notazione 6.4. Nel caso in cui $X=\mathbb{R}^{\mathrm{d}}$ e $Y=\mathbb{R}$ useremo le prime due notazioni di (6.1), altrimenti useremo la terza.

Esempio 6.5 (Esempio di calcolo di derivata direzionale)**.** 1. Se $(X,Y)=(\mathbb{R},\mathbb{R})$ il differenziale di Gateaux in (6.1) coincide con il concetto di derivata;

2. $(X,Y)=\left(\mathbb{R}^{\mathtt{d}},\mathbb{R}\right)$, per esempio $\mathtt{d}=3$. Consideriamo la funzione $f\left(x,y,z\right):=\frac{yz}{1+x^2}$. Fissiamo $v:=(v_1,v_2,v_3)\in\mathbb{R}^3$ ed $\epsilon>0$, calcoliamo

$$\begin{aligned}
f\left(x+\epsilon v_1,y+\epsilon v_2,z+\epsilon v_3\right)-f\left(x,y,z\right)&\\
&=\frac{\left(y+\epsilon v_2\right)\left(z+\epsilon v_3\right)}{1+\left(x+\epsilon v_1\right)^2}-\frac{yz}{1+x^2}\\
&=\frac{\epsilon^2 v_2v_3x^2+\epsilon^2 v_2v_3+\left(\epsilon v_3x^2+\epsilon v_3\right)y+\left(\epsilon v_2x^2+\epsilon v_2-\left(\epsilon^2v_1^2+2\epsilon v_1x\right)y\right)z}{2\epsilon v_1x^3+\epsilon^2v_1^2+x^4+2\epsilon v_1x+\left(\epsilon^2v_1^2+2\right)x^2+1}.
\end{aligned}$$

Pertanto ottengo che

$$\begin{aligned}
\frac{f\left(x+\epsilon v_1,y+\epsilon v_2,z+\epsilon v_3\right)-f\left(x,y,z\right)}{\epsilon}&=\frac{v_3\left(x^2+1\right)y+\left(v_2\left(x^2+1\right)-2v_1xy\right)z+O(\epsilon)}{\left(1+x^2\right)^2+O(\epsilon)}\\
&\xrightarrow{\epsilon\to0}-\frac{2v_1xyz}{\left(x^2+1\right)^2}+\frac{v_3y}{x^2+1}+\frac{v_2z}{x^2+1}=\partial_vf\left(x,y,z\right). \qquad (6.3)
\end{aligned}$$

3. $(X,Y)=\left(\mathbb{R}^{\mathtt{N}},\mathbb{R}^{\mathtt{M}}\right)$ (per esempio $\mathtt{N}=3,\mathtt{M}=2$). In questo caso si ha che

$$f\left(x,y,z\right)=\begin{pmatrix}f_1\left(x,y,z\right)\\ f_2\left(x,y,z\right)\end{pmatrix}.$$

Possiamo quindi ragionare componente per componente analogamente all'esempio precedente.

Esempio 6.6 (Calcolo di derivate direzionali tra spazi di dimensione infinita)**.** 1. Sia $X=\mathcal{C}\left([0,1];\mathbb{R}\right)$ con la norma del sup (cf. (5.2)), la coppia $(X,\|\cdot\|_\infty)$ è di Banach come dimostrato nel Teorema 5.11. Consideraimo la funzione

$$\begin{array}{rccc} f & : & B_X(0,1) & \to & X\\ & & u(x) & \mapsto & \dfrac{1}{1+u(x)} \end{array}.$$

È chiaro che $f(u)\in X$ solamente se $\|u\|_\infty<1$ (altrimenti si ha un punto d'esplosione). Sia $h\in X$, calcoliamo

$$\begin{aligned}
\lim_{\epsilon\to0}\frac{f(u+\epsilon h)-f(u)}{\epsilon}(x)=\lim_{\epsilon\to0}\frac{\frac{1}{1+u(x)+\epsilon h(x)}-\frac{1}{1+u(x)}}{\epsilon}&=\lim_{\epsilon\to0}\frac{-h(x)}{(1+u(x)+\epsilon h(x))(1+u(x))}\\
&=-\frac{1}{(1+u(x))^2}\,h(x)=\mathrm{d_G}f(u;h)(x). \qquad (6.4)
\end{aligned}$$

2. Consideriamo lo spazio, $\mathcal{C}^1([0,1];\mathbb{R})=\mathcal{C}^1$. La norma

$$\|u\|_{\mathcal{C}^1}:=\|u\|_\infty+\left\|u'\right\|_\infty,$$

rende $\mathcal{C}^1$ uno spazio di Banach. Consideriamo ora l'*operatore derivata*

$$\begin{array}{rccc} D & : & \mathcal{C}^1 & \to & \mathcal{C}^0\\ & & u & \mapsto & u' \end{array}.$$

La mappa D è una mappa tra spazi di Banach, calcoliamo la sua derivata come mappa tra spazi di Banach: Sia $h\in\mathcal{C}^1$

$$\frac{D(u+\epsilon h)-Du}{\epsilon}=h',$$

pertanto $\mathrm{d_G}D(u;h)=h'$.

Osservazione 6.7. Dal punto di vista applicativo il calcolo della derivata direzionale in (6.1) non è molto efficiente. Sia $\Omega \subset X$, $f : \Omega \to Y$ G-differenziabile e tale che esiste un $h \in X$ ed un $\epsilon_0 > 0$ tali che l'applicazione

$$\epsilon \in [0, \epsilon_0] \mapsto \mathrm{d_G} f\left(x^0 + \epsilon h; h\right) \in Y,$$

è continua in zero. Allora

$$\mathrm{d_G} f\left(x^0; h\right) = \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon} f\left(x^0 + \varepsilon h\right)\right|_{\varepsilon=0}. \tag{6.5}$$

Dimostriamo l'identità in (6.5), definendo $g(\varepsilon) := f\left(x^0 + \varepsilon h\right)$ ho che $\mathrm{d_G} f\left(x^0; h\right) = g'(0)$ e $\frac{\mathrm{d}}{\mathrm{d}\varepsilon} f\left(x^0 + \varepsilon h\right) = g'(\varepsilon)$. Se posso calcolare $g'(0)$ come limite per $\varepsilon \to 0$ di $g'(\varepsilon)$ ho quindi concluso. Calcoliamo

$$g'(\varepsilon) = \lim_{\eta \to 0} \frac{f\left(x^0 + (\varepsilon + \eta) h\right) - f\left(x^0 + \varepsilon h\right)}{\eta} = \mathrm{d_G} f\left(x^0 + \varepsilon h; h\right),$$

che per ipotesi è continua, passo dunque al limite per $\varepsilon \to 0$ e concludo.

Esempio 6.8. Utlizziamo l'Osservazione 6.7 per calcolare il differenziale di Gateaux negli Esempi 6.5 e 6.6.

- $f\left(x, y, z\right) := \frac{yz}{1+x^2}$, calcoliamo

$$\begin{aligned}
\lim_{\varepsilon \to 0} \frac{\mathrm{d}}{\mathrm{d}\varepsilon} f\left(x + \varepsilon v_1, y + \varepsilon v_2, z + \varepsilon v_3\right) &= \lim_{\varepsilon \to 0} \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \frac{\left(y + \varepsilon v_2\right)(z + \varepsilon v_3)}{1 + (x + \varepsilon v_1)^2} \\
&= \lim_{\varepsilon \to 0} \left[-\frac{2(\varepsilon v_1 + x)\left(\varepsilon v_2 + y\right)(\varepsilon v_3 + z) v_1}{\left((\varepsilon v_1 + x)^2 + 1\right)^2} + \frac{(\varepsilon v_3 + z) v_2}{(\varepsilon v_1 + x)^2 + 1} + \frac{\left(\varepsilon v_2 + y\right) v_3}{(\varepsilon v_1 + x)^2 + 1} \right] \\
&= -\frac{2 v_1 x y z}{\left(x^2 + 1\right)^2} + \frac{v_3 y}{x^2 + 1} + \frac{v_2 z}{x^2 + 1},
\end{aligned}$$

  che coincide con l'espressione in (6.3).

- $f(u)(x) := \frac{1}{1+u(x)}$, calcoliamo

$$\begin{aligned}
\lim_{\varepsilon \to 0} \frac{\mathrm{d}}{\mathrm{d}\varepsilon} f(u + \varepsilon h)(x) = \lim_{\varepsilon \to 0} \frac{\mathrm{d}}{\mathrm{d}\varepsilon} \frac{1}{1 + u(x) + \varepsilon h(x)} &= -\lim_{\varepsilon \to 0} \frac{1}{(1 + u(x) + \varepsilon h(x))^2} h(x) \\
&= -\frac{1}{(1 + u(x))^2} h(x),
\end{aligned}$$

  che coincide con (6.4).

Definizione 6.9 (Derivata parziale)**.** Sia $(X, Y) = \left(\mathbb{R}^{\mathrm{d}}, \mathbb{R}\right)$ e $h := e_i$, con e_i vettore della base canonica di $\mathbb{R}^{\mathrm{d}}$, la derivata direzionale $\frac{\partial f}{\partial e_i}\left(x^0\right)$ è detta derivata parziale rispetto alla variabile x_i e si usano equivalentemente le seguenti notazioni

$$\frac{\partial f}{\partial x_i}\left(x^0\right), \qquad f_{x_i}\left(x^0\right), \qquad \partial_i f\left(x^0\right), \qquad \partial_{x_i} f\left(x^0\right).$$

Esempio 6.10 (Calcolo esplicito di una derivata parziale)**.** Data una funzione $f : \Omega \subset \mathbb{R}^{\mathrm{d}} \to \mathbb{R}$ per calcolare la derivata parziale $\partial_i f$ è sufficiente derivare nella variabile x_i considerando le restanti variabili come se fossero costanti. Vediamo degli esempi

1. $f\left(x, y\right) := xy$. Si ha che $\partial_x f\left(x, y\right) = y$ e $\partial_y f\left(x, y\right) = x$.
2. $f\left(x, y, z\right) := \log\left(x^2 + y^3 + z^4\right)$. Si ha che (nel dominio della funzione)

$$\partial_x f\left(x, y, z\right) = \frac{2x}{x^2 + y^3 + z^4}, \qquad \partial_y f\left(x, y, z\right) = \frac{3y^2}{x^2 + y^3 + z^4}, \qquad \partial_z f\left(x, y, z\right) = \frac{4z^3}{x^2 + y^3 + z^4}.$$

Un funzione può essere G-differenziabile in un punto x^0 ma non continua, come illustrato dal seguente esempio:

Esempio 6.11 (G-differenzabilità non implica la continuità)**.** Sia $f:\mathbb{R}^2\to\mathbb{R}$, definita da

$$f\left(x,y\right)=\begin{cases}\dfrac{x^4y^2}{\left(x^4+y^2\right)^2}, & \text{se } \left(x,y\right)\neq(0,0)\\ 0, & \text{se } \left(x,y\right)=(0,0)\end{cases} \tag{6.6}$$

Per ogni direzione $v\in\mathbb{R}^2$, esiste la derivata direzionale

$$\frac{\partial f}{\partial v}(0,0)=\lim_{t\to0}\frac{f\left(tv_1,tv_2\right)}{t}=\lim_{t\to0}\frac{tv_1^4v_2^2}{\left(t^2v_1^4+v_2^2\right)^2}=0.$$

Consideriamo la successione $z_n=\left(\frac{1}{n},\frac{1}{n^2}\right)$, che converge a $(0,0)$. Si ha che $f\left(z_n\right)\equiv\frac{1}{4}\neq f(0,0)=0$. Dunque f non è continua nell'origine.

Definizione 6.12 (Operatori lineari limitati)**.** Dati due $\mathbb{K}$-spazi vettoriali normati X e Y denotiamo con $\mathcal{L}\left(X;Y\right)$ l'insieme delle applicazioni lineari da X a Y tali che, dato $T\in\mathcal{L}\left(X;Y\right)$, esiste $C>0$ tale che per ogni $x\in X$ si ha che $\|Tx\|_Y\leq C\|x\|_X$. Lo spazio $\mathcal{L}\left(X;Y\right)$ dotato della norma

$$\|T\|_{\mathcal{L}(X;Y)}:=\sup_{\|x\|_X=1}\|Tx\|_Y,$$

è uno spazio di Banach (cf. Definizione 4.48). Se $X=Y$ scriviamo $\mathcal{L}\left(X\right)$.

Osservazione 6.13. Dato $x\in X$ e $T\in\mathcal{L}\left(X;Y\right)$, sia $\hat{x}:=\frac{x}{\|x\|_X}\in\mathbb{S}_X$ notiamo che

$$\|Tx\|_Y=\|\|x\|_X\ T\hat{x}\|_Y=\|x\|_X\|T\hat{x}\|_Y\leq\|T\|_{\mathcal{L}(X;Y)}\|x\|_X.$$

Esempio 6.14 (Esempi canonici di operatori limitati)**.**

1. Caso $(X,Y)=\left(\mathbb{R}^{\mathtt{N}},\mathbb{R}^{\mathtt{M}}\right)$. Sappiamo che tutti e soli gli operatori da $\mathbb{R}^{\mathtt{N}}$ in $\mathbb{R}^{\mathtt{M}}$ sono della forma

   $$Tx=Ax,\qquad\qquad A\in\mathbb{R}^{\mathtt{M}\times\mathtt{N}}.$$

   Verifichiamo che l'applicazione $x\mapsto Ax$ è limitata come da Definizione 6.12.

   $$\|Ax\|_{\mathbb{R}^{\mathtt{M}}}=\sqrt{\sum_{i=1}^{\mathtt{M}}(Ax)_i^2}=\sqrt{\sum_{i=1}^{\mathtt{M}}\left(\sum_{j=1}^{\mathtt{N}}A_{ij}x_j\right)^2}\leq\mathtt{N}\sqrt{\mathtt{M}}\max_{\substack{i=1,\dots,\mathtt{M}\\ j=1,\dots,\mathtt{N}}}\left|A_{ij}\right|\max_{j=1,\dots,\mathtt{N}}\left|x_j\right|.$$

   Da (4.7) otteniamo che $\max_{j=1,\dots,\mathtt{N}}\left|x_j\right|\leq\|x\|_{\mathbb{R}^{\mathtt{N}}}$ e definendo $C_A:=\mathtt{N}\sqrt{\mathtt{M}}\max_{\substack{i=1,\dots,\mathtt{M}\\ j=1,\dots,\mathtt{N}}}\left|A_{ij}\right|$ otteniamo

   $$\|Ax\|_{\mathbb{R}^{\mathtt{M}}}\leq C_A\|x\|_{\mathbb{R}^{\mathtt{N}}}.$$

   Pertanto l'applicazione è limitata. In particolare se $\mathtt{N}=\mathtt{M}=1$ ho che

   $$T\in\mathcal{L}\left(\mathbb{R}\right)\qquad\Leftrightarrow\qquad Tx=\lambda x,\ \lambda\in\mathbb{R},\quad\forall x\in\mathbb{R}.$$

2. Caso infinito-dimensionale: l'operatore di moltiplicazione. Sia $\Omega\subset\mathbb{R}^{\mathrm{d}}$ e $X=Y=\mathcal{B}\left(\Omega;\mathbb{R}\right)$. Fissiamo $U_0\in X$ e definiamo l'applicazione

   $$\begin{array}{cccc}\mathsf{M}_{U_0} & : & X & \to & X\\ & & v & \mapsto & U_0\,v\end{array}.$$

   L'applicazione M_{U_0} è certamente lineare, verifichiamo che è limitata. Si ha che per ogni $v\in X$

   $$\left\|\mathsf{M}_{U_0}v\right\|_X=\sup_{x\in\Omega}\left|U_0\left(x\right)\ v\left(x\right)\right|\leq\sup_{x\in\Omega}\left|U_0\left(x\right)\right|\sup_{x\in\Omega}\left|v\left(x\right)\right|=\|U_0\|_X\|v\|_X.$$

Esempio 6.15 (Esempio di applicazione lineare e non limitata)**.** Consideriamo lo spazio di Banach $(\mathcal{C}([0,1];\mathbb{R}), \|\cdot\|_\infty)$. Consideriamo l'operatore di differenziazione $D : \mathcal{C}([0,1];\mathbb{R}) \to \mathcal{C}([0,1];\mathbb{R})$. Ovviamente si ha che per ogni $\delta > 0$

$$\left\| D\left(\delta x^n\right)\right\|_\infty = \sup_{x\in[0,1]} \left|\delta n\, x^{n-1}\right| = \delta n \left\|x^n\right\|_\infty . \tag{6.7}$$

Pertanto l'operatore di differenziazione non soddisfa la Definizioni 6.12, ed in particolare la stima in Equazione (6.7) mostra come l'operatore di differenziazione *non è continuo nello zero*, infatti manda elementi arbitrariamente piccoli di $(\mathcal{C}([0,1];\mathbb{R}), \|\cdot\|_\infty)$ in elementi arbitrariamente grandi di $(\mathcal{C}([0,1];\mathbb{R}), \|\cdot\|_\infty)$.

Definizione 6.16 (Differenziabilità secondo Fréchet)**.** Siano $(X, \|\cdot\|_X)$ e $(Y, \|\cdot\|_Y)$ due spazi di Banach (cf. Definizione 4.48). Sia $f : \Omega \subset X \to Y$ e sia $x^0 \in \Omega$ e $\delta > 0$ tale che $B\left(x^0,\delta\right) \subset \Omega$. La funzione f si dice *(Fréchet) differenziabile in* x^0 se esiste un'applicazione lineare $L \in \mathcal{L}(X;Y)$ tale che

$$\left\| f\left(x^0+h\right) - f\left(x^0\right) - Lh\right\|_Y = o\left(\|h\|_X\right), \qquad h \in B_X(0,\delta). \tag{6.8}$$

- Il funzionale lineare L si chiama *differenziale* di f in x^0 nella direzione h e si indica con $\mathrm{d}f\left(x^0\right)h$ oppure $\mathrm{d}f\left(x^0\right)[h]$;
- La quantità $f\left(x^0\right) + \mathrm{d}f\left(x^0\right)[h]$ si dice *approssimante lineare di f in x^0*;
- Denotiamo con $R_f\left(x^0;h\right) = R\left(x^0;h\right) := f\left(x^0+h\right) - f\left(x^0\right) - \mathrm{d}f\left(x^0\right)[h]$ la parte di resto superlineare in h.

Se la mappa $x^0 \in \Omega \mapsto \mathrm{d}f\left(x^0\right)$ è un'applicazione continua tra gli spazi normati $(X, \|\cdot\|_X)$ e $\left(\mathcal{L}(X;Y), \|\cdot\|_{\mathcal{L}(X;Y)}\right)$ diremo che $f \in \mathcal{C}^1(\Omega;Y)$.

Esercizio 6.17. Siano X, Y di Banach e $\Omega \subset X$, dimostrare che l'applicazione

$$\begin{array}{cccc} \mathrm{d} & : & \mathcal{C}^1(\Omega;Y) & \to & \{\Omega \to \mathcal{L}(X;Y)\} \\ & & f & \mapsto & \left\{x^0 \mapsto \mathrm{d}f\left(x^0\right)\right\} \end{array},$$

è lineare.

Osservazione 6.18. Compariamo le definizioni di differenziabilità date nelle Definizioni 6.3 e 6.16 rispettivamente. Notiamo che il differenziale di Gateaux *non è detto che sia un operatore lineare* come da Definizione 6.12. Semplici calcoli mostrano che è sì una funzione omogenea, ossia $\mathrm{d_G}f\left(x^0;\lambda h\right) = \lambda\, \mathrm{d_G}f\left(x^0;h\right)$, tuttavia non è additiva. Per tale ragione la notazione per il differenziale di Gateux è diversa da quella del differenziale di Fréchet. Prendiamo per esempio $f : \mathbb{R}^2 \to \mathbb{R}$, definita da

$$f(x,y) = \begin{cases} \dfrac{x^3}{x^2+y^2}, & \text{se } (x,y) \neq (0,0) \\ 0, & \text{se } (x,y) = (0,0) \end{cases}$$

Un calcolo esplicito mostra che

$$\partial_{(a,b)} f(0,0) = \begin{cases} \dfrac{a^3}{a^2+b^2}, & \text{se } (a,b) \neq (0,0) \\ 0, & \text{se } (a,b) = (0,0) \end{cases},$$

che non è lineare in (a,b).

Un'altra fondamentale differenza è evidenziata dal seguente risultato immediato, che stabilisce che la differenziabilità secondo Fréchet è più forte della continuità.

Lemma 6.19. *Sia f come nella Definizione 6.16. Se f è differenziabile in x^0, allora è continua in x^0.*

Dimostrazione. Se f è differenziabile in x^0 allora per ogni $x \in B_X\left(x^0,\delta\right)$, con $\delta > 0$ sufficientemente piccolo, si ha che

$$f(x) = f\left(x^0\right) + \mathrm{d}f\left(x^0\right)\left[x-x^0\right] + R_f\left(x^0;x-x^0\right).$$

Possiamo dunque passare al limite per $x \to x^0$ ed otteniamo la tesi. □

Lemma 6.20. *Una funzione Fréchet differenziabile è Gateaux differenziabile.*

Dimostrazione. La dimostrazione è immediata, sia $x^0 \in \Omega$, $B_X\left(x^0,\delta\right) \subset \Omega$, $h \in X$ e $t \in (0,\delta)$, siccome f è differenziabile in x^0 ho che

$$\frac{f\left(x^0+\varepsilon h\right)-f\left(x^0\right)}{\varepsilon}=\mathrm{d}f\left(x^0\right)h+\frac{R_f\left(x^0;\varepsilon h\right)}{\varepsilon},$$

pertanto passando al limite per $\varepsilon \to 0$ otteniamo il risultato. □

Osservazione 6.21 (Calcolo esplicito del differenziale di Fréchet in casi semplificati)**.** Vediamo ora esplicitamente come si calcola il differenziale di Fréchet in alcuni casi specifici:

1. Sia $f:\Omega\subset\mathbb{R}^{\mathrm{d}}\to\mathbb{R}$ differenziabile secondo Fréchet in Ω. Vogliamo dimostrare che

$$\mathrm{d}f\left(x^0\right)[v]=\frac{\partial f}{\partial v}\left(x^0\right)=\sum_{i=1}^{\mathrm{d}} v_i\frac{\partial f}{\partial x_i}\left(x^0\right). \tag{6.9}$$

   Sia $v \in \mathbb{R}^{\mathrm{d}}$. Ponendo $h = tv$ in (6.8) (pertanto assumiamo t "piccolo"), si ha, dalla definizione di differenziabilità

$$\lim_{t\to 0}\frac{\left|f\left(x^0+tv\right)-f\left(x^0\right)-\mathrm{d}f\left(x^0\right)[tv]\right|}{|t|}=0$$

   e quindi anche

$$\lim_{t\to 0}\frac{f\left(x^0+tv\right)-f\left(x^0\right)-\mathrm{d}f\left(x^0\right)[tv]}{t}=0$$

   Dato che $\mathrm{d}f\left(x^0\right)[tv]=t\,\mathrm{d}f\left(x^0\right)[v]$, si ha che esiste la derivata direzionale

$$\frac{\partial f}{\partial v}\left(x^0\right)=\lim_{t\to 0}\frac{f\left(x^0+tv\right)-f\left(x^0\right)}{t}=\mathrm{d}f\left(x^0\right)[v]$$

   Dalla linearità di $\mathrm{d}f\left(x^0\right)$ segue che

$$\frac{\partial f}{\partial v}\left(x^0\right)=\mathrm{d}f\left(x^0\right)\left[\sum_{i=1}^{\mathrm{d}} v_ie_i\right]=\sum_{i=1}^{\mathrm{d}} v_i\mathrm{d}f\left(x^0\right)[e_i]=\sum_{i=1}^{\mathrm{d}} v_i\frac{\partial f}{\partial x_i}\left(x^0\right),$$

   provando (6.9). Notiamo che la formula (6.9) presenta vari punti interessanti:

   (a) ci da una maniera esplicita per calcolare il differenziale *di una funzione differenziabile*. Notiamo in particolare che tale formula può venire applicata pure a funzioni che sono solo G-differenziabili, come in (6.6), che tuttavia non essendo continua non dà un differenziale di Fréchet, cf. Lemma 6.19;

   (b) il vettore

$$\nabla f\left(x^0\right):=\begin{pmatrix}\partial_1 f\left(x^0\right)\\ \vdots\\ \partial_{\mathrm{d}} f\left(x^0\right)\end{pmatrix}, \tag{6.10}$$

   è il *vettore gradiente* di una funzione $f:D\subset\mathbb{R}^{\mathrm{d}}\to\mathbb{R}$.

2. Consideriamo ora $f:\Omega\subset\mathbb{R}^{\mathrm{N}}\to\mathbb{R}^{\mathrm{M}}$ con $f(x)=\left(f_1(x_1,\ldots,x_{\mathrm{N}}),\ldots,f_{\mathrm{M}}(x_1,\ldots,x_{\mathrm{N}})\right)^{\mathsf{T}}$, $x^0\in\Omega$, definiamo il *Jacobiano di f*

$$Jf\left(x^0\right):=\begin{pmatrix}\nabla f_1\left(x^0\right)^{\mathsf{T}}\\ \vdots\\ \nabla f_{\mathrm{M}}\left(x^0\right)^{\mathsf{T}}\end{pmatrix}=\begin{pmatrix}\partial_1 f_1\left(x^0\right) & \cdots & \partial_{\mathrm{N}} f_1\left(x^0\right)\\ \vdots & & \vdots\\ \partial_1 f_{\mathrm{M}}\left(x^0\right) & \cdots & \partial_{\mathrm{N}} f_{\mathrm{M}}\left(x^0\right)\end{pmatrix}, \tag{6.11}$$

   vogliamo dimostrare che

$$\mathrm{d}f\left(x^0\right)=Jf\left(x^0\right),$$

ossia che

$$R_f\left(x^0;h\right) := f\left(x^0+h\right) - f\left(x^0\right) - Jf\left(x^0\right)h, \qquad \left\|R_f\left(x^0;h\right)\right\|_{\mathbb{R}^{\mathtt{M}}} = o\left(\|h\|_{\mathbb{R}^{\mathtt{N}}}\right). \tag{6.12}$$

Consideriamo l'identità in (6.12) per componenti, da (6.11) otteniamo che per ogni $i = 1,\ldots,\mathtt{N}$

$$R_f\left(x^0;h\right)_i = f_i\left(x^0+h\right) - f_i\left(x^0\right) - \nabla f_i\left(x^0\right)\cdot h,$$

tuttavia utilizzando il punto 1 qui sopra sappiamo che $\mathrm{d}f_i\left(x^0\right) = \nabla f_i\left(x^0\right)$ e pertanto $\left|R_f\left(x^0;h\right)_i\right| = o\left(\|h\|_{\mathbb{R}^{\mathtt{N}}}\right)$. L'identità in (6.12) segue da

$$\left\|R_f\left(x^0;h\right)\right\|_{\mathbb{R}^{\mathtt{M}}} = \sqrt{\sum_{i=1}^{\mathtt{M}}\left|R_f\left(x^0;h\right)_i\right|^2}$$

Definizione 6.22. Il punto x^0 si dice critico o stazionario per f se $\nabla f\left(x^0\right)$ è il vettore nullo; altrimenti si dice regolare.

Osservazione 6.23. La formula (6.10) ci permette di calcolare tutte le derivate direzionali a partire dalla conoscenza delle derivate parziali e ci dice in particolare che se x^0 è un punto critico allora tutte le derivate direzionali in x^0 sono nulle. Se invece x^0 è un punto regolare, ricordando la disuguaglianza di Cauchy-Schwarz, si ha che la direzione in cui la derivata direzionale (e quindi la velocità di crescita) è massima è $v = \frac{\nabla f(x^0)}{\|\nabla f(x^0)\|}$ mentre la direzione in cui la derivata direzionale è minima (e quindi la velocità di decrescita è massima) è $v = -\frac{\nabla f(x^0)}{\|\nabla f(x^0)\|}$. Pertanto il gradiente di f in x^0, qualora non nullo, è il vettore la cui direzione è quella per cui è massima la velocità di crescita di f e la cui norma coincide con il valore di tale velocità di crescita, dato che $\nabla f\left(x^0\right)\cdot\frac{\nabla f(x^0)}{\|\nabla f(x^0)\|} = \left\|\nabla f\left(x^0\right)\right\|$.

Osservazione 6.24. Nel caso $\mathtt{d} = 2$, il grafico di f è una superficie nello spazio cartesiano $O(x,y,z)$

$$\Sigma = \{(x,y,f(x,y)), (x,y)\in D\}$$

Se f è differenziabile in $P_0\left(x_0,y_0\right)\in\mathring{D}$, allora per ogni direzione v il grafico della restrizione di f alla retta passante per P_0 avente direzione v è una curva dotata di retta tangente in $Q_0\left(x_0,y_0,f\left(x_0,y_0\right)\right.$ avente pendenza $m = \frac{\partial f}{\partial v}\left(x_0,y_0\right)$; i vettori tangenti a tale curva sono tutti e soli i multipli di $t_v\left(v_1,v_2,\frac{\partial f}{\partial v}\left(x_0,y_0\right)\right)$. In particolare, le curve grafico delle restrizioni di f alle rette parallele agli assi x e y hanno per vettori tangenti i multipli di $t_x\left(1,0,\frac{\partial f}{\partial x}\left(x_0,y_0\right)\right)$ e $t_y\left(0,1,\frac{\partial f}{\partial y}\left(x_0,y_0\right)\right)$ rispettivamente, che sono linearmente indipendenti. Tenendo conto di (13), si ha che, per ogni direzione v, $t_v = v_1 t_x + v_2 t_y$, quindi i vettori tangenti a tutte le curve ottenute intersecando la superficie Σ, grafico di f, con tutti i piani ortogonali al piano $z=0$ passanti per Q_0 costituiscono un sottospazio di dimensione 2 che ha come base i vettori t_x e t_y. Questo garantisce che le rette tangenti a tutte queste curve costituiscono un piano passante per Q_0, detto piano tangente al grafico di f in Q_0, avente equazione

$$\pi := \left\{(x,y,z)\in\mathbb{R}^3 \;\middle|\; z = f\left(x_0,y_0\right) + \frac{\partial f}{\partial x}\left(x_0,y_0\right)\left(x-x_0\right) + \frac{\partial f}{\partial y}\left(x_0,y_0\right)\left(y-y_0\right)\right\}$$

I concetti di differenziabilità secondo Gateaux e Fréchet sono stati introdotti rispettivamente nelle Definizioni 6.3 e 6.16. I seguenti punti riassuntivi sono degni di nota:

1. La G-differenziabilità non è la corretta estensione multi-dimensionale del concetto di derivata, cf. esempio 6.11, tuttavia si può calcolare il G-differenziale di una funzione con sufficiente regolarità in maniera relativamente semplice, si veda l'Osservazione 6.7.

2. L'F-differenziabilità è si una giusta estensione multidimensionale del concetto di derivata, cf. Lemma 6.19, ed una funzione F-differenziabile è pure G-differenziabile, cf. Lemma 6.20, tuttavia non abbiamo un criterio semplice come quello proposto nell'Osservazione 6.7 per calcolare il differenziale di Fréchet di una funzione.

Domanda. Quando la differenziabilità di Gateaux implica la differenziabilità secondo Fréchet?

Se possiamo rispondere alla domanda qui sopra otteniamo dunque un criterio semplice di differenziabilità che ci permette di calcolare esplicitamente il differenziale di una funzione. Tale risultato prende il nome di *Teorema del differenziale totale* e garantisce che una funzione G-differenziabile con derivate direzionali continue è pure F-differenziabile.

Tale risultato verrà enunciato per funzioni tra spazi di Banach, ma dimostrato nel semplice caso $(X,Y) = \left(\mathbb{R}^{\mathrm{d}},\mathbb{R}\right)$, si veda il Teorema 6.27 sotto. La dimostrazione del caso generico si può trovare in [1] ma richiede l'utilizzo del Teorema di Hahn-Banach, [2].

Teorema 6.25 (del differenziale totale in spazi di Banach)**.** *Sia $\Omega \subset X$, e $f : \Omega \to Y$ una funzione G-differenziabile in Ω tale che la mappa*

$$\begin{array}{ccc} \Omega & \to & \mathcal{L}(X;Y) \\ x^0 & \mapsto & \mathrm{d}_{\mathrm{G}} f\left(x^0\right) \end{array},$$

è continua. Allora f è Fréchet differenziabile in x^0 e si ha che $\mathrm{d}_{\mathrm{G}} f\left(x^0;\bullet\right) = \mathrm{d}f\left(x^0\right)[\bullet]$.

Osservazione 6.26.

1. Notiamo che nell'enunciato del Teorema 6.25 si richiede, aggiuntivamente a ciò che accade nella Definizione 6.3, cf. Osservazione 6.18, che il differenziale di Gateaux di f in x^0 sia un operatore lineare limitato.

2. Notiamo che sotto le ipotesi del Teorema 6.25 possiamo utilizzare l'Osservazione 6.7 per calcolare esplicitamente il differenziale di Fréchet di una funzione. Pertanto nell'esempio 6.8 si è calcolato il differenziale di Fréchet delle funzioni considerate, notiamo infatti che gli operatori dedotti sono lineari in v ed h rispettivamente come sottolineato nel punto 1 qui sopra.

Teorema 6.27 (Teorema del differenziale totale in $\mathbb{R}^{\mathrm{d}}$)**.** *Sia $f : \Omega \subset \mathbb{R}^{\mathrm{d}} \to \mathbb{R}$ dotata di tutte le derivate parziali in un intorno di x^0, continue in x^0. Allora $f \in \mathcal{C}^1(\Omega;\mathbb{R})$ e $\mathrm{d}f\left(x^0\right) = \nabla f\left(x^0\right)$ (cf.* (6.10)*) per ogni $x^0 \in \Omega$.*

Dimostrazione. Per semplicità dimostriamo il teorema nel caso $\mathrm{d} = 2$, il caso generale segue in modo analogo. Sia $\delta > 0$ tale che f è dotata di tutte le derivate parziali in $B\left(x^0,\delta\right)$. Definiamo

$$\mathrm{d}_{\mathrm{G}} f\left(x^0\right) v = \frac{\partial f}{\partial v}\left(x^0\right) := \nabla f\left(x^0\right)\cdot v,$$

dove $\nabla f\left(x^0\right)$ è dato da (6.10). Siccome le derivate parziali sono continue in x^0, assumendo δ sufficientemente piccolo, si ha che, applicando (4.3)

$$\left|\nabla f\left(x^0\right)\cdot v\right| \leq \left\|\nabla f\left(x^0\right)\right\| \|v\| \leq \underbrace{2\sup_{i=1,2}\left|\frac{\partial f}{\partial x_i}\left(x^0\right)\right|}_{\leq C<\infty} \|v\|,$$

pertanto $\nabla f\left(x^0\right) \in \mathcal{L}\left(\mathbb{R}^2;\mathbb{R}\right)$. Per ogni $x \in B\left(x^0,\delta\right)$, si ha

$$f(x) - f\left(x^0\right) = f(x_1,x_2) - f\left(x_1^0,x_2\right) + f\left(x_1^0,x_2\right) - f\left(x_1^0,x_2^0\right). \tag{6.13}$$

Il teorema di Lagrange garantisce che esistono ξ_1 compreso tra x_1^0 e x_1 e ξ_2 compreso tra x_2^0 e x_2 (entrambi dipendenti da x ed x^0) tali che

$$\begin{aligned} f(x_1,x_2) - f\left(x_1^0,x_2\right) &= \frac{\partial f}{\partial x_1}(\xi_1,x_2)\left(x_1 - x_1^0\right), \\ f\left(x_1^0,x_2\right) - f\left(x_1^0,x_2^0\right) &= \frac{\partial f}{\partial x_2}\left(x_1^0,\xi_2\right)\left(x_2 - x_2^0\right). \end{aligned} \tag{6.14}$$

Le Equazioni (6.13) e (6.14) dimostrano dunque che

$$f(x) - f\left(x^0\right) = \frac{\partial f}{\partial x_1}(\xi_1,x_2)\left(x_1 - x_1^0\right) + \frac{\partial f}{\partial x_2}\left(x_1^0,\xi_2\right)\left(x_2 - x_2^0\right). \tag{6.15}$$

Applichamo dunque (6.15) ottenendo

$$\begin{aligned}R\left(x^0;x-x^0\right) &:= f(x)-f\left(x^0\right)-\nabla f\left(x^0\right)\cdot\left(x-x^0\right)\\ &= \left(\frac{\partial f}{\partial x_1}(\xi_1,x_2)-\frac{\partial f}{\partial x_1}\left(x^0\right)\right)\left(x_1-x_1^0\right)+\left(\frac{\partial f}{\partial x_2}\left(x_1^0,\xi_2\right)-\frac{\partial f}{\partial x_2}\left(x^0\right)\right)\left(x_2-x_2^0\right), \quad (6.16)\end{aligned}$$

da cui

$$\frac{\left|R\left(x^0;x-x^0\right)\right|}{\left\|x-x^0\right\|}\leq\left|\frac{\partial f}{\partial x_1}(\xi_1,x_2)-\frac{\partial f}{\partial x_1}\left(x^0\right)\right|+\left|\frac{\partial f}{\partial x_2}\left(x_1^0,\xi_2\right)-\frac{\partial f}{\partial x_2}\left(x^0\right)\right|.$$

Dalla continuità delle derivate parziali in x^0, si ha

$$\lim_{x\to x^0}\frac{R\left(x^0;x-x^0\right)}{\left\|x-x^0\right\|}=0.$$

Dalla formula (6.16) segue che $\mathrm{d}f\left(x^0\right)=\nabla f\left(x^0\right)$, la continuità della mappa $x^0\mapsto \mathrm{d}f\left(x^0\right)=\nabla f\left(x^0\right)$ segue dalla continuità delle derivate parziali. □

Osservazione 6.28. Il Teorema 6.27 gioca un ruolo fondamentale nel contesto semplificato dove $(X,Y)=\left(\mathbb{R}^{\mathrm{d}},\mathbb{R}\right)$. Questo teorema stabilisce una corrispondenza tra funzioni Fréchet differenziabili e funzioni che possiedono derivate parziali continue. Pertanto, nel prosieguo di questo testo, useremo la presenza di derivate parziali continue come criterio per definire la Fréchet differenziabilità di una funzione $f:\mathbb{R}^{\mathrm{d}}\to\mathbb{R}$.

Esercizio 6.29. Dimostrare che le derivate parziali di (6.6) non sono continue in zero.

Notazione 6.30. Dati due punti x_1,x_2 di un $\mathbb{K}=\mathbb{R},\mathbb{C}$ spazio vettoriale X denotiamo con

$$[x_1,x_2]:=\left\{x\in X \mid x=tx_1+(1-t)\,x_2,\ t\in[0,1]\right\}.$$

Teorema 6.31 (degli incrementi finiti)**.** *Sia Ω un aperto di $\mathbb{R}^{\mathrm{d}}$ e $f\in\mathcal{C}^1(\Omega;\mathbb{R})$. Siano $x,y\in\Omega$, $x\neq y$, tali che $[x,y]\subset\Omega$. Allora esiste $\xi\in[x,y]$ tale che*

$$f(y)-f(x)=\nabla f(\xi)\cdot(y-x), \qquad (6.17)$$

$$\left|f(y)-f(x)\right|\leq\sup_{\xi\in[x,y]}\left\|\nabla f(\xi)\right\|\left\|y-x\right\|. \qquad (6.18)$$

Dimostrazione. Consideriamo la direzione $v=\frac{y-x}{\|y-x\|}\in\mathbb{S}^{\mathrm{d}-1}$ e sia $g(t)=f(x+tv)$. Si ha che $g(0)=f(x)$, $g(\|y-x\|)=f(y)$. Inoltre g è derivabile in $[0,\|y-x\|]$ e, per ogni $t\in[0,\|y-x\|]$ si ha che

$$g'(t)=\lim_{h\to0}\frac{g(t+h)-g(t)}{h}=\lim_{h\to0}\frac{f(x+tv+hv)-f(x+tv)}{h}=\frac{\partial f}{\partial v}(x+tv).$$

Applicando il teorema di Lagrange a g, si ha che esiste $s\in(0,\|y-x\|)$ tale che

$$f(y)-f(x)=g'(s)\left\|y-x\right\|=\frac{\partial f}{\partial v}(\xi)\left\|y-x\right\|=\nabla f(\xi)\cdot(y-x),$$

dove $\xi:=x+sv$, dimostrando (6.17).

La disuguaglianza (6.18) segue immediatamente da (6.17). □

Le disuguaglianza (6.18) è valida pure nel caso di applicazioni tra spazi di Banach generici, la cui dimostrazione è omessa in quanto richiede l'utilizzo del Teorema di Hahn-Banach, [2].

Proposizione 6.32. *Sia $f:\Omega\subset X\to Y$ G-differenziabile in Ω* e tale che $\mathrm{d}_{\mathrm{G}}f\left(x^0\right)\in\mathcal{L}(X;Y)$ per ogni $x^0\in\Omega$. *Siano $u,v\in\Omega$ tali che $[u,v]\subset\Omega$, allora*

$$\left\|f(u)-f(v)\right\|_Y\leq\sup_{w\in[u,v]}\|\mathrm{d}_{\mathrm{G}}(w)\|_{\mathcal{L}(X;Y)}\ \|u-v\|_X$$

Osservazione 6.33. La Proposizione 6.32 viene generalmente applicata quando una funzione è F-differenziabile in in un aperto Ω.

Vale la seguente generalizzazione della regola della catena

Proposizione 6.34 (Differenziale della funzione composta)**.** *Siano X, Y e Z di Banach, Ω e Θ aperti rispettivamente di X e Y tali che $x^0 \in \Omega \subset X$, $f : \Omega \to Y$, $f(x^0) \in \Theta \subset Y$ e $g : \Theta \to Z$. Siano f differenziabile in x^0 e g in $f(x^0)$. Allora si ha che*

$$\mathrm{d}(g \circ f)(x^0) = \mathrm{d}g(f(x^0))\ \mathrm{d}f(x^0).$$

Dimostrazione. Supponiamo $\delta, \epsilon > 0$ sufficientemente piccoli in modo che $B_X(x^0;\delta) \subset \Omega$ e $f(B_X(x^0;\delta)) \subset B_Y(f(x^0), \epsilon) \subset \Theta$. Dobbiamo dimostrare che per ogni $h \in B_X(0;\delta)$ si ha che

$$R_{g\circ f}(x^0;h) := g \circ f(x^0+h) - g \circ f(x^0) - \mathrm{d}g(f(x^0))\ \mathrm{d}f(x^0)h,$$

è tale che

$$\left\| R_{g\circ f}(x^0;h) \right\|_Z = o(\|h\|_X). \tag{6.19}$$

Definiamo

$$y^0 := f(x^0), \qquad k := f(x^0+h) - f(x^0). \tag{6.20}$$

La parte destra di (6.19) diventa

$$g(y^0+k) - g(y^0) - \mathrm{d}g(f(x^0))\ \mathrm{d}f(x^0)h = g(y^0+k) - g(y^0) - \mathrm{d}g(y^0)k + \mathrm{d}g(y^0)[k - \mathrm{d}f(x^0)h]. \tag{6.21}$$

Per differenziabilità di g in y^0 ho che (cf. Osservazione 6.13)

$$\left\| g(y^0+k) - g(y^0) - \mathrm{d}g(y^0)k \right\|_Z = o(\|k\|_Y), \tag{6.22}$$

$$\left\| \mathrm{d}g(y^0)[k - \mathrm{d}f(x^0)h] \right\|_Z \le \left\| \mathrm{d}g(y^0) \right\|_{\mathcal{L}(Y;Z)} \left\| k - \mathrm{d}f(x^0)h \right\|_Y. \tag{6.23}$$

Notiamo da (6.20) e dalla differenziabilità di f in x^0 che

$$\|k\|_Y = \left\| f(x^0+h) - f(x^0) \right\|_Y = \left\| \mathrm{d}f(x^0)h + R_f(x^0;h) \right\|_Y \le \left\| \mathrm{d}f(x^0) \right\|_{\mathcal{L}(X;Y)} \|h\|_X + o(\|h\|_X).$$

Nuovamente, dalla differenziabilità di f in x^0 che $\left\| \mathrm{d}f(x^0) \right\|_{\mathcal{L}(X;Y)} < \infty$ e pertanto usando le proprietà della notazione o-piccolo deduco che

$$o(\|k\|_Y) = o(\|h\|_X), \tag{6.24}$$

quindi (6.24) in (6.22) da

$$\left\| g(y^0+k) - g(y^0) - \mathrm{d}g(y^0)k \right\|_Z = o(\|h\|_X). \tag{6.25}$$

Studiamo ora (6.23), per differenziabilità di g in y^0 ho che $\left\| \mathrm{d}g(y^0) \right\|_{\mathcal{L}(Y;Z)} < \infty$ e da (6.20) e dalla differenziabilità di f in x^0 ho che

$$\left\| k - \mathrm{d}f(x^0)h \right\|_Y = \left\| R_f(x^0;h) \right\|_Y = o(\|h\|_X). \tag{6.26}$$

Le Eq. (6.25) e (6.26) dimostrano che la parte sinistra di Eq. (6.21) è un $o(\|h\|_X)$, dimostrando (6.19), il che conclude la dimostrazione. □

Esempio 6.35 (Esempio di applicazione della regola della catena)**.** Sia $g : \mathbb{R}^{\mathtt{M}} \to \mathbb{R}$ e $f : \mathbb{R}^{\mathtt{N}} \to \mathbb{R}^{\mathtt{M}}$ (ci si può restringere ad aperti con domini e codomini inclusi) differenziabili rispettivamente in $y^0 := f(x^0)$ ed x^0. Consideriamo l'applicazione

$$g \circ f(x) = g\left(f_1(x_1, \ldots, x_{\mathtt{N}}), \ldots, f_{\mathtt{M}}(x_1, \ldots, x_{\mathtt{N}})\right).$$

Applicando Proposizione 6.34 ottengo che per ogni $h \in \mathbb{R}^{\mathtt{N}}$

$$\mathrm{d}(g \circ f)(x^0)h = \mathrm{d}g(f(x^0))\mathrm{d}f(x^0)h.$$

Applichiamo i risultati provati nell'Osservazione 6.21 ottenendo che

$$\mathrm{d}g(y^0)k = \nabla g(y^0) \cdot k, \quad k \in \mathbb{R}^{\mathtt{M}} \qquad \mathrm{d}f(x^0)h = Jf(x^0)h, \quad h \in \mathbb{R}^{\mathtt{N}},$$

pertanto

$$\mathrm{d}(g \circ f)(x^0)h = \nabla g(f(x^0))^{\mathsf{T}} \cdot Jf(x^0)h = \sum_{j=1}^{\mathtt{M}} \sum_{i=1}^{\mathtt{N}} \partial_j g(f(x^0)) \partial_i f_j(x^0) h_i.$$

6.2 Derivate di ordine superiore, teorema di Schwartz e formula di Taylor

6.2.1 Differenziale di ordine superiore e Teorema di Schwartz

Definizione 6.36 (Differenziale secondo)**.** Siano X, Y di Banach, $\Omega \subset X$ e $f \in \mathcal{C}^1(\Omega; Y)$. Diciamo che f *è doppiamente differenziabile in* Ω se la mappa

$$\begin{array}{ccc} \Omega & \rightarrow & \mathcal{L}(X;Y) \\ x^0 & \mapsto & \mathrm{d}f\left(x^0\right) \end{array},$$

è differenziabile secondo la Definizione 6.16. La risultante mappa

$$\begin{array}{ccc} \Omega & \rightarrow & \mathcal{L}(X;\mathcal{L}(X;Y)) \\ x^0 & \mapsto & \mathrm{d}^2 f\left(x^0\right) \end{array}, \tag{6.27}$$

è detto *differenziale secondo di f in* x^0. Se la mappa in (6.27) è continua diremo che $f \in \mathcal{C}^2(\Omega; Y)$.

La Definizioni 6.36 dà un'estensione logica del concetto di differenziale di Fréchet di ordine superiore, tuttavia lo spazio $\mathcal{L}(X;\mathcal{L}(X;Y))$ nel quale è definito il differenziale secondo non è molto intuitivo, vediamo ora di caratterizzare $\mathcal{L}(X;\mathcal{L}(X;Y))$ in termini più immediati.

Definizione 6.37 (Topologia prodotto)**.** Sia $p \in \mathbb{N} \setminus 0$ e X_i, $i = 1, \ldots, p$ spazi di Banach. Consideriamo lo spazio prodotto $\prod_{i=1}^{p} X_i$, e la funzione

$$\left\| \vec{x}_p \right\|_{\prod_{i=1}^{p} X_i} := \sqrt{\sum_{i=1}^{p} \|x_i\|_{X_i}^2}, \qquad \forall\, \vec{x}_p := \left(x_1, \ldots, x_p\right) \in \prod_{i=1}^{p} X_i,$$

la coppia $\left(\prod_{i=1}^{p} X_i, \|\cdot\|_{\prod_{i=1}^{p} X_i}\right)$ è di Banach.

Definizione 6.38 (Operatori multilineari limitati)**.** Siano X, Y di Banach e $p \in \mathbb{N} \setminus 0$. Lo spazio degli *operatori p-lineari* $\mathcal{L}_p(X;Y)$ è composto dalle applicazioni

$$M_p : X^p \rightarrow Y,$$

che sono

1. *lineari* in ognuna dei p argomenti;
2. *limitate*, ossia esiste un $C > 0$ tale che per ogni $\left(x_1, \ldots, x_p\right) \in X^p$ si ha che

$$\left\| M_p\left(x_1, \ldots, x_p\right) \right\|_Y \leq C \prod_{i=1}^{p} \|x_i\|_X .$$

Lo spazio $\mathcal{L}_p(X;Y)$ con la norma

$$\left\| M_p \right\|_{\mathcal{L}_p(X;Y)} := \sup_{\left\|\left(x_1,\ldots,x_p\right)\right\|_{X^p} \leq 1} \left\| M_p\left(x_1, \ldots, x_p\right) \right\|_Y < \infty,$$

è di Banach. Se $X = Y$ denotiamo $\mathcal{L}_p(X;X) =: \mathcal{L}_p(X)$.

Esempio 6.39.

1. Siano $\mathtt{d} \in \mathbb{N} \setminus 0$, $A \in \mathbb{R}^{\mathtt{d} \times \mathtt{d}}$, l'applicazione

$$B_A\left(x, y\right) := (Ax) \cdot y = y^{\mathsf{T}} \cdot (Ax) = \sum_{i=1}^{\mathtt{d}} \sum_{j=1}^{\mathtt{d}} A_{ij} x_j y_i, \qquad \forall x, y \in \mathbb{R}^{\mathtt{d}},$$

   è bilineare ed appartiene allo spazio $\mathcal{L}_2\left(\mathbb{R}^{\mathtt{d}};\mathbb{R}\right)$.

2. Siano $u, v \in \mathcal{C}^1(\mathbb{R})$, l'applicazione

$$(u, v) \mapsto u'v' \in \mathcal{L}_2\left(\mathcal{C}^1(\mathbb{R});\mathcal{C}^0(\mathbb{R})\right).$$

Proposizione 6.40. *Siano X, Y di Banach, si ha la seguente isometria*

$$\mathcal{L}(X;\mathcal{L}(X;Y)) \simeq \mathcal{L}_2(X;Y).$$

Dimostrazione. L'applicazione

$$\begin{array}{rcccl} \tilde{\Phi} & : & \mathcal{L}(X;\mathcal{L}(X;Y)) & \to & \mathcal{L}_2(X;Y) \\ & & A & \mapsto & \{(u,v)\mapsto A(u)[v]\} \end{array},$$

definisce l'isometria cercata, con inversa

$$\begin{array}{rcccl} \tilde{A} & : & \mathcal{L}_2(X;Y) & \to & \mathcal{L}(X;\mathcal{L}(X;Y)) \\ & & \Phi & \mapsto & \{u\mapsto\{v\mapsto\Phi(u,v)\}\} \end{array}.$$

Verificare i dettagli come esercizio. □

una conseguenza immediata della Proposizione 6.40 è il seguente corollario:

Corollario 6.41. *Si ha che* (6.27) *è equivalente a*

$$\begin{array}{ccc} \Omega & \to & \mathcal{L}_2(X;Y) \\ x^0 & \mapsto & \mathrm{d}^2 f\left(x^0\right) \end{array}, \tag{6.28}$$

Il Corollario 6.41 ci permette dunque di identificare il differenziale secondo con un'applicazione bilineare.

Esempio 6.42 (Calcolo del differenziale secondo in alcuni caso specifici)**.**

1. Sia $f\in\mathcal{C}^2(\Omega;\mathbb{R})$. Da (6.9) sappiamo che $\mathrm{d}f\left(x^0\right)h=\nabla f\left(x^0\right)\cdot h$, applichiamo Osservazione 6.26, 2 ed otteniamo

$$\mathrm{d}^2 f\left(x^0\right)[k,h]=\lim_{\epsilon\to 0}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\nabla f\left(x^0+\epsilon k\right)\cdot h=\lim_{\epsilon\to 0}\frac{\mathrm{d}}{\mathrm{d}\epsilon}\sum_{i=1}^{\mathrm{d}}\partial_i f\left(x^0+\epsilon k\right)h_i$$

Definiamo $\phi:[0,\epsilon_0]\to\mathbb{R}^{\mathrm{d}}$ come $\phi(\epsilon):=x^0+\epsilon k$, applichiamo l'esempio 6.35 ed otteniamo

$$\frac{\mathrm{d}}{\mathrm{d}\epsilon}\partial_i f\left(x^0+\epsilon k\right)=\nabla\partial_i f\left(\phi(\epsilon)\right)^{\intercal}\cdot J\phi(\epsilon),$$

e calcolando esplicitamente il Jacobiano di ϕ otteniamo

$$J\phi(\epsilon)=\begin{pmatrix}\phi_1'(\epsilon)\\ \vdots\\ \phi_{\mathrm{d}}'(\epsilon)\end{pmatrix}=k,$$

pertanto

$$\mathrm{d}^2 f\left(x^0\right)[k,h]=\lim_{\epsilon\to 0}\sum_{i,j=1}^{\mathrm{d}}\partial_j\partial_i f\left(x^0+\epsilon k\right)k_j h_i.$$

Per ipotesi $f\in\mathcal{C}^2$ e pertanto le sue derivate seconde sono continue quindi otteniamo che

$$\mathrm{d}^2 f\left(x^0\right)[k,h]=\sum_{i,j=1}^{\mathrm{d}}\partial_j\partial_i f\left(x^0\right)k_j h_i.$$

2. Sia $X:=\mathcal{C}([0,1];\mathbb{R})$ e $f(u):=\frac{1}{1+u}$ come in esempio 6.6. abbiamo visto che $\mathrm{d}f(u)$ è l'operatore moltiplicativo (cf. esempio 6.14, punto 2)

$$h\mapsto\frac{-h}{(1+u)^2}.$$

Calcoliamo ora le derivata direzionale (cf. Definizione 6.3)

$$\partial_k \mathrm{d}f(u)[h] = \lim_{\epsilon\to 0} \frac{\mathrm{d}f(u+\epsilon k)\,h - \mathrm{d}f(u)\,h}{\epsilon}.$$

Se il risultato che troviamo è una mappa continua da un inotrno di $u \in X$ a $\mathcal{L}_2(X)$ allora applicando Teorema 6.25 otteniamo il differenziale secondo. Calcoliamo

$$\big(\mathrm{d}f(u+\epsilon k) - \mathrm{d}f(u)\big)\,h = \left[-\frac{1}{(1+u+\epsilon k)^2} + \frac{1}{(1+u)^2}\right] h = \frac{2\epsilon u k h + O\big(\epsilon^2\big)}{(1+u+\epsilon k)^2\,(1+u)^2},$$

quindi

$$\left(\frac{\mathrm{d}f(u+\epsilon k) - \mathrm{d}f(u)}{\epsilon}\right) h \xrightarrow{\epsilon\to 0} \underbrace{\frac{2u}{(1+u)^4}\,k}_{=\partial_k \mathrm{d}f(u)}\,h. \tag{6.29}$$

Certamente la mappa $(k,h) \mapsto \partial_k \mathrm{d}f(u)\,h$ definita in (6.29) è bilineare, vediamo ora che è continua in un intorno di u:

$$\left\|\partial_k \mathrm{d}f(u+\delta u)\,h - \partial_k \mathrm{d}f(u)\,h\right\|_X = \left\|\left(\frac{2(u+\delta u)}{(1+u+\delta u)^4} - \frac{2u}{(1+u)^4}\right) hk\right\|_X.$$

Notiamo che

$$\left\|\left(\frac{2(u+\delta u)}{(1+u+\delta u)^4} - \frac{2u}{(1+u)^4}\right) hk\right\|_X \le \left\|\frac{2(u+\delta u)}{(1+u+\delta u)^4} - \frac{2u}{(1+u)^4}\right\|_X \|h\|_X \|k\|_X,$$

ed inoltre

$$\frac{2(u+\delta u)}{(1+u+\delta u)^4} - \frac{2u}{(1+u)^4} = O(\delta u),$$

otteniamo quindi che

$$\left\|\partial_k \mathrm{d}f(u+\delta u)\,h - \partial_k \mathrm{d}f(u)\,h\right\|_X = O(\|\delta u\|_X)\,\|k\|_X\,\|h\|_X \xrightarrow{\|\delta u\|_X\to 0} 0,$$

pertanto applicando il Teorema 6.25 ottengo che

$$\mathrm{d}^2 f(u)[k,h] = \frac{2u}{(1+u)^4}\,kh.$$

Possiamo quindi generalizzare il concetto di differenziale ad un ordine arbitrario

Definizione 6.43 (Differenziale p-esimo)**.** Siano X, Y di Banach, $p \in \mathbb{N}\setminus 0$, $\Omega \subset X$ e $f \in \mathcal{C}^{p-1}(\Omega; Y)$. Diciamo che *$f$ è p volte differenziabile in* Ω se la mappa

$$\begin{array}{ccc} \Omega & \to & \mathcal{L}_{p-1}(X;Y) \\ x^0 & \mapsto & \mathrm{d}^{p-1} f\big(x^0\big) \end{array},$$

è differenziabile secondo la Definizione 6.16. La risultante mappa

$$\begin{array}{ccc} \Omega & \to & \mathcal{L}_{p}(X;Y) \\ x^0 & \mapsto & \mathrm{d}^{p} f\big(x^0\big) \end{array}, \tag{6.30}$$

è detto *differenziale p-esimo di f in x^0*. Se la mappa in (6.30) è continua diremo che $f \in \mathcal{C}^p(\Omega; Y)$. Diremo che $f \in \mathcal{C}^\infty(\Omega; Y)$ se $f \in \bigcap_{p\in\mathbb{N}} \mathcal{C}^p(\Omega; Y)$.

Definizione 6.44 (Operatore multilineare simmetrico)**.** Sia $p \in \mathbb{N}\setminus 0$, X, Y di Banach, $M_p \in \mathcal{L}_p(X;Y)$, si dice *simmetrico* se per ogni σ permutazione dell'insieme $\{1,\dots,p\}$ e $\big(x_1,\dots,x_p\big) \in X^p$ si ha che

$$M_p\big(x_1,\dots,x_p\big) = M_p\Big(x_{\sigma(1)},\dots,x_{\sigma(p)}\Big).$$

Osservazione 6.45. Nel caso si consideri $p = 2$ un operatore bilineare è simmetrico se $B\big(x,y\big) = B\big(y,x\big)$.

Teorema 6.46 (di Schwartz)**.** *Siano X, Y di Banach, $x^0 \in \Omega \subset X$ e $f : \Omega \to Y$ tale che f è doppiamente differenziabile in Ω, allora $\mathrm{d}^2 f\left(x^0\right)$ è un operatore limitato bilineare simmetrico da X^2 a Y.*

Dimostrazione. Sia $\delta > 0$ tale che $B_X\left(x^0, \delta\right) \subset \Omega$ e $h, k \in B_X(0, \delta)$.

Step 1 (Approssimate del differenziale secondo)**.** Definiamo l'approssimante discreto del differenziale secondo

$$\triangle^2 f\left(x^0; k, h\right) := f\left(x^0 + k + h\right) - f\left(x^0 + k\right) - f\left(x^0 + h\right) + f\left(x^0\right), \tag{6.31}$$

detta anche *differenza finita.*

Step 2 (Stima dell'errore)**.** Definiamo l'errore

$$\mathcal{E} f\left(x^0; h, k\right) := \triangle^2 f\left(x^0; k, h\right) - \mathrm{d}^2 f\left(x^0\right)[k, h], \tag{6.32}$$

ovviamente si ha che $\mathcal{E} f : \Omega \times B_X(0, \delta)^2 \to Y$, e vogliamo calcolare

$$\left\| \mathcal{E} f\left(x^0; h, k\right) \right\|_Y.$$

Definiamo la seguente curva a valori in Y

$$\begin{array}{rcccl} \gamma^1_{k,h} & : & [0,1] & \to & Y \\ & & \tau & \mapsto & f\left(x^0 + k + \tau h\right) - f\left(x^0 + \tau h\right) - \mathrm{d}^2 f\left(x^0\right)[k, \tau h], \end{array}$$

notiamo che $\mathcal{E} f\left(x^0; h, k\right) = \gamma^1_{k,h}(1) - \gamma^1_{k,h}(0)$. Utilizzando Proposizione 6.34 ed il fatto che $f \in \mathcal{C}^1(\Omega; Y)$ possiamo calcolare il differenziale di $\gamma^1_{k,h}$ in punto arbitrario di $[0,1]$:

$$\mathrm{d}\gamma^1_{k,h}(\tau)\,\tau' = \mathrm{d}f\left(x^0 + k + \tau h\right)\tau' h - \mathrm{d}f\left(x^0 + \tau h\right)\tau' h - \left(\mathrm{d}^2 f\left(x^0\right) k\right)\tau' h, \qquad \forall \left(\tau, \tau'\right) \in [0,1]^2, \tag{6.33}$$

pertanto $\gamma^1_{k,h} \in \mathcal{C}^1([0,1]; Y)$ e quindi, per differenziabilità, si ha che

$$\mathcal{E} f\left(x^0; h, k\right) = \mathrm{d}\gamma^1_{k,h}(0)\,1 + R_1(0;1), \qquad \left\| R_1(0;\eta) \right\|_Y = o(|\eta|). \tag{6.34}$$

Un calcolo esplicito ci mostra che

$$\mathrm{d}\gamma^1_{k,h}(0)\,1 = \mathrm{d}f\left(x^0 + k\right) h - \mathrm{d}f\left(x^0\right) h - \left(\mathrm{d}^2 f\left(x^0\right) k\right) h = \gamma^2_{k,h}(1) - \gamma^2_{k,h}(0), \tag{6.35}$$

dove

$$\begin{array}{rcccl} \gamma^2_{k,h} & : & [0,1] & \to & Y \\ & & \vartheta & \mapsto & \mathrm{d}f\left(x^0 + \vartheta k\right) h - \left(\mathrm{d}^2 f\left(x^0\right) \vartheta k\right) h \end{array}, \tag{6.36}$$

quindi, visto che f è doppiamente differenziabile in x^0, ottengo che $\gamma^2_{k,h}$ è differenziabile in 0 ed applicando la Proposizione 6.34 ottengo

$$\begin{array}{c} \gamma^2_{k,h}(1) - \gamma^2_{k,h}(0) = \mathrm{d}\gamma^2_{k,h}(0)\,1 + R_2(0;1), \quad \left\| R_2(0,\eta) \right\|_Y = o(|\eta|), \\ \mathrm{d}\gamma^2_{k,h}(0)\,\vartheta' = 0\,. \end{array} \tag{6.37}$$

Inserendo le Equazioni (6.35) e (6.37) in (6.34) ottengo

$$\mathcal{E} f\left(x^0; h, k\right) = R_1(0;1) + R_2(0;1). \tag{6.38}$$

Calcoliamo ora esplicitamente i termini $R_j(0;1)$, $j = 1, 2$, ottenendo dalle Equazioni (6.36) e (6.37)

$$R_2(0;1) = \gamma^2_{k,h}(1) - \gamma^2_{k,h}(0) - \mathrm{d}\gamma^2_{k,h}(0)\,1 = \left(\mathrm{d}f\left(x^0 + k\right) - \mathrm{d}f\left(x^0\right) - \mathrm{d}^2 f\left(x^0\right) k\right) h. \tag{6.39}$$

Ma, f è differenziabile due volte in x^0, quindi

$$\mathrm{d}f\left(x^0 + k\right) - \mathrm{d}f\left(x^0\right) - \mathrm{d}^2 f\left(x^0\right) k = R_{\mathrm{d}f}\left(x^0; k\right), \qquad \left\| R_{\mathrm{d}f}\left(x^0; k\right) \right\|_{\mathcal{L}(X;Y)} = o(\|k\|_X), \tag{6.40}$$

pertanto le Equazioni (6.39) e (6.40) danno

$$\|R_2(0,1)\|_Y = \left\|R_{\mathrm{d}f}\left(x^0;k\right)h\right\|_Y = o\left(\|k\|_X\|h\|_X\right). \tag{6.41}$$

Procediamo in maniera analoga con il termine $R_1(0;1)$ definito in (6.34). Usando le Equazioni (6.32) e (6.33) ottengo

$$R_1(0;1) = \left[f\left(x^0+k+h\right) - f\left(x^0+k\right) - \mathrm{d}f\left(x^0+k\right)h\right] - \left[f\left(x^0+h\right) - f\left(x^0\right) - \mathrm{d}f\left(x^0\right)h\right] \quad ,$$

quindi siccome f è doppiamente differenziabile in Ω[1] applicando la Proposizione 6.32 ho che

$$\|R_1(0;1)\|_Y \le \sup_{\vartheta\in[0,1]} \left\|\mathrm{d}f\left(x^0+\vartheta k+h\right) - \mathrm{d}f\left(x^0+\vartheta k\right) - \mathrm{d}^2 f\left(x^0+\vartheta k\right)h\right\|_{\mathcal{L}(X;Y)}\|k\|_X,$$

e quindi dalla definizione di differenziabilità ho che per ogni $\vartheta\in[0,1]$

$$R_{\mathrm{d}f}\left(x^0+\vartheta k;h\right) := \mathrm{d}f\left(x^0+\vartheta k+h\right) - \mathrm{d}f\left(x^0+\vartheta k\right) - \mathrm{d}^2 f\left(x^0+\vartheta k\right)h, \quad \left\|R_{\mathrm{d}f}\left(x^0+\vartheta k;h\right)\right\|_{\mathcal{L}(X;Y)} = o\left(\|h\|_X\right),$$

ottenendo

$$\|R_1(0;1)\|_Y = o\left(\|k\|_X\|h\|_X\right). \tag{6.42}$$

Inserendo Eq. (6.41) e (6.42) in (6.38) ottengo

$$\left\|\mathcal{E}f\left(x^0;h,k\right)\right\|_Y = o\left(\|k\|_X\|h\|_X\right),$$

ossia esiste un $\varepsilon = \varepsilon(\delta) \xrightarrow{\delta\to 0} 0$ tale che

$$\left\|\mathcal{E}f\left(x^0;h,k\right)\right\|_Y \le \varepsilon\|k\|_X\|h\|_X. \tag{6.43}$$

Step 3 (Simmetrizzazione)**.** Usiamo ora il fatto che l'applicazione (cf. (6.31)) $(k,h)\mapsto \triangle^2 f\left(x^0;k,h\right)$ è simmetrica per calcolare

$$\left\|\mathrm{d}^2 f\left(x^0\right)[k,h] - \mathrm{d}^2 f\left(x^0\right)[h,k]\right\|_Y \le \left\|\mathcal{E}f\left(x^0;h,k\right)\right\|_Y + \left\|\mathcal{E}f\left(x^0;k,h\right)\right\|_Y,$$

quindi applicando (6.43) ottengo

$$\left\|\mathrm{d}^2 f\left(x^0\right)[k,h] - \mathrm{d}^2 f\left(x^0\right)[h,k]\right\|_Y \le \varepsilon\|k\|_X\|h\|_X, \qquad \forall\, k,h\in B_X(0,\delta). \tag{6.44}$$

Step 4 (Riscalamento)**.** La stima in Eq. (6.44) vale solo per h,k piccoli. Consideriamo ora $k,h\in X$ arbitrari, e scriviamo $(k,h) = \left(\|k\|_X\hat{k}, \|h\|_X\hat{h}\right)$ dove $\hat{k},\hat{h}\in\mathbb{S}_X$ e $\mathcal{N}(k,h) := \mathrm{d}^2 f\left(x^0\right)[k,h] - \mathrm{d}^2 f\left(x^0\right)[h,k]$. Per bilinearità di $\mathcal{N}$ ho che

$$\mathcal{N}(k,h) = \frac{\|k\|_X\|h\|_X}{\delta^2}\mathcal{N}\left(\delta\hat{k},\delta\hat{h}\right),$$

applicando (6.44) otteniamo

$$\|\mathcal{N}(k,h)\|_Y \le \varepsilon(\delta)\|k\|_X\|h\|_X. \tag{6.45}$$

Notiamo che la parte sinistra di (6.45) non dipende da $\delta>0$, pertanto siccome $\varepsilon \xrightarrow{\delta\to 0} 0$ ottengo

$$\|\mathcal{N}(k,h)\|_Y \le \inf_{\delta>0}\varepsilon(\delta)\|k\|_X\|h\|_X = 0,$$

il che conclude la dimostrazione.

□

[1] Notare che questo è l'unico punto in cui si richiede doppia differenziabilità in un punto diverso da x^0

6.2.2 Differenziale secondo nel caso di dimensione finita

Notazione 6.47. Data $f:\mathbb{R}^{\mathtt{d}}\to\mathbb{R}$ denotiamo con

$$\partial_{ij}f := f_{x_i x_j} := \frac{\partial^2 f}{\partial x_i \partial x_j} := \frac{\partial}{\partial x_i}\frac{\partial f}{\partial x_j}, \qquad i,j=1,\dots,\mathtt{d}. \tag{6.46}$$

Definizione 6.48 (Matrice Hessiana)**.** Data $f:\Omega\subset\mathbb{R}^{\mathtt{d}}$ per la quale gli elementi in (6.46) esistono per $x^0\in\Omega$ definiamo la *matrice Hessiana*

$$Hf\left(x^0\right) := \left(\partial_{ij}f\left(x^0\right)\right)_{i,j=1,\dots,\mathtt{d}} = \begin{pmatrix} \partial_{11}f\left(x^0\right) & \cdots & \partial_{1\mathtt{d}}f\left(x^0\right) \\ \vdots & & \vdots \\ \partial_{\mathtt{d}1}f\left(x^0\right) & \cdots & \partial_{\mathtt{dd}}f\left(x^0\right) \end{pmatrix} \in \mathbb{R}^{\mathtt{d}\times\mathtt{d}}.$$

Lemma 6.49. *Sia $\Omega\subset\mathbb{R}^{\mathtt{d}}$ e $f\in\mathcal{C}^1(\Omega;\mathbb{R})$ tale che $Hf\in\mathcal{C}\left(\Omega;\mathbb{R}^{\mathtt{d}\times\mathtt{d}}\right)$, allora $f\in\mathcal{C}^2(\Omega;\mathbb{R})$ e*

$$\mathrm{d}^2 f\left(x^0\right)[x,y] = y\cdot\left(Hf\left(x^0\right)x\right), \qquad \forall x,y\in\mathbb{R}^{\mathtt{d}},\ x^0\in\Omega.$$

Dimostrazione. Siccome f è differenziabile applicando il Teorema 6.27 otteniamo che la condizione di differenziabilità è equivalente a

$$\nabla f\in\mathcal{C}\left(\Omega;\mathbb{R}^{\mathtt{d}}\right),$$

dove il gradiente viene considerato nella sua formulazione finito-dimensionale data in (6.10). Vogliamo dunque dimostrare che (la norma senza pedice indica la norma Euclidea in $\mathbb{R}^{\mathtt{d}}$)

$$R_{\mathrm{d}f}\left(x^0;x-x^0\right) := \nabla f(x)-\nabla f\left(x^0\right)-Hf\left(x^0\right)\cdot\left(x-x^0\right) = o\left(\left\|x-x^0\right\|\right). \tag{6.47}$$

Notiamo che

$$\left\|R_{\mathrm{d}f}\left(x^0;x-x^0\right)\right\| \le \mathtt{d}\max_{i=1,\dots,\mathtt{d}}\left|R_{f_i}\left(x-x^0\right)\right| = \mathtt{d}\max_{i=1,\dots,\mathtt{d}}\left|\partial_i f(x)-\partial_i f\left(x^0\right)-\nabla\partial_i f\left(x^0\right)\cdot\left(x-x^0\right)\right|. \tag{6.48}$$

Possiamo dunque ridurci a considerare le singole componenti del gradiente $\partial_i f:\Omega\to\mathbb{R}$ dove $i=1,\dots,\mathtt{d}$, applicando il Teorema 6.27 ottengo che $\partial_i f\in\mathcal{C}^1(\Omega;\mathbb{R})$ se e solamente se esiste

$$\nabla\partial_i f := \begin{pmatrix}\partial_{1i}f \\ \vdots \\ \partial_{\mathtt{d}i}f\end{pmatrix} \in \mathcal{C}\left(\Omega;\mathbb{R}^{\mathtt{d}}\right). \tag{6.49}$$

La condizione in (6.49) è automaticamente verificata grazie all'ipotesi di continuità dell'Hessiana, pertanto $\partial_i f\in\mathcal{C}^1(\Omega;\mathbb{R})$, con $\mathrm{d}\partial_i f\left(x^0\right)=\nabla\partial_i f\left(x^0\right)$ e dalla definizione di differenziabilità fer la funzione $\partial_i f$ in x^0 ho che

$$\left|\partial_i f(x)-\partial_i f\left(x^0\right)-\nabla\partial_i f\left(x^0\right)\cdot\left(x-x^0\right)\right| = o\left(\left\|x-x^0\right\|\right). \tag{6.50}$$

Inserendo (6.50) in (6.48) ottengo (6.47), la continuità dell'Hessiana implica che il differenziale secondo è continuo, questo conclude la dimostrazione. □

Corollario 6.50 (Teorema di Schwartz nel caso di dimensione finita)**.** *Sia f come nel Lemma 6.49, allora la matrice $Hf\left(x^0\right)$ è simmetrica.*

Dimostrazione. Sotto le ipotesi del Lemma 6.49 ho che

$$\mathrm{d}^2 f\left(x^0\right)[x,y] = \sum_{i,j=1}^{\mathtt{d}}\partial_{ij}f\left(x^0\right)x_j y_i,$$

inoltre il Lemma 6.49 ci garantisce che $f\in\mathcal{C}^2$, pertanto posso applicare il Teorema 6.46 ottenendo che

$$\sum_{i,j=1}^{\mathtt{d}}\partial_{ij}f\left(x^0\right)x_j y_i = \sum_{i,j=1}^{\mathtt{d}}\partial_{ji}f\left(x^0\right)x_i y_j, \qquad \forall x,y\in\mathbb{R}^{\mathtt{d}}.$$

Seleziono pertanto $x=\mathrm{e}_k$ e $y=\mathrm{e}_l$ dove e_ℓ, $\ell=1,\dots,\mathtt{d}$ sono i vettori della base canonica di $\mathbb{R}^{\mathtt{d}}$ ed ottengo che $\partial_{lk}f\left(x^0\right)=\partial_{kl}f\left(x^0\right)$. □

Osservazione 6.51. Il Corollario 6.50 ci garantisce che , sotto l'ipotesi di continuità delle derivate seconde, *l'operazione di derivazione è commutativa.*

6.2.3 Formula di Taylor

Per facilitare la presentazione del problema in questa sezione ci riduciamo al caso $(X,Y) = (\mathbb{R}^{\mathrm{d}}, \mathbb{R})$.

Definizione 6.52 (Multi-indice)**.** $p = (p_1, \ldots, p_{\mathrm{d}}) \in \mathbb{N}^{\mathrm{d}}$ è detto *multi-indice*. Denotiamo con $|p| := p_1 + \ldots + p_{\mathrm{d}}$ la *lunghezza* del multi-indice e con $p! := p_1! \cdots p_{\mathrm{d}}!$.

Definizione 6.53. Sia p un multi-indice, denotiamo con

$$\partial^p f := \frac{\partial^{|p|} f}{\partial^{p_1} x_1 \cdots \partial^{p_{\mathrm{d}}} x_{\mathrm{d}}},$$

e dato un vettore $w \in \mathbb{R}^{\mathrm{d}}$, si indica con

$$w^p = w_1^{p_1} \cdots w_{\mathrm{d}}^{p_{\mathrm{d}}}.$$

Teorema 6.54 (Formula di Taylor)**.** *Siano Ω un aperto di $\mathbb{R}^{\mathrm{d}}$, $f \in \mathcal{C}^n(\Omega;\mathbb{R})$, $x^0, x \in \Omega$, $x \neq x^0$, tali che $[x_0, x] \subset \Omega$. Allora vale la seguente formula*

$$f(x) = \sum_{|p| \le n} \frac{\partial^p f(x^0)}{p!} \left(x - x^0\right)^p + R_f^{[n]}\left(x^0; x - x^0\right) \tag{6.51}$$

con

$$\left| R_f^{[n]}\left(x^0; x - x^0\right) \right| = o\left(\left\| x - x^0 \right\|^n \right). \tag{6.52}$$

Inoltre, se $f \in \mathcal{C}^{n+1}(\Omega;\mathbb{R})$,

$$R_f^{[n]}\left(x^0; x - x^0\right) = \sum_{|p| = n+1} \frac{\partial^p f(\xi)}{p!} \left(x - x^0\right)^p, \tag{6.53}$$

con $\xi \in \left[x^0, x\right]$.

Dimostrazione. Posto $w = \frac{x - x^0}{\|x - x^0\|}$, definiamo la funzione

$$g(t) = f\left(x^0 + tw\right), \qquad t \in \left[0, \left\|x - x^0\right\|\right],$$

Si ha, applicando la Proposizione 6.34

$$\begin{aligned} g'(t) &= \frac{\partial f}{\partial w}\left(x^0 + tw\right) = \sum_{i=1}^{\mathrm{d}} \partial_i f\left(x^0 + tw\right) w_i \\ g''(t) &= \sum_{i,j=1}^{\mathrm{d}} \partial_j \partial_i f\left(x^0 + tw\right) w_i w_j \end{aligned} \tag{6.54}$$

e, in generale,

$$g^{(h)}(t) = \sum_{i_1, i_2, \ldots, i_h = 1}^{\mathrm{d}} \partial_{i_1} \partial_{i_2} \ldots \partial_{i_h} f\left(x^0 + tw\right) w_{i_1} w_{i_2} \ldots w_{i_h}$$

Se $p = (p_1, \ldots, p_{\mathrm{d}})$ è un multiindice di lunghezza h, il termine $\partial^p f\left(x^0 + tw\right) w^p$ compare nella somma precedente quando, degli indici $i_1, i_2, \ldots, i_h$, ce ne sono p_1 uguali a 1, p_2 uguali a $2, \ldots, p_{\mathrm{d}}$ uguali a d, ovvero $\frac{h!}{p!}$ volte. Si ha allora

$$g^{(h)}(t) = h! \sum_{|p| = h} \frac{\partial^p f\left(x^0 + tw\right) w^p}{p!} \tag{6.55}$$

Ricordiamo la formula di Taylor di ordine n nel punto $t = 0$ per $g(t)$:

$$g(t) = \sum_{h=0}^{n} \frac{g^{(h)}(0)}{h!} t^h + R_n(t, 0)$$

con

$$\lim_{t \to 0} \frac{R_n(t,0)}{t^n} = 0 \tag{6.56}$$

Sostituendo (6.55) in (6.56) e scegliendo $t := \left\|x - x^0\right\|$, si ottengono le Eq. (6.51) e (6.52). Se, inoltre $f \in \mathcal{C}^{n+1}(\Omega;\mathbb{R})$, allora $g \in \mathcal{C}^{n+1}\left(\left[0, \left\|x - x^0\right\|\right]; \mathbb{R}\right)$ e (6.53) segue dalla formula di Taylor per g con resto alla Lagrange. $\square$

6.3 Funzione implicita ed invertibilità locale

Consideriamo l'equazione

$$F(x,y) = 0 \tag{6.57}$$

con $F : \Omega \subset \mathbb{R}^2 \to \mathbb{R}$. Siamo interessati a descrivere l'insieme delle sue soluzioni (o zeri di F)

$$\mathcal{Z} := \mathcal{Z}(F) = \{(x,y) \in \Omega \mid F(x,y) = 0\}$$

Definizione 6.55 (Grafico di una funzione)**.** Sia $f : A \subset \mathbb{R}^N \to \mathbb{R}^M$ definiamo il *grafico di f* nel dominio A

$$\Gamma(f;A) := \{(x,y) \in A \times \mathbb{R}^M \mid y = f(x)\}.$$

In alcuni casi si può esplicitare (6.57) nella forma

$$y = f(x), \quad x \in I \subset \mathbb{R}$$

e in tali casi si ha

$$\mathcal{Z} = \{(x, f(x)),\ x \in I\} = \Gamma(f;I),$$

e si dice che (6.57) definisce implicitamente la funzione $y = f(x)$ o, equivalentemente, che $y = f(x)$ è definita implicitamente da (6.57).

Vediamo alcuni esempi.

Esempio 6.56.

1. Sia $F(x,y) = ax+by+c$, con $a^2+b^2 > 0$. $\mathcal{Z}$ è una retta e (6.57) definisce implicitamente una funzione $y = f(x)$ o $x = h(y)$ (entrambe se $a \neq 0$ e $b \neq 0$).

2. Sia $F(x,y) = xe^y - 3y^2 + 1$. Possiamo facilmente esplicitare x in funzione di $y : x = (3y^2 - 1)e^{-y}$. E' possibile esplicitare anche y in funzione di x ?

3. Sia $F(x,y) = x^2 + y^2 - 1$. $\mathcal{Z}$ è la circonferenza di centro l'origine e raggio 1. Non è il grafico di alcuna funzione nè del tipo $y = f(x)$, nè del tipo $x = h(y)$. Possiamo però definire $y = f(x) = \sqrt{1-x^2}$, per $x \in [-1,1]$ e $y = -f(x) = -\sqrt{1-x^2}$, per $x \in [-1,1]$. Si ha che $\mathcal{Z} = \Gamma(f) \cup \Gamma(-f)$, cioè $\mathcal{Z}$ è unione di due grafici. Un'analisi più dettagliata mostra che per ogni punto $P_0 \in \mathcal{Z}$ non appartenente agli assi cartesiani, esiste un suo intorno U in cui si può esplicitare ciascuna variabile in funzione dell'altra, mentre i quattro punti in cui $\mathcal{Z}$ interseca gli assi hanno un intorno in cui si può esplicitare una variabile, ma nessun intorno in cui si possa esplicitare anche l'altra variabile. In conclusione per ogni punto $P_0 \in \mathcal{Z}$, esiste un suo intorno U tale che $\mathcal{Z} \cap U$ è il grafico di una funzione.

Definizione 6.57. Diremo che $\mathcal{Z}$ è *localmente grafico* se per ogni $P_0 \in \mathcal{Z}$ esiste U intorno di P_0 e f funzione tale che $\mathcal{Z} \cap U = \Gamma(f)$.

Esempio 6.58.

i Sia $F(x,y) = x^2 + y^2 + 1$, $\mathcal{Z} = \emptyset$.

ii Sia $F(x,y) = x^2 + y^2$. $\mathcal{Z} = \{O = (0,0)\}$, è un punto isolato.

iii Sia $F(x,y) = x^2 - y^2$. $\mathcal{Z}$ è unione di due rette passanti per l'origine (dette rami). Non è localmente grafico perchè non è grafico di una funzione in nessun intorno dell'origine (appartenente ad $\mathcal{Z}$). In questo caso diremo che (6.57) definisce implicitamente due funzioni in ogni intorno dell'origine.

iv Sia $F(x,y) = xe^y + ye^x$. Non si vede come esplicitare una variabile, si nota facilmente che $F(0,0) = 0$, cioè che $P_0 = (0,0) \in \mathcal{Z}$. Come sarà $\mathcal{Z}$ in un intorno dell'origine? Per rispondere a questa domanda avremo bisogno di conoscere la teoria delle funzioni implicite.

6.3.1 Il teorema di Dini.

Definizione 6.59. Sia $F:\Omega\subset\mathbb{R}^2\to\mathbb{R}$. Diremo che una funzione $y=f(x), x\in I\subset\mathbb{R}$ è definita implicitamente dall'equazione (6.57) se $\Gamma\left(f;I\right)\subset Z$.

Teorema 6.60 (del Dini)**.** *Sia $\Omega\subset\mathbb{R}^2$ aperto e $F\in\mathcal{C}^1(\Omega;\mathbb{R})$. Sia $\left(x_0,y_0\right)\in\Omega$ tale che $F\left(x_0,y_0\right)=0$ e $F_y\left(x_0,y_0\right)\neq 0$. Allora esistono un intorno U di x_0 ed un intorno V di y_0 ed una funzione $f\in\mathcal{C}^1(U;V)$ tali che $y_0=f(x_0)$, $F\left(x,f(x)\right)=0$ per ogni $x\in U$ e*

$$f'(x)=-\frac{F_x\left(x,f(x)\right)}{F_y\left(x,f(x)\right)},\qquad\forall\, x\in U. \tag{6.58}$$

Dimostrazione. Supponiamo ad esempio che $F_y\left(x_0,y_0\right)>0$. Essendo F_y continua, per il teorema di permanenza del segno esiste un rettangolo chiuso $R:=\bar{W}\times\bar{V}:=[x_0-k,x_0+k]\times\left[y_0-h,y_0+h\right]$ tale che $F_y>0$ in R. Quindi per ogni $x\in W$ la funzione

$$\begin{array}{cccc} F(x,\cdot) & : & V & \to & \mathbb{R} \\ & & y & \mapsto & F\left(x,y\right) \end{array}$$

è strettamente crescente. In particolare, per $x=x_0$, si ha

$$F\left(x_0,y_0-h\right)<F\left(x_0,y_0\right)=0<F\left(x_0,y_0+h\right).$$

Per ogni $\bar{y}\in\bar{V}\setminus\left\{y^0\right\}$ posso applicare il teorema della permanenza del segno alla funzione $x\mapsto F\left(x,\bar{y}\right)$ ottenendo che esiste un $U_{\bar{y}}\subset W$ di x^0 tale che $F\left(x,\bar{y}\right)\big|_{U_{\bar{y}}}>0$. Pertanto esiste un intorno $U:=[x_0-l,x_0+l]\subset W$ tale che $F\left(x,y_0-h\right)<0$ e $F\left(x,y_0+h\right)>0$ per ogni $x\in U$. Applicando il teorema dell'esistenza degli zeri a $F(x,\cdot)$ in V e ricordando che $F(x,\cdot)$ è strettamente crescente in V, si ha che per ogni $x\in U$ esiste un unico $y\in V$ tale che $F\left(x,y\right)=0$. Resta così definita la funzione f. Proviamo che f è continua. Dati $x_1,x_2\in U$, consideriamo i punti $P_1:=\left(x_1,f(x_1)\right)\in U\times V$ e $P_2:=\left(x_2,f(x_2)\right)\in U\times V$. Il segmento chiuso $[P_1,P_2]$ è contenuto nel rettangolo R. Usando (6.17), esiste un punto $(\xi,\eta)\in[P_1,P_2]$ tale che

$$F(P_2)-F(P_1)=\nabla F(\xi,\eta)\cdot(P_2-P_1).$$

Dato che $F(P_1)=F(P_2)=0$, si ha

$$F_x(\xi,\eta)(x_2-x_1)+F_y(\xi,\eta)\left(f(x_2)-f(x_1)\right)=0$$

da cui

$$f(x_2)-f(x_1)=-\frac{F_x(\xi,\eta)(x_2-x_1)}{F_y(\xi,\eta)}$$

Applicando il teorema di Weierstrass alle funzioni F_x e F_y, continue sul compatto R, definisco $M=\max_R|F_x|$ e $m=\min_R F_y>0$. Si ha dunque

$$\left|f(x_2)-f(x_1)\right|\leq\frac{M}{m}|x_2-x_1|$$

e pertanto $f\in\mathcal{C}^{0,1}(U;V)$ (cf. Definizione 5.20). Consideriamo ora il rapporto incrementale

$$\frac{f(x_2)-f(x_1)}{x_2-x_1}=-\frac{F_x(\xi,\eta)}{F_y(\xi,\eta)}$$

Per la continuità di f, se x_2 tende a x_1 si ha che $f(x_2)$ tende a $f(x_1)$, cioè P_2 tende a P_1 e quindi anche $(\xi,\eta)\in[P_1,P_2]$ tende a P_1. Dunque, per la continuità di F_x e F_y,

$$\exists\lim_{x_2\to x_1}\frac{f(x_2)-f(x_1)}{x_2-x_1}=-\frac{F_x\left(x_1,f(x_1)\right)}{F_y\left(x_1,f(x_1)\right)},\qquad\forall x_1\in U,$$

dimostrando che f è derivabile in U. Siccome $F\in\mathcal{C}^1(\Omega;\mathbb{R})$ ed $F_y\big|_{U\times V}>0$ otteniamo pure che la funzione derivata f' è continua e pertanto $f\in\mathcal{C}^1(U;V)$, dimostrando (6.58) e concludendo la dimostrazione del Teorema 6.60. □

Osservazione 6.61. 1. E' evidente che se l'ipotesi $F_y\left(x_0,y_0\right)\neq 0$ è sostituita da $F_x\left(x_0,y_0\right)\neq 0$, allora, invertendo il ruolo delle variabili x e y, si ha una seconda versione del teorema di Dini, simmetrica alla precedente, e cioè in tal caso esistono un intorno U di x_0 ed un intorno V di y_0 tali che per ogni $y\in V$ esiste un unico $x\in U$ tale che $F\left(x,y\right)=0$. Indicato con $h\left(y\right)$ tale x, resta definita una funzione $h:V\to U$, tale che $h\in C^1(V)$ e

$$h'\left(y\right)=-\frac{F_y(h\left(y\right),y)}{F_x(h\left(y\right),y)},\quad \text{per ogni } y\in V$$

2. Se $F\in C^1(\Omega), P_0=\left(x_0,y_0\right)\in \mathcal{Z}(F)$ e $\nabla F\left(x_0,y_0\right)\neq 0$, allora, in un opportuno intorno di $P_0, \mathcal{Z}(F)$ è il grafico di una funzione.

3. Se $\nabla F(P)\neq 0$, per ogni $P\in \mathcal{Z}(F)$, allora $\mathcal{Z}(F)$ è localmente grafico.

6.3.2 Inversione locale e funzione implicita negli spazi di Banach

Il Teorema 6.60 è la formulazione più semplice possibile di un vasto insieme di teoremi che identificano condizioni di invertibilità per funzioni tra spazi di Banach. Tali teoremi sono estremamente importanti nell studio delle equazioni alle derivate parziali, dove si ha bisogno di teoremi di inversione tra spazi che sono infinito-dimensionali.

Definizione 6.62 (Diffeomorfismo)**.** Sia $k\in\mathbb{N}$, X,Y di Banach, $\Omega\times\Theta\subset X\times Y$. Diciamo che $f\in\mathcal{C}^k\left(\Omega;\Theta\right)$ è un *k-diffeomorfismo* tra Ω ed Θ se f ammette un'inversa $f^{-1}\in\mathcal{C}^k\left(\Theta;\Omega\right)$. Denotiamo con $\mathrm{Diff}_k\left(\Omega;\Theta\right)$ l'insieme dei k-diffeomorfismi tra Ω e Θ. Un 0-diffeomorfismo è detto *omeomorfismo* e denotiamo $\mathrm{Hom}\left(\Omega;\Theta\right):=\mathrm{Diff}_0\left(\Omega;\Theta\right)$.

Teorema 6.63 (di inversione locale in spazi di Banach)**.** *Siano X,Y di Banach, $x^0\in\Omega\subset X$, $k\geq 1$, $f\in\mathcal{C}^k\left(\Omega;Y\right)$ con $\mathrm{d}f\left(x^0\right)$ invertibile*[2]*. Allora esiste U di x^0 e V di $y^0:=f\left(x^0\right)$ tali che $f|_U\in\mathrm{Diff}_k\left(U;V\right)$ e si ha che per ogni $y\in V$ vale $\mathrm{d}g\left(y\right)=\mathrm{d}f\left(g\left(y\right)\right)^{-1}$.*

La dimostrazione del Teorema 6.63 è piuttosto semplice, indirizziamo il lettore interessato a [1, Thm. 1.2].

Osservazione 6.64. È facile visualizzare il risultato del Teorema 6.63 nel caso in cui $X=Y=\mathbb{R}$. In tal caso la condizione di invertibilità di $\mathrm{d}f\left(x^0\right)$ diventa $f'\left(x^0\right)\neq 0$, siccome $f\in\mathcal{C}^1$ allora f' è continua e pertanto esiste un intorno U di x^0 nel quale $f'\left(x\right)\neq 0$, quindi la funzione f è localmente monotona, quindi iniettiva ed ovviamente suriettiva se ristretta nel dominio U e codominio $V:=f\left(U\right)$ e pertanto localmente invertibile. La formula di derivazione della funzione inversa dà quindi la formula cercata.

Esempio 6.65. Sia $\Omega\subset\mathbb{R}^{\mathrm{d}}$ un dominio liscio[3], consideriamo *l'operatore Laplaciano* $\Delta=\sum_{i=1}^{\mathrm{d}}\partial_i^2$, sia $h\in\mathcal{C}\left(\Omega;\mathbb{R}\right)$ e la seguente equazione ellittica semilineare

$$\begin{cases}\Delta u+\lambda u-u^3=h & \text{in } \Omega\\ u=0 & \text{in } \partial\Omega\end{cases},\tag{6.59}$$

dove $\lambda\in\mathbb{R}$. Sia $k\in\mathbb{N}$ e $X_k:=\left\{u\in\mathcal{C}^k\left(\bar{\Omega};\mathbb{R}\right)\ \middle|\ u|_{\partial\Omega}\equiv 0\right\}$, definiamo $T_\lambda\left(u\right):=\Delta u+\lambda u-u^3$, certamente $T_\lambda: X_{k+2}\to X_k$ per ogni k, dunque l'equazione (6.59) è risolta dall'elemento (se esiste)

$$u:=T_\lambda^{-1}\left(h\right).$$

L'esistenza della funzione T_λ^{-1} (come applicazione tra gli spazi di Banach X_k e X_{k+2}) può essere dedotta usando il Teorema 6.63.

Teorema 6.66 (Teorema della Funzione Implicita, [1], Teorema 2.3)**.** *Sia $F\in C^k\left(\Lambda\times U;Y\right)$, $k\geq 1$, dove Y è uno spazio di Banach e Λ (rispettivamente U) è un sottoinsieme aperto dello spazio di Banach T (rispettivamente X). Supponiamo che $F\left(\lambda^*,u^*\right)=0$ e che $\mathrm{d}_uF\left(\lambda^*,u^*\right)\in\mathcal{L}\left(U;Y\right)$ sia invertibile. Allora esistono intorni Θ di λ^* in T e U^* di u^* in X e una mappa $f\in\mathcal{C}^k\left(\Theta;X\right)$ tale che*

[2] Ossia esiste $B\in\mathcal{L}\left(Y;X\right)$ tale che $\mathrm{d}f\left(x^0\right)\circ B=\mathrm{Id}_X$ e $B\circ\mathrm{d}f\left(x^0\right)=\mathrm{Id}_Y$.

[3] Il bordo di Ω è localmente il grafico di una funzione liscia

1. $F(\lambda, f(\lambda)) = 0$ *per ogni* $\lambda \in \Theta$,
2. $F(\lambda, u) = 0$, $(\lambda, u) \in \Theta \times U^*$, *implica* $u = f(\lambda)$,
3. $\mathrm{d}f(\lambda) = -\left[\mathrm{d}_u F(\lambda, f(\lambda))\right]^{-1} \circ \mathrm{d}_\lambda F(\lambda, f(\lambda))$, *con* $\lambda \in \Theta$.

7 Massimi e minini per funzioni di più variabili

Definizione 7.1 (Massimi e minimi assoluti e relativi)**.** Data una funzione a valori reali, $f : \Omega \subset \mathbb{R}^{\mathrm{d}} \to \mathbb{R}$, un punto $x^0 \in \Omega$ si dice punto di *massimo assoluto o globale* se $f(x^0) \geq f(x)$ per ogni $x \in \Omega$; si dice punto di *massimo relativo o locale* se esiste un intorno U di x^0 tale che $f(x^0) \geq f(x)$ per ogni $x \in \Omega \cap U$. Il valore $f(x^0)$ è detto *massimo assoluto* (o semplicemente massimo) di f nel primo caso e *massimo relativo* di f nel secondo caso. Analoghe definizioni si danno per i minimi. Un punto di massimo o di minimo assoluto/-relativo è detto punto di estremo assoluto/relativo; analogamente si parla di estremi assoluti/relativi per i valori assunti dalla funzione nei punti di estremo.

Lemma 7.2. *Sia* $f : \Omega \subset \mathbb{R}^{\mathrm{d}} \to \mathbb{R}$ *e sia* x^0 *un punto di estremo relativo per* f *interno a* Ω*. Se* f *è differenziabile in* x^0*, allora* x^0 *è un* punto critico*, ovvero* $\nabla f(x^0) = 0$.

Dimostrazione. Se x^0 è estremo relativo allora $\partial_v f(x^0) = 0$ per ogni $v \in \mathbb{S}^{\mathrm{d}-1}$, pertanto se $v = \mathrm{e}_i$ otteniamo la tesi. □

Osservazione 7.3. Come nel caso uno dimensionale non è vero il contrario, basti considerare la funzione $f(x, y) = x^2 - y^2$ ed il punto $(0, 0)$.

Definizione 7.4. Data una funzione $f : \Omega \subset \mathbb{R}^{\mathrm{d}} \to \mathbb{R}$, un punto $x^0 \in \Omega$, critico per f, si dice *punto di sella* di f se in ogni intorno di $x^0 f$ assume sia valori strettamente maggiori di $f(x^0)$, sia valori strettamente minori di $f(x^0)$.

Il punto $(0, 0)$ nell'Osservazione 7.3 è di sella per f.

Lemma 7.5. *Sia* $A \in \mathbb{R}^{\mathrm{d} \times \mathrm{d}}$ *simmetrica si ha che*

$$m\|x\|^2 \leq x \cdot Ax \leq M\|x\|^2, \qquad \forall x \in \mathbb{R}^{\mathrm{d}}, \tag{7.1}$$

con m *e* M *il minimo e il massimo autovalore di* A.

Dimostrazione. Consideriamo la funzione

$$\begin{array}{cccc} F & : & \mathbb{R}^{\mathrm{d}} \setminus \{0\} & \to & \mathbb{R} \\ & & x & \mapsto & \dfrac{x \cdot Ax}{\|x\|^2} = \dfrac{\sum_{i,j=1}^{\mathrm{d}} a_{ij} x_i x_j}{\sum_{i=1}^{\mathrm{d}} x_i^2}, \end{array}$$

si ha che $F \in \mathcal{C}^\infty\left(\mathbb{R}^{\mathrm{d}} \setminus \{0\}; \mathbb{R}\right)$ (cf. Definizione 6.43). Si ha che

$$\frac{\partial F}{\partial x_k}(x) = \frac{2\sum_{j=1}^{\mathrm{d}} a_{kj} x_j \|x\|^2 - 2x_k x \cdot Ax}{\|x\|^4} = \frac{2\sum_{j=1}^{\mathrm{d}} a_{kj} x_j - 2x_k F(x)}{\|x\|^2}$$

da cui

$$\nabla F(x) = 2\frac{Ax - F(x)\, x}{\|x\|^2}$$

□

Dal Lemma 7.2 segue che se x^0 è un punto critico di F, allora $F(x^0)$ è un autovalore della matrice A e x^0 è un corrispondente autovettore pertanto vogliamo dimostrare ora che F assume massimo e minimo assoluti. Per Weierstrass esistono $\hat{x}_m, \hat{x}_M \in \mathbb{S}^{\mathrm{d}-1}$ per i quali $F|_{\mathbb{S}^{\mathrm{d}-1}}$ assume massimo e minimo, tuttavia F è 0-omogenea pertanto $F(x) = F(\hat{x})$ con $\hat{x} = \frac{x}{\|x\|}$, quindi $F(\hat{x}_m)$ e $F(\hat{x}_M)$ sono gli autovalori associati a $\hat{x}_m$ e $\hat{x}_M$. Dimostriamo m e M sono il massimo ed il minimo autovalore, supponiamo esista $\lambda < m$ e $\hat{x}_\lambda \in \mathbb{S}^{\mathrm{d}-1}$ tale che $A\hat{x}_\lambda = \lambda \hat{x}_\lambda$, in tal caso si ha che $F(\hat{x}_\lambda) = \lambda < m = F(\hat{x}_m)$, che è assurdo.

Definizione 7.6. Sia $A \in \mathbb{R}^{d\times d}$ simmetrica si dice *semidefinita positiva* se $x \cdot Ax \geq 0$, per ogni $x \in \mathbb{R}^d$; si dice *definita positiva* se $x \cdot Ax > 0$, per ogni $x \in \mathbb{R}^d$, $x \neq 0$.

Lemma 7.7. *Una matrice simmetrica A:*

1. *è semidefinita positiva se e solo se tutti i suoi autovalori sono non negativi;*
2. *è definita positiva se e solo se tutti i suoi autovalori sono positivi;*
3. *è semidefinita negativa se e solo se tutti i suoi autovalori sono non positivi;*
4. *è definita negativa se e solo se tutti i suoi autovalori sono negativi.*

Dimostrazione. Se gli autovalori di A sono tutti non negativi, allora il minimo di essi, m, verifica $m \geq 0$ e dalla prima disuguaglianza in (7.1), segue che $x \cdot Ax \geq 0$, per ogni $x \in \mathbb{R}^d$. Sia A semidefinita positiva e supponiamo, per assurdo, che A abbia un autovalore negativo λ. Allora, se $x \in \mathbb{R}^d \setminus \{0\}$ è un autovettore relativo a λ, si ha $x \cdot Ax = \lambda \|x\|^2 < 0$, contraddicendo l'ipotesi. La dimostrazione è analoga per matrici semidefinite negative e definite negative. □

Proposizione 7.8 (Criterio della matrice Hessiana)**.** *Sia* $f \in \mathcal{C}^2\left(\mathbb{R}^d;\mathbb{R}\right)$. *Sia* $x^0 \in \Omega$ *e* $\nabla f\left(x^0\right) = 0$.

1. *Affinché* x^0 *sia punto di minimo relativo per* f *:*
   (a) *è necessario che* $Hf\left(x^0\right)$ *sia semidefinita positiva;*
   (b) *è sufficiente che* $Hf\left(x^0\right)$ *sia definita positiva.*
2. *Affinché* x^0 *sia punto di massimo relativo* per f *:*
   (a) *è necessario che* $Hf\left(x^0\right)$ *sia semidefinita negativa;*
   (b) *è sufficiente che* $Hf\left(x^0\right)$ *sia definita negativa.*

Dimostrazione. Consideriamo il caso del minimo, l'altro essendo analogo.

a Se x^0 è un punto di minimo relativo per f, allora per ogni direzione $v \in \mathbb{R}^d$, la funzione $g(t) = f\left(x^0 + tv\right)$ ha un minimo relativo per $t = 0$, per cui deve essere $g''(0) \geq 0$. D'altra parte, da (6.54) si ha che $g''(0) = v \cdot Hf\left(x^0\right) v$, concludendo.

b Consideriamo la formula di Taylor di ordine 2 di centro x^0 :

$$f(x) = f\left(x^0\right) + \frac{1}{2}\left(x - x^0\right) \cdot Hf\left(x^0\right)\left(x - x^0\right) + R_f^{[2]}\left(x - x^0; x^0\right),$$

con $R_f^{[2]}\left(x - x^0; x^0\right) = o\left(\left\|x - x^0\right\|^2\right)$. Se $m > 0$ è il più piccolo autovalore di $Hf\left(x^0\right)$, si ha

$$f(x) - f\left(x^0\right) \geq \frac{1}{2} m \left\|x - x^0\right\|^2 + R_f^{[2]}\left(x - x^0; x^0\right).$$

Esiste inoltre $\delta > 0$ tale che $\left|R_f^{[2]}\left(x - x^0; x^0\right)\right| < \frac{m}{2}\left\|x - x^0\right\|^2$ per ogni $x \in B\left(x^0, \delta\right)$, $x \neq x_0$. Pertanto per ogni $x \in B\left(x^0, \delta\right)$, $x \neq x_0$, si ha che $f(x) > f\left(x^0\right)$. Quindi x^0 è punto di minimo relativo stretto.

□

Se gli autovalori di $Hf\left(x^0\right)$ hanno tutti lo stesso segno ma almeno uno di essi è nullo, il criterio della matrice Hessiana non fornisce informazioni.

7.1 Estremi vincolati.

Considereremo ora situazioni in cui si ricercano gli estremi (massimo/minimo) di una funzione di più variabili, le cui variabili non sono indipendenti l'una dall'altra, ma sono soggette a delle condizioni, espresse da una o più equazioni, dette vincoli.

Definizione 7.9 (Vincolo)**.** Sia Ω un aperto di $\mathbb{R}^m$, con $m \geq 2$. Siano $f : \Omega \subset \mathbb{R}^m \to \mathbb{R}$ e $F : \Omega \subset \mathbb{R}^m \to \mathbb{R}^n$, con $n < m$. Dunque $F = (F_1, \dots, F_n)^\intercal$. Sia $\mathcal{Z}(F) := \{x \in \Omega : F(x) = 0\} = \bigcap_{i=1}^n \mathcal{Z}(F_i)$. L'insieme $\mathcal{Z}(F)$ è detto *vincolo*. Il termine vincolo è utilizzato anche per indicare l'equazione $F = 0$ o la funzione F.

Definizione 7.10 (Estremo vincolato)**.** Un punto $x_0 \in \Omega$ si dice punto di *massimo vincolato* (o condizionato) per f con vincolo $F = 0$ se $F(x_0) = 0$ e $f(x) \leq f(x_0)$ per ogni $x \in \mathcal{Z}(F)$.
Un punto $x_0 \in \Omega$ si dice punto di *massimo vincolato* (o condizionato) relativo (o locale) per f con vincolo $F = 0$ se $F(x_0) = 0$ ed esiste un intorno U di x_0 tale che $f(x) \leq f(x_0)$ per ogni $x \in \mathcal{Z}(F) \cap U$. Il valore $f(x_0)$ è detto massimo vincolato.
Analoghe definizioni si danno per i punti di minimo vincolati. Un punto di estremo vincolato è un punto di massimo o minimo vincolato.

Teorema 7.11 (dei moltiplicatori di Lagrange)**.** *Sia $f : \Omega \subset \mathbb{R}^m \to \mathbb{R}$, differenziabile su Ω, con Ω aperto e $m \geq 2$. Sia $F : \Omega \subset \mathbb{R}^m \to \mathbb{R}^n$, con $n < m$, $F \in \mathcal{C}^1(\Omega; \mathbb{R}^n)$. Sia $x_0 \in \Omega$ un punto di estremo relativo vincolato per f con vincolo $F = 0$, tale che* $\mathrm{rk}(JF(x_0)) = n$*. Allora esistono $\lambda_1, \dots, \lambda_n \in \mathbb{R}$, detti* moltiplicatori di Lagrange *tali che*

$$F(x_0) = 0 \qquad e \qquad \nabla f(x_0) + \sum_{i=1}^n \lambda_i \nabla F_i(x_0) = 0. \tag{7.2}$$

Osservazione 7.12.

1. Posto $\lambda = (\lambda_1, \dots, \lambda_n)$, si ha che $(x_0, \lambda) \in \mathbb{R}^{m+n}$ verifica (7.2) che è un sistema di $m + n$ equazioni in $m + n$ incognite

$$\begin{cases} F_1(x_1, \dots, x_m) = 0, \\ \vdots \\ F_n(x_1, \dots, x_m) = 0 \\ \left(f + \sum_{i=1}^n \lambda_i F_i\right)_{x_1}(x_1, \dots, x_m) = 0, \\ \vdots \\ \left(f + \sum_{i=1}^n \lambda_i F_i\right)_{x_m}(x_1, \dots, x_m) = 0. \end{cases} \tag{7.3}$$

2. Il numero dei moltiplicatori è uguale al numero dei vincoli. La seconda equazione in (7.2) esprime $\nabla f(x_0)$ come combinazione lineare dei vettori (linearmente indipendenti per ipotesi) $\nabla F_i(x_0)$. Perciò i moltiplicatori di Lagrange sono univocamente determinati.

3. La condizione di estremo vincolato relativo è meno restrittiva rispetto a quella di estremo libero relativo e, corrispondentemente, questo teorema fornisce una condizione necessaria affinchè un punto sia di estremo vincolato relativo, che è meno restrittiva rispetto alla condizione necessaria $\nabla f(x_0) = 0$ soddisfatta dagli estremi liberi.

4. Un metodo per ricordare il sistema (7.2) consiste nell'introdurre la funzione $\varphi(x, \lambda) = f(x) + \sum_{i=1}^n \lambda_i F_i(x)$ e imporre $\nabla_x \varphi = 0$.

5. Il teorema fornisce una condizione necessaria, che è utile per cercare gli estremi vincolati assoluti (quando esistono). Precisamente, se sappiamo che f assume massimo e minimo assoluti su $\mathcal{Z}(F)$ (ad esempio grazie all'applicazione del teorema di Weierstrass), si determinano le soluzioni del sistema (7.3) oltre agli eventuali punti di $\mathcal{Z}(F)$ in cui non è soddisfatta la condizione di rango massimo. Dal confronto dei valori assunti da f in tali punti, si determinano gli estremi assoluti.

Dimostrazione. Siccome per ipotesi $\mathrm{rk}(JF(x_0)) = n$ non è restrittivo, definendo $k := m - n$, supporre che

$$\det\left(\frac{\partial(F_1, \dots, F_n)}{\partial(x_{k+1}, \dots, x_m)}(x_0)\right) \neq 0. \tag{7.4}$$

Riscriviamo il generico punto $x = (x_1, \dots, x_k, x_{k+1}, \dots, x_m)$ nella forma $(x, y) = (x_1, \dots, x_k, y_1, \dots, y_n)$. In tal modo il punto x_0 viene riscritto come (x_0, y_0) e la condizione (7.4) diventa

$$\det\left(\frac{\partial F}{\partial y}(x_0, y_0)\right) \neq 0. \tag{7.5}$$

L'Eq. (7.5) garantisce che il differenziale nella variabile y di F è invertibile, pertanto posso applicare Teorema 6.66 ottenendo che esistono un intorno W di x_0, un intorno V di y_0 ed una funzione $g \in \mathcal{C}^1(W; V)$, tale che $g(x_0) = y_0$ e $F(x, g(x)) \equiv 0$ in W. Equivalentemente, $(x, g(x)) \in \mathcal{Z}(F)$ per ogni $x \in W$. Essendo $(x_0, g(x_0))$ punto di estremo relativo vincolato per f con vincolo $\mathcal{Z}(F)$, allora x_0 è punto di estremo relativo libero per la funzione $\varphi(x) = f(x, g(x))$, definita in W. Quindi

$$\nabla\varphi(x_0) = \left(\frac{\partial f}{\partial x} + \frac{\partial f}{\partial y}\frac{\partial g}{\partial x}\right)(x_0, y_0) = 0.$$

Dal Teorema 6.66 abbiamo che $\frac{\partial g}{\partial x} = -\left[\frac{\partial F}{\partial y}\right]^{-1}\frac{\partial F}{\partial x}\Big|_{(x,g(x))}$ per ogni $x \in W$ e quindi $\left(\frac{\partial f}{\partial x} - \frac{\partial f}{\partial y}\left[\frac{\partial F}{\partial y}\right]^{-1}\frac{\partial F}{\partial x}\right)(x_0, y_0) = 0$. Dunque (x_0, y_0) risolve un sistema di k equazioni in m incognite, ma, se aggiungiamo le n equazioni contenute nella condizione $F = 0$, otteniamo il seguente sistema di m equazioni in m incognite[4]

$$\begin{cases} \left(\dfrac{\partial f}{\partial x} - \dfrac{\partial f}{\partial y}\left[\dfrac{\partial F}{\partial y}\right]^{-1}\dfrac{\partial F}{\partial x}\right)(x_0, y_0) = 0 \\ F(x_0, y_0) = 0 \end{cases}$$

Definiamo

$$\lambda := (\lambda_1, \dots, \lambda_n) := -\frac{\partial f}{\partial y}\left[\frac{\partial F}{\partial y}\right]^{-1}\Bigg|_{(x_0,y_0)},$$

da cui

$$\left(\frac{\partial f}{\partial y} + \lambda\frac{\partial F}{\partial y}\right)(x_0, y_0) = 0,$$

pertanto otteniamo

$$\left(\frac{\partial f}{\partial x} + \lambda\frac{\partial F}{\partial x}\right)(x_0, y_0) = 0.$$

□

8 Successioni e serie di funzioni

8.1 Successioni di funzioni

Sia I un intervallo, eventualmente illimitato, e sia f_n una successione di funzioni a valori reali definite in I.

Definizione 8.1 (Convergenza puntuale)**.** Diremo che la successione f_n converge puntualmente in I a f se per ogni $x \in I$ si ha

$$f(x) = \lim_{n\to\infty} f_n(x).$$

Osservazione 8.2. La Definizione 8.1 implica che per ogni $x \in I$ e per ogni $\epsilon > 0$ esiste un indice $\bar{n} = \bar{n}(\epsilon, x)$ tale che per ogni $n \geq \bar{n}$

$$|f_n(x) - f(x)| < \epsilon.$$

[4] Questa formula ha vari svantaggi: comporta il calcolo dell'inversa di una matrice e del prodotto di tre matrici; inoltre essa è stata ottenuta supponendo che, riutilizzando la notazione iniziale, il minore costituito dalle ultime n righe della matrice $JF(x_0)$ avesse determinante non nullo e, nell'applicazione pratica del teorema tale minore può variare al variare del punto $x_0 \in \Omega$. Ci troveremmo quindi, in generale, a dover risolvere più di un sistema di tale tipo.

Definizione 8.3 (Convergenza uniforme)**.** Diremo che la successione f_n è uniformemente convergente in I a una funzione f se per ogni $\epsilon > 0$ esiste un indice $\bar{n} = \bar{n}(\epsilon)$ tale che per ogni $n \geq \bar{n}$

$$\sup_{x\in I} \left| f_n(x) - f(x) \right| \leq \epsilon,$$

oppure utilizzando la notazione in (5.2)

$$\lim_{n\to\infty} \left\| f_n - f \right\|_\infty = 0.$$

Osservazione 8.4. Notiamo la differenza nella definizione di $\bar{n}$ in Osservazione 8.2 e Definizione 8.3: nel primo caso dipende a $x \in I$, nes secondo caso no.

Esempio 8.5 (Convergenza puntuale ma non uniforme)**.** Sia $I = [0,1]$. Si consideri la successione di funzioni f_n definite da

$$f_n(x) = 1 - x^n, x \in I.$$

Si verifica immediatamente che f_n converge alla funzione

$$f(x) = \begin{cases} 1, & \text{se } 0 \leq x < 1 \\ 0, & \text{se } x = 1 \end{cases}$$

Dunque,

$$f(x) - f_n(x) = \begin{cases} x^n, & \text{se } 0 \leq x < 1 \\ 0, & \text{se } x = 1 \end{cases}$$

Si ha allora che, per ogni $n \in \mathbb{N}$,

$$\sup_{0\leq x\leq 1} \left| f_n(x) - f(x) \right| = \sup_{0\leq x<1} x^n = \left(\sup_{0\leq x<1} x \right)^n = 1$$

per cui la convergenza di f_n a f non è uniforme.

Proposizione 8.6 (Continuità del limite uniforme di funzioni continue)**.** *Sia f_n una successione di funzioni uniformemente convergente in I a una funzione f. Se tutte le funzioni della successione sono continue in un punto $x_0 \in I$, allora f è continua in x_0. In particolare, se le f_n sono continue in I, la funzione limite f è continua in I.*

Dimostrazione. Sappiamo che per ogni $\epsilon > 0$ esiste un indice $\bar{n} = \bar{n}(\epsilon)$ tale che per ogni $n \geq \bar{n}$

$$\left| f_n(x) - f(x) \right| < \epsilon, \forall x \in I.$$

Dato che $f_{\bar{n}}$ è continua in x_0,

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x \in I, \quad |x - x_0| < \delta \Rightarrow \left| f_{\bar{n}}(x) - f_{\bar{n}}(x_0) \right| < \epsilon$$

Allora, per ogni $x \in I$ tale che $|x - x_0| < \delta$, si ha

$$\left| f(x) - f(x_0) \right| \leq \left| f(x) - f_{\bar{n}}(x) \right| + \left| f_{\bar{n}}(x) - f_{\bar{n}}(x_0) \right| + \left| f_{\bar{n}}(x_0) - f(x_0) \right| < 3\epsilon,$$

quindi f è continua in x_0. □

Proposizione 8.7 (Passaggio al limite sotto segno di derivata)**.** *Sia f_n una successione di funzioni convergente puntualmente nell'intervallo chiuso $I = [a,b]$ ad una funzione f. Inoltre le funzioni f_n siano tutte derivabili in $[a,b]$ e la successione f_n' sia uniformemente convergente in $[a,b]$ ad una funzione g. Allora f è derivabile in $[a,b]$ e*

$$f'(x) = g(x), \qquad \forall x \in [a,b].$$

Dimostrazione. Fisso $x_0 \in [a,b]$, definiamo

$$\varphi_n(h) := \begin{cases} \dfrac{f_n(x_0+h) - f_n(x_0)}{h}, & \text{se } h \neq 0 \\ f_n'(x_0), & \text{se } h = 0 \end{cases} \tag{8.1}$$

Le funzioni $\varphi_n \in \mathcal{C}(J;\mathbb{R})$ dove $J := [a - x_0, b - x_0]$ e convergono puntualmente per $n \to \infty$ alla funzione

$$\varphi(h) := \begin{cases} \dfrac{f(x_0+h) - f(x_0)}{h}, & \text{se } h \neq 0 \\ g(x_0), & \text{se } h = 0 \end{cases}$$

Proviamo che la successione φ_n è uniformemente convergente in J. Applicando il teorema di Lagrange, si ha

$$\begin{aligned} \left|\varphi_m(h) - \varphi_n(h)\right| &= \left|\frac{\left(f_m(x_0+h) - f_m(x_0)\right) - \left(f_n(x_0+h) - f_n(x_0)\right)}{h}\right| \\ &= \left|\frac{\left(f_m - f_n\right)(x_0+h) - \left(f_m - f_n\right)(x_0)}{h}\right| = \left|\left(f_m - f_n\right)'(x_0 + \theta h)\right|, \end{aligned} \tag{8.2}$$

dove $\theta = \theta(m,n,x_0) \in (0,1)$. Dalla convergenza uniforme della successione f_n' si ha che per ogni $\epsilon > 0$ esiste $\bar{n} \in \mathbb{N}$ tale che per ogni $m, n \geq \bar{n}$,

$$\left|f_m'(x) - f_n'(x)\right| \leq \left|f_m'(x) - g(x)\right| + \left|g(x) - f_n'(x)\right| < \epsilon, \qquad \forall x \in [a,b]. \tag{8.3}$$

Pertanto inserendo (8.3) in (8.2) ottengo che esiste $\bar{n}$ tale che per $m, n \geq \bar{n}$,

$$\left|\varphi_m(h) - \varphi_n(h)\right| < \epsilon, \qquad \forall h \in J.$$

Pertanto $\left(\varphi_n\right)_n$ è di Cauchy nello spazio normato $(\mathcal{C}(J;\mathbb{R}), \|\cdot\|_\infty)$, che è completo grazie al risultato nel Teorema 5.11 e quindi $\left(\varphi_n\right)_n$ converge uniformemente ad una funzione limite che deve essere la funzione esplicitata in (8.1) dall'unicità del limite. Dalla Proposizione 8.6 so che il limite φ è continuo in zero, ossia

$$\lim_{h \to 0} \varphi(h) = \lim_{h \to 0} \frac{f(x_0+h) - f(x_0)}{h} = g(x_0),$$

e ciò vuol dire che esiste $f'(x_0) = g(x_0)$. □

Teorema 8.8 (Passaggio al limite sotto segno di integrale)**.** *Sia $I := [a,b]$ e $\left(f_n\right)_n \in \mathcal{C}(I;\mathbb{R})$ uniformemente convergente in I a una funzione f. Allora si ha*

$$\lim_{n\to\infty} \int_a^b f_n(x)\,\mathrm{d}x = \int_a^b \lim_{n\to\infty} f_n(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x$$

Dimostrazione. Per la Proposizione 8.6 f è continua in I e pertanto integrabile in I. Per ipotesi di uniforme convergenza, per ogni $\epsilon > 0$, esiste $\bar{n} \in \mathbb{N}$ tale che per ogni $n \geq \bar{n}$ $\left|f_n(x) - f(x)\right| < \epsilon$ per ogni $x \in I$. Quindi per ogni $n \geq \bar{n}$ si ha

$$\left|\int_a^b f_n(x) - \int_a^b f(x)\mathrm{d}x\right| \leq \int_a^b \left|f_n(x) - f(x)\right|\mathrm{d}x < \epsilon(b-a).$$

□

8.2 Serie di funzioni

Sia $f_k : I \subset \mathbb{R} \to \mathbb{R}$, per ogni $k \in \mathbb{N} \setminus 0$. Definiamo la successione di funzioni φ_n

$$\varphi_n := \sum_{k=1}^{n} f_k. \tag{8.4}$$

Analogamente a quanto visto per le serie numeriche, l'espressione

$$\sum_{k=1}^{\infty} f_k \tag{8.5}$$

è detta *serie di funzioni* ed è definita convergente, assolutamente convergente, divergente in un punto $x \in I$ a seconda che tale sia la serie numerica $\sum_{k=1}^{\infty} f_k(x)$.

Se la serie (8.5) è convergente per ogni $x \in I$, rimane definita la funzione somma della serie

$$\varphi = \sum_{k=1}^{\infty} f_k = \lim_{n\to\infty} \varphi_n.$$

Definizione 8.9. Diremo che la serie (8.5) converge uniformemente con somma φ se la successione (8.4) converge uniformemente in I alla funzione φ.

Lemma 8.10 (Criterio di Cauchy per le serie). *Sia $I := [a, b]$. La serie in* (8.5) *converge uniformemente in I se e solo se*

$$\forall \epsilon > 0, \exists \bar{n} : \forall n \geq \bar{n}, \forall p \in \mathbb{N}, \left| \sum_{k=n+1}^{n+p} f_k(x) \right| < \epsilon, \forall x \in I. \tag{8.6}$$

Dimostrazione. La (8.6) implica che la successione $\left(\varphi_n\right)_n$ è di Cauchy in $(\mathcal{C}(I;\mathbb{R}), \|\cdot\|_\infty)$, che è completo grazie al Teorema 5.11, e quindi convergente. □

Definizione 8.11 (Convergenza assoluta uniforme). Se la serie $\sum_{k=1}^{\infty} \left|f_k\right|$ converge uniformemente in I, diremo che la serie $\sum_{k=1}^{\infty} f_k$ è dotata di *convergenza assoluta uniforme* ed in tal caso, dal Lemma 8.10, segue che è anche uniformemente convergente.

Scegliendo $p = 1$ in (8.6), si deduce che se una serie di funzioni converge uniformemente in I allora la successione f_k converge uniformemente alla funzione identicamente nulla in I.

Definizione 8.12 (Convergenza totale). Diremo che la serie (8.5) *converge totalmente* in I se esiste una successione di numeri reali l_k tale che

i $\left|f_k(x)\right| \leq l_k, \forall x \in I$;

ii $\sum_{k=1}^{\infty} l_k$ è convergente.

Dal Lemma 8.10 segue che una serie convergente totalmente è dotata di convergenza assoluta uniforme.

Definizione 2.3.

8.3 Serie di Taylor

Sia f una funzione infinitamente derivabile in un intervallo I e sia x_0 un punto interno a I. Per ogni $n \in \mathbb{N}$, possiamo allora scrivere la formula di Taylor di f di punto iniziale x_0 e ordine n, con resto alla Lagrange:

$$f(x) = P_n(x) + R_n(x), \tag{8.7}$$

dove

$$P_n(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k, \qquad R_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}, \tag{8.8}$$

con $\xi \in [x_0, x]$. Fissato n e facendo tendere x a x_0, si ha che

$$\lim_{x\to x_0} R_n(x) = 0$$

da cui deriva che

$$\forall \epsilon > 0, \exists \delta > 0 : \forall x, |x - x_0| < \delta \Rightarrow \left|f(x) - P_n(x)\right| < \epsilon$$

ovvero il polinomio di Taylor è una buona approssimazione di f in un intorno di x_0.

È naturale chiedersi se, fissato $x \in I$, e facendo tendere di n a $+\infty$, si verifichi

$$\lim_{n\to\infty} R_n(x) = 0 \Leftrightarrow \lim_{n\to\infty} P_n(x) = f(x)$$

in altre parole se la serie seguente, detta serie di Taylor di punto iniziale x_0,

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k$$

converga con somma $f(x)$, per ogni $x \in I$.

Definizione 8.13 (Sviluppabilità in serie di una funzione)**.** Sia $f \in \mathcal{C}^\infty(I;\mathbb{R})$, $x_0 \in \mathring{I}$ si dice che f è (localmente) *sviluppabile in serie di Taylor* se dato P_n come in (8.8) si ha che

$$\lim_{n\to\infty} \left\|f - P_n\right\|_\infty = \lim_{n\to\infty} \sup_{x\in I} \left|f(x) - P_n(x)\right| = 0.$$

Il seguente esempio illustra che una funzione infinitamente derivabile non è necessariamente sviluppabile in serie di Taylor:

Esempio 8.14. Consideriamo la funzione definita su $\mathbb{R}$ da

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}}, & \text{se } x \neq 0 \\ 0, & \text{se } x = 0 \end{cases}$$

Si verifica facilmente che f è infinitamente derivabile e $f^{(k)}(0) = 0$ per ogni $k \in \mathbb{N}$. Pertanto la serie di Taylor di f di punto iniziale 0 risulta convergente, ma ha per somma la funzione identicamente nulla.

Lemma 8.15. *Sia $f \in \mathcal{C}^\infty(I;\mathbb{R})$. Sia $x_0 \in \mathring{I}$ e $r > 0$ tale che $(x_0 - r, x_0 + r) \subset I$ ed esistano due numeri reali L, M tali che*

$$\left|f^{(k)}(x)\right| \leq ML^k, \qquad \forall x \in (x_0 - r, x_0 + r), \quad \forall k \in \mathbb{N}.$$

Allora f è sviluppabile in serie di Taylor di punto iniziale x_0 in $(x_0 - r, x_0 + r)$.

Dimostrazione. Per ogni $x \in (x_0 - r, x_0 + r)$ e per ogni $n \in \mathbb{N}$, si ha che

$$|R_n(x)| \leq \frac{M(Lr)^{n+1}}{(n+1)!} \to 0, \text{ per } n \to \infty\,.$$

□

Teorema 8.16. *Sia f infinitamente derivabile nell'intervallo I. Sia x_0 un punto interno a I e sia $r > 0$ tale che $(x_0 - r, x_0 + r) \subset I$ ed esista $M > 0$ tali che*

$$\left|f^{(k)}(x)\right| \leq Mk!r^{-k}, \forall x \in (x_0 - r, x_0 + r), \forall k \in \mathbb{N}$$

Allora f è sviluppabile in serie di Taylor di punto iniziale x_0 in $(x_0 - r, x_0 + r)$.

Dimostrazione. Per ogni $x \in (x_0 - r, x_0 + r)$ e per ogni $n \in \mathbb{N}$, si ha che

$$|R_n(x)| \leq M\left(\frac{|x - x_0|}{r}\right)^{n+1} \to 0, \text{ per } n \to \infty.$$

□

Definizione 8.17. Una $f \in \mathcal{C}^\infty(I;\mathbb{R})$ è detta *analitica* in un punto $x_0 \in I$ se è sviluppabile in serie di Taylor di punto iniziale x_0 in un intorno di x_0. Si dice analitica in I se è analitica in ogni punto di I. Infichiamo con $\mathcal{C}^\omega(I;\mathbb{R})$ l'insieme delle funzioni analitiche da I a $\mathbb{R}$.

Osservazione 8.18. L'esempio 8.14 ci dimostra che una funzione può essere $\mathcal{C}^\infty$ ma non $\mathcal{C}^\omega$. Una funzione $\mathcal{C}^\omega$ è sempre $\mathcal{C}^\infty$ per definizione.

8.4 4 Serie di potenze

Definizione 8.19. Sia $x_0 \in \mathbb{R}$, le serie di funzioni della forma

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k, \qquad x_0, x, a_k \in \mathbb{R}, \tag{8.9}$$

sono dette *serie di potenze di punto iniziale x_0 e coefficienti a_k.*

Ovviamente ogni serie di potenze converge con somma a_0 in $x = x_0$. Abbiamo visto nel precedente paragrafo esempi di serie di potenze: le serie di Taylor associate alle funzioni infinitamente derivabili.

Definizione 8.20 (Raggio di convergenza)**.** Posto $L := \limsup_{k\to\infty} \sqrt[k]{|a_k|}(0 \le L \le +\infty)$, definiamo *raggio di convergenza* della serie di potenze (8.9)

$$r := \begin{cases} 0 & \text{se } L = +\infty \\ \dfrac{1}{L} & \text{se } L \in (0, \infty) \\ +\infty & \text{se } L = 0 \end{cases} . \tag{8.10}$$

Teorema 8.21. *Sia r come in* (8.10)*, se $r > 0$, per ogni ρ tale che $0 < \rho < r$, la serie* (8.9) *converge totalmente (cf. Definizione 8.12) nell'intervallo chiuso $[x_0 - \rho, x_0 + \rho]$. Se $r < +\infty$, la serie* (8.9) *non è convergente in nessun punto x tale che $|x - x_0| > r$.*

Dimostrazione. Sia $r > 0$ (ovvero $0 \le L < +\infty$) e sia $\rho \in (0, r)$. Si ha che

$$\left|f_k(x)\right| = \left|a_k (x - x_0)^k\right| \le |a_k|\, \rho^k, \forall x \in [x_0 - \rho, x_0 + \rho]$$

Applicando il Teorema 2.31 alla serie numerica

$$\sum_{k=0}^{\infty} |a_k|\, \rho^k$$

si ha

$$\limsup_{k\to\infty} \sqrt[k]{|a_k|\, \rho^k} = \rho \limsup_{k\to\infty} \sqrt[k]{|a_k|} = \rho L < 1,$$

e quindi tale serie converge. Di conseguenza la serie (8.9) converge totalmente in $[x_0 - \rho, x_0 + \rho]$.
Se $r < +\infty$ (ovvero $0 < L \le +\infty$) e x è tale che $|x - x_0| > r$, si ha

$$\limsup_{k\to\infty} \sqrt[k]{\left|a_k (x - x_0)^k\right|} = |x - x_0| \limsup_{k\to\infty} \sqrt[k]{|a_k|} > 1$$

Pertanto il termine generale della serie (8.9) non è infinitesimo e quindi la serie non converge. □

Osservazione 8.22. In base alla Proposizione 8.6, la somma di una serie di potenze avente raggio positivo r è una funzione continua in $(x_0 - r, x_0 + r)$.

Osservazione 4.2.

Dal teorema precedente segue che una serie di potenze: se $r = 0$, converge solo in x_0; se $r = +\infty$, converge per ogni $x \in \mathbb{R}$; se $0 < r < +\infty$ converge per $|x - x_0| < r$, non converge per $|x - x_0| > r$, mentre non ci dà informazioni sul carattere della serie negli estremi dell'intervallo di convergenza $x = x_0 \pm r$. L'analisi delle serie di potenze

$$\sum_{k=0}^{\infty} (x - x_0)^k, \quad \sum_{k=1}^{\infty} \frac{1}{k} (x - x_0)^k, \quad \sum_{k=1}^{\infty} \frac{1}{k^2} (x - x_0)^k$$

che hanno tutte raggio di convergenza $r = 1$, prova come negli estremi dell'intervallo di convergenza si può avere convergenza assoluta o convergenza semplice o divergenza o indeterminazione.

Proposizione 8.23. *Supponiamo che la serie* (8.9) *abbia raggio di convergenza* $r > 0$ *e sia*

$$f(x) := \sum_{k=0}^{\infty} a_k (x - x_0)^k, \qquad x \in (x_0 - r, x_0 + r) eqdef I. \tag{8.11}$$

Allora le serie

$$\sum_{k=m}^{\infty} k(k-1)\dots(k-m+1) a_k (x-x_0)^{k-m}, \quad m \in \mathbb{N}, \qquad \sum_{k=0}^{\infty} \frac{a_k}{k+1} (x-x_0)^{k+1}, \tag{8.12}$$

hanno raggio di convergenza r*. Inoltre* $f \in \mathcal{C}^{\infty}(I;\mathbb{R})$ *è dotata delle derivate di ogni ordine in* $(x_0 - r, x_0 + r)$ *e si ha*

$$f^{(m)}(x) = \sum_{k=m}^{\infty} k(k-1)\dots(k-m+1) a_k (x-x_0)^{k-m}, \qquad \int_{x_0}^{x} f(t)\mathrm{d}t = \sum_{k=0}^{\infty} \frac{a_k}{k+1} (x-x_0)^{k+1}, \tag{8.13}$$

per ogni $x \in I$.

Dimostrazione. Proviamo (8.13) per $m = 1$. Si ha

$$\limsup_{k\to\infty} \sqrt[k-1]{|k a_k|} = \lim_{k\to\infty} \sqrt[k-1]{k} \limsup_{k\to\infty} \sqrt[k-1]{|a_k|} = \limsup_{k\to\infty} \sqrt[k]{|a_k|},$$

quindi la serie (8.12) ha lo stesso raggio di convergenza della serie (8.11). Si osservi che la serie (8.12) è la serie ottenuta derivando termine a termine la serie (8.11). Per ogni fissato $\rho \in (0, r)$, la serie (8.12) converge uniformemente nell'intervallo $\left[x_0 - \rho, x_0 + \rho\right]$. Per la Proposizione 8.7, si ha dunque che f è derivabile in $\left[x_0 - \rho, x_0 + \rho\right]$ e che ivi $f'(x) = \sum_{k=1}^{\infty} k a_k (x-x_0)^{k-1}$. Il caso delle derivate di ordine superiore segue per induzione.
Per quanto riguarda la seconda serie in (8.12) si ha

$$\limsup_{k\to\infty} \sqrt[k+1]{\left|\frac{a_k}{k+1}\right|} = \lim_{k\to\infty} \sqrt[k+1]{\frac{1}{k+1}} \limsup_{k\to\infty} \sqrt[k+1]{|a_k|} = \limsup_{k\to\infty} \sqrt[k]{|a_k|},$$

quindi la serie (8.12) ha lo stesso raggio di convergenza di (8.11). Sia g la sua somma. Poichè la serie ottenuta derivando termine a termine la serie (8.12) coincide con la serie (8.11), da quanto appena dimostrato, segue che g è derivabile e

$$g'(x) = \sum_{k=0}^{\infty} a_k (x-x_0)^k$$

Ne segue che g è la primitiva di f in $(x_0 - r, x_0 + r)$ verificante $g(x_0) = 0$, il che prova la (8.13). □

8.5 5 Serie di Fourier

Definizione 5.1 (Funzioni periodiche). Sia $T > 0$. Una funzione $f : \mathbb{R} \to \mathbb{R}$ si dice periodica di periodo T (T-periodica) se $f(t + T) = f(t)$ per ogni $t \in \mathbb{R}$.

Evidentemente, se f ha periodo T, ha anche periodo kT, per ogni intero positivo k. Ad esempio le funzioni $\sin t$ e $\cos t$ hanno periodo 2π, ma anche $4\pi, 6\pi$, ecc; per queste funzioni 2π è il periodo minimo.

Le funzioni costanti sono T-periodiche per ogni $T > 0$, quindi non hanno un periodo minimo. In generale, l'esistenza di un periodo minimo per una funzione periodica non costante non è scontata; ad esempio la funzione di Dirichlet (che assume valore 1 sui razionali e 0 sugli irrazionali) ha periodo q, per ogni $q \in \mathbb{Q}^+$, quindi non ammette un periodo minimo.

Proposizione 5.2. Sia $f : \mathbb{R} \to \mathbb{R}$ continua, periodica e non costante. Allora f ammette periodo minimo.

Dimostrazione. Sia $H = \{T > 0 : T \text{ è periodo di } f\}$ e sia $\tau = \inf H, \tau \geq 0$. Per la seconda proprietà dell'estremo inferiore, esiste una scuccessione T_n, $T_n \in H$, tale che $\lim_{n\to\infty} T_n = \tau$. Allora per ogni $t \in \mathbb{R}$ e per ogni $n \in \mathbb{N}$, si ha $f(t + T_n) = f(t)$. Per la continuità di f, si ha che $\lim_{n\to\infty} f(t + T_n) = f(t+\tau)$, quindi $f(t+\tau) = f(t)$ per ogni $t \in \mathbb{R}$. Se $\tau > 0$, allora $\tau \in H$ ed è perciò il periodo minimo di f. Rimane quindi da provare che $\tau > 0$. Essendo f non costante, esistono $s, t \in \mathbb{R}, s < t$, tali che $f(s) \neq f(t)$. Poniamo $\epsilon = |f(s) - f(t)| > 0$. Dalla continuità di f in t si ha che esiste $\delta > 0$ tale che per ogni $w \in R, |w - t| < \delta$, si ha $|f(w) - f(t)| < \epsilon$. Supponiamo per assurdo che sia $\tau = 0$. Esiste allora un periodo T verificante $0 < T < \delta$. Esiste $k_0 \in \mathbb{N}$ tale che $s+k_0T \le t < s+(k_0+1)\,T$. Dunque $|t-(s+k_0T)| < T < \delta$, da cui $|f(s)-f(t)| = \left|f(s+k_0T)-f(t)\right| < \epsilon$, contraddicendo la definizione di ϵ.

Osserviamo che una funzione T-periodica è determinata dai suoi valori in un qualsiasi intervallo $[a,a+T)$ e che, se è integrabile, $\int_a^{a+T} f(t)\mathrm{d}t = \int_0^T f(t)\mathrm{d}t$ (segue dall'additività dell'integrale e operando un cambio di variabile).

Data $f:[a,a+T)\to\mathbb{R}$, la si può estendere per periodicità su $\mathbb{R}$ ottenendo una funzione T-periodica.

Le funzioni del tipo

$$s_n(t) = a_0 + \sum_{k=1}^{n} a_k \cos\left(\frac{2\pi k}{T}t\right) + \sum_{k=1}^{n} b_k \sin\left(\frac{2\pi k}{T}t\right)$$

al variare di $a_0, a_k, b_k \in \mathbb{R}$, sono T-periodiche; sono dette polinomi trigonometrici di grado n.

Si vuole studiare se le funzioni periodiche sono approssimabili (in un qualche senso) mediante i polinomi trigonometrici, ovvero se sono esprimibili come somma di una serie trigonometrica, cioè una serie della forma

$$a_0 + \sum_{k=1}^{\infty} a_k \cos\left(\frac{2\pi k}{T}t\right) + \sum_{k=1}^{\infty} b_k \sin\left(\frac{2\pi k}{T}t\right)$$

per opportuni valori dei coefficienti $a_0, a_k, b_k \in \mathbb{R}$.

Cominciamo con l'osservare che, per $j,k = 1,2,\ldots$,

$$\int_0^T \cos\left(\frac{2\pi k}{T}t\right)\cos\left(\frac{2\pi j}{T}t\right) = \frac{T}{2}\delta_{kj},$$

$$\int_0^T \sin\left(\frac{2\pi k}{T}t\right)\sin\left(\frac{2\pi j}{T}t\right) = \frac{T}{2}\delta_{kj},$$

$$\int_0^T \sin\left(\frac{2\pi k}{T}t\right)\cos\left(\frac{2\pi j}{T}t\right) = 0.$$

La dimostrazione delle precedenti uguaglianze si ottiene facilmente usando le formule di Werner.

Considereremo due diverse modalità di convergenza delle serie trigonometriche: la convergenza puntuale e la convergenza in media quadratica.

Supponiamo, per ora, che la serie (38) sia uniformemente convergente su $\mathbb{R}$ e sia f la sua somma:

$$f(t) = a_0 + \sum_{k=1}^{\infty} a_k \cos\left(\frac{2\pi k}{T}t\right) + \sum_{k=1}^{\infty} b_k \sin\left(\frac{2\pi k}{T}t\right)$$

Essendo i termini della serie funzioni continue, applicando il teorema 2.6 possiamo integrare termine a termine e, usando le (39)-(41), si ha

$$\int_0^T f(t)\mathrm{d}t = a_0 T,$$

$$\int_0^T f(t)\cos\left(\frac{2\pi j}{T}t\right)\mathrm{d}t = \frac{T}{2}a_j, \quad j = 1,2,\ldots,$$

$$\int_0^T f(t)\sin\left(\frac{2\pi j}{T}t\right)\mathrm{d}t = \frac{T}{2}b_j, \quad j = 1,2,\ldots,.$$

Dalle precedenti uguaglianze ricaviamo i coefficienti della serie (42) in funzione della sua somma f:

$$a_0 = \frac{1}{T}\int_0^T f(t)\mathrm{d}t$$

$$a_k = \frac{2}{T}\int_0^T f(t)\cos\left(\frac{2\pi k}{T}t\right)\mathrm{d}t, \quad k = 1,2,\ldots,$$

$$b_k = \frac{2}{T}\int_0^T f(t)\sin\left(\frac{2\pi k}{T}t\right)\mathrm{d}t, \quad k = 1,2,\ldots$$

Alla luce di ciò, data una funzione f, T-periodica, limitata e integrabile su $[0,T]$, si definisce serie di Fourier di f la serie (38) avente i coefficienti dati da (46)-(48), detti coefficienti di Fourier di f. Indicheremo inoltre con S_n la somma parziale n-esima della serie di Fourier, ovvero

$$S_n(t) = a_0 + \sum_{k=1}^{n} a_k \cos\left(\frac{2\pi k}{T} t\right) + \sum_{k=1}^{n} b_k \sin\left(\frac{2\pi k}{T} t\right)$$

con i coefficienti dati da (46)-(48).

Osservazione 5.3. Considerando come intervallo di integrazione $\left[-\frac{T}{2}, \frac{T}{2}\right]$, e osservando che l'integrale di una funzione dispari su tale intervallo è nullo, si ricava che se f è dispari $a_k = 0$, per ogni $k \in \mathbb{N} \cup \{0\}$, mentre se f è pari $b_k = 0$, per ogni $k \in \mathbb{N}$. Pertanto la serie di Fourier di una funzione pari contiene solo coseni (oltre alla costante) e quella di una funzione dispari solo seni.

8.5.1 5.1 Convergenza in media quadratica delle serie di Fourier

Sia $\mathcal{P}_n$ l'insieme dei polinomi trigonometrici s_n di grado n e periodo T.

Teorema 5.4. Sia f una funzione T-periodica, limitata e integrabile su $[0, T]$. Allora

$$\frac{1}{T}\int_0^T \left|f(t) - s_n(t)\right|^2 \mathrm{d}t$$

assume minimo sull'insieme $\mathcal{P}_n$ in corrispondenza al polinomio trigonometrico S_n i cui coefficienti sono i coefficienti di Fourier di f, cioè quelli dati dalle relazioni (46)-(48). Inoltre

$$\int_0^T \left|f(t) - S_n(t)\right|^2 \mathrm{d}t = \int_0^T |f(t)|^2 \mathrm{d}t - \frac{T}{2}\left(2a_0^2 + \sum_{k=1}^{n}\left(a_k^2 + b_k^2\right)\right),$$

$$2a_0^2 + \sum_{k=1}^{\infty}\left(a_k^2 + b_k^2\right) \le \frac{2}{T}\int_0^T |f(t)|^2 \mathrm{d}t, \qquad \text{(disuguaglianza di Bessel).}$$

Dimostrazione. Sia

$$S_n = a_0 + \sum_{k=1}^{n} a_k \cos\left(\frac{2\pi k}{T} t\right) + \sum_{k=1}^{n} b_k \sin\left(\frac{2\pi k}{T} t\right)$$

con a_0, a_k, b_k i coefficienti di Fourier di f e sia

$$s_n = \alpha_0 + \sum_{k=1}^{n} \alpha_k \cos\left(\frac{2\pi k}{T} t\right) + \sum_{k=1}^{n} \beta_k \sin\left(\frac{2\pi k}{T} t\right)$$

il generico polinomio trigonometrico di grado n.

Tenendo conto delle (39)-(41) e delle (46)-(48), si ha

$$\begin{gathered}
\int_0^T \left|f(t) - s_n(t)\right|^2 \mathrm{d}t = \int_0^T |f(t)|^2 \mathrm{d}t + T\alpha_0^2 + \frac{T}{2}\sum_{k=1}^{n}\alpha_k^2 + \frac{T}{2}\sum_{k=1}^{n}\beta_k^2 + \\
-2T\alpha_0 a_0 - T\sum_{k=1}^{n}\alpha_k a_k - T\sum_{k=1}^{n}\beta_k b_k = \\
\int_0^T |f(t)|^2 \mathrm{d}t + T(\alpha_0 - a_0)^2 + \frac{T}{2}\sum_{k=1}^{n}(\alpha_k - a_k)^2 + \frac{T}{2}\sum_{k=1}^{n}\left(\beta_k - b_k\right)^2 \\
-Ta_0^2 - \frac{T}{2}\sum_{k=1}^{n} a_k^2 - \frac{T}{2}\sum_{k=1}^{n} b_k^2,
\end{gathered}$$

che assume il valore minimo per $\alpha_k = a_k, k = 0, 1, 2\ldots, \beta_k = b_k, k = 1, 2\ldots$ D'altra parte, per $s_n = S_n$, si ha

$$0 \le \int_0^T \left|f(t) - S_n(t)\right|^2 \mathrm{d}t = \int_0^T |f(t)|^2 \mathrm{d}t - \frac{T}{2}\left(2a_0^2 + \sum_{k=1}^{n}\left(a_k^2 + b_k^2\right)\right)$$

da cui, passando al limite per $n \to \infty$, segue la (52).

Dalla disuguaglianza di Bessel deriva il seguente Corollario.

Corollario 5.5 (Riemann-Lebesgue). Nelle ipotesi del teorema 5.4,

$$\lim_{k\to\infty}\int_0^T f(t)\cos\left(\frac{2\pi k}{T}t\right)\mathrm{d}t=0$$
$$\lim_{k\to\infty}\int_0^T f(t)\sin\left(\frac{2\pi k}{T}t\right)\mathrm{d}t=0.$$

Si può in realtà dimostrare (ma la dimostrazione richiede nozioni più avanzate) che in (52) sussiste l'uguaglianza (identità di Parseval). Passando quindi al limite per $n\to\infty$ in (51), si ricava che S_n converge a f in media quadratica, cioè $\lim_{n\to\infty}\int_0^T\left|f(t)-S_n(t)\right|^2\mathrm{d}t=0$.

8.5.2 5.2 Convergenza puntuale delle serie di Fourier

Definizione 5.6. Una funzione $f:[a,b]\to\mathbb{R}$ si dice continua a tratti se è continua tranne al più in un numero finito di punti nei quali esistono finiti il limite destro e sinistro.

Osserviamo che se f è continua a tratti in $[a,b]$, allora è limitata e integrabile in $[a,b]$.

Definizione 5.7. Sia $f:\mathbb{R}\to\mathbb{R}$. Diremo che f soddisfa la condizione di Dirichlet in un punto t_0 se

i) è derivabile in t_0;

oppure

ii) è continua in t_0 e dotata delle derivate destra $f'_+(t_0)$ e sinistra $f'_-(t_0)$; oppure

iii) esistono finiti in t_0 i limiti destro $f\left(t_0^+\right)$e sinistro $f\left(t_0^-\right)$di f e i seguenti limiti

$$\widetilde{f}'_+(t_0):=\lim_{t\to t_0^+}\frac{f(t)-f\left(t_0^+\right)}{t-t_0},\quad \widetilde{f}'_-(t_0):=\lim_{t\to t_0^-}\frac{f(t)-f\left(t_0^-\right)}{t-t_0}.$$

Teorema 5.8. Sia f una funzione T - periodica e continua a tratti in $[0,T]$. La serie di Fourier di f converge in ogni punto t_0 in cui è soddisfatta la condizione di Dirichlet e la sua somma $s(t_0)$ è data da

$$s(t_0)=\frac{f\left(t_0^-\right)+f\left(t_0^+\right)}{2}.$$

Per alleggerire le notazioni della dimostrazione osserviamo che, operando il cambio di variabile $x=\frac{2\pi}{T}t$, ci si può ricondurre al caso delle funzioni 2π periodiche e utilizzare come intervallo $[-\pi,\pi]$.

Lemma 5.9 (Formula di Dirichlet). Per ogni $n\in\mathbb{N}$, e per ogni $t\in\mathbb{R}$, $t\neq 2m\pi(m\in\mathbb{Z})$,

$$\frac{1}{2}+\cos t+\cos 2t+\ldots+\cos nt=\frac{\sin\left(n+\frac{1}{2}\right)t}{2\sin\frac{t}{2}}$$

Dimostrazione. La dimostrazione si svolge per induzione. Per $n=0$, la (56) è ovvia. Suppostala vera per $n-1$, si ha

$$\begin{aligned}\frac{1}{2}+\cos t+\cos 2t+\ldots+\cos(n-1)t+\cos nt&=\frac{\sin\left(n-\frac{1}{2}\right)t}{2\sin\frac{t}{2}}+\cos nt=\\&=\frac{\sin\left(n-\frac{1}{2}\right)t+2\sin\frac{t}{2}\cos nt}{2\sin\frac{t}{2}}\end{aligned}$$

Dato che

$$\sin\left(n-\frac{1}{2}\right)t=\sin nt\cos\frac{t}{2}-\cos nt\sin\frac{t}{2}$$

sostituendo la precedente espressione in (57) si ha la tesi.

Dimostrazione del teorema 5.8. Ricordando le formule (46)-(48), si ha che la somma parziale ennesima $S_n(t_0)$ della serie di Fourier di f in t_0 è data da

$$S_n(t_0)=\frac{1}{\pi}\int_{-\pi}^{\pi}f(t)\left(\frac{1}{2}+\sum_{k=1}^{n}(\cos kt\cos kt_0+\sin kt\sin kt_0)\right)\mathrm{d}t$$

Dalle formule di addizione degli angoli e operando un cambio di variabile, si ha

$$
\begin{aligned}
S_n(t_0) &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left(\frac{1}{2}+\sum_{k=1}^{n}\cos k(t-t_0)\right)\mathrm{d}t = \\
&= \frac{1}{\pi}\int_{-\pi-t_0}^{\pi-t_0} f(u+t_0)\left(\frac{1}{2}+\sum_{k=1}^{n}\cos ku\right)du = \frac{1}{\pi}\int_{-\pi}^{\pi} f(u+t_0)\left(\frac{1}{2}+\sum_{k=1}^{n}\cos ku\right)du.
\end{aligned}
$$

Rinominando t la variabile di integrazione, si ha

$$
S_n(t_0) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t+t_0)\left(\frac{1}{2}+\sum_{k=1}^{n}\cos kt\right)\mathrm{d}t
$$

Si verifica banalmente che

$$
\frac{1}{\pi}\int_{0}^{\pi}\left(\frac{1}{2}+\sum_{k=1}^{n}\cos kt\right)\mathrm{d}t = \frac{1}{\pi}\int_{-\pi}^{0}\left(\frac{1}{2}+\sum_{k=1}^{n}\cos kt\right)\mathrm{d}t = \frac{1}{2}
$$

Pertanto possiamo calcolare

$$
\begin{aligned}
S_n(t_0) - \frac{f\left(t_0^-\right)+f\left(t_0^+\right)}{2} &= \frac{1}{\pi}\int_{-\pi}^{0}\left(f(t_0+t)-f\left(t_0^-\right)\right)\left(\frac{1}{2}+\sum_{k=1}^{n}\cos kt\right)\mathrm{d}t + \\
&+ \frac{1}{\pi}\int_{0}^{\pi}\left(f(t_0+t)-f\left(t_0^+\right)\right)\left(\frac{1}{2}+\sum_{k=1}^{n}\cos kt\right)\mathrm{d}t.
\end{aligned}
$$

Usando la formula di Dirichlet (56), si ha

$$
S_n(t_0) - \frac{f\left(t_0^-\right)+f\left(t_0^+\right)}{2} = \frac{1}{\pi}\int_{-\pi}^{\pi} F(t)\sin\left(n+\frac{1}{2}\right)t\,\mathrm{d}t,
$$

dove

$$
F(t) = \begin{cases}
\frac{f(t_0+t)-f\left(t_0^-\right)}{2\sin\left(\frac{t}{2}\right)}, & \text{per } -\pi \le t < 0,\\
0, & \text{per } t = 0,\\
\frac{f(t_0+t)-f\left(t_0^+\right)}{2\sin\left(\frac{t}{2}\right)}, & \text{per } 0 < t < \pi.
\end{cases}
$$

Esistono finiti i limiti

$$
\begin{aligned}
\lim_{t\to 0^+} F(t) &= \lim_{t\to 0^+} \frac{f(t_0+t)+f\left(t_0^+\right)}{t}\,\frac{\frac{t}{2}}{\sin\left(\frac{t}{2}\right)} = \widetilde{f}'_+(t_0),\\
\lim_{t\to 0^-} F(t) &= \lim_{t\to 0^-} \frac{f(t_0+t)+f\left(t_0^-\right)}{t}\,\frac{\frac{t}{2}}{\sin\left(\frac{t}{2}\right)} = \widetilde{f}'_-(t_0).
\end{aligned}
$$

Dunque la funzione F è continua a tratti, limitata e integrabile in $[-\pi,\pi)$ e si può quindi applicare il Corollario di Riemann Lebesgue 5.5 alle funzioni $F(t)\sin\left(\frac{t}{2}\right)$ e $F(t)\cos\left(\frac{t}{2}\right)$. Da (62) si ha dunque

$$
\begin{aligned}
S_n(t_0) - \frac{f\left(t_0^-\right)+f\left(t_0^+\right)}{2} &= \frac{1}{\pi}\int_{-\pi}^{\pi} F(t)\sin\left(n+\frac{1}{2}\right)t\,\mathrm{d}t = \\
&= \frac{1}{\pi}\int_{-\pi}^{\pi} F(t)\sin\left(\frac{t}{2}\right)\cos nt\,\mathrm{d}t + \frac{1}{\pi}\int_{-\pi}^{\pi} F(t)\cos\left(\frac{t}{2}\right)\sin nt\,\mathrm{d}t \to 0,
\end{aligned}
$$

al tendere di n a $+\infty$.

Riferimenti bibliografici

[1] Antonio Ambrosetti and Giovanni Prodi, *A primer of nonlinear analysis*, Cambridge Studies in Advanced Mathematics, vol. 34, Cambridge University Press, Cambridge, 1993.

[2] Haim Brezis, *Functional analysis, Sobolev spaces and partial differential equations*, Universitext, Springer, New York, 2011.