

Indice

1	Nozioni Preliminari	3
1.1	Spazi metrici e topologia	3
1.2	Compattezza e completezza in spazi metrici	8
1.3	Lo spazio $\mathbb{R}^n$ e le rette parametriche	12
2	Funzioni di più variabili	15
2.1	Calcolo differenziale in più variabili	15
2.1.1	Derivabilità	17
2.1.2	Differenziabilità	21
2.1.3	Derivate successive	25
2.1.4	Composizione di funzioni di più variabili	27
2.1.5	Massimi, minimi e punti stazionari di funzioni di più variabili	30
3	Integrazione	35
3.1	Integrale di Riemann	35
3.2	La misura di Peano-Jordan e alcune funzioni integrabili	37
3.2.1	Insiemi normali	41
3.3	Teorema di Fubini e cambio di variabili	43
3.3.1	Coordinate polari, cilindriche e sferiche	46
3.4	Integrali impropri	47
4	Successioni e serie di funzioni	51
4.1	Successioni di funzioni	51
4.2	Serie di funzioni	56
4.2.1	Serie di potenze	57
4.2.2	Serie di Fourier	60
5	Curve e superfici	67
5.1	Curve	67
5.2	Superfici	72
5.3	Massimi e minimi vincolati	74
6	Equazioni Differenziali e Problema di Cauchy	79
6.1	Problema di Cauchy	80
6.2	Soluzioni massimali e globali	84
6.3	Teoremi di confronto	86

Capitolo 1

Nozioni Preliminari

1.1 Spazi metrici e topologia

Denotiamo con X un generico insieme non vuoto.

Definizione 1.1.1. Una funzione $d : X \times X \rightarrow [0, +\infty]$ si dice *metrica* su X (o anche *distanza* su X) se valgono le seguenti proprietà:

- (i) $d(x, y) \geq 0$ per ogni $x, y \in X$ e $d(x, y) = 0$ se e solo se $x = y$;
- (ii) $d(x, y) = d(y, x)$ per ogni $x, y \in X$;
- (iii) $d(x, y) \leq d(x, z) + d(z, y)$ per ogni $x, y, z \in X$.

Un insieme X dotato di metrica d si chiama *spazio metrico*, e si indica con (X, d) .

Osservazione 1.1.2. Se $Y \subseteq X$ è un sottoinsieme non vuoto di X e (X, d) è uno spazio metrico, allora risulta che anche (Y, d) è uno spazio metrico.

Esempio 1.1.3. L'insieme dei numeri reali $\mathbb{R}$ dotato della distanza $d(x, y) := |x - y|$, per ogni $x, y \in \mathbb{R}$, è uno spazio metrico. Come conseguenza dell'osservazione precedente si ha anche che $(\mathbb{N}, d)$, $(\mathbb{Z}, d)$, $(\mathbb{Q}, d)$ sono spazi metrici, e più in generale (E, d) è uno spazio metrico per ogni sottoinsieme non vuoto $E \subset \mathbb{R}$.

Esempio 1.1.4. Lo spazio vettoriale $\mathbb{R}^n$ è usualmente dotato della metrica euclidea d_e data da

$$d_e(x, y) = \|x - y\| = \sqrt{\sum_{i=1}^n (x_i - y_i)^2} \quad (1.1.1)$$

per ogni $x, y \in \mathbb{R}^n$, dove x_i indica la i -esima coordinata di $x = (x_1, x_2, \dots, x_n)$. Si noti però che possiamo definire altre metriche su $\mathbb{R}^n$, come per esempio

$$d_1(x, y) := \sum_{i=1}^n |x_i - y_i|, \quad (1.1.2)$$

$$d_\infty(x, y) := \max\{|x_1 - y_1|, |x_2 - y_2|, \dots, |x_n - y_n|\}, \quad (1.1.3)$$

definite per ogni $x, y \in \mathbb{R}^n$.

Osserviamo preliminarmente che la disuguaglianza triangolare implica la seguente: dati $x, y, z \in X$, si ha

$$|d(x, z) - d(z, y)| \leq d(x, y). \quad (1.1.4)$$

Infatti per la disuguaglianza triangolare abbiamo

$$d(x, z) - d(z, y) \leq d(x, y)$$

e anche

$$d(z, y) - d(x, z) \leq d(x, y)$$

le quali, messe insieme, implicano (1.1.4).

L'introduzione di una metrica su un insieme X fornisce automaticamente una topologia su X , e così X diventa uno spazio topologico. Nello specifico, avremo a disposizione una base per gli insiemi aperti in (X, d) che è data dalle palle aperte, definite come segue:

Definizione 1.1.5. Sia (X, d) uno spazio metrico, allora la palla aperta centrata in $x \in X$ e con raggio $r > 0$ è l'insieme

$$B_r(x) := \{y \in X : |x - y| < r\}.$$

Un insieme $A \subseteq X$ si dice aperto se per ogni $x \in A$ esiste una palla $B_r(x)$ contenuta in A , ovvero $B_r(x) \subseteq A$. Un insieme $C \subseteq X$ si dirà chiuso se il suo complementare $C^c := X \setminus C$ è aperto.

Per definizione si considera l'insieme vuoto $\emptyset$ aperto. Insiemi aperti e chiusi godono delle seguenti proprietà:

- Se A è aperto, allora $A^c := X \setminus A$ è chiuso, e viceversa se C è chiuso allora C^c è aperto;
- Se $(A_i)_{i \in I}$ è famiglia qualsiasi di insiemi aperti A_i , dove i è un indice che varia in un insieme di indici I , allora l'insieme

$$A := \bigcup_{i \in I} A_i,$$

ottenuto facendo l'unione di tutti gli A_i , è ancora un insieme aperto;

- Se $(C_i)_{i \in I}$ è famiglia qualsiasi di insiemi chiusi C_i allora l'insieme

$$C := \bigcap_{i \in I} C_i, \quad (1.1.5)$$

ottenuto facendo l'intersezione di tutti gli C_i , è ancora un insieme chiuso.

Introduciamo ora il concetto di convergenza:

Definizione 1.1.6. Sia $(x_n) \subset X$ una successione e sia $x \in X$. Diremo che la successione (x_n) tende a x (quando $n \rightarrow \infty$) se e solo se

$$\lim_{n \rightarrow \infty} d(x_n, x) = 0. \quad (1.1.6)$$

In tal caso scriveremo

$$x_n \rightarrow x$$

per indicare che (x_n) tende a x quando $n \rightarrow \infty$.

Diremo che una successione $(x_n) \subseteq X$ è convergente se esiste $x \in X$ tale che $x_n \rightarrow x$ quando $n \rightarrow +\infty$.

Equivalentemente, avremmo potuto dire che la successione (x_n) tende a x se e solo se vale la condizione seguente

$$\forall \varepsilon > 0 \quad \exists N_\varepsilon \in \mathbb{N} \quad \text{tale che} \quad d(x_n, x) < \varepsilon \quad \forall n \geq N_\varepsilon. \quad (1.1.7)$$

Ovvero, usando la nozione di palla $B_r(x)$, se e solo se vale la seguente condizione

$$\forall r > 0 \quad \exists N_r \in \mathbb{N} \quad \text{tale che} \quad x_n \in B_r(x) \quad \forall n \geq N_r. \quad (1.1.8)$$

Definizione 1.1.7. Sia $E \subseteq X$; un punto $x \in X$ si dice *punto di accumulazione* per E se per ogni $r > 0$ esiste un punto $y \neq x$ con $y \in E \cap B_r(x)$.

Osserviamo che se x è un punto di accumulazione per E , non necessariamente $x \in E$. Tuttavia esiste sempre una successione $(x_n) \subset E \setminus \{x\}$ tale che $x_n \rightarrow x$ quando $n \rightarrow +\infty$. Infatti, per definizione di punto di accumulazione, per ogni $n > 0$ esisterà un punto $y \neq x$, $y \in E \cap B_{\frac{1}{n}}(x)$. Denotando $y = x_n$, troviamo così una successione che tende a x . Fissato infatti $r > 0$, esiste un $N_r \in \mathbb{N}$ tale $\frac{1}{N_r} < r$, e inoltre $x_n \in B_{\frac{1}{n}}(x) \subset B_r(x)$ per ogni $n \geq N_r$. Questo implica che $x_n \rightarrow x$ grazie alla definizione (1.1.8).

Definizione 1.1.8. Sia $E \subseteq X$; sia $\mathcal{C}(E)$ la famiglia costituita da tutti gli insiemi chiusi che contengono E :

$$\mathcal{C}(E) := \{C \subseteq X \text{ tale che } C \text{ è chiuso e } C \supseteq E\}.$$

Allora chiameremo chiusura di E l'insieme $\overline{E}$ definito come

$$\overline{E} := \bigcap_{C \in \mathcal{C}(E)} C,$$

l'intersezione di tutti gli insiemi chiusi contenenti E .

Si osserva che per la proprietà (1.1.5), $\overline{E}$ è un insieme chiuso. Inoltre, ovviamente $\overline{E}$ contiene E . Possiamo anche dire che $\overline{E}$ è “il più piccolo” insieme chiuso contenente E , nel senso che se C è un insieme chiuso contenente E , allora $\overline{E} \subseteq C$ (questo segue dal fatto che $C \in \mathcal{C}(E)$).

Definizione 1.1.9. Sia $E \subseteq X$ e sia $\mathcal{A}(E)$ la famiglia di tutti gli insiemi aperti contenuti in E , ovvero

$$\mathcal{A}(E) := \{A \subseteq X \text{ tale che } A \text{ è aperto e } A \subseteq E\}.$$

Allora la *parte interna* di E , denotata con $\overset{\circ}{E}$, è definita da

$$\overset{\circ}{E} := \bigcup_{A \in \mathcal{A}(E)} A,$$

l'unione di tutti gli insiemi aperti contenuti in E .

Anche in questo caso osserviamo che $\overset{\circ}{E}$ è necessariamente un insieme aperto e che è contenuto in E . Possiamo anche dire che $\overset{\circ}{E}$ è “il più grande” insieme aperto contenuto in E , nel senso che se A è un insieme aperto contenuto in E , allora $A \subseteq \overset{\circ}{E}$ (infatti $A \in \mathcal{A}(E)$).

Lemma 1.1.10. *Sia $C \subseteq X$ un insieme chiuso e supponiamo che (x_n) sia una successione con $x_n \in C$ per ogni n e tale che $x_n \rightarrow x$ quando $n \rightarrow \infty$. Allora $x \in C$.*

Dimostrazione. Supponiamo per assurdo $x \notin C$. Allora $x \in C^c$, e C^c è un insieme aperto, quindi esiste $r > 0$ tale che $B_r(x) \subseteq C^c$. Ma questo è assurdo, perchè visto che $x_n \rightarrow x$, per la (1.1.8) troveremo sempre punti x_n in $B_r(x) \subseteq C^c$, contraddicendo il fatto che $x_n \in C$. $\square$

Lemma 1.1.11. *Sia $E \subseteq X$; allora $x \in \overline{E}$ se e solo se esiste una successione $(x_n) \in E$ tale che $x_n \rightarrow x$ quando $n \rightarrow +\infty$.*

Dimostrazione. Essendo $\overline{E}$ un insieme chiuso, se $x_n \in E \subseteq \overline{E}$, allora per il lemma precedente $x \in \overline{E}$. Viceversa supponiamo $x \in \overline{E}$ e prendiamo $n \in \mathbb{N}$; allora esisterà un elemento $x_n \in B_{\frac{1}{n}}(x) \cap E$. Infatti, qualora avessimo $B_{\frac{1}{n}}(x) \cap E = \emptyset$, si avrà che $B_{\frac{1}{n}}(x)^c$ è un chiuso contenente E , e quindi $\overline{E} \subseteq B_{\frac{1}{n}}(x)^c$, che implica $x \notin \overline{E}$, assurdo. Quindi per ogni $n \in \mathbb{N}$ esiste $x_n \in B_{\frac{1}{n}}(x)$ con $x_n \in E$, e pertanto la successione (x_n) converge a x (verificarlo) e soddisfa la tesi del lemma. $\square$

Corollario 1.1.12. *Ogni punto di accumulazione per $E \subseteq X$ appartiene a $\overline{E}$.*

Corollario 1.1.13. *Un insieme $C \subseteq X$ è chiuso se e solo se per ogni successione $(x_n) \subseteq C$ tale che $x_n \rightarrow x$ quando $n \rightarrow \infty$, si ha $x \in C$.*

Dimostrazione. Il “solo se” è implicato dal Lemma 1.1.10; viceversa, supponiamo per ogni successione $(x_n) \subseteq C$ con $x_n \rightarrow x$ quando $n \rightarrow \infty$, si abbia $x \in C$. Per il Lemma 1.1.11 abbiamo che ogni $x \in \overline{C}$ ammette una successione $(x_n) \subseteq C$ con $x_n \rightarrow x$, e quindi per l’ipotesi fatta avremo anche $x \in C$. In altre parole, ogni punto di $\overline{C}$ appartiene anche a C , e quindi necessariamente $C = \overline{C}$. Essendo $\overline{C}$ chiuso, anche C lo è, e abbiamo concluso. $\square$

Definizione 1.1.14. Sia $E \subseteq X$ un insieme non vuoto; definiamo il *diametro* di E il numero (eventualmente pari a $+\infty$)

$$\text{diam}(E) := \sup\{d(x, y) : x, y \in E\}.$$

Un insieme $E \subseteq X$ si dice *limitato* se esiste $r > 0$ e un punto $x \in X$ tale che $E \subseteq B_r(x)$.

Grazie alla disuguaglianza triangolare possiamo dimostrare che il diametro di una palla di raggio r è sempre minore o uguale a $2r$. Infatti, se z è il centro della palla e $x, y \in B_r(z)$, allora

$$d(x, y) \leq d(x, z) + d(z, y) \leq r + r = 2r.$$

Visto che questa disuguaglianza vale per ogni $x, y \in B_r(z)$, anche l’estremo superiore di $d(x, y)$ su tutti gli $x, y \in B_r(z)$ sarà $\leq 2r$.

In particolare deduciamo che se $E \subseteq X$ è un insieme limitato, allora il suo diametro è finito. Infatti $E \subset B_r(x)$ e quindi $\text{diam}(E) \leq \text{diam}(B_r(x)) \leq 2r$.

Si può anche dimostrare che se $\text{diam}(E) < +\infty$ allora E è un insieme limitato. Lasciamo quest’ultima prova per esercizio.

Definizione 1.1.15. Sia (X, d) uno spazio metrico, e sia $f : X \rightarrow \mathbb{R}$ una funzione. Allora f si dice *continua in $\bar{x} \in X$* se

$$\forall \varepsilon > 0 \exists \delta > 0 \text{ per cui } |f(y) - f(\bar{x})| \leq \varepsilon \quad \forall y \in X \text{ tale che } d(\bar{x}, y) \leq \delta. \quad (1.1.9)$$

Una funzione $f : X \rightarrow \mathbb{R}$ si dice *continua su* $E \subseteq X$ se è continua in ogni $\bar{x} \in E$.

Una funzione $f : X \rightarrow \mathbb{R}$ si dice *uniformemente continua su* $E \subseteq X$ se

$$\forall \varepsilon > 0 \exists \delta > 0 \text{ per cui } |f(y) - f(x)| \leq \varepsilon \quad \forall x, y \in E \text{ tali che } d(x, y) \leq \delta. \quad (1.1.10)$$

Osservazione 1.1.16. Dire che f è continua in $\bar{x}$ equivale a richiedere che per ogni successione (x_n) che converge a $\bar{x}$, si ha

$$f(x_n) \rightarrow f(\bar{x}).$$

Dimostrare questa equivalenza per esercizio.

Osservazione 1.1.17. Osserviamo che l'uniforme continuità di f su E implica la continuità di f su E , ma le due nozioni non sono equivalenti. Potremmo infatti dire che se f è continua su E , allora vale la seguente condizione:

$$\forall x \in E, \forall \varepsilon > 0 \exists \delta = \delta(x) > 0 \text{ per cui } |f(y) - f(x)| \leq \varepsilon \quad \forall y \in X : d(x, y) \leq \delta. \quad (1.1.11)$$

Questa condizione contiene la formula $\forall x$ all'inizio, mentre nella (1.1.10) tale formula è presente alla fine. Questo significa, nella (1.1.11) che l'esistenza di $\delta > 0$ dipende da quale x abbiamo fissato. Ovvero il valore di δ dipende da x , e per questo abbiamo scritto $\delta = \delta(x)$. Nella (1.1.10) il valore di $\delta > 0$ è invece lo stesso per ogni $x \in E$.

Osservazione 1.1.18. Notiamo infine che se f è una funzione continua su X , allora controimmagine $f^{-1}(A)$ di un insieme aperto $A \subseteq \mathbb{R}$ è sempre un aperto in X . Viceversa, una funzione f è continua su X se per ogni aperto $A \subseteq \mathbb{R}$ la sua controimmagine $f^{-1}(A)$ è aperta in X . Lasciamo questa verifica per esercizio.

Come conseguenza si nota che se $f : X \rightarrow \mathbb{R}$ è continua allora la controimmagine di un insieme chiuso $C \subseteq \mathbb{R}$ è sempre chiusa in X .

Lemma 1.1.19. Sia $\bar{x} \in X$. Allora la funzione $d(\bar{x}, \cdot) : X \rightarrow [0, +\infty]$ (quella che associa ad ogni $x \in X$ il valore $d(\bar{x}, x)$) è continua su X .

Dimostrazione. Facciamo vedere, utilizzando l'Osservazione 1.1.16, che se $x_n \rightarrow x$ quando $n \rightarrow \infty$, allora

$$d(\bar{x}, x_n) \rightarrow d(\bar{x}, x).$$

Per la formula (1.1.4), si ha

$$|d(\bar{x}, x_n) - d(\bar{x}, x)| \leq d(x_n, x),$$

ed ora, se tende a zero il membro di destra, lo farà anche il membro di sinistra essendo non negativo. $\square$

Lemma 1.1.20. Siano (x_n) e (y_n) due successioni in X tali che $x_n \rightarrow x$ e $y_n \rightarrow y$ quando $n \rightarrow \infty$. Allora

$$d(x_n, y_n) \rightarrow d(x, y) \quad \text{quando } n \rightarrow \infty.$$

Dimostrazione. Per la disuguaglianza triangolare e per la (1.1.4) possiamo scrivere

$$\begin{aligned} |d(x_n, y_n) - d(x, y)| &= |d(x_n, y_n) - d(x_n, y) + d(x_n, y) - d(x, y)| \\ &\leq |d(x_n, y_n) - d(x_n, y)| + |d(x_n, y) - d(x, y)| \\ &\leq d(y_n, y) + d(x_n, x). \end{aligned}$$

Ora il membro di destra tende a zero, e quindi lo farà anche il membro di sinistra essendo non negativo. $\square$

Definizione 1.1.21. Se $x \in X$ e $r > 0$ definiamo il disco di centro x e raggio r l'insieme

$$D_r(x) := \{y \in X : d(x, y) \leq r\}.$$

Ovviamente si ha $B_r(x) \subseteq D_r(x)$ per ogni $x \in X$ e $r > 0$. Inoltre, si osserva che $D_r(x)$ è chiuso. Quest'affermazione può essere verificata in due modi, il primo è utilizzando il Corollario 1.1.13 e sfruttando la continuità della funzione $d(\cdot, x)$ (lasciamo la verifica per esercizio), il secondo è osservando che $D_r(x)$ è la controimmagine dell'intervallo chiuso $[0, r]$ rispetto la funzione continua $d(\cdot, x)$, e quindi sarà chiuso (si veda Osservazione 1.1.18).

Come conseguenza avremo $\overline{B_r(X)} \subseteq D_r(x)$ per ogni $r > 0$ e $x \in X$. Tuttavia non è sempre vero che tale inclusione sia in realtà un'uguaglianza (dipende dallo spazio metrico in considerazione).

Esercizio 1.1.22. Si trovi un esempio di spazio metrico (X, d) in cui può accadere che $\overline{B_r(X)} \neq D_r(x)$.

Anche per i dischi $D_r(x)$ si verifica che $\text{diam}(D_r(x)) \leq 2r$.

1.2 Compattezza e completezza in spazi metrici

Definizione 1.2.1. Una successione $(x_n) \subseteq X$ si dice di Cauchy se vale la seguente proprietà:

$$\forall \varepsilon > 0 \exists N_\varepsilon \in \mathbb{N} \text{ tale che } d(x_n, x_m) \leq \varepsilon \quad \forall n, m \geq N_\varepsilon. \quad (1.2.1)$$

Si verifica immediatamente che ogni successione convergente è una successione di Cauchy (esercizio). Il viceversa non è sempre vero, a meno che (X, d) non sia uno spazio completo:

Definizione 1.2.2. Uno spazio metrico (X, d) si dice *completo* se ogni successione di Cauchy è convergente. Un sottoinsieme E di X si dice completo se lo spazio metrico (E, d) risulta completo.

Lasciamo per esercizio la verifica del seguente lemma:

Lemma 1.2.3. Sia $(x_n) \subset X$ una successione di Cauchy e supponiamo che ci sia una sottosuccessione (x_{n_k}) convergente a x . Allora necessariamente tutta la successione (x_n) converge a x .

Mostriamo invece il seguente fatto:

Lemma 1.2.4. Sia (X, d) uno spazio metrico completo, e sia $E \subseteq X$ un sottoinsieme. Allora E è completo se e solo se è chiuso.

Dimostrazione. Supponiamo E è completo e prendiamo una successione $(x_n) \subset E$ convergente. Allora essa è di Cauchy, ed essendo E completo il suo limite sarà in E . Pertanto E è chiuso per il Corollario 1.1.13.

Viceversa, se $(x_n) \subset E$ è una successione di Cauchy, allora essa ammette limite $x \in X$ (X è completo), e sempre per lo stesso Corollario si ha $x \in E$. Quindi E è completo. $\square$

La nozione di compattezza in spazi topologici è la seguente:

Definizione 1.2.5. Dato $E \subseteq X$, diremo che una famiglia di insiemi aperti $\{A_i : i \in I\}$ forma un ricoprimento aperto di E se

$$E \subseteq \bigcup_{i \in I} A_i.$$

Un insieme $K \subseteq X$ non vuoto si dice *compatto* se vale la seguente condizione: per ogni ricoprimento aperto $\{A_i : i \in I\}$ di K esiste un (sotto-)ricoprimento $\{A_i : i \in \hat{I}\}$ con $\hat{I} \subseteq I$ e $\hat{I}$ di cardinalità finita, tale che

$$K \subseteq \bigcup_{i \in \hat{I}} A_i.$$

Osservazione 1.2.6. Si vede facilmente che se $K_1, K_2, \dots, K_n$ sono compatti, allora anche la loro unione $K = \bigcup_{i=1}^n K_i$ risulta essere un insieme compatto (esercizio).

In presenza di una metrica si utilizza la seguente nozione:

Definizione 1.2.7. Un insieme $K \subseteq X$ non vuoto si dice *compatto per successioni* se vale la seguente condizione: per ogni successione $(x_n) \subset K$ esiste una sottosuccessione (x_{n_k}) convergente ad un punto $x \in K$.

Per caratterizzare gli insiemi compatti in uno spazio metrico abbiamo bisogno della seguente nozione:

Definizione 1.2.8. Un insieme $E \subseteq X$ si dice *totalmente limitato* se vale la seguente condizione: per ogni $\varepsilon > 0$ esiste $n \in \mathbb{N}$ e un numero finito di punti $\{x_1, \dots, x_n\} \subset X$ tali che

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

Osservazione 1.2.9. Si verifica immediatamente (esercizio) che ogni sottoinsieme $F \subseteq E$ di un insieme E totalmente limitato, è totalmente limitato.

Osservazione 1.2.10. Si verifica anche che se E è totalmente limitato, allora deve essere anche limitato. Il viceversa non è sempre vero.

In uno spazio metrico vale il seguente risultato:

Teorema 1.2.11. Sia (X, d) uno spazio metrico e sia $K \subseteq X$. Allora le seguenti affermazioni sono equivalenti:

- (i) K è compatto per successioni;
- (ii) K è totalmente limitato e completo;
- (iii) K è compatto.

Come conseguenza le due nozioni di compattezza e compattezza per successioni sono equivalenti in spazi metrici.

Dimostrazione. Per dimostrare il teorema facciamo vedere le tre implicazioni (i) $\Rightarrow$ (ii), (ii) $\Rightarrow$ (iii), e (iii) $\Rightarrow$ (i).

(i) $\Rightarrow$ (ii): supponiamo K compatto per successioni, e prendiamo $(x_n) \subset K$ una successione di Cauchy. Essendo K compatto per successioni, esiste una sottosuccessione (x_{n_k}) convergente ad un punto $x \in K$. Per il Lemma 1.2.3 tutta la successione x_n tende a $x \in K$, e quindi K risulta completo. Mostriamo ora che K deve essere totalmente limitato. Supponiamo per assurdo non lo sia; negando la proprietà di totale limitatezza, deduciamo che

$$\exists \bar{\varepsilon} > 0 : \forall n \in \mathbb{N} \text{ non esiste alcuna famiglia di } n \text{ palle di raggio } \bar{\varepsilon} \text{ che ricopre } K. \quad (1.2.2)$$

Prendiamo ora $x_i \in K$; per la proprietà precedente deve esistere necessariamente $x_2 \in K \setminus B_{\bar{\varepsilon}}(x_1)$ (altrimenti $B_{\bar{\varepsilon}}(x_1)$ ricoprirebbe K). Similmente, deve esistere $x_3 \in K \setminus B_{\bar{\varepsilon}}(x_1) \setminus B_{\bar{\varepsilon}}(x_2)$, e così via, per ogni $k > 0$ esiste $x_k \in K \setminus (\cup_{i=1}^{k-1} B_{\bar{\varepsilon}}(x_i))$. In questo modo abbiamo costruito una successione $(x_k) \subset K$ che verifichiamo non ammettere sottosuccessioni convergenti (e questo contraddice il fatto che K è compatto per successioni). Infatti, per come è definita la successione, avremo che

$$\forall n, m \in \mathbb{N}, n \neq m, \quad \text{si ha} \quad d(x_n, x_m) \geq \bar{\varepsilon}.$$

Pertanto, anche se prendiamo una sottosuccessione (x_{k_j}) di (x_k) , essa avrà elementi che distano a due a due almeno $\bar{\varepsilon}$. In particolare si vede che (x_{k_j}) non è una successione di Cauchy, e quindi non può essere convergente. Pertanto abbiamo trovato un assurdo con l'ipotesi che K fosse compatto per successioni.

(ii) $\Rightarrow$ (iii): supponiamo K totalmente limitato e completo. In quanto totalmente limitato, per ogni $\varepsilon > 0$ esistono $N_\varepsilon \in \mathbb{N}$ e dischi $D_\varepsilon(p_i)$, con $i = 1, \dots, N_\varepsilon$, tali che $K \subseteq \bigcup_{i=1}^{N_\varepsilon} D_\varepsilon(p_i)$. In particolare, ponendo $\varepsilon = 1$, ci sono $N \in \mathbb{N}$ dischi $D_1(p_i)$ tali che

$$K \subseteq D_1(p_1) \cup \dots \cup D_1(p_N).$$

Supponiamo ora per assurdo che K non sia compatto. Esisterebbe allora un ricoprimento aperto $\{A_i : i \in I\}$ di K tale che nessun suo sottoricoprimento $\{A_i : i \in \hat{I}\}$, con $\hat{I} \subset I$ di cardinalità finita, soddisfa $K \subseteq \bigcup_{i \in \hat{I}} A_i$. Ne deduciamo che almeno uno tra gli insiemi $K \cap D_1(p_1)$, $K \cap D_1(p_2)$, $\dots$, $K \cap D_1(p_N)$, NON¹ può essere ricoperto con un numero finito di aperti in $\{A_i : i \in I\}$. Senza perdita di generalità diciamo che tale insieme è $K \cap D_1(p_1)$ e rinominiamolo $K_1 := K \cap D_1(p_1)$. Anche quest'ultimo è totalmente limitato (Osservazione 1.2.9), quindi possiamo ricoprirlo con dischi di raggio $\frac{1}{2}$, in modo che

$$K_1 \subseteq D_{\frac{1}{2}}(p_1) \cup \dots \cup D_{\frac{1}{2}}(p_N),$$

dove i punti p_i non sono gli stessi del caso precedente. Anche in questo caso, visto che per K_1 non possiamo trovare un sottoricoprimento finito di $\{A_i : i \in I\}$, necessariamente almeno uno tra gli insiemi $K_1 \cap D_{\frac{1}{2}}(p_1)$, $K_1 \cap D_{\frac{1}{2}}(p_2)$, $\dots$, $K_1 \cap D_{\frac{1}{2}}(p_N)$ non può essere ricoperto con

¹Infatti, se per ogni $h = 1, \dots, N$ esistesse un sottoricoprimento finito degli $\{A_i : i \in I\}$, (chiamiamolo $\{A_i : i \in I_h\}$, con I_h insieme finito) tale che $(K \cap D_1(p_h)) \subseteq \bigcup_{i \in I_h} A_i$, allora si avrebbe

$$K \subseteq \bigcup_{i \in \hat{I}} A_i,$$

con $\hat{I} = I_1 \cup \dots \cup I_N$ un insieme finito; questo contraddirebbe quanto detto in precedenza.

un numero finito di aperti A_i . Supponiamo tale insieme sia $K_2 := K_1 \cap D_{\frac{1}{2}}(p_1)$. Possiamo allora procedere come prima, ovvero trovare dei dischi di raggio $\frac{1}{4}$ tali che

$$K_2 \subseteq D_{\frac{1}{4}}(p_1) \cup \dots \cup D_{\frac{1}{4}}(p_N),$$

e almeno uno tra gli insiemi $K_2 \cap D_{\frac{1}{4}}(p_1)$, $K_2 \cap D_{\frac{1}{4}}(p_2)$, $\dots$, $K_2 \cap D_{\frac{1}{4}}(p_N)$ non può essere ricoperto con un numero finito di aperti A_i , e chiamiamo esso K_3 . Procedendo ricorsivamente, per ogni $n \in \mathbb{N}$ definiamo K_n quell'insieme tra $K_{n-1} \cap D_{\frac{1}{2^{n-1}}}(p_1)$, $K_{n-1} \cap D_{\frac{1}{2^{n-1}}}(p_2)$, $\dots$, $K_{n-1} \cap D_{\frac{1}{2^{n-1}}}(p_N)$ che non può essere ricoperto con un numero finito di aperti A_i . Una volta definita questa successione di insiemi K_n , prendiamo per ogni n un punto $x_n \in K_n$. Per definizione avremo, per ogni $n \in \mathbb{N}$,

$$\begin{aligned} K_{n+1} &\subseteq K_n, \\ K_{n+1} &\subseteq D_{\frac{1}{2^n}}(p), \end{aligned}$$

per qualche punto p . Se ne deduce che $|x_k - x_m| \leq \frac{1}{2^n}$ per ogni $k, m \geq n$. In particolare, la successione $(x_n) \subseteq K$ sarà di Cauchy, ed essendo K completo, essa tenderà ad un punto $\bar{x} \in K$. A questo punto, esisterà un aperto $A \in \{A_i : i \in I\}$ per cui $\bar{x} \in A$. Essendo A aperto, esso conterrà una palla $B_r(\bar{x})$ centrata in $\bar{x}$, e di conseguenza, se n è tale che $\frac{1}{2^{n-1}} < r$, risulterà che K_{n+1} è contenuto in $B_r(\bar{x}) \subseteq A$. Ma questo implica che K_{n+1} può essere ricoperto da A , ovvero esiste un sottoricoprimento finito di $\{A_i : i \in I\}$ che copre K_{n+1} (tale sottoricoprimento è quello costituito dal solo aperto A). Pertanto abbiamo trovato una contraddizione, che deriva dall'ipotesi fatta che K non fosse compatto.

(iii) $\Rightarrow$ (i): supponiamo K compatto e, sempre per assurdo, assumiamo che K non sia compatto per successioni. Esisterà di conseguenza una successione $(x_n) \subseteq K$ che non ammette sottosuccessioni convergenti. Da questa ipotesi discende il seguente fatto:

$$\forall x \in K \exists r(x) > 0 \quad \text{tale che} \quad \#\{n : x_n \in B_{r(x)}(x)\} < +\infty. \quad (1.2.3)$$

In altre parole, per ogni $x \in K$ esiste una pallina centrata in x che contiene solo un numero finito di punti della successione (x_n) . Infatti, se esistesse un punto $\bar{x}$ tale che per ogni pallina $B_r(x)$ ci fossero infiniti elementi in (x_n) , sarebbe possibile estrarre una sottosuccessione da (x_n) convergente a x (dimostrarlo per esercizio). Ora dalla condizione (1.2.3) deduciamo che la famiglia di palline $\{B_{r(x)}(x) : x \in K\}$ forma un ricoprimento aperto di K . Essendo K compatto possiamo trovare un sottoricoprimento finito, ovvero un numero finito di punti $p_1, p_2, \dots, p_N$ tali che

$$K \subseteq \bigcup_{i=1}^N B_{r(p_i)}(p_i).$$

Questo porta ad un assurdo, perchè se ognuna delle palline a membro di destra contiene un numero finito di punti (x_n) , anche $\bigcup_{i=1}^N B_{r(p_i)}(p_i)$ contiene solo un numero finito di punti in (x_n) ; quindi anche K non potrebbe contenere tutta la successione (x_n) , assurdo. Questo conclude la dimostrazione. $\square$

Osservazione 1.2.12. La compattezza per successioni di un insieme E implica la chiusura di E . Di conseguenza, ogni insieme compatto E (equivalentemente compatto per successioni) risulta anche chiuso.

Nel caso speciale in cui siamo nello spazio metrico $\mathbb{R}^n$ dotato della distanza Euclidea² si dimostra che la limitatezza di un insieme E è equivalente alla limitatezza per successioni di E . Come conseguenza si deduce il seguente teorema:

Teorema 1.2.13. *Sia $E \subseteq \mathbb{R}^n$. Allora E è compatto se e solo se è chiuso e limitato.*

Dimostrazione. Se E è compatto allora è chiuso e totalmente limitato, e quindi, essendo in $\mathbb{R}^n$, sarà limitato. Viceversa, supponiamo E chiuso e limitato. Essendo $\mathbb{R}^n$ completo, ogni suo sottoinsieme chiuso è completo, e quindi anche E . Di conseguenza E è completo e limitato, quindi anche totalmente limitato, e pertanto compatto, come dimostra il Teorema 1.2.11. $\square$

1.3 Lo spazio $\mathbb{R}^n$ e le rette parametriche

Sia n un numero naturale ≥ 1 . Denotiamo con $\mathbb{R}^n := \mathbb{R} \times \mathbb{R} \times \dots \times \mathbb{R}$ il prodotto cartesiano di n copie dell'insieme dei numeri reali $\mathbb{R}$. Nel caso $n = 1$ tale prodotto si riduce ad $\mathbb{R}$ stesso.

Ogni punto $x \in \mathbb{R}^n$ è rappresentato da una n -upla di numeri reali

$$x = (x_1, \dots, x_n), \quad x_i \in \mathbb{R}, \quad \forall i = 1, \dots, n.$$

Lo spazio $\mathbb{R}^n$ dotato della distanza Euclidea $d_e(x, y) = \|x - y\|$ definita in (1.1.1), risulta essere uno spazio metrico completo e soddisfa la proprietà che un insieme $M \subset \mathbb{R}^n$ è limitato se e solo se è totalmente limitato. Segue quindi che un insieme $K \subset \mathbb{R}^n$ è compatto se e solo se è chiuso e limitato.

Sia $\bar{q} = (q_1, \dots, q_n) \in \mathbb{R}^n$ un punto fissato, e sia $\bar{v} = (v_1, \dots, v_n) \in \mathbb{R}^n$ un vettore non nullo (ovvero $\bar{v} \neq (0, \dots, 0)$). Denoteremo con $r_{\bar{q}, \bar{v}}$ la retta passante per il punto $\bar{q}$ e con direzione $\bar{v}$. Ogni punto $\bar{p} \in r_{\bar{q}, \bar{v}}$ di tale retta sarà della forma

$$\bar{p} = \bar{q} + t\bar{v}, \tag{1.3.1}$$

dove $t \in \mathbb{R}$ rappresenta un opportuno numero reale, detto parametro. Se $\bar{p} = (p_1, \dots, p_n)$ sono le coordinate del punto $\bar{p}$, allora l'espressione precedente coincide con

$$(p_1, \dots, p_n) = (q_1 + tv_1, \dots, q_n + tv_n).$$

Al variare del parametro $t \in \mathbb{R}$, varierà la posizione del punto $\bar{p}$ nella formula (1.3.1); per questo indicheremo il punto $\bar{p}$ con $\bar{p}(t)$, per evidenziare che il suo valore dipende dalla scelta del parametro $t \in \mathbb{R}$. Se il vettore $\bar{v}$ è normalizzato (ovvero $\|\bar{v}\| = 1$) allora la distanza tra il punto $\bar{p}(t)$ e il punto $\bar{q}$ sarà esattamente $|t|$. Infatti

$$\|\bar{p}(t) - \bar{q}\| = \|\bar{q} + t\bar{v} - \bar{q}\| = \|t\bar{v}\| = |t|\|\bar{v}\| = |t|.$$

L'espressione

$$\bar{p}(t) = \bar{q} + t\bar{v}, \quad t \in \mathbb{R}, \tag{1.3.2}$$

è detta espressione parametrica della retta $r_{\bar{q}, \bar{v}}$ (o retta in forma parametrica), $\bar{q}$ è detto punto base della retta, $\bar{v}$ la sua direzione.

²In realtà qualsiasi altra distanza equivalente.

Vediamo ora come una retta r nel piano Cartesiano $\mathbb{R}^2$ espressa dall'equazione

$$ax + by + c = 0, \quad (1.3.3)$$

si può sempre esprimere in forma parametrica. Tale espressione è detta forma implicita della retta r , e sta a significare che ogni punto $p = (x, y) \in \mathbb{R}^2$ appartiene alla retta se e solo se le sue coordinate (x, y) soddisfano l'equazione in (1.3.3). In altre parole $r \subset \mathbb{R}^2$ coincide con

$$r := \{(x, y) \in \mathbb{R}^2 : (1.3.3) \text{ è soddisfatta}\}.$$

Inoltre, nella (1.3.3), si suppone che a e b non siano contemporaneamente nulli. Supponiamo dunque $a \neq 0$ e poniamo

$$\bar{q} = (q_1, q_2) := \begin{cases} (-\frac{c}{a}, 0) & \text{se } a \neq 0, \\ (0, -\frac{c}{b}) & \text{se } a = 0. \end{cases}$$

In ambo i casi il punto $\bar{q}$ ha coordinate che soddisfano la (1.3.3), e quindi $\bar{q}$ appartiene alla retta. Definiamo inoltre

$$\bar{v} = (-b, a) \in \mathbb{R}^2,$$

che sarà un vettore non nullo. Supponiamo $a \neq 0$, e prendiamo $t \in \mathbb{R}$ qualsiasi, così il punto $\bar{p}(t) = \bar{q} + t\bar{v} = (-\frac{c}{a} - bt, at)$ si verifica immediatamente che appartiene a r (le sue coordinate soddisfano (1.3.3)). Viceversa supponiamo $\bar{p} = (p_1, p_2) \in r$. Si avrà

$$ap_1 + bp_2 + c = 0 \quad \Rightarrow \quad p_1 = \frac{-c}{a} - \frac{p_2}{a}b = q_1 - tb$$

dove abbiamo posto $t = \frac{p_2}{a}$, mentre

$$p_2 = a\frac{p_2}{a} = at = q_2 + at$$

(perchè $q_2 = 0$). Quindi per qualsiasi $\bar{p} \in r$ abbiamo trovato che esiste $t \in \mathbb{R}$ tale che

$$\bar{p} = \bar{p}(t) = \bar{q} + t\bar{v}.$$

Abbiamo quindi dimostrato che $r = r_{\bar{q}, \bar{v}}$.

Una verifica simile si fa nel caso in cui supponiamo inizialmente che $a = 0$ (farlo per esercizio).

Osservazione 1.3.1. Nel caso in cui $b \neq 0$ sappiamo, dal corso di Analisi 1, che ogni retta r della forma (1.3.3) si può scrivere come luogo di punti del piano le cui coordinate soddisfano

$$y = mx + q.$$

Viceversa, ogni retta che ammette questa espressione si esprime anche in forma implicita come in (1.3.3). Pertanto ogni retta nella forma precedente ammette un'espressione in forma parametrica.

Capitolo 2

Funzioni di più variabili

Considereremo in questo capitolo funzioni f definite su un aperto $A \subseteq \mathbb{R}^n$ e che prendono valori in $\mathbb{R}^k$, con $n, k \geq 1$. In questo caso, una funzione $f : A \rightarrow \mathbb{R}^k$ associa ad ogni punto $x \in A$ il punto $f(x) \in \mathbb{R}^k$. Scriveremo

$$x = (x_1, x_2, \dots, x_n) \qquad f(x) = (f_1(x), f_2(x), \dots, f_k(x)),$$

per indicare le coordinate di x e $f(x)$. In particolare, per ogni $i = 1, \dots, k$, la coordinata f_i di f può essere vista come una funzione $f_i : A \rightarrow \mathbb{R}$ che ad ogni $x \in A$ associa un numero reale $f_i(x)$.

2.1 Calcolo differenziale in più variabili

Definizione 2.1.1. Sia $f : A \rightarrow \mathbb{R}^k$ una funzione definita su un aperto A di $\mathbb{R}^n$, e sia $x \in A$. La funzione f si dice continua in x se per ogni $\varepsilon > 0$ esiste un intorno $U_\varepsilon \subset A$ di x tale che per ogni $y \in U_\varepsilon$ si ha $\|f(x) - f(y)\| \leq \varepsilon$. Inoltre f si dice continua su A se è continua in ogni punto $x \in A$.

Visto che una base per gli aperti di $\mathbb{R}^n$ è fornita dalle palle aperte, possiamo anche riformulare la nozione di continuità di f in x nel seguente modo: per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che se $y \in B_\delta(x)$ allora $\|f(x) - f(y)\| \leq \varepsilon$. Equivalentemente, f è continua in x se vale

$$(C) \quad \forall \varepsilon > 0 \quad \exists \delta > 0 \text{ t.c. } \forall y : \|y - x\| \leq \delta \text{ si ha } \|f(x) - f(y)\| \leq \varepsilon.$$

Visto che A risulta essere uno spazio metrico (con la metrica indotta da $\mathbb{R}^n$), la continuità di una funzione è equivalente alla continuità per successioni. Pertanto (C) è equivalente alla seguente:

$$(C') \quad \text{per ogni successione } (x^n) \subset A \text{ convergente ad } x, \text{ si ha } f(x^n) \rightarrow f(x).$$

Come già osservato prima, se $f : A \rightarrow \mathbb{R}^k$ allora ogni sua componente $f_i : A \rightarrow \mathbb{R}$ è una funzione a valori reali. Visto che per ogni indice $i \in \{1, \dots, k\}$ si ha $|f_i(x) - f_i(y)| \leq \|f(x) - f(y)\|$ si ottiene dalla proprietà (C) che se f è continua in x allora anche f_i è continua in x . E' vero anche il viceversa, come espresso dal seguente esercizio:

Esercizio 2.1.2. Data $f : A \rightarrow \mathbb{R}^k$ come sopra, e dato $x \in A$, dimostrare che f è continua in x se e solo se ogni sua componente f_i risulta continua in x .

Esempio 2.1.3. Facciamo qualche esempio:

- (1) In questo caso $n = 2$ e $k = 1$; definiamo $A = \mathbb{R}^2$ e $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ come

$$f(x) = \|x\| = \sqrt{x_1^2 + x_2^2}.$$

Verifichiamo che in ogni $x \in \mathbb{R}^2$ la funzione f risulta continua. Infatti, sia $y \in \mathbb{R}^2$, allora $f(y) = \|y\|$. Abbiamo, per la disuguaglianza triangolare

$$\|x\| \leq \|x - y\| + \|y\| \quad \Rightarrow \quad \|x\| - \|y\| \leq \|x - y\|$$

e al tempo stesso

$$\|y\| \leq \|x - y\| + \|x\| \quad \Rightarrow \quad \|y\| - \|x\| \leq \|x - y\|.$$

Mettendo insieme le due espressioni trovate otteniamo

$$|\|y\| - \|x\|| \leq \|x - y\| \quad \Rightarrow \quad |f(y) - f(x)| \leq \|x - y\|.$$

Questo implica che, per ogni $\varepsilon > 0$ possiamo prendere $\delta := \varepsilon$ in modo che ogni volta che $\|x - y\| \leq \delta = \varepsilon$ si ha, per la disuguaglianza precedente,

$$|f(y) - f(x)| \leq \varepsilon.$$

Questo dimostra (C), ovvero f è continua in x . Essendo tale argomento valido per ogni $x \in \mathbb{R}^2$, la funzione f è continua su $A = \mathbb{R}^2$.

- (2) In questo esempio $n = 2$ e $k = 2$; definiamo $A = \mathbb{R}^2$ e $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ come

$$f(x) = (f_1(x), f_2(x)) := (\|x\|, 1 - (x_2)^2).$$

Abbiamo già verificato nell'esempio precedente che la prima componente $f_1(x) = \|x\|$ è continua su tutto $\mathbb{R}^2$. Verifichiamo che anche la seconda componente $f_2(x) = 1 - (x_2)^2$ è continua su tutto $A = \mathbb{R}^2$. Questo è evidente dal fatto che essa non dipende dalla variabile x_1 e come funzione della sola variabile x_2 , è continua. In altre parole, se $x^n \rightarrow x$ quando $n \rightarrow \infty$, allora sappiamo che $x_1^n \rightarrow x_1$ e $x_2^n \rightarrow x_2$, e allora

$$f_2(x^n) = 1 - (x_2^n)^2 \rightarrow 1 - (x_2)^2 = f_2(x).$$

- (3) Consideriamo la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ (quindi $n = 2$, $k = 1$) definita da

$$f(x) := \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2} & \text{se } x \neq (0, 0), \\ 0 & \text{se } x = (0, 0). \end{cases}$$

Se consideriamo $x \neq (0, 0)$ allora f risulta continua in x (verificarlo per esercizio). Vediamo se f è continua in $x = (0, 0)$. Utilizzando la condizione (C') dovremmo verificare che ogni volta che x^n è una successione tale che $x^n \rightarrow (0, 0)$ quando $n \rightarrow \infty$, allora $\lim_{n \rightarrow \infty} f(x^n) = f(0, 0) = 0$. Prendendo la successione particolare $x^n = (x_1^n, x_2^n) := (\frac{1}{n}, \frac{1}{n})$, che tende a $(0, 0)$ quando $n \rightarrow \infty$, otterremmo

$$f(x^n) = f\left(\frac{1}{n}, \frac{1}{n}\right) = \frac{\frac{1}{n^2}}{\frac{2}{n^2}} = \frac{1}{2}.$$

Questo valore però non tende a zero quando $n \rightarrow \infty$ (è un valore costante e non dipende neppure da n). Quindi la funzione f non risulta continua in $(0, 0)$.

(4) Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ (quindi $n = 2$, $k = 1$) definita da

$$f(x) := \begin{cases} \frac{x_1^2 x_2}{x_1^2 + x_2^2} & \text{se } x \neq (0, 0), \\ 0 & \text{se } x = (0, 0). \end{cases}$$

Come nell'esempio precedente f risulta continua in ogni $x \neq (0, 0)$. Vediamo che è anche continua in $(0, 0)$. In questo caso, facciamo vedere che per qualsiasi successione $x^n \rightarrow (0, 0)$ si avrà

$$\lim_{n \rightarrow \infty} f(x^n) = 0,$$

il che dimostrerà che f è continua in $(0, 0)$. Scriviamo, utilizzando che $x_1 x_2 \leq \frac{x_1^2}{2} + \frac{x_2^2}{2}$,

$$|f(x^n)| = \left| \frac{(x_1^n)^2 x_2^n}{(x_1^n)^2 + (x_2^n)^2} \right| = |x_1^n| \frac{|x_1^n x_2^n|}{(x_1^n)^2 + (x_2^n)^2} \leq \frac{|x_1^n|}{2} \frac{(x_1^n)^2 + (x_2^n)^2}{(x_1^n)^2 + (x_2^n)^2} = \frac{|x_1^n|}{2},$$

e l'ultimo termine tende a zero quando $n \rightarrow \infty$, qualsiasi sia la successione x^n che tende a $(0, 0)$. Quindi tenderà a zero anche il membro di sinistra, ovvero $f(x^n)$.

2.1.1 Derivabilità

Ricordiamo che $\{e_1, \dots, e_n\}$ denota la base canonica di $\mathbb{R}^n$, dove il vettore e_i coincide con

$$e_i := (0, \dots, 1, \dots, 0),$$

dove tutte le coordinate sono nulle tranne l' i -esima che è uguale ad 1.

Sia $f : A \rightarrow \mathbb{R}^k$ una funzione definita su un aperto A di $\mathbb{R}^n$ e consideriamo un punto $\bar{x} \in A$ fissato. Fissiamo anche un indice $j \in \{1, \dots, n\}$.

Definizione 2.1.4. Diremo che f ammette derivata parziale nella direzione j (o lungo il versore e_j , o rispetto la coordinata j -esima) nel punto $\bar{x}$ se esiste finito il limite

$$\lim_{h \rightarrow 0} \frac{f(\bar{x} + h e_j) - f(\bar{x})}{h}.$$

In tal caso indicheremo con $\frac{\partial f}{\partial x_j}(\bar{x})$ il valore limite.

Si noti che se $f : A \rightarrow \mathbb{R}^k$ è una funzione che ammette derivata rispetto a x_i allora il limite

$$\frac{\partial f}{\partial x_j}(\bar{x}) = \lim_{h \rightarrow 0} \frac{f(\bar{x} + h e_j) - f(\bar{x})}{h},$$

è un vettore in $\mathbb{R}^k$. Si noti inoltre che l'esistenza di tale limite implica che esistono, per ogni $i = 1, \dots, n$, i limiti

$$\frac{\partial f_i}{\partial x_j}(\bar{x}) := \lim_{h \rightarrow 0} \frac{f_i(\bar{x} + h e_j) - f_i(\bar{x})}{h}, \quad (2.1.1)$$

e quindi il vettore $\frac{\partial f}{\partial x_j}(\bar{x})$ coinciderà con il vettore di coordinate $\frac{\partial f_i}{\partial x_j}(\bar{x})$, ovvero

$$\frac{\partial f}{\partial x_j}(\bar{x}) = \left(\frac{\partial f_1}{\partial x_j}(\bar{x}), \dots, \frac{\partial f_k}{\partial x_j}(\bar{x}) \right).$$

Definizione 2.1.5. Sia $f : A \rightarrow \mathbb{R}^k$ una funzione definita su $A \subset \mathbb{R}^n$. Diremo che f è derivabile in un punto $\bar{x} \in A$ se f ammette derivata parziale $\frac{\partial f}{\partial x_j}(\bar{x})$ rispetto ad ogni direzione e_j , per $j = 1, \dots, n$. Inoltre diremo che f è derivabile in A se risulta derivabile in ogni punto di A .

Nel caso in cui f è derivabile in $\bar{x}$, raccoglieremo tutti i vettori $\frac{\partial f}{\partial x_j}$ in una matrice che denoteremo con $\nabla f(\bar{x})$, in cui sulla colonna j -esima inseriremo il vettore colonna $\frac{\partial f}{\partial x_j}$, nel seguente modo

$$\nabla f(\bar{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\bar{x}) & \frac{\partial f_1}{\partial x_2}(\bar{x}) & \cdots & \frac{\partial f_1}{\partial x_n}(\bar{x}) \\ \frac{\partial f_2}{\partial x_1}(\bar{x}) & \frac{\partial f_2}{\partial x_2}(\bar{x}) & \cdots & \frac{\partial f_2}{\partial x_n}(\bar{x}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1}(\bar{x}) & \frac{\partial f_k}{\partial x_2}(\bar{x}) & \cdots & \frac{\partial f_k}{\partial x_n}(\bar{x}) \end{pmatrix} \quad (2.1.2)$$

La componente (i, j) della matrice $\nabla f(\bar{x})$ è la derivata parziale

$$(\nabla f(\bar{x}))_{ij} = \frac{\partial f_i}{\partial x_j}(\bar{x}),$$

per ogni $i = 1, \dots, k$, $j = 1, \dots, n$. In particolare $\nabla f(\bar{x})$ è una matrice $k \times n$, con k righe ed n colonne.

Esempio 2.1.6. Riprendiamo gli esempi in 2.1.3, e discutiamone la derivabilità:

(1) Data $f : \mathbb{R}^2 \rightarrow \mathbb{R}$,

$$f(x) = \|x\| = \sqrt{x_1^2 + x_2^2},$$

mostriamo che essa è derivabile in ogni punto $\bar{x} \neq (0, 0)$. Prendiamo $h \neq 0$ e calcoliamo la derivata di f rispetto a x_1 facendo il limite del rapporto incrementale;

$$\lim_{h \rightarrow 0} \frac{f(\bar{x}_1 + h, \bar{x}_2) - f(\bar{x}_1, \bar{x}_2)}{h} = \lim_{h \rightarrow 0} \frac{\sqrt{(\bar{x}_1 + h)^2 + \bar{x}_2^2} - \sqrt{\bar{x}_1^2 + \bar{x}_2^2}}{h}.$$

Per concludere il conto basta osservare che, quando $h \rightarrow 0$, si ha

$$\begin{aligned} \frac{\sqrt{(\bar{x}_1 + h)^2 + \bar{x}_2^2} - \sqrt{\bar{x}_1^2 + \bar{x}_2^2}}{h} &= \frac{\sqrt{\bar{x}_1^2 + \bar{x}_2^2 + h^2 + 2h\bar{x}_1} - \sqrt{\bar{x}_1^2 + \bar{x}_2^2}}{h} \\ &= \frac{h^2 + 2h\bar{x}_1}{h\sqrt{\bar{x}_1^2 + \bar{x}_2^2 + h^2 + 2h\bar{x}_1} + h\sqrt{\bar{x}_1^2 + \bar{x}_2^2}} \\ &= \frac{h + 2\bar{x}_1}{\sqrt{\bar{x}_1^2 + \bar{x}_2^2 + h^2 + 2h\bar{x}_1} + \sqrt{\bar{x}_1^2 + \bar{x}_2^2}} \\ &\rightarrow \frac{\bar{x}_1}{\sqrt{\bar{x}_1^2 + \bar{x}_2^2}} = \frac{\bar{x}_1}{\|x\|}, \end{aligned}$$

dove i passaggi sono giustificati dal fatto che $\bar{x} \neq (0, 0)$ e quindi il denominatore non è nullo. Osserviamo anche che il limite precedente, essendo un rapporto incrementale, è di fatto il calcolo di una derivata come siamo abituati a fare in un corso di Analisi

1. Precisamente, se trattiamo $\bar{x}_2$ come una costante C e $\bar{x}_1$ come la variabile rispetto a cui si deriva, il limite

$$\lim_{h \rightarrow 0} \frac{\sqrt{(\bar{x}_1 + h)^2 + \bar{x}_2^2} - \sqrt{\bar{x}_1^2 + \bar{x}_2^2}}{h}$$

è esattamente il calcolo della derivata prima della funzione $x_1 \mapsto \sqrt{x_1^2 + C^2}$, che fa infatti $\frac{x_1}{\sqrt{x_1^2 + C^2}}$. Questo ci suggerisce che il calcolo delle derivate parziali può essere fatto semplicemente derivando la funzione in questione vista come funzione della sola variabile rispetto a cui si deriva, e trattando le altre come fossero costanti.

Nell'esempio in questione però dobbiamo ancora trattare il caso in cui $\bar{x} = (0, 0)$. In tal caso, se $h > 0$ il limite del rapporto incrementale diventa

$$\lim_{h \rightarrow 0^+} \frac{\sqrt{(\bar{x}_1 + h)^2 + \bar{x}_2^2} - \sqrt{\bar{x}_1^2 + \bar{x}_2^2}}{h} = \lim_{h \rightarrow 0^+} \frac{\sqrt{h^2}}{h} = \lim_{h \rightarrow 0^+} \frac{|h|}{h} = 1.$$

Invece il limite da sinistra viene

$$\lim_{h \rightarrow 0^-} \frac{\sqrt{(\bar{x}_1 + h)^2 + \bar{x}_2^2} - \sqrt{\bar{x}_1^2 + \bar{x}_2^2}}{h} = \lim_{h \rightarrow 0^-} \frac{|h|}{h} = -1.$$

Questo dimostra che i limiti destro e sinistro non coincidono, quindi non esiste il limite del rapporto incrementale per $h \rightarrow 0$, e pertanto f non è derivabile in $(0, 0)$ rispetto alla variabile x_1 .

Se volessimo calcolare la derivata di f rispetto alla variabile x_2 avremmo una situazione del tutto simile, e verrebbe

$$\frac{\partial f}{\partial x_2}(\bar{x}) = \frac{\bar{x}_2}{\|\bar{x}\|},$$

per gli $\bar{x} \neq (0, 0)$, mentre f non risulterebbe derivabile in $(0, 0)$. In conclusione il gradiente di f esiste solo nei punti $x \neq (0, 0)$ e coincide con

$$\nabla f(x) = \left(\frac{\bar{x}_1}{\|x\|}, \frac{\bar{x}_2}{\|x\|} \right).$$

(2) In questo esempio $f : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ è la funzione

$$f(x) = (f_1(x), f_2(x)) := (\|x\|, 1 - (x_2)^2),$$

ovvero $f_1(x) = \|x\|$ e si tratta come nell'esempio precedente, mentre $f_2(x) = 1 - (x_2)^2$. Per il calcolo del gradiente di f_2 osserviamo che essa non dipende dalla variabile x_1 , e quindi risulta sempre derivabile rispetto a x_1 e la derivata è nulla (f_2 è una costante rispetto x_1),

$$\frac{\partial f_2}{\partial x_1}(x) = 0.$$

Invece, rispetto a x_2 , sarà sufficiente calcolare la derivata nel modo classico, ovvero

$$\frac{\partial f_2}{\partial x_2}(x) = \frac{d}{dx_2}(1 - (x_2)^2) = -2x_2.$$

Quindi se $x \neq (0, 0)$ la f risulta derivabile e il gradiente sarà

$$\nabla f(x) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x) & \frac{\partial f_1}{\partial x_2}(x) \\ \frac{\partial f_2}{\partial x_1}(x) & \frac{\partial f_2}{\partial x_2}(x) \end{pmatrix} = \begin{pmatrix} \frac{x_1}{\|x\|} & \frac{x_2}{\|x\|} \\ 0 & -2x_2 \end{pmatrix}.$$

Invece, se $x = (0, 0)$, la funzione f non è derivabile in x , perchè non è derivabile la sua prima componente f_1 .

(3) Ricordiamo che $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ è definita da

$$f(x) := \begin{cases} \frac{x_1 x_2}{x_1^2 + x_2^2} & \text{se } x \neq (0, 0), \\ 0 & \text{se } x = (0, 0). \end{cases}$$

Se consideriamo $x \neq (0, 0)$ allora f risulta derivabile (verificarlo per esercizio) e derivando prima rispetto a x_1 e poi rispetto a x_2 si trova il gradiente

$$\nabla f(x) = \left(\frac{x_2(x_1^2 - x_2^2)}{(x_1^2 + x_2^2)^2}, \frac{x_1(x_2^2 - x_1^2)}{(x_1^2 + x_2^2)^2} \right).$$

Vediamo in $(0, 0)$ se la funzione è derivabile: calcoliamo a questo scopo il solito rapporto incrementale

$$\frac{\partial f}{\partial x_1}((0, 0)) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0 \cdot h - 0}{h} = 0,$$

e similmente

$$\frac{\partial f}{\partial x_2}((0, 0)) = 0.$$

Pertanto f risulta anche derivabile in $(0, 0)$, nonostante non sia continua!

(4) Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ la funzione

$$f(x) := \begin{cases} \frac{x_1^2 x_2}{x_1^2 + x_2^2} & \text{se } x \neq (0, 0), \\ 0 & \text{se } x = (0, 0). \end{cases}$$

Anche in questo caso, per $x \neq (0, 0)$ la f risulta derivabile (verificarlo per esercizio) e derivando prima rispetto a x_1 e poi rispetto a x_2 si trova il gradiente

$$\nabla f(x) = \left(\frac{2x_1 x_2^3}{(x_1^2 + x_2^2)^2}, \frac{x_1^4 - x_1^2 x_2^2}{(x_1^2 + x_2^2)^2} \right).$$

Vediamo in $(0, 0)$ se la funzione è derivabile: calcoliamo a questo scopo il solito rapporto incrementale

$$\frac{\partial f}{\partial x_1}((0, 0)) = \lim_{h \rightarrow 0} \frac{f(h, 0) - f(0, 0)}{h} = \lim_{h \rightarrow 0} \frac{0 \cdot h^2 - 0}{h} = 0,$$

e similmente

$$\frac{\partial f}{\partial x_2}((0, 0)) = 0.$$

Pertanto f risulta anche derivabile in $(0, 0)$.

Così come, data una funzione $f : A \rightarrow \mathbb{R}^k$, possiamo chiederci se essa ammette derivata nella direzione e_j , allo stesso modo possiamo fissare una direzione $v \in \mathbb{R}^n$ non necessariamente parallela ad uno degli assi coordinati, e chiederci se f ammette derivata in quella direzione.

Definizione 2.1.7. Sia $f : A \rightarrow \mathbb{R}^k$ con $A \subset \mathbb{R}^n$ aperto, e sia $\bar{x} \in A$ e $v \in \mathbb{R}^n$, $v \neq 0$. Diremo che f ammette derivata direzionale in $\bar{x}$ nella direzione v se esiste finito il limite

$$\frac{\partial f}{\partial v}(\bar{x}) := \lim_{h \rightarrow 0} \frac{f(\bar{x} + hv) - f(\bar{x})}{h}.$$

Nel caso specifico in cui $v = e_j$ si ritrova ovviamente la derivata parziale rispetto a x_j . In generale però v può essere una combinazione lineare non banale delle e_j .

Esempio 2.1.8. (1) Prendiamo $f : \mathbb{R}^3 \rightarrow \mathbb{R}$ definita come

$$f(x) = x_1 x_2 - 3x_3,$$

e calcoliamo $\frac{\partial f}{\partial v}(\bar{x})$ per $x \in \mathbb{R}^3$, e

$$v := (1, -2, -1).$$

Dobbiamo allora calcolare il limite

$$\begin{aligned} \frac{\partial f}{\partial v}(x) &:= \lim_{h \rightarrow 0} \frac{f(x + hv) - f(x)}{h} = \lim_{h \rightarrow 0} \frac{f(x_1 + h, x_2 - 2h, x_3 - h) - f(x_1, x_2, x_3)}{h} \\ &= \lim_{h \rightarrow 0} \frac{(x_1 + h)(x_2 - 2h) - 3(x_3 - h) - x_1 x_2 + 3x_3}{h} \\ &= \lim_{h \rightarrow 0} (-2x_1 + x_2 - 2h + 3) = -2x_1 + x_2 + 3. \end{aligned}$$

(2) Consideriamo la funzione f nell'esempio 2.1.3 punto (4) e calcoliamo la derivata direzionale in $(0, 0)$ rispetto al vettore $v = (1, -1)$. Si ha

$$\frac{\partial f}{\partial v}((0, 0)) = \lim_{h \rightarrow 0} \frac{f((h, -h)) - f((0, 0))}{h} = \lim_{h \rightarrow 0} \frac{-h^3 - 0}{2h^3} = -\frac{1}{2}.$$

2.1.2 Differenziabilità

Il concetto di differenziabilità descrive una proprietà ancora più forte della derivabilità.

Definizione 2.1.9. Sia $f : A \rightarrow \mathbb{R}^k$ con $A \subset \mathbb{R}^n$ aperto, e sia $\bar{x} \in A$; supponiamo f sia derivabile nel punto $\bar{x}$. Allora diremo che f è differenziabile in $\bar{x}$ se vale

$$\lim_{x \rightarrow \bar{x}} \frac{f(x) - f(\bar{x}) - \nabla f(\bar{x})(x - \bar{x})}{\|x - \bar{x}\|} = 0. \quad (2.1.3)$$

Si noti che al numeratore della frazione che compare nel membro di destra compare il prodotto di $\nabla f(\bar{x})$ con il vettore $(x - \bar{x})$. Esso è un prodotto tra una matrice $\mathbb{R}^{k \times n}$ e un vettore con n componenti, e va quindi inteso come un prodotto "riga per colonna". Nello specifico, il risultato sarà un vettore in $\mathbb{R}^k$ (quindi con k componenti) e avremo

$$\nabla f(\bar{x})(x - \bar{x}) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\bar{x})(x_j - \bar{x}_j),$$

dove ognuno degli $\frac{\partial f}{\partial x_j}$ è un vettore con k componenti. La componente i -esima dunque di $\nabla f(\bar{x})(x - \bar{x})$ è data da

$$(\nabla f(\bar{x})(x - \bar{x}))_i = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(\bar{x})(x_j - \bar{x}_j).$$

Quindi il numeratore della frazione che compare in (2.1.3) è un vettore in $\mathbb{R}^k$. Il denominatore è uno scalare, quindi la frazione intera è un vettore in $\mathbb{R}^k$ e dire che il suo limite è 0 significa che come vettore tende all'origine (ovvero con 0 si denota l'origine in $\mathbb{R}^k$, ovvero il vettore $(0, \dots, 0) \in \mathbb{R}^k$). Per questo motivo, la (2.1.3) è equivalente a richiedere che

$$\lim_{x \rightarrow \bar{x}} \frac{\|f(x) - f(\bar{x}) - \nabla f(\bar{x})(x - \bar{x})\|}{\|x - \bar{x}\|} = 0, \quad (2.1.4)$$

dove questa volta la frazione è uno scalare che tende a 0 inteso come numero reale. Si noti che in quest'ultima espressione al numeratore compare la norma di un vettore in $\mathbb{R}^k$, mentre al denominatore la norma di un vettore in $\mathbb{R}^n$.

Esempio 2.1.10. Riprendiamo l'analisi della funzione al punto (4) dell'Esempio 2.1.3. Per verificare la differenziabilità in $\bar{x} = (0, 0)$, dovremmo mostrare la validità della (2.1.3) per $\bar{x} = (0, 0)$. Visto che le derivate direzionali in $(0, 0)$ sono

$$\frac{\partial f}{\partial x_1}((0, 0)) = 0, \quad \frac{\partial f}{\partial x_2}((0, 0)) = 0,$$

(verificarlo per esercizio), avremo $\nabla f((0, 0)) = (0, 0)$ e quindi bisognerà studiare il limite per $x \rightarrow (0, 0)$ della quantità

$$\frac{f(x) - f((0, 0)) - \nabla f((0, 0))x}{\|x\|} = \frac{\frac{x_1^2 x_2}{x_1^2 + x_2^2}}{\|x\|} = \frac{x_1^2 x_2}{(x_1^2 + x_2^2)^{3/2}}.$$

Tale limite non esiste, infatti prendendo la successione $(x)_n = (\frac{1}{n}, 0)$ che tende a $(0, 0)$ per $n \rightarrow \infty$, avremo

$$\frac{f((x)_n) - f((0, 0)) - \nabla f((0, 0))x}{\|(x)_n\|} = 0 \quad \forall n,$$

e quindi

$$\lim_{n \rightarrow \infty} \frac{f((x)_n) - f((0, 0)) - \nabla f((0, 0))x}{\|(x)_n\|} = 0.$$

Invece scegliendo la successione $(x)_n = (\frac{1}{n}, \frac{1}{n})$ (tende a $(0, 0)$ anch'essa), abbiamo

$$\frac{f((x)_n) - f((0, 0)) - \nabla f((0, 0))x}{\|(x)_n\|} = \frac{1/n^3}{(2/n^2)^{3/2}} = \frac{1}{2^{3/2}} \rightarrow \frac{1}{2^{3/2}}.$$

Quindi avremo due successioni che tendono a $(0, 0)$ lungo le quali il limite della quantità sopra è diverso, e pertanto il limite per $x \rightarrow (0, 0)$ di tale quantità non esiste. Concludiamo che la funzione f non è differenziabile in $(0, 0)$.

Teorema 2.1.11. *Sia $A \subset \mathbb{R}^n$ aperto, sia $f : A \rightarrow \mathbb{R}^k$ e sia $\bar{x} \in A$; se f è differenziabile in $\bar{x}$ allora valgono le seguenti conclusioni:*

(a) f è continua in $\bar{x}$;

(b) per ogni vettore $v \in \mathbb{R}^n$, $v \neq 0$, esiste la derivata direzionale di f rispetto a v , nel punto $\bar{x}$, e vale

$$\frac{\partial f}{\partial v}(\bar{x}) = \nabla f(\bar{x})v.$$

Dimostrazione. Per verificare il punto (a) scriviamo

$$\begin{aligned} \|f(x) - f(\bar{x})\| &= \|f(x) - f(\bar{x}) - \nabla f(\bar{x})(x - \bar{x}) + \nabla f(\bar{x})(x - \bar{x})\| \\ &\leq \|f(x) - f(\bar{x}) - \nabla f(\bar{x})(x - \bar{x})\| + \|\nabla f(\bar{x})(x - \bar{x})\| \\ &= \frac{\|f(x) - f(\bar{x}) - \nabla f(\bar{x})(x - \bar{x})\|}{\|x - \bar{x}\|} \|x - \bar{x}\| + \|\nabla f(\bar{x})(x - \bar{x})\| \\ &\leq \frac{\|f(x) - f(\bar{x}) - \nabla f(\bar{x})(x - \bar{x})\|}{\|x - \bar{x}\|} \|x - \bar{x}\| + \|\nabla f(\bar{x})\| \|x - \bar{x}\|, \end{aligned}$$

e adesso, facendo il limite per x che tende a $\bar{x}$, il membro di destra tende a zero, visto che $\|x - \bar{x}\| \rightarrow 0$, la frazione tende a zero per la (2.1.4), e il numero $\|\nabla f(\bar{x})\|$ è una costante che non dipende da x . Quindi anche il membro di sinistra (che è non negativo) tende a zero, e la continuità è dimostrata.

Facciamo vedere ora il punto (b): ricordiamo che $\nabla f(\bar{x})v$ è un vettore in $\mathbb{R}^k$ ottenuto moltiplicando la matrice $\nabla f(\bar{x})$ per il vettore v . Prendiamo $h \neq 0$ e $x := \bar{x} + hv$ in modo tale che quando $h \rightarrow 0$ si ha $x \rightarrow \bar{x}$. Quindi utilizzando il limite (2.1.3), con questa scelta di x , si avrà

$$0 = \lim_{h \rightarrow 0} \frac{f(\bar{x} + hv) - f(\bar{x}) - \nabla f(\bar{x})(hv)}{|h||v|} = \lim_{h \rightarrow 0} \left(\frac{f(\bar{x} + hv) - f(\bar{x})}{|h||v|} - \frac{h}{|h||v|} \nabla f(\bar{x})v \right).$$

Per $h > 0$, si ha $|h| = h$ e l'espressione precedente diventa

$$\lim_{h \rightarrow 0^+} \frac{f(\bar{x} + hv) - f(\bar{x})}{h} = \nabla f(\bar{x})v,$$

mentre se $h < 0$ si ha $|h| = -h$ e allora

$$\lim_{h \rightarrow 0^-} \frac{f(\bar{x} + hv) - f(\bar{x})}{h} = \nabla f(\bar{x})v.$$

Gli ultimi due limiti ci dicono che esiste la derivata parziale di f in $\bar{x}$ in direzione v , e che coincide con $\nabla f(\bar{x})v$. $\square$

Vediamo ora una condizione sufficiente per garantire la differenziabilità di una funzione.

Teorema 2.1.12 (Teorema del differenziale totale). *Sia $A \subset \mathbb{R}^n$ aperto, sia $f : A \rightarrow \mathbb{R}$ e sia $\bar{x} \in A$; supponiamo che esista $\delta > 0$ tale che f risulta derivabile in $B_\delta(\bar{x})$ e supponiamo che le derivate parziali $\frac{\partial f}{\partial x_j}$, per $j = 1, \dots, n$, siano continue in $\bar{x}$. Allora f risulta differenziabile in $\bar{x}$.*

Dimostrazione. Per semplificare i conti, supponiamo che $n = 2$ (ovvero la funzione dipende da sole due variabili e $x = (x_1, x_2)$). Dobbiamo mostrare la validità di (2.1.3), dove ricordiamo

$$\nabla f(\bar{x})(x - \bar{x}) = \sum_{j=1}^n \frac{\partial f}{\partial x_j}(\bar{x})(x_j - \bar{x}_j) = \frac{\partial f}{\partial x_1}(\bar{x})(x_1 - \bar{x}_1) + \frac{\partial f}{\partial x_2}(\bar{x})(x_2 - \bar{x}_2).$$

Quindi mostriamo che

$$\lim_{x \rightarrow \bar{x}} \frac{f(x) - f(\bar{x}) - \frac{\partial f}{\partial x_1}(\bar{x})(x_1 - \bar{x}_1) - \frac{\partial f}{\partial x_2}(\bar{x})(x_2 - \bar{x}_2)}{\|x - \bar{x}\|} = 0. \quad (2.1.5)$$

Consideriamo la quantità $f(x) - f(\bar{x}) = f((x_1, x_2)) - f((\bar{x}_1, \bar{x}_2))$, e supponendo senza perdita di generalità che $x_1 > \bar{x}_1$ e $x_2 > \bar{x}_2$ riscriviamo tale quantità come

$$f((x_1, x_2)) - f((\bar{x}_1, \bar{x}_2)) = f((x_1, x_2)) - f((x_1, \bar{x}_2)) + f((x_1, \bar{x}_2)) - f((\bar{x}_1, \bar{x}_2)). \quad (2.1.6)$$

Essendo la funzione $t \mapsto f((t, \bar{x}_2))$ continua sull'intervallo $[\bar{x}_1, x_1]$ e derivabile in $(\bar{x}_1, x_1)$ (la sua derivata sarà $\frac{\partial f}{\partial x_1}(t, \bar{x}_2)$) possiamo applicare ad essa il teorema di Lagrange deducendo l'esistenza di un punto $\xi \in (\bar{x}_1, x_1)$ tale che

$$f((x_1, \bar{x}_2)) - f((\bar{x}_1, \bar{x}_2)) = \frac{\partial f}{\partial x_1}(\xi, \bar{x}_2)(x_1 - \bar{x}_1).$$

Similmente anche la funzione $s \mapsto f((x_1, s))$ è continua su $[\bar{x}_2, x_2]$ e derivabile in $(\bar{x}_2, x_2)$ (la sua derivata sarà $\frac{\partial f}{\partial x_2}(x_1, s)$) e sempre il teorema di Lagrange ci fornisce $\zeta \in (\bar{x}_2, x_2)$ tale che

$$f((x_1, x_2)) - f((x_1, \bar{x}_2)) = \frac{\partial f}{\partial x_2}(x_1, \zeta)(x_2 - \bar{x}_2).$$

Inserendo le ultime due espressioni in (2.1.6) otteniamo

$$f((x_1, x_2)) - f((\bar{x}_1, \bar{x}_2)) = \frac{\partial f}{\partial x_1}(\xi, \bar{x}_2)(x_1 - \bar{x}_1) + \frac{\partial f}{\partial x_2}(x_1, \zeta)(x_2 - \bar{x}_2),$$

e così inserendo tale espressione in (2.1.5), lo scopo della dimostrazione sarà far vedere che la quantità

$$\begin{aligned} & \frac{\frac{\partial f}{\partial x_1}(\xi, \bar{x}_2)(x_1 - \bar{x}_1) + \frac{\partial f}{\partial x_2}(x_1, \zeta)(x_2 - \bar{x}_2) - \frac{\partial f}{\partial x_1}(\bar{x}_1, \bar{x}_2)(x_1 - \bar{x}_1) - \frac{\partial f}{\partial x_2}(\bar{x}_1, \bar{x}_2)(x_2 - \bar{x}_2)}{\|x - \bar{x}\|} \\ &= \frac{\left(\frac{\partial f}{\partial x_1}(\xi, \bar{x}_2) - \frac{\partial f}{\partial x_1}(\bar{x}_1, \bar{x}_2) \right) (x_1 - \bar{x}_1) + \left(\frac{\partial f}{\partial x_2}(x_1, \zeta) - \frac{\partial f}{\partial x_2}(\bar{x}_1, \bar{x}_2) \right) (x_2 - \bar{x}_2)}{\|x - \bar{x}\|} \\ &= \left(\frac{\partial f}{\partial x_1}(\xi, \bar{x}_2) - \frac{\partial f}{\partial x_1}(\bar{x}_1, \bar{x}_2) \right) \frac{(x_1 - \bar{x}_1)}{\|x - \bar{x}\|} + \left(\frac{\partial f}{\partial x_2}(x_1, \zeta) - \frac{\partial f}{\partial x_2}(\bar{x}_1, \bar{x}_2) \right) \frac{(x_2 - \bar{x}_2)}{\|x - \bar{x}\|} \end{aligned}$$

tende a zero quando $x \rightarrow \bar{x}$. Ma ora le quantità $\frac{x_1 - \bar{x}_1}{\|x - \bar{x}\|}$ e $\frac{x_2 - \bar{x}_2}{\|x - \bar{x}\|}$ sono, in valore assoluto, minori o uguali a 1; quindi basterà far vedere che entrambi i termini

$$\left(\frac{\partial f}{\partial x_1}(\xi, \bar{x}_2) - \frac{\partial f}{\partial x_1}(\bar{x}_1, \bar{x}_2) \right), \quad \left(\frac{\partial f}{\partial x_2}(x_1, \zeta) - \frac{\partial f}{\partial x_2}(\bar{x}_1, \bar{x}_2) \right),$$

tendono a zero quando $x \rightarrow \bar{x}$. Il primo termine tende a zero perchè $\frac{\partial f}{\partial x_1}$ è continua in $\bar{x}$ e, quando $x \rightarrow \bar{x}$, anche $x_1 \rightarrow \bar{x}_1$ e così, essendo $\xi \in (\bar{x}_1, x_1)$, anche $\xi \rightarrow \bar{x}_1$. Stesso ragionamento per il secondo termine, in cui basta osservare che se $x \rightarrow \bar{x}$ allora $(x_1, \zeta) \rightarrow (\bar{x}_1, \bar{x}_2)$. La dimostrazione è conclusa. $\square$

2.1.3 Derivate successive

Così come abbiamo calcolato le derivate parziali di una funzione $f : A \rightarrow \mathbb{R}^k$, con $A \subset \mathbb{R}^n$ aperto, possiamo anche calcolare le derivate parziali di $\frac{\partial f}{\partial x_j}$. Infatti, se la derivata parziale $\frac{\partial f}{\partial x_j}$ esiste in A , allora essa sarà una funzione $\frac{\partial f}{\partial x_j} : A \rightarrow \mathbb{R}^k$, e come tale potrebbe ammettere derivate parziali.

Definizione 2.1.13. Sia $f : A \rightarrow \mathbb{R}^k$, con $A \subset \mathbb{R}^n$ aperto, e supponiamo che, per qualche $j \in \{1, \dots, n\}$, esista la derivata parziale $\frac{\partial f}{\partial x_j}$ in un punto $\bar{x} \in A$. Sia $i \in \{1, \dots, n\}$. Diremo che f ammette derivata seconda rispetto a x_i e x_j in $\bar{x}$ se $\frac{\partial f}{\partial x_j}$ ammette derivata parziale rispetto a x_i in $\bar{x}$. In tal caso scriveremo

$$\frac{\partial^2 f}{\partial x_j \partial x_i}(\bar{x}) := \frac{\partial}{\partial x_i} \left(\frac{\partial f}{\partial x_j} \right) (\bar{x}) = \lim_{h \rightarrow 0} \frac{\frac{\partial f}{\partial x_j}(\bar{x} + h e_i) - \frac{\partial f}{\partial x_j}(\bar{x})}{h}.$$

Nel caso in cui $i = j$ scriveremo anche

$$\frac{\partial^2 f}{(\partial x_i)^2}(\bar{x}) := \frac{\partial^2 f}{\partial x_i \partial x_i}(\bar{x}).$$

Supponiamo per semplicità che $f : A \rightarrow \mathbb{R}$ assuma valori reali (quindi non in $\mathbb{R}^k$ con $k > 1$).

Definizione 2.1.14. Nel caso in cui $f : A \rightarrow \mathbb{R}$ ammetta in $\bar{x}$ tutte le derivate parziali seconde $\frac{\partial^2 f}{\partial x_j \partial x_i}(\bar{x})$ per ogni $i, j = 1, \dots, n$, diremo che f è derivabile due volte in $\bar{x}$. In tal caso si definisce l'Hessiano di f in $\bar{x}$ la matrice $\nabla^2 f(\bar{x})$ contenente tutte le derivate parziali seconde di f in $\bar{x}$, ovvero

$$\nabla^2 f(\bar{x}) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\bar{x}) & \frac{\partial^2 f}{\partial x_1 \partial x_2}(\bar{x}) & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(\bar{x}) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(\bar{x}) & \frac{\partial^2 f}{\partial x_2^2}(\bar{x}) & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n}(\bar{x}) \\ \cdots & \cdots & \cdots & \cdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(\bar{x}) & \frac{\partial^2 f}{\partial x_n \partial x_2}(\bar{x}) & \cdots & \frac{\partial^2 f}{\partial x_n^2}(\bar{x}) \end{pmatrix} \quad (2.1.7)$$

In altre parole $\nabla^2 f(\bar{x})$ è quella matrice che in posizione della riga i -esima e colonna j -esima ha

$$\nabla^2 f(\bar{x})_{ij} = \frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}).$$

Nel caso in cui $f : A \rightarrow \mathbb{R}^k$ l'Hessiano sarà un tensore $k \times n \times n$. In altre parole, per ogni $i = 1, \dots, n$, la funzione i -esima componente di f sarà $f_i : A \rightarrow \mathbb{R}$ e ammetterà Hessiano

$$\nabla^2 f_i(\bar{x})_{kj} = \frac{\partial^2 f_i}{\partial x_k \partial x_j}(\bar{x}),$$

$k, j = 1, \dots, n$.

Il seguente teorema fornisce condizioni sufficienti affinché le derivate miste $\frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x})$ e $\frac{\partial^2 f}{\partial x_j \partial x_i}(\bar{x})$ ottenute derivando prima rispetto a x_i e poi rispetto a x_j , e viceversa, coincidano.

Teorema 2.1.15 (Teorema di Schwarz). *Sia $A \subset \mathbb{R}^n$ un aperto e $f : A \rightarrow \mathbb{R}$ funzione che ammette derivate seconde miste $\frac{\partial^2 f}{\partial x_i \partial x_j}$ e $\frac{\partial^2 f}{\partial x_j \partial x_i}$ per alcuni indici $i \neq j$. Supponiamo che tali derivate siano continue in $\bar{x} \in A$; allora*

$$\frac{\partial^2 f}{\partial x_i \partial x_j}(\bar{x}) = \frac{\partial^2 f}{\partial x_j \partial x_i}(\bar{x}).$$

Il teorema è enunciato nel caso in cui $A \subset \mathbb{R}^n$ e quindi f dipende da n variabili. La dimostrazione la diamo nel caso in cui $n = 2$, essendo tuttavia estendibile al caso generale in modo immediato. Pertanto stiamo supponendo f dipenda da sole due variabili x_1 e x_2 e vogliamo dimostrare che sotto le ipotesi del teorema vale

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(\bar{x}) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\bar{x}).$$

Dimostrazione. Consideriamo un punto $x = (x_1, x_2) \in A$ e denotiamo con $h := x_1 - \bar{x}_1$ e $k := x_2 - \bar{x}_2$. Visto che vorremo mandare al limite x a $\bar{x}$, possiamo supporre che ci sia una palla $B_\delta(\bar{x}) \subset A$ e $x \in B_\delta(\bar{x})$. Inoltre, supponiamo per semplicità che $h > 0$ e $k > 0$ (gli altri casi si tratteranno in modo analogo).

Definiamo la funzione G come

$$G(h, k) := f(\bar{x}_1 + h, \bar{x}_2 + k) - f(\bar{x}_1 + h, \bar{x}_2) - f(\bar{x}_1, \bar{x}_2 + k) + f(\bar{x}_1, \bar{x}_2). \quad (2.1.8)$$

Introduciamo anche la funzione di una variabile $F : [\bar{x}_1, \bar{x}_1 + h] \rightarrow \mathbb{R}$ definita come

$$F(t) := f(t, \bar{x}_2 + k) - f(t, \bar{x}_2). \quad (2.1.9)$$

Tale funzione risulta continua su $[\bar{x}_1, \bar{x}_1 + h]$ e derivabile nell'intervallo $(\bar{x}_1, \bar{x}_1 + h)$, con derivata

$$F'(t) = \frac{\partial f}{\partial x_1}(t, \bar{x}_2 + k) - \frac{\partial f}{\partial x_1}(t, \bar{x}_2). \quad (2.1.10)$$

Con l'ausilio di F , la formula in (2.1.8) diventa

$$G(h, k) = F(\bar{x}_1 + h) - F(\bar{x}_1),$$

e allora possiamo applicare il teorema di Lagrange, che ci fornisce un numero $\xi_1 \in (\bar{x}_1, \bar{x}_1 + h)$ tale che

$$G(h, k) = F'(\xi_1)h = \left(\frac{\partial f}{\partial x_1}(\xi_1, \bar{x}_2 + k) - \frac{\partial f}{\partial x_1}(\xi_1, \bar{x}_2) \right) h.$$

Possiamo adesso studiare la quantità a membro destro: la funzione $t \mapsto \frac{\partial f}{\partial x_1}(\xi_1, t)$ risulta derivabile in t grazie alle ipotesi che esistono le derivate seconde parziali, e visto che infatti la derivata rispetto a t di $\frac{\partial f}{\partial x_1}(\xi_1, t)$ coincide con $\frac{\partial}{\partial x_2} \frac{\partial f}{\partial x_1}(\xi_1, x_2)$. Quindi ancora il Teorema di Lagrange ci garantisce che esiste un numero $\xi_2 \in (\bar{x}_2, \bar{x}_2 + k)$ tale che

$$\frac{\partial f}{\partial x_1}(\xi_1, \bar{x}_2 + k) - \frac{\partial f}{\partial x_1}(\xi_1, \bar{x}_2) = \frac{\partial}{\partial x_2} \frac{\partial f}{\partial x_1}(\xi_1, \xi_2)k = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\xi_1, \xi_2).$$

Quindi tornando all'espressione precedente, otteniamo

$$G(h, k) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\xi_1, \xi_2)kh. \quad (2.1.11)$$

Ora ripartiamo dalla (2.1.8) e introduciamo la mappa $H : [\bar{x}_2, \bar{x}_2 + k] \rightarrow \mathbb{R}$ definita come

$$H(t) := f(\bar{x}_1 + h, t) - f(\bar{x}_1, t). \quad (2.1.12)$$

In tal modo osserviamo che $G(h, k) = H(\bar{x}_2 + k) - H(\bar{x}_2)$, e ragionando in modo analogo a prima, troviamo un numero $\eta_2 \in (\bar{x}_2, \bar{x}_2 + k)$ tale che

$$G(h, k) = H'(\eta_2)k = \left(\frac{\partial f}{\partial x_2}(\bar{x}_1 + h, \eta_2) - \frac{\partial f}{\partial x_2}(\bar{x}_1, \eta_2) \right) k.$$

Procedendo come nel caso precedente, si trova un $\eta_1 \in (\bar{x}_1, \bar{x}_1 + h)$ tale che

$$G(h, k) = \left(\frac{\partial f}{\partial x_2}(\bar{x}_1 + h, \eta_2) - \frac{\partial f}{\partial x_2}(\bar{x}_1, \eta_2) \right) k = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\eta_1, \eta_2) h k.$$

Da questa espressione e dalla (2.1.11) si trova

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(\xi_1, \xi_2) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\eta_1, \eta_2),$$

per ogni h, k scelti come all'inizio. Ora, se facciamo tendere $(h, k) \rightarrow (0, 0)$, vediamo, per come sono stati trovati i numeri $\xi_1, \xi_2, \eta_1, \eta_2$, che $(\xi_1, \xi_2) \rightarrow (\bar{x}_1, \bar{x}_2)$ e $(\eta_1, \eta_2) \rightarrow (\bar{x}_1, \bar{x}_2)$. Pertanto, essendo le derivate continue in $(\bar{x}_1, \bar{x}_2)$, i concludiamo

$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(\bar{x}) = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\bar{x}),$$

ovvero la tesi. □

Definizione 2.1.16. Una funzione $f : A \rightarrow \mathbb{R}^k$ si dice di classe C^0 se è continua su A . In tal caso scriveremo $f \in C^0(A; \mathbb{R}^k)$ dove $C^0(A; \mathbb{R}^k)$ è l'insieme di tutte le funzioni continue da A in $\mathbb{R}^k$.

Sia $m \geq 1$ un numero naturale. Diremo che una funzione $f : A \rightarrow \mathbb{R}^k$ è di classe C^m se f ammette tutte le derivate di ordine m ed esse risultano continue su A . Se f è di classe C^m scriveremo $f \in C^m(A; \mathbb{R}^k)$, dove $C^m(A; \mathbb{R}^k)$ denota l'insieme di tutte le funzioni da A in $\mathbb{R}^k$ di classe C^m .

Infine, diremo che $f : A \rightarrow \mathbb{R}^k$ è di classe C^∞ se f ammette tutte le derivate di qualsiasi ordine $m \geq 1$ ed esse risultano continue su A . Anche in questo caso denoteremo con $C^\infty(A; \mathbb{R}^k)$ l'insieme di tutte le funzioni da A in $\mathbb{R}^k$ di classe C^∞ .

2.1.4 Composizione di funzioni di più variabili

Consideriamo una funzione $g : A \rightarrow B$ dove $A \subset \mathbb{R}^n$ e $B \subset \mathbb{R}^m$ sono insiemi aperti, e sia $f : B \rightarrow \mathbb{R}^k$ una funzione data. Possiamo considerare allora la composizione di g ed f come la funzione che ad ogni $x \in A$, associa il valore $f(g(x)) \in \mathbb{R}^k$. Indicheremo tale funzione con $f \circ g : A \rightarrow \mathbb{R}^k$ e sarà chiamata funzione composta di f con g .

Esempio 2.1.17. (1) Siano $A = \mathbb{R}$ e $B = \mathbb{R}^m$, con $m > 1$. Siano $\bar{x} \in \mathbb{R}^m$ e $v \in \mathbb{R}^m \setminus \{0\}$ due vettori prefissati e definiamo $g(t) := \bar{x} + tv$ per ogni $t \in \mathbb{R}$. Sia inoltre $f : \mathbb{R}^m \rightarrow \mathbb{R}^k$ una funzione generica; la funzione composta $f \circ g : \mathbb{R} \rightarrow \mathbb{R}^k$ sarà allora

$$f(g(t)) = f(\bar{x} + tv), \quad \forall t \in \mathbb{R}.$$

Possiamo allora calcolare il gradiente di $f \circ g$ (qualora risulti derivabile). In tal caso, visto che $f \circ g$ è funzione di una sola variabile t , dovremmo calcolare un'unica derivata, che consiste in

$$(f \circ g)'(t) = \lim_{h \rightarrow 0} \frac{f(g(t+h)) - f(g(t))}{h} = \lim_{h \rightarrow 0} \frac{f(\bar{x} + tv + hv) - f(\bar{x} + tv)}{h}.$$

Ponendo $\bar{y} := \bar{x} + tv$, tale limite coincide con

$$\lim_{h \rightarrow 0} \frac{f(\bar{y} + hv) - f(\bar{y})}{h} = \frac{\partial f}{\partial v}(\bar{y}) = \nabla f(\bar{y})v.$$

La prima uguaglianza è valida chiaramente se il limite esiste finito (e in tal caso coincide con la derivata parziale di f rispetto a v in $\bar{y}$), mentre la seconda uguaglianza è valida, ad esempio, se f è differenziabile in $\bar{y}$ (e in tal caso l'uguaglianza segue dal punto (b) del Teorema 2.1.11). In definitiva abbiamo $(f \circ g)'(t) = \nabla f(\bar{x} + tv)v$. Ricordando che $\bar{x} + tv = g(t)$ l'ultima espressione si riscrive come

$$(f \circ g)'(t) = \nabla f(g(t))v.$$

D'altro canto, $g'(t) = v$, e quindi abbiamo ottenuto la formula di derivata per la funzione composta

$$(f \circ g)'(t) = \nabla f(g(t))g'(t),$$

dove il prodotto tra $\nabla f(g(t))$ e $g'(t)$ è da intendersi come prodotto tra una matrice e un vettore, riga per colonna.

- (2) Proseguendo come nell'esempio precedente, abbiamo verificato che la funzione $t \rightarrow f \circ g(t) = f(\bar{x} + tv)$ è derivabile (in realtà differenziabile) qualora tale risulti f . Questa differenziabilità è valida per ogni $t \in \mathbb{R}$, e quindi la quantità

$$f(\bar{x} + tv) - f(\bar{x}) = f(g(t)) - f(g(0))$$

può essere espressa, tramite il Teorema di Lagrange, nel seguente modo

$$f(\bar{x} + tv) - f(\bar{x}) = f(g(t)) - f(g(0)) = t(f \circ g)'(\xi),$$

dove $\xi \in (0, t)$ è un numero opportuno (stiamo supponendo $t > 0$, altrimenti avremmo scritto $\xi \in (t, 0)$). Quindi

$$f(\bar{x} + tv) - f(\bar{x}) = t \nabla f(g(\xi))g'(\xi) = t \frac{\partial f}{\partial v}(g(\xi)) = t \frac{\partial f}{\partial v}(\bar{x} + \xi v). \quad (2.1.13)$$

Questa formula tornerà comoda nel seguito.

Teorema 2.1.18 (di derivazione della funzione composta). *Siano $A \subset \mathbb{R}^n$ e $B \subset \mathbb{R}^m$ insiemi aperti, siano $g : A \rightarrow B$ ed $f : B \rightarrow \mathbb{R}^k$ funzioni date di classe C^1 . Allora la funzione composta $f \circ g : A \rightarrow \mathbb{R}^k$ è di classe C^1 e il suo gradiente è*

$$\nabla(f \circ g)(x) = \nabla f(g(x))\nabla g(x).$$

La formula precedente va interpretata come un prodotto (riga per colonna) tra matrici. La prima matrice $\nabla f(g(x))$ ha dimensione $k \times m$, mentre $\nabla g(x)$ ha dimensione $m \times n$, e quindi il risultato sarà una matrice $k \times n$. Ricordiamo, denotando $h = f \circ g$, che per definizione di gradiente

$$\nabla h(x) = \begin{pmatrix} \frac{\partial h_1}{\partial x_1}(x) & \frac{\partial h_1}{\partial x_2}(x) & \dots & \frac{\partial h_1}{\partial x_n}(x) \\ \frac{\partial h_2}{\partial x_1}(x) & \frac{\partial h_2}{\partial x_2}(x) & \dots & \frac{\partial h_2}{\partial x_n}(x) \\ \dots & \dots & \dots & \dots \\ \frac{\partial h_k}{\partial x_1}(x) & \frac{\partial h_k}{\partial x_2}(x) & \dots & \frac{\partial h_k}{\partial x_n}(x) \end{pmatrix} \quad (2.1.14)$$

e inoltre, denotando con il simbolo y la variabile in $\mathbb{R}^m$,

$$\nabla f(g(x)) = \begin{pmatrix} \frac{\partial f_1}{\partial y_1}(g(x)) & \frac{\partial f_1}{\partial y_2}(g(x)) & \dots & \frac{\partial f_1}{\partial y_m}(g(x)) \\ \frac{\partial f_2}{\partial y_1}(g(x)) & \frac{\partial f_2}{\partial y_2}(g(x)) & \dots & \frac{\partial f_2}{\partial y_m}(g(x)) \\ \dots & \dots & \dots & \dots \\ \frac{\partial f_k}{\partial y_1}(g(x)) & \frac{\partial f_k}{\partial y_2}(g(x)) & \dots & \frac{\partial f_k}{\partial y_m}(g(x)) \end{pmatrix} \quad (2.1.15)$$

e

$$\nabla g(x) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(x) & \frac{\partial g_1}{\partial x_2}(x) & \dots & \frac{\partial g_1}{\partial x_n}(x) \\ \frac{\partial g_2}{\partial x_1}(x) & \frac{\partial g_2}{\partial x_2}(x) & \dots & \frac{\partial g_2}{\partial x_n}(x) \\ \dots & \dots & \dots & \dots \\ \frac{\partial g_k}{\partial x_1}(x) & \frac{\partial g_k}{\partial x_2}(x) & \dots & \frac{\partial g_k}{\partial x_n}(x) \end{pmatrix}. \quad (2.1.16)$$

Quindi, svolgendo il prodotto riga per colonna delle ultime due matrici, la tesi del teorema afferma che

$$(\nabla h(x))_{ij} = (\nabla(g \circ f)(x))_{ij} = \sum_{k=1}^m \frac{\partial f_i}{\partial y_k}(g(x)) \frac{\partial g_k}{\partial x_j}(x), \quad (2.1.17)$$

per ogni $i = 1, \dots, k$ e $j = 1, \dots, n$.

Dimostrazione. Fissiamo $i \in \{1, \dots, k\}$ e iniziamo a calcolare le derivate direzionali della componente $(f \circ g)_i$. Per comodità, come prima denotiamo $h = f \circ g$, e fissiamo $v \in \mathbb{R}^n$ un vettore non nullo. Per ogni numero reale $t > 0$ scriviamo il rapporto incrementale

$$\frac{h_i(x + tv) - h_i(x)}{t} = \frac{(f \circ g)_i(x + tv) - (f \circ g)_i(x)}{t} = \frac{f_i(g(x + tv)) - f_i(g(x))}{t}.$$

Ponendo $w := \frac{g(x+tv)-g(x)}{s}$ e $s := \|g(x + tv) - g(x)\|$, riscriviamo l'ultima espressione come

$$\frac{f_i(g(x) + g(x + tv) - g(x)) - f_i(g(x))}{t} = \frac{f_i(g(x) + sw) - f_i(g(x))}{t}$$

ed ora applichiamo la formula (2.1.13) che ci fornisce di un numero $\xi \in (0, s)$ tale che

$$\frac{f_i(g(x) + sw) - f_i(g(x))}{t} = \frac{s}{t} \frac{\partial f_i}{\partial w}(g(x) + \xi w) = \frac{s}{t} \nabla f_i(g(x) + \xi w) w.$$

Pertanto, abbiamo ottenuto

$$\frac{h_i(x + tv) - h_i(x)}{t} = \nabla f_i(g(x) + \xi w) \frac{g(x + tv) - g(x)}{t}$$

ed ora mandando $t \rightarrow 0$ arriviamo a

$$\frac{\partial h_i}{\partial v}(x) = \frac{\partial(f \circ g)_i}{\partial v}(x) = \nabla f_i(g(x)) \frac{\partial g}{\partial v}(x) = \nabla f_i(g(x)) (\nabla g(x)v).$$

Nel caso particolare in cui $v = e_j$, avremo

$$\frac{\partial(f \circ g)_i}{\partial x_j}(x) = \nabla f_i(g(x)) \frac{\partial g}{\partial x_j}(x) = \sum_{k=1}^m \frac{\partial f_i}{\partial y_k}(g(x)) \frac{\partial g_k}{\partial x_j}(x). \quad (2.1.18)$$

Abbiamo provato che le derivate parziali di $f \circ g$ (e quindi le componenti del gradiente) hanno forma (2.1.17). Rimane da far vedere che $f \circ g$ è differenziabile. Usiamo allora il teorema del differenziale totale, che ci consente di dire che $f \circ g$ è differenziabile in $\bar{x}$ se le derivate parziali sono continue in un intorno di $\bar{x}$. Ma questo è evidente dalla (2.1.18) grazie all'ipotesi che f e g sono di classe C^1 . $\square$

Torniamo all'esempio 2.1.17, abbiamo verificato che se $g(t) := \bar{x} + tv$ ed $f : \mathbb{R}^m \rightarrow \mathbb{R}$ allora

$$(f \circ g)'(t) = \nabla f(g(t))g'(t) = \nabla f(\bar{x} + tv)v = \sum_{k=1}^n \frac{\partial f}{\partial x_k}(\bar{x} + tv)v_k.$$

Supponendo f di classe C^2 in un intorno di $\bar{x}$ possiamo calcolare la derivata seconda di $f \circ g$ in $t = 0$ che è

$$\begin{aligned} (f \circ g)''(0) &= \lim_{t \rightarrow 0} \frac{(f \circ g)'(t) - (f \circ g)'(0)}{t} = \lim_{t \rightarrow 0} \frac{(\nabla f(\bar{x} + tv) - \nabla f(\bar{x}))v}{t} \\ &= \lim_{t \rightarrow 0} \frac{\sum_{k=1}^n \left(\frac{\partial f}{\partial x_k}(\bar{x} + tv) - \frac{\partial f}{\partial x_k}(\bar{x}) \right) v_k}{t} = \sum_{k=1}^n \frac{\partial}{\partial v} \frac{\partial f}{\partial x_k}(\bar{x}) v_k \\ &= \sum_{k=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_k}(\bar{x}) v_j v_k = v(\nabla^2 f(\bar{x})v). \end{aligned} \quad (2.1.19)$$

2.1.5 Massimi, minimi e punti stazionari di funzioni di più variabili

Supponiamo $f : A \rightarrow \mathbb{R}$ sia una funzione definita su un aperto $A \subset \mathbb{R}^n$.

Definizione 2.1.19. Diremo che $\bar{x} \in A$ è un punto di minimo relativo se esiste una palla $B_r(\bar{x})$ tale che

$$f(\bar{x}) \leq f(x) \quad \forall x \in A \cap B_r(\bar{x}). \quad (2.1.20)$$

Similmente, diremo che è un punto di massimo relativo se esiste una palla $B_r(\bar{x})$ tale che

$$f(\bar{x}) \geq f(x) \quad \forall x \in A \cap B_r(\bar{x}). \quad (2.1.21)$$

Inoltre, se $\bar{x}$ è un punto di minimo relativo (rispettivamente massimo relativo), diremo che $\bar{x}$ è un minimo relativo stretto (rispettivamente, massimo relativo stretto) se la condizione in (2.1.20) (rispettivamente (2.1.21)) vale con disuguaglianza stretta ogni volta che $x \neq \bar{x}$.

Infine, diremo che $\bar{x}$ è un minimo assoluto per f se

$$f(\bar{x}) \leq f(x) \quad \forall x \in A, \quad (2.1.22)$$

e diremo che è un massimo assoluto per f se

$$f(\bar{x}) \geq f(x) \quad \forall x \in A. \quad (2.1.23)$$

Diremo che il minimo assoluto (rispettivamente massimo assoluto) è stretto se le disuguaglianze precedenti valgono strettamente se $x \neq \bar{x}$.

Supponiamo ora $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ sia di classe C^1 e che $\bar{x} \in A$ sia un punto di minimo relativo. Esiste allora una palla $B_\rho(\bar{x})$, intorno di $\bar{x}$, tale che $f(\bar{x}) \leq f(x)$ per ogni $x \in B_\rho(\bar{x})$. In particolare, per ogni $h \in (-\rho, \rho)$ e ogni $v \in \mathbb{R}^n$ con $\|v\| = 1$, si avrà

$$f(\bar{x} + hv) - f(\bar{x}) \geq 0.$$

Ciò implica che

$$\frac{\partial f}{\partial v}(\bar{x}) = \lim_{h \rightarrow 0^+} \frac{f(\bar{x} + hv) - f(\bar{x})}{h} \geq 0, \quad (2.1.24)$$

ma allo stesso tempo

$$\frac{\partial f}{\partial v}(\bar{x}) = \lim_{h \rightarrow 0^-} \frac{f(\bar{x} + hv) - f(\bar{x})}{h} \leq 0. \quad (2.1.25)$$

Quindi necessariamente $\frac{\partial f}{\partial v}(\bar{x}) = 0$. In particolare, essendo f di classe C^1 e quindi differenziabile in $\bar{x}$, risulterà

$$\nabla f(\bar{x}) = 0.$$

Lo stesso risultato sarebbe stato ottenuto se avessimo supposto $\bar{x}$ massimo relativo.

Definizione 2.1.20. Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$. Un punto $\bar{x} \in A$ si dice stazionario se $\nabla f(\bar{x}) = 0$.

Il ragionamento precedente porta al seguente teorema:

Teorema 2.1.21. Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ di classe C^1 e sia $\bar{x} \in A$. Se $\bar{x}$ è un punto di massimo o minimo relativo, allora $\bar{x}$ è un punto stazionario.

Supponiamo ora $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ di classe C^2 e andiamo a studiare cosa succede alla derivata seconda in un punto $\bar{x} \in A$ di minimo relativo. A tale scopo, come fatto prima, prendiamo $v \in \mathbb{R}^n$ con $\|v\| = 1$, e la funzione $g(t) := \bar{x} + tv$, $t \in (-\rho, \rho)$. La funzione composta $f \circ g : (-\rho, \rho) \rightarrow \mathbb{R}$ soddisferà

$$f \circ g(t) = f(\bar{x} + tv) \geq f(\bar{x}) = f \circ g(0), \quad \forall t \in (-\rho, \rho),$$

e quindi avrà un minimo relativo in $t = 0$. Di conseguenza la sua derivata seconda in $t = 0$ sarà necessariamente non negativa:

$$(f \circ g)''(0) = v(\nabla^2 f(\bar{x})v) = \sum_{k=1}^n \sum_{j=1}^n \frac{\partial^2 f}{\partial x_j \partial x_k}(\bar{x}) v_j v_k \geq 0, \quad (2.1.26)$$

dove abbiamo usato la formula (2.1.19).

Introduciamo allora la seguente definizione:

Definizione 2.1.22. Sia $M \in \mathbb{R}^{n \times n}$ una matrice $n \times n$. Allora diremo che M è definita positiva se per ogni $v \in \mathbb{R}^n$ con $\|v\| = 1$ si ha $v(Mv) > 0$, ovvero

$$\sum_{k=1}^n \sum_{j=1}^n M_{jk} v_j v_k > 0, \quad \forall v \in \mathbb{R}^n : \|v\| = 1.$$

Diremo che M è semidefinita positiva se la disuguaglianza precedente vale con il $\geq$ invece del maggiore stretto ed esiste v con $\|v\| = 1$ tale che $v(Mv) = 0$.

Similmente diremo che M è definita negativa se

$$\sum_{k=1}^n \sum_{j=1}^n M_{jk} v_j v_k < 0, \quad \forall v \in \mathbb{R}^n : \|v\| = 1,$$

e semidefinita negativa se la disuguaglianza precedente vale con il $\leq$ invece del minore stretto ed esiste v con $\|v\| = 1$ tale che $v(Mv) = 0$.

Osservazione 2.1.23. Notare che M è definita positiva se e solo se

$$\min\{v(Mv) : \|v\| = 1\} = \inf\{v(Mv) : \|v\| = 1\} > 0,$$

e similmente

$$\max\{v(Mv) : \|v\| = 1\} = \sup\{v(Mv) : \|v\| = 1\} < 0$$

se è definita negativa. Questo perchè l'insieme dei vettori v con $\|v\| = 1$ è un compatto di $\mathbb{R}^n$ e la funzione $v \mapsto v(Mv)$ è continua.

La discussione precedente porta allora al seguente teorema:

Teorema 2.1.24. Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ di classe C^2 e sia $\bar{x} \in A$ un punto di minimo relativo (massimo relativo, rispettivamente). Allora $\nabla^2 f(\bar{x})$ è definita o semidefinita positiva (negativa, rispettivamente).

Abbiamo quindi trovato una condizione necessaria perchè $\bar{x}$ sia un punto di massimo o minimo relativo. Servirà però trovare anche delle condizioni sufficienti. A tale scopo, introduciamo il seguente risultato:

Teorema 2.1.25. Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ di classe C^2 e sia $\bar{x} \in A$; allora esiste una funzione $R_{\bar{x}} : A \rightarrow \mathbb{R}$ tale che per ogni $x \in A$ si ha

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})(x - \bar{x}) + \frac{1}{2}(x - \bar{x}) \left(\nabla^2 f(\bar{x})(x - \bar{x}) \right) + R_{\bar{x}}(x), \quad (2.1.27)$$

e inoltre

$$\lim_{x \rightarrow \bar{x}} \frac{|R_{\bar{x}}(x)|}{\|x - \bar{x}\|^2} = 0. \quad (2.1.28)$$

Dimostrazione. Consideriamo un vettore $v \in \mathbb{R}^n$ con $\|v\| = 1$ e la funzione $g(t) = f(\bar{x} + tv)$, che è definita per t reali e in un intorno di 0. Essendo g la composizione di f con la funzione $t \mapsto \bar{x} + tv$ abbiamo già osservato che g risulterà di classe C^2 e di conseguenza ammette sviluppo di Taylor al secondo ordine in 0 della forma

$$g(t) = g(0) + g'(0)t + \frac{1}{2}g''(\tau)t^2,$$

dove $\tau \in (0, t)$ (se $t > 0$) o $\tau \in (t, 0)$ (se $t < 0$). La formula precedente può essere riscritta come

$$g(t) = g(0) + g'(0)t + \frac{1}{2}g''(0)t^2 + \left(\frac{1}{2}g''(\tau)t^2 - \frac{1}{2}g''(0)t^2\right),$$

e visto che $g'(0) = \nabla f(\bar{x})v$, $g''(0) = v(\nabla^2 f(\bar{x})v)$, e $g''(\tau) = v(\nabla^2 f(\bar{x} + \tau v)v)$, otteniamo

$$f(\bar{x} + tv) = f(\bar{x}) + \nabla f(\bar{x})vt + \frac{1}{2}v(\nabla^2 f(\bar{x})v)t^2 + \frac{1}{2}t^2\left(v(\nabla^2 f(\bar{x} + \tau v)v) - v(\nabla^2 f(\bar{x})v)\right).$$

Scegliendo $v := \frac{x - \bar{x}}{\|x - \bar{x}\|}$ e $t = \|x - \bar{x}\|$ l'espressione precedente diventa

$$f(x) = f(\bar{x}) + \nabla f(\bar{x})(x - \bar{x}) + \frac{1}{2}(x - \bar{x})(\nabla^2 f(\bar{x})(x - \bar{x})) + R_{\bar{x}}(x),$$

dove abbiamo definito

$$R_{\bar{x}}(x) := \frac{1}{2}\|x - \bar{x}\|^2 \frac{x - \bar{x}}{\|x - \bar{x}\|} \left((\nabla^2 f(\bar{x} + \tau \frac{x - \bar{x}}{\|x - \bar{x}\|}) - \nabla^2 f(\bar{x})) \frac{x - \bar{x}}{\|x - \bar{x}\|} \right),$$

e dove τ dipende da x e soddisfa $\tau \in (0, \|x - \bar{x}\|)$. Rimane da verificare la proprietà (2.1.28): dividendo l'espressione precedente per $\|x - \bar{x}\|^2$ otteniamo

$$\begin{aligned} \frac{|R_{\bar{x}}(x)|}{\|x - \bar{x}\|^2} &= \frac{1}{2} \left| \frac{x - \bar{x}}{\|x - \bar{x}\|} \left((\nabla^2 f(\bar{x} + \tau \frac{x - \bar{x}}{\|x - \bar{x}\|}) - \nabla^2 f(\bar{x})) \frac{x - \bar{x}}{\|x - \bar{x}\|} \right) \right| \\ &\leq \frac{1}{2} \left\| \frac{x - \bar{x}}{\|x - \bar{x}\|} \right\| \left\| (\nabla^2 f(\bar{x} + \tau \frac{x - \bar{x}}{\|x - \bar{x}\|}) - \nabla^2 f(\bar{x})) \frac{x - \bar{x}}{\|x - \bar{x}\|} \right\| \\ &\leq \frac{1}{2} \left\| \nabla^2 f(\bar{x} + \tau \frac{x - \bar{x}}{\|x - \bar{x}\|}) - \nabla^2 f(\bar{x}) \right\| \left\| \frac{x - \bar{x}}{\|x - \bar{x}\|} \right\| \\ &= \frac{1}{2} \left\| \nabla^2 f(\bar{x} + \tau \frac{x - \bar{x}}{\|x - \bar{x}\|}) - \nabla^2 f(\bar{x}) \right\|, \end{aligned}$$

e ora il membro di destra tende a zero quando $x \rightarrow \bar{x}$ perchè $\tau \rightarrow 0$ e quindi

$$\bar{x} + \tau \frac{x - \bar{x}}{\|x - \bar{x}\|} \rightarrow \bar{x}$$

e le derivate seconde sono continue in $\bar{x}$ (essendo f di classe C^2). □

Grazie a questo risultato possiamo dimostrare il seguente teorema:

Teorema 2.1.26. *Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ di classe C^2 e sia $\bar{x} \in A$ un punto stazionario (ovvero $\nabla f(\bar{x}) = 0$); se $\nabla^2 f(\bar{x})$ è definita negativa, allora $\bar{x}$ è un punto di massimo relativo stretto. Se invece $\nabla^2 f(\bar{x})$ è definita positiva, allora $\bar{x}$ è un punto di minimo relativo stretto.*

Dimostrazione. Mettiamoci nel caso in cui $\nabla^2 f(\bar{x})$ è definita negativa. Allora per il Teorema 2.1.25 abbiamo

$$f(x) - f(\bar{x}) = \nabla f(\bar{x})(x - \bar{x}) + \frac{1}{2}(x - \bar{x})(\nabla^2 f(\bar{x})(x - \bar{x})) + R_{\bar{x}}(x),$$

e visto che $\nabla f(\bar{x}) = 0$, dividendo per $\|x - \bar{x}\|^2$ otteniamo

$$\frac{f(x) - f(\bar{x})}{\|x - \bar{x}\|^2} = \frac{1}{2} \frac{x - \bar{x}}{\|x - \bar{x}\|} \left(\nabla^2 f(\bar{x}) \frac{x - \bar{x}}{\|x - \bar{x}\|} \right) + \frac{R_{\bar{x}}(x)}{\|x - \bar{x}\|^2}.$$

Grazie all'Osservazione 2.1.23, si ha

$$\frac{1}{2} \frac{x - \bar{x}}{\|x - \bar{x}\|} \left(\nabla^2 f(\bar{x}) \frac{x - \bar{x}}{\|x - \bar{x}\|} \right) \leq \frac{m}{2} < 0,$$

dove $m = \sup\{v(\nabla^2 f(\bar{x})v) : \|v\| = 1\}$. Inoltre, grazie alla formula (2.1.28), esiste $\delta > 0$ tale che se $\|x - \bar{x}\| < \delta$ allora $\frac{|R_{\bar{x}}(x)|}{\|x - \bar{x}\|^2} < \frac{|m|}{4}$, e di conseguenza dall'espressione precedente deduciamo (essendo $m < 0$)

$$\frac{f(x) - f(\bar{x})}{\|x - \bar{x}\|^2} \leq \frac{m}{2} + \frac{|m|}{4} = \frac{m}{4} < 0,$$

per ogni $x \in B_\delta(\bar{x})$. Questo implica che $f(x) - f(\bar{x}) < 0$, per ogni $x \in B_\delta(\bar{x})$, $x \neq \bar{x}$, e quindi $\bar{x}$ è un massimo relativo stretto. $\square$

Capitolo 3

Integrazione

La teoria dell'integrale di Riemann è valida in $\mathbb{R}^n$ per ogni $n \geq 1$. Tuttavia, per semplificare le notazioni, svolgeremo la maggior parte della nostra discussione nel caso $n = 2$.

3.1 Integrale di Riemann

Definizione 3.1.1. Un rettangolo $R \subset \mathbb{R}^2$ è un insieme della forma $[a, b) \times [c, d)$ dove $a, b, c, d \in \mathbb{R}$ sono numeri con $a < b$, $c < d$. Un plurirettangolo $P \subset \mathbb{R}^2$ è un qualsiasi insieme che può essere scritto come unione finita di rettangoli disgiunti. Ovvero P è un plurirettangolo se esistono rettangoli R_i , $i = 1, \dots, k$ a due a due disgiunti, tali che $P = \cup_{i=1}^k R_i$.

Dato un rettangolo $R = [a, b) \times [c, d)$, chiameremo partizione di R una sua suddivisione in rettangoli più piccoli. Ovvero, una partizione di R si identifica con due sequenze di numeri reali $a = x_0 < x_1 < \dots < x_n = b$ e $c = y_0 < y_1 < \dots < y_m = d$, e denoteremo con

$$R_{ij} := [x_{i-1}, x_i) \times [y_{j-1}, y_j), \quad i \in \{1, \dots, n\}, j \in \{1, \dots, m\}$$

i rettangolini la cui unione è R , ossia

$$R = \bigcup_{i=1}^n \bigcup_{j=1}^m R_{ij}.$$

In particolare, R può essere visto come un plurirettangolo.

Definizione 3.1.2. Sia $E \subseteq \mathbb{R}^2$ un insieme, denoteremo con $\chi_E : \mathbb{R}^2 \rightarrow \mathbb{R}$ la funzione caratteristica dell'insieme E , definita come

$$\chi_E(x) = \begin{cases} 1 & \text{se } x \in E, \\ 0 & \text{se } x \notin E. \end{cases} \quad (3.1.1)$$

In particolare χ_E assume solo i valori 0 e 1.

Definizione 3.1.3. Sia $R \subset \mathbb{R}^2$ un rettangolo fissato. Una funzione $f : R \rightarrow \mathbb{R}$ si dice semplice se esiste una partizione di R in rettangolini $R = \bigcup_{i=1}^n \bigcup_{j=1}^m R_{ij}$ tale che f è costante su ogni rettangolino R_{ij} . In altre parole, f è semplice se esistono numeri reali λ_{ij} , $i = 1, \dots, n$, $j = 1, \dots, m$ tali che

$$f(x) = \lambda_{ij} \quad \text{se } x \in R_{ij}, \quad i \in \{1, \dots, n\}, j \in \{1, \dots, m\}.$$

La formula sopra può essere riscritta in forma compatta nel seguente modo

$$f = \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij} \chi_{R_{ij}}. \quad (3.1.2)$$

Per ogni rettangolo $R = [a, b] \times [c, d]$ possiamo definire la sua area come $m(R) = (b - a)(d - c)$. La quantità $m(R)$ prende il nome di misura di Peano-Jordan di R . Vogliamo ora definire l'integrale di una funzione semplice.

Definizione 3.1.4. Sia $R = [a, b] \times [c, d]$ un rettangolo e sia $f : R \rightarrow \mathbb{R}$ una funzione semplice della forma (3.1.2). Allora si definisce l'integrale di f su R la quantità

$$\int_R f(x, y) dx dy := \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij} m(R_{ij}) = \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij} (x_i - x_{i-1})(y_j - y_{j-1}). \quad (3.1.3)$$

Vogliamo ora definire l'integrale di funzioni generiche. A questo scopo introduciamo la seguente definizione:

Definizione 3.1.5. Sia $R = [a, b] \times [c, d]$ un rettangolo e sia $f : R \rightarrow \mathbb{R}$ una funzione limitata. Definiamo le seguenti due classi di funzioni

$$\mathcal{F}^+ := \{\varphi : R \rightarrow \mathbb{R} : \varphi \text{ è semplice e } \varphi \geq f \text{ su } R\}, \quad (3.1.4)$$

$$\mathcal{F}^- := \{\psi : R \rightarrow \mathbb{R} : \psi \text{ è semplice e } \psi \leq f \text{ su } R\}. \quad (3.1.5)$$

Il fatto che f è una funzione limitata garantisce che le classi $\mathcal{F}^+$ e $\mathcal{F}^-$ sono non vuote. Osserviamo inoltre che per ogni funzione $\varphi \in \mathcal{F}^+$ e ogni $\psi \in \mathcal{F}^-$ si ha $\varphi \geq \psi$ e

$$\int_R \psi(x, y) dx dy \leq \int_R \varphi(x, y) dx dy. \quad (3.1.6)$$

Definizione 3.1.6. Sia $R = [a, b] \times [c, d]$ un rettangolo e sia $f : R \rightarrow \mathbb{R}$ una funzione limitata. Definiamo l'integrale superiore di f su R il valore

$$\int^* f(x, y) dx dy := \inf \left\{ \int_R \varphi(x, y) dx dy : \varphi \in \mathcal{F}^+ \right\}. \quad (3.1.7)$$

Similmente l'integrale inferiore di f su R è definito come

$$\int_* f(x, y) dx dy := \sup \left\{ \int_R \psi(x, y) dx dy : \psi \in \mathcal{F}^- \right\}. \quad (3.1.8)$$

Dalla formula (3.1.4) segue che, per ogni $f : R \rightarrow \mathbb{R}$ limitata, si ha

$$\int_* f(x, y) dx dy \leq \int^* f(x, y) dx dy.$$

Inoltre dalla definizione segue che se $\varphi \in \mathcal{F}^+$ e $\psi \in \mathcal{F}^-$ allora

$$\int_R \psi(x, y) dx dy \leq \int_* f(x, y) dx dy \leq \int^* f(x, y) dx dy \leq \int_R \varphi(x, y) dx dy. \quad (3.1.9)$$

Definizione 3.1.7. Diremo che $f : R \rightarrow \mathbb{R}$ limitata è integrabile secondo Riemann su R se

$$\int_* f(x, y) dx dy = \int^* f(x, y) dx dy. \quad (3.1.10)$$

In tal caso si definisce l'integrale di f su R come

$$\int_R f(x, y) dx dy := \int_* f(x, y) dx dy = \int^* f(x, y) dx dy. \quad (3.1.11)$$

Lemma 3.1.8. Sia $f : R \rightarrow \mathbb{R}$ una funzione limitata. Allora i due fatti seguenti sono equivalenti:

- f è integrabile secondo Riemann su R ;
- per ogni $\varepsilon > 0$ esistono due funzioni semplici $\varphi \in \mathcal{F}^+$ e $\psi \in \mathcal{F}^-$ tali che

$$\int_R \varphi(x, y) dx dy - \int_R \psi(x, y) dx dy \leq \varepsilon. \quad (3.1.12)$$

Dimostrazione. Se f è integrabile secondo Riemann allora, per definizione, $\int_* f(x, y) dx dy = \int^* f(x, y) dx dy$. Essendo l'integrale inferiore un estremo superiore, avremo che per ogni $\varepsilon > 0$ esiste una funzione semplice $\psi \in \mathcal{F}^-$ tale che

$$\int_R \psi(x, y) dx dy \geq \int_* f(x, y) dx dy - \frac{\varepsilon}{2}.$$

Similmente per l'integrale superiore, esisterà $\varphi \in \mathcal{F}^+$ tale che

$$\int_R \varphi(x, y) dx dy \leq \int^* f(x, y) dx dy + \frac{\varepsilon}{2}.$$

Sottraendo l'ultima espressione alla precedente, abbiamo trovato $\varphi \in \mathcal{F}^+$ e $\psi \in \mathcal{F}^-$ tali che (3.1.12) vale.

Viceversa, supponiamo per ogni $\varepsilon > 0$ esistano $\varphi \in \mathcal{F}^+$ e $\psi \in \mathcal{F}^-$ tali che (3.1.12) valga. Grazie alla (3.1.9), troviamo immediatamente

$$\int^* f(x, y) dx dy - \int_* f(x, y) dx dy \leq \int_R \varphi(x, y) dx dy - \int_R \psi(x, y) dx dy \leq \varepsilon.$$

Pertanto, $\int^* f(x, y) dx dy - \int_* f(x, y) dx dy \leq \varepsilon$ per ogni $\varepsilon > 0$, e quindi necessariamente $\int^* f(x, y) dx dy - \int_* f(x, y) dx dy = 0$, ovvero f è integrabile. $\square$

3.2 La misura di Peano-Jordan e alcune funzioni integrabili

Definizione 3.2.1. Sia $f : \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione data. Chiameremo supporto di f l'insieme

$$\text{supp}(f) := \overline{\{x \in \mathbb{R}^n : f(x) \neq 0\}}.$$

Diremo che f è a supporto compatto se $\text{supp}(f)$ risulta essere un insieme compatto in $\mathbb{R}^n$.

Visto che il supporto di una funzione è sempre chiuso (per definizione), esso risulterà compatto se e solo se risulta limitato.

Ad esempio, se $E \subset \mathbb{R}^n$ è un insieme limitato, allora χ_E è una funzione a supporto compatto. Se una funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ è a supporto compatto, allora esiste un rettangolo R che contiene il suo supporto,

$$\text{supp}(f) \subset R.$$

Definizione 3.2.2. Sia $E \subset \mathbb{R}^2$ un insieme limitato, e sia R un rettangolo tale che $E \subset R$. Diremo che E è misurabile secondo Peano-Jordan se la funzione $\chi_E : R \rightarrow \mathbb{R}$ risulta integrabile secondo Riemann su R . Il suo integrale sarà anche denotato con $m(E)$ che prende il nome di misura di Peano-Jordan di E :

$$m(E) := \int_R \chi_E(x, y) dx dy.$$

Non è difficile verificare che la definizione appena data è indipendente dalla scelta del rettangolo R che contiene E . In particolare, la misura di Peano-Jordan $m(E)$ di E non dipende dalla scelta del rettangolo R su cui si integra χ_E . Lasciamo questa verifica per esercizio.

Supponiamo E sia un insieme limitato misurabile secondo Peano-Jordan. Allora, se R è un rettangolo contenente E e denotiamo con $f = \chi_E$, sappiamo che per ogni $\varepsilon > 0$ esistono funzioni semplici $\varphi \in \mathcal{F}^+$ e $\psi \in \mathcal{F}^-$ tali che (3.1.12) vale. Siano R_{ij} rettangolini che formano una partizione di R tali che

$$\varphi = \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij} \chi_{R_{ij}}, \quad \psi = \sum_{i=1}^n \sum_{j=1}^m \mu_{ij} \chi_{R_{ij}}.$$

per alcuni valori di λ_{ij} e μ_{ij} reali. Ora, se $R_{ij} \subset E$, allora, visto che $\lambda_{ij} \geq \chi_E \geq \mu_{ij}$, si ha $\lambda_{ij} \geq 1$ e $\mu_{ij} \leq 1$. A meno di diminuire il valore di φ e di aumentare il valore di ψ , possiamo supporre che $\lambda_{ij} = 1 = \mu_{ij}$ in questo rettangolino R_{ij} , e (3.1.12) vale ancora. Analogamente se $R_{ij} \cap E = \emptyset$ possiamo invece supporre che $\lambda_{ij} = 0 = \mu_{ij}$. E invece, se $R_{ij} \cap E \neq \emptyset$ e allo stesso tempo $R_{ij} \setminus E \neq \emptyset$, avremo necessariamente $\lambda_{ij} \geq 1$ e $\mu_{ij} \leq 0$. Supponendo uguaglianze in questo caso, avremo che φ e ψ coincidono su tutti i rettangolini R_{ij} tranne in quelli che intersecano sia E che $R_{ij} \setminus E$, e in tal caso $\lambda_{ij} = 1 > 0 = \mu_{ij}$. Questi ultimi rettangolini contengono quelli che intersecano ∂E . Denotando l'insieme di questi ultimi rettangolini con $\mathcal{R}_{\partial E}$ avremo

$$\int_R \varphi(x, y) dx dy - \int_R \psi(x, y) dx dy = \sum_{R_{ij} \in \mathcal{R}_{\partial E}} m(R_{ij}) \leq \varepsilon. \quad (3.2.1)$$

Ora, i rettangolini in $\mathcal{R}_{\partial E}$ sono tutti e soli quelli che ricoprono ∂E ; pertanto, abbiamo scoperto che se E è un insieme limitato misurabile secondo Peano-Jordan, per ogni $\varepsilon > 0$ possiamo trovare un ricoprimento di ∂E composto da rettangolini (quindi un plurirettangolo che contiene ∂E) la cui area è minore o uguale a ε . Quanto osservato implica che la funzione caratteristica $\chi_{\partial E}$ è integrabile secondo Riemann e il suo integrale è nullo: ovvero, abbiamo ottenuto il seguente lemma:

Lemma 3.2.3. *Sia $E \subset \mathbb{R}^2$ un insieme limitato misurabile secondo Peano-Jordan; allora ∂E è un insieme misurabile secondo Peano-Jordan e risulta $m(\partial E) = 0$.*

Sia ora $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ una funzione. Studiamo ora alcune condizioni sufficienti affinché f risulti integrabile:

Teorema 3.2.4. *Sia $R \subset \mathbb{R}^2$ un rettangolo e sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ una funzione continua su $\overline{R}$. Allora f risulta integrabile su R .*

Dimostrazione. La dimostrazione utilizza il fatto che una funzione continua su un compatto è anche uniformemente continua.

Essendo uniformemente continua, sappiamo che per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che, se $v, w \in \overline{R}$ soddisfano $\|w - v\| \leq \delta$, allora $|f(v) - f(w)| \leq \varepsilon$. Per ogni $\varepsilon > 0$ allora scegliamo una partizione R_{ij} di R tale che il diametro di R_{ij} sia minore di δ , per ogni i, j . Inoltre, la funzione $f : \overline{R}_{ij} \rightarrow \mathbb{R}$ è a sua volta continua, e quindi ammette massimo e minimo assoluto su $\overline{R}_{ij}$, che denoteremo con λ_{ij}^+ e λ_{ij}^- , rispettivamente. Per quanto detto prima, risulterà $\lambda_{ij}^+ - \lambda_{ij}^- \leq \varepsilon$. Di conseguenza, se definiamo le due funzioni semplici $\varphi, \psi : R \rightarrow \mathbb{R}$ come

$$\varphi := \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij}^+ \chi_{R_{ij}} \quad \psi := \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij}^- \chi_{R_{ij}},$$

avremo che $\psi \leq f \leq \varphi$ su R , e gli integrali soddisfano

$$\int_R \varphi dx dy = \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij}^+ m(R_{ij}) \quad \int_R \psi dx dy = \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij}^- m(R_{ij}).$$

Sottraendo le due espressioni precedenti, abbiamo

$$\int_R \varphi dx dy - \int_R \psi dx dy = \sum_{i=1}^n \sum_{j=1}^m (\lambda_{ij}^+ - \lambda_{ij}^-) m(R_{ij}) \leq \sum_{i=1}^n \sum_{j=1}^m \varepsilon m(R_{ij}) = \varepsilon m(R). \quad (3.2.2)$$

Pertanto, per ogni $\varepsilon > 0$ abbiamo trovato che esistono due funzioni semplici $\varphi \in \mathcal{F}^+$ e $\psi \in \mathcal{F}^-$ tali che vale (3.2.2). Essendo $m(R)$ una costante, la condizione (3.1.12) è soddisfatta, e il teorema è dimostrato. $\square$

Teorema 3.2.5. *Sia $f : R \rightarrow \mathbb{R}$ una funzione limitata integrabile, e sia $E \subset R$ un insieme misurabile secondo Peano-Jordan. Allora la funzione $f \chi_E$ risulta integrabile su R .*

Dimostrazione. Essendo f limitata esiste una costante $M > 0$ tale che $|f| \leq M$ su $R = [a, b) \times [c, d)$. Essendo f integrabile su R per ogni $\varepsilon > 0$ esistono funzioni semplici $\varphi = \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij} \chi_{R_{ij}}$ e $\psi = \sum_{i=1}^n \sum_{j=1}^m \mu_{ij} \chi_{R_{ij}}$ su una partizione comune $a = x_0 < x_1 < \dots < x_n = b$ e $c = y_0 < y_1 < \dots < y_m = d$, tali che (3.1.12) vale. Inoltre E essendo misurabile, possiamo applicare il Lemma 3.2.3 e trovare una partizione $a = x'_0 < x'_1 < \dots < x'_{n'} = b$ e $c = y'_0 < y'_1 < \dots < y'_{m'} = d$ tale che, denotando con $R'_{ij} := [x'_{i-1}, x'_i) \times [y'_{j-1}, y'_j)$ un suo rettangolino e con $\mathcal{R}_{\partial E}$ la famiglia contenente i rettangolini R'_{ij} che intersecano ∂E , si ha

$$\sum_{R'_{ij} \in \mathcal{R}_{\partial E}} m(R'_{ij}) \leq \varepsilon. \quad (3.2.3)$$

Consideriamo gli insiemi di punti $X = \{x_0, x_1, \dots, x_n\} \cup \{x'_0, x'_1, \dots, x'_{n'}\}$ e $Y = \{y_0, y_1, \dots, y_m\} \cup \{y'_0, y'_1, \dots, y'_{m'}\}$, possiamo ordinare in ordine crescente gli elementi di X e Y e denotarli con

$X = \{\hat{x}_0, \hat{x}_1, \dots, \hat{x}_N\}$ e $Y = \{\hat{y}_0, \hat{y}_1, \dots, \hat{y}_M\}$, in modo che essi formano una partizione di R . Denotiamo anche con $\hat{R}_{ij}$ il generico rettangolino $[\hat{x}_{i-1}, \hat{x}_i) \times [\hat{y}_{j-1}, \hat{y}_j)$, per $i = 1, \dots, N$, $j = 1, \dots, M$.

Denotiamo con $\hat{\mathcal{R}}_{\partial E}$ la famiglia di rettangolini $\hat{R}_{ij}$ che intersecano ∂E . Ora, per come abbiamo costruito la partizione X e Y è facile vedere che (esercizio)

$$\bigcup_{\hat{R}_{ij} \in \hat{\mathcal{R}}_{\partial E}} \hat{R}_{ij} \subseteq \bigcup_{R'_{ij} \in \mathcal{R}_{\partial E}} R'_{ij},$$

e quindi dalla (3.2.3) si deduce che

$$\sum_{\hat{R}_{ij} \in \hat{\mathcal{R}}_{\partial E}} m(\hat{R}_{ij}) \leq \varepsilon. \quad (3.2.4)$$

Adesso modifichiamo le funzioni semplici φ e ψ nel seguente modo: anzitutto notiamo che esse possono essere scritte nella forma

$$\varphi = \sum_{i=1}^N \sum_{j=1}^M \tilde{\lambda}_{ij} \chi_{\hat{R}_{ij}}, \quad \psi = \sum_{i=1}^N \sum_{j=1}^M \tilde{\mu}_{ij} \chi_{\hat{R}_{ij}},$$

dove abbiamo definito $\tilde{\lambda}_{ij} = \lambda_{kh}$ se $\hat{R}_{ij} \subseteq R_{kh}$, e similmente $\tilde{\mu}_{ij} = \mu_{kh}$ se $\hat{R}_{ij} \subseteq R_{kh}$. Detto ciò, definiamo

$$\hat{\varphi} := \sum_{i=1}^N \sum_{j=1}^M \hat{\lambda}_{ij} \chi_{\hat{R}_{ij}}, \quad \hat{\psi} = \sum_{i=1}^N \sum_{j=1}^M \hat{\mu}_{ij} \chi_{\hat{R}_{ij}},$$

dove

$$\hat{\lambda}_{ij} = \begin{cases} \tilde{\lambda}_{ij} & \text{se } \hat{R}_{ij} \subseteq E, \\ 0 & \text{se } \hat{R}_{ij} \subseteq (R \setminus E), \\ M & \text{se } \hat{R}_{ij} \in \hat{\mathcal{R}}_{\partial E}, \end{cases} \quad \hat{\mu}_{ij} = \begin{cases} \tilde{\mu}_{ij} & \text{se } \hat{R}_{ij} \subseteq E, \\ 0 & \text{se } \hat{R}_{ij} \subseteq (R \setminus E), \\ -M & \text{se } \hat{R}_{ij} \in \hat{\mathcal{R}}_{\partial E}. \end{cases}$$

In tal modo è immediato verificare che $\hat{\psi} \leq f \chi_E \leq \hat{\varphi}$ in E . Inoltre,

$$\begin{aligned} \int_R \hat{\varphi} dx dy - \int_R \hat{\psi} dx dy &= \sum_{i=1}^N \sum_{j=1}^M (\hat{\lambda}_{ij} - \hat{\mu}_{ij}) m(\hat{R}_{ij}) = \\ &= \sum_{\hat{R}_{ij} \in \hat{\mathcal{R}}_{\partial E}} (\hat{\lambda}_{ij} - \hat{\mu}_{ij}) m(\hat{R}_{ij}) + \sum_{\hat{R}_{ij} \subseteq E} (\hat{\lambda}_{ij} - \hat{\mu}_{ij}) m(\hat{R}_{ij}) \\ &= 2M \sum_{\hat{R}_{ij} \in \hat{\mathcal{R}}_{\partial E}} m(\hat{R}_{ij}) + \sum_{\hat{R}_{ij} \subseteq E} (\tilde{\lambda}_{ij} - \tilde{\mu}_{ij}) m(\hat{R}_{ij}) \\ &\leq 2M\varepsilon + \sum_{i=1}^N \sum_{j=1}^M (\tilde{\lambda}_{ij} - \tilde{\mu}_{ij}) m(\hat{R}_{ij}) \end{aligned} \quad (3.2.5)$$

$$= 2M\varepsilon + \int_R \varphi dx dy - \int_R \psi dx dy \leq (2M+1)\varepsilon, \quad (3.2.6)$$

dove abbiamo usato la (3.2.4) e la (3.1.12). Di conseguenza, per ogni ε abbiamo trovato due funzioni semplici tali che

$$\int_R \hat{\varphi} dx dy - \int_R \hat{\psi} dx dy \leq (2M+1)\varepsilon,$$

ed essendo $2M + 1$ una costante indipendente da ε , questa condizione è equivalente alla (3.1.12), ovvero $f\chi_E$ è integrabile. $\square$

Definizione 3.2.6. Se $E \subset R$ è un insieme misurabile secondo Peano-Jordan in un rettangolo R e $f : R \rightarrow \mathbb{R}$ una funzione tale che $f\chi_E$ è integrabile. Allora f si dirà integrabile su E e l'integrale di f su E si definisce come

$$\int_E f dx dy := \int_R f\chi_E dx dy.$$

Dai teoremi 3.2.4 e 3.2.5 seguono immediatamente i seguenti corollari:

Corollario 3.2.7. Sia $f : \bar{R} \rightarrow \mathbb{R}$ una funzione continua su un rettangolo chiuso $\bar{R}$, allora per ogni insieme misurabile $E \subset R$ risulta che f è integrabile su E .

Corollario 3.2.8. Sia $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ una funzione continua a supporto compatto e sia E un insieme limitato misurabile secondo Peano-Jordan. Allora f è integrabile su E .

3.2.1 Insiemi normali

Analizziamo ora alcuni casi particolari di insiemi misurabili secondo Peano-Jordan.

Definizione 3.2.9. Un insieme $E \subset \mathbb{R}^n$ si dice normale rispetto alla variabile x_i se esiste un insieme $E' \subset \mathbb{R}^{n-1}$ misurabile secondo Peano-Jordan e due funzioni limitate e integrabili $f, g : E' \rightarrow \mathbb{R}$ tali che $g(x') < f(x')$ per ogni $x' \in E'$ e

$$E = \{x \in \mathbb{R}^n : g(x') \leq x_i \leq f(x')\},$$

dove abbiamo denotato $x' = (x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_n) \in \mathbb{R}^{n-1}$.

In $\mathbb{R}^2$ ad esempio, dato un intervallo $E' = (a, b)$ e due funzioni integrabili $f, g : (a, b) \rightarrow \mathbb{R}$ limitate e tali che $g(x) < f(x)$ per ogni $x \in (a, b)$, il dominio $E = \{(x, y) : x \in (a, b), g(x) \leq y \leq f(x)\}$ si dirà normale rispetto l'asse y . Viceversa, il dominio $E = \{(x, y) : y \in (c, d), g(y) \leq x \leq f(y)\}$ si dirà normale rispetto l'asse x , ogni volta che $f, g : (c, d) \rightarrow \mathbb{R}$ sono come sopra.

Mostriamo ora che tali insiemi sono misurabili secondo Peano-Jordan. Dimostreremo tale proprietà nel caso di domini bidimensionali, anche se una piccola modifica della dimostrazione garantisce che il risultato sia valido in $\mathbb{R}^n$.

Lemma 3.2.10. Sia $E = \{(x, y) : x \in (a, b), g(x) \leq y \leq f(x)\}$ un dominio normale. Allora E è misurabile secondo Peano-Jordan.

Dimostrazione. Per definizione le funzioni f e g sono integrabili su (a, b) e quindi per ogni $\varepsilon > 0$ esistono funzioni semplici $\varphi_1, \varphi_2, \psi_1, \psi_2 : (a, b) \rightarrow \mathbb{R}$ tali che $\psi_1 \leq f \leq \varphi_1$ e $\psi_2 \leq g \leq \varphi_2$ su (a, b) e

$$\int_a^b (\varphi_1 - \psi_1) dx \leq \varepsilon, \quad \int_a^b (\varphi_2 - \psi_2) dx \leq \varepsilon.$$

Essendo $f \geq g$ possiamo anche supporre che $\varphi_2 \leq \psi_1$ su (a, b) . Inoltre, possiamo supporre che le funzioni $\varphi_1, \varphi_2, \psi_1, \psi_2$ sono definite su una medesima partizione $a = x_0 < x_1 < \dots < x_n = b$

(perchè?). In tal caso risulterà

$$\begin{aligned}\varphi_1 &= \sum_{i=1}^n \lambda_i^+ \chi_{[x_{i-1}, x_i]}, & \psi_1 &= \sum_{i=1}^n \lambda_i^- \chi_{[x_{i-1}, x_i]}, \\ \varphi_2 &= \sum_{i=1}^n \mu_i^+ \chi_{[x_{i-1}, x_i]}, & \psi_2 &= \sum_{i=1}^n \mu_i^- \chi_{[x_{i-1}, x_i]}.\end{aligned}$$

Inoltre, essendo f e g limitate, possiamo supporre che $\max_{i=1, \dots, n} \lambda_i^+ = \sup_{x \in (a, b)} \{f(x)\}$, e $\min_{i=1, \dots, n} \mu_i^- = \inf_{x \in (a, b)} \{g(x)\}$. Con questa notazione a disposizione, troviamo

$$\int_a^b (\varphi_1 - \psi_1) dx = \sum_{i=1}^n (\lambda_i^+ - \lambda_i^-)(x_i - x_{i-1}) \leq \varepsilon, \quad (3.2.7)$$

e similmente

$$\int_a^b (\varphi_2 - \psi_2) dx = \sum_{i=1}^n (\mu_i^+ - \mu_i^-)(x_i - x_{i-1}) \leq \varepsilon. \quad (3.2.8)$$

Denotato con $M > 0$ una costante tale che $-M \leq g \leq f \leq M$ su (a, b) , prendiamo ora un rettangolo $R := [a, b] \times [-M, M]$, in modo che $E \subseteq R$. Costruiamo due funzioni semplici $\Psi, \Phi : R \rightarrow \mathbb{R}$ tali che $\Psi \leq \chi_E \leq \Phi$ su R e $\int_R (\Phi - \Psi) dx dy$ è sufficientemente piccolo. Definiamo dunque

$$\Psi := \sum_{i=1}^n \chi_{[x_{i-1}, x_i] \times [\mu_i^+, \lambda_i^-]}, \quad \Phi := \sum_{i=1}^n \chi_{[x_{i-1}, x_i] \times [\mu_i^-, \lambda_i^+]}.$$

In questo modo, essendo $[x_{i-1}, x_i] \times [\mu_i^+, \lambda_i^-] \subseteq [x_{i-1}, x_i] \times [\mu_i^-, \lambda_i^+]$ si vede subito che $\Psi \leq \Phi$. Inoltre, visto che $\varphi_1 \geq f \geq g \geq \psi_2$ risulterà

$$\bigcup_{i=1}^n [x_{i-1}, x_i] \times [\mu_i^+, \lambda_i^-] \subseteq E,$$

e similmente

$$E \subseteq \bigcup_{i=1}^n [x_{i-1}, x_i] \times [\mu_i^-, \lambda_i^+],$$

così da avere $\Psi \leq \chi_E \leq \Phi$. Calcoliamo ora gli integrali, ottenendo

$$\begin{aligned}\int_R (\Phi - \Psi) dx dy &= \sum_{i=1}^n (\lambda_i^+ - \mu_i^- - \lambda_i^- + \mu_i^+)(x_i - x_{i-1}) \\ &= \sum_{i=1}^n \left((\lambda_i^+ - \lambda_i^-) + (\mu_i^+ - \mu_i^-) \right) (x_i - x_{i-1}) \leq 2\varepsilon.\end{aligned}$$

Questo mostra che χ_E è integrabile secondo Riemann su R e quindi E è misurabile secondo Peano-Jordan. $\square$

3.3 Teorema di Fubini e cambio di variabili

Vediamo ora un metodo per calcolare integrali multipli. Mettiamoci prima nel caso planare, $n = 2$.

Teorema 3.3.1 (Fubini in dimensione 2). *Sia $R := [a, b] \times [c, d]$ un rettangolo e sia $f : R \rightarrow \mathbb{R}$ una funzione limitata integrabile. Supponiamo che per ogni $x \in [a, b]$ la funzione $f_x : [c, d] \rightarrow \mathbb{R}$ definita come $f_x(y) := f(x, y)$ sia integrabile su $[c, d]$. Allora la funzione $F : [a, b] \rightarrow \mathbb{R}$ definita da $F(x) = \int_c^d f(x, y) dy$ risulta integrabile su $[a, b]$ e si ha*

$$\int_a^b F(x) dx = \int_R f(x, y) dx dy. \quad (3.3.1)$$

La formula precedente afferma quindi che

$$\int_a^b \int_c^d f(x, y) dy dx = \int_R f(x, y) dx dy, \quad (3.3.2)$$

ovvero per calcolare l'integrale di f su R è possibile svolgere i due integrali iterati, prima nella variabile y e poi nella x .

Prima di dimostrare l'enunciato precedente, vediamo che esso vale per funzioni semplici: se f è una funzione semplice su R abbiamo il seguente risultato.

Proposizione 3.3.2. *Sia $f : R \rightarrow \mathbb{R}$ una funzione semplice. Allora le ipotesi del teorema 3.3.1 sono soddisfatte e vale la formula (3.3.2). Inoltre, la funzione F sarà semplice.*

Dimostrazione. Essendo f semplice, si scriverà nella forma

$$f(x, y) = \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij} \chi_{R_{ij}}(x, y),$$

con $R_{ij} := [x_{i-1}, x_i] \times [y_{j-1}, y_j]$ per opportune partizioni $a = x_0 < \dots < x_n = b$ e $c = y_0 < \dots < y_m = d$. Notiamo subito che per $x \in [a, b]$ fissato, diciamo $x \in [x_{\bar{i}-1}, x_{\bar{i}}]$ per qualche $\bar{i}$, avremo

$$f_x(y) = f(x, y) = \sum_{j=1}^m \lambda_{\bar{i}j} \chi_{R_{\bar{i}j}}(x, y) = \sum_{j=1}^m \lambda_{\bar{i}j} \chi_{[y_{j-1}, y_j]}(y).$$

In particolare, f_x è una funzione semplice sull'intervallo $[c, d]$ e in quanto tale è integrabile su tale intervallo. Inoltre, il suo integrale è

$$\int_c^d f_x(y) dy = \sum_{j=1}^m \lambda_{\bar{i}j} (y_j - y_{j-1}).$$

Ovviamente tale integrale dipende dal valore di x che abbiamo fissato, perchè è l'indice $\bar{i}$ a dipendere dalla scelta di $x \in [a, b]$. In particolare, $F(x) := \int_c^d f_x(y) dy$ avrà la seguente espressione

$$F(x) = \begin{cases} \sum_{j=1}^m \lambda_{1j} (y_j - y_{j-1}) & \text{se } x \in [x_0, x_1), \\ \sum_{j=1}^m \lambda_{2j} (y_j - y_{j-1}) & \text{se } x \in [x_1, x_2), \\ \dots & \\ \sum_{j=1}^m \lambda_{nj} (y_j - y_{j-1}) & \text{se } x \in [x_{n-1}, x_n). \end{cases}$$

In forma concisa, si avrà

$$F(x) = \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij} (y_j - y_{j-1}) \chi_{[x_{i-1}, x_i)}(x).$$

Risulterà che anche F è una funzione semplice, e in quanto tale sarà integrabile su $[a, b)$, con integrale pari a

$$\int_a^b F(x) dx = \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij} (y_j - y_{j-1}) (x_i - x_{i-1}) = \sum_{i=1}^n \sum_{j=1}^m \lambda_{ij} m(R_{ij}) = \int_R f dx dy.$$

Questo conclude la dimostrazione. $\square$

Dimostrazione del Teorema 3.3.1. Supponiamo f come nelle ipotesi, e in quanto integrabile, per ogni $\varepsilon > 0$ esisteranno due funzioni semplici $\varphi \in \mathcal{F}^+$ e $\psi \in \mathcal{F}^-$ tali che

$$\int_R \varphi dx dy - \int_R \psi dx dy \leq \varepsilon. \quad (3.3.3)$$

Per ogni x fissato, avremo

$$\psi_x(y) \leq f_x(y) \leq \varphi_x(y) \quad \forall y \in [c, d]$$

e inoltre, le funzioni introdotte sono integrabili sull'intervallo $[c, d)$ (la f_x lo è per ipotesi, mentre le altre due perchè sono semplici). Pertanto risulterà

$$\int_c^d \psi_x(y) dy \leq \int_c^d f_x(y) dy \leq \int_c^d \varphi_x(y) dy, \quad \forall x \in [a, b).$$

Denotando con $\Psi(x)$, $F(x)$, e $\Phi(x)$ i tre integrali sopra, rispettivamente, abbiamo che per la Proposizione 3.3.2, Ψ e Φ sono funzioni semplici su $[a, b)$, quindi integrabili e i loro integrali su $[a, b)$ coincidono con gli integrali di ψ e φ su R ; in altre parole

$$\int_R \psi dx dy = \int_a^b \Psi(x) dx \leq \int_a^b \Phi(x) dx = \int_R \varphi dx dy.$$

Inoltre, per la (3.3.3), si ha

$$\int_a^b \Phi(x) dx - \int_a^b \Psi(x) dx \leq \varepsilon. \quad (3.3.4)$$

Abbiamo quindi trovato, per ogni $\varepsilon > 0$, due funzioni semplici Ψ e Φ soddisfacenti $\Psi(x) \leq F(x) \leq \Phi(x)$, per ogni $x \in [a, b)$ e i cui integrali soddisfano (3.3.4). Questo implica che F è integrabile e il suo integrale è compreso tra $\int_a^b \Psi(x) dx$ e $\int_a^b \Phi(x) dx$, ovvero

$$\int_R \psi dx dy \leq \int_a^b F(x) dx \leq \int_R \varphi dx dy.$$

Visto che anche $\int_R f(x, y) dx dy$ soddisfa

$$\int_R \psi dx dy \leq \int_R f(x, y) dx dy \leq \int_R \varphi dx dy,$$

avremo,

$$\left| \int_R f(x, y) dx dy - \int_a^b F(x) dx \right| \leq \varepsilon.$$

Essendo tale disuguaglianza valida per ogni $\varepsilon > 0$ risulterà necessariamente $\int_R f(x, y) dx dy = \int_a^b F(x) dx$, e la dimostrazione è conclusa. $\square$

Enunciamo ora il teorema di Fubini in dimensione n generica. Dato un rettangolo $R = R := [a_1, b_1] \times \cdots \times [a_n, b_n]$ in $\mathbb{R}^n$ denotiamo con $R_{n-1} \subset \mathbb{R}^{n-1}$ il rettangolo $R_{n-1} := [a_1, b_1] \times \cdots \times [a_{n-1}, b_{n-1}]$. In generale denotiamo con $R_k \subset \mathbb{R}^k$ il rettangolo $R_k := [a_1, b_1] \times \cdots \times [a_k, b_k]$, per ogni $k = 1, \dots, n-1$.

Teorema 3.3.3. (Fubini) Sia $R := [a_1, b_1] \times \cdots \times [a_n, b_n]$ un rettangolo in $\mathbb{R}^n$ e sia $f : R \rightarrow \mathbb{R}$ una funzione limitata integrabile. Supponiamo che la funzione $x_n \mapsto f(x_1, \dots, x_n)$ sia integrabile su $[a_n, b_n]$ per ogni scelta di $(x_1, \dots, x_{n-1}) \in R_{n-1}$. Allora la funzione $F_n : R_{n-1} \rightarrow \mathbb{R}$ definita da

$$F_n(x_1, \dots, x_{n-1}) = \int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n$$

risulta integrabile su R_{n-1} e si ha

$$\int_{R_{n-1}} F_n(x_1, \dots, x_{n-1}) dx_1 \dots dx_{n-1} = \int_R f(x) dx. \quad (3.3.5)$$

In generale per ogni $k \in \{2, \dots, n-1\}$ supponiamo che la funzione $x_k \mapsto F_{k+1}(x_1, \dots, x_k)$ risulta integrabile su $[a_k, b_k]$ per ogni scelta di $(x_1, \dots, x_{k-1}) \in R_{k-1}$, allora

$$F_k(x_1, \dots, x_{k-1}) := \int_{a_k}^{b_k} F_{k+1}(x_1, \dots, x_k) dx_k$$

risulta integrabile su R_{k-1} e avremo

$$\int_{R_{k-1}} F_k(x_1, \dots, x_{k-1}) dx_1 \dots dx_{k-1} = \int_{R_k} F_{k+1}(x_1, \dots, x_k) dx_1 \dots dx_k.$$

Come conseguenza, nelle ipotesi precedenti avremo

$$\int_{a_1}^{b_1} \int_{a_2}^{b_2} \cdots \int_{a_n}^{b_n} f(x_1, \dots, x_n) dx_n dx_{n-1} \dots dx_1 = \int_R f(x) dx.$$

Introduciamo ora un teorema che ci permette di calcolare integrali multipli con una formula di cambio di variabili.

A questo scopo specifichiamo prima cos'è una mappa di cambio di variabile. Supponiamo $A, B \subset \mathbb{R}^n$ siano due insiemi non vuoti e sia $\Phi : A \rightarrow B$ una funzione biunivoca che soddisfa le seguenti proprietà:

(C1) Φ è di classe C^1 , e scriveremo $\Phi \in C^1(A; B)$;

(C2) il gradiente di Φ non si annulla mai in A , ovvero $\nabla \Phi(x) \neq 0$ per ogni $x \in A$.

Ricordiamo che essendo Φ una mappa da $A \subset \mathbb{R}^n$ in $B \subset \mathbb{R}^n$, il suo gradiente sarà una matrice $n \times n$, con componenti

$$\nabla \Phi_{ij} = \frac{\partial \Phi_i}{\partial x_j}$$

e in quanto tale il significato di (C2) è per ogni x si ha che le componenti $\frac{\partial \Phi_i}{\partial x_j}$ non sono tutte contemporaneamente nulle (in altre parole, $\nabla \Phi$ non è mai la matrice identicamente nulla). Inoltre si può calcolare il determinante di $\nabla \Phi$, detto determinante Jacobiano. Richiederemo inoltre che

(C3) $\det(\nabla \Phi(x)) \neq 0$ per ogni $x \in A$.

Si noti a questo punto, che (C3) implica (C2).

Teorema 3.3.4. *Supponiamo A, B sottoinsiemi aperti di $\mathbb{R}^n$ e sia $\Phi : A \rightarrow B$ una mappa che soddisfa (C1)-(C3). Sia $f : B \rightarrow \mathbb{R}$ una funzione integrabile su $E \subset B$, allora*

$$\int_E f(y) dy = \int_{\Phi^{-1}(E)} f(\Phi(x)) |\det(\nabla \Phi(x))| dx. \quad (3.3.6)$$

Non dimostreremo questo teorema, ma lo utilizzeremo spesso per risolvere esercizi di integrali multipli.

3.3.1 Coordinate polari, cilindriche e sferiche

Nel caso planare, in $\mathbb{R}^2$, abbiamo la nozione di coordinate polari che sono spesso utili per risolvere integrali multipli.

Sia $\Phi : (0, +\infty) \times [0, 2\pi) \rightarrow \mathbb{R}^2 \setminus \{0\}$ la funzione di cambio di variabile

$$\Phi(\rho, \theta) = (\rho \cos \theta, \rho \sin \theta). \quad (3.3.7)$$

Un facile conto mostra che il determinante di $\nabla \Phi$ coincide con ρ ,

$$\det(\nabla \Phi(\rho, \theta)) = \rho. \quad (3.3.8)$$

Le coordinate ρ, θ vengono chiamate coordinate polari, e parametrizzano $\Phi_1(\rho, \theta) = x$ e $\Phi_2(\rho, \theta) = y$, quindi il sistema di cambio di variabili si può anche riscrivere come

$$\begin{cases} x = \rho \cos \theta, \\ y = \rho \sin \theta. \end{cases} \quad (3.3.9)$$

Nel caso consideriamo anche la coordinata verticale z , e quindi ci mettiamo in $\mathbb{R}^3$, possiamo considerare il cambio di variabili

$$\begin{cases} x = \rho \cos \theta, \\ y = \rho \sin \theta, \\ z = z'. \end{cases} \quad (3.3.10)$$

Le coordinate (ρ, θ, z') vengono chiamate coordinate cilindriche e si ottengono dal cambio di variabili $\Phi : (0, +\infty) \times [0, 2\pi) \times \mathbb{R} \rightarrow \mathbb{R}^3 \setminus \{(x, y, z) : x = y = 0\}$, data da

$$\Phi(\rho, \theta, z') = (\rho \cos \theta, \rho \sin \theta, z'). \quad (3.3.11)$$

Anche in questo caso lasciamo per esercizio il calcolo del gradiente e del determinante, che come nel caso precedente verrà

$$\det(\nabla\Phi(\rho, \theta, z')) = \rho. \quad (3.3.12)$$

Lievemente più complicate sono le coordinate sferiche (r, θ, φ) , che portano al sistema

$$\begin{cases} x = r \cos \varphi \cos \theta, \\ y = r \cos \varphi \sin \theta, \\ z = r \sin \varphi. \end{cases} \quad (3.3.13)$$

La mappa di cambio di variabile sarà $\Phi : (0, +\infty) \times [0, 2\pi) \times (-\frac{\pi}{2}, \frac{\pi}{2}) \rightarrow \mathbb{R}^3 \setminus \{(x, y, z) : x = y = 0\}$, data da

$$\Phi(r, \theta, \varphi) = (r \cos \varphi \cos \theta, r \cos \varphi \sin \theta, r \sin \varphi). \quad (3.3.14)$$

Il determinante di $\nabla\Phi$ in questo caso viene

$$\det(\nabla\Phi(r, \theta, \varphi)) = r^2 \cos \varphi. \quad (3.3.15)$$

3.4 Integrali impropri

Abbiamo finora visto come integrare funzioni limitate su domini limitati, trattiamo ora il caso in cui le funzioni sono illimitate o il dominio di integrazione è tale.

Definizione 3.4.1. Sia $E \subset \mathbb{R}^n$ un sottoinsieme non limitato e tale che, per ogni $R > 0$ risulti che l'insieme $E \cap B_R(0)$, quando non vuoto, è un insieme misurabile secondo Peano-Jordan.

Sia $f : E \rightarrow \mathbb{R}$ una funzione limitata non negativa, e tale che f risulti integrabile su $E \cap B_R(0)$ per ogni $R > 0$. Allora f si dice integrabile in senso improprio su E se il limite

$$\lim_{R \rightarrow \infty} \int_{E \cap B_R(0)} f(x) dx, \quad (3.4.1)$$

esiste finito. In tal caso si definisce tale limite come l'integrale di f su E , ovvero

$$\int_E f(x) dx = \lim_{R \rightarrow \infty} \int_{E \cap B_R(0)} f(x) dx.$$

Nel caso in cui f non è non negativa, si scompone $f = f^+ - f^-$ dove $f^+(x) := \max\{f(x), 0\}$ e $f^-(x) := \max\{-f(x), 0\}$ sono la parte positiva e negativa di f , rispettivamente. Allora diremo che f è integrabile in senso improprio su E se f^+ e f^- lo sono, e in tal caso definiamo

$$\int_E f(x) dx = \int_E f^+(x) dx - \int_E f^-(x) dx.$$

Nel caso in cui invece f va integrata su un dominio limitato, ma non risulta essere una funzione limitata, ci appelliamo alla seguente definizione.

Definizione 3.4.2. Sia $E \subset \mathbb{R}^n$ un sottoinsieme limitato misurabile secondo Peano-Jordan, e sia $\bar{x} \in \mathbb{R}^n$ un punto di accumulazione per E . Sia $f : E \rightarrow \mathbb{R}$ una funzione non negativa e tale che f risulti limitata e integrabile su $E \setminus B_r(\bar{x})$ per ogni $r > 0$. Allora f si dice integrabile in senso improprio su E se il limite

$$\lim_{r \rightarrow 0^+} \int_{E \setminus B_r(\bar{x})} f(x) dx, \quad (3.4.2)$$

esiste finito. In tal caso si definisce tale limite come l'integrale di f su E , ovvero

$$\int_E f(x) dx = \lim_{r \rightarrow 0^+} \int_{E \setminus B_r(\bar{x})} f(x) dx.$$

Nel caso in cui f non è non negativa, anche in questo caso si scompone $f = f^+ - f^-$ e diremo che f è integrabile in senso improprio su E se f^+ e f^- lo sono, e in tal caso definiamo

$$\int_E f(x) dx = \int_E f^+(x) dx - \int_E f^-(x) dx.$$

Osservazione 3.4.3. Un'osservazione importante è che i limiti in (3.4.1) e (3.4.2) esistono sempre, perchè la funzione f è stata presa non negativa e pertanto la quantità $\int_{E \cap B_R(0)} f(x) dx$ è monotona non decrescente in R (similmente, $\int_{E \setminus B_r(\bar{x})} f(x) dx$ è monotona non crescente in r). Il problema è che entrambi possono essere $+\infty$.

Un'altra osservazione importante è che tali limiti possono essere svolti anche se invece di rimuovere (o intersecare) delle palle, prendiamo altre figure geometriche. Questa osservazione è formalizzata nel prossimo lemma.

Definizione 3.4.4. Sia $F \subseteq \mathbb{R}^n$ un insieme qualsiasi contenente l'origine. Per ogni $r > 0$ definiamo il riscalato F_r di F di fattore r come

$$F_r := \{rx : x \in F\}.$$

Inoltre, per ogni $\bar{x} \in \mathbb{R}^n$, definiamo $F(\bar{x})$ il traslato di F come

$$F(\bar{x}) := \{x + \bar{x} : x \in F\}.$$

Ad esempio se $F = B_R(0)$, risulta $F_r = B_{Rr}(0)$ e $F(\bar{x}) = B_R(\bar{x})$.

Lemma 3.4.5. Sia $F \subset \mathbb{R}^n$ un insieme misurabile secondo Peano-Jordan che contiene l'origine e risulta essere limitato. Supponiamo che esistano due costanti positive c_1, c_2 tali che

$$B_{c_1}(0) \subseteq F \subseteq B_{c_2}(0).$$

Sia inoltre $f : E \rightarrow \mathbb{R}$ una funzione non negativa definita su E come nella Definizione 3.4.1. Allora vale

$$\lim_{R \rightarrow \infty} \int_{E \cap B_R(0)} f(x) dx = \lim_{R \rightarrow \infty} \int_{E \cap F_R} f(x) dx.$$

Se invece $f : E \rightarrow \mathbb{R}$ è una funzione non negativa e E è come nella Definizione 3.4.2 allora

$$\lim_{r \rightarrow 0^+} \int_{E \setminus B_r(\bar{x})} f(x) dx = \lim_{r \rightarrow 0^+} \int_{E \setminus F_r(\bar{x})} f(x) dx.$$

Dimostrazione. Mettiamoci nel primo caso: grazie all'ipotesi avremo, per ogni $R > 0$,

$$B_{c_1 R}(0) \subseteq F_R \subseteq B_{c_2 R}(0).$$

Di conseguenza avremo, essendo f non negativa,

$$\int_{E \cap B_{c_1 R}(0)} f(x) dx \leq \int_{E \cap F_R} f(x) dx \leq \int_{E \cap B_{c_2 R}(0)} f(x) dx.$$

Passando allora al limite per $R \rightarrow \infty$ la disuguaglianza precedente, otteniamo subito la tesi grazie al teorema dei carabinieri. In modo analogo si dimostra la seconda uguaglianza (esercizio). $\square$

Capitolo 4

Successioni e serie di funzioni

Così come si studiano successioni e serie numeriche, possiamo generalizzare tale concetto a successioni e serie in cui ogni termine dipende da una variabile x . In tal caso, ogni termine della successione o serie in questione potrà essere visto come una funzione della variabile x , definita su un opportuno dominio.

4.1 Successioni di funzioni

Definizione 4.1.1. Sia $f_n : E \rightarrow \mathbb{R}$, $n \in \mathbb{N}$, una successione di funzioni definite su un dominio comune $E \subseteq \mathbb{R}^m$. Diremo che la successione di funzioni f_n converge puntualmente ad una funzione $f : E \rightarrow \mathbb{R}$ se per ogni $x \in E$ la successione numerica $f_n(x)$ tende a $f(x)$, ovvero

$$\forall x \in E, \quad f_n(x) \rightarrow f(x).$$

Diremo che f_n converge ad f uniformemente se vale la seguente condizione

$$\forall \varepsilon > 0 \quad \exists n_\varepsilon \in \mathbb{N} \text{ tale che } |f_n(x) - f(x)| \leq \varepsilon, \quad \forall x \in E, \quad \forall n \geq n_\varepsilon. \quad (4.1.1)$$

E' facile verificare che la convergenza uniforme implica quella puntuale. Il viceversa è falso in generale. Possiamo esprimere la condizione di convergenza puntuale nel seguente modo: f_n converge puntualmente ad f se e solo se

$$\forall x \in E, \quad \forall \varepsilon > 0 \quad \exists n_\varepsilon \in \mathbb{N} \text{ tale che } |f_n(x) - f(x)| \leq \varepsilon, \quad \forall n \geq n_\varepsilon. \quad (4.1.2)$$

La presenza del quantificatore $\forall x \in E$ all'inizio dell'affermazione fa sì che in questo caso $n_\varepsilon = n_\varepsilon(x)$ dipende da x e quindi può essere diverso a seconda del punto x in cui stiamo guardando la convergenza. In (4.1.1) invece, n_ε non dipende da x , e quindi è "uniforme" rispetto a $x \in E$.

Definizione 4.1.2. Sia $f : E \rightarrow \mathbb{R}$ con $E \subseteq \mathbb{R}^m$. Definiamo la norma infinito $\|f\|_\infty$ di f come

$$\|f\|_\infty := \sup\{|f(x)| : x \in E\}. \quad (4.1.3)$$

Se $A \subset E$ possiamo definire la norma infinito $\|f\|_{\infty, A}$ di f su A come

$$\|f\|_{\infty, A} := \sup\{|f(x)| : x \in A\}. \quad (4.1.4)$$

Osservazione 4.1.3. Si dimostra con una verifica immediata (esercizio) che $f_n : E \rightarrow \mathbb{R}$ converge uniformemente a $f : E \rightarrow \mathbb{R}$ se e solo se

$$\forall \varepsilon > 0 \quad \exists n_\varepsilon \in \mathbb{N} \text{ tale che } \|f_n - f\|_\infty \leq \varepsilon \quad \forall n \geq n_\varepsilon. \quad (4.1.5)$$

Esempio 4.1.4. Sia $f_n : [0, 1] \rightarrow \mathbb{R}$, $n \geq 2$, la funzione

$$f_n(x) = \begin{cases} nx & \text{se } x \in [0, \frac{1}{n}] \\ 2 - nx & \text{se } x \in [\frac{1}{n}, \frac{2}{n}] \\ 0 & \text{se } x \in [\frac{2}{n}, 1]. \end{cases}$$

Si verifica che per ogni $x \in [0, 1]$ la successione $f_n(x) \rightarrow 0$ quando $n \rightarrow \infty$. Quindi $f_n \rightarrow f$ puntualmente, dove $f : [0, 1] \rightarrow \mathbb{R}$ è la funzione identicamente nulla. D'altra parte non c'è convergenza uniforme, infatti, grazie all'Osservazione 4.1.3 possiamo vedere che $\|f_n\|_\infty = 1$ per ogni n , e pertanto $\|f_n - f\|_\infty = \|f_n\|_\infty$ non tende a zero.

Un'altra verifica immediata porta al seguente:

Lemma 4.1.5. Sia $f : E \rightarrow \mathbb{R}$. La quantità $\|f\|_\infty$ è finita se e solo se f è limitata.

Definizione 4.1.6. Sia $E \subset \mathbb{R}^m$. Definiamo i seguenti insiemi di funzioni:

$$\begin{aligned} F_b(E) &:= \{f : E \rightarrow \mathbb{R} : f \text{ è limitata}\}, \\ C(E) &:= \{f : E \rightarrow \mathbb{R} : f \text{ è continua su } E\}, \\ C_b(E) &:= C(E) \cap F_b(E), \end{aligned}$$

ovvero $C_b(E)$ consiste delle funzioni $f : E \rightarrow \mathbb{R}$ continue e limitate su E .

Lemma 4.1.7. Gli insiemi $F_b(E)$, $C(E)$ e $C_b(E)$ sono spazi vettoriali su $\mathbb{R}$. Inoltre la quantità $\|\cdot\|_\infty$ è una norma su $F_b(E)$ e $C_b(E)$.

Dimostrazione. La verifica che $F_b(E)$, $C(E)$ e $C_b(E)$ sono spazi vettoriali è immediata e segue dal fatto che sono insiemi chiusi per combinazioni lineari a coefficienti reali.

Inoltre, per il Lemma 4.1.5 la quantità $\|f\|_\infty$ è finita per ogni $f \in F_b(E)$ e $f \in C_b(E)$. Per verificare che $\|\cdot\|_\infty$ è una norma, dobbiamo controllare che siano valide le seguenti condizioni, per ogni funzione f e g :

- $\|f\|_\infty \geq 0$, e $\|f\|_\infty = 0$ se e solo se $f \equiv 0$;
- per ogni $\lambda \in \mathbb{R}$, $\|\lambda f\|_\infty = |\lambda| \|f\|_\infty$;
- $\|f + g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$.

Lasciamo al lettore tale semplice esercizio. Per l'ultima condizione, la disuguaglianza triangolare, basta sapere che per ogni $x \in E$ si ha

$$|f(x) + g(x)| \leq |f(x)| + |g(x)| \leq \|f\|_\infty + \|g\|_\infty,$$

da cui segue la tesi passando all'estremo superiore su $x \in E$ nel membro di sinistra. $\square$

Lemma 4.1.8. *Siano $f_n \in F_b(E)$ e sia $f : E \rightarrow \mathbb{R}$ tale che $f_n \rightarrow f$ uniformemente. Allora $f \in F_b(E)$ e*

$$\|f_n\|_\infty \rightarrow \|f\|_\infty.$$

Dimostrazione. Per definizione di convergenza uniforme avremo $\|f_n - f\|_\infty \rightarrow 0$ quando $n \rightarrow \infty$. Inoltre, per ogni n avremo

$$\|f\|_\infty \leq \|f_n\|_\infty + \|f_n - f\|_\infty,$$

e allo stesso tempo

$$\|f_n\|_\infty \leq \|f\|_\infty + \|f_n - f\|_\infty.$$

La prima delle due disuguaglianze dice che necessariamente $\|f\|_\infty < +\infty$ che ci dice che f è limitata. Mettendo invece insieme le due disuguaglianze precedente si evince che

$$|\|f_n\|_\infty - \|f\|_\infty| \leq \|f_n - f\|_\infty,$$

da cui segue immediatamente che $\|f_n\|_\infty \rightarrow \|f\|_\infty$. □

Mostriamo ora la seguente proprietà:

Lemma 4.1.9. *Sia $f_n : E \rightarrow \mathbb{R}$ una successione di funzioni continue tali che $f_n \rightarrow f$ uniformemente. Allora f è una funzione continua.*

Dimostrazione. Sia $x \in E$; facciamo vedere che f è continua in x . Sia $\varepsilon > 0$, allora esiste $n_\varepsilon \in \mathbb{N}$ tale che

$$|f_n(y) - f(y)| \leq \frac{\varepsilon}{3} \quad \text{per ogni } y \in E, \forall n \geq n_\varepsilon.$$

In particolare, questo varrà per $n = n_\varepsilon$, ed essendo f_{n_ε} continua in x avremo anche che esiste $\delta > 0$ tale che

$$|f_{n_\varepsilon}(y) - f_{n_\varepsilon}(x)| \leq \frac{\varepsilon}{3} \quad \text{per ogni } y \in B_\delta(x) \cap E.$$

Allora, per ogni $y \in B_\delta(x) \cap E$ avremo anche

$$|f(x) - f(y)| \leq |f(x) - f_{n_\varepsilon}(x)| + |f_{n_\varepsilon}(x) - f_{n_\varepsilon}(y)| + |f_{n_\varepsilon}(y) - f(y)| \leq \frac{\varepsilon}{3} + \frac{\varepsilon}{3} + \frac{\varepsilon}{3} = \varepsilon.$$

Quindi abbiamo scoperto che per ogni $\varepsilon > 0$ possiamo trovare $\delta > 0$ tale che $|f(x) - f(y)| \leq \varepsilon$ per ogni $y \in E$ con $\|x - y\| < \delta$, ovvero f è continua in x . □

Definizione 4.1.10. *Dato uno spazio vettoriale normato, è naturale introdurre una distanza nel seguente modo: Se X denota lo spazio vettoriale e $\|\cdot\|$ la norma su di esso, definiamo la distanza tra $x, y \in X$ come*

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

La verifica che la distanza precedente è di fatto una distanza è lasciata per esercizio.

Applicando la precedente definizione agli spazi $F_b(E)$ e $C_b(E)$ otteniamo che tali spazi sono spazi metrici dotati della distanza

$$d(f, g) := \|f - g\|_\infty, \quad \forall f, g.$$

Tale distanza induce, come discusso nel Capitolo 1, il concetto di convergenza: $f_n \rightarrow f$ in tale spazio se e solo se $d(f_n, f) \rightarrow 0$, ovvero se e solo se $\|f_n - f\|_\infty \rightarrow 0$, e quindi se e solo se $f_n \rightarrow f$ uniformemente. Come conseguenza del Lemma precedente abbiamo il seguente:

Corollario 4.1.11. *Lo spazio vettoriale $C_b(E)$ è chiuso rispetto alla convergenza uniforme.*

Mostriamo ora che $C_b(E)$ è anche uno spazio completo.

Teorema 4.1.12. *Lo spazio $C_b(E)$ è completo.*

Dimostrazione. Dobbiamo far vedere che ogni successione di Cauchy converge (ovvero esiste $f \in C_b(E)$ tale che $f_n \rightarrow f$ in $C_b(E)$). A questo scopo, sia f_n una successione di Cauchy, ovvero tale che per ogni $\varepsilon > 0$ esiste $n_\varepsilon \in \mathbb{N}$ tale che $\|f_n - f_m\|_\infty \leq \varepsilon$ per ogni $n, m \geq n_\varepsilon$. Per definizione di norma $\|\cdot\|_\infty$ sarà allora vero che, per ogni $x \in E$,

$$|f_n(x) - f_m(x)| \leq \varepsilon \quad \text{per ogni } n, m \geq n_\varepsilon.$$

In altre parole, per ogni $x \in E$ la successione (numerica) $f_n(x)$ è una successione di Cauchy. In quanto tale, essa converge ad un valore reale che chiamiamo $f(x)$. Pertanto, abbiamo scoperto che $f_n \rightarrow f$ puntualmente. Ora, la convergenza sarà anche uniforme, dal momento che, passando al limite per $m \rightarrow \infty$ nella espressione precedente troviamo che, per ogni $\varepsilon > 0$ esiste $n_\varepsilon \in \mathbb{N}$ tale che, per ogni $x \in E$ si ha

$$|f_n(x) - f(x)| \leq \varepsilon \quad \text{per ogni } n, m \geq n_\varepsilon.$$

Ma questa equivale alla convergenza uniforme di f_n ad f . A sua volta, essa è equivalente a richiedere che $\|f_n - f\|_\infty \rightarrow 0$ quando $n \rightarrow \infty$. Ora, per il Lemma 4.1.9, $f \in C_b(E)$ e pertanto abbiamo trovato una funzione $f \in C_b(E)$ tale che $f_n \rightarrow f$ rispetto la distanza in $C_b(E)$. La tesi segue. $\square$

Mostriamo ora che la convergenza uniforme preserva anche l'integrabilità delle funzioni.

Lemma 4.1.13. *Siano $f_n : E \rightarrow \mathbb{R}$ funzioni limitate e integrabili secondo Riemann su un insieme E misurabile secondo Peano-Jordan, con $E \subset R$ e R un rettangolo. Supponiamo $f_n \rightarrow f$ uniformemente, allora f risulta integrabile e*

$$\int_E f_n(x) dx \rightarrow \int_E f(x) dx.$$

Prima di dimostrare il Lemma premettiamo la seguente osservazione:

Osservazione 4.1.14. Sia $f : E \rightarrow \mathbb{R}$ limitata con E come nel lemma precedente: allora f è integrabile se e solo se per ogni $\varepsilon > 0$ possiamo trovare due funzioni integrabili $g, h : E \rightarrow \mathbb{R}$ con $g \leq f \leq h$ e tali che

$$\int_E h(x) - g(x) dx \leq \varepsilon.$$

Infatti, da una parte se f è integrabile la condizione sopra è sempre verificata perchè possiamo trovare g, h semplici (e quindi integrabili) che soddisfano tale condizione. Viceversa, supponiamo la condizione sopra sia soddisfatta; essendo g ed h integrabili possiamo trovare funzioni semplici $\psi^\pm, \varphi^\pm$ tali che $\psi^- \leq g \leq \psi^+$ e $\varphi^- \leq h \leq \varphi^+$ e

$$\int_E \psi^+(x) - \psi^-(x) dx \leq \varepsilon, \quad \int_E \varphi^+(x) - \varphi^-(x) dx \leq \varepsilon.$$

Sarà allora anche vero che

$$\int_E g(x) - \psi^-(x) dx \leq \varepsilon, \quad \int_E \varphi^+(x) - h(x) dx \leq \varepsilon,$$

e quindi

$$\int_E \varphi^+(x) - \psi^-(x) dx = \int_E \varphi^+(x) - h(x) dx + \int_E h(x) - g(x) dx + \int_E g(x) - \psi^-(x) dx \leq 3\varepsilon.$$

Pertanto, per ogni $\varepsilon > 0$ possiamo trovare funzioni semplici $\varphi^+, \psi^- : E \rightarrow \mathbb{R}$ tali che $\psi^- \leq f \leq \varphi^+$ e $\int_E \varphi^+(x) - \psi^-(x) dx \leq \varepsilon$, condizione sufficiente (e necessaria, in realtà) per garantire l'integrabilità di f .

Dimostrazione del Lemma 4.1.13. Visto che $f_n \rightarrow f$ uniformemente per ogni $\varepsilon > 0$ esiste n_ε tale che $\|f_n - f\|_\infty \leq \varepsilon$ per ogni $n \geq n_\varepsilon$. Da ciò $|f(x) - f_n(x)| \leq \varepsilon$ per ogni $x \in \mathbb{R}$ e quindi $f(x) \leq f_n(x) + \varepsilon =: h(x)$ e allo stesso tempo $g(x) := f_n(x) - \varepsilon \leq f(x)$ per ogni $x \in E$. Quindi per ogni $\varepsilon > 0$ troviamo due funzioni g, h come nell'Osservazione precedente, implicando che f è integrabile. Per mostrare che l'integrale di f_n converge a quello di f basta osservare che

$$\int_E f_n(x) dx - m(E)\varepsilon = \int_E f_n(x) - \varepsilon dx \leq \int_E f(x) dx \leq \int_E f_n(x) + \varepsilon dx = m(E)\varepsilon + \int_E f_n(x) dx$$

che implica

$$\left| \int_E f(x) dx - \int_E f_n(x) dx \right| \leq \varepsilon m(E).$$

Essendo tale disuguaglianza valida per ogni $n \geq n_\varepsilon$ abbiamo mostrato che $\int_E f_n(x) dx \rightarrow \int_E f(x) dx$. $\square$

Definizione 4.1.15. Sia $f : [a, b] \rightarrow \mathbb{R}^k$. Diremo che f è di classe C^1 sull'intervallo chiuso $[a, b]$ se

- f è continua su $[a, b]$;
- $f \in C^1((a, b); \mathbb{R}^k)$;
- esiste la derivata destra $f'(a^+)$ di f in a ed esiste la derivata sinistra $f'(b^-)$ di f in b ;
- ponendo

$$f'(x) = \begin{cases} f'(a^+) & \text{se } x = a \\ f'(x) & \text{se } x \in (a, b) \\ f'(b^-) & \text{se } x = b \end{cases}$$

allora f' è continua su $[a, b]$.

Mostriamo ora il seguente:

Teorema 4.1.16. Siano $f_n : [a, b] \rightarrow \mathbb{R}$ funzioni di classe C^1 sull'intervallo chiuso $[a, b]$. Siano $f, g : [a, b] \rightarrow \mathbb{R}$ tali che f_n converge uniformemente ad f ed f'_n converge uniformemente a g su $[a, b]$. Allora f è di classe C^1 su $[a, b]$ e $f' = g$.

Dimostrazione. Sappiamo che per ogni n vale il teorema fondamentale del calcolo integrale:

$$f_n(x) = f_n(a) + \int_a^x f'_n(t) dt, \quad \forall x \in [a, b].$$

Grazie al fatto che f'_n sono continue deduciamo che anche g è continua, e inoltre f'_n essendo integrabile su $[a, b]$ soddisfa, grazie al Lemma 4.1.13,

$$\int_a^x f'_n(t) dt \rightarrow \int_a^x g(t) dt.$$

Inoltre avremo $f_n \rightarrow f$ puntualmente, e quindi passando al limite per $n \rightarrow \infty$ nella formula sopra otteniamo

$$f(x) = f(a) + \int_a^x g(t) dt, \quad \forall x \in [a, b].$$

Da questa formula, di nuovo il teorema fondamentale del calcolo integrale ci permette di dedurre che $f'(x) = g(x)$ per ogni $x \in [a, b]$, da cui si deduce facilmente la tesi. $\square$

Definizione 4.1.17. Introduciamo il seguente spazio vettoriale

$$C^1([a, b]) := \{f : [a, b] \rightarrow \mathbb{R} : f \text{ è di classe } C^1 \text{ su } [a, b]\}.$$

Ogni funzione $f \in C^1([a, b])$ risulta limitata con f' limitata. Di conseguenza per ogni tale funzione risulterà finita la norma

$$\|f\|_{C^1} := \|f\|_{\infty} + \|f'\|_{\infty}.$$

Come prima, risulterà che $C^1([a, b])$ è uno spazio vettoriale e $\|\cdot\|_{C^1}$ è una norma su tale spazio. Grazie al teorema precedente possiamo mostrare ora che $C^1([a, b])$ è uno spazio metrico completo rispetto alla distanza indotta da tale norma.

Teorema 4.1.18. *Lo spazio $C^1([a, b])$ è completo se dotato della norma $\|\cdot\|_{C^1}$.*

Dimostrazione. Sia $f_n \in C^1([a, b])$ una successione di Cauchy: ovvero, per ogni $\varepsilon > 0$ esiste n_{ε} per cui

$$\|f_n - f_m\|_{\infty} + \|f'_n - f'_m\|_{\infty} = \|f_n - f_m\|_{C^1} \leq \varepsilon, \quad \forall n, m \geq n_{\varepsilon}.$$

Ma allora si vede subito che f_n e f'_n sono entrambe successioni di Cauchy in $C_b([a, b])$. Ne deduciamo dal Teorema 4.1.12 che esistono $f, g \in C_b([a, b])$ tali che

$$f_n \rightarrow f, \quad f'_n \rightarrow g \quad \text{uniformemente per } n \rightarrow \infty.$$

Concludiamo dal Teorema 4.1.16 che $f \in C^1([a, b])$, e $f' = g$. Infine, grazie alle convergenze appena mostrate, si vede subito che $\|f_n - f\|_{C^1} \rightarrow 0$ quando $n \rightarrow \infty$, che conclude la dimostrazione. $\square$

4.2 Serie di funzioni

Serie di funzioni sono particolari successioni, espresse come somme parziali.

Definizione 4.2.1. Siano $f_n : E \rightarrow \mathbb{R}$ funzioni date; la serie di funzioni $\sum_{n=0}^{\infty} f_n$ si dice che converge puntualmente ad $f : E \rightarrow \mathbb{R}$ se la successione delle somme parziali $S_n := \sum_{i=0}^n f_i$ converge puntualmente ad f . In tal caso scriveremo $f(x) = \sum_{n=0}^{\infty} f_n(x)$. Similmente diremo che la serie converge uniformemente ad f se la successione di somme parziali S_n converge ad f uniformemente. Diremo inoltre che la serie $\sum_{n=0}^{\infty} f_n$ converge totalmente se $\sum_{n=0}^{\infty} \|f_n\|_{\infty}$ è una serie convergente.

Lemma 4.2.2. *Supponiamo la serie $\sum_{n=0}^{\infty} f_n$ converga totalmente. Allora esiste una funzione $f : E \rightarrow \mathbb{R}$ tale che $\sum_{n=0}^{\infty} f_n$ converge uniformemente ad f .*

Dimostrazione. Da una parte abbiamo

$$\sum_{n=0}^{\infty} |f_n(x)| \leq \sum_{n=0}^{\infty} \|f_n\|_{\infty} < +\infty,$$

quindi per ogni $x \in E$ la serie $\sum_{n=0}^{\infty} |f_n(x)|$ converge, e quindi anche $\sum_{n=0}^{\infty} f_n(x)$ converge; chiamiamo il limite $f(x)$. Mostriamo che le somme parziali convergono uniformemente ad f . Infatti

$$|f(x) - S_n(x)| = \left| \sum_{i=n+1}^{\infty} f_i(x) \right| \leq \sum_{i=n+1}^{\infty} |f_i(x)| \leq \sum_{i=n+1}^{\infty} \|f_i\|_{\infty},$$

ed essendo tale disuguaglianza valida per ogni $x \in E$ otteniamo anche

$$\|f - S_n\|_{\infty} \leq \sum_{i=n+1}^{\infty} \|f_i\|_{\infty}.$$

Il membro di destra andando a 0 quando $n \rightarrow \infty$, concludiamo che $\|f - S_n\|_{\infty} \rightarrow 0$ quando $n \rightarrow \infty$. $\square$

Teorema 4.2.3. *Sia $f = \sum_{n=0}^{\infty} f_n$ una serie convergente totalmente, dove $f_n : E \rightarrow \mathbb{R}$ sono funzioni limitate integrabili; allora f è limitata, integrabile e*

$$\int_E f dx = \sum_{n=0}^{\infty} \int_E f_n dx.$$

Dimostrazione. Sappiamo dal lemma precedente che le somme parziali $S_n \rightarrow f$ uniformemente. Dal Lemma 4.1.13 sappiamo quindi che $\int_E S_n dx \rightarrow \int_E f dx$. Pertanto

$$\int_E f dx = \lim_{n \rightarrow \infty} \sum_{i=0}^n \int_E f_i dx = \sum_{n=0}^{\infty} \int_E f_n dx.$$

$\square$

Teorema 4.2.4. *Siano $f_n : [a, b] \rightarrow \mathbb{R}$ funzioni di classe $C^1([a, b])$, e supponiamo che $\sum_{n=0}^{\infty} f_n$ e $\sum_{n=0}^{\infty} f'_n$ siano serie convergenti totalmente ad f e g , rispettivamente. Allora $f \in C^1([a, b])$ e $f' = g$.*

Dimostrazione. Applicando il Teorema 4.1.16 alle somme parziali $S_n = \sum_{i=0}^n f_i$ e alle loro derivate $S'_n = \sum_{i=0}^n f'_i$ otteniamo subito la tesi. $\square$

4.2.1 Serie di potenze

Le serie di potenze sono particolari serie di funzioni in cui $f_n(y) = a_n(y - y_0)^n$; qui $a_n \in \mathbb{R}$ sono numeri reali e $y_0 \in \mathbb{R}$ è fissato. Con il cambio di variabili $x = y - y_0$ la serie di potenze si scrive:

$$\sum_{n=0}^{\infty} a_n x^n.$$

Sicuramente osserviamo che tale serie converge per il valore $x = 0$. Tuttavia, per valori diversi da 0 potrebbe non convergere mai, come ad esempio nel caso in cui $a_n = n^n$.

Vale comunque il fatto seguente:

Lemma 4.2.5. *Sia $\sum_{n=0}^{\infty} a_n x^n$ una serie di potenze: supponiamo che essa converga per $x = y \neq 0$. Allora per ogni $x \in (-|y|, |y|)$ si ha che $\sum_{n=0}^{\infty} a_n x^n$ converge.*

Dimostrazione. Sappiamo che $\sum_{n=0}^{\infty} a_n y^n$ converge, e quindi il suo termine generale sarà infinitesimo, e la successione $a_n y^n$ limitata. Esiste dunque $M = \sup_n \{|a_n y^n|\} < +\infty$. Sia ora $x \in (-|y|, |y|)$ e mostriamo che $\sum_{n=0}^{\infty} a_n x^n$ converge assolutamente:

$$\sum_{n=0}^{\infty} |a_n x^n| = \sum_{n=0}^{\infty} |a_n y^n| \left| \frac{x^n}{y^n} \right| \leq \sum_{n=0}^{\infty} M \frac{|x|^n}{|y|^n} = M \frac{1}{1-r} < +\infty, \quad (4.2.1)$$

dove abbiamo posto $r = \left| \frac{x}{y} \right| < 1$. Questo conclude la dimostrazione. $\square$

Osservazione 4.2.6. Come conseguenza del Lemma appena enunciato abbiamo il fatto seguente: se la serie di potenze $\sum_{n=0}^{\infty} a_n x_0^n$ non converge, allora essa non converge per nessun valore di x tale che $|x| > x_0$. Infatti, se per assurdo essa convergesse per un certo y con $|y| > x_0$, convergerebbe anche per tutti i valori x compresi in $(-|y|, |y|)$, che includono x_0 , assurdo.

Definizione 4.2.7. *Sia $\sum_{n=0}^{\infty} a_n x^n$ una serie di potenze. Si definisce raggio di convergenza R della serie il numero (eventualmente uguale a $+\infty$)*

$$R := \sup \left\{ |x| : \sum_{n=0}^{\infty} a_n x^n \text{ converge} \right\}.$$

Come conseguenza del Lemma 4.2.5 e l'Osservazione 4.2.6 deduciamo immediatamente il seguente:

Teorema 4.2.8. *Sia $\sum_{n=0}^{\infty} a_n x^n$ una serie di potenze e sia R il suo raggio di convergenza. Se $R = 0$ allora la serie converge solo per $x = 0$; se $R = +\infty$ allora $\sum_{n=0}^{\infty} a_n x^n$ converge per ogni $x \in \mathbb{R}$. Se invece $R \in (0, \infty)$ allora la serie converge per ogni $x \in (-R, R)$ e non converge se $|x| > R$.*

In generale, a priori non si può dedurre la convergenza della serie nel caso in cui $|x| = R$. Le due serie $\sum_{n=0}^{\infty} \frac{1}{n+1} x^n$ e $\sum_{n=0}^{\infty} \frac{1}{n^2+1} x^n$ hanno entrambe raggio di convergenza $R = 1$, la prima non converge per $x = 1$, mentre la seconda sì.

Teorema 4.2.9. *Sia $\sum_{n=0}^{\infty} a_n x^n$ una serie di potenze, e sia $R > 0$ il suo raggio di convergenza. Allora $f(x) = \sum_{n=0}^{\infty} a_n x^n$ risulta di classe C^1 su $(-R, R)$, la serie di potenze $g(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$ ha raggio di convergenza R , e $f' = g$ su $(-R, R)$.*

Dimostrazione. Facciamo vedere come prima cosa che $g(x) = \sum_{n=1}^{\infty} n a_n x^{n-1}$ converge in $(-R, R)$; siano $x \in (-R, R)$ tali che $y > |x|$; allora la serie $\sum_{n=0}^{\infty} a_n y^n$ converge, ed è quindi limitata la successione $|a_n y^n| \leq M < +\infty$. Mostriamo che la serie di funzioni $\sum_{n=1}^{\infty} n a_n x^{n-1}$ converge totalmente su $[-r, r]$ con $r = |x|$ (come conseguenza convergerà puntualmente, e quindi anche in x .) Si ha

$$\sum_{n=1}^{\infty} \|n a_n x^{n-1}\|_{\infty, [-r, r]} = \sum_{n=1}^{\infty} |n a_n| \|x^{n-1}\|_{\infty, [-r, r]} = \sum_{n=1}^{\infty} |n a_n| r^n \leq \sum_{n=1}^{\infty} M n \frac{r^{n-1}}{y^{n-1}}$$

e adesso l'ultima serie converge grazie al criterio della radice, e grazie al fatto che $|\frac{r}{y}| < 1$. Abbiamo dunque dedotto che $\sum_{n=1}^{\infty} na_n x^{n-1}$ converge totalmente su $[-r, r]$ per ogni $r < R$, in che implica anche che converge puntualmente per ogni $x < R$ (come conseguenza il raggio di convergenza di $\sum_{n=1}^{\infty} na_n x^{n-1}$ sarà $\geq R$; mostrare per esercizio che non può essere strettamente maggiore). Inoltre, convergenza totale implica quella uniforme, quindi le somme parziali $S'_n(x) = \sum_{i=1}^n ia_i x^{i-1}$ convergono uniformemente a g su $[-r, r]$.

Ora un argomento simile mostra che anche $\sum_{n=0}^{\infty} a_n x^n$ converge totalmente su $[-r, r]$, e quindi anche $S_n(x) = \sum_{i=0}^n a_i x^i$ convergono uniformemente ad f . Essendo S_n ed S'_n continue deduciamo che f e g sono continue su $[-r, r]$. Il Teorema 4.2.4 allora implica $f \in C^1([-r, r])$, che $f' = g$ su $[-r, r]$, ed essendo $r < R$ arbitrario, concludiamo la tesi. $\square$

Visto che la serie $f'(x) = g(x) = \sum_{n=1}^{\infty} na_n x^{n-1}$ ha raggio di convergenza R , come la sua primitiva $f(x)$, possiamo riapplicare il teorema precedente alla serie g , deducendo che essa è C^1 e la sua derivata ha stesso raggio di convergenza. Possiamo iterare il ragionamento per ogni derivata di f , e se ne deduce il seguente:

Teorema 4.2.10. *Sia $f(x) = \sum_{n=0}^{\infty} a_n x^n$ una serie di potenze, e sia $R > 0$ il suo raggio di convergenza. Allora f è di classe C^∞ su $(-R, R)$ e per ogni $k > 0$ si ha*

$$f^k(x) = \sum_{n=k}^{\infty} \frac{n!}{(n-k)!} a_n x^{n-k}.$$

Mostriamo ora un criterio per il calcolo del raggio di convergenza.

Teorema 4.2.11. *Sia $\sum_{n=0}^{\infty} a_n x^n$ una serie di potenze. Supponiamo esista il limite $L := \lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}$; allora*

- (i) *Se $L = 0$ allora il raggio di convergenza R è $+\infty$;*
- (ii) *Se $L = +\infty$ allora $R = 0$ e la serie converge solo per $x = 0$;*
- (iii) *Se $L \in (0, \infty)$, allora si ha $R = \frac{1}{L}$.*

Dimostrazione. Mettiamoci subito nel terzo caso (iii): il criterio della radice applicato alla serie $\sum_{n=0}^{\infty} a_n x^n$ ci garantisce che essa converge se $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n x^n|} < 1$, mentre non convergerà se tale limite fosse > 1 . Tale limite coincide con

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n x^n|} = |x|L.$$

Pertanto se $|x| < \frac{1}{L}$ converge, mentre non convergerà se $|x| > \frac{1}{L}$. Questo implica $R = \frac{1}{L}$. Nel caso (ii) avremo sempre

$$\lim_{n \rightarrow \infty} \sqrt[n]{|a_n x^n|} = +\infty, \quad \forall x \neq 0,$$

e non può esserci convergenza per tali x . In modo analogo (esercizio), si tratta il caso (i). $\square$

4.2.2 Serie di Fourier

Consideriamo le funzioni trigonometriche $\cos(kx)$ e $\sin(kx)$, $k \in \mathbb{N}$; considereremo anche la funzione costante 1. Come notazione globale useremo la seguente: per ogni $k \in \mathbb{Z}$ definiamo $e_k : \mathbb{R} \rightarrow \mathbb{R}$ come

$$e_k(x) = \begin{cases} 1 & \text{se } k = 0, \\ \cos(kx) & \text{se } k > 0, \\ \sin(|k|x) & \text{se } k < 0. \end{cases} \quad (4.2.2)$$

Lemma 4.2.12. *Le funzioni e_k appena definite soddisfano la seguente proprietà:*

$$\int_{-\pi}^{\pi} e_k(x) e_h(x) dx = \begin{cases} 0 & \text{se } k \neq h \\ 2\pi & \text{se } k = h = 0 \\ \pi & \text{se } k = h \neq 0. \end{cases} \quad (4.2.3)$$

Inoltre

$$\int_{-\pi}^{\pi} e_k(x) dx = \begin{cases} 2\pi & \text{se } k = 0 \\ 0 & \text{altrimenti.} \end{cases} \quad (4.2.4)$$

Lasciamo per esercizio la semplice verifica.

Definizione 4.2.13. Sia $f : \mathbb{R} \rightarrow \mathbb{R}$; f è periodica di periodo $T > 0$ se per ogni $x \in \mathbb{R}$ risulta $f(x) = f(x + T)$. Sia $f : [-\pi, \pi) \rightarrow \mathbb{R}$; denoteremo con $\bar{f} : \mathbb{R} \rightarrow \mathbb{R}$ la funzione periodica di periodo 2π definita come

$$\bar{f}(x) = f(x - 2m\pi) \quad \text{se } x \in [(2m - 1)\pi, (2m + 1)\pi) \quad \forall m \in \mathbb{Z}.$$

$\bar{f}$ prende nome di estensione 2π -periodica di f .

Definizione 4.2.14. Una funzione $f : [a, b] \rightarrow \mathbb{R}$ si dirà continua a tratti se esiste una partizione di $[a, b]$ della forma $a = t_0 < t_1 < \dots < t_n = b$ tali che f risulta continua su (t_i, t_{i+1}) per ogni $i = 0, \dots, n - 1$. Diremo che f è C^1 a tratti se è continua a tratti e se su ogni intervallo $[t_i, t_{i+1}]$ valgono le seguenti proprietà:

- esistono i limiti $\lim_{x \rightarrow t_i^+} f(x) =: f(t_i^+)$ e $\lim_{x \rightarrow t_{i+1}^-} f(x) =: f(t_{i+1}^-)$;
- esistono i limiti dei rapporti incrementali

$$f'(t_i^+) := \lim_{h \rightarrow 0^+} \frac{f(t_i + h) - f(t_i^+)}{h} \quad f'(t_{i+1}^-) := \lim_{h \rightarrow 0^-} \frac{f(t_{i+1} + h) - f(t_{i+1}^-)}{h};$$

- la funzione

$$f'(x) := \begin{cases} f'(t_i^+) & \text{se } x = t_i^+ \\ f'(x) & \text{se } x \in (t_i, t_{i+1}) \\ f'(t_{i+1}^-) & \text{se } x = t_{i+1}^- \end{cases}$$

risulta continua su $[t_i, t_{i+1}]$.

Definizione 4.2.15. Sia $f : [-\pi, \pi) \rightarrow \mathbb{R}$ una funzione integrabile: allora definiamo i coefficienti di Fourier di f i valori

$$\begin{aligned} a_0 &:= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) e_0(x) dx = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) dx \\ a_k &:= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) e_k(x) dx = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \cos(kx) dx \quad \text{per } k > 0, \\ b_k &:= \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) e_{-k}(x) dx = \frac{1}{\pi} \int_{-\pi}^{\pi} f(x) \sin(kx) dx \quad \text{per } k > 0. \end{aligned} \quad (4.2.5)$$

Data una funzione integrabile $f : [-\pi, \pi) \rightarrow \mathbb{R}$ si definisce la sua serie di Fourier la seguente

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx)) \quad (4.2.6)$$

Ci chiediamo ora quand'è che la serie di Fourier associata ad $f : [-\pi, \pi) \rightarrow \mathbb{R}$ converge, e in caso affermativo se converge (almeno puntualmente) alla funzione f .

Lemma 4.2.16. Per ogni $t \in \mathbb{R}$, $n \in \mathbb{N}$, risulta

$$\frac{1}{2} + \cos t + \cos(2t) + \cdots + \cos(nt) = \frac{\sin(t(n + \frac{1}{2}))}{2 \sin(\frac{t}{2})}$$

Dimostrazione. Mostriamo il lemma per induzione su n . Per $n = 0$ la formula è corretta, quindi bisogna mostrare il passo induttivo. Supponiamo allora il lemma sia vero per un certo valore di $n \in \mathbb{N}$. Facciamo vedere che

$$\frac{1}{2} + \cos t + \cos(2t) + \cdots + \cos(nt) + \cos((n+1)t) = \frac{\sin(t(n + \frac{3}{2}))}{2 \sin(\frac{t}{2})}. \quad (4.2.7)$$

Per ipotesi induttiva

$$\frac{1}{2} + \cos t + \cos(2t) + \cdots + \cos(nt) + \cos((n+1)t) = \frac{\sin(t(n + \frac{1}{2}))}{2 \sin(\frac{t}{2})} + \cos((n+1)t);$$

il membro destro è allora

$$\begin{aligned} & \frac{\sin(t(n + \frac{1}{2})) + 2 \sin(\frac{t}{2}) \cos((n+1)t)}{2 \sin(\frac{t}{2})} = \frac{\sin(t(n+1) - \frac{t}{2}) + 2 \sin(\frac{t}{2}) \cos((n+1)t)}{2 \sin(\frac{t}{2})} \\ &= \frac{\sin(t(n+1)) \cos(\frac{t}{2}) - \sin(\frac{t}{2}) \cos((n+1)t) + 2 \sin(\frac{t}{2}) \cos((n+1)t)}{2 \sin(\frac{t}{2})} \\ &= \frac{\sin(t(n+1)) \cos(\frac{t}{2}) + \sin(\frac{t}{2}) \cos((n+1)t)}{2 \sin(\frac{t}{2})} = \frac{\sin(t(n+1) + \frac{t}{2})}{2 \sin(\frac{t}{2})} \end{aligned} \quad (4.2.8)$$

che implica la (4.2.7). □

Proposizione 4.2.17 (Disuguaglianza di Bessel). Sia $f : [-\pi, \pi) \rightarrow \mathbb{R}$ una funzione limitata e integrabile e siano a_0, a_k, b_k , $k > 0$, i suoi coefficienti di Fourier. Allora

$$\frac{a_0^2}{2} + \sum_{k>0} (a_k^2 + b_k^2) \leq \frac{1}{\pi} \int_{-\pi}^{\pi} f(x)^2 dx. \quad (4.2.9)$$

Dimostrazione. Denotiamo con $S_n(x) = \frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx))$ la somma parziale di Fourier. Avremo $\int_{-\pi}^{\pi} |f(x) - S_n(x)|^2 dx \geq 0$. Sviluppando il quadrato troviamo

$$\int_{-\pi}^{\pi} f(x)^2 dx \geq 2 \int_{-\pi}^{\pi} f(x) S_n(x) dx - \int_{-\pi}^{\pi} S_n(x)^2 dx. \quad (4.2.10)$$

Da una parte abbiamo

$$\begin{aligned} 2 \int_{-\pi}^{\pi} f(x) S_n(x) dx &= a_0 \int_{-\pi}^{\pi} f(x) dx + \sum_{k=1}^n \left(2a_k \int_{-\pi}^{\pi} f(x) \cos(kx) dx + 2b_k \int_{-\pi}^{\pi} f(x) \sin(kx) dx \right) \\ &= \pi a_0^2 + 2\pi \sum_{k=1}^n (a_k^2 + b_k^2) \end{aligned} \quad (4.2.11)$$

dove si è usata la definizione di coefficienti di Fourier. D'altra parte si ha

$$\begin{aligned} \int_{-\pi}^{\pi} S_n(x)^2 dx &= \int_{-\pi}^{\pi} \left(\frac{a_0}{2} + \sum_{k=1}^n (a_k \cos(kx) + b_k \sin(kx)) \right)^2 dx \\ &= \int_{-\pi}^{\pi} \left(\frac{a_0^2}{4} + \sum_{k=1}^n (a_k^2 \cos^2(kx) + b_k^2 \sin^2(kx)) \right) dx \\ &\quad + \sum_{\substack{i,j=1 \\ i \neq j}}^n \int_{-\pi}^{\pi} (a_i \cos(ix) + b_i \sin(ix))(a_j \cos(jx) + b_j \sin(jx)) dx \\ &\quad + 2 \sum_{k=1}^n a_k b_k \int_{-\pi}^{\pi} \cos(kx) \sin(kx) dx + a_0 \sum_{k=1}^n \int_{-\pi}^{\pi} a_k \cos(kx) + b_k \sin(kx) dx. \end{aligned}$$

Utilizzando ora il Lemma 4.2.12 l'ultima espressione ci dà

$$\int_{-\pi}^{\pi} S_n(x)^2 dx = \frac{\pi}{2} a_0^2 + \pi \sum_{k=1}^n (a_k^2 + b_k^2).$$

Questa, insieme alla (4.2.11), implica nella (4.2.10)

$$\int_{-\pi}^{\pi} f(x)^2 dx \geq \frac{\pi}{2} a_0^2 + \pi \sum_{k=1}^n (a_k^2 + b_k^2), \quad (4.2.12)$$

che è valida per ogni $n \in \mathbb{N}$. Quindi possiamo passare al limite per $n \rightarrow \infty$ e ottenere la tesi. $\square$

Osservazione 4.2.18. Come conseguenza della disuguaglianza di Bessel otteniamo che la serie dei quadrati dei coefficienti di Fourier è convergente, e in particolare infinitesima:

$$a_k \rightarrow 0, \quad b_k \rightarrow 0 \quad \text{quando } k \rightarrow \infty.$$

In particolare, otteniamo che se f è una funzione integrabile e limitata allora

$$\int_{-\pi}^{\pi} f(x) \cos(kx) dx \rightarrow 0, \quad \int_{-\pi}^{\pi} f(x) \sin(kx) dx \rightarrow 0 \quad \text{quando } k \rightarrow \infty.$$

Possiamo ora dimostrare il seguente Teorema:

Teorema 4.2.19. *Sia $f : [-\pi, \pi) \rightarrow \mathbb{R}$ una funzione C^1 a tratti limitata. Allora la serie di Fourier $\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx))$ di f converge puntualmente a $\frac{f(x^+) + f(x^-)}{2}$, dove*

$$f(x^+) = \lim_{y \rightarrow x^+} f(y), \quad f(x^-) = \lim_{y \rightarrow x^-} f(y).$$

Dimostrazione. Ricordando la definizione dei coefficienti di Fourier possiamo scrivere

$$\begin{aligned} S_n(x) &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \left(\frac{1}{2} + \sum_{i=1}^n \cos(ix) \cos(iy) + \sin(ix) \sin(iy) \right) dy \\ &= \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \left(\frac{1}{2} + \sum_{i=1}^n \cos(i(y-x)) \right) dy = \frac{1}{\pi} \int_{-\pi}^{\pi} f(y) \frac{\sin((n+\frac{1}{2})(y-x))}{2 \sin(\frac{y-x}{2})} dy \\ &= \frac{1}{\pi} \int_{-\pi-x}^{\pi-x} f(t+x) \frac{\sin((n+\frac{1}{2})t)}{2 \sin(\frac{t}{2})} dt \end{aligned} \quad (4.2.13)$$

dove abbiamo usato il Lemma 4.2.16 e il cambio di variabili $y-x=t$. Usando la definizione di estensione 2π -periodica $\bar{f}$ possiamo scrivere l'ultimo termine come

$$S_n(x) = \frac{1}{\pi} \int_{-\pi}^{\pi} \bar{f}(x+t) \frac{\sin((n+\frac{1}{2})t)}{2 \sin(\frac{t}{2})} dt.$$

Osserviamo ora che, sempre per il Lemma 4.2.16,

$$\int_0^{\pi} \frac{\sin((n+\frac{1}{2})t)}{2 \sin(\frac{t}{2})} dt = \frac{\pi}{2} = \int_{-\pi}^0 \frac{\sin((n+\frac{1}{2})t)}{2 \sin(\frac{t}{2})} dt.$$

Quindi anche

$$\frac{1}{\pi} \int_0^{\pi} f(x^+) \frac{\sin((n+\frac{1}{2})t)}{2 \sin(\frac{t}{2})} dt + \frac{1}{\pi} \int_{-\pi}^0 f(x^-) \frac{\sin((n+\frac{1}{2})t)}{2 \sin(\frac{t}{2})} dt = \frac{f(x^+) + f(x^-)}{2}.$$

Pertanto abbiamo ottenuto

$$\begin{aligned} S_n(x) - \frac{f(x^+) + f(x^-)}{2} &= \frac{1}{\pi} \int_0^{\pi} \frac{\bar{f}(x+t) - f(x^+)}{2 \sin(\frac{t}{2})} \sin((n+\frac{1}{2})t) dt \\ &\quad + \frac{1}{\pi} \int_{-\pi}^0 \frac{\bar{f}(x+t) - f(x^-)}{2 \sin(\frac{t}{2})} \sin((n+\frac{1}{2})t) dt. \end{aligned} \quad (4.2.14)$$

Definendo

$$G(x, t) := \begin{cases} \frac{\bar{f}(x+t) - f(x^+)}{2 \sin(\frac{t}{2})} & \text{se } t \in [0, \pi) \\ \frac{\bar{f}(x+t) - f(x^-)}{2 \sin(\frac{t}{2})} & \text{se } t \in [-\pi, 0), \end{cases}$$

l'ultima espressione è pari a

$$S_n(x) - \frac{f(x^+) + f(x^-)}{2} = \frac{1}{\pi} \int_{-\pi}^{\pi} G(x, t) \sin((n+\frac{1}{2})t) dt.$$

Dalla definizione di G segue subito che essa è continua a tratti sugli intervalli aperti $(-\pi, 0)$ e $(0, \pi)$. Inoltre

$$\lim_{t \rightarrow 0^+} G(x, t) = \lim_{t \rightarrow 0^+} \frac{\bar{f}(x+t) - f(x^+)}{t} \frac{\frac{t}{2}}{\sin(\frac{t}{2})} = f'(x^+),$$

e anche

$$\lim_{t \rightarrow 0^-} G(x, t) = \lim_{t \rightarrow 0^-} \frac{\bar{f}(x+t) - f(x^-)}{t} \frac{\frac{t}{2}}{\sin(\frac{t}{2})} = f'(x^-),$$

che implica che G è limitata e continua a tratti su tutto $[-\pi, \pi)$. Come conseguenza essa è integrabile e limitata; scrivendo allora

$$\sin((n + \frac{1}{2})t) = \cos(nt) \sin(t/2) + \sin(nt) \cos(t/2),$$

avremo

$$S_n(x) - \frac{f(x^+) + f(x^-)}{2} = \frac{1}{\pi} \int_{-\pi}^{\pi} G(x, t) \cos(nt) \sin(t/2) dt + \frac{1}{\pi} \int_{-\pi}^{\pi} G(x, t) \sin(nt) \cos(t/2) dt,$$

ed essendo anche le funzioni $G(x, t) \sin(t/2)$ e $G(x, t) \cos(t/2)$ integrabili e limitate su $[-\pi, \pi)$, concludiamo dall'Osservazione 4.2.18 che

$$S_n(x) - \frac{f(x^+) + f(x^-)}{2} \rightarrow 0, \text{ quando } n \rightarrow \infty. \quad (4.2.15)$$

Questa è la tesi. □

Raffiniamo ora la convergenza appena ottenuta.

Teorema 4.2.20. *Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ una funzione di classe C^1 e 2π -periodica. Allora la sua serie di Fourier converge totalmente su $[-\pi, \pi)$ (e quindi anche uniformemente ad f).*

Dimostrazione. Notiamo anzitutto che $f' : \mathbb{R} \rightarrow \mathbb{R}$ sarà una funzione continua e 2π -periodica, quindi per la disuguaglianza di Bessel

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f'(x)^2 dx \geq \frac{\alpha_0^2}{2} + \sum_{k \geq 1} (\alpha_k^2 + \beta_k^2),$$

dove abbiamo indicato con $\alpha_0, \alpha_k, \beta_k$ i coefficienti di Fourier di f' . Possiamo ora ottenere una relazione tra questi coefficienti e i coefficienti di Fourier di f : si ha

$$\begin{aligned} \pi \alpha_0 &= \int_{-\pi}^{\pi} f'(x) dx = f(\pi) - f(-\pi) = 0, \\ \pi a_k &= \int_{-\pi}^{\pi} f(x) \cos(kx) dx = \frac{1}{k} f(x) \sin(kx) \Big|_{x=-\pi}^{x=\pi} - \frac{1}{k} \int_{-\pi}^{\pi} f'(x) \sin(kx) dx = -\pi \frac{\beta_k}{k} \\ \pi b_k &= \int_{-\pi}^{\pi} f(x) \sin(kx) dx = -\frac{1}{k} f(x) \cos(kx) \Big|_{x=-\pi}^{x=\pi} + \frac{1}{k} \int_{-\pi}^{\pi} f'(x) \cos(kx) dx = \pi \frac{\alpha_k}{k}, \end{aligned}$$

da cui $|a_k| = \frac{|\beta_k|}{k}$ e $|b_k| = \frac{|\alpha_k|}{k}$ per ogni $k > 0$.

Dobbiamo ora mostrare che

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} (a_k \cos(kx) + b_k \sin(kx))$$

converge totalmente su $[-\pi, \pi)$. A questo scopo stimiamo

$$\|a_k \cos(kx) + b_k \sin(kx)\|_{\infty, [-\pi, \pi]} \leq |a_k| + |b_k|,$$

e quindi

$$\begin{aligned} \sum_{k=1}^{\infty} \|a_k \cos(kx) + b_k \sin(kx)\|_{\infty, [-\pi, \pi]} &\leq \sum_{k=1}^{\infty} (|a_k| + |b_k|) = \sum_{k=1}^{\infty} \frac{1}{k} (|\alpha_k| + |\beta_k|) \\ &\leq \frac{1}{2} \sum_{k=1}^{\infty} \frac{2}{k^2} + |\alpha_k|^2 + |\beta_k|^2 < +\infty. \end{aligned}$$

Questo conclude la dimostrazione. $\square$

Corollario 4.2.21. *Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ una funzione di classe C^1 e 2π -periodica. Allora la disuguaglianza di Bessel è un'uguaglianza, ovvero*

$$\frac{1}{\pi} \int_{-\pi}^{\pi} f(x)^2 dx = \frac{a_0^2}{2} + \sum_{k \geq 1} (a_k^2 + b_k^2). \quad (4.2.16)$$

L'uguaglianza (4.2.16) prende il nome di identità di Parseval.

Dimostrazione. Dal Teorema 4.2.20 deduciamo che S_n , somma parziale di Fourier relativa ad f , converge ad f uniformemente su $[-\pi, \pi)$. Di conseguenza anche S_n^2 tende uniformemente ad f^2 , da cui, per il Lemma 4.1.13, abbiamo

$$\lim_{n \rightarrow \infty} \int_{-\pi}^{\pi} S_n(x)^2 dx = \int_{-\pi}^{\pi} f(x)^2 dx.$$

La tesi allora segue osservando che $\int_{-\pi}^{\pi} S_n(x)^2 dx = \frac{a_0^2}{2} + \sum_{k=1}^n (a_k^2 + b_k^2)$. $\square$

Capitolo 5

Curve e superfici

5.1 Curve

Definizione 5.1.1. Una curva parametrizzata in $\mathbb{R}^n$ è una funzione $\gamma : [a, b] \rightarrow \mathbb{R}^n$ di classe C^1 . Diremo che la curva è di classe C^1 sull'intervallo chiuso $[a, b]$, e scriveremo $\gamma \in C^1([a, b])$, se

- esistono i limiti dei rapporti incrementali

$$\gamma'(a) := \lim_{t \rightarrow a^+} \frac{\gamma(t) - \gamma(a)}{t - a} \quad \gamma'(b) := \lim_{t \rightarrow b^-} \frac{\gamma(t) - \gamma(b)}{t - b};$$

- la funzione $\gamma' : [a, b] \rightarrow \mathbb{R}^n$ è continua su $[a, b]$;
- per ogni $t \in (a, b)$ risulta $\gamma'(t) \neq 0$.

Data una curva regolare $\gamma : [a, b] \rightarrow \mathbb{R}^n$ chiameremo talvolta velocità della curva la sua derivata $\gamma'(t)$.

Un esempio già visto di curva regolare è quello delle rette parametriche: dati $p, v \in \mathbb{R}^n$, $v \neq 0$, la funzione $\gamma : \mathbb{R} \rightarrow \mathbb{R}^n$ data da $\gamma(t) = p + tv$ è una curva regolare (con $[a, b] = \mathbb{R}$) dato che $\gamma'(t) = v \neq 0$.

Definizione 5.1.2. Data una curva regolare $\gamma : [a, b] \rightarrow \mathbb{R}^n$ chiameremo sostegno della curva l'insieme $\gamma([a, b]) \in \mathbb{R}^n$.

Un altro esempio di curve regolari sono i grafici di funzioni (opportunamente parametrizzati, come segue): sia $f : [a, b] \rightarrow \mathbb{R}^n$ una funzione di classe C^1 su $[a, b]$; allora la curva $\gamma : [a, b] \rightarrow \mathbb{R}^{n+1}$ definita come

$$\gamma(t) := (t, f(t)), \quad t \in [a, b], \quad (5.1.1)$$

è una curva regolare, dal momento che sarà di classe $C^1([a, b])$ (verificarlo) e la derivata

$$\gamma'(t) = (1, f'(t)), \quad t \in [a, b],$$

non si annulla mai (la sua prima componente è sempre $\neq 0$). In questo caso il sostegno della curva γ coinciderà con il grafico della funzione f , infatti

$$\gamma([a, b]) = \{(x, y) \in [a, b] \times \mathbb{R}^n : y = f(x)\}.$$

In questo caso specifico, osserviamo facilmente che la velocità della curva rappresenta un vettore tangente al sostegno della curva stessa. Precisamente, dato $\bar{t} \in [a, b]$, il vettore $\gamma'(\bar{t}) = (1, f'(\bar{t}))$ è tangente a $\gamma([a, b])$ nel punto $\gamma(\bar{t}) = (\bar{t}, f(\bar{t}))$, e la retta

$$r(t) := \gamma(\bar{t}) + t\gamma'(\bar{t}) = (\bar{t} + t, f(\bar{t}) + tf'(\bar{t}))$$

ha come sostegno la retta tangente al $\gamma([a, b])$ nel punto $\gamma(\bar{t})$.

Mostriamo ora che tale proprietà geometrica vale per ogni curva regolare, non solo per i grafici di funzione. A questo scopo mostriamo il seguente:

Proposizione 5.1.3. *Sia $\gamma : [a, b] \rightarrow \mathbb{R}^2$ una curva regolare, e sia $t \in (a, b)$. Allora esiste un intorno $I \subset (a, b)$ di t tale che γ può essere espressa come il grafico di una funzione f della variabile x , oppure di una funzione g della variabile y .*

Dimostrazione. Essendo γ regolare abbiamo $\gamma'(t) \neq 0$, e quindi una delle due componenti di $\gamma'(t)$ non si annulla. Supponiamo per esempio $\gamma'_2(t) \neq 0$, e come tale assumiamo $\gamma'_2(t) > 0$. Esisterà allora un intorno I di t tale che $\gamma'_2(s) > 0$ per ogni $s \in I$, e quindi la funzione $\gamma_2 : I \rightarrow \mathbb{R}$ è strettamente crescente, e pertanto invertibile. Essendo I un intervallo, anche $J := \gamma_2(I)$ è un intervallo, e $\gamma_2^{-1} : J \rightarrow I$ risulterà crescente. Chiamando $y = \gamma_2(s)$, avremo $s = \gamma_2^{-1}(y)$, e allora $\gamma(s) = (\gamma_1(s), \gamma_2(s)) = (\gamma_1(\gamma_2^{-1}(y)), y) = \gamma \circ \gamma_2^{-1}(y)$ per ogni $s \in I$. Perciò, la mappa $\hat{\gamma} := \gamma \circ \gamma_2^{-1} : J \rightarrow \mathbb{R}^2$ ha stesso sostegno di γ su I , e definendo $g := \gamma_1 \circ \gamma_2^{-1}$ si avrà

$$\hat{\gamma}(y) := (g(y), y) \quad \forall y \in J.$$

In modo analogo si può procedere nel caso in cui $\gamma'_1(t) \neq 0$ e mostrare che esiste una funzione $f : J \rightarrow \mathbb{R}$ tale che $\gamma(I)$ coincide con il sostegno di

$$\hat{\gamma}(x) := (x, f(x)) \quad \forall x \in J.$$

□

Essendo quindi ogni curva regolare “localmente” grafico di funzioni, risulterà che $\gamma'(t)$ rappresenta un vettore tangente alla curva nel punto $\gamma(t)$.

Esempio 5.1.4. Prendiamo la curva $\gamma(t) := (t^3, |t|^3)$ per $t \in \mathbb{R}$. Risulta che γ è di classe C^1 su $\mathbb{R}$, ma non è una curva regolare in quanto

$$\gamma'(t) = (3t^2, 3\frac{t^3}{|t|})$$

si annulla (entrambe le componenti si annullano) in $t = 0$. Inoltre, se denotiamo con $x := t^3$, ci accorgiamo che la curva γ ha come sostegno il grafico della funzione $f(x) = |x|$, in quanto

$$\gamma(t) = (x, |x|), \quad \forall x \in \mathbb{R}.$$

Tale grafico non ammette vettore tangente nel punto $(0, 0)$, e infatti in tale punto la derivata $\gamma'(0)$ si annulla.

Definizione 5.1.5. *Sia $\gamma : [a, b] \rightarrow \mathbb{R}^n$ una curva regolare, allora per ogni $t \in [a, b]$ si definisce il versore tangente alla curva in $\gamma(t)$ il vettore*

$$\tau_\gamma(t) := \frac{\gamma'(t)}{\|\gamma'(t)\|}.$$

Definizione 5.1.6. Siano $\gamma : [a, b] \rightarrow \mathbb{R}^n$ e $\sigma : [c, d] \rightarrow \mathbb{R}^n$ due curve. Diremo che esse sono equivalenti se esiste una funzione $p : [a, b] \rightarrow [c, d]$ biunivoca e di classe C^1 , tale che $p'(t) \neq 0$ per ogni t , e

$$\gamma(t) = \sigma(p(t)) \quad \forall t \in [a, b].$$

Diremo che γ e σ hanno stessa direzione se $p'(t) > 0$ per ogni t , mentre direzione opposta se $p'(t) < 0$ per ogni t .

Lemma 5.1.7. Siano $\gamma : [a, b] \rightarrow \mathbb{R}^n$ e $\sigma : [c, d] \rightarrow \mathbb{R}^n$ due curve regolari. Allora se hanno stessa direzione si avrà $\tau_\gamma(t) = \tau_\sigma(p(t))$ per ogni t , mentre se hanno direzione opposta avremo $\tau_\gamma(t) = -\tau_\sigma(p(t))$.

Dimostrazione. Nel caso in cui γ e σ hanno stessa direzione si avrà, dalla definizione di equivalenza,

$$\gamma(t) = \sigma(p(t)) \Rightarrow \gamma'(t) = \sigma'(p(t))p'(t),$$

da cui $\|\gamma'(t)\| = \|\sigma'(p(t))\||p'(t)|$, e quindi

$$\tau_\gamma(t) = \frac{\gamma'(t)}{\|\gamma'(t)\|} = \frac{\sigma'(p(t))}{\|\sigma'(p(t))\|} \cdot \frac{p'(t)}{|p'(t)|} = \frac{\sigma'(p(t))}{\|\sigma'(p(t))\|} = \tau_\sigma(p(t)).$$

Nel caso in cui la direzione è opposta si avrà $\frac{p'(t)}{|p'(t)|} = -1$, e quindi

$$\tau_\gamma(t) = -\frac{\sigma'(p(t))}{\|\sigma'(p(t))\|} = -\tau_\sigma(p(t)).$$

□

Definizione 5.1.8. Sia $\gamma : [a, b] \rightarrow \mathbb{R}^n$ una curva di classe C^1 , allora si definisce la lunghezza di γ il valore $L(\gamma)$ dato da

$$L(\gamma) := \int_a^b \|\gamma'(t)\| dt.$$

Prendiamo come esempio una curva $\gamma : [0, 1] \rightarrow \mathbb{R}^n$ che parametrizza il segmento di vertici $P, Q \in \mathbb{R}^n$. Avremo, ad esempio,

$$\gamma(t) := Pt + Q(1 - t)$$

in modo che $\gamma(0) = Q$, $\gamma(1) = P$, γ è regolare e $\gamma'(t) = P - Q$. In particolare, avremo

$$L(\gamma) = \int_0^1 \|\gamma'(t)\| dt = \int_0^1 \|P - Q\| dt = \|P - Q\|,$$

che è la lunghezza del segmento $\overline{PQ}$.

Vediamo ora che la lunghezza di una curva non dipende dalla specifica parametrizzazione, ma solo dal sostegno della stessa.

Lemma 5.1.9. Siano $\gamma : [a, b] \rightarrow \mathbb{R}^n$ e $\sigma : [c, d] \rightarrow \mathbb{R}^n$ due curve di classe C^1 equivalenti. Allora $L(\gamma) = L(\sigma)$.

Dimostrazione. Supponiamo $p : [a, b] \rightarrow [c, d]$ sia strettamente decrescente, e quindi $p'(t) < 0$ per ogni t (il caso contrario si tratta in maniera analoga). Allora, dovendo essere bigettiva, p soddisferà $p(a) = d$ e $p(b) = c$. Pertanto

$$\int_a^b \|\gamma'(t)\| dt = \int_a^b \left\| \frac{d}{dt} \sigma(p(t)) \right\| dt = \int_a^b \|\sigma'(p(t))p'(t)\| dt = - \int_a^b \|\sigma'(p(t))\| p'(t) dt$$

e facendo ora il cambio di variabile $p(t) = s$ avremo

$$\int_a^b \|\gamma'(t)\| dt = - \int_d^c \|\sigma'(s)\| ds = \int_c^d \|\sigma'(s)\| ds,$$

che è la tesi. $\square$

Lemma 5.1.10. *Sia $\gamma : [a, b] \rightarrow \mathbb{R}^n$ una curva di classe C^1 , allora*

$$\left\| \int_a^b \gamma'(t) dt \right\| = \|\gamma(b) - \gamma(a)\| \leq \int_a^b \|\gamma'(t)\| dt.$$

Dimostrazione. Se $\|\gamma(b) - \gamma(a)\| = 0$ non c'è nulla da dimostrare, quindi assumiamo $\|\gamma(b) - \gamma(a)\| > 0$. In tal caso, per ogni vettore $v \in \mathbb{R}^n$ avremo

$$v \cdot (\gamma(b) - \gamma(a)) = v \cdot \left(\int_a^b \gamma'(t) dt \right) = \int_a^b v \cdot \gamma'(t) dt \leq \int_a^b \|v\| \|\gamma'(t)\| dt = \|v\| \int_a^b \|\gamma'(t)\| dt,$$

e ponendo $v = \gamma(b) - \gamma(a)$ si ottiene subito la tesi. $\square$

Definizione 5.1.11. *Sia $\gamma : [a, b] \rightarrow \mathbb{R}^n$ una curva di classe C^1 , e siano $t_i \in [a, b]$ punti tali che $a = t_0 < t_1 < \dots < t_k = b$. La famiglia $\{t_i : i = 0, \dots, k\}$ prende nome di partizione di $[a, b]$. Per ogni partizione di $[a, b]$ si definisce la spezzata $\sigma : [a, b] \rightarrow \mathbb{R}^n$ di vertici $\gamma(t_i)$ la curva definita come*

$$\sigma(t) = \gamma(t_i) + \frac{t - t_i}{t_{i+1} - t_i} (\gamma(t_{i+1}) - \gamma(t_i)) \quad \text{se } t \in [t_i, t_{i+1}], \quad (5.1.2)$$

per ogni $i = 0, \dots, k-1$. La lunghezza della spezzata σ si definisce come

$$L(\sigma) := \sum_{i=0}^{k-1} L(\sigma|_{[t_i, t_{i+1}]}) = \sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} \|\sigma'(t)\| dt = \sum_{i=0}^{k-1} \|\gamma(t_{i+1}) - \gamma(t_i)\|.$$

Si noti che una spezzata non è in generale di classe C^1 .

Teorema 5.1.12. *Sia $\gamma : [a, b] \rightarrow \mathbb{R}^n$ una curva di classe C^1 , allora per ogni $k \in \mathbb{N}$ e ogni partizione $\{t_i : i = 0, \dots, k\}$, la spezzata σ di vertici $\gamma(t_i)$ soddisfa $L(\sigma) \leq L(\gamma)$. Inoltre, vale*

$$L(\gamma) = \sup\{L(\sigma) : k \in \mathbb{N}, t_i : i = 0, \dots, k, \sigma \text{ è la spezzata di vertici } \gamma(t_i)\}.$$

Per dimostrare il teorema premettiamo il seguente lemma:

Lemma 5.1.13. *Sia $v : [\alpha, \beta] \rightarrow \mathbb{R}^n$ una funzione continua. Allora per ogni $\tau \in (\alpha, \beta)$ si ha*

$$\int_\alpha^\beta \|v(t)\| dt \leq \left\| \int_\alpha^\beta v(t) dt \right\| + 2 \int_\alpha^\beta \|v(t) - v(\tau)\| dt. \quad (5.1.3)$$

Dimostrazione. Fissato $\tau \in (\alpha, \beta)$, abbiamo

$$v(\tau)(\beta - \alpha) = \int_{\alpha}^{\beta} v(\tau) dt = \int_{\alpha}^{\beta} v(t) dt + \int_{\alpha}^{\beta} (v(\tau) - v(t)) dt$$

da cui, usando anche il Lemma 5.1.10,

$$\|v(\tau)\|(\beta - \alpha) \leq \left\| \int_{\alpha}^{\beta} v(t) dt \right\| + \int_{\alpha}^{\beta} \|v(\tau) - v(t)\| dt. \quad (5.1.4)$$

D'altra parte, la disuguaglianza triangolare implica che $\|v(\tau)\| \geq \|v(t)\| - \|v(\tau) - v(t)\|$, e quindi

$$\|v(\tau)\|(\beta - \alpha) = \int_{\alpha}^{\beta} \|v(\tau)\| dt \geq \int_{\alpha}^{\beta} \|v(t)\| dt - \int_{\alpha}^{\beta} \|v(\tau) - v(t)\| dt.$$

Mettendo insieme quest'ultima disuguaglianza con (5.1.4) concludiamo

$$\int_{\alpha}^{\beta} \|v(t)\| dt - \int_{\alpha}^{\beta} \|v(\tau) - v(t)\| dt \leq \left\| \int_{\alpha}^{\beta} v(t) dt \right\| + \int_{\alpha}^{\beta} \|v(\tau) - v(t)\| dt,$$

ovvero la tesi. $\square$

Dimostrazione del Teorema 5.1.12. Utilizzando sempre il Lemma 5.1.10 abbiamo

$$L(\sigma) = \sum_{i=0}^{k-1} \|\gamma(t_{i+1}) - \gamma(t_i)\| \leq \sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} \|\gamma'(t)\| dt = \int_a^b \|\gamma'(t)\| dt = L(\gamma),$$

per ogni spezzata σ . Di conseguenza avremo anche

$$\sup\{L(\sigma) : k \in \mathbb{N}, t_i : i = 0, \dots, k, \sigma \text{ è la spezzata di vertici } \gamma(t_i)\} \leq L(\gamma).$$

Mostriamo la disuguaglianza opposta e fissiamo $\varepsilon > 0$ arbitrario; essendo γ' continua su $[a, b]$ sarà uniformemente continua ed esisterà $\delta > 0$ tale che se $|t - s| < \delta$ allora $\|\gamma'(t) - \gamma'(s)\| \leq \varepsilon$. Prendiamo allora una partizione t_i , con $i = 0, \dots, k$, di $[a, b]$ tale che $t_{i+1} - t_i < \delta$ per ogni $i = 0, \dots, k$. Come conseguenza avremo

$$\|\gamma'(t) - \gamma'(s)\| \leq \varepsilon, \quad \forall t, s \in [t_i, t_{i+1}], \quad \forall i = 0, \dots, k-1.$$

Per il Lemma 5.1.13 avremo

$$\begin{aligned} \int_{t_i}^{t_{i+1}} \|\gamma'(t)\| dt &\leq \left\| \int_{t_i}^{t_{i+1}} \gamma'(t) dt \right\| - 2 \int_{t_i}^{t_{i+1}} \|\gamma'(t) - \gamma'(\tau_i)\| dt \\ &= \|\gamma(t_{i+1}) - \gamma(t_i)\| - 2 \int_{t_i}^{t_{i+1}} \|\gamma'(t) - \gamma'(\tau_i)\| dt \\ &\geq \|\gamma(t_{i+1}) - \gamma(t_i)\| - 2\varepsilon(t_{i+1} - t_i), \end{aligned}$$

dove $\tau_i \in (t_i, t_{i+1})$ è un qualsiasi punto. Sommando su $i = 0, \dots, k-1$, otteniamo

$$L(\gamma) = \sum_{i=0}^{k-1} \int_{t_i}^{t_{i+1}} \|\gamma'(t)\| dt \leq \sum_{i=0}^{k-1} \|\gamma(t_{i+1}) - \gamma(t_i)\| - 2\varepsilon = L(\sigma) - 2\varepsilon.$$

Prendendo quindi l'estremo superiore su tutte le possibili partizioni σ abbiamo ottenuto

$$L(\gamma) + 2\varepsilon \leq \sup\{L(\sigma) : k \in \mathbb{N}, t_i : i = 0, \dots, k, \sigma \text{ è la spezzata di vertici } \gamma(t_i)\},$$

che implica la tesi essendo $\varepsilon > 0$ arbitrario. $\square$

5.2 Superfici

Definizione 5.2.1. *Dati due vettori $a, b \in \mathbb{R}^3$ si definisce il loro prodotto vettoriale $a \times b$ il vettore di $\mathbb{R}^3$ che ha componenti*

$$a \times b = (a_2b_3 - a_3b_2, a_3b_1 - a_1b_3, a_1b_2 - a_2b_1).$$

Un modo di ricordarsi il risultato del prodotto vettore $a \times b$ è calcolando il determinante della matrice 3×3

$$M = \begin{pmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ \bar{e}_1 & \bar{e}_2 & \bar{e}_3 \end{pmatrix} \quad (5.2.1)$$

che produce

$$\det(M) = (a_2b_3 - a_3b_2)\bar{e}_1 + (a_3b_1 - a_1b_3)\bar{e}_2 + (a_1b_2 - a_2b_1)\bar{e}_3,$$

in modo che se $\bar{e}_i$ rappresenta l' i -esimo vettore della base canonica, il vettore ottenuto coincide con $a \times b$.

Elenchiamo ora alcune proprietà del prodotto vettoriale:

- (1) per ogni $a, b \in \mathbb{R}^3$ si ha

$$a \times b = -b \times a;$$

- (2) se $a \times b = (0, 0, 0)$ allora a è parallelo a b . Ovvero, se $a \neq 0$, esiste $\lambda \in \mathbb{R}$ tale che $\lambda a = b$;

- (3) per ogni $a, b, c \in \mathbb{R}^3$ si ha

$$(a \times b) \cdot c = (b \times c) \cdot a = (c \times a) \cdot b.$$

Lasciamo per esercizio la verifica delle proprietà sopraelencate. Vediamo invece alcune conseguenze di esse: dalla proprietà (1) segue che $a \times a = -a \times a$ e quindi

$$a \times a = (0, 0, 0)$$

per ogni $a \in \mathbb{R}^3$. Similmente, se $\lambda \in \mathbb{R}$ risulterà anche

$$(\lambda a) \times a = (0, 0, 0).$$

Dalla proprietà (3) prendendo $c = a$ risulta

$$(a \times b) \cdot a = (b \times a) \cdot a = (a \times a) \cdot b = 0,$$

questo mostra che il vettore $a \times b$ è ortogonale ad a , e in modo analogo sarà anche ortogonale a b .

Ricordiamo che un insieme $A \subset \mathbb{R}^2$ si dice sconnesso se esistono due aperti $U, V \subset \mathbb{R}^2$ tali che

- $U \cap A \neq \emptyset \neq V \cap A$;

- $U \cap V = \emptyset$;
- $A \subseteq U \cup V$.

L'insieme A si dirà connesso che non è sconnesso.

Definizione 5.2.2. Sia $A \subset \mathbb{R}^2$ un aperto limitato e connesso. Denotiamo con $K := \overline{A}$ la chiusura di A (che risulterà compatta). Una superficie parametrizzata è una mappa $\varphi : K \rightarrow \mathbb{R}^3$ che soddisfa le seguenti proprietà:

- $\varphi \in C^1(A; \mathbb{R}^3)$;
- $\varphi \in C^0(K; \mathbb{R}^3)$.

Diremo che φ è una superficie regolare se valgono anche:

- (i) φ è iniettiva su A ;
- (ii) $\frac{\partial \varphi}{\partial x} \times \frac{\partial \varphi}{\partial y} \neq (0, 0, 0)$ su tutto A .

Infine, chiameremo sostegno della curva l'immagine di K attraverso φ , ovvero l'insieme $\varphi(K) \subset \mathbb{R}^3$.

Per le superfici regolari esiste la seguente definizione di area:

Definizione 5.2.3. Sia $\varphi : K \rightarrow \mathbb{R}^3$ una superficie regolare. L'area di φ si definisce come quel valore $a(\varphi)$ ottenuto come

$$a(\varphi) := \int_K \left\| \frac{\partial \varphi}{\partial x} \times \frac{\partial \varphi}{\partial y} \right\| dx dy. \quad (5.2.2)$$

Consideriamo ora una superficie regolare $\varphi : K \rightarrow \mathbb{R}^3$ e se $(x_0, y_0) \in A$ è un punto assegnato, prendiamo un intorno rettangolare di esso $(x_0 - \delta, x_0 + \delta) \times (y_0 - \eta, y_0 + \eta)$ che sia contenuto in A . Allora possiamo considerare le curve $\gamma : (x_0 - \delta, x_0 + \delta) \rightarrow \mathbb{R}^3$ e $\sigma : (y_0 - \eta, y_0 + \eta) \rightarrow \mathbb{R}^3$ come

$$\gamma(t) := \varphi(t, y_0), \quad \sigma(s) := \varphi(x_0, s).$$

Entrambe le curve hanno sostegno nel sostegno della superficie φ e passano per il punto $\varphi(x_0, y_0)$. Infatti $\gamma(x_0) = \sigma(y_0) = \varphi(x_0, y_0)$; in particolare avremo che il vettore tangente alla curva γ in $\varphi(x_0, y_0)$ sarà anche tangente alla superficie, e similmente il vettore tangente a σ (perchè le curve sono contenute nella superficie). Risulta quindi che i due vettori

$$\gamma'(x_0) = \frac{d}{dt} \varphi(t, y_0)|_{t=x_0} = \frac{\partial \varphi}{\partial x}(x_0, y_0), \quad (5.2.3)$$

$$\sigma'(y_0) = \frac{d}{ds} \varphi(x_0, s)|_{s=y_0} = \frac{\partial \varphi}{\partial y}(x_0, y_0), \quad (5.2.4)$$

sono tangenti alla superficie nel punto $\varphi(x_0, y_0)$. Questo ci permette di definire il piano tangente alla superficie nel punto $\varphi(x_0, y_0)$, che sarà a sua volta una superficie che denotiamo

con $T_{\varphi(x_0, y_0)}S$, dove $S = \varphi(K)$ è il sostegno di φ . Allora $T_{\varphi(x_0, y_0)}S$ sarà definito come il piano passante per $\varphi(x_0, y_0)$ e contenente i vettori $\frac{\partial \varphi}{\partial x}(x_0, y_0)$ e $\frac{\partial \varphi}{\partial y}(x_0, y_0)$; ovvero

$$T_{\varphi(x_0, y_0)}S(u, v) := \varphi(x_0, y_0) + u \frac{\partial \varphi}{\partial x}(x_0, y_0) + v \frac{\partial \varphi}{\partial y}(x_0, y_0), \quad (u, v) \in \mathbb{R}^2. \quad (5.2.5)$$

Questo piano avrà anche un vettore ortogonale (in realtà infiniti vettori ortogonali, che però avranno stessa direzione a meno di un segno). Ricordando che $a \times b$ è ortogonale sia ad a che a b , e considerando che i vettori $\frac{\partial \varphi}{\partial x}(x_0, y_0)$ e $\frac{\partial \varphi}{\partial y}(x_0, y_0)$ sono contenuti nel piano (sono paralleli ad esso), avremo che un vettore normale al piano $T_{\varphi(x_0, y_0)}S$ sarà $\frac{\partial \varphi}{\partial x}(x_0, y_0) \times \frac{\partial \varphi}{\partial y}(x_0, y_0)$. Ovvero

$$\frac{\partial \varphi}{\partial x}(x_0, y_0) \times \frac{\partial \varphi}{\partial y}(x_0, y_0) \perp T_{\varphi(x_0, y_0)}S. \quad (5.2.6)$$

5.3 Massimi e minimi vincolati

Veniamo ora al problema di studiare massimi e minimi di funzioni relativamente a sottoinsiemi specifici dei propri domini.

Teorema 5.3.1 (Dini). *Sia $F : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ una funzione di classe C^1 e sia $(x_0, y_0) \in A$ un punto tale che $F(x_0, y_0) = 0$. Supponiamo che $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$, allora esiste un rettangolo $I \times J := [x_0 - \delta, x_0 + \delta] \times [y_0 - \eta, y_0 + \eta]$ incluso in A tale che per ogni $x \in I$ esiste un unico $y \in J$ tale che $F(x, y) = 0$. Denotando tale y con $y = g(x)$, avremo che la funzione $g : I \rightarrow J$ è di classe C^1 e*

$$g'(x) = -\frac{\frac{\partial F}{\partial x}(x, g(x))}{\frac{\partial F}{\partial y}(x, g(x))} \quad \forall x \in I. \quad (5.3.1)$$

Supponendo invece che $\frac{\partial F}{\partial x}(x_0, y_0) \neq 0$ esisterà sempre un rettangolo $I \times J := [x_0 - \delta, x_0 + \delta] \times [y_0 - \eta, y_0 + \eta]$ incluso in A tale che per ogni $y \in J$ esiste un unico $x \in I$ tale che $F(x, y) = 0$. Denotando in tal caso $x = h(y)$, avremo che $h : J \rightarrow I$ è di classe C^1 e

$$h'(y) = -\frac{\frac{\partial F}{\partial y}(h(y), y)}{\frac{\partial F}{\partial x}(h(y), y)} \quad \forall y \in J. \quad (5.3.2)$$

Il Teorema del Dini ci permette di dire quando il luogo di zeri della funzione F , denotato con $\Gamma := \{(x, y) \in A : F(x, y) = 0\}$ può essere espresso localmente come grafico di una funzione (o della x o della y) e quindi come sostegno di una curva parametrica.

Dimostrazione. Mettiamoci nell'ipotesi in cui $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$, ad esempio $\frac{\partial F}{\partial y}(x_0, y_0) > 0$, e dimostriamo il teorema in questo caso specifico (gli altri casi si trattano in modo analogo). Essendo $\frac{\partial F}{\partial y}(x_0, y_0) > 0$ e $\frac{\partial F}{\partial y}$ è continua, esisterà un intorno (che prendiamo rettangolare) $\hat{I} \times J \subset A$ di (x_0, y_0) su cui $\frac{\partial F}{\partial y} > 0$. Prendiamo $J = [y_0 - \eta, y_0 + \eta]$ con $\eta > 0$, e allora avremo che la funzione $s \mapsto F(x_0, s)$ con $s \in J$ ha derivata in s che coincide con $\frac{\partial F}{\partial y}$ e sarà quindi strettamente positiva. Di conseguenza tale funzione è strettamente crescente, ed essendo $F(x_0, y_0) = 0$ avremo che $F(x_0, s) < 0$ per $s < y_0$ e $F(x_0, s) > 0$ per $s > y_0$. In particolare

$F(x_0, y_0 - \eta) < 0$ e $F(x_0, y_0 + \eta) > 0$. Scegliamo ora $\delta > 0$ in modo che $I := [x_0 - \delta, x_0 + \delta] \subset \hat{I}$ e

$$F(x, y_0 - \eta) < 0 \quad \forall x \in I, \quad F(x, y_0 + \eta) > 0 \quad \forall x \in I.$$

Concludiamo che per ogni $(x, y) \in I \times J$ si avrà che $\frac{\partial F}{\partial y}(x, y) > 0$ e allo stesso tempo vale la condizione precedente. Conseguentemente la funzione $s \mapsto F(x, s)$ con $s \in J$ ha derivata rispetto ad s strettamente positiva, e ha valore negativo per $s = y_0 - \eta$ e positivo per $s = y_0 + \eta$. Ne deduciamo che per ogni $x \in I$ deve esistere un $y \in J$ tale che la funzione $s \mapsto F(x, s)$ si annulla in $s = y$, ovvero la prima parte della tesi. Denotando ora tale y con $g(x)$ dimostriamo che $g : I \rightarrow J$ è di classe C^1 e soddisfa (5.3.1). A tale scopo fissiamo $x_1, x_2 \in I$ e consideriamo la funzione di $t \in [0, 1]$ definita da

$$H(t) := F(x_1 + t(x_2 - x_1), g(x_1) + t(g(x_2) - g(x_1))).$$

Avremo $H(0) = F(x_1, g(x_1)) = 0$ e $H(1) = F(x_2, g(x_2)) = 0$. Inoltre H è continua (perchè lo è F) e derivabile con

$$\begin{aligned} H'(t) &= \frac{d}{dt} F(x_1 + t(x_2 - x_1), g(x_1) + t(g(x_2) - g(x_1))) \\ &= (x_2 - x_1) \frac{\partial F}{\partial x}(x_1 + t(x_2 - x_1), g(x_1) + t(g(x_2) - g(x_1))) \\ &\quad + (g(x_2) - g(x_1)) \frac{\partial F}{\partial y}(x_1 + t(x_2 - x_1), g(x_1) + t(g(x_2) - g(x_1))) \end{aligned}$$

(dove abbiamo usato la formula di derivazione della funzione composta). Per il teorema di Rolle esisterà un punto $\xi \in (0, 1)$ tale che $H'(\xi) = 0$, e quindi

$$\begin{aligned} (x_2 - x_1) \frac{\partial F}{\partial x}(x_1 + \xi(x_2 - x_1), g(x_1) + \xi(g(x_2) - g(x_1))) &= \\ = -(g(x_2) - g(x_1)) \frac{\partial F}{\partial y}(x_1 + \xi(x_2 - x_1), g(x_1) + \xi(g(x_2) - g(x_1))), \end{aligned}$$

il che implica, ponendo $M := \min\{\frac{\partial F}{\partial y}(x, y) : (x, y) \in I \times J\}$ e $m = \max\{\frac{\partial F}{\partial x}(x, y) : (x, y) \in I \times J\}$,

$$\begin{aligned} |g(x_2) - g(x_1)| &= \frac{|\frac{\partial F}{\partial x}(x_1 + \xi(x_2 - x_1), g(x_1) + \xi(g(x_2) - g(x_1)))|}{|\frac{\partial F}{\partial y}(x_1 + \xi(x_2 - x_1), g(x_1) + \xi(g(x_2) - g(x_1)))|} |x_2 - x_1| \\ &\leq \frac{m}{M} |x_2 - x_1|, \end{aligned}$$

ovvero g è Lipschitziana di costante $\frac{m}{M}$, e quindi continua. Dalla stessa formula otteniamo anche che

$$\frac{g(x_2) - g(x_1)}{x_2 - x_1} = - \frac{\frac{\partial F}{\partial x}(x_1 + \xi(x_2 - x_1), g(x_1) + \xi(g(x_2) - g(x_1)))}{\frac{\partial F}{\partial y}(x_1 + \xi(x_2 - x_1), g(x_1) + \xi(g(x_2) - g(x_1)))},$$

da cui mandando $x_1 \rightarrow x_2$ e usando che in tal caso avremo $\xi(x_2 - x_1) \rightarrow 0$ e $\xi(g(x_2) - g(x_1)) \rightarrow 0$, otteniamo la formula (5.3.1) con $x = x_2$. Da qui segue la tesi visto che da tale formula si evince anche la continuità di g' e quindi il fatto che g è di classe C^1 . $\square$

In ambo i casi (5.3.1) e (5.3.2) il Teorema precedente ci assicura che il luogo di zeri $\Gamma = \{(x, y) \in A : F(x, y) = 0\}$ della funzione F può essere espresso come curva parametrizzata (localmente), e in particolare come grafico di una funzione. Se $(x_0, y_0) \in \Gamma$ e sono soddisfatte le ipotesi del teorema, allora in un intorno di (x_0, y_0) avremo che Γ coincide con il sostegno della curva

$$\gamma(x) = (x, g(x)) \quad \text{oppure} \quad \sigma(y) = (h(y), y).$$

Le velocità delle curve sopra saranno

$$\gamma'(x) = (1, g'(x)) = \left(1, -\frac{\frac{\partial F}{\partial x}(x, g(x))}{\frac{\partial F}{\partial y}(x, g(x))}\right) \quad \sigma'(y) = (h'(y), 1) = \left(-\frac{\frac{\partial F}{\partial y}(h(y), y)}{\frac{\partial F}{\partial x}(h(y), y)}, 1\right).$$

Dalle osservazioni fatte nella sessione precedente sappiamo che $\gamma'(x_0)$ e $\sigma'(y_0)$ sono tangenti alla curva (quindi a Γ) nel punto (x_0, y_0) . Calcolando il prodotto scalare di $\gamma'(x_0)$ e $\sigma'(y_0)$ con $\nabla F(x_0, y_0) = (\frac{\partial F}{\partial x}(x_0, y_0), \frac{\partial F}{\partial y}(x_0, y_0))$ e usando che $g(x_0) = y_0$ e $h(y_0) = x_0$ ci accorgiamo che tale prodotto scalare è nullo. Questo significa che i vettori $\gamma'(x_0)$ e $\sigma'(y_0)$ (che sono tangenti a Γ) risultano ortogonali a $\nabla F(x_0, y_0)$. Ne segue che

$$\nabla F(x_0, y_0) \perp \Gamma \quad \text{in } (x_0, y_0), \quad (5.3.3)$$

ovvero il gradiente di F è ortogonale al luogo di zeri di F .

Un teorema analogo a quello sopra vale quando cerchiamo di descrivere il luogo di zeri di funzioni definite su domini di $\mathbb{R}^3$.

Teorema 5.3.2 (Dini 2). *Sia $F : A \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ una funzione di classe C^1 e sia $(x_0, y_0, z_0) \in A$ un punto tale che $F(x_0, y_0, z_0) = 0$. Supponiamo che $\frac{\partial F}{\partial z}(x_0, y_0, z_0) \neq 0$, allora esiste un parallelepipedo (rettangolo in $\mathbb{R}^3$) $I \times J \times K := [x_0 - \delta, x_0 + \delta] \times [y_0 - \eta, y_0 + \eta] \times [z_0 - \zeta, z_0 + \zeta]$ incluso in A tale che per ogni $(x, y) \in I \times J$ esiste un unico $z \in K$ tale che $F(x, y, z) = 0$. Denotando tale z con $z = f(x, y)$, avremo che la funzione $f : I \times J \rightarrow K$ è di classe C^1 e*

$$\nabla f(x, y) = \left(-\frac{\frac{\partial F}{\partial x}(x, y, f(x, y))}{\frac{\partial F}{\partial z}(x, y, f(x, y))}, -\frac{\frac{\partial F}{\partial y}(x, y, f(x, y))}{\frac{\partial F}{\partial z}(x, y, f(x, y))} \right) \quad \forall (x, y) \in I \times J. \quad (5.3.4)$$

Anche in questo caso possiamo esprimere localmente il luogo di zeri S di F ,

$$S := \{(x, y, z) \in A : F(x, y, z) = 0\},$$

come grafico di una funzione di due variabili (e quindi come sostegno di una superficie)

$$\varphi(x, y) := (x, y, f(x, y)),$$

le cui derivate parziali saranno

$$\frac{\partial \varphi}{\partial x}(x, y) = \left(1, 0, -\frac{\frac{\partial F}{\partial x}(x, y, f(x, y))}{\frac{\partial F}{\partial z}(x, y, f(x, y))}\right) \quad \frac{\partial \varphi}{\partial y}(x, y) = \left(0, 1, -\frac{\frac{\partial F}{\partial y}(x, y, f(x, y))}{\frac{\partial F}{\partial z}(x, y, f(x, y))}\right).$$

Anche in questo caso, svolgendo i prodotti scalari di queste derivate con $\nabla F(x_0, y_0, z_0)$ troviamo 0 come risultato. Questo significa che il gradiente di F è ortogonale a tali derivate, che sappiamo però, dalla sessione precedente, essere tangenti ad S . Ne consegue che

$$\nabla F(x_0, y_0, z_0) \perp S, \quad (5.3.5)$$

ovvero anche in questo caso ∇F è ortogonale al luogo di zeri S di F .

Occupiamoci ora di studiare massimi e minimi relativi di funzioni ristrette a curve o superfici, o in generale a insiemi che sono descritti come zeri di funzioni scalari.

Denotiamo con Z il luogo di zeri di una funzione $F : A \rightarrow \mathbb{R}$, dove assumiamo $A \subset \mathbb{R}^n$ (dove $n \geq 2$, e terremo in mente i casi trattati prima in cui $n = 2, 3$). Nel caso in cui $n = 2$, supporremo che il luogo di zeri Z è parametrizzato (localmente) da una curva (sappiamo che questo può essere assunto grazie al teorema del Dini). Scriviamo $Z = \gamma(I)$ dove $\gamma : I \rightarrow \mathbb{R}^2$ parametrizza Z ed I è un intervallo. Se vogliamo studiare i punti di massimo o minimo di una funzione $f : A \rightarrow \mathbb{R}$ ristretta a Z dovremmo studiare prima i punti stazionari di

$$f(\gamma(t)) \quad t \in I.$$

A questo scopo imponiamo che la derivata di $f \circ \gamma$ si annulla. Avremo

$$\frac{d}{dt} f \circ \gamma(t) = \nabla f(\gamma(t)) \cdot \gamma'(t) = 0.$$

Pertanto, la condizione necessaria affinché il punto $\gamma(t) \in Z$ sia un punto di massimo o minimo relativo per la funzione f sul vincolo Z è che il gradiente di f sia ortogonale a $\gamma'(t)$. Se $\gamma(t) = (x_0, y_0) \in Z$, questa condizione si traduce con

$$\nabla f(x_0, y_0) \perp Z \quad \text{in } (x_0, y_0).$$

Confrontando questa espressione con (5.3.3), anche ∇F deve essere ortogonale a Z , e quindi risulterà che ∇f e ∇F sono paralleli in (x_0, y_0) : esisterà allora $\lambda \in \mathbb{R}$ tale che

$$\nabla f(x_0, y_0) = \lambda \nabla F(x_0, y_0). \quad (5.3.6)$$

In maniera analoga, nel caso in cui Z è il luogo di zeri di una funzione $F : A \rightarrow \mathbb{R}$ con $A \subset \mathbb{R}^3$, possiamo supporre che (localmente) Z sia parametrizzato da una superficie (Teorema del Dini 2) $\varphi : K \rightarrow A$. Pertanto, punti di massimo o minimo di f su Z dovranno soddisfare la condizione

$$\nabla(f \circ \varphi)(u, v) = \nabla f(\varphi(u, v)) \nabla \varphi(u, v) = 0.$$

Ricordando che

$$\nabla \varphi = \begin{pmatrix} \frac{\partial \varphi_1}{\partial u} & \frac{\partial \varphi_1}{\partial v} \\ \frac{\partial \varphi_2}{\partial u} & \frac{\partial \varphi_2}{\partial v} \end{pmatrix}$$

la condizione precedente si traduce nel richiedere che $\nabla f(\varphi(u, v))$ sia ortogonale sia a $\nabla \varphi_1(u, v)$ che a $\nabla \varphi_2(u, v)$: ma i $\nabla \varphi_i(u, v)$ determinano il piano tangente a Z in $\varphi(u, v) = (x_0, y_0, z_0)$, e quindi la condizione precedente ci dice che

$$\nabla f(x_0, y_0, z_0) \perp Z \quad \text{in } (x_0, y_0, z_0).$$

Ricordando infine la (5.3.5), concludiamo anche in questo caso che ∇F e ∇f devono essere paralleli tra loro, e quindi esisterà $\lambda \in \mathbb{R}$ tale che

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla F(x_0, y_0, z_0). \quad (5.3.7)$$

Possiamo allora enunciare il seguente teorema:

Teorema 5.3.3. (*Moltiplicatori di Lagrange*) Sia $A \subset \mathbb{R}^n$ un aperto e siano $F, f : A \rightarrow \mathbb{R}$ funzioni di classe C^1 . Sia $Z := \{x \in A : F(x) = 0\}$ e supponiamo ∇F non si annulla su Z . Allora condizione necessaria affinché il punto $x_0 \in Z$ sia di massimo o minimo relativo per la funzione f ristretta a Z è che esista $\lambda \in \mathbb{R}$ tale che

$$\begin{cases} \nabla f(x_0) = \lambda \nabla F(x_0) \\ F(x_0) = 0. \end{cases}$$

Il teorema precedente può anche essere enunciato dicendo che condizione necessaria affinché il punto $x_0 \in Z$ sia di massimo o minimo relativo per la funzione f ristretta a Z è esista $\lambda \in \mathbb{R}$ per cui il punto (λ, x_0) sia un punto stazionario della funzione Lagrangiana

$$L(\lambda, x) := \lambda F(x) - f(x).$$

Il teorema può essere anche generalizzato a più vincoli: se ad esempio vogliamo studiare i punti di massimo e minimo relativo di f sul luogo $Z := \{x \in A : F(x) = G(x) = 0\}$ dove si annullano le funzioni F e G , dobbiamo cercare i punti stazionari della Lagrangiana $L : \mathbb{R}^2 \times A \rightarrow \mathbb{R}$ definita come

$$L(\lambda_1, \lambda_2, x) := \lambda_1 F(x) + \lambda_2 G(x) - f(x).$$

Capitolo 6

Equazioni Differenziali e Problema di Cauchy

Tratteremo per lo più il caso in cui l'incognita dell'equazione differenziale è una funzione di una variabile a valori reali. Quindi ci scontriamo con un'equazione del tipo

$$u^{(n)}(t) = F(t, u, u', u'', \dots, u^{(n-1)}) \quad (6.0.1)$$

dove $u^{(k)} = \frac{d^k}{dt^k}u$ indica la derivata di ordine k nella variabile t della funzione u , ed $F : I \times \mathbb{R}^n \rightarrow \mathbb{R}$ è una data funzione, $I \subseteq \mathbb{R}$ un intervallo. Di fronte ad un'equazione come sopra ci chiediamo se esiste una funzione $u : I \rightarrow \mathbb{R}$ che la soddisfa per ogni $t \in I$. L'equazione (6.0.1) prende nome di equazione differenziale di ordine n .

Ogni equazione differenziale di ordine n della forma (6.0.1) può essere ricondotta ad un sistema di equazioni differenziali del primo ordine, nel seguente modo: definiamo il vettore $U \in \mathbb{R}^n$ come

$$U_1 := u, \quad U_2 := u', \quad \dots, U_k := u^{(k-1)}, \dots, \quad U_n := u^{(n-1)}. \quad (6.0.2)$$

Facendo la derivata prima di U otteniamo, usando anche la (6.0.1),

$$\begin{cases} U'_1 = U_2 \\ U'_2 = U_3 \\ \dots \\ U'_{n-1} = U_n \\ U'_n = F(t, U_1, \dots, U_n). \end{cases} \quad (6.0.3)$$

sistema che si può sintetizzare nell'espressione

$$U' = f(t, U), \quad (6.0.4)$$

dove la funzione $f : I \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ avrà componenti

$$f_i(t, U) = U_{i+1} \quad \text{se } i \leq n-1, \quad f_n(t, U) = F(t, U).$$

L'equazione (6.0.4) è un'equazione differenziale del prim'ordine vettoriale (cioè in cui la funzione incognita è un vettore $U : I \rightarrow \mathbb{R}^n$), ed nel caso generale può essere vista come sistema

di equazioni differenziali del prim'ordine per n funzioni scalari,

$$\begin{cases} U'_1 = f_1(t, U) \\ U'_2 = f_2(t, U) \\ \dots \\ U'_{n-1} = f_{n-1}(t, U) \\ U'_n = f_n(t, U). \end{cases} \quad (6.0.5)$$

Ci occuperemo ora di studiare le equazioni differenziali al prim'ordine, e per semplicità ci metteremo nel caso scalare, ovvero quello in cui la funzione incognita u è a valori reali.

6.1 Problema di Cauchy

Il problema di Cauchy consiste nel cercare una soluzione u di un'equazione differenziale del primo ordine, imponendo che in un certo tempo t_0 la funzione u assumi un determinato valore prefissato u_0 . Ovvero, cerchiamo soluzioni del sistema

$$\begin{cases} u'(t) = f(t, u(t)) \\ u(t_0) = u_0 \end{cases} \quad (6.1.1)$$

Teorema 6.1.1. *Sia $A \subset \mathbb{R}^2$ un insieme aperto e sia $f : A \rightarrow \mathbb{R}$ una funzione di classe C^1 . Sia $(t_0, u_0) \in A$. Allora esiste un intervallo $I = (t_0 - r_0, t_0 + r_0)$ ed una funzione $u : I \rightarrow \mathbb{R}$ che soddisfa (6.1.1). Inoltre, tale soluzione del problema di Cauchy è unica.*

Osservazione 6.1.2. Cercare una soluzione di (6.1.1) presuppone richiedere che la funzione da cercare sia quantomeno derivabile nella variabile t (altrimenti l'equazione differenziale non avrebbe senso). Per funzioni di una variabile, la derivabilità implica la continuità, e quindi la funzione cercata u sarà continua. Di conseguenza, essendo f una funzione continua, anche $f(t, u(t))$ risulterà continua (composizione di funzioni continue), e ne deduciamo che se $u'(t) = f(t, u(t))$ anche u' sarà continua. Concludendo, ogni soluzione dell'equazione differenziale $u'(t) = f(t, u(t))$ sarà necessariamente di classe C^1 .

Per funzioni di classe C^1 vale in teorema fondamentale del calcolo integrale, che implica che se integriamo ambo i membri dell'equazione differenziale su un intervallo (t_0, t) otteniamo

$$u(t) - u(t_0) = \int_{t_0}^t u'(s) ds = \int_{t_0}^t f(s, u(s)) ds.$$

Infine, visto che $u(t_0) = u_0$, abbiamo ottenuto che ogni soluzione del problema di Cauchy (6.1.1) deve essere di classe C^1 e soddisfare, per ogni t ,

$$u(t) = u_0 + \int_{t_0}^t f(s, u(s)) ds. \quad (6.1.2)$$

Viceversa, se $u \in C^1$ soddisfa (6.1.2), allora (ponendo $t = t_0$) deve soddisfare $u(t_0) = u_0$, e (derivando la l'espressione precedente) anche l'equazione differenziale $u'(t) = f(t, u(t))$. In definitiva, u soddisfa il problema di Cauchy (6.1.1) se e solo se u è di classe C^1 e soddisfa (6.1.2).

Dimostrazione. Iniziamo a discutere l'esistenza della soluzione: a questo scopo, grazie all'osservazione precedente, cerchiamo soluzioni u che soddisfino (6.1.2).

Step 1: Iniziamo nel fissare un rettangolo $R := [t_0 - r, t_0 + r] \times [u_0 - \rho, u_0 + \rho]$ in modo che sia tutto contenuto in A . Essendo R compatto ed f e ∇f continui, esisteranno finiti i due valori $M, L \geq 0$ definiti da

$$M := \max\{|f(t, u)| : (t, u) \in R\} \quad L := \max\left|\frac{\partial f}{\partial u}(t, u)\right| : (t, u) \in R. \quad (6.1.3)$$

Poniamo allora

$$r_0 := \min\left\{r, \frac{\rho}{M}\right\} \quad I := (t_0 - r_0, t_0 + r_0). \quad (6.1.4)$$

Facciamo ora vedere che possiamo definire una successione di funzioni $u_k : I \rightarrow [u_0 - \rho, u_0 + \rho]$ il cui limite uniforme è una soluzione del problema di Cauchy. Definiamo, per ogni $k \geq 1$,

$$\begin{aligned} u_0(t) &:= u_0 && \text{per ogni } t \in I, \\ u_k(t) &:= u_0 + \int_{t_0}^t f(s, u_{k-1}(s))ds && \text{per ogni } t \in I. \end{aligned} \quad (6.1.5)$$

Step 2: Facciamo vedere che per ogni $k \geq 0$ e per ogni $t \in I$ si ha

$$|u_k(t) - u_0| \leq \rho. \quad (6.1.6)$$

Per $k = 0$ è chiaramente vero; procediamo per induzione e mostriamo che se vale per k , allora anche per $k + 1$. Scriviamo, utilizzando la definizione di u_{k+1} ,

$$|u_{k+1}(t) - u_0| = \left| \int_{t_0}^t f(s, u_k(s))ds \right|.$$

Supponendo $t > t_0$ avremo

$$|u_{k+1}(t) - u_0| \leq \int_{t_0}^t |f(s, u_k(s))|ds$$

ed essendo, per ipotesi induttiva, $u_k(s) \in [u_0 - \rho, u_0 + \rho]$, possiamo stimare

$$\int_{t_0}^t |f(s, u_k(s))|ds \leq \int_{t_0}^t Mds = |t - t_0|M \leq r_0 M \leq \rho.$$

Essendo anche valida una stima identica nel caso in cui $t < t_0$, abbiamo mostrato (6.1.6) nel caso $k + 1$, e per induzione per ogni $k \in \mathbb{N}$.

Step 3: Mostriamo, sempre per induzione, che per ogni $k \in \mathbb{N}$ e $t \in I$, si ha

$$|u_k(t) - u_{k-1}(t)| \leq ML^{k-1} \frac{|t - t_0|^k}{k!}. \quad (6.1.7)$$

Nel caso $k = 1$ abbiamo $|u_1(t) - u_0(t)| \leq \int_{t_0}^t |f(s, u_0(s))|ds \leq M|t - t_0|$, che è la tesi nel caso $t > t_0$ (l'altro si fa allo stesso modo). Procediamo con il passo induttivo, e supponendo (6.1.7), facciamo vedere che

$$|u_{k+1}(t) - u_k(t)| \leq ML^k \frac{|t - t_0|^{k+1}}{(k+1)!}. \quad (6.1.8)$$

Anche in questo caso supponiamo $t > t_0$ (la verifica nell'altro caso è identica). Scriviamo

$$|u_{k+1}(t) - u_k(t)| = \left| \int_{t_0}^t f(s, u_k(s)) - f(s, u_{k-1}(s)) ds \right| \leq \int_{t_0}^t |f(s, u_k(s)) - f(s, u_{k-1}(s))| ds.$$

Per il teorema di Lagrange sappiamo che, per ogni s , esiste un valore $\xi(s)$ compreso tra $u_k(s)$ e $u_{k-1}(s)$ tale che $f(s, u_k(s)) - f(s, u_{k-1}(s)) = \frac{\partial f}{\partial u}(s, \xi(s))(u_k(s) - u_{k-1}(s))$. Da qui si deduce

$$|f(s, u_k(s)) - f(s, u_{k-1}(s))| \leq \left| \frac{\partial f}{\partial u}(s, \xi(s)) \right| |u_k(s) - u_{k-1}(s)| \quad (6.1.9)$$

da cui

$$\begin{aligned} |u_{k+1}(t) - u_k(t)| &\leq \int_{t_0}^t \left| \frac{\partial f}{\partial u}(s, \xi(s)) \right| |u_k(s) - u_{k-1}(s)| ds \leq L \int_{t_0}^t |u_k(s) - u_{k-1}(s)| ds \\ &\leq L \int_{t_0}^t M L^{k-1} \frac{(s - t_0)^k}{k!} ds = M L^k \frac{|t - t_0|^{k+1}}{(k+1)!}, \end{aligned}$$

che è la tesi.

Step 4: Prendendo ora l'estremo superiore nella (6.1.7) per $t \in I$ otteniamo anche

$$\|u_k - u_{k-1}\|_{\infty, I} \leq M L^{k-1} \frac{r_0^k}{k!}.$$

Questo implica che la serie

$$\sum_{k=1}^{\infty} (u_k - u_{k-1})$$

converge totalmente su I , dal momento che $\sum_{k=1}^{\infty} M L^{k-1} \frac{r_0^k}{k!} = \frac{M}{L} e^{L r_0} < +\infty$. Di conseguenza, le somme parziali $\sum_{k=1}^n (u_k - u_{k-1})$ convergono uniformemente a qualche funzione $\tilde{u} : I \rightarrow \mathbb{R}$. Ma adesso, $\sum_{k=1}^n (u_k - u_{k-1}) = u_n - u_0$, e quindi scopriamo che

$$u_n \rightarrow u := \tilde{u} + u_0 \quad \text{uniformemente su } I.$$

Dallo Step 2 scopriamo anche che $|u - u_0| \leq \rho$, e quindi $u : I \rightarrow [u_0 - \rho, u_0 + \rho]$. Inoltre, le u_n sono tutte continue e tale sarà la u . Infine, scrivendo

$$u_n(t) = u_0 + \int_{t_0}^t f(s, u_{n-1}(s)) ds, \quad (6.1.10)$$

osserviamo che essendo f continua su R , anche le funzioni integrande $f(s, u_{n-1}(s))$ convergeranno uniformemente a $f(s, u(s))$ (verifica lasciata per esercizio, guardare la (6.1.9)), e quindi passando al limite per $n \rightarrow \infty$ nell'espressione precedente concludiamo

$$u(t) = u_0 + \int_{t_0}^t f(s, u(s)) ds. \quad (6.1.11)$$

Questo ci dice che u è di classe C^1 ed è soluzione di (6.1.1) (grazie all'Osservazione 6.1.2).

Passiamo ora alla dimostrazione dell'unicità della soluzione.

Step 5: Supponiamo u e v siano entrambe soluzioni di (6.1.2). Allora

$$u(t) - v(t) = \int_{t_0}^t f(s, u(s)) - f(s, v(s)) ds.$$

Se u e v non coincidessero ovunque, definendo

$$\begin{aligned} t_1 &:= \sup\{t \in [t_0, t_0 + r_0] : u = v \text{ su } [t_0, t]\}, \\ t_2 &:= \inf\{t \in [t_0 - r_0, t_0] : u = v \text{ su } [t, t_0]\}, \end{aligned}$$

si avrebbe che o $t_1 < t_0 + r_0$, oppure $t_2 > t_0 - r_0$ (o entrambi). Supponiamo $t_2 > t_0 - r_0$ (l'altro caso è simile) e quindi avremmo $u = v$ su $[t_2, t_0]$ mentre esse non coinciderebbero su $[\bar{t}, t_0]$, per ogni $\bar{t} < t_2$. Scriviamo allora, per ogni $t \in [\bar{t}, t_0]$,

$$\begin{aligned} |u(t) - v(t)| &= \left| \int_{t_0}^t f(s, u(s)) - f(s, v(s)) ds \right| \leq \int_t^{t_0} |f(s, u(s)) - f(s, v(s))| ds \\ &= \int_t^{t_2} |f(s, u(s)) - f(s, v(s))| ds \leq L \int_t^{t_2} |u(s) - v(s)| ds, \end{aligned}$$

dove abbiamo usato la (6.1.9). Stimando $|u - v|$ con la sua norma infinito, abbiamo per ogni $t \in [\bar{t}, t_2]$

$$|u(t) - v(t)| \leq L \|u - v\|_{\infty, [\bar{t}, t_2]} (t_2 - \bar{t}),$$

e prendendo l'estremo superiore su tutti gli $t \in [\bar{t}, t_2]$ nel membro di sinistra, concludiamo

$$\|u - v\|_{\infty, [\bar{t}, t_2]} \leq L \|u - v\|_{\infty, [\bar{t}, t_2]} (t_2 - \bar{t}).$$

Ora, se per assurdo $\|u - v\|_{\infty, [\bar{t}, t_2]} \neq 0$, concluderemmo $(t_2 - \bar{t}) \geq \frac{1}{L}$, il che è assurdo perchè $\bar{t}$ è un numero arbitrario $< t_2$, e quindi potremmo scegliere $\bar{t} := t_2 - \frac{1}{2L}$ ottenendo una contraddizione.

Concludiamo che necessariamente $\|u - v\|_{\infty, [\bar{t}, t_2]} = 0$, ovvero $u = v$ su $[\bar{t}, t_2]$ portando ad una ulteriore contraddizione con la definizione di t_2 . Segue che t_2 deve essere necessariamente uguale a $t_0 - r_0$, ovvero $u = v$ su tutto I . $\square$

Così come abbiamo mostrato il teorema precedente, avremmo potuto dimostrare il seguente:

Teorema 6.1.3. *Sia $A \subset \mathbb{R}^2$ un insieme aperto e sia $f : A \rightarrow \mathbb{R}$ una funzione continua localmente Lipschitziana nella seconda variabile. Sia $(t_0, u_0) \in A$. Allora esiste un intervallo $I = (t_0 - r_0, t_0 + r_0)$ ed una funzione $u : I \rightarrow \mathbb{R}$ che soddisfa (6.1.1). Inoltre, tale soluzione del problema di Cauchy è unica.*

La locale Lipschitzianità ci permette di dire che nel rettangolo R scelto nella dimostrazione precedente esiste un $L > 0$ per il quale

$$|f(t, u) - f(t, v)| \leq L|u - v|. \quad (6.1.12)$$

Questa proprietà permette di evitare di passare per la (6.1.9) (che ci porta esattamente alla stessa disuguaglianza). Il resto della dimostrazione è identica al caso precedente.

Un simile teorema vale per le funzioni a valori vettoriali:

Teorema 6.1.4. *Sia $A \subset \mathbb{R}^{n+1}$ un insieme aperto e sia $f : A \rightarrow \mathbb{R}$ una funzione continua localmente Lipschitziana nella seconda variabile. Sia $(t_0, u_0) \in A$. Allora esiste un intervallo $I = (t_0 - r_0, t_0 + r_0)$ ed una funzione $u : I \rightarrow \mathbb{R}^n$ che soddisfa*

$$\begin{cases} u' = f(t, u) \\ u(0) = u_0. \end{cases} \quad (6.1.13)$$

Inoltre, tale soluzione del problema di Cauchy è unica.

Quest'ultimo teorema ci garantisce l'esistenza di una soluzione dell'equazione differenziale (6.0.1) di ordine n . Infatti abbiamo visto che essa è equivalente al sistema del prim'ordine (6.0.3), che a sua volta è un caso particolare del sistema (6.0.5). Se imponiamo una condizione iniziale per il sistema (6.0.3) esso diventa un caso particolare del problema di Cauchy (6.1.13). Si osservi che fissare un dato iniziale per u in (6.1.13) significa fissare un vettore $u_0 := (u_0^1, u_0^2, \dots, u_0^n)$ e imporre che $u_i(t_0) = u_0^i$. Adesso ri-interpretando il vettore u come in (6.0.2), otteniamo che il sistema (6.0.3) diventa

$$\begin{cases} u^{(n)} = F(t, u, u', u'', \dots, u^{(n-1)}) \\ u(0) = u_0^1 \\ u'(0) = u_0^2 \\ \dots \\ u^{(n-1)}(0) = u_0^n. \end{cases} \quad (6.1.14)$$

Quindi il Teorema (6.1.4) ci garantisce l'esistenza e unicità di una soluzione dell'equazione (6.0.4) a patto di imporre n dati iniziali, uno per ogni derivata di ordine inferiore all'ennesima.

6.2 Soluzioni massimali e globali

Proposizione 6.2.1. *Siano $u : (a, b) \rightarrow \mathbb{R}$ e $v : (c, d) \rightarrow \mathbb{R}$ funzioni che risolvono la medesima equazione differenziale $u'(t) = f(t, u(t))$ e $v'(t) = f(t, v(t))$ sui loro domini, e supponiamo $u = v$ su $(a, b) \cap (c, d)$. Allora la funzione $w : (a, b) \cup (c, d) \rightarrow \mathbb{R}$ data da*

$$w(t) := \begin{cases} u(t) & \text{se } t \in (a, b) \\ v(t) & \text{se } t \in (c, d) \end{cases}$$

è ancora soluzione di $w'(t) = f(t, w(t))$ sul suo dominio.

Dimostrazione. Intanto osserviamo che w è ben definita, visto che $u = v$ su $(a, b) \cap (c, d)$. Ora, se $t \in (a, b)$ avremo $w'(t) = u'(t) = f(t, u(t)) = f(t, w(t))$, e similmente se $t \in (c, d)$. Quindi w risolve l'equazione differenziale. $\square$

Proposizione 6.2.2. *Siano $u : (a, b) \rightarrow \mathbb{R}$ e $v : (a, b) \rightarrow \mathbb{R}$ funzioni che risolvono la medesima equazione differenziale $u'(t) = f(t, u(t))$ e $v'(t) = f(t, v(t))$, dove supponiamo f continua e localmente Lipschitziana nella seconda variabile. Se $u(t_0) = v(t_0)$ per qualche $t_0 \in (a, b)$, allora $u = v$ su (a, b) .*

Dimostrazione. Supponiamo falsa la tesi e poniamo

$$\bar{t} := \sup\{t \in [t_0, b) : u = v \text{ su } [t_0, t]\};$$

supponiamo allora $\bar{t} < b$. Per continuità avremo $u(\bar{t}) = v(\bar{t}) =: w_0$, e allora u e v risolvono entrambe il problema di Cauchy

$$\begin{cases} w'(t) = f(t, w(t)) \\ w(\bar{t}) = w_0, \end{cases}$$

così coincideranno in un intorno di $\bar{t}$ grazie all'unicità della soluzione. In particolare, esisterà $r_0 > 0$ tale che $u = v$ su $[\bar{t} - r_0, \bar{t} + r_0]$ che porta ad una contraddizione con la definizione di $\bar{t}$. Concludiamo che allora deve essere $\bar{t} = b$. Un discorso analogo mostra che $u = v$ su $[a, t_0]$, e abbiamo concluso. $\square$

Definizione 6.2.3. Siano $u : (a, b) \rightarrow \mathbb{R}$ e $v : (c, d) \rightarrow \mathbb{R}$ funzioni che risolvono la medesima equazione differenziale $u'(t) = f(t, u(t))$ e $v'(t) = f(t, v(t))$ sui loro domini. Diremo che u è un prolungamento di v se

- $(c, d) \subsetneq (a, b)$;
- $u = v$ su (c, d) .

Diremo che u è una soluzione massimale dell'equazione differenziale $u'(t) = f(t, u(t))$ se u non ammette prolungamenti.

Veniamo ora alla discussione dell'esistenza di soluzioni massimali.

Teorema 6.2.4. Sia $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ una funzione continua e localmente Lipschitziana nella seconda variabile. Supponiamo $u : (a, b) \rightarrow \mathbb{R}$ risolva $u' = f(t, u)$; allora o u è una soluzione massimale, oppure essa ammette un prolungamento massimale.

Dimostrazione. Supponiamo u non sia massimale e dimostriamo che esiste un prolungamento massimale. Sia $\mathcal{U}$ la famiglia di funzione che sono prolungamenti di u :

$$\mathcal{U} := \{w : (a_w, b_w) \rightarrow \mathbb{R} : w \text{ è un prolungamento di } u\}.$$

Sia $I := \bigcup_{w \in \mathcal{U}} (a_w, b_w)$, e definiamo $v : I \rightarrow \mathbb{R}$ come

$$v(t) = w(t) \quad \text{se } t \in (a_w, b_w).$$

Anzitutto notiamo che v è ben definita, perchè se w_1 e w_2 sono funzioni tali che $t \in (a_{w_i}, b_{w_i})$ per $i = 1, 2$, avremo che $w_1(t) = w_2(t)$. Infatti, entrambe w_1 e w_2 prolungano u , quindi coincidono su (a, b) e per le Proposizioni 6.2.1 e 6.2.2 coincideranno anche su $(a_{w_1}, b_{w_1}) \cap (a_{w_2}, b_{w_2})$ e quindi in t . Essendo ora che v coincide con w su (a_w, b_w) , essa risolve la medesima equazione differenziale di w , e quindi $v' = f(t, v)$. Infine, v non ammette prolungamenti, perchè se $\bar{v} : J \rightarrow \mathbb{R}$ fosse un suo prolungamento, avremo $I \subsetneq J$ e $\bar{v}$ sarebbe anche un prolungamento di u , quindi $\bar{v} \in \mathcal{U}$ e per definizione di I avremo $J \subset I$, contraddizione. La tesi è dimostrata. $\square$

Analizziamo ora un caso in cui soluzioni massimali sono definite globalmente.

Teorema 6.2.5. Sia $A = I \times \mathbb{R}$ un aperto con I intervallo (può essere anche tutto $\mathbb{R}$), e sia $f : A \rightarrow \mathbb{R}$ una funzione continua, localmente Lipschitziana nella seconda variabile, e tale che esiste $K > 0$ per cui

$$|f(t, u)| \leq K(|u| + 1) \quad \forall (t, u) \in A.$$

Allora ogni soluzione massimale dell'equazione differenziale $u' = f(t, u)$ è definita su tutto I .

Dimostrazione. Sia $u : (\alpha, \beta) \rightarrow \mathbb{R}$ una soluzione massimale, e supponiamo per assurdo che $(\alpha, \beta) \subsetneq I$. Se $I = (a, b)$ possiamo allora supporre che $\beta < b$. Prendiamo un punto $t_0 < \beta$ tale che $t_0 + \frac{1}{2k} > \beta$. Consideriamo il problema di Cauchy

$$\begin{cases} w'(t) = f(t, w(t)) \\ w(t_0) = u(t_0) =: u_0, \end{cases} \quad (6.2.1)$$

prendiamo allora $r \leq \frac{1}{2k}$ tale che $R := [t_0 - r, t_0 + r] \times [u_0 - \rho, u_0 + \rho] \subset A$, e tale che $t_0 + r > \beta$ (questo si può fare perchè $t_0 + \frac{1}{2k} > \beta$). Il Teorema 6.1.1 (in realtà il Teorema 6.1.3) ci dice che esiste una soluzione $w : [t_0 - r_0, t_0 + r_0] \rightarrow [u_0 - \rho, u_0 + \rho]$ dove

$$r_0 = \min\left\{r, \frac{\rho}{M}\right\}, \quad M := \max\{|f(t, u)| : (t, u) \in R\}.$$

Ricordiamo che ρ può essere scelto a piacere, e scegliamo $\rho := 1 + |u_0|$, in modo che se $(t, u) \in R$ si abbia $|u| \leq |u_0| + \rho = 1 + 2|u_0|$, e

$$|f(t, u)| \leq K(1 + |u|) \leq 2K(1 + |u_0|),$$

da cui $M \leq 2K(1 + |u_0|)$ e $\frac{\rho}{M} \geq \frac{1+|u_0|}{2K(1+|u_0|)} = \frac{1}{2K}$. Pertanto $r_0 = r$, e quindi la soluzione w è definita sull'intervallo $[t_0 - r, t_0 + r]$ con $t_0 + r > \beta$. Questo significa che possiamo estendere la soluzione u sul dominio $(\alpha, \beta) \cup [t_0 - r, t_0 + r]$ che contiene strettamente (α, β) e pertanto la u non sarebbe stata massimale, assurdo. Grazie a questa contraddizione abbiamo scoperto che necessariamente $b = \beta$; un discorso analogo mostra che $a = \alpha$ e quindi u è definita su tutto I . $\square$

6.3 Teoremi di confronto

Iniziamo con la seguente proprietà:

Lemma 6.3.1. Siano $u, v : I \rightarrow \mathbb{R}$ funzioni di classe C^1 , e supponiamo $u(t_0) = v(t_0)$ per qualche $t_0 \in I$. Se $u'(t_0) > v'(t_0)$ allora esiste un intorno J di t_0 tale che per ogni $t \in J$ si abbia

$$u(t) > v(t) \text{ se } t > t_0, \quad u(t) < v(t) \text{ se } t < t_0. \quad (6.3.1)$$

La dimostrazione del lemma è lasciata per esercizio.

Teorema 6.3.2. Sia $f : A \rightarrow \mathbb{R}$ funzione continua e siano $u, v : I \rightarrow \mathbb{R}$ funzioni di classe C^1 tali che

$$u'(t) \geq f(t, u(t)) \quad v'(t) \leq f(t, v(t)) \quad \forall t \in I. \quad (6.3.2)$$

Supponiamo inoltre che una delle due disuguaglianze precedenti valga in modo stretto. Allora se $v(t_0) \leq u(t_0)$ per qualche $t_0 \in I$ si ha anche che $v(t) < u(t)$ per ogni $t > t_0$.

Dimostrazione. Supponiamo per assurdo che $u(t) \leq v(t)$ per qualche $t > t_0$. Definiamo

$$\bar{t} := \inf\{t > t_0 : u(t) \leq v(t)\}.$$

Per continuità si avrà $u(\bar{t}) = v(\bar{t})$, ma per ipotesi avremo anche $u'(\bar{t}) > v'(\bar{t})$ perchè

$$v'(\bar{t}) \leq f(\bar{t}, v(\bar{t})) = f(\bar{t}, u(\bar{t})) \leq u'(\bar{t}),$$

dove una delle due disuguaglianze è stretta. Allora per il lemma precedente esisterà un intorno di $\bar{t}$ tale che se $t < \bar{t}$ avremo $u(t) < v(t)$, il che contraddice la definizione di $\bar{t}$. Segue che necessariamente $u(t) \geq v(t)$ per ogni $t > t_0$, e il teorema è dimostrato. $\square$

Teorema 6.3.3. *Sia $f : A \rightarrow \mathbb{R}$ funzione continua e localmente Lipschitziana nella seconda variabile, e siano $u, v : I \rightarrow \mathbb{R}$ funzioni di classe C^1 che soddisfano*

$$\begin{cases} u'(t) = f(t, u(t)) \\ u(t_0) = y_0 \end{cases} \quad \begin{cases} v'(t) \leq f(t, v(t)) \\ v(t_0) \leq y_0, \end{cases} \quad (6.3.3)$$

per qualche $(t_0, y_0) \in A$. Allora $v(t) \leq u(t)$ per ogni $t \geq t_0$.