# Appunti di Analisi Matematica 2, A.A 2015 - 2016

1 giugno 2016

## 1 Elementi di topologia di $\mathbb{R}^n$ e $\mathbb{C}^n$

**Definizione 1.** Un *gruppo* è una coppia $(\mathscr{G}, *)$ ove con $\mathscr{G}$ si indica un insieme non vuoto e con

$$\begin{aligned} *\colon \mathscr{G} \times \mathscr{G} &\to \mathscr{G} \\ (g_1, g_2) &\mapsto g_1 * g_2 \end{aligned}$$

un'operazione binaria che soddisfi le seguenti proprietà:

1. Associatività: $\forall g_1, g_2, g_3 \in \mathscr{G}$ vale che $(g_1 * g_2) * g_3 = g_1 * (g_2 * g_3)$;
2. Esistenza dell'elemento neutro: $\exists e \in \mathscr{G} \mid \forall g \in \mathscr{G}$ vale che $g * e = e * g = g$;
3. Esistenza dell'inverso bilatero: $\forall g \in \mathscr{G}\ \exists g^{-1} \in \mathscr{G} : g * g^{-1} = g^{-1} * g = e$.

**Definizione 2.** Se $\forall g_1, g_2 \in \mathscr{G}$ vale che $g_1 * g_2 = g_2 * g_1$, la coppia $(\mathscr{G}, *)$ è un *gruppo abeliano* o *commutativo.*

Esempi di gruppi abeliani e non possono essere:

- $(\mathbb{R}^n, +), (\mathbb{Z}, +), (\mathbb{Q}, +)$;
- L'insieme delle traslazioni è un gruppo abeliano;
- L'insieme delle rotazioni di centro fisso è un gruppo abeliano;
- $(B_{ii}(X), \circ)$ è un gruppo non abeliano (aka l'insieme delle funzioni biiettive definite su $X$ a valori in $X$);
- L'insieme delle rototraslazioni è un gruppo non abeliano. Una semplice dimostrazione/spiegazione può essere ottenuta grazie al teorema di isomorfismo: essendo le trasformazioni delle applicazioni lineari da $\mathbb{R}^n$ in $\mathbb{R}^n$, ad ognuna di esse è associata in modo univoco una matrice $M_n(f) \in M_n(\mathbb{R})$. Poiché il prodotto fra matrici non è commutativo, così non è commutativa la composizione di rotazioni e traslazioni.

**Definizione 3.** Si definisce *prodotto scalare* nello spazio vettoriale $\mathbb{R}^n$ una forma bilineare simmetrica definita positiva

$$\begin{aligned} \langle , \rangle : \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R} \\ (\mathbf{v}, \mathbf{w}) &\mapsto \langle \mathbf{v}, \mathbf{w} \rangle \end{aligned}$$

che soddisfi le seguenti proprietà:

1. Positività: $\forall \mathbf{v} \in \mathbb{R}^n\ \langle \mathbf{v}, \mathbf{v} \rangle \geq 0 \land \langle \mathbf{v}, \mathbf{v} \rangle = 0 \iff \mathbf{v} = \mathbf{0}$;
2. Linearità rispetto al primo (e al secondo, qui non riportata) termine: $\forall \mathbf{v_1}, \mathbf{v_2}, \mathbf{w} \in \mathbb{R}^n\ \langle \mathbf{v_1} + \mathbf{v_2}, \mathbf{w} \rangle = \langle \mathbf{v_1}, \mathbf{w} \rangle + \langle \mathbf{v_2}, \mathbf{w} \rangle$;
3. Simmetria: $\forall \mathbf{v}, \mathbf{w} \in \mathbb{R}^n\ \langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$.

**Definizione 4.** Sia $\mathbb{C}^n$ uno spazio vettoriale. Una *forma sesquilineare* sul campo $\mathbb{C}$ è una mappa

$$\begin{aligned}\phi\colon \mathbb{C}^n \times \mathbb{C}^n &\to \mathbb{C}\\ (\mathbf{v},\mathbf{w}) &\mapsto \phi(\mathbf{v},\mathbf{w})\end{aligned}$$

che soddisfi le seguenti proprietà:

1. $\forall \mathbf{v_1},\mathbf{v_2},\mathbf{w_1},\mathbf{w_2} \in \mathbb{C}^n\ \phi(\mathbf{v_1}+\mathbf{v_2},\mathbf{w_1}+\mathbf{w_2}) = \phi(\mathbf{v_1},\mathbf{w_1})+\phi(\mathbf{v_1},\mathbf{w_2})+\phi(\mathbf{v_2},\mathbf{w_1})+\phi(\mathbf{v_2},\mathbf{w_2})$;
2. $\forall \mathbf{v},\mathbf{w} \in \mathbb{C}^n \wedge \forall \lambda \in \mathbb{C}\ \phi(\lambda\mathbf{v},\mathbf{w}) = \lambda\phi(\mathbf{v},\mathbf{w})$;
3. $\forall \mathbf{v},\mathbf{w} \in \mathbb{C}^n \wedge \forall \lambda \in \mathbb{C}\ \phi(\mathbf{v},\lambda\mathbf{w}) = \bar{\lambda}\phi(\mathbf{v},\mathbf{w})$.

Se la forma sesquilineare è simmetrica, cioè se $\forall \mathbf{v},\mathbf{w} \in \mathbb{C}^n\ \phi(\mathbf{v},\mathbf{w}) = \overline{\phi(\mathbf{w},\mathbf{v})}$, viene chiamata *forma hermitiana.*

Il prodotto scalare standard su $\mathbb{R}^n$, data la base canonica $\mathscr{C} := \{\mathbf{e_1},\dots,\mathbf{e_n}\}$, è così definito:

$$\langle \mathbf{x},\mathbf{y}\rangle := \sum_{i=1}^{n} x_i y_i$$

ove con $x_i, y_i$ si intendono le componenti rispettivamente dei vettori $\mathbf{x},\mathbf{y}$ rispetto all'iesimo vettore della base ortonormale $\mathscr{C}$. La forma hermitiana standard su $\mathbb{C}^n$, data una base ortonormale $\mathscr{C} := \{\mathbf{e_1},\dots,\mathbf{e_n}\}$, è invece:

$$\phi(\mathbf{x},\mathbf{y}) = \sum_{i=1}^{n} x_i \bar{y}_i$$

La definizione di un prodotto scalare su $\mathbb{R}^n$ permette anche di definire la norma euclidea/standard.

**Definizione 5.** Sia $\mathbf{x} \in \mathbb{R}^n$ un vettore dello spazio vettoriale $\mathbb{R}^n$, spazio in cui abbiamo definito un prodotto scalare $\langle \mathbf{v},\mathbf{w}\rangle$. Si definisce *norma euclidea* del vettore $\mathbf{x}$

$$\|\mathbf{x}\| := \sqrt{\langle \mathbf{x},\mathbf{x}\rangle} = \sqrt{\sum_{i=1}^{n} x_i^2}$$

Valgono le seguenti proprietà:

1. $\forall \mathbf{x} \in \mathbb{R}^n\ \|\mathbf{x}\| \geq 0 \wedge \|\mathbf{x}\| = 0 \iff \mathbf{x} = \mathbf{0}$;
2. $\forall \mathbf{x} \in \mathbb{R}^n \wedge \forall \lambda \in \mathbb{R}\ \|\lambda\mathbf{x}\| = |\lambda|\,\|\mathbf{x}\|$;
3. $\forall \mathbf{x},\mathbf{y} \in \mathbb{R}^n\ \|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$;
4. $\forall \mathbf{x},\mathbf{y} \in \mathbb{R}^n\ \big|\|\mathbf{x}\| - \|\mathbf{y}\|\big| \leq \|\mathbf{x}-\mathbf{y}\|$.

**Teorema 1** (Di Cauchy-Schwarz)**.** $\forall \mathbf{x},\mathbf{y} \in \mathbb{R}^n\ \big|\langle \mathbf{x},\mathbf{y}\rangle\big| \leq \|\mathbf{x}\|\|\mathbf{y}\|$

*Dimostrazione.* Siano $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$ due vettori non nulli. Prendiamo $\lambda \in \mathbb{R}$. Abbiamo che

$$\begin{aligned}\|\mathbf{x}+\lambda\mathbf{y}\|^2 &= \langle \mathbf{x}+\lambda\mathbf{y},\mathbf{x}+\lambda\mathbf{y}\rangle = \langle \mathbf{x},\mathbf{x}\rangle + \lambda\langle \mathbf{y},\mathbf{x}\rangle + \lambda\langle \mathbf{x},\mathbf{y}\rangle + \lambda^2\langle \mathbf{y},\mathbf{y}\rangle\\ &= \|\mathbf{x}\|^2 + 2\lambda\langle \mathbf{x},\mathbf{y}\rangle + \lambda^2\|\mathbf{y}\|^2\end{aligned}$$

Scegliendo opportunamente $\lambda$ (in questo caso ponendolo uguale a $-\frac{\langle \mathbf{x},\mathbf{y}\rangle}{\|\mathbf{y}\|^2}$) si ottiene

$$\|\mathbf{x}+\lambda\mathbf{y}\|^2 = \|\mathbf{x}\|^2 - 2\frac{\langle \mathbf{x},\mathbf{y}\rangle^2}{\|\mathbf{y}\|^2} + \frac{\langle \mathbf{x},\mathbf{y}\rangle^2}{\|\mathbf{y}\|^2} = \|\mathbf{x}\|^2 - \frac{\langle \mathbf{x},\mathbf{y}\rangle^2}{\|\mathbf{y}\|^2}$$

Ricordando che $\|\mathbf{x}+\lambda\mathbf{y}\|^2 \geq 0$, abbiamo che

$$\|\mathbf{x}\|^2 - \frac{\langle \mathbf{x},\mathbf{y}\rangle^2}{\|\mathbf{y}\|^2} \geq 0; \quad \|\mathbf{x}\|^2 \geq \frac{\langle \mathbf{x},\mathbf{y}\rangle^2}{\|\mathbf{y}\|^2}; \quad \|\mathbf{x}\|^2\|\mathbf{y}\|^2 \geq \langle \mathbf{x},\mathbf{y}\rangle^2$$

Elevando ambo i membri alla 0.5 otteniamo, con i dovuti accorgimenti

$$\|\mathbf{x}\|\|\mathbf{y}\| \geq \big|\langle \mathbf{x}, \mathbf{y}\rangle\big|$$

L'asserto è dimostrato. $\square$

**Definizione 6.** Una *distanza* o *metrica* su $\mathbb{R}^n$ è una funzione

$$\begin{aligned} d\colon \mathbb{R}^n \times \mathbb{R}^n &\to \mathbb{R} \\ (\mathbf{x}, \mathbf{y}) &\mapsto d(\mathbf{x}, \mathbf{y}) \end{aligned}$$

definita, data una norma, nel modo seguente

$$d(\mathbf{x}, \mathbf{y}) := \|\mathbf{x} - \mathbf{y}\|$$

che verifica le seguenti proprietà:

1. $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n \ d(\mathbf{x}, \mathbf{y}) \geq 0 \wedge d(\mathbf{x}, \mathbf{y}) = 0 \iff \mathbf{x} = \mathbf{y}$;
2. $\forall \mathbf{x}, \mathbf{x} \in \mathbb{R}^n \ d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$;
3. *Disuguaglianza triangolare*: $\forall \mathbf{x}, \mathbf{y}, \mathbf{z} \in \mathbb{R}^n \ d(\mathbf{x}, \mathbf{y}) \leq d(\mathbf{x}, \mathbf{z}) + d(\mathbf{z}, \mathbf{y})$.

La coppia $(\mathbb{R}^n, d)$ viene chiamata *spazio metrico.* In particolare, in $\mathbb{R}^n$, data una base $\mathscr{C}$, la metrica è così definita:

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

**Definizione 7.** Sia $\mathbb{R}^n$ uno spazio metrico, e sia $\mathbf{x_0} \in \mathbb{R}^n$. Si dice *intorno sferico* di centro $\mathbf{x_0}$ e raggio $\rho > 0$ l'insieme dei punti di $\mathbb{R}^n$ che hanno distanza da $\mathbf{x_0}$ minore di $\rho$

$$B(x_0, \rho) := \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x_0}, \mathbf{x}) < \rho\}$$

In generale, un *intorno* di $\mathbf{x_0} \in \mathbb{R}^n$ è un sottoinsieme di $U(\mathbf{x_0}) \subset \mathbb{R}^n$ tale che

$$\mathbf{x_0} \in U(\mathbf{x_0}) \wedge U(\mathbf{x_0}) \supseteq B(\mathbf{x_0}, \rho) \text{ per un certo } \rho > 0$$

Abbiamo così definito una base $\mathcal{B}$[1] della topologia (detta standard) indotta dalla metrica euclidea su $\mathbb{R}^n$. È importante notare che variando la definizione di distanza (ad esempio, definendo la *p-distanza* $d_p\colon \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$, che manda la coppia $(\mathbf{x}, \mathbf{y})$ in $d_p(\mathbf{x}, \mathbf{y}) := \sqrt[p]{\sum_{i=1}^{n} |x_i - y_i|^p}$, oppure l'$\infty$*-distanza*, che manda la coppia $(\mathbf{x}, \mathbf{y})$ in $d_\infty := \max_{1 \leq i \leq n}\{|x_1 - y_i|\}$) cambia la topologia indotta su $\mathbb{R}^n$. Esempi di intorni sferici nei casi consueti $\mathbb{R}, \mathbb{R}^2$ e $\mathbb{R}^3$ sono:

- In $\mathbb{R}$, il segmento (i.e. estremi esclusi) di centro $\mathbf{x_0}$ e semilunghezza $\rho$;
- In $\mathbb{R}^2$, il cerchio di centro $\mathbf{x_0}$ e raggio $\rho$ esclusa la circonferenza di centro $x_0$ e raggio $\rho$;
- In $\mathbb{R}^3$, la sfera di centro $\mathbf{x_0}$ e raggio $\rho$, esclusa la superficie sferica di centro $\mathbf{x_0}$ e raggio $\rho$.

Si può anche dare una definizione assiomatica di intorno, in particolare nel modo seguente:

1. Ogni punto $\mathbf{x} \in \mathbb{R}^n$ ha un intorno $U(\mathbf{x})$ e $\mathbf{x} \in U(\mathbf{x})$;
2. $U_1(\mathbf{x}) \cap U_2(\mathbf{x})$ contiene un intorno di $\mathbf{x}$;
3. Se $\mathbf{y} \in B(\mathbf{x})$, allora $\exists U(\mathbf{y}) \subset U(\mathbf{x})$;
4. Dati $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n : \mathbf{x} \neq \mathbf{y}$, allora esistono $U(\mathbf{x})$ e $U(\mathbf{y})$ tali che $U(\mathbf{x}) \cap U(\mathbf{y})$ è nota (i.e. è l'insieme vuoto $\emptyset$)(*proprietà di separazione di Hausdorff*).

[1] Sia $\mathcal{T}$ una topologia su un insieme $X$. Una collezione $\mathcal{B} \subset \mathcal{T}$ viene chiamata *base* di $\mathcal{T}$ se ogni aperto $A \in \mathcal{T}$ può essere scritto come l'unione di elementi di $\mathcal{B}$.

**Definizione 8.** Un punto $\mathbf{x} \in \mathbb{R}^n$ si dirà *interno* ad $E \subseteq \mathbb{R}^n$ se $\exists \rho > 0 : B(\mathbf{x}, \rho) \subseteq E$. Si dirà altresì *esterno* ad $E$ se è interno al suo complemento $E^c := \mathbb{R}^n \setminus E$

**Definizione 9.** Dato un insieme $E \subseteq \mathbb{R}^n$, si dice che $\mathbf{x} \in \mathbb{R}^n$ è un *punto di frontiera* di $E$ se

$$\forall \rho > 0 \ B(\mathbf{x}, \rho) \cap E \neq \emptyset \wedge B(\mathbf{x}, \rho) \cap E^c \neq \emptyset$$

cioè ogni intorno di $\mathbf{x}$ interseca sia $E$, sia il suo complemento $E^c$.

Si suole indicare con $E^\circ$ l'insieme dei punti interni di $E$ e con $\partial E = \partial E^c$ l'insieme dei punti di frontiera di $E$. Esempi di punti interni ed esterni in questo caso sono:

- Dato un sottoinsieme $E \subset \mathbb{R} : E = (a, b)$, tutti i punti di $E$ sono punti interni: $E^\circ = E$ e $\partial E = \{a, b\}$;
- Dato un sottoinsieme $E \subset \mathbb{R} : E = [a, b]$, $E^\circ = (a, b)$ e $\partial E = \{a, b\}$;
- Dato un sottoinsieme $E \subset \mathbb{R} : E = [a, +\infty)$, $E^\circ = (a, +\infty)$ e $\partial E = \{a\}$;
- Sia $\mathbb{Q} \subset \mathbb{R}$; ogni intorno $U(x)$ di un qualsiasi $x \in \mathbb{R}$ contiene sia razionali, sia irrazionali. Quindi $\mathbb{Q}^\circ = \emptyset$ e $\partial\mathbb{Q} = \mathbb{R}$;
- Sia dato un intorno sferico $B(\mathbf{x}, \rho)$ di centro $\mathbf{x}$ e raggio $\rho$; $B^\circ = B(\mathbf{x}, \rho)$ e $\partial B = \{\mathbf{y} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{y}) = \rho\}$;
- Sia data una retta $E = \{(x, y) \in \mathbb{R}^2 : y = \alpha x\}$; $E^\circ = \emptyset$ e $\partial E = E$.

**Definizione 10.** Un punto $\mathbf{x} \in \mathbb{R}^n$ si dirà *punto di accumulazione* per l'insieme $E \subseteq \mathbb{R}^n$ se ogni intorno $U(\mathbf{x})$ di $\mathbf{x}$ contiene almeno un punto di $E \setminus \{\mathbf{x}\}$. Se un punto $\mathbf{x} \in E$ non è un punto di accumulazione per $E$, viene detto *punto isolato*.

Ogni punto interno di $E$ è un punto di accumulazione, mentre i punti di frontiera possono essere sia punti di accumulazione, sia punti isolati (è questo il caso del punto $\mathbf{c}$ dell'insieme $E := \{\mathbf{x} \in \mathbb{R}^n : d(\mathbf{x}, \mathbf{x_0}) \leq \rho\} \cup \{\mathbf{c}\}$). L'insieme dei punti di accumulazione di $E$ viene detto *derivato* di $E$ e viene indicato con $E'$. Se $E' = E$, di dice che $E$ è *perfetto*.
Se $E' = \emptyset$, l'insieme $E$ è detto *discreto*. N.B. Se un insieme è discreto, tutti i punti appartenenti all'insieme sono punti isolati, ma l'implicazione inversa non vale. Esempio: consideriamo la successione $\{S_n\}_{n \in \mathbb{N}} = \{x \in \mathbb{R} : x = \frac{1}{n}, n \in \mathbb{N}\} \subset (0, 1]$; in questo caso, ogni punto della successione è un punto isolato, ma $E' \neq \emptyset$, in quanto 0 è un punto di accumulazione! Un altro esempio peculiare e pregnante è costituito da $\mathbb{N} \subset \mathbb{R}^*$, ove con $\mathbb{R}^*$ si intende l'insieme dei reali esteso, ovvero $\mathbb{R} \cup \{-\infty, +\infty\}$. In questo caso infatti $+\infty$ costituisce un punto di accumulazione per $\mathbb{N}$!

**Teorema 2.** *Sia $E \subseteq \mathbb{R}^n$ un insieme, e sia $\mathbf{x}$ un suo punto di accumulazione. Ogni intorno $U(\mathbf{x})$ contiene infiniti punti di $E$.*

*Dimostrazione.* Supponiamo per assurdo che esista un intorno $U(\mathbf{x})$ che contenga un numero finto $n$ di punti di $E$, che indichiamo con $\mathbf{x_1}, \mathbf{x_2}, \ldots, \mathbf{x_n}$. Consideriamo l'intorno sferico $B(\mathbf{x}, \rho)$ di centro $\mathbf{x}$ e raggio $\rho$ definito nel modo seguente:

$$\rho = \min_{1 \leq i \leq n} d(\mathbf{x}, \mathbf{x_i})$$

Questo intorno sferico non contiene punti di $E$, a parte $\mathbf{x}$. Di conseguenza, considerando l'intorno $U(\mathbf{x})$ conincidente con il suddetto intorno sferico, abbiamo trovato un intorno di $\mathbf{x}$ che non contiene punti di $E$. Ma questo è un assurdo, in quanto $\mathbf{x}$ è un punto di accumulazione per $E$. L'assero è quindi dimostrato. $\square$

**Definizione 11.** Un insieme $E \subseteq \mathbb{R}^n$ viene detto *aperto* se ogni punto $\mathbf{x} \in E$ è un punto interno ad $E$ ($E^\circ = E$). $E$ è altresì *chiuso* se $E^c$ è aperto.

N.B. Questa è una definizione basata sulla topologia indotta dalla metrica standard!

**Teorema 3.** *$E \subseteq \mathbb{R}^n$ è aperto $\iff$ $E^c = \mathbb{R}^n \setminus E$ è chiuso.*

Sarebbe opportuno indicare subito una convenzione: $\emptyset$ è per convenzione aperto; ciò implica che $\mathbb{R}^n$ è chiuso. L'esperienza (e una dimostrazione a detta di Romeo difficile) ci dice che $\mathbb{R}^n$ è aperto, e quindi $\emptyset$ è anche chiuso.
Esempi pregnanti:

- L'intervallo $E \subset \mathbb{R} : E = (a, b)$ è aperto, mentre $\mathbb{R} \setminus E = (-\infty, a] \cup [b, +\infty)$ è chiuso;
- $\mathbb{Q}$ non è né aperto, né chiuso: infatti, anche $\mathbb{R} \setminus \mathbb{Q}$ non è né aperto, né chiuso (cfr. §Definizione 9, esempi);
- Sia data una retta $E = \{(x, y) \in \mathbb{R}^2 : y = \alpha x\}$; in questo caso, i semipiani individuati dalla retta sono aperti, e di conseguenza $E$ è chiuso.

**Teorema 4.** *1. Sia $\mathscr{F}$ una famiglia*[2] *di aperti di $\mathbb{R}^n$. Allora $\bigcup_i \mathscr{F}_i$ è un aperto di $\mathbb{R}^n$.*

*2. Sia $\mathscr{F}$ una famiglia* finita *di aperti di $\mathbb{R}^n$. Allora $\bigcap_i \mathscr{F}_i$ è un aperto di $\mathbb{R}^n$.*

*3. Sia $\mathscr{F}$ una famiglia di chiusi di $\mathbb{R}^n$. Allora $\bigcap_i \mathscr{F}_i$ è un chiuso di $\mathbb{R}^n$.*

*4. Sia $\mathscr{F}$ una famiglia* finita *di chiusi di $\mathbb{R}^n$. Allora $\bigcup_i \mathscr{F}_i$ è un chiuso di $\mathbb{R}^n$.*

**Teorema 5.** *Sia $E \subseteq \mathbb{R}^n$. Le seguenti affermazioni sono equivalenti:*

*1. $E$ è chiuso:*

*2. $\partial E \subseteq E$;*

*3. $E' = E$.*

**Definizione 12.** Sia $E \subseteq \mathbb{R}^n$. Chiamo *chiusura* di $E$ l'insieme $\bar{E} := E \cup \partial E$.

L'insieme è chiuso $\iff$ coincide con la sua chiusura.

**Definizione 13.** Sia $A \subseteq E$. $A$ è *denso* in $E$ se $\bar{A} = \bar{E}$.

Ad esempio, $\mathbb{Q}$ è denso in $\mathbb{R}$.

**Definizione 14.** Sia $E \subseteq \mathbb{R}^n$. $E$ si dice *limitato* se esiste un intorno sferico $B(\mathbf{x}, \rho)$ di centro $\mathbf{x}$ e raggio $\rho$ tale che $E \subseteq B(\mathbf{x}, \rho)$

La definizione da un'altra prospettiva: sia $D := \{d(\mathbf{x}, \mathbf{y}) : \mathbf{x}, \mathbf{y} \in E\} \subseteq \mathbb{R}$. Se $E$ è limitato, $\exists M \in \mathbb{R} : M \geq d \ \forall d \in D$. Per il teorema di completezza dei reali, $D$ ammette un estremo superiore. Definiamo $\sup D$ *diametro* di $E$.

**Teorema 6** (di Bolzano-Weierstrass)**.** *Un sottoinsieme $E \subseteq \mathbb{R}^n$ limitato ed infinito (i.e. $\#E \geq \aleph_0$) ammette almeno un punto di accumulazione.*

*Dimostrazione.* Per pigrizia (*"è una virtù"*) e semplicità, supponiamo di essere in $\mathbb{R}^2$. Consideriamo il rettangolo $T_0 := [p_0, q_0] \times [r_0, s_0]$ che contiene $E$ (che esiste in quanto $E$ è limitato). Consideriamo ora i 4 rettangoli ottenuti tracciando gli assi dei segmenti (gli intervalli $[p_0, q_0]$ e $[r_0, s_0]$). In almeno uno di questi sottoinsiemi, poiché per ipotesi $E$ è infinito, si troveranno infiniti elementi; indichiamo il suddetto sottoinsieme con $T_1 := [p_1, q_1] \times [r_1, s_1]$, ove con $p_1, q_1, r_1, s_1$ indichiamo, a seconda della scelta, uno dei punti inziali $p_0, q_0$ e il punto medio $\frac{p_0+q_0}{2}$ per il primo intervallo, e $r_0, s_0$ e $\frac{r_0+s_0}{2}$ per il secondo. Consideriamo le due, delle quattro che ne risultano (il procedimento è analogo per l'altro intervallo), successioni

$$\{P_n\}_{n \in \mathbb{N}} = \{p_0, p_1, \ldots, p_n, \ldots\} \qquad \{Q_n\}_{n \in \mathbb{N}} = \{q_0, q_1, \ldots, q_n, \ldots\}$$

[2]Siano $A$ e $\Omega$ due insiemi, e si supponga che ad ogni elemento $\alpha \in A$ sia associato un sottoinsieme $E_\alpha$ di $\Omega$. L'insieme i cui elementi sono gli insiemi $E_\alpha$ viene detto *famiglia* di insiemi. Più formalmente, una famiglia è una tripletta $(A, \Omega, i)$ costituita da due insiemi $A, \Omega$ e da una mappa $i \colon A \to \Omega$.

Esse sono entrambe monotone, in particolare la prima crescente e la seconda decrescente (non necessariamente strettamente), per costruzione. Si ha quindi che

$$p_0 \leq p_1 \leq \dots \qquad q_0 \geq q_1 \geq \dots$$

Siano dati $h, k \in \mathbb{N}$; $\forall n \geq h \wedge n \geq k$ vale che $p_h \leq p_n \leq q_n \leq q_k$. Abbiamo così dimostrato che la successione $\{P_n\}_{n\in\mathbb{N}}$ ammette un maggiorante; di conseguenza, essendo la successione $\{P_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$, per l'assioma di Dedekind (cfr. assioma di continuità o assioma di completezza) essa ammette un estremo superiore $\sup\{P_n\}_{n\in\mathbb{N}}$. Sappiamo quindi che $\sup\{P_n\}_{n\in\mathbb{N}} \leq q_k \ \forall k \in \mathbb{N}$; l'insieme $\{Q_n\}_{n\in\mathbb{N}}$ ammette quindi un estremo inferiore $\inf\{Q_n\}_{n\in\mathbb{N}}$. Si ha così che

$$\sup\{P_n\}_{n\in\mathbb{N}} \leq \inf\{Q_n\}_{n\in\mathbb{N}}$$

Dato che avevamo precedentemente identificato il comportamento delle successioni come rispettivamente crescente e decrescente, possiamo dedurre che $\lim_{n\to\infty} P_n = \sup\{P_n\}_{n\in\mathbb{N}}$ e $\lim_{n\to\infty} Q_n = \inf\{Q_n\}_{n\in\mathbb{N}}$. Possiamo altresì osservare che

$$\forall n \in \mathbb{N} \ q_n - p_n \geq \inf\{Q_n\}_{n\in\mathbb{N}} - \sup\{P_n\}_{n\in\mathbb{N}} \geq 0$$

Poiché per come abbiamo costruito le successioni $q_n - p_n = \frac{q-p}{2^n}$, abbiamo che

$$\frac{q-p}{2^n} \geq \inf\{Q_n\}_{n\in\mathbb{N}} - \sup\{P_n\}_{n\in\mathbb{N}} \geq 0$$

Per $n \to \infty$ la quantità a sinistra della catena di disuguaglianze tende a 0, e quindi si ha che $\inf\{Q_n\}_{n\in\mathbb{N}} = \sup\{P_n\}_{n\in\mathbb{N}}$. Indichiamo questo valore con $x_1$, mentre quello ottenuto ripetendo il procedimento per l'altro intervallo verrà denominato $x_2$. Sia $\mathbf{x} \in \mathbb{R}^2 : \mathbf{x} = (x_1, x_2)$: possiamo osservare che $\mathbf{x} = \bigcap_{n=0}^{\infty} T_n$ (in un certo senso può essere considerato un intorno degenere; osservazione per quanto viene detto dopo). Consideriamo l'intorno sferico $B(\mathbf{x}, \rho)$ di centro $\mathbf{x}$ e di raggio $\rho$. Per l'osservazione precedente, posso trovare $n(\rho)$ tale che $T_n(\rho) \subseteq B(\mathbf{x}, \rho)$. In particolare, $n$ deve essere tale da garantire che la massima possibile distanza all'interno di $T_n$ (maggiorata dalla diagonale del quadrato avente come lato il maggiore fra $\frac{q-p}{2^n}$ e $\frac{s-r}{2^n}$) $\sqrt{2}\frac{q-p}{2^n} < \rho$. Risolvendo si ottiene che $n \geq [\log_2(\sqrt{2}\frac{q-p}{\rho})] + 1$. Poiché per costruzione i $T_n$ contengono infiniti punti di $E$, l'asserto è dimostrato. □

**Definizione 15.** Sia $E \subseteq \mathbb{R}^n$. Sia $\mathscr{F}$ una famiglia di aperti di $\mathbb{R}^n$ tale che $\bigcup_i \mathscr{F}_i \supseteq E$. Diremo allora che $\mathscr{F}$ è un *ricoprimento* di $E$.

**Definizione 16.** Un *sottoricoprimento* di $X$ è una famiglia $\mathscr{G} \subseteq \mathscr{F}$ che ricopre $X$.

**Definizione 17.** Uno spazio topologico $X$ si definisce *compatto* se da ogni suo ricoprimento $\mathscr{F}$ è possibile estrarre un sottoricoprimento *finito* $\mathscr{G}$ tale che $\bigcup_{i=1}^{n} \mathscr{G}_i \supset X$.

È assai utile studiare gli spazi compatti perché sono molto simili a degli spazi finiti; infatti, il fatto che siano contenuti in ricoprimenti finiti consente sempre di "approssimare" l'intero spazio con un numero finito di punti, permettendo l'estensione agli spazi compatti di molti risultati dimostrabili negli insiemi finiti.

**Teorema 7** (di Heine-Borel)**.** *Sia $E \subseteq \mathbb{R}^n$. Allora $E$ è compatto $\iff$ $E$ è chiuso e limitato. Questa è definita la* caratterizzazione dei compatti in $\mathbb{R}^n$.

*Dimostrazione.* Per pigrizia e semplicità poniamoci in $\mathbb{R}^2$. La generalizzazione a $n$ qualsiasi non richiede altro se non iterazioni del procedimento.

1. Dimostriamo la condizione necessaria.

   - Verifichiamo prima l'implicazione della *limitatezza*. Considero gli intorni sferici $B(\mathbf{0}, n)$ di centro $\mathbf{0}$ e di raggio $n \in \mathbb{N}$. Si ha che $\bigcup_{n\in\mathbb{N}} B(\mathbf{0}, n) = \mathbb{R}^2$: quindi la famiglia $\mathscr{B} := \{B_1, \dots, B_n, \dots\}$ è un ricoprimento di $\mathbb{R}^2$. Poiché $E \subseteq \mathbb{R}^2$, $\mathscr{B}$ ricopre $E$. Ma $E$ per ipotesi è compatto, indi per cui da $\mathscr{B}$ posso estrarre un sottoricoprimento finito di aperti $\mathscr{E}$ che ricopre $E$. In altre parole, considero il sottoinsieme $\{k_1, \dots, k_N\} \subset \mathbb{N}$ tale che $\bigcup_{i=1}^{N} \mathscr{E}_{k_i} \supseteq E$. Posso spingermi oltre: poiché $\{k_1, \dots, k_N\}$ è un sottoinsieme finito di $\mathbb{N}$, posso facilmente individuarne il massimo $\bar{k} := \max\{k_1, \dots, k_N\}$. Vale che $B(\mathbf{0}, \bar{k}) \supseteq \bigcup_{i=1}^{N} \mathscr{E}_{k_i} \supseteq E$. Di conseguenza, $E$ è contenuto in un intorno sferico $B(\mathbf{0}, \bar{k})$ di centro $\mathbf{0}$ e raggio $\bar{k}$, ed è quindi limitato (cfr. Definizione 14).

- Verifichiamo ora l'implicazione della *chiusura*. Dimostro che il complementare in $\mathbb{R}^n$ di $E$ $E^c$ è aperto, ovvero che $\forall \mathbf{x} \in E^c\ \exists B(\mathbf{x}, \rho) : B(\mathbf{x}, \rho) \subseteq E^c$. Fissiamo $\mathbf{x}$. Per ogni $\mathbf{y} \in E$ definisco $\delta(\mathbf{y}) := \frac{\|\mathbf{x}-\mathbf{y}\|}{2}$. Considero l'intorno sferico $B(\mathbf{y}, \delta(\mathbf{y}))$ al variare di $\mathbf{y}$ in $E$. Ottengo una famiglia $\mathscr{B}$ di aperti che ricopre $E$. Poiché $E$ è compatto, da $\mathscr{B}$ posso estrarre un sottoricoprimento finito di $E$: esistono $\mathbf{y_1}, \dots, \mathbf{y_N} \in E$ : $\bigcup_{i=1}^{N} B(\mathbf{y_i}, \delta(\mathbf{y_i})) \supseteq E$. Considero il minimo delle distanze $\delta := \min\{\delta(\mathbf{y_1}), \dots, \delta(\mathbf{y_N})\}$. A $\delta$ corrisponde, in quanto funzione degli $\mathbf{y_i}$, un certo $\bar{\mathbf{y}}$. Considero ora l'intersezione $B(\mathbf{x}, \delta) \cap B(\bar{\mathbf{y}}, \delta)$ che è uguale all'insieme vuoto, in quanto sono entrambi intorni sferici di raggio delta, che ho definito come la semilunghezza del segmento $[\mathbf{x}, \mathbf{y}]$. Ciò vale anche per qualsiasi $\mathbf{y_i}, i = 1, \dots, N$. Quindi $B(\mathbf{x}, \delta) \cap \bigcup_{i=1}^{N} B(\mathbf{y_i}, \delta(\mathbf{y_i})) = \emptyset$. Ma allora $B(\mathbf{x}, \delta) \subset E^c$. Poiché vale per un qualsiasi $\mathbf{x} \in E$, $E^c$ è aperto, e quindi $E$ è chiuso.

2. Devo ora dimostrare la condizione sufficiente; per farlo, passo attravero la dimostrazione che un qualsiasi rettangolo $T := [a, b] \times [c, d]$ è un compatto e che ogni sottoinsieme chiuso di un compatto è un compatto.

   - Supponiamo che da un ricoprimento $\mathscr{F}$ del rettangolo $T$ non sia possibile estrarre un sottoricoprimento finito di $T_0$. Procediamo in un modo analogo a quello che si utilizza nella dimostrazione del teorema di Bolzano-Weierstrass (cfr. Teorema 6): dati $m_j := \frac{b_j - a_j}{2}$ e $n_j := \frac{d_j - c_j}{2}$, si considerino i segmenti $[a_j, m_j]$, $[m_j, b_j]$, $[c_j, n_j]$ e $[n_j, d_j]$, che individuano 4 rettangoli di cui almeno uno (sia esso $T_j$) non può essere ricoperto da una sottofamiglia finita di $\mathscr{F}$ (per l'ipotesi iniziale, ndr). Si ripeta il procedimento per $T_j$. Si ottiene una successione $\{T_n\} := \{T_0 \supset T_1 \supset \dots \supset T_n \supset \dots\}$ di sottoinsiemi di $T_0$ non ricopribili da una famiglia finita di aperti la cui intersezione è il punto $\mathbf{x} = (x_1, x_2) = \bigcap_{n \in \mathbb{N}} T_n$. Poiché $\mathbf{x} \in T_0$, posso trovare un aperto $U \in \mathscr{F}$ tale che $\mathbf{x} \in U$. Poiché $U$ è aperto, posso scegliere un certo $\rho > 0$ tale che $B(\mathbf{x}, \rho) \subset U$. A questo punto, posso trovare dei rettangoli $T_n$ contenuti in questo intorno: $\exists \bar{n}(\rho)$ : $T_n \subset B(\mathbf{x}, \rho)\ \forall n \geq \bar{n}$. Si ha così che $T_n \subset U$, e quindi esiste una sottofamiglia finita di aperti che ricopre $T_n$. Abbiamo un assurdo, e di conseguenza il rettangolo $T_0$ è compatto.
   - Siano $E \subset F$ un chiuso e $F$ un compatto. Per definizione di compattezza, esiste una famiglia finita di aperti $\mathscr{F} := \{\mathscr{F}_1, \dots, \mathscr{F}_n\}$ che ricopre $F$. Considero ora un generico ricoprimento $\mathscr{E}$ di $E$ e il complementare del suddetto insieme $E^c := \mathbb{R}^n \setminus E$. Vale che $\bigcup_i \mathscr{E}_i \cup E^c \supset F$ (il ricoprimento di $E$ contiene $E$, mentre il complementare di $E$ contiene tutti gli elementi appartenenti a $F \setminus E$). Ho quindi individuato un'altra famiglia di aperti (il ricoprimento è una famiglia di aperti e $E^c$ è aperto perché $E$ è chiuso per ipotesi) che ricopre $F$. Ma quindi, poiché $F$ è compatto, da quest'ultima famiglia è possibile estrarre un sottoricoprimento finito di $F$. Poiché la cardinalità della famiglia dipende (quasi) esclusivamente da $\mathscr{E}$, ciò implica che esiste un sottoricoprimento finito $\mathscr{E}' \subset \mathscr{E}$ di $E$, e quindi $E$ è compatto (ricordiamo che $\mathscr{E}$ è un ricoprimento generico di $E$!).

Per tornare all'enunciato iniziale, essendo $E$ limitato è inscrivibile, dati un certo $\rho > 0$ e un $\mathbf{x} \in E$, in un intorno sferico $B(\mathbf{x}, \rho)$, e di conseguenza in un rettangolo $T(\rho)$ sufficientemente grande (che è un compatto). Essendo $E$ poi chiuso, ed essendo contenuto in un compatto, è a sua volta un compatto, e l'asserto è dimostrato.

□

Esempi di insiemi compatti sono quindi:

- Ogni intervallo chiuso $[a, b] \subset \mathbb{R}$;
- Ogni chiusura di un intorno sferico $\bar{B}(\mathbf{x_0}, \rho)$ di centro $\mathbf{x_0}$ e raggio $\rho$ (un intorno sferico invece non è un compatto! $\partial B \not\subset B$).

**Definizione 18.** Siano $E_1, E_2 \subseteq \mathbb{R}^n$. Si dice che $E_1$ e $E_2$ sono *separati* se valgono contemporaneamente le seguenti proprietà:

1. $\bar{E}_1 \cap E_2 = \emptyset$
2. $\bar{E}_2 \cap E_1 = \emptyset$

**Definizione 19.** Sia $E \subseteq \mathbb{R}^n$. Diremo che uno spazio topologico $X$ è *sconnesso* se esistono due insiemi $A, B$ aperti, non vuoti e disgiunti tali che $X = A \cup B$. Altrimenti si dice *connesso.*

**Definizione 20.** Uno spazio topologico $X$ si dice *connesso per archi* se $\forall (x, y) \in X$ esiste una funzione continua $\alpha \colon [0, 1] \to X$ tale che $\alpha(0) = x \wedge \alpha(1) = y$

**Definizione 21.** Sia $E \subseteq \mathbb{R}^n$. $E$ è detto *dominio* se è non vuoto, aperto, connesso.

# 2 Funzioni da $\mathbb{R}^n$ a elementi in $\mathbb{R}^m$

**Definizione 22.** Dato un insieme $X \subseteq \mathbb{R}^n$, una *funzione scalare* è una funzione $f \colon X \to \mathbb{R}$. Sono particolarmente interessanti perché, oltre a rappresentare i *campi scalari*, godono di molte delle proprietà delle funzioni $g \colon \mathbb{R} \to \mathbb{R}$, in quanto $f(X) \subseteq \mathbb{R}$.

Esempi di funzioni scalari possono essere:

- la norma euclidea $\| \quad \| \colon \mathbb{R}^n \to \mathbb{R}$;
- la funzione di proiezione $\pi_k \colon \mathbb{R}^n \to \mathbb{R}$ che associa ad ogni vettore $\mathbf{x}$ la sua k-esima componente $x_k$;
- *fissato* $\mathbf{x_0} \in \mathbb{R}^n$, il prodotto scalare standard $\langle , \rangle$ può essere visto come una funzione scalare, così come la distanza $d$;
- la funzione gaussiana $f \colon \mathbb{R}^n \to \mathbb{R}$ che associa ad ogni elemento $\mathbf{x} \in \mathbb{R}^n$ l'elemento $f(\mathbf{x}) := e^{-\|\mathbf{x}\|^2}$.

**Definizione 23.** Siano $X \subseteq \mathbb{R}^n$, $f$ una funzione $f \colon X \to \mathbb{R}$ e $\mathbf{x_0} \in \dot{\mathbb{R}}^n$ un punto di accumulazione per $X$. Consideriamo $l \in \mathbb{R}^*(\dot{\mathbb{R}})$. Diremo che $f(\mathbf{x}) \to l$ per $\mathbf{x} \to \mathbf{x_0}$ o

$$\lim_{\mathbf{x} \to \mathbf{x_0}} f(\mathbf{x}) = l$$

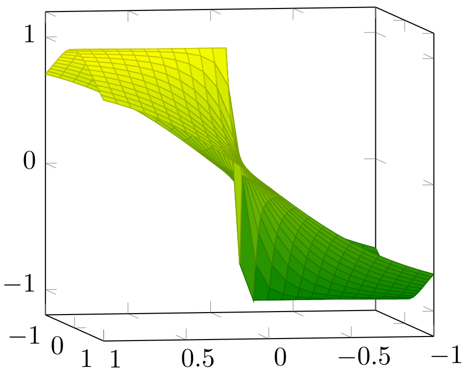
se $\forall V(l)\ \exists U(\mathbf{x_0}) : f(\mathbf{x}) \in V(l) \setminus \{l\}\ \forall \mathbf{x} \in (U(\mathbf{x_0}) \setminus \{\mathbf{x_0}\}) \cap X$.

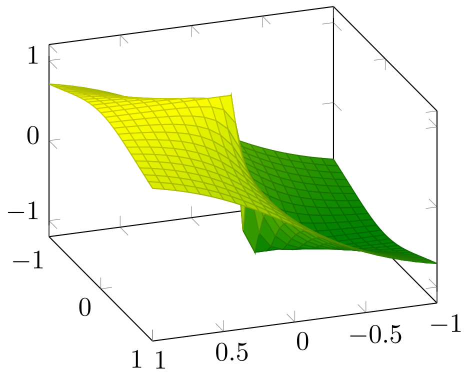
Da questa definizione è possibile risalire a quella classica, se si considerano $\mathbf{x_0} \in X' \wedge l \in \mathbb{R}$:

$$\forall B(l, \epsilon)\ \exists B(\mathbf{x_0}, \rho(\epsilon)) : d_{\mathbb{R}}(f(\mathbf{x}), l) < \epsilon\ \forall x \in X : d_{\mathbb{R}^n}(\mathbf{x}, \mathbf{x_0}) < \rho(\epsilon)$$

È di vitale importanza notare però che l'esistenza del limite per $\mathbb{R}^n$ con $n \geq 2$ è molto differente da quella di $n = 1$: infatti il limite esiste se e solo se esistono (e coincidono) i limiti in *qualsiasi* direzione; di conseguenza, questa direzione viene generalmente utilizzata per dimostrare la non-esistenza del limite. Ad esempio, si consideri la funzione

$$\begin{aligned} f \colon \mathbb{R}^2 \setminus \{\mathbf{0}\} &\to \mathbb{R} \\ (x_1, x_2) &\mapsto f(x_1, x_2) := \frac{x_1}{\|\mathbf{x}\|} \end{aligned}$$






In questo caso:

- per i punti del tipo $(0, x_2), x_2 \in \mathbb{R},\ f(\mathbf{x}) = 0$;
- per i punti del tipo $(x_1, 0), x_1 \in \mathbb{R},\ f(\mathbf{x}) = \pm 1$.

Come si può facilmente intuire, determinare l'esistenza o meno del limite in questo modo non è un processo immediato; richiederebbe infatti un numero di iterazioni infinito. Esiste tuttavia un trucco (per $\mathbb{R}^2$) utile per semplificare il tutto, e consiste nell'utilizzare le coordinate polari $x_1 = \rho\cos(\theta), x_2 = \rho\cos(\theta), \theta \in [0, 2\pi]$. In questo caso infatti, nel $\lim_{\mathbf{x}\to\mathbf{0}} f(\mathbf{x})$ la distanza $d(\mathbf{x}, \mathbf{0}) \to 0$ e, poiché la distanza altri non è se non $\rho$, si ha che

$$\lim_{\mathbf{x}\to\mathbf{0}} f(\mathbf{x}) = \lim_{\rho\to 0} \tilde{f}(\rho, \theta)$$

In particolare, se dimostro che il $\lim_{\rho\to 0}$ può essere fatto uniformemente (i.e. indipendentemente) rispetto a $\theta$, il limite esiste.

$$\forall \epsilon > 0\ \exists \delta(\epsilon) > 0 : \tilde{f}(\rho, \theta) - l < \epsilon\ \forall \rho \in (0, \delta(\epsilon))$$

**Definizione 24.** Consideriamo una funzione $f\colon X \subseteq \mathbb{R}^2 \to \mathbb{R}$, e sia $\mathbf{x_0} = (0, 0)$ (per pigrizia e semplicità) un punto di accumulazione per $X$. Se esiste una funzione $g\colon U(0) \to \mathbb{R} : \left|f(\rho, \theta) - l\right| < g(\rho)\ \forall \rho \in U(0) \wedge \lim_{\rho\to 0} g(\rho) = 0$, allora $\lim_{\rho\to 0} f(\rho, \theta) = \lim_{(x,y)\to(0,0)} f(x, y) = l$.

A titolo di esempio, si consideri la funzione $f\colon \mathbb{R}^2 \to \mathbb{R}$ che manda l'elemento $\mathbf{x} = (x, y)$ nell'elemento $f(\mathbf{x}) = \frac{x^3}{x^2+y^2}$. Si consideri innanzitutto la restrizione della funzione al dominio $X := \{(x, y) \in \mathbb{R}^2 : y = mx\}$, in modo da capire quale possa essere il limite nel caso esista:

$$\lim_{(x,y)\to(0,0)} \frac{x^3}{x^2 + y^2} = \lim_{x\to 0} \frac{x^3}{x^2 + (mx)^2} = \lim_{x\to 0} \frac{x^3}{x^2(1 + m^2)} = \lim_{x\to 0} \frac{x}{1 + m^2} = 0$$

Ricorrendo ora alle coordinate polari, si ha che

$$\left|f(x, y)\right| = \left|\tilde{f}(\rho, \theta)\right| = \left|\frac{\rho^3 \cos^3(\theta)}{\rho^2(\cos^2(\theta) + \sin^2(\theta))}\right| = \left|\rho\cos^3(\theta)\right| \leq |\rho|$$

Abbiamo quindi trovato una funzione $g(\rho)$ che tende a 0 per $\rho$ che tende a 0. Il $\lim_{\rho\to 0} \tilde{f}(\rho, \theta)$ converge uniformemente rispetto a $\theta$, e quindi

$$\lim_{(x,y)\to(0,0)} f(x, y) = \lim_{\rho\to 0} \tilde{f}(\rho, \theta) = 0$$

**Definizione 25.** Una funzione $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$ che manda, date una base $\mathscr{C} = \{\mathbf{e_1}, \ldots, \mathbf{e_m}\}$ di $\mathbb{R}^m$, l'elemento $\mathbf{x} \in X$ nell'elemento $f(\mathbf{x}) = \mathbf{y} = (y_1, y_2, \ldots, y_m) = (f_1(\mathbf{x}), \ldots, f_m(\mathbf{x}))$ può essere vista come una m-upla di funzioni scalari $f_1, \ldots, f_m\colon X \to \mathbb{R}$.

Data la definizione di numero complesso come coppia $(a, b)$ di numeri reali, è possibile creare una corrispondenza naturale tra $f\colon \mathbb{C} \to \mathbb{C}$ e $f\colon \mathbb{R}^2 \to \mathbb{R}^2$, tendendo a mente che le operazioni sono sempre definite su $\mathbb{C}$. Ad esempio, si consideri

$$f(z) = z^2 = (x + iy)^2 = (x^2 - y^2 + 2ixy)$$
$$f(x, y) = (x, y)(x, y) = (u(x, y), v(x, y)) = (x^2 - y^2, 2xy)$$

**Definizione 26.** Siano $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{x_0} \in \dot{\mathbb{R}}^n$ un punto di accumulazione per $X$ e $\mathbf{l} \in \dot{\mathbb{R}}^m$. Diremo che $f(\mathbf{x}) \to \mathbf{l}$ per $\mathbf{x} \to \mathbf{x_0}$ o

$$\lim_{\mathbf{x}\to\mathbf{x_0}} f(\mathbf{x}) = \mathbf{l}$$

se $\forall V(\mathbf{l}) \subseteq \dot{\mathbb{R}}^n\ \exists U(\mathbf{x_0}) \subseteq \dot{\mathbb{R}}^n : f(\mathbf{x}) \in V(\mathbf{l}) \forall \mathbf{x} \in (U(\mathbf{x_0}) \setminus \{\mathbf{x_0}\}) \cap X$

**Teorema 8.** *Siano $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $\mathbf{x_0} \in \dot{\mathbb{R}}^n$ un punto di accumulazione per $X$ e $\mathbf{l} \in \dot{\mathbb{R}}^m$. Data una base $\mathscr{C} = \{\mathbf{e_1}, \dots, \mathbf{e_m}\}$ di $\dot{\mathbb{R}}^m$ si ha che*

$$\lim_{\mathbf{x}\to\mathbf{x_0}} f(\mathbf{x}) = \mathbf{l} \iff \lim_{\mathbf{x}\to\mathbf{x_0}} f_i(\mathbf{x}) = l_i\ \forall i$$

**Teorema 9.** *Sia $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ e sia $\mathbf{x_0} \in X$. Vale che*

$$\lim_{\mathbf{x}\to\mathbf{x_0}} f(\mathbf{x}) = l \iff \lim_{\mathbf{x}\to\mathbf{x_0}} f|_A(\mathbf{x}) = l\ \forall A \subseteq X \wedge \mathbf{x_0} \in A'$$

Questo teorema viene generalmente utilizzato in una ben precisa castistica:

1. Per trovare un candidato limite in modo più semplice restringendo ad un dominio comodo, per poi andare a controllare se il numero trovato sia effettivamente il limite attraverso la definizione;
2. Per verificare la non-esistenza di un limite, trovando, restringendo a due domini differenti, due limiti diversi.

**Definizione 27.** Una *successione* in $\mathbb{R}^n$ è un'applicazione $x\colon \mathbb{N} \to \mathbb{R}^n$, dove per convenzione si denota $x(n)$ con $x_n$ e si utilizza $\{x_n\}_{n\in\mathbb{N}}$ per indicare l'immagine dell'applicazione stessa.

**Definizione 28.** Sia $\{x_n\}_{n\in\mathbb{N}}$ una successione di punti di $\mathbb{R}^n$. Diremo che $\{x_n\}_{n\in\mathbb{N}}$ *converge* al punto $\mathbf{x_0} \in \mathbb{R}^n$ e scriveremo $\lim_{n\to\infty} x_n = \mathbf{x_0}$ se

$$\forall \epsilon > 0\ \exists n^*(\epsilon) \in \mathbb{N} : d(x_n, \mathbf{x_0}) < \epsilon\ \forall n > n^*$$

**Teorema 10.** *La successione $\{x_n\}_{n\in\mathbb{N}}$ converge $\iff$ la successione di ogni singola componente $\{x_n^i\}_{n\in\mathbb{N}}$ converge alla corrispondente componente $x_0^i$.*

*Dimostrazione.* Supponiamo che la successione $\{x_n\}_{n\in\mathbb{N}}$ converga a $\mathbf{x_0}$ (vedasi la definizione di convergenza). Siccome si ha che, dato $\mathbf{v} \in \mathbb{R}^n$, $|v_i| \le \|\mathbf{v}\| \le \sqrt{n} \max_{1\le i\le n}|v_i|$, possiamo scrivere che

$$\left|x_n^i - x_0^i\right| \le \|x_n - \mathbf{x_0}\| \le \epsilon\ \forall \epsilon \wedge \forall i$$

Abbiamo quindi dimostrato la prima implicazione.
Si supponga ora che la successione di ogni singola componente $\{x_n^i\}_{n\in\mathbb{N}}$ converga alla corrispondente componente ${x_0}^i$. Poiché in questo caso le successioni sono funzioni $x^i\colon \mathbb{N} \to \mathbb{R}$, si ha che $\forall \epsilon > 0\ \exists \bar{n}(\epsilon) \in \mathbb{N} : \left|x_n^i - x_0^i\right| < \epsilon\ \forall n > \bar{n}$. Si prenda ora $h = \max_{1\le i\le n}\left|x_n^i - x_0^i\right|$. Si ha che

$$\left|x_n^i - x_0^i\right| \le \|x_n - \mathbf{x_0}\| \le \sqrt{n}h < \sqrt{n}\epsilon\ \forall \epsilon \wedge \forall i$$

Poiché si ha che $\|x_n - \mathbf{x_0}\| < \sqrt{n}\epsilon$, la successione $\{x_n\}_{n\in\mathbb{N}}$ converge, e l'asserto è dimostrato. □

**Teorema 11.** *Ogni successione a valori in $\mathbb{R}^n$ convergente è limitata, ove con limitata si intende dire che $\forall k \in \mathbb{N}\ \|x_k\| < \rho$ per un certo $\rho > 0$.*

*Dimostrazione.* Sia data la successione convergente a $\mathbf{x_0} \in \mathbb{R}^n$ $\{x_n\}_{n\in\mathbb{N}}$. Per il teorema precedente, si ha che la successione di ogni singola componente $\{x_n^i\}_{n\in\mathbb{N}}$ converge alla rispettiva componente $x_0^i$. Queste sono successioni $x^i\colon \mathbb{N} \to \mathbb{R}$, e in questo caso, poiché sono convergenti, sono successioni di Cauchy. Vale che una successione di Cauchy è limitata:

$$\forall \epsilon > 0\ \exists \bar{n}(\epsilon) \in \mathbb{N} : d(x_n^i, x_0^i) < \epsilon \iff x_0^i - \epsilon < x_n^i < x_0^i + \epsilon\ \forall n > \bar{n}$$

Considero ora $m := \max_{n\le\bar{n}} d(x_n^i, x_0^i)$, e definisco $r^i := \max\left(\left|x_0^i\right| + \epsilon, m\right)$. So quindi che $\left|x_n^i\right| < r^i\ \forall n$. Posso quindi dire che $\|x_n\|^2 < \sum_{i=1}^n (r^i)^2$; definisco di conseguenza $\rho := \sqrt{\sum_{i=1}^n (r^i)^2}$. L'asserto è dimostrato. □

**Teorema 12.** *Sia $E \subseteq \mathbb{R}^n$, e sia $y \in E'$ un punto di accumulazione per $E$. Esiste una successione $\{x_n\}_{n\in\mathbb{N}} \subseteq E$ che converge a $y$.*

*Dimostrazione.* Sia dato un certo $\rho_1 > 0$. Poiché $\mathbf{x_0}$ è un punto di accumulazione per $E$, per definizione ogni intorno sferico $B(\mathbf{x_0}, \rho)$ di centro $\mathbf{x_0}$ e raggio $\rho$ contiene un punto $\mathbf{x} \in E \setminus \{\mathbf{x_0}\}$. Indico con $x_1$ suddetto punto, e definisco $\rho_2 := d(\mathbf{x}, \mathbf{x_0})$. Iterando il processo sopra riportato $n$ volte, si otterrà una successione $\{x_n\}_{n\in\mathbb{N}}$ di punti di $E$ il cui $\lim_{n\to\infty} x_n$ è $\mathbf{x_0}$. □

**Teorema 13.** *Da una successione limitata posso estrarre una sottosuccessione convergente.*

**Teorema 14.** *Siano $E \subseteq \mathbb{R}^n$ e $\mathbf{y} \in \mathbb{R}^n$. Allora $\mathbf{y} \in \bar{E} \iff$ esiste una successione $\{x_n\}_{n\in\mathbb{N}} \subset E$ tale che $\lim_{n\to\infty} x_n = \mathbf{y}$.*

*Dimostrazione.*

1. Iniziamo con il dimostrare la condizione necessaria. Possono presentarsi due casi:
   - $\mathbf{y} \in E$ (se E coincide con la chiusura), ed in questo caso la soluzione è triviale: prendo la successione $\{x_n\}_{n\in\mathbb{N}}$ tale che $x_n = \mathbf{y}$ $\forall n \in \mathbb{N}$;
   - $\mathbf{y} \in E'$, e in questo caso per il teorema 12 esiste la suddetta successione.
2. Sia $\{x_n\}_{n\in\mathbb{N}}$ una successione di punti di $E$ tale che $x_n \to \mathbf{y}$ per $n \to \infty$. Supponiamo che $\mathbf{y} \in \bar{E}^c$, ove $\bar{E}^c$ indica il complementare in $\mathbb{R}^n$ della chiusura di $E$. Tale insieme è aperto: esiste quindi per defnzione un intorno sferico $U(\mathbf{y}, \rho) \subset \bar{E}^c$. Ma allora non esiste una successione di punti di $E$ convergente a $\mathbf{y}$, in quanto questi definitivamente (i.e. dopo un certo $n(\rho)$) dovrebbero appartenere a $B(\mathbf{y}, \rho)$ e quindi a $E^c \subset \bar{E}^c$. Siamo giunti ad un assurdo, e l'asserto è dimostrato.

□

**Teorema 15.** *Sia $E \subseteq \mathbb{R}^n$. $E$ è chiuso in $\mathbb{R}^n \iff$ ogni successione convergente $\{x_n\}_{n\in\mathbb{N}} \subset E$ ha come punto limite un elemento di $E$.*

*Dimostrazione.*

1. Se $E$ è chiuso $E \equiv \bar{E}$, e quindi per il teorema 14 posso creare una successione convergente per ogni punto $\mathbf{x} \in E$;
2. Prendiamo $\mathbf{y} \in E'$. Allora esiste una successione $\{x_n\}_{n\in\mathbb{N}} \subset E$ tale che $\lim_{n\to\infty} x_n = \mathbf{y}$. Ma per ipotesi il punto limite di $\{x_n\}_{n\in\mathbb{N}}$ appartiene ad $E$: quindi $\forall \mathbf{y} \in E'$ $\mathbf{y} \in E$. Di conseguenza $E \equiv \bar{E}$.

□

**Definizione 29.** Sia $K \subseteq \mathbb{R}^n$. Diremo che $K$ è *compatto per successioni* o *sequenzialmente compatto* se da ogni successione $\{x_n\}_{n\in\mathbb{N}} \subseteq K$ è possibile estrarre una sottosuccessione convergente ad un elemento di $K$.

**Teorema 16.** *Sia $K \subseteq \mathbb{R}^n$. $K$ è compatto $\iff$ $K$ è sequenzialmente compatto.*

*Dimostrazione.*

1. Si supponga che $K$ sia compatto (il che vuol dire, per il Teorema di Heine-Borel, che $K$ è chiuso e limitato), e si consideri una successione $\{x_n\}_{n\in\mathbb{N}}$ (che è limitata in quanto $K$ è limitato). Se l'insieme costituito dai punti della successione è finito, consideriamo la successione (i.e. la sottosuccessione $\{p_n\}_{n\in\mathbb{N}} \subset \{x_n\}_{n\in\mathbb{N}}$) tale che $p_1 = \cdots = p_n = \cdots = p$. Se l'insieme è invece infinito, per il Teorema di Bolzano-Weierstrass ammette almeno un punto di accumulazione $\mathbf{y_0}$. Esiste allora una successione di elementi di $\{p_n\}_{n\in\mathbb{N}}$ che converge a $\mathbf{y_0}$. Ho quindi trovato una sottosuccessione di $\{x_n\}_{n\in\mathbb{N}}$ convergente ad un elemento di $K$. Poiché la successione $\{x_n\}_{n\in\mathbb{N}}$ è una successione generica, $K$ è sequenzialmente compatto.
2. Si supponga che $K$ sia sequenzialmente compatto. Sia $\mathbf{y} \in \bar{K}$; per il teorema 14, esiste una successione $\{x_n\}_{n\in\mathbb{N}} \subset K$ tale che $\lim_{n\to\infty} x_n = \mathbf{y}$. Poiché $K$ è sequenzialmente compatto, da $\{x_n\}_{n\in\mathbb{N}}$ posso estrarre una sottosuccessione $\{x_i\}_{i\in\mathbb{N}}$ tale che $\lim_{i\to\infty} x_i = \mathbf{y_0}$, con $\mathbf{y_0} \in K$. Essendo $\{x_i\}_{i\in\mathbb{N}}$ una sottosuccessione di una successione convergente, $\mathbf{y_0} = \mathbf{y}$. Per l'arbitrarietà di $\mathbf{y}$, $K$ è chiuso. Supponiamo ora che $K$ sia illimitato (superiormente

o inferiormente, la scelta è irrilevante). Sia $\{x_n\}_{n\in\mathbb{N}}$ la successione costruita nel modo seguente: $x_n := \max\{x_{n-1}, n\}$. Da questa successione non posso estrarre una sottosuccessione convergente; si ha quindi un assurdo, e di conseguenza $K$ è limitato. Per il Teorema di Heine-Borel, $K$ è compatto.

□

**Definizione 30.** Sia data una successione $\{x_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^n$. Diremo che $\{x_n\}_{n\in\mathbb{N}}$ è una *successione di Cauchy* se $\forall\epsilon > 0\ \exists\bar{n}(\epsilon) : d(x_n, x_m) < \epsilon\ \forall n, m > \bar{n}$.

# 3 Continuità

**Definizione 31.** Siano $X \subseteq \mathbb{R}^n$ e $f\colon X \to \mathbb{R}^m$. Se un punto $\mathbf{x_0} \in X$ è isolato diremo che $f$ è continua in $\mathbf{x_0}$. Altrimenti, sia $x_0 \in (X \cap X')$. Diremo che $f$ è *continua* in $\mathbf{x_0}$ se valgono le seguenti condizioni equialenti:

1. $\forall U(f(\mathbf{x_0})) \subseteq \mathbb{R}^m\ \exists U(\mathbf{x_0}) \subseteq \mathbb{R}^n$ tale che $\forall \mathbf{x} \in U(\mathbf{x_0})\ f(\mathbf{x}) \in U(f(\mathbf{x_0}))$;
2. $\forall\epsilon > 0\ \exists\delta(\epsilon) > 0$ tale che $\forall\mathbf{x} : d(\mathbf{x}, \mathbf{x_0}) < \delta\ d(f(\mathbf{x}), f(\mathbf{x_0})) < \epsilon$;
3. $\lim_{\mathbf{x}\to\mathbf{x_0}} f(\mathbf{x}) = f(\mathbf{x_0})$;
4. $\forall\{x_n\}_{n\in\mathbb{N}} \subset X$ che converge a $\mathbf{x_0}$, il $\lim_{\mathbf{x}\to\mathbf{x_0}} f(\mathbf{x}) = f(\mathbf{x_0})$.

Si consideri ad esempio, fissata una base $\mathscr{C}$ di $\mathbb{R}^n$, la funzione di proiezione $\pi_k\colon \mathbb{R}^n \to \mathbb{R}$. Proviamo a dimostrare che il $\lim_{\mathbf{x}\to\mathbf{x_0}} \pi_k(\mathbf{x}) = x_{0_k}$. Usando la definizione, $\forall\epsilon > 0\ \exists\delta(\epsilon) : \|\mathbf{x} - \mathbf{x_0}\| < \delta \implies |x_k - x_{0_k}| < \epsilon$. Ricordando che vale che $|x_i| \leq \|\mathbf{x}\| \leq \sqrt{n}\max_{i=1,\dots,n}\{|x_i|\}$, vale che $|x_k - x_{0_k}| \leq \|\mathbf{x} - \mathbf{x_0}\| < \delta$. Prendendo $\delta = \epsilon$ la condizione di continuità è verificata (per ogni $\mathbf{x} \in \mathbb{R}^n$).
Siano $f, g$ due funzioni continue in $X \subseteq \mathbb{R}^n$. Vale che:

- $(f + g)(\mathbf{x}) = f(\mathbf{x}) + g(\mathbf{x})$ è continua;
- $(\lambda f)(\mathbf{x}) = \lambda f(\mathbf{x})$ è continua;
- $(fg)(\mathbf{x}) = f(\mathbf{x})g(\mathbf{x})$ è continua;
- $\frac{f}{g}(\mathbf{x}) = \frac{f(\mathbf{x})}{g(\mathbf{x})}$ è continua.

N.B. Le ultime due affermazioni valgono per $g\colon X \to \mathbb{R}$!

- Siano $f\colon X \to \mathbb{R}^n$ e $g\colon Y \to \mathbb{R}^n$ due funzioni continue nei rispettivi domini e tali che $f(X) \subseteq Y$. Vale che $(g \circ f)(\mathbf{x}) = g(f(\mathbf{x}))$ è una funzione continua.

In virtù di quanto scritto precedentemente, risulta chiaro che i polinomi $p(\mathbf{x}) \in k[\mathbf{x}]$ siano funzioni continue in $\mathbb{R}^n$. Si consideri infatti un generico polinomio $p(\mathbf{x}) = p(x_1, \dots, x_n)$: esso altro non è se non una combinazione lineare di funzioni di proiezione, e di conseguenza è una funzione continua. Da questa osservazione risulta altresì che le funzioni razionali $R(\mathbf{x}) = \frac{p(\mathbf{x})}{q(\mathbf{x})}$ sono funzioni continue (nel loro naturale dominio di definizione si intende).

**Teorema 17** (caratterizzazione della continuità in termini di topologia indotta)**.** *Siano $X \subseteq \mathbb{R}^n$ e $f\colon X \to \mathbb{R}^m$. $f$ è continua in $X \iff \forall A \subseteq \mathbb{R}^m$ aperto della topologia (standard) di $\mathbb{R}^m$ $f^{-1}(A)$ è un aperto di $X$.*

*Dimostrazione.* 1. Siano $f$ una funzione continua in $X$ e $A \subseteq \mathbb{R}^m$ un insieme aperto della topologia standard di $\mathbb{R}^m$. Supponiamo che $A \subseteq f(X)$: sia dato un punto $f(\mathbf{x_0}) \in A : \mathbf{x_0} \in f^{-1}(A)$. Poiché $A$ è aperto, $\exists B(f(\mathbf{x_0}), \rho) \subseteq A$. Per la continuità di $f$, $\exists B(\mathbf{x_0}, \rho) : \forall\mathbf{x} \in B(\mathbf{x_0}, \rho) \cap X f(\mathbf{x}) \in B(f(\mathbf{x_0}), \rho)$. In particolare esiste quindi un intorno, nella topologia indotta su $X$, interamente contenuto in $f^{-1}(A)$; dato che questa proprietà vale per ogni $\mathbf{x_0} \in f^{-1}(A)$, $f^{-1}(A)$ è aperto.

2. Fissiamo un punto $\mathbf{x_0} \in X$. Consideriamo la sua immagine $f(\mathbf{x_0})$ ed un suo intorno $A(f(\mathbf{x_0}), \rho)$. Per ipotesi, $f^{-1}(A)$ è un aperto, e $\mathbf{x_0} \in f^{-1}(A)$. Poiché vale quanto appena detto, $\exists B(\mathbf{x_0}, \rho) \subseteq X$. Abbiamo di conseguenza verificato la definizione di continuità.

$\square$

**Teorema 18.** *Sia $K \subseteq \mathbb{R}^n$ un compatto, e sia $f \colon K \to \mathbb{R}^n$ una funzione continua in $K$. Si ha che $f(K)$ è un compatto in $\mathbb{R}^m$.*

*Dimostrazione.* Dimostriamo che $f(K)$ è compatto per successioni. Sia $\{y_n\}_{n\in\mathbb{N}}$ una successione qualsiasi ad elementi in $f(K)$. Esiste una successione $\{x_n\}_{n\in\mathbb{N}} : y_n = f(x_n)$. Poiché $K$ è compatto (per successioni), da ogni successione ad elementi in $K$ posso estrarre una sottosuccessione $\{x_i\}_{i\in\mathbb{N}}$ che converga ad un elemento di $K$. Consideriamo ora l'immagine della sopraccitata successione $f(\{x_i\}_{i\in\mathbb{N}}) \subseteq y_{n_{n\in\mathbb{N}}}$. Per definizione di continuità, sappiamo che $\lim_{i\to\infty} x_i = \mathbf{x_0} \implies \lim_{i\to\infty} y_i = f(\mathbf{x_0})$. Abbiamo quindi trovato una sottosuccessione convergente di una generica successione $\{y_n\}_{n\in\mathbb{N}} \subset \mathbb{R}^m$, e di conseguenza $f(K)$ è compatto. $\square$

**Teorema 19** (di Weierestrass)**.** *Sia $K \subset \mathbb{R}^n$ un compatto in $\mathbb{R}^n$, e sia $f \colon K \to \mathbb{R}^m$ una funzione continua. $f$ ammette almeno un minimo e almeno un massimo.*

*Dimostrazione.* Per il Teorema 18, si ha che $f(K)$ è un compatto; di conseguenza, per il Teorema di Heine-Borel, $f(K)$ è chiuso e limitato. Essendo limitato, esso ammette $\inf f(K)$ e $\sup f(K)$, ed essendo chiuso si ha che $\inf f(K) = \min f(K)$ e $\sup f(K) = \max f(K)$. $\square$

**Teorema 20.** *Sia $K \subseteq \mathbb{R}^n$ un insieme compatto, e sia $f \colon K \to \mathbb{R}^m$ una funzione continua ed iniettiva (quindi, abusando un po' di linguaggio, invertibile. Cfr. Appunti di Serapioni). Si ha che $f^{-1} \colon f(K) \to K$ è una funzione continua.*

*Dimostrazione.* Utilizziamo la definizione alternativa di continuità e di chiusura. Una $f \colon X \to Y$ è continua se, dato un insieme $A \subset X$ e un punto $x \in X$ di aderenza per $A$, $f(x)$ è un punto di aderenza per $f(A)$. Un insieme $X$ si dice chiuso se coincide con l'insieme dei punti di aderenza di $X$. Sappiamo che $K$ è compatto, e quindi per il Teorema di Heine-Borel è chiuso e limitato; sappiamo inoltre che la funzione $f$ è continua, e quindi $f(K)$ è un compatto (i.e. chiuso e limitato). Inoltre, poiché è iniettiva $\forall y \in f(K)\ \#f^{-1}(y) = 1$. Supponiamo ora che la funzione $f^{-1} \colon f(K) \subseteq \mathbb{R}^m \to K$ non sia continua: esiste un punto $y \in \mathbb{R}^m$ di aderenza per $f(K)$ che non è di aderenza per $K$. Poiché $f(K)$ è chiuso, $y \in f(K)$, e di conseguenza $f^{-1}(y) \in K$. Ma poiché $K$ è chiuso, ogni punto di $K$ è un punto di aderenza di $K$. Di conseguenza $f^{-1}(y)$ è un punto di aderenza per $K$: abbiamo un assurdo. $\square$

**Definizione 32.** Siano $X \subseteq \mathbb{R}^n$ e $f \colon X \to \mathbb{R}^m$. $f$ è detta *uniformemente continua* se $\forall \epsilon > 0\ \exists \delta(\epsilon) > 0 : \forall (\mathbf{x}, \mathbf{y}) : d(\mathbf{x}, \mathbf{y}) < \delta\ d(f(\mathbf{x}), f(\mathbf{y})) < \epsilon$.

La continuità uniforme è una condizione più forte della continuità (è meno forte della condizione di Lipschitz[3]): essa infatti descrive una proprietà globale della funzione (indipendentemente dai punti), e l'essere uniformemente continua implica l'essere continua.

**Teorema 21** (Heine-Cantor)**.** *Sia $X \subseteq \mathbb{R}^n$ un compatto e sia $f \colon X \to \mathbb{R}^m$ una funzione continua in $X$. Allora $f$ è uniformemente continua.*

**Teorema 22.** *Sia $X \subseteq \mathbb{R}^n$ un insieme connesso e sia $f \colon X \to \mathbb{R}^n$ una funzione continua. $f(X)$ è connesso.*

[3]Una funzione $f \colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$ è una *funzione lipschitziana* se esiste una costante $K$ tale che $d(f(\mathbf{x_1}), f(\mathbf{x_2})) \leq K d(\mathbf{x_1}, \mathbf{x_2})$.

# 4 Derivabilità e differenzialibità di funzioni scalari

Vogliamo cercare di estendere il concetto di *derivata* al caso di funzioni a più variabili. Cominciamo analizzando il comportamento delle funzioni scalari $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ con un vincolo rispetto al dominio di definizione di $f$: assumiamo infatti che $X$ sia un aperto nella topologia standard di $\mathbb{R}^n$, questo per evitare problemi dovuti ai punti di frontiera, in cui è necessario scegliere in modo oculato quali valori trascurare per evitare di uscire dal naturale dominio di definizione della funzione presa in esame.

**Definizione 33.** Sia $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione scalare definita su un insieme $X$ aperto. Fissiamo un punto $\mathbf{x} \in X$ e un versore $\mathbf{v} \in \mathbb{R}^n$ (i.e. $\|\mathbf{v}\| = 1$). Consideriamo l'insieme dei punti $\mathbf{x} + t\mathbf{v} \in X$, $t \in \mathbb{R}$ ($t$ è quindi vincolato!). Definisco la *derivata direzionale nella direzione di* $\mathbf{v}$ *di* $f$ *calcolata nel punto* $\mathbf{x}$, se esiste finito, il

$$\lim_{t\to 0} \frac{f(\mathbf{x} + t\mathbf{v}) - f(\mathbf{x})}{t}$$

che viene indicato con $D_{\mathbf{v}} f(\mathbf{x})$.

Avendo noi fissato precedentemente sia $\mathbf{x}$, sia $\mathbf{v}$, la derivata direzionale $D_{\mathbf{v}} f(\mathbf{x})$ altro non è se non una funzione di $t$: posso quindi indicare con $\phi(t) = f(\mathbf{x}+t\mathbf{v})$, e riscrivere il limite precedente come

$$\lim_{t\to 0} \frac{\phi(t) - \phi(0)}{t}$$

riducendo un problema pluridimensionale ad un problema unidimensionale, essendo $\phi$ una funzione $\phi\colon \mathbb{R} \to \mathbb{R}$. A titolo di esempio, si consideri la funzione

$$\begin{aligned} f\colon \mathbb{R}^n &\to \mathbb{R} \\ \mathbf{x} &\mapsto \|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle = \sum_{i=1}^{n} x_i^2 \end{aligned}$$

Consideriamo un generico punto $\mathbf{x} \in \mathbb{R}^n$ ed un generico versore $\mathbf{v} \in \mathbb{R}^n$. Vale che $\phi(t) = f(\mathbf{x} + t\mathbf{v}) = \langle \mathbf{x} + t\mathbf{v}, \mathbf{x} + t\mathbf{v} \rangle = \langle \mathbf{x}, \mathbf{x} \rangle + 2t \langle \mathbf{x}, \mathbf{v} \rangle + t^2 \langle \mathbf{v}, \mathbf{v} \rangle$. Si ha quindi che

$$\begin{aligned} &\lim_{t\to 0} \frac{\phi(t) - \phi(0)}{t} = \lim_{t\to 0} \frac{\langle \mathbf{x}, \mathbf{x} \rangle + 2t \langle \mathbf{x}, \mathbf{v} \rangle + t^2 \langle \mathbf{v}, \mathbf{v} \rangle - \langle \mathbf{x}, \mathbf{x} \rangle}{t} = \\ &\lim_{t\to 0} (2 \langle \mathbf{x}, \mathbf{v} \rangle + t \langle \mathbf{v}, \mathbf{v} \rangle) = 2 \langle \mathbf{x}, \mathbf{v} \rangle = 2 \sum_{i=1}^{n} x_i v_i \end{aligned}$$

Abbiamo quindi trovato un'espressione della derivata direzionale valida $\forall \mathbf{x} \wedge \forall \mathbf{v} \in \mathbb{R}^n$. È un caso fortuito: in genere infatti non si è così fortunati. Si consideri per istanza

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

prendendo $\mathbf{x} = (0, 0)$ e come versori direzionali i versori $\mathbf{e_1}, \mathbf{e_2} \in \mathbb{R}^n$ della base canonica. In questo caso si ha che

$$D_{\mathbf{e_1}} f(\mathbf{x}) = \lim_{t\to 0} \frac{f(\mathbf{x} + t\mathbf{e_1}) - f(\mathbf{x})}{t} = \lim_{t\to 0} \frac{f(t, 0)}{t} = \lim_{t\to 0} \frac{1}{t} \frac{t^2}{t^2} = \lim_{t\to 0} \frac{1}{t} = \pm\infty$$

e quindi non esiste la derivata direzionale $D_{\mathbf{e_1}} f(\mathbf{x})$, mentre vale che

$$D_{\mathbf{e_2}} f(\mathbf{x}) = \lim_{t\to 0} \frac{f(\mathbf{x} + t\mathbf{e_2}) - f(\mathbf{x})}{t} = \lim_{t\to 0} \frac{f(0, t)}{t} = \lim_{t\to 0} \frac{1}{t} \frac{0}{t^2} = 0$$

e di conseguenza esiste $D_{\mathbf{e_2}} f(\mathbf{x}) = 0$.

**Definizione 34.** Sia $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione scalare definita su un aperto $X$, e sia $\mathscr{C} := \{\mathbf{e_1}, \dots, \mathbf{e_n}\}$ la base canonica di $\mathbb{R}^n$. Chiamiamo *derivata parziale rispetto alla coordinata j-esima* e la indichiamo con

$$\frac{\partial f}{\partial x_j}(\mathbf{x}) \equiv \partial_j f(\mathbf{x}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{e_j}) - f(\mathbf{x})}{t}$$

la derivata direzionale nella direzione di $\mathbf{e_j}$ calcolata nel punto $\mathbf{x}$.

Sviluppando il rapporto incrementale precedente, si nota che

$$\frac{\partial f}{\partial x_j}(\mathbf{x}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{e_j}) - f(\mathbf{x})}{t} = \lim_{t \to 0} \frac{f(x_i, \dots, x_j + t, \dots .x_n) - f(\mathbf{x})}{t}$$

altro non è se non la derivata di $f$ rispetto alla variabile $x_j$, mentre le altre vengono lette come costanti. Si consideri infatti la funzione $f(x, y, z) = x^2\sqrt{y^2 + z^2}$. Calcolando le derivate parziali rispetto alle tre coordinate (tenendo a mente che l'origine è un punto patologico) si ottiene che

$$\begin{cases} \frac{\partial f}{\partial x} = 2x\sqrt{y^2 + z^2} \\ \frac{\partial f}{\partial y} = \frac{2yx^2}{2\sqrt{y^2+z^2}} = \frac{yx^2}{\sqrt{y^2+z^2}} \\ \frac{\partial f}{\partial z} = \frac{2zx^2}{2\sqrt{y^2+z^2}} = \frac{zx^2}{\sqrt{y^2+z^2}} \end{cases}$$

**Definizione 35.** Sia $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione scalare definita su un aperto $X$ e sia $\mathscr{C} := \{\mathbf{e_1}, \dots, \mathbf{e_n}\}$ la base canonica di $\mathbb{R}^n$. Se esistono tutte le derivate parziali $\frac{\partial f}{\partial x_i}$ calcolate in $\mathbf{x} \in X$, è possibile definire un vettore che viene chiamato *gradiente di $f$ calcolato nel punto* $\mathbf{x}$:

$$\nabla f(\mathbf{x}) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x})\mathbf{e_i} = \left(\frac{\partial f}{\partial x_1}(\mathbf{x}), \dots, \frac{\partial f}{\partial x_n}(\mathbf{x})\right)$$

La direzione del gradiente $\nabla f(\mathbf{x})$ indica la retta lungo la quale si ha il massimo valore di incremento della $f$ in un intorno del punto $\mathbf{x}$

Quanto si è precedentemente discusso non generalizza tuttavia il concetto di derivata già visto nelle funzioni $f\colon X \subseteq \mathbb{R} \to \mathbb{R}$, per via dell'assenza dell'implicazione della continuità da parte dell'esistenza delle derivate parziali/direzionali, come si può osservare dal seguente esempio. Si consideri la funzione

$$f(x, y) := \begin{cases} xe^{\frac{x}{y}} & \text{per } y \neq 0 \\ 0 & \text{per } y = 0 \end{cases}$$

Siano dati un generico versore $\mathbf{v} \in \mathbb{R}^2$ e $\mathbf{x} = (0, 0)$, il primo dei quali espresso attraverso le coordinate polari come $(\cos\theta, \sin\theta)$, con $\theta \in [0, 2\pi)$. Si ha che

$$D_{\mathbf{v}} f(\mathbf{x}) = \lim_{t \to 0} \frac{f(\mathbf{x} + t\mathbf{v}) - f(\mathbf{x})}{t} = \lim_{t \to 0} \frac{t\cos\theta e^{\cot\theta}}{t} = \cos\theta e^{\cot\theta}$$

La derivata direzionale esiste quindi $\forall\theta \neq 0, \pi$. Analizzaimo i casi potenzialmente patologici:

$$\theta = 0 \implies \mathbf{v} = (1, 0) \quad \lim_{t \to 0} \frac{f(t, 0) - f(0, 0)}{t} = \lim_{t \to 0} \frac{0}{t} = 0$$
$$\theta = \pi \implies \mathbf{v} = (-1, 0) \quad \lim_{t \to 0} \frac{f(-t, 0) - f(0, 0)}{t} = \lim_{t \to 0} \frac{0}{t} = 0$$

La derivata direzionale esiste quindi in ogni direzione; la funzione tuttavia non è continua. Si consideri infatti la restrizione della funzione al dominio $A := \{(x, y) \in \mathbb{R}^2 : y = x^3\}$, e si calcoli il

$$\lim_{(x,y) \to (0,0)} f(x, y) = \lim_{x \to 0} xe^{\frac{x}{x^3}} = \lim_{x \to 0} xe^{\frac{1}{x^2}} = \pm\infty$$

La funzione non è quindi continua in $(0, 0)$. Qual'è il motivo di questa apparente incongruenza? Non c'è trucco, non c'è inganno come direbbe qualcuno: questa discrepanza è dovuta al fatto che per funzioni $f\colon \mathbb{R} \to \mathbb{R}$ vengono a coincidere due concetti: quelli di derivabilità e *differenziabilità*.

**Definizione 36.** Sia $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione definita su un aperto $X$. $f$ si dice *differenziabile* in un punto $\mathbf{x} \in X$ se esiste un'applicazione lineare[4] $L\colon \mathbb{R}^n \to \mathbb{R}$ tale che

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = L(\mathbf{h}) + o(\mathbf{h})$$

per $\mathbf{h} \to \mathbf{0}$.

Ricordando che esiste una corrispondenza biunivoca tra applicazioni lineari e matrici, l'applicazione lineare in questione è descritta da una matrice $A \in M_{1,n}(\mathbb{R})$ (i.e. un vettore riga), e, poiché l'applicazione altro non è se non $A\mathbf{x}$, il prodotto tra matrici può essere descritto dal prodotto scalare $\langle \mathbf{a}, \mathbf{x} \rangle$.

**Definizione 37.** Sia $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione definita su un aperto $X$. $f$ si dice differenziabile in un punto $\mathbf{x}$ se esiste un vettore $\mathbf{a} \in \mathbb{R}^n$ tale che

$$f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = \langle \mathbf{a}, \mathbf{h} \rangle + o(\|\mathbf{h}\|)$$

per ogni $\mathbf{h} \in \mathbb{R}^n : \mathbf{x}+\mathbf{h} \in X$.(?)

**Teorema 23.** *Sia $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione definita su un aperto $X$ e differenziabile in un punto $\mathbf{x} \in X$. Allora*

1. *$f$ è continua in $\mathbf{x}$;*
2. *$f$ è derivabile in $\mathbf{x}$ lungo qualsiasi direzione, ed in particolare esistono le derivate parziali calcolate in $\mathbf{x}$.*

*Vale inoltre che $\mathbf{a} = \nabla f(\mathbf{x})$ e che $D_{\mathbf{v}} f(\mathbf{x}) = \langle \nabla f(\mathbf{x}), \mathbf{v} \rangle$.*

*Dimostrazione.*

1. Sappiamo che $f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = \langle \mathbf{a}, \mathbf{h} \rangle + o(\|\mathbf{h}\|)$ per $\mathbf{h} \to \mathbf{0}$. Vale quindi che $\lim_{\mathbf{h}\to 0} f(\mathbf{x}+\mathbf{h}) - f(\mathbf{x}) = 0$, e l'asserto è dimostrato.
2. Sia data la base canonica $\mathscr{C}$ di $\mathbb{R}^n$, e si consideri $\mathbf{h} = t\mathbf{e_j}$. Per $t \to 0$ vale che

$$\begin{aligned} f(\mathbf{x}+t\mathbf{e_j}) - f(\mathbf{x}) &= \langle \mathbf{a}, t\mathbf{e_j} \rangle + o(\|t\mathbf{e_j}\|) \\ f(\mathbf{x}+t\mathbf{e_j}) - f(\mathbf{x}) &= a_j t + o(|t|) \\ \frac{f(\mathbf{x}+t\mathbf{e_j}) - f(\mathbf{x})}{t} &= a_j + \frac{o(|t|)}{t} \\ \lim_{t\to 0} \frac{f(\mathbf{x}+t\mathbf{e_j}) - f(\mathbf{x})}{t} &= a_j \end{aligned}$$

$a_j$ altro non è quindi se non $\partial_j f(\mathbf{x})$. Abbiamo quindi dimostrato la seconda implicazione, nonché il fatto che $\mathbf{a} = \nabla f(\mathbf{x})$.

□

Consideriamo la funzione di proiezione $\pi_k\colon \mathbb{R}^n \to \mathbb{R}$, e calcoliamo $d\pi_k(\mathbf{x})(\mathbf{h}) := \langle \nabla f(\mathbf{x}), \mathbf{h} \rangle = \langle (0, \ldots, \frac{\partial \pi_j}{\partial x_j}(\mathbf{x}) = 1, \ldots, 0), \mathbf{h} \rangle = h_j$, che talvolta viene indicato come $dx_j$. Questa notazione viene talvolta utilizzata per scrivere

$$df(\mathbf{x}) = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(\mathbf{x}) dx_j$$

In virtù del teorema precedente vale inoltre che $D_{\mathbf{v}} f(\mathbf{x}) = \langle \nabla f(\mathbf{x}), \mathbf{v} \rangle$ è scrivibile come $\|\nabla f(\mathbf{x})\| \|\mathbf{v}\| \cos\theta$, ove con $\theta$ si indica l'angolo compreso tra il vettore gradiente e il versore $\mathbf{v}$. Vale quindi che $D_{\mathbf{v}} f(\mathbf{x}) = \|\nabla f(\mathbf{x})\| \cos\theta$. È quindi facilmente intuibile come la direzione e il modulo del vettore gradiente rappresentino rispettivamente la direzione di massima variazione della funzione (i.e.

[4] Il differenziale può essere visto come una funzione che manda da uno spazio vettoriale ad uno spazio di funzioni (in questo caso di applicazioni lineari) $df_{\mathbf{x}}\colon X \to L(\mathbb{R}^n, \mathbb{R})$ $\mathbf{x} \mapsto \langle \nabla f(\mathbf{x}), \mathbf{h} \rangle$ $\forall \mathbf{h} \in \mathbb{R}^n$ che rispetti la regola di Leibniz (cfr. Appunti di Serapioni).

la direzione in cui la derivata direzionale è massima) e il modulo della suddetta derivata: per $\theta = 0$ si ha infatti che $D_{\mathbf{v}} f(\mathbf{x}) = \|\nabla f(\mathbf{x})\|$.
Può risultare interessante individuare un'analogia fra il differenziale di una funzione $f\colon (a, b) \subset \mathbb{R} \to \mathbb{R}$ e il differenziale di una funzione scalare $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$. A questo scopo, si considerino una fuzione differenziabile $f$ e un punto $\bar{\mathbf{x}} \in X$. Posto $\mathbf{h} = \mathbf{x} - \bar{\mathbf{x}}$ (con $\mathbf{x}$ un generico punto di $X$) si ha che $f(\bar{\mathbf{x}} + \mathbf{h}) - f(\bar{\mathbf{x}}) = f(\mathbf{x}) - f(\bar{\mathbf{x}}) = \langle \nabla f(\bar{\mathbf{x}}), \mathbf{x} - \bar{\mathbf{x}} \rangle + o(\|\mathbf{x} - \bar{\mathbf{x}}\|)$. Il risultato finale è quindi, a meno di un termine trascurabile (*"bruscolini"*), un *iperpiano* tangente al grafico della funzione nel punto $\bar{\mathbf{x}}$

$$f(\mathbf{x}) = f(\bar{\mathbf{x}}) + \langle \nabla f(\bar{\mathbf{x}}), \mathbf{x} - \bar{\mathbf{x}} \rangle$$

in cui $\langle \nabla f(\bar{\mathbf{x}}), \mathbf{x} - \bar{\mathbf{x}} \rangle$ rappresenta un piano di $\mathbb{R}^n$ passante per l'origine, che viene traslato di un fattore $f(\bar{\mathbf{x}})$. Sviluppando il prodotto scalare (ponendoci in $\mathbb{R}^2$ e considerando il punto $\bar{\mathbf{x}} = (x_0, y_0)$ a titolo di esempio) si ottiene

$$\begin{aligned} f(\mathbf{x}) = f(\bar{\mathbf{x}}) + \langle \nabla f(\bar{\mathbf{x}}), \mathbf{x} - \bar{\mathbf{x}} \rangle = f(x_0, y_0) + \frac{\partial f}{\partial x}(\bar{\mathbf{x}})(x - x_0) + \frac{\partial f}{\partial y}(\bar{\mathbf{x}})(y - y_0) \\ - \frac{\partial f}{\partial x}(\bar{\mathbf{x}})(x - x_0) - \frac{\partial f}{\partial y}(\bar{\mathbf{x}})(y - y_0) + f(x, y) - f(x_0, y_0) = 0 \end{aligned}$$

ove $\mathbf{n} = \big( - \frac{\partial f}{\partial x}(\bar{\mathbf{x}}), - \frac{\partial f}{\partial y}(\bar{\mathbf{x}}), 1 \big)$ rappresenta il vettore normale al piano tangente alla curva in $\bar{\mathbf{x}}$ (cfr. Appunti di Fontanari).

**Teorema 24** (del Differenziale Totale)**.** *Sia $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione scalare definita su un aperto $X$ e sia $\mathbf{x} \in X$. Se le derivate parziali $\frac{\partial f}{\partial x_i}$ esistono tutte in un intorno $U(\mathbf{x})$ e se sono continue in $\mathbf{x}$ allora $f$ è differenziabile in $\mathbf{x}$.*

*Dimostrazione.* Poniamoci, come sempre, nel caso $n = 2$ (la pigrizia regna sovrana in questo corso). Siano $\mathbf{x} = (x_1, x_2)$ e $\mathbf{h} = (h_1, h_2)$. Vale che $f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = f(x_1 + h_1, x_2 + h_2) - f(x_1, x_2) = f(x_1 + h_1, x_2 + h_2) - f(x_1 + h_1, x_2) + f(x_1 + h_1, x_2) - f(x_1, x_2)$. Consideriamo i primi due termini a destra dell'ultimo simbolo di uguaglianza; poiché $f$ ammette per ipotesi le derivate parziali, è possibile applicare il teorema di Lagrange (o Teorema del Valor Medio):

$$f(x_1 + h_1, x_2 + h_2) - f(x_1 + h_1, x_2) = \frac{\partial f}{\partial x_2}(x_1 + h_1, x_2 + \theta h_2) h_2 \text{ con } \theta \in (0, 1)$$

Sappiamo però che la derivata parziale è continua in $\mathbf{x}$: vale quindi che

$$\frac{\partial f}{\partial x_2}(x_1 + h_1, x_2 + \theta h_2) = \frac{\partial f}{\partial x_2}(x_1, x_2) + \epsilon(h_1, h_2) \text{ per } (h_1, h_2) \to (0, 0)$$

con $\epsilon(h_1, h_2) \to 0$ per $(h_1, h_2) \to 0$ (poiché nell'espressione originale della derivata parziale avevamo entrambe le componenti di $\mathbf{h}$ entrano entrambe nel calcolo dell'errore). Inoltre, sempre in virtù dell'esistenza delle derivate parziali[5], si ha

$$f(x_1 + h_1, x_2) - f(x_1, x_2) = \frac{\partial f}{\partial x_2}(x_1, x_2) h_1 + h_1 \eta(h_1)$$

con $\eta(h_1) \to 0$ per $(h_1, h_2) \to 0$ (qui invece l'unica componente di $\mathbf{h}$ da prendere in considerazione è $h_1$). Combinando le due osservazioni, si ha, per $(h_1, h_2) \to (0, 0)$ che

$$\begin{aligned} f(\mathbf{x} + \mathbf{h}) - f(\mathbf{x}) = &\frac{\partial f}{\partial x_2}(x_1, x_2) h_2 + \frac{\partial f}{\partial x_2}(x_1, x_2) h_1 + h_1 \eta(h_1) + h_2 \epsilon(h_1, h_2) = \\ &\langle \nabla f(\mathbf{x}), \mathbf{h} \rangle + h_1 \eta(h_1) + h_2 \epsilon(h_1, h_2) \end{aligned}$$

Affinché $f$ sia differenziabile in $\mathbf{x}$ è necessario che $h_1 \eta(h_1) + h_2 \epsilon(h_1, h_2)$ sia un $o(\|\mathbf{h}\|)$. Andiamo quindi a verificare questa condizione:

$$\left| \frac{h_1 \eta(h_1) + h_2 \epsilon(h_1, h_2)}{\sqrt{h_1^2 + h_2^2}} \right| \leq \left| \frac{h_1 \eta(h_1)}{\sqrt{h_1^2 + h_2^2}} \right| + \left| \frac{h_2 \epsilon(h_1, h_2)}{\sqrt{h_1^2 + h_2^2}} \right| \leq \frac{|h_1| |\eta(h_1)|}{\sqrt{h_1^2 + h_2^2}} + \frac{|h_2| |\epsilon(h_1, h_2)|}{\sqrt{h_1^2 + h_2^2}}$$

[5]Questo è un passaggio al limite mascherato: togliendo il limite, è necessario aggiungere un errore $\eta(h_1)$ per far sì che l'uguaglianza sia verificata.

In virtù dell'ormai nota disuguaglianza $|v_i| \leq \|\mathbf{v}\|$ si ha che[6]

$$\frac{|h_1|\big|\eta(h_1)\big|}{\sqrt{h_1^2+h_2^2}} + \frac{|h_2|\big|\epsilon(h_1,h_2)\big|}{\sqrt{h_1^2+h_2^2}} \leq \big|\eta(h_1)\big| + \big|\epsilon(h_1,h_2)\big|$$

ove i maggioranti per costruzione tendono a 0 al tendere di $(h_1, h_2) \to (0,0)$. È quindi un $o(\|\mathbf{h}\|)$ e l'asserto è dimostrato. $\square$

**Definizione 38.** Siano $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione definita su un aperto $X$ e $\mathbf{x} \in X$; supponiamo inoltre che $D_{\mathbf{v}}(\mathbf{x})$ esista e sia definita in un intorno $U(\mathbf{x})$. Consideriamo il vettore $\mathbf{w} \in \mathbb{R}^n$ e il punto $\mathbf{y} \in (U(\mathbf{x}) \setminus \{\mathbf{x}\})$. Possiamo definire la derivata direzionale lungo $\mathbf{w}$ calcolata in $\mathbf{y}$ e la indicheremo $D_{\mathbf{w}}(D_{\mathbf{v}}f(\mathbf{x}))(\mathbf{y})$ (oppure $\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}\big(\frac{\partial f}{\partial x_i}(\mathbf{x})\big)(\mathbf{y})$, qualora le direzioni siano lungo i vettori della base canonica $\mathbf{e_j}$ e $\mathbf{e_i}$). Essa è pari a

$$\begin{aligned}&\lim_{t\to 0}\frac{D_{\mathbf{v}}f(\mathbf{y}+t\mathbf{w}) - D_{\mathbf{v}}f(\mathbf{y})}{t} = \\ &\lim_{t\to 0}\frac{f(\mathbf{y}+t\mathbf{w}+t\mathbf{v}) - f(\mathbf{y}+t\mathbf{w}) - f(\mathbf{y}+t\mathbf{v}) + f(\mathbf{y})}{t^2}\end{aligned}$$

Tale derivata è anche detta *derivata mista.*

Consideriamo la funzione $f\colon \mathbb{R}^n \setminus \{\mathbf{0}\} \to \mathbb{R}$ che manda l'elemento $\mathbf{x}$ in $f(\mathbf{x}) = \frac{1}{\|\mathbf{x}\|}$. Calcoliamone le derivate parziali lungo le direzioni $\mathbf{e_j}$ e $\mathbf{e_k}$:

$$\begin{aligned}\frac{\partial f}{\partial x_j} &= \frac{\partial}{\partial x_j}\left(\frac{1}{\sqrt{x_1^2+\cdots+x_n^2}}\right) = -\frac{1}{2}\frac{2x_j}{\|\mathbf{x}\|^3} = -\frac{x_j}{\|\mathbf{x}\|^3} \\ \frac{\partial^2 f}{\partial x_k \partial x_j} &= \frac{\partial}{\partial x_k}(-x_j(x_1^2+\cdots+x_n^2)^{-\frac{3}{2}}) = -\frac{\delta_{kj}}{\|\mathbf{x}\|^3} - \frac{3x_j x_k}{\|\mathbf{x}\|^5}\end{aligned}$$

Si potrebbe quindi pensare che le derivate miste siano invarianti rispetto all'ordine in cui vengono effettuate; tuttavia non è così. Si consideri la funzione $f\colon \mathbb{R}^2 \to \mathbb{R}$ così definita:

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

Calcoliamo ora, attraverso il limite del rapporto incrementale, la derivata direzionale lungo i due versori della base canonica in $(x, 0)$ e $(0, y)$:

$$\begin{aligned}\lim_{t\to 0}\frac{f(t,y)}{t} &= \lim_{t\to 0} y\frac{t^2-y^2}{t^2+y^2} = -y \\ \lim_{t\to 0}\frac{f(x,t)}{t} &= \lim_{t\to 0} x\frac{x^2-t^2}{x^2+t^2} = x\end{aligned}$$

Come si può facilmente intuire, derivando ancora si otterranno due risultati diversi.

**Teorema 25** (della derivazione a catena)**.** *Sia $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione definita su un aperto $X$ differenziabile in $\mathbf{x_0} \in X$. Sia $g_i\colon [a,b] \to \mathbb{R}$, con $i = 1,\ldots,n$, una funzione derivabile in $t_0 \in [a,b]$ tale che $g = (g_1,\ldots,g_n)$, $g[a,b] \subseteq A$ e $g(t_0) = \mathbf{x_0}$. La funzione $F := f \circ g$ è derivabile in $t_0$ e vale che*

$$F'(t_0) = \langle \nabla f(g(t_0)), g'(t_0)\rangle$$

*Dimostrazione.* Sia $h\colon X \to \mathbb{R}$ la funzione così definita:

$$h(\mathbf{y}) = \begin{cases} \frac{f(\mathbf{y}) - f(\mathbf{x_0}) - \langle \nabla f(\mathbf{x_0}), \mathbf{y}-\mathbf{x_0}\rangle}{\|\mathbf{y}-\mathbf{x_0}\|} & \mathbf{y} \neq \mathbf{x_0} \\ 0 & \mathbf{y} = \mathbf{x_0} \end{cases}$$

[6]Si è qui maggiorato $\frac{|h_1|}{\sqrt{h_1^2+h_2^2}} = \frac{|h_1|}{\|\mathbf{h}\|}$ con 1.

Poiché $h$ è una combinazione di funzioni differenziabili (e di conseguenza continue), è continua in $X$. Consideriamo $h \circ g$: essendo una composizione di funzioni continue in $\mathbf{x_0}$, è continua in $\mathbf{x_0}$. Considero la quantità

$$\lim_{\tau \to 0} h(g(t_0+\tau)) \frac{\|g(t_0+\tau) - g(t_0)\|}{\tau} =$$
$$\lim_{\tau \to 0} \frac{f(g(t_0+\tau)) - f(g(t_0)) - \langle \nabla f(g(t_0)), g(t_0+\tau) - g(t_0)\rangle}{\tau} =$$
$$\lim_{\tau \to 0} \left( \frac{F(t_0+\tau) - F(t_0)}{\tau} - \frac{\langle \nabla f(g(t_0)), g(t_0+\tau) - g(t_0)\rangle}{\tau} \right) = 0$$

Ricordando le proprietà del prodotto scalare, si ha che

$$F'(t_0) = \lim_{\tau \to 0} \frac{\langle \nabla f(g(t_0)), g(t_0+\tau) - g(t_0)\rangle}{\tau} =$$
$$\lim_{\tau \to 0} \langle \nabla f(g(t_0)), \frac{g(t_0+\tau) - g(t_0)}{\tau} \rangle = \langle \nabla f(g(t_0)), g'(t_0)\rangle$$

□

**Teorema 26** (di Schwarz)**.** *Sia $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$. Se $\frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x})$ e $\frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x})$ esistono in un intorno $U(\mathbf{x})$ e se $f \in \mathscr{C}^2(\mathbf{x})$, allora sono simmetriche.*

*Dimostrazione.* Consideriamo $\Delta := f(\mathbf{x} + t\mathbf{e_k} + t\mathbf{e_j}) - f(\mathbf{x} + t\mathbf{e_k}) - f(\mathbf{x} + t\mathbf{e_j}) + f(\mathbf{x})$, con $t$ sufficientemente piccolo affinché $\Delta$ sia definita in un intorno di $\mathbf{x}$. Definiamo la funzione $g(\tau) := f(\mathbf{x} + t\mathbf{e_k} + \tau\mathbf{e_j}) - f(\mathbf{x} + \tau\mathbf{e_j})$, con $\tau \in [0, t]$. $\Delta$ può quindi essere riscritto come $g(t) - g(0)$. Consideriamo le funzioni $\psi_{\mathbf{x_0}+t\mathbf{e_k}}(\tau) = \mathbf{x_0} + t\mathbf{e_k} + \tau\mathbf{e_j}$ e $\psi_{\mathbf{x_0}}(\tau) = \mathbf{x_0} + \tau\mathbf{e_j}$, derivabili in quanto funzioni polinomiali: risulta quindi che

$$g(\tau) = (f \circ \psi_{\mathbf{x_0}+t\mathbf{e_j}})(\tau) - (f \circ \psi_{\mathbf{x_0}})(\tau)$$

Possiamo quindi applicare il Teorema della Derivazione a Catena:

$$\frac{dg}{d\tau} = \langle \nabla f(\mathbf{x_0} + t\mathbf{e_k} + \tau\mathbf{e_j}), \mathbf{e_j}\rangle - \langle \nabla f(\mathbf{x_0} + \tau\mathbf{e_j}), \mathbf{e_j}\rangle =$$
$$\left( \frac{\partial f}{\partial x_j}(\mathbf{x} + t\mathbf{e_k} + \tau\mathbf{e_j}) - \frac{\partial f}{\partial x_j}(\mathbf{x} + \tau\mathbf{e_j}) \right)$$

Inoltre, la continuità delle derivate ci garantisce che anche $g$ sia continua (ci garantiscono la differenziabilità in un intorno $U(\mathbf{x})$ e quindi la continuità della funzione in quell'intorno). È quindi possibile applicare il Teorema di Lagrange:

$$\Delta = g(t) - g(0) = \left( \frac{\partial f}{\partial x_j}(\mathbf{x} + t\mathbf{e_k} + \theta t\mathbf{e_j}) - \frac{\partial f}{\partial x_j}(\mathbf{x} + \theta t\mathbf{e_j}) \right) t \text{ con } \theta \in (0,1)$$

Definiamo ora $\phi(\tau) := \frac{\partial f}{\partial x_j}(\mathbf{x} + \tau\mathbf{e_k} + \theta\tau\mathbf{e_j})$, con $\tau \in [0, t]$. Si ha che $\Delta = (\phi(t) - \phi(0))t$. Anche qui, dopo aver verificato l'applicabilità del Teorema della Derivazione a Catena, è possibile applicare il Teorema di Lagrange (per delucidazioni vedi sopra):

$$\Delta = (\phi(t) - \phi(0))t = \left( \frac{\partial f}{\partial x_k \partial x_j}(\mathbf{x} + \theta\tau\mathbf{e_k} + \eta\tau\mathbf{e_j}) \right) t^2 \text{ con } \eta \in (0,1)$$

Passando al limite[7] accorpando i coefficienti dei versori, si ha

$$\lim_{t \to 0} \frac{\Delta}{t^2} = \lim_{t \to 0} \frac{\partial f}{\partial x_k \partial x_j}(\mathbf{x} + \theta\mathbf{e_k} + \eta\mathbf{e_j}) = \frac{\partial f}{\partial x_k \partial x_j}(\mathbf{x})$$

Noto che avrei potuto eseguire il processo in modo analogo invertendo la scelta dei versori presi in considerazioni nelle funzioni ausiliare: avrei avuto dei coefficienti $\tilde{\theta}$ e $\tilde{\eta}$ diversi, mantenendo però analogo il risultato finale. L'asserto è dimostrato. □

[7] È qui che entra la continuità delle derivate seconde: non fossero continue in $\mathbf{x}$, il limite non convergerebbe a $\frac{\partial^2 f}{\partial x_k \partial x_j}(\mathbf{x})$.

**Teorema 27** (di Taylor)**.** *Sia $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$, ove con $X$ si indica un insieme aperto e convesso di $\mathbb{R}^n$, e siano $\mathbf{x_0}, \mathbf{x} \in X$. Vale che*

1. *Se $f \in \mathscr{C}^2(X)$ allora*

$$f(\mathbf{x}) = f(\mathbf{x_0}) + \langle \nabla f(\mathbf{x}), \mathbf{x} - \mathbf{x_0} \rangle + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\xi)(x_i - x_{0_i})(x_j - x_{0_j})$$

2. *Se $f \in \mathscr{C}^3(X)$ allora*

$$\begin{aligned} f(\mathbf{x}) =& f(\mathbf{x}) + \langle \nabla f(\mathbf{x}), \mathbf{x} - \mathbf{x_0} \rangle + \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x})(x_i - x_{0_i})(x_j - x_{0_j}) + \\ & \frac{1}{6} \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\xi)(x_i - x_{0_i})(x_j - x_{0_j})(x_k - x_{0_k}) \end{aligned}$$

*In particolare, gli ultimi termini sono rispettivamente un $R(x,h) = o(\|\mathbf{x}\|)$ e $R^2(x,h) = o(\|\mathbf{x}\|^2)$, e corrispondono al Resto secondo Lagrange (con $\xi$ appartenente al segmento congiungente $\mathbf{x_0}$ e $\mathbf{x}$).*

*Dimostrazione.* Sia $\phi\colon [0,1] \to \mathbb{R}$ una funzione definita come $\phi(t) = (f \circ \psi)(t) - (f \circ \psi)(0)$, con $\psi_{\mathbf{x_0}} = \mathbf{x_0} + t\mathbf{h}$.

$$\phi'(t) = \langle \nabla f(\psi_{\mathbf{x_0}}(t)), \psi'_{\mathbf{x_0}}(t) \rangle = \langle \nabla f(\mathbf{x_0} + t\mathbf{h}), \mathbf{h} \rangle = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathbf{x_0} + t\mathbf{h}) h_i$$

$$\phi''(t) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x_0} + t\mathbf{h}) h_i h_j$$

Applicando quindi il Teorema di Taylor per funzioni definite in $\mathbb{R}$ a valori in $\mathbb{R}$ (calcolando il polinomio di Taylor $T_2\phi(0)$), si ottiene che

$$\phi(t) = \phi(0) + \langle \nabla f(\mathbf{x_0}), \mathbf{h} \rangle\, t + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x_0} + \mathbf{h}) h_i h_j t^2$$

Calcolando il sopraccitato sviluppo in $t = 1$ si ottiene l'asserto:

$$\phi(1) = f(\mathbf{x_0} + \mathbf{h}) - f(\mathbf{x_0}) = \langle \nabla f(\mathbf{x_0}), \mathbf{h} \rangle + \frac{1}{2} \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_i}(\mathbf{x_0} + \mathbf{h}) h_i h_j$$

L'asserto è dimostrato. □

**Definizione 39.** Sia $f\colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ una funzione definita su un aperto $X$ in cui è anche differenziabile e sia $\mathbf{x} \in X$. Diremo che $f$ è *differenziabile due volte in* $\mathbf{x}$ se l'applicazione $df$ è differenziabile in $\mathbf{x}$, ottenendo l'applicazione lineare

$$d^2 f\colon X \to L(\mathbb{R}^n, \mathbb{R}^n) \text{ definita come } d^2 f(\mathbf{x}) = \sum_{i,j=1}^{n} \partial_{x_i x_j} f(\mathbf{x}) h_i h_j$$

Tale applicazione viene detta *differenziale secondo di $f$ in* $\mathbf{x}$.

La matrice che rappresenta $d^2 f(\mathbf{x})$ è la matrice quadrata

$$\begin{pmatrix} \partial_{x_1 x_1} f(\mathbf{x}) & \partial_{x_1 x_2} f(\mathbf{x}) & \dots & \partial_{x_1 x_n} f(\mathbf{x}) \\ \partial_{x_2 x_1} f(\mathbf{x}) & \partial_{x_2 x_2} f(\mathbf{x}) & \dots & \partial_{x_2 x_n} f(\mathbf{x}) \\ \vdots & \vdots & \vdots & \vdots \\ \partial_{x_n x_1} f(\mathbf{x}) & \partial_{x_n x_2} f(\mathbf{x}) & \dots & \partial_{x_n x_n} f(\mathbf{x}) \end{pmatrix}$$

che viene detta *matrice Hessiana di $f$ in* $\mathbf{x}$, indicata con $\mathbf{H}_f(\mathbf{x})$.

**Teorema 28.** *Se $f$ è differenziabile due volte in $\mathbf{x}$ vale che $\partial_{x_i x_j} f(\mathbf{x}) = \partial_{x_j x_i} f(\mathbf{x})$.*

*Dimostrazione.* Procediamo come nella dimostrazione del Teorema di Schwarz: sia $\Delta := f(\mathbf{x} + t\mathbf{e_j} + t\mathbf{e_k}) - f(\mathbf{x} + t\mathbf{e_k}) - f(\mathbf{x} + t\mathbf{e_j}) + f(\mathbf{x})$, e a questo $\Delta$, una volta definita la funzione $g(\tau) := f(\mathbf{x} + t\mathbf{e_j} + \tau\mathbf{e_k}) - f(\mathbf{x} + \tau\mathbf{e_k})$ (ottenendo così che $\Delta = g(t) - g(0)$), applichiamo il Teorema di Lagrange:

$$\Delta = \frac{dg}{d\tau} t = \left( \frac{\partial f}{\partial k}(\mathbf{x} + t\mathbf{e_j} + \theta t\mathbf{e_k}) - \frac{\partial f}{\partial k}(\mathbf{x} + \theta t\mathbf{e_k}) \right) t$$

A questo punto, non possiamo nuovamente applicare il Teorema di Lagrange in quanto $\frac{\partial f}{\partial k}$ è differenziabile solamente nel punto $\mathbf{x}$, e non è detto che lo sia in qualche suo intorno. Si può tuttavia utilizzare la definizione della differenziabilità in $\mathbf{x}$

$$\begin{aligned} \frac{\partial f}{\partial k}(\mathbf{x} + t\mathbf{e_j} + \theta t\mathbf{e_k}) = & \frac{\partial f}{\partial k}(\mathbf{x}) + \langle \nabla \frac{\partial f}{\partial k}(\mathbf{x}), (t\mathbf{e_j} + \theta t\mathbf{e_k}) \rangle + \eta(\mathbf{x}, t) = \\ & \frac{\partial f}{\partial k}(\mathbf{x}) + \frac{\partial^2 f}{\partial j \partial k}(\mathbf{x}) t + \frac{\partial^2 f}{\partial k^2}(\mathbf{x}) \theta t + \eta(\mathbf{x}, t) \end{aligned}$$

$$\begin{aligned} \frac{\partial f}{\partial k}(\mathbf{x} + \theta t\mathbf{e_k}) = & \frac{\partial f}{\partial k}(\mathbf{x}) + \langle \nabla \frac{\partial f}{\partial k}(\mathbf{x}), \theta t\mathbf{e_k} \rangle + \tilde{\eta}(\mathbf{x}, t) = \\ & \frac{\partial f}{\partial k}(\mathbf{x}) + \frac{\partial^2 f}{\partial k^2}(\mathbf{x}) \theta t + \tilde{\eta}(\mathbf{x}, t) \end{aligned}$$

Si ottiene quindi che

$$\frac{\Delta}{t^2} = \frac{\partial^2 f}{\partial j \partial k}(\mathbf{x}) + \frac{\tilde{\eta}(\mathbf{x}, t)}{t} + \frac{\eta(\mathbf{x}, t)}{t}$$

Per $t \to 0$ si ha che $\frac{\Delta}{t^2} = \frac{\partial^2 f}{\partial j \partial k}(\mathbf{x})$ e, poiché la definizione iniziale della $g(\tau)$ è simmetrica rispetto alle due direzioni, si ottiene che

$$\frac{\partial^2 f}{\partial j \partial k}(\mathbf{x}) = \frac{\partial^2 f}{\partial k \partial j}(\mathbf{x})$$

□

# 5 Ottimizzazione

**Definizione 40.** Sia $f \colon X \subseteq \mathbb{R}^n \to \mathbb{R}$ e sia $\mathbf{x_0} \in X$. Diremo che $\mathbf{x_0}$ è un *punto di massimo (minimo) locale* di $f$ se esiste un intorno $B(\mathbf{x_0}, \rho)$ tale che $\forall \mathbf{x_0} \in B(\mathbf{x_0}, \rho) \cap X$ si ha che $f(\mathbf{x}) \leq f(\mathbf{x_0})$ ($f(\mathbf{x}) \geq f(\mathbf{x_0})$). Se quanto detto vale $\forall \mathbf{x} \in X$, $\mathbf{x_0}$ viene detto di massimo (minimo) *globale.*

Se, per ogni $\mathbf{x} \neq \mathbf{x_0}$ la disuguaglianza vale in senso stretto, diremo che $\mathbf{x_0}$ è un punto di massimo (minimo) locale *forte*. Se $X$ è un insieme aperto ($X = X^\circ$), gli estremi di $f$ vengono detti *liberi.*

**Teorema 29.** *Sia $\mathbf{x_0}$ un punto estremante di $f$. Se esiste un versore $\mathbf{v} \in \mathbb{R}^n$ e $D_\mathbf{v} f(\mathbf{x_0})$ esiste, allora $D_\mathbf{v} f(\mathbf{x_0}) = 0$.*

*Dimostrazione.* Supponiamo che $\mathbf{x_0}$ sia un punto di massimo locale debole: $\forall \mathbf{x} \in B(\mathbf{x_0}, \rho) \cap X$ $f(\mathbf{x}) \leq f(\mathbf{x_0})$. Consideriamo la derivata direzionale lungo $\mathbf{v}$, in particolare per $(\mathbf{x} + t\mathbf{v}) \in B(\mathbf{x_0}, \rho)$ ($t \in [-\rho, \rho]$):

$$\lim_{t \to 0} \frac{f(\mathbf{x_0} + t\mathbf{v}) - f(\mathbf{x_0})}{t}$$

Consideriamo $t \in (-\rho, 0)$: in questo caso

$$\lim_{t \to 0} \frac{f(\mathbf{x_0} + t\mathbf{v}) - f(\mathbf{x_0})}{t} \geq 0$$

in quanto $f(\mathbf{x_0} + t\mathbf{v}) \leq f(\mathbf{x_0})$ per ipotesi e $t < 0$. Prendendo invece $t \in (0, \rho)$ vale che

$$\lim_{t \to 0} \frac{f(\mathbf{x_0} + t\mathbf{v}) - f(\mathbf{x_0})}{t} \leq 0$$

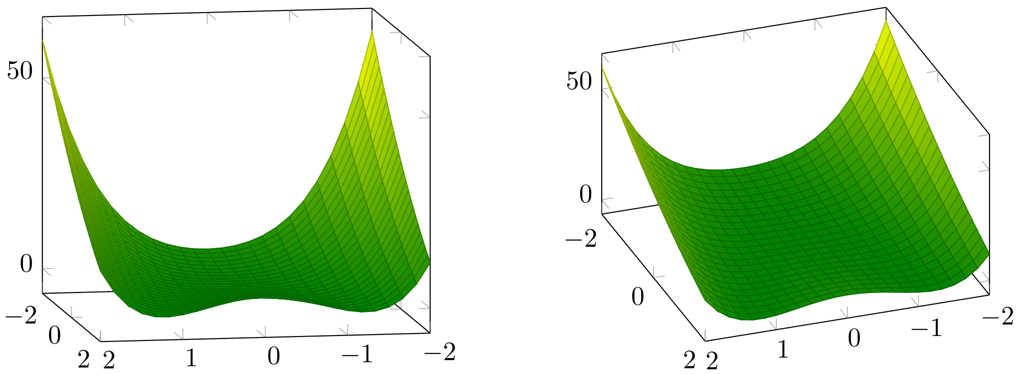
Poiché il limite esiste per ipotesti, l'unico possibile output del rapporto incrementale è 0. L'asserto è dimostrato. □

**Teorema 30** (di Fermat)**.** *Se $f$ è differenziabile in $\mathbf{x_0}$, punto estremante locale, allora tutte le derivate direzionali sono nulle, ovvero $\nabla f(\mathbf{x_0}) = \mathbf{0}$. I punti in cui il gradiente si annulla vengono detti* punti critici o stazionari.

Non è tuttavia detto che i punti critici siano di minimo o di massimo. Si consideri infatti la funzione $f(x, y) = y^2 - 3x^2y + 2x^4$. Troviamo un punto il cui il gradiente si annulla:

$$\begin{cases} \partial_x f = -6xy + 8x^3 = 0 \\ \partial_y f = 2y - 3x^2 = 0 \end{cases}$$

$(0, 0)$ è quindi un punto critico. Tuttavia, non è né un punto di massimo, né un punto di minimo. Una verifica diretta può essere fatta studiando il segno della funzione. Si consideri la stessa funzione, questa volta espressa come $f(x, y) = (y - x^2)(y - 2x^2)$. La regione compresa tra le due parabole in $(0, 0)$ non ammette minimo, bensì un massimo (in ogni intorno di $\mathbf{x_0}$ la funzione cambia segno). In virtù di questa doppia natura, il punto $(0, 0)$ viene detto *punto di sella*.




Un modo per determinare la natura di un punto critico $\mathbf{x_0}$ è quello di utilizzare la forumla di Taylor per analizzare il segno di $f(\mathbf{x_0} + \mathbf{h}) - f(\mathbf{x_0})$:

$$f(\mathbf{x_0} + \mathbf{h}) - f(\mathbf{x_0}) = \langle \nabla f(\mathbf{x_0}), \mathbf{h} \rangle + \frac{1}{2} \langle \mathbf{H}_f(\mathbf{x_0})\mathbf{h}, \mathbf{h} \rangle$$

La determinazione del segno di $f(\mathbf{x_0}+\mathbf{h})-f(\mathbf{x_0})$ conduce quindi all'analisi della forma quadratica $\langle \mathbf{H}_f(\mathbf{x_0})\mathbf{h}, \mathbf{h} \rangle$.

**Definizione 41.** Sia data una forma bilinare $b \colon V \times V \to \mathbb{R}$. Si definisce *forma*[8] *quadratica associata* a $b$ l'applicazione $q \colon V \to \mathbb{R}$ che associa ad ogni elemento $\mathbf{v} \in V$ l'elemento $q(\mathbf{v}) = b(\mathbf{v}, \mathbf{v})$. In particolare, la forma quadratica $q$ associata a $b$ soddisfa le seguenti condizioni:

- $q(k\mathbf{v}) = k^2 q(\mathbf{v})$;
- $2b(\mathbf{v}, \mathbf{w}) = q(\mathbf{v} + \mathbf{w}) - q(\mathbf{v}) - q(\mathbf{w})$.

Se è data una base $\mathscr{C} = \{\mathbf{e_1}, \ldots, \mathbf{e_n}\}$ di $V$ e se $A \in M_n(\mathbb{R})$ è la matrice che rappresenta la forma bilineare simmetrica $b$ si ha[9], per ogni $\mathbf{v} \in V$

$$q(\mathbf{v}) = \mathbf{v}^t A \mathbf{v} = \sum_{i,j=1}^{n} A_{ij} v_i v_j$$

[8] Il termine *forma* fa riferimento ad un polinomio omogeno a più variabili. Le figate che si imparano in giro!
[9] $\mathbf{v}$ va qui inteso come vettore colonna; $\mathbf{v}^t$ è invece il vettore riga.

**Teorema 31** (di Sylvester). *Sia $V$ uno spazio vettoriale reale, con $\dim(V) = n \geq 1$, e sia $b\colon V \times V \to \mathbb{R}$ una forma bilineare simmetrica. Esistono un numero intero non negativo $p \leq r$, ove $r$ è il rango di $b$, dipendente solo da $b$ e una base $\mathscr{E} = \{\mathbf{v_1}, \dots, \mathbf{v_n}\}$ di $V$ tali che rispetto ad $\mathscr{E}$ la forma $b$ abbia la seguent matrice*

$$\begin{pmatrix} D_p & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & -D_{r-p} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \end{pmatrix}$$

*Equivalentemente, ogni matrice simmetrica è congruene ad una matrice diagonale di questa forma.*
*La coppia (p, r-p) viene detta* segnatura *di $b$ e $q$.*

**Definizione 42.** Una forma quadratica $q$ sullo spazio vettoriale $V$ si dice:

1. *definita positiva* se $q(\mathbf{v}) > 0$ per ogni $\mathbf{v} \neq \mathbf{0}$;
2. *definita negativa* se $q(\mathbf{v}) < 0$ per ogni $\mathbf{v} \neq \mathbf{0}$;
3. *semidefinita positiva* se $q(\mathbf{v}) \geq 0$ per ogni $\mathbf{v}$;
4. *semidefinita negativa* se $q(\mathbf{v}) \leq 0$ per ogni $\mathbf{v}$.

**Teorema 32** (Primo Teorema di Debreu). *La matrice simmetrica $A \in M_n(\mathbb{R})$ è:*

*1.* definita positiva *se e solo se tutti i suoi autovalori sono strettamente positivi;*

*2.* definita negativa *se e solo se tutti i suoi autovalori sono strettamente negativi;*

*3.* semidefinita positiva *se e solo se tutti i suoi autovalori sono positivi e ne esiste almeno uno uguale a 0;*

*4.* semidefinita negativa *se e solo se tutti i suoi autovalori sono negativi e ne esiste almeno uno uguale a 0.*

**Teorema 33** (Secondo Teorema di Debreu). *La matrice simmetrica $A \in M_n(\mathbb{R})$ è:*

*1.* definita positiva *se e solo se tutti i minori principali di Nord-Ovest sono strettamente maggiori di 0;*

*2.* definita negativa *se e solo se tutti i suoi minori principali di Nord-Ovest di orinde $k$ hanno segno $(-1)^k$;*

*3.* semidefinita positiva *se e solo se tutti i minori principali sono maggiori o uguali 0;*

*4.* semidefinita negativa *se e solo se tutti i suoi minori principali di ordine $k$ hanno segno $(-1)^k$ o 0.*

**Teorema 34.** *Sia $f \in \mathscr{C}^2(X)$ e sia $\mathbf{x_0} \in X$ un punto critico per $f$. Se $\mathbf{H}_f(\mathbf{x_0})$ è definita positiva, allora $\mathbf{x_0}$ è un punto di minimo locale forte; se $\mathbf{H}_f(\mathbf{x_0})$ è definita negativa, allora $\mathbf{x_0}$ è un punto di massimo locale forte; se $\mathbf{H}_f(\mathbf{x_0})$ è indefinita, $\mathbf{x_0}$ è un punto di sella.*

# 6 Derivabilità e differenzialibità di funzioni a valori vettori

**Definizione 43.** Una funzione $A\colon X \to Y$ (definita su un $\mathbb{R}$-spazio vettoriale a valori in un $\mathbb{R}$-spazio vettoriale) è un'*applicazione lineare* se $A(\lambda_1\mathbf{x_i} + \lambda_2\mathbf{x_2}) = \lambda_1 A(\mathbf{x_1}) + \lambda_2 A(\mathbf{x_2})$ $\forall \mathbf{x_1}, \mathbf{x_2} \in X \wedge \forall \lambda_1, \lambda_2 \in \mathbb{R}$.

Si osservi che, data una base $\mathscr{C} := \{\mathbf{v_1}, \ldots, \mathbf{v_n}\}$ di $X$, la linearità di un'applicazione lineare così come le sue caratteristiche sono defenite in modo totale dall'azione della stessa sulla base $\mathscr{C}$: si ha infatti che

$$A(\mathbf{x}) = A\left(\sum_{i=1}^{n} x_i \mathbf{v_i}\right) = \sum_{i=1}^{n} x_i A(\mathbf{v_i})$$

Chiameremo lo spazio vettoriale delle trasformazioni lineari da $X$ a $Y$ $L(X,Y)$. È facilmente verificabile che sia uno spazio vettoriale (*"È elementare, se proprio Le risulta difficile mi scriva una mail"*). All'interno di questo spazio, è possibile individuare un sottospazio vettoriale, chiamato $\mathrm{End}(X)$, che consiste nello spazio delle funzioni lineari da $X$ a valori in $X$. All'interno di $\mathrm{End}(X)$ è possibile inoltre individuare un altro sottospazio vettoriale, quello degli endomorfismi biiettivi, indicato con $\mathrm{Aut}(X)$. Gli automorfismi di $X$ formano un gruppo, detto *gruppo di automorfismi*, rispetto alla composizione. È facile verificarlo:

- Siano $A, B \in \mathrm{Aut}(X)$. La definizione della composizione in $L(X,X)$ induce la proprietà dell'associatività su $X$;
- Sia $A \in \mathrm{Aut}(X)$; esiste la funzione $id_X : X \to X$ tale che $(id_X \circ A)(\mathbf{x}) = (A \circ id_x)(\mathbf{x}) = A(\mathbf{x})$;
- Poiché $\mathrm{Aut}(X)$ è il gruppo degli endomorfismi biiettivi, ammettono l'inversa: $\exists A^{-1} \in \mathrm{Aut}(X)$ tale che $(A^{-1} \circ A)(\mathbf{x}) = (A \circ A^{-1})(\mathbf{x}) = \mathbf{x}$.

Siano $\dim X = n$ e $\mathscr{C} := \{\mathbf{e_1}, \ldots, \mathbf{e_n}\}$ una base di $X$, $\dim Y = m$ e $\mathscr{C}' := \{\mathbf{e_1}', \ldots, \mathbf{e_m}'\}$ una base di $Y$. Ogni $A \in L(X,Y)$ individua un insieme di numeri tali che

$$A(\mathbf{e_j}) = \sum_{i=1}^{m} a_{ij} \mathbf{e_i}'$$

Per un vettore $\mathbf{x}$ generico, risulta, in virtù della linearità di $A$

$$A(\mathbf{x}) = A\left(\sum_{i=1}^{n} x_i \mathbf{e_i}\right) = \sum_{i=1}^{n} x_i A(\mathbf{e_i}) = \sum_{i=1}^{n}\sum_{j=1}^{m} x_i a_{ij} \mathbf{e_j}'$$

Risulta conveniente visualizzare questi elementi come matrice $A_{\mathscr{C}',\mathscr{C}} \in M_{m,n}(\mathbb{R})$:

$$\begin{pmatrix} a_{11} & \ldots & a_{1j} & \ldots & a_{1n} \\ a_{21} & \ldots & a_{2j} & \ldots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{m1} & \ldots & a_{mj} & \ldots & a_{mn} \end{pmatrix}$$

La funzione $A$ può quindi essere indicata come l'applicazione lineare che manda l'elemento $\mathbf{x} \in X$ nell'elemento $A\mathbf{x} \in Y$. Ciò risulta evidente se si sviluppa il prodotto riga per colonna fra $A$ e $\mathbf{x}$. Risulta quindi evidente come ci sia una corrispondenza biunivoca fra applicazioni lineari $L(X,Y)$ e matrici $M_{m,n}(\mathbb{R})$: in particolare, questa corrispondenza è un isomorfismo[10]. È inoltre possibile individuare un altro isomorfismo fra $M_{m,n}(\mathbb{R})$ e $\mathbb{R}^{m\times n}$: $\mathscr{C}\colon A \mapsto (A_1^t, \ldots, A_j^t, \ldots, A_n^j)$, ove con $A_j^t$ viene indicato il trasposto del $j$-esimo vettore colonna. Avendo individuato questa relazione, è possibile definire in modo naturale (?) la norma di una mappa lineare.

**Definizione 44.** Sia $A \in L(\mathbb{R}^n, \mathbb{R}^m)$ un'applicazione lineare, e sia $A_{\mathscr{C}',\mathscr{C}} \in M_{m,n}(\mathbb{R})$ la matrice ad essa associata nelle basi $\mathscr{C}$ di $X$ e $\mathscr{C}'$ di $Y$. Definiamo *norma* di $A$ (e la indichiamo con $\|A\|_{Eucl}$) la norma in $\mathbb{R}^{m\times n}$ del vettore ottenuto trasponendo i vettori colonna della matrice $A$

$$\|A\|_{Eucl} = \sqrt{\sum_{j=1}^{n} \|A_j^t\|^2} = \sqrt{\sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}^2}$$

[10] Un isomorfismo è un morfismo (i.e. ...) biiettivo. Come dice il buon tutor, preserva la struttura degli spazi.

**Definizione 45.** Sia $A \in L(X, Y)$ un'applicazione lineare. Definiamo *operatore aggiunto* di $A$ l'applicazione $A^t$ tale che $\langle A\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, A^t\mathbf{y}\rangle\ \forall \mathbf{x} \in \mathbb{R}^n \wedge \forall \mathbf{y} \in \mathbb{R}^m$.

Si considerino $\mathbf{x} = \mathbf{e_j}$ e $\mathbf{y} = \mathbf{e_i}'$, con $\mathbf{e_j} \in \mathscr{C}$ e $\mathbf{e_i}' \in \mathscr{C}'$, nella definizione precedente. Si ha che

$$\langle A\mathbf{e_j}, \mathbf{e_i}\rangle = \langle \sum_{k=1}^{m} a_{kj}\mathbf{e_k}', \mathbf{e_i}'\rangle = \sum_{k=1}^{n} a_{kj}\, \langle \mathbf{e_k}', \mathbf{e_i}'\rangle = a_{kj}\delta_{ki} = a_{ij}$$
$$\langle \mathbf{e_j}, A^t\mathbf{e_i}'\rangle = \langle \mathbf{e_j}, \sum_{k=1}^{n} a^t_{ki}\mathbf{e_k}\rangle = \sum_{k=1}^{m} a^t_{ki}\, \langle \mathbf{e_j}, \mathbf{e_k}\rangle = a^t_{ki}\delta_{kj} = a^t_{jk}$$

Risulta quindi evidente come la matrice associata all'operatore aggiunto di $A$ altro non sia se non la trasposta $A^t$ della matrice associata ad $A$.

**Teorema 35.** *$\forall A \in L(\mathbb{R}^n, \mathbb{R}^m) \wedge \forall \mathbf{h} \in \mathbb{R}^n$ vale che $\|A\mathbf{h}\|_{\mathbb{R}^m} \leq \|A\|_{Eucl}\|\mathbf{h}\|_{\mathbb{R}^n}$.*

*Dimostrazione.* Consideriamo $\|A\mathbf{h}\|^2$:

$$\|A\mathbf{h}\|^2 = \sum_{i=1}^{n}\sum_{j=1}^{m} a$$

□

**Teorema 36.** *L'insieme $Aut(\mathbb{R}^n)$ è un aperto di $End(\mathbb{R}^n)$. Equivalentemente, $GL_n(\mathbb{R})$ è un aperto di $M_n(\mathbb{R})$.*

*Dimostrazione.* Sappiamo che $A \in GL_n(\mathbb{R}) \iff \det A \neq 0$. Consideriamo la funzione $\det\colon \mathbb{R}^{n^2} \to \mathbb{R}$: è una funzione continua. Consideriamo l'insieme $(0, +\infty) \subset \mathbb{R}$, e in particolare consideriamo la sua controimmagine $\det^{-1}(0, +\infty)$, che coincide con $GL_n(\mathbb{R})$. Poiché det è una funzione continua e $(0, +\infty)$ è un aperto in $(\mathbb{R}, \tau_\epsilon)$, $GL_n(\mathbb{R})$ è un aperto. □

**Definizione 46.** Sia $A \in \mathrm{End}(\mathbb{R}^n)$. $A$ è un *operatore autoaggiunto* se $\langle A\mathbf{x}, \mathbf{y}\rangle = \langle \mathbf{x}, A\mathbf{y}\rangle$ $\forall \mathbf{x}, \mathbf{y} \in \mathbb{R}^n$. Per quanto detto precedentemente in merito agli operatori aggiunti, risulta evidente che $A = A^t$.
Per converso, un operatore si dice *anti-autoaggiunto* se la matrice ad esso associato è antisimmetrica.

Denotiamo con $\mathrm{End}^+(\mathbb{R}^n)$ l'insieme degli operatori autoaggiunti su $\mathbb{R}^n$, e con $\mathrm{End}^-(\mathbb{R}^n)$ l'insieme degli operatori anti-autoaggiunti su $\mathbb{R}^n$.

**Teorema 37.** *$End(\mathbb{R}^n) = End^+(\mathbb{R}^n) \oplus End^-(\mathbb{R}^n)$.*

*Dimostrazione.* Sia $A \in \mathrm{End}(\mathbb{R}^n)$ un generico endomorfismo di $\mathbb{R}^n$; scriviamola come $\frac{1}{2}(A + A^t) + \frac{1}{2}(A - A^t)$. Vogliamo ora provare che $\frac{1}{2}(A + A^t) \in \mathrm{End}^+(\mathbb{R}^n)$ e che $\frac{1}{2}(A + A^t) \in \mathrm{End}^-(\mathbb{R}^n)$. Supponiamo che □

**Definizione 47.** Sia $\mathbf{f}\colon X \subseteq \mathbb{R}^n \to \mathbb{R}^m$, e sia $\mathbf{x_n} \in X$. Diremo che $\mathbf{f}$ è *differenziabile* in $\mathbf{x_n}$ quando esiste un'applicazione $d\mathbf{f}\colon X \to L(\mathbb{R}^n, \mathbb{R}^m)$ tale che

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\|\mathbf{f}(\mathbf{x_0} + \mathbf{h}) - \mathbf{f}(\mathbf{x_0}) - d\mathbf{f}(\mathbf{x_0})\mathbf{h}\|}{\|\mathbf{h}\|} = 0$$

ove con $d\mathbf{f}(\mathbf{x_0})$ si indica un'applicazione lineare $d\mathbf{f}(\mathbf{x_0})\colon \mathbb{R}^n \to \mathbb{R}^m$.
La relazione può essere riscritta nella forma

$$\mathbf{f}(\mathbf{x_0} + \mathbf{h}) - \mathbf{f}(\mathbf{x_0}) = d\mathbf{f}(\mathbf{x_0})\mathbf{h} + \epsilon_{\mathbf{x_0}}(\mathbf{h})$$

ove con $\epsilon_{\mathbf{x_0}}(\mathbf{h})$ si indica una quantità tale che $\lim_{\mathbf{h}\to\mathbf{0}} \frac{\epsilon_{\mathbf{x_0}}(\mathbf{h})}{\mathbf{h}} = 0$.

Si considerino i seguenti esempi:

1. Sia $A \in L(\mathbb{R}^n, \mathbb{R}^m)$. Applicando la definizione, si ottiene che $dA(\mathbf{x}) = A$, ovvero il differenziale di $X$ è identicamente $A$ per ogni $\mathbf{x} \in \mathbb{R}^n$.

2. Sia $id_{\mathbb{R}^n} \in \mathrm{End}(\mathbb{R}^n)$. Poiché gli endomorfismi di $\mathbb{R}^n$ sono applicazioni lineari da $\mathbb{R}^n$ a valori in $\mathbb{R}^n$, possiamo applicare il risultato precedente: si ha quindi che $did_{\mathbb{R}^n} = id_{\mathbb{R}^n}$.

3. Si consideri $\langle , \rangle : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$. Si ha che

$$\langle \mathbf{x_1} + \mathbf{h_1}, \mathbf{x_2} + \mathbf{h_2} \rangle - \langle \mathbf{x_1}, \mathbf{x_2} \rangle = \langle \mathbf{x_1}, \mathbf{h_2} \rangle + \langle \mathbf{x_1}, \mathbf{h_1} \rangle + \langle \mathbf{h_1}, \mathbf{h_2} \rangle$$

I primi due termini altro non sono se non l'applicazione lineare $dp(\mathbf{x_1}, \mathbf{x_2})$ valutata in $(\mathbf{h_1}, \mathbf{h_2})$. Vogliamo invece dimostrare che la quantità $\langle \mathbf{h_1}, \mathbf{h_2} \rangle$ è un $o(\|(\mathbf{h_1}, \mathbf{h_2})\|)$ per $\mathbf{h} \to \mathbf{0}$. Consideriamo la quantità

$$\frac{|\langle \mathbf{h_1}, \mathbf{h_2} \rangle|}{\|(\mathbf{h_1}, \mathbf{h_2})\|} \leq \frac{\|\mathbf{h_1}\| \|\mathbf{h_2}\|}{\|\mathbf{h_2}\|} \leq \|\mathbf{h_1}\|$$

ove si è utilizzata la Disuguaglianza di Cauchy-Schwarz e si è operata la seguente maggiorazione

$$\|(\mathbf{h_1}, \mathbf{h_2})\| \geq \|\mathbf{h_2}\|$$

La quantità tende a 0 per $\mathbf{h1} \to \mathbf{0}$, e quindi è tutto a posto.

**Definizione 48.** Sia $A \subseteq \mathbb{R}^n$ un aperto di $\mathbb{R}^n$, sia $\mathbf{x_0} \in A$ e sia $\mathbf{f} \colon A \to \mathbb{R}^m$ una funzione differenziabile in $\mathbf{x_0}$ tale che $d\mathbf{f}(\mathbf{x_0}) \in \mathrm{Aut}(\mathbb{R}^n)$. Diremo che $\mathbf{x_0}$ è uno *zero isolato* per la funzione $\mathbf{f} - \mathbf{f}(\mathbf{x_0})$ se esiste un intorno $U(\mathbf{x_0})$ tale che $\forall \mathbf{x} \in U(\mathbf{x_0})\ \mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x_0}) \implies \mathbf{x} = \mathbf{x_0}$.

Dimostriamo ora l'unicità dello zero isolato. Poiché $d\mathbf{f}(\mathbf{x_0}) \in \mathrm{Aut}(\mathbb{R}^n)$, sappiamo che esiste $d\mathbf{f}(\mathbf{x_0})^{-1}$, e che suddetta funzione è continua (in quanto lineare). Per questi motivi sappiamo che

$$\lim_{\mathbf{h} \to \mathbf{0}} \frac{\|d\mathbf{f}(\mathbf{x_0})^{-1}(\epsilon_{\mathbf{x_0}}(\mathbf{h}))\|}{\|\mathbf{h}\|} = \|d\mathbf{f}(\mathbf{x_0})^{-1}\left(\lim_{\mathbf{h} \to \mathbf{0}} \frac{\epsilon_{\mathbf{x_0}}(\mathbf{h})}{\|\mathbf{h}\|}\right)\| = \|d\mathbf{f}(\mathbf{x_0})^{-1}(\mathbf{0})\| = 0$$

Possiamo quindi affermare che, dato un $\delta \in \mathbb{R}$ e preso $\mathbf{h} \in \mathbb{R}^n$ tale che $0 < \|\mathbf{h}\| < \delta$, si ha

$$\|d\mathbf{f}(\mathbf{x_0})^{-1}(\epsilon_{\mathbf{x_0}}(\mathbf{h}))\| < \|\mathbf{h}\|$$

Supponiamo ora che esista un certo $\mathbf{x_0} + \mathbf{h} \in U(\mathbf{x_0})$, con $\|h\| < \delta$, tale che $\mathbf{f}(\mathbf{x_0} + \mathbf{h}) = \mathbf{f}(\mathbf{x_0})$. Consideriamo la differenza

$$\mathbf{f}(\mathbf{x_0} + \mathbf{h}) - \mathbf{f}(\mathbf{x_0}) = d\mathbf{f}(\mathbf{x_0})\mathbf{h} + \epsilon_{\mathbf{x_0}}(\mathbf{h}) = \mathbf{0}$$

**Teorema 38** (Lemma di Hadamard)**.** *Siano $A \subseteq \mathbb{R}^n$ un aperto di $\mathbb{R}^n$, $\mathbf{x_0} \in A$ e $\mathbf{f} \colon A \to \mathbb{R}^m$. Le seguenti asserzioni sono equivalenti:*

1. *$\mathbf{f}$ è differenziabile in $\mathbf{x_0}$;*

2. *esiste una mappa a valori operatori lineari $\phi_{\mathbf{x_0}} \colon A \to L(\mathbb{R}^n, \mathbb{R}^m)$ continua in $\mathbf{x_0}$ tale che $\mathbf{f}(\mathbf{x}) = \mathbf{f}(\mathbf{x_0}) + \phi_{\mathbf{x_0}}(\mathbf{x_0})(\mathbf{x} - \mathbf{x_0})$.*

*In particolare, se una delle due è valida allora $\phi_{\mathbf{x_0}}(\mathbf{x_0}) = d\mathbf{f}(\mathbf{x_0})$.*

*Dimostrazione.* ...Through the darkness, and into the void □

Si consideri una funzione $\mathbf{f} \colon A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, e siano $\mathscr{C}_1 = \{\mathbf{e_1}, \dots, \mathbf{e_n}\}$ e $\mathscr{C}_2 = \{\mathbf{u_1}, \dots, \mathbf{u_1}\}$ le basi canoniche rispettivamente di $\mathbb{R}^n$ e di $\mathbb{R}^m$. Le componenti di $\mathbf{f}$ sono le funzioni scalari $f_i, \dots, f_m$ definite in modo che

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^{m} f_i(\mathbf{x})\mathbf{u_i}$$

**Definizione 49.** Siano $A \subseteq \mathbb{R}^n$ un aperto di $\mathbb{R}^n$, $\mathbf{x_0} \in A$ e $\mathbf{f} \colon A \to \mathbb{R}^m$. Diremo che $\mathbf{f}$ è *parzialmente derivabile* in $\mathbf{x_0}$ nella direzione $\mathbf{h} \in \mathbb{R}^n$ se la funzione $g \colon \mathbb{R} \to \mathbb{R}^m$ che manda $t \mapsto \mathbf{f}(\mathbf{x_0} + t\mathbf{h})$ è differenziabile in $t = 0$.

**Teorema 39.** *Siano $A \subseteq \mathbb{R}^n$ un aperto di $\mathbb{R}^n$, $\mathbf{x_0} \in A$ e $\mathbf{f}\colon A \to \mathbb{R}^m$. Se $\mathbf{f}$ è differenziabile in $\mathbf{x_0}$, allora le derivate parziali $D_j f_i(\mathbf{x_0})$ esistono, e, per $\mathbf{h} = \sum_{i=1}^n h_i \mathbf{e_i}$ si ha che*

$$d\mathbf{f}(\mathbf{x_0})\mathbf{h} = \sum_{i=1}^m \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(\mathbf{x_0}) h_j \mathbf{u_i}$$

*Dimostrazione.* Poiché $\mathbf{f}$ è differenziabile in $\mathbf{x_0}$, sappiamo che $\mathbf{f}(\mathbf{x_0}+\mathbf{h}) - \mathbf{f}(\mathbf{x_0}) = d\mathbf{f}(\mathbf{x_0})\mathbf{h} + \epsilon_{\mathbf{x_0}}(\mathbf{h})$. Data $\mathscr{C}_1 = \{\mathbf{e_1}, \dots, \mathbf{e_n}\}$ la base canonica di $\mathbb{R}^n$ Consideriamo il vettore $\mathbf{h} = t\mathbf{e_j}$: si ha che

$$\mathbf{f}(\mathbf{x_0} + t\mathbf{e_j}) - \mathbf{f}(\mathbf{x_0}) = d\mathbf{f}(\mathbf{x_0}) t\mathbf{e_j} + \epsilon_{\mathbf{x_0}}(t\mathbf{e_j})$$

Per la linearità del differenziale è possibile estrarre $t$ e, facendo il passaggio al limite, si ha che

$$\lim_{t\to 0} \frac{\mathbf{f}(\mathbf{x_0} + t\mathbf{e_j}) - \mathbf{f}(\mathbf{x_0})}{t} = d\mathbf{f}(\mathbf{x_0})\mathbf{e_j}$$

Poiché la funzione $\mathbf{f}$ può essere vista come un vettore le cui componenti sono le funzioni scalari $f_i$, si ha che

$$\lim_{t\to 0} \frac{\sum_{i=1}^m f_i(\mathbf{x_0} + t\mathbf{e_j})\mathbf{u_i} - \sum_{i=1}^m f_i(\mathbf{x_0})\mathbf{u_i}}{t}$$

che diventa, per le proprietà dei limiti

$$\lim_{t\to 0} \sum_{i=1}^m \frac{f_i(\mathbf{x_0} + t\mathbf{e_j}) - f_i(\mathbf{x_0})}{t} \mathbf{u_i} = \lim_{t\to 0} \sum_{i=1}^m \frac{\partial f_i}{\partial x_j}(\mathbf{x_0})\mathbf{u_i} = d\mathbf{f}(\mathbf{x_0})\mathbf{e_j}$$

Reiterando il procedimento per ogni $\mathbf{e_j}$ appartenente alla base canonica $\mathscr{C}$ si ottiene l'asserto. □

Soffermiamoci per qualche istante sulla dimostrazione precedente, ed in particolare consideriamo il vettore $d\mathbf{f}(\mathbf{x_0})\mathbf{e_j}$ nell'ottica della formulazione matriciale degli operatori lineari: esso altro non è se non la $j$-esima colonna della matrice associata all'operatore lineare $d\mathbf{f}(\mathbf{x_0})$ nelle basi canoniche di $\mathbb{R}^n$ e $\mathbb{R}^m$. Si ha quindi che

$$d\mathbf{f}(\mathbf{x_0}) = \begin{pmatrix} D_1 f_1(\mathbf{x_0}) & \dots & D_n f_1(\mathbf{x_0}) \\ D_1 f_2(\mathbf{x_0}) & \dots & D_n f_2(\mathbf{x_0}) \\ \vdots & \vdots & \vdots \\ D_1 f_m(\mathbf{x_0}) & \dots & D_n f_m(\mathbf{x_0}) \end{pmatrix} = \begin{pmatrix} \nabla f_1(\mathbf{x_0}) \\ \nabla f_2(\mathbf{x_0}) \\ \vdots \\ \nabla f_m(\mathbf{x_0}) \end{pmatrix}$$

Tale matrice viene chiamata *matrice jacobiana.*

**Teorema 40** (Regola della Derivazione a Catena)**.** *Siano $A \subseteq \mathbb{R}^n$ un aperto di $\mathbb{R}^n$, $B \subseteq \mathbb{R}^m$ un aperto di $\mathbb{R}^m$, $\mathbf{x_0} \in A$, $\mathbf{f}\colon A \to \mathbb{R}^m$, $\mathbf{f}(\mathbf{x_0}) \in B$, $\mathbf{g}\colon B \to \mathbb{R}^k$, con $\mathbf{f}$ e $\mathbf{g}$ due funzioni differenziabili rispettivamente in $\mathbf{x_0}$ e $\mathbf{f}(\mathbf{x_0})$. La funzione $\mathbf{g} \circ \mathbf{f}\colon A \to \mathbb{R}^k$ è differenziabile in $\mathbf{x_0}$ e si ha che $d(\mathbf{g} \circ \mathbf{f})(\mathbf{x_0}) = d\mathbf{g}(\mathbf{f}(\mathbf{x_0}))d\mathbf{f}(\mathbf{x_0})$.*

*Dimostrazione.* Sia $\mathbf{y_0} = \mathbf{f}(\mathbf{x_0})$, e siano definite le seguenti funzioni

$$\mathbf{u}(\mathbf{h}) = \mathbf{f}(\mathbf{x_0} + \mathbf{h}) - \mathbf{f}(\mathbf{x_0}) - d\mathbf{f}(\mathbf{x_0})\mathbf{h}$$
$$\mathbf{v}(\mathbf{k}) = \mathbf{g}(\mathbf{y_0} + \mathbf{k}) - \mathbf{g}(\mathbf{y_0}) - d\mathbf{g}(\mathbf{y_0})\mathbf{k}$$

Essendo le funzioni $\mathbf{f}$ e $\mathbf{g}$ differenziabili rispettivamente in $\mathbf{x_0}$ e $\mathbf{y_0}$, le funzioni sopra definite altro non sono se non $\epsilon_{\mathbf{x_0}}(\mathbf{h})\mathbf{h}$ e $\eta_{\mathbf{y_0}}(\mathbf{k})\mathbf{k}$. Fissiamo $\mathbf{h}$ e definiamo $\mathbf{k} = \mathbf{f}(\mathbf{x_0}+\mathbf{h}) - \mathbf{f}(\mathbf{x_0}) = d\mathbf{f}(\mathbf{x_0})\mathbf{h} + \epsilon_{\mathbf{x_0}}(\mathbf{h})\mathbf{h}$. Si ha quindi che

$$\mathbf{g}(\mathbf{f}(\mathbf{x_0}+\mathbf{h})) - \mathbf{g}(\mathbf{f}(\mathbf{x_0})) = d\mathbf{g}(\mathbf{f}(\mathbf{x_0}))d\mathbf{f}(\mathbf{x_0})\mathbf{h} + d\mathbf{g}(\mathbf{f}(\mathbf{x_0}))\epsilon_{\mathbf{x_0}}(\mathbf{h})\mathbf{h} + \eta_{\mathbf{y_0}}(\mathbf{k})\mathbf{k}$$

Dobbiamo ora dimostrare che la quantità $d\mathbf{g}(\mathbf{f}(\mathbf{x_0}))\epsilon_{\mathbf{x_0}}(\mathbf{h})\mathbf{h} + \eta_{\mathbf{y_0}}(\mathbf{k})\mathbf{k}$ è un $o(\|\mathbf{h}\|)$. Operando delle maggiorazioni si riesce ad ottenere che

$$\|\mathbf{k}\| = \|d\mathbf{f}(\mathbf{x_0})\mathbf{h} + \epsilon_{\mathbf{x_0}}(\mathbf{h})\mathbf{h}\| \le \|d\mathbf{f}(\mathbf{x_0}) + \epsilon_{\mathbf{x_0}}(\mathbf{h})\|\|\mathbf{h}\|$$

Sostituendo nella precedente e facendo degli abili passaggi algebrici (tra cui il passaggio al limite) si ottiene che

$$\begin{aligned}
&\lim_{\|\mathbf{h}\|\to 0} \frac{\mathbf{g}(\mathbf{f}(\mathbf{x_0}+\mathbf{h})) - \mathbf{g}(\mathbf{f}(\mathbf{x_0})) - d\mathbf{g}(\mathbf{f}(\mathbf{x_0}))d\mathbf{f}(\mathbf{x_0})\mathbf{h}}{\|\mathbf{h}\|} \\
&\quad \leq \lim_{\|\mathbf{h}\|\to 0} d\mathbf{g}(\mathbf{f}(\mathbf{x_0}))\epsilon_{\mathbf{x_0}}(\mathbf{h}) + \eta_{\mathbf{y_0}}(\mathbf{k})\|d\mathbf{f}(\mathbf{x_0}) + \epsilon_{\mathbf{x_0}}(\mathbf{h})\|
\end{aligned}$$

Poiché sia $\epsilon_{\mathbf{x_0}}(\mathbf{h})$, sia $\eta_{\mathbf{y_0}}(\mathbf{k})$ tendono a 0 per $\|\mathbf{h}\| \to 0$, l'asserto è dimostrato. □

Da notare che la relazione di cui sopra può essere espressa attraverso le matrici di Jacobi associate alle applicazioni lineari sopra indicate: $J_{\mathbf{g}\circ\mathbf{f}}(\mathbf{x_0}) = J_g(\mathbf{f}(\mathbf{x_0}))J_f(\mathbf{x_0})$, ove fra le due matrici si è applicato il prodotto matriciale.

**Definizione 50.** Siano $A, B \subseteq \mathbb{R}^n$ due aperti di $\mathbb{R}^n$, sia $\mathbf{f}\colon A \to B$ una funzione biiettiva e differenziabile in $A$ e sia $\mathbf{f}^{-1}\colon B \to A$ una funzione differenziabile in $B$[11]. Allora $\mathbf{f}$ è un *diffeomorfismo.*

Una proprietà interessante dei diffeomorfismi che si può facilmente ricavare è la seguente. Si consideri $d(\mathbf{f}\circ\mathbf{f}^{-1})(\mathbf{x}) = d(id_{\mathbb{R}^n})(\mathbf{x}) = id_{\mathbb{R}^n}$. Applicando la Regola della Derivazione a Catena si ottiene che $d\mathbf{f}^{-1}(\mathbf{f}(\mathbf{x}))d\mathbf{f}(\mathbf{x}) = id_{\mathbb{R}^n}$: quindi non solo $\mathbf{f}$, ma anche $d\mathbf{f}$ è invertibile! That's some cool shit, ain't it?
Vogliamo ora verificare la veridicità del Teorema del Valor Medio (o Teorema di Lagrange) per funzioni a valori vettori. Si consideri ad esempio la funzione

$$\begin{aligned}
\mathbf{f}\colon [0,2\pi] &\to \mathbb{R}^2 \\
t &\mapsto (\cos(t), \sin(t))
\end{aligned}$$

E si consideri l'intervallo $[0, 2\pi]$. Vogliamo trovare un valore $\mathbf{f}'(\theta)$ tale che $\mathbf{f}(2\pi)-\mathbf{f}(0) = \mathbf{f}'(\theta)2\pi$. Ci si può facilmente accorgere che non esiste un unico valore che mi permetta di avere $\mathbf{f}'(\theta) = \mathbf{0}$. Dobbiamo quindi trovare un'altra generalizzazione.

**Definizione 51.** Siano $A \subseteq \mathbb{R}^n$ un aperto di $\mathbb{R}^n$, $\mathbf{f}\colon A \to \mathbb{R}^m$ una funzione differenziabile. Supponiamo che $d\mathbf{f}\colon A \to L(\mathbb{R}^n, \mathbb{R}^m)$ sia limitato: esiste un $M > 0$ tale che $\|d\mathbf{f}(\mathbf{x})\mathbf{h}\| \leq M\|\mathbf{h}\|$ per ogni $\mathbf{x} \in \mathbb{R}^n$ e per ogni $\mathbf{h} \in \mathbb{R}^n$. Allora vale che

$$\|\mathbf{f}(\mathbf{x}+\mathbf{h}) - \mathbf{f}(\mathbf{x})\| \leq M\|\mathbf{h}\|$$

La funzione $\mathbf{f}$ viene detta *lipschitziana*, con costante di Lipschitz $M$.

**Teorema 41.** *Sia $\mathbf{f}\colon \mathbb{R}^n \to \mathbb{R}^m$ una funzione di classe $\mathscr{C}^1(\mathbb{R}^n)$ e sia $K \subset \mathbb{R}^n$ un compatto in $\mathbb{R}^n$. Allora $\mathbf{f}|_K$ è lipschitziana.*

# 7 Spazi di funzioni

**Definizione 52.** Sia dato un insieme $X$. Chiameremo *distanza* su $X$ una funzione $d\colon X\times X \to \mathbb{R}$ che soddisfa tre condizioni:

1. $d(x,y) \geq 0\ \forall x,y \in X \wedge d(x,y) = 0 \iff x = y$;
2. $d(x,y) = d(y,x)$;
3. $d(x,y) \leq d(x,z) + d(z,y)$.

Chiameremo la coppia $(X, d)$ *spazio metrico.*

Consideriamo i seguenti esempi di distanza:

[11] È interessante notare che $A$ e $B$ sono necessariamente sottoinsiemi di insiemi aventi la medesima dimensione, in quanto la funzione $\mathbf{f}$ è un omeomorfismo.

- Dato un generico insieme $X$,la distanza discreta

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

Da notare come questa metrica induca una topologia particolare su $X$ (che Ulde, Davide ed io abbiamo denominato $\tau_{duv}$, anche se è nota come topologia discreta), in cui la base della topologia $\mathcal{B}$ è costituita dai singoletti $\{x\}, x \in X$. e in cui gli aperti sono $\bigcup_i \{x_i\}$. Questa è l'unica topologia che ci è venuta in mente in cui la distanza discreta rispetta la richiesta di continuità della distanza: nella topologia euclidea $\tau_\epsilon$ la suddetta distanza non è infatti continua.

- Dato lo spazio delle funzioni continue $\mathscr{C}^0(I)$ su un intervallo $I = [a,b]$, definiamo l'infinito-distanza

$$d_\infty(f,g) = \max_{x\in I} \left| f(x) - g(x) \right|$$

È possibile generalizzare questa definizione a $\mathscr{C}^k(I)$, ponendo l'infinito-distanza

$$d_\infty(f,g) = \max_{x\in I} \left| f(x) - g(x) \right| + \cdots + \max_{x\in I} \left| f^{(k)}(x) - g^{(k)}(x) \right|$$

- Sia $\mathscr{R}(I)$ lo spazio delle funzioni Riemann-integrabili su $I$. Definiamo

$$d(f,g) = \int_a^b \left| f(x) - g(x) \right| \, dx$$

Dato uno spazio metrico $(X,d)$, per mezzo della distanza $d$ possiamo anche indurre una topologia su $X$ definendo gli intorni sferici (cfr. Definizione 7).

**Definizione 53.** Sia dato uno spazio metrico $(X,d)$, Una successione $\{x_n\}_{n\in\mathbb{N}}$ di elementi di $X$ è *convergente* in $x \in X$ se è verificata una delle condizioni equivalenti:

1. Sia dato un intorno $U(x)$; $\{x_n\}_{n\in\mathbb{N}}$ è definitivamente in $U(x)$;
2. $\forall \epsilon > 0 \ \exists \bar{n} = \bar{n}(\epsilon) \in \mathbb{N} : d(x_n, x) < \epsilon \ \forall n > \bar{n}$;
3. $\lim_{n\to\infty} x_n = x$.

**Definizione 54.** Sia data una successione $\{x_n\}_{n\in\mathbb{N}} \subset X$. Tale successione si dirà *fondamentale o di Cauchy* se $\forall \epsilon > 0 \ \exists \bar{n} = \bar{n}(\epsilon) \in \mathbb{N} : d(x_m, x_n) < \epsilon \ \forall n, m > \bar{n}$.

Se ogni successione fondamentale converge in $(X,d)$, allora il suddetto spazio metrico viene detto *completo*.

**Definizione 55.** Sia $I \subseteq \mathbb{R}$ e siano $f_n \colon I \to \mathbb{R}$. Consideriamo la successione di funzioni $\{f_n\}_{n\in\mathbb{N}}$. Diremo che $\{f_n\}$ *converge puntualmente* in $I$ a $f \colon I \to \mathbb{R}$ se la successione di numeri reali $\{f_n(x)\}_{n\in\mathbb{N}}$ converge per ogni $x \in I$: $\forall \epsilon > 0 \wedge \forall t \in I \ \exists \bar{n} = \bar{n}(\epsilon, t) \in \mathbb{N} : \left| f_n(t) - f(t) \right| < \epsilon \ \forall n > \bar{n}$

Consideriamo la successioni di funzioni $\{x^n\}_{n\in\mathbb{N}}$, con $x^n \colon [0,1] \to \mathbb{R}$. La funzione limite della successione è la funzione

$$f(x) = \begin{cases} 0 & x < 1 \\ 1 & x = 1 \end{cases}$$

che non è continua in tutto il suo dominio di definizione. Vogliamo allora trovare una nozione di convergenza per le successioni di funzioni più forte, e possiamo farlo utilizzando la definizione di infinito-distanza.

**Definizione 56.** Sia $I \subseteq \mathbb{R}$ e siano $f_n \colon I \to \mathbb{R}$. Consideriamo la successione di funzioni $\{f_n\}_{n\in\mathbb{N}}$. Diremo che $\{f_n\}$ *converge uniformemente* in $I$ a $f \colon I \to \mathbb{R}$ se $\forall \epsilon > 0 \ \exists \bar{n} = \bar{n}(\epsilon) \in \mathbb{N} :$ $\left| f_n(x) - f(x) \right| < \epsilon \ \forall n > \bar{n} \wedge \forall x \in I$.

In questo caso la funzione di prima non converge uniformemente in $I$: infatti $\max_{x\in I}\left|f_n(x) - f(x)\right| = 1\ \forall n \in \mathbb{N}$. Se invece considerassimo la medesima funzione definita sull'insieme $[0, b], b < 1$ si avrebbe che $\max_{x\in I}\left|f_n(x) - f(x)\right| = |b^n| \to 0$ per $n \to \infty$.

**Teorema 42.** *Una successione di funzioni $\{f_n\}_{n\in\mathbb{N}}$ definite su $I$ converge uniformemente in $I$ se e solo se $\{f_n\}_{n\in\mathbb{N}}$ è una successione fondamentale.*

**Teorema 43.** *Sia $\{f_n\}_{n\in\mathbb{N}}$ una successione di funzioni limitate che converge uniformente a $f : I \to \mathbb{R}$. Allora $f$ è limitata.*

*Dimostrazione.* Poiché la successione $\{f_n\}_{n\in\mathbb{N}}$ converge, è una successione di Cauchy: esiste allora un naturale $\bar{n}$ tale per cui $\left|f_n(x) - f_m(x)\right| < 1\ \forall n, m > \bar{n}$. □

**Teorema 44.** *Siano $\{l_n\}_{n\in\mathbb{N}} \subset \mathbb{R}$, $f_n : I \subseteq \mathbb{R} \to \mathbb{R}$, $n \in \mathbb{N}$ e $t_0 \in I$. Si supponga che*

1. *$\{f_n\}_{n\in\mathbb{N}}$ converga uniformemente in $I$ alla funzione $f$;*
2. *$\lim_{t\to t_0} f_n(t) = l_n$*

*Ne consegue che esistono, finiti, i limiti $\lim_{t\to t_0} f(t)$ e $\lim_{n\to\infty} l_n$, e vale che $\lim_{n\to\infty}\lim_{t\to t_0} f_n(t) = \lim_{t\to t_0}\lim_{n\to\infty} f_n(t)$.*

*Dimostrazione.* Sia dato un certo $\epsilon > 0$. Poiché $\{f_n\}$ converge uniformemente in $I$, esiste un $\bar{n}(\epsilon) \in \mathbb{N}$ tale che $\forall n, m > \bar{n} \wedge \forall x \in I$ vale che $\left|f_n(x) - f_m(x)\right| < \epsilon$. Passando al limite per $t \to t_0$, si ha che $|l_n - l_m| < \epsilon\ \forall n, m > \bar{n}$. Abbiamo quindi dimostrato che $\{l_n\}$ è una successione di Cauchy, e quindi converge in $\mathbb{R}$ a $l$. Si consideri ora $\left|f(t) - l\right| = \left|f(t) - f_n(t) + f_n(t) - l_n + l_n - l\right| \le \left|f(t) - f_n(t)\right| + \left|f_n(t) - l_n\right| + |l_n - l|$. Sia dato un certo $\bar{n}$ tale che $\left|f(t) - f_n(t)\right| \le \frac{\epsilon}{3}, \left|f_n(t) - l_n\right| \le \frac{\epsilon}{3}$ e $|l_n - l| \le \frac{\epsilon}{3}$ per ogni $n > \bar{n}$. Accorpando le varie disuguaglianze, si ottiene che $\left|f(t) - l\right| \le \epsilon\ \forall n > \bar{n}$. Abbiamo quindi l'asserto. □

**Teorema 45.** *Se $\{f_n\}$ è una successione di funzioni $f_n : I \to \mathbb{R}$ continue in $I$ che converge uniformemente a $f$, allora $f$ è continua.*

**Teorema 46.** *Sia $\{f_n\}_{n\in\mathbb{N}}$ una successione di funzioni tale che $f_n \in \mathscr{R}([a, b])$ e che $f_n \to f$ uniformemente in $[a, b]$. Allora $f \in \mathscr{R}([a, b])$ e vale che*

$$\lim_{n\to\infty}\int_a^b f_n\ dx = \int_a^b \lim_{n\to\infty} f_n\ dx = \int_a^b f\ dx$$

*Dimostrazione.* Definiamo $\epsilon_n = \sup_{[a,b]}\left|f_n(t) - f(t)\right|$; vale che

$$f_n(t) - \epsilon_n \le f(t) \le f_n(t) + \epsilon_n \qquad \forall t \in [a, b]$$

Integriamo i membri della disuguaglianza precedente: otteniamo che

$$\int_a^b f_n(t) - \epsilon_n\, dt \le \underline{\int_a^b} f(t)\, dt \le \overline{\int_a^b} f(t)\, dt \le \int_a^b f_n(t) + \epsilon_n\, dt$$

Con degli abili passaggi algebrici ci riduciamo alla forma

$$0 \le \overline{\int_a^b} f(t)\, dt - \underline{\int_a^b} f(t)\, dt \le 2\epsilon_n(b - a)$$

Poiché $\{f_n\}$ converge uniformemente a $f$, $\lim_{n\to\infty}\epsilon_n = 0$ per la definizione di $\epsilon_n$ e di convergenza uniforme. Vale quindi che

$$0 \le \overline{\int_a^b} f(t)\, dt - \underline{\int_a^b} f(t)\, dt \le 0$$

I due integrali coincidono, e per la definizione di integrale secondo Riemann $f \in \mathscr{R}([a, b])$. L'altra parte dell'asserto deriva in modo naturale dai passaggi utilizzati per la dimostrazione. □

Si consideri ora la relazione che intercorre tra convergenza uniforme e differenziabilità. Ad esempio, si prenda in considerazione la successione di funzioni $\{f_n\}_{n\in\mathbb{N}}$, con

$$f_n(t) = \frac{\sin(nt)}{n}$$

La successione converge uniformemente alla funzione costante $f(t) = 0$, in quanto $\left|\sin(nt)\right| \leq 1$ (e conseguentemente $d_\infty(f_n, f) = \sup_{t\in[a,b]}\left|f_n(t)\right| = \left|\frac{1}{n}\right| \to 0$ per $n \to \infty$). Consideriamo ora la successione delle derivate prime $f'_n(t) = \cos(nt)$, verificando in particolare che valga $\lim_{n\to\infty} \frac{d}{dt} f_n \overset{?}{=} \frac{d}{dt} \lim_{n\to\infty} f_n$. Date queste ipotesi e questa funzione la condizione non vale: infatti $\cos(nt)$ non tende uniformemente a 0 (che è $\frac{d}{dt}\lim_{n\to\infty} f_n$), come si può ben notare considerando la successione $\{f'_n(0)\}$ che è la successione costituita da 1.

**Teorema 47.** *Sia $\{f_n\}$ una successione di funzioni $f_n \colon [a, b] \to \mathbb{R}$ tale che $\{f_n(x_0)\}$ converga in $[a, b]$ per un qualche $x_0 \in [a, b]$ e tale che $\{f'_n\}$ converga uniformente in $[a, b]$. Allora $\{f_n\}$ converge uniformemente in $[a, b]$ alla funzione limite $f \colon [a, b] \to \mathbb{R}$ e vale che*

$$\lim_{n\to\infty} f'_n(x) = \frac{d}{dt} \lim_{n\to\infty} f(x) = f'(x)$$

*Dimostrazione.* Sia dato $\epsilon > 0$. Per la convergenza uniforme di $\{f'_n\}$ e per la convergenza puntuale di $\{f_n\}$ esiste un $\bar{n}$ tale che per ogni $n, m > \bar{n}$ vale che

$$\left|f_n(x_0) - f_m(x_0)\right| < \frac{\epsilon}{2}$$

$$\left|f'_n(t) - f'_m(t)\right| < \frac{\epsilon}{2(b-a)} \qquad \forall t \in [a, b]$$

Consideriamo ora il termine $\left|f_n(t) - f_n(x) - f_m(t) + f_m(x)\right|$: applichiamo il Teorema del Valor Medio ai termini nel valore assoluto, ottenendo che

$$\left|f_n(t) - f_n(x) - f_m(t) + f_m(x)\right| = \left|f'_n(\xi) - f'_m(\xi)\right|\left|t - x\right| \leq \frac{\epsilon}{2}\frac{|t - x|}{(b - a)} \leq \frac{\epsilon}{2}$$

Tale disuguaglianza è valida per ogni $t$ e per ogni $x$ in $[a, b]$ se $n, m > \bar{n}$, e quindi vale in particolare per $x_0$. Consideriamo ora $\left|f_n(t) - f_m(t)\right|$: vale che

$$\begin{aligned}\left|f_n(t) - f_m(t)\right| &= \left|f_n(t) - f_n(x_0) - f_m(t) + f_m(x_0) + f_n(x_0) - f_m(x_0)\right| \leq \\ &\left|f_n(t) - f_n(x_0) - f_m(t) + f_m(x_0)\right| + \left|f_n(x_0) - f_m(x_0)\right| < \epsilon\end{aligned}$$

per quanto detto antecedentemente. Poiché quest'ultima disuguaglianza vale, dato un certo $\epsilon > 0$, per ogni $n, m > \bar{n}$ e per ogni $t \in [a, b]$, la successione $\{f_n\}$ converge uniformemente a $f \colon [a, b] \to \mathbb{R}$ in $[a, b]$. Definiamo ora le funzione ausiliarie

$$\phi_n(t) = \frac{f_n(t) - f_n(x)}{t - x} \qquad \phi(t) = \frac{f(t) - f(x)}{t - x} \qquad \forall t \neq x$$

notando che $\lim_{t\to x} \phi_n(t) = f'_n(x)$. Riprendendo la disuguaglianza precedente con la nuova funzione, abbiamo che

$$\left|\phi_n(t) - \phi_m(t)\right| \leq \frac{\epsilon}{2(b-a)}$$

per ogni $n, m > \bar{n}$ e per ogni $t \in [a, b]$; di conseguenza, la successione $\{\phi_n\}$ è uniformemente convergente in $[a, b]$ e, per la convergenza uniforme di $\{f_n\}$ vale che

$$\lim_{n\to\infty} \phi_n(t) = \phi(t)$$

Possiamo quindi applicare il Teorema dello Scambio dei Limiti, ottenendo che

$$\lim_{t\to x} \lim_{n\to\infty} \phi_n(t) = f'(x) = \lim_{n\to\infty} \lim_{t\to x} \phi_n(t) = \lim_{n\to\infty} f'_n(x)$$

E otteniamo così l'asserto. □

**Teorema 48.** *La coppia* $(\mathscr{C}^k([a,b]), d_{\infty,k})$ *è uno spazio metrico completo.*

**Teorema 49.** *La coppia* $(\mathscr{R}([a,b]), d_2)$, *con* $d_2 := \sqrt{\int_a^b (f(x)-g(x))^2\ dx}$ *la distanza integrale di ordine 2, non è uno spazio metrico completo; può essere completato dallo spazio* $\mathscr{L}_2([a,b])$.

Siano $(X, d_x)$ e $(Y, d_y)$ due spazi metrici generici, e consideriamo funzioni $f\colon X \to Y$ che mandano elementi dello spazio metrico $X$ nello spazio metrico $Y$. Funzioni di questo tipo possono essere:

- L'integrazione secondo Riemann, che manda un elemento $f \in \mathscr{R}([a,b])$ nell'elemento $\int_a^b f(x)\ dx \in \mathbb{R}$, e che oltretutto è una funzione lineare;
- La derivazione, che manda un elemento $f \in \mathscr{C}^1([a,b])$ nell'elemento $Df \in \mathscr{C}^0([a,b])$.

Vogliamo ora generalizzare alcuni concetti delle funzioni tra spazi vettoriali alle funzioni tra spazi metrici.

**Definizione 57.** Sia $f\colon X \to Y$. Diremo che $f$ è *limitata* se $f(X)$ è limitata in $(Y, \tau_{d_y})$, ove con $\tau_{d_y}$ si indica la topologia indotta su $Y$ dalla metrica $d_y$, ovvero esiste un intorno sferico $B_y(y,\rho) : B_y(y,\rho) \supseteq f(x)$.

**Definizione 58.** Sia $f\colon X \to Y$. $f$ è detta continua se per ogni $\{x_n\}_{n\in\mathbb{N}} \subset X$ convergente a $x_0 \in X$ si ha che $\lim_{n\to\infty} f(x_n) = f(\lim_{n\to\infty} x_n) = f(x_0)$.

**Definizione 59.** Siano $(X, d_x)$ e $(Y, d_y)$ due spazi metrici, e sia $f$ una funzione tra i due spazi metrici. Diremo che $f$ è *lipschitziana* se esiste $M \in \mathbb{R}$ tale che $\forall x_1, x_2 \in X$ vale che

$$d_y(f(x_1), f(x_2)) \leq M d_x(x_1, x_2)$$

Se $M < 1$ e $f\colon X \to X$ diremo che $f$ è una *contrazione* di $X$ in $X$.

**Teorema 50** (di Banach-Caccioppoli)**.** *Sia* $(X, d_x)$ *uno spazio metrico completo (cfr. Definizione 55), e sia* $f\colon X \to X$ *una contrazione di costante* $\rho$. *Allora* $f$ *ammette un unico punto fisso*[12] $\bar{x} \in X$. *Inoltre, da ogni* $x_0 \in X$ *fissato la successione delle iterate* $x_{n+1} = f(x_n)$ *converge al punto fisso* $\bar{x}$.

*Dimostrazione.* Verifichiamo l'unicità del punto fisso. Supponiamo che esistano due punti fissi per $f$ $\bar{x_1}, \bar{x_2}$, per cui per definizione vale che $f(\bar{x_1}) = \bar{x_1}$ e $f(\bar{x_2}) = \bar{x_2}$. Sappiamo che la funzione $f$ è una contrazione: quindi

$$d_x(f(\bar{x_1}), f(\bar{x_2})) = d_x(\bar{x_1}, \bar{x_2}) \leq \rho d_x(\bar{x_1}, \bar{x_2}) < d_x(\bar{x_1}, \bar{x_2})$$

Ma la condizione è verificata solo per $d_x(\bar{x_1}, \bar{x_2}) = 0$, e quindi per $\bar{x_1} = \bar{x_2}$. Verifichiamo ora che questo punto fisso esista. Sia $x_0 \in X$ un generico punto di $X$. Definiamo la successione $\{x_n\}$ in modo che $x_{n+1} = f(x_n)$. Poiché $f$ è una contrazione, vale che

$$d_x(x_{n+1}, x_n) = d_x(f(x_n), f(x_{n-1})) \leq \rho d_x(x_n, x_{n-1}) = \cdots \leq \rho^n d_x(x_1, x_0)$$

Supponiamo ora di avere un certo $n > m$. Vale che

$$d_x(x_n, x_m) \leq d_x(x_n, x_{n-1}) + d_x(x_{n-1}, x_m) \leq \cdots \leq \sum_{i=m}^{n-1} d_x(x_{i+1}, x_i)$$

$$\sum_{i=m}^{n-1} d_x(x_{i+1}, x_i) \leq \sum_{i=m}^{n-1} \rho^i d_x(x_1, x_0) = d_x(x_1, x_0) \sum_{i=m}^{n-1} \rho^i$$

La serie $\sum_{i=m}^{n-1} \rho^i$ è una serie geometrica di ragione $\rho$, ed è uguale a $\frac{\rho^m}{1-\rho}$. Ritornando all'asserto, si ha che

$$d_x(x_n, x_m) \leq \frac{\rho^m}{1-\rho} d_x(x_1, x_0)$$

12

Dato un $\epsilon > 0$, è sempre possibile trovare un $\bar{n}$ tale che $\frac{\rho^{\bar{n}}}{1-\rho} d_x(x_1, x_0) < \epsilon$. Di conseguenza $\{x_n\}$ è una successione di Cauchy. Poiché $(X, d_x)$ è completo $\{x_n\}$ converge, e quindi $\lim_{n\to\infty} x_n = \bar{x}$. Poiché $f$ è una contrazione, $f$ è continua (in realtà uniformemente continua), e quindi si ha che

$$f(\bar{x}) = \lim_{n\to\infty} f(x_n) = \lim_{n\to\infty} x_{n+1} = \bar{x}$$

L'asserto è quindi dimostrato. □

**Definizione 60.** Sia data una serie di funzioni $\sum_{n=0}^{\infty} f_n(x)$, con $f_n \colon I \subseteq \mathbb{R} \to \mathbb{R}$. Diremo che la serie *converge puntualmente (uniformemente)* in $I$ se la successione delle somme parziali $\{S_N(x)\}$, con $S_N(x) = \sum_{n=0}^{N} f_n(x)$, converge puntualmente (uniformemente).

Si consideri a titolo di esempio la serie

$$\sum_{n=0}^{\infty} x^n \qquad S_N = \sum_{n=0}^{\infty} x^n = \frac{1 - x^{N+1}}{1 - x}$$

che converge puntualmente per $x \in (-1, 1)$ alla funzione $S(x) = \frac{1}{1-x}$. Tuttavia la convergenza non è uniforme, in quanto la successione delle ridotte non converge uniformemente a $S$. Infatti, essendo gli elementi $S_N$ polinomi per ogni $N \in \mathbb{N}$, sono limitati sull'intervallo di convergenza della serie. Se la successione convergesse uniformemente, la funzione limite $S$ sarebbe limitata, cosa non vera. All'interno dell'intervallo di convergenza è tuttavia possibile trovare un intervallo in cui la successione delle somme parziali converge uniformemente alla funzione limite. Si consideri un generico $b \in [0, 1[$, e si analizzi il comportamento della successione in $I = [-b, b]$. Considerando l'infinito-distanza, si ha

$$\sup_{x\in I} \left| S_N(x) - S(x) \right| = \sup_{x\in I} \left| \frac{-x^{N+1}}{1 - x} \right|$$

**Definizione 61** (Criterio di Cauchy)**.** Sia $\sum_{n=0}^{\infty} f_n(x)$ una serie di funzioni $f_n \colon I \subseteq \mathbb{R} \to \mathbb{R}$. La serie converge puntualmente in $I$ se e solo se $\forall \epsilon > 0 \wedge \forall x \in I \; \exists \bar{n}(\epsilon, x) : \forall p > \bar{n} \wedge q > 0 \; \left| S_{p+q} - S_p \right| < \epsilon$.

**Teorema 51** (Criterio di Weierstrass)**.** *Sia $\{f_n\}$ una successione di funzioni $f_n \colon I \subseteq \mathbb{R} \to \mathbb{R}$. Diremo che la serie $\sum_{n=0}^{\infty}$ converge uniformemente in $I$ se esistono dei $c_n \in \mathbb{R}^+$ tali per cui $\left| f_n(x) \right| \leq c_n \; \forall x \in I$ e $\sum_{n=0}^{\infty} c_n$ è convergente.*

*Dimostrazione.* Poiché la serie $\sum_{n=0}^{\infty}$ converge, la successione delle somme parziali $\{C_N\}$ è una successione fondamentale: $\forall \epsilon > 0 \; \exists \bar{n} : \forall p > \bar{n} \wedge q > 0 \; \left| C_{p+q} - C_p \right| < \epsilon$. Espandendo il termine all'interno del valore assoluto si ottiene che

$$\left| C_{p+q} - C_p \right| = \left| c_{p+q} + \cdots + c_{p+1} \right| < \epsilon$$

Sappiamo che $c_n > \left| f_n(x) \right| \; \forall x \in I$: abbiamo quindi

$$\left| f_{p+q}(x) + \cdots + f_{p+1}(x) \right| \leq \left| f_{p+q} \right| + \ldots f_{p+1} \leq c_{p+q} + \cdots + c_{p+1} < \epsilon$$

Per il criterio di Cauchy la serie di funzioni converge. □

**Definizione 62.** Si $X$ un generico insieme. Chiameremo *norma* una funzione $\|.\| \colon X \to \mathbb{R}$ che rispetti i seguenti assiomi:

1. $\|x\| \geq 0 \; \forall x \in X \wedge \|x\| = 0 \iff x = 0$;
2. $\|\lambda x\| = |\lambda| \, \|x\| \; \forall x \in X \wedge \forall \lambda \in \mathbb{C}$;
3. $\|x + y\| \leq \|x\| + \|y\| \; \forall x, y \in X$.

Una coppia $(X, \|.\|)$ viene detta *spazio di Banach.*

Ad esempio, sia $\mathscr{C}([a,b])$ lo spazio delle funzioni continue su $[a,b]$. Ad ogni $f \in \mathscr{C}([a,b])$ associamo la norma uniforme (o norma di Chebyshev) $\|f\|_\infty = \sup_{x\in[a,b]}\left|f(x)\right|$. Poiché $f$ è continua in un intervallo chiuso e limitato, $\|f\|_\infty < \infty$. Inoltre, $\|f\|_\infty = 0$ solamente se $f(x) = 0$ per ogni $x \in [a,b]$. Se consideriamo la funzione $h = f + g$, con $f, g \in \mathscr{C}([a,b])$, si ha che $\left|h(x)\right| \leq \left|f(x)\right| + \left|g(x)\right|$ per ogni $x \in [a,b]$. Di conseguenza $\|f+g\|_\infty \leq \|f\|_\infty + \|g\|_\infty$. Abbiamo quindi dimostrato che la norma uniforme è effettivamente una norma. Ogni spazio di Banach è uno spazio metrico, in quanto la norma permette di definire una distanza. Ritornando al caso precedente, la norma uniforme induce un particolare tipo di distanza, la *distanza di Chebyshev* (cfr. ).

Sia $(X, \|.\|)$ un generico spazio di Banach. Ci chiediamo, data una successione $\{x_n\} \subset X$, il significato della convergenza della serie $\sum_{n=0}^\infty x_n$. La risposta si trova nel Criterio di Cauchy: $\forall\epsilon > 0\ \exists\bar{n}(\epsilon) : \forall p > \bar{n} \wedge q > 0\ \|x_{p+q} - x_p\| < \epsilon$.

**Teorema 52** (Della Convergenza Totale)**.** *Sia $\{x_n\} \subset X$ una successione in un generico spazio di Banach. Se la successione $\{S_N\}$ delle ridotte delle norme ($S_N = \sum_{n=0}^N \|x_n\|$) degli elementi della successione è convergente, allora la serie $\sum_{n=0}^\infty x_n$ converge e vale che $\|\sum_{n=0}^\infty x_n\| \leq \sum_{n=0}^\infty \|x_n\|$.*

Se la serie $\sum_{n=0}^\infty f_n(t)$ converge uniformemente, allora $\lim_{t\to t_0} \sum_{n=0}^\infty f_n(t) = \sum_{n=0}^\infty \lim_{t\to t_0} f_n(t)$.

**Teorema 53.** *Siano $f_n\colon I \subseteq \mathbb{R} \to \mathbb{R}$ delle funzioni limitate appartenenti a $\mathscr{R}(I)$. Se la serie $\sum_{n=0}^\infty f_n(x)$ converge uniformemente in $I$, la funzione $f(x) = \sum_{n=0}^\infty f_n(x)$ è limitata e appartiene a $\mathscr{R}(I)$, e vale che*

$$\int_a^b \sum_{n=0}^\infty f_n(t)\ dt = \sum_{n=0}^\infty \int_a^b f_n(t)\ dt$$

Ad esempio, si consideri la funzione $u(t) = \frac{t^3}{e^t - 1}$. Definiamo

$$\epsilon = \int_0^\infty \frac{t^3}{e^t - 1}\ dt = \int_0^\infty \frac{t^3 e^{-t}}{1 - e^{-t}}\ dt$$

Possiamo porre

$$\frac{t^3}{1 - e^{-t}} = t^3 \sum_{n=0}^\infty e^{-nt}$$

poiché $\frac{1}{1-e^{-t}}$ è la somma della serie sopra riportata. Consideriamo ora la serie $\sum_{n=0}^\infty t^3 e^{-nt}$; in particolare verifichiamo che converga uniformemente. Osserviamo infatti che

$$\left|f_n(x)\right| = \left|t^3 e^{-nt}\right| \leq \max_{x\in[0,+\infty)} t^3 e^{-nt}$$

Derivando $t^3 e^{-nt}$ si trova che il massimo di $f_n(x)$ è in $t = \frac{3}{n}$; si ottiene quindi che $\left|f_n(x)\right| \leq \frac{27}{n^3} e^{-3}$. La serie $\sum_{n=1}^\infty \frac{27}{n^3} e^{-3}$ converge, e di conseguenza la serie delle $f_n(x)$ converge uniformemente per il Criterio di Weierstrass. Sostituendo, si ha che

$$\epsilon = \int_0^\infty \sum_{n=0}^\infty t^3 e^{-(n+1)t}\ dt \overset{U}{=} \sum_{n=0}^\infty \int_0^\infty t^3 e^{-(n+1)t}\ dt = \sum_{n=0}^\infty \lim_{\omega\to\infty} \int_0^\omega t^3 e^{-(n+1)t} =$$
$$\sum_{n=0}^\infty \lim_{\omega\to\infty} \frac{-6}{(n+1)^4} e^{-(n+1)t}\Bigg|_0^\omega = \sum_{n=0}^\infty \frac{6}{(n+1)^4} = \frac{\pi^4}{15}$$

**Definizione 63.** Sia data una successione $\{c_n\}_{n\in\mathbb{N}} \subset \mathbb{C}$; la serie $\sum_{n=0}^\infty c_n z^n$ viene detta *serie di potenze*, e i numeri $c_n$ vengono detti coefficienti della serie.

In generale, la serie potrebbe convergere o divergere a seconda della scelta di $z$. In particolare, ad ogni serie di potenze è associato un cerchio, detto *cerchio di convergenza*, tale per cui la serie converge se $z$ è all'interno del cerchio (per generalizzare si suole considerare il piano un cerchio di raggio infinito).

**Teorema 54.** *Sia $\sum_n^\infty c_n z^n$ una serie di potenze; definiamo*

$$\alpha := \lim_{n\to\infty} \sup \sqrt[n]{|c_n|} = \lim_{n\to\infty} \sup \left|\frac{c_{n+1}}{c_n}\right| \qquad \rho := \frac{1}{\alpha}$$

*La serie di potenze converge se $|z| < \rho$, e diverge se $|z| > \rho$.*

**Teorema 55.** *Sia data una serie di potenze $\sum_{n=0}^\infty c_n x^n$. Se la serie converge per un certo $x_0$, allora converge per ogni $x$ tale che $|x| < |x_0|$.*

*Dimostrazione.* Poiché $\sum_{n=0}^\infty c_n x_0^n$ converge, $\lim_{n\to\infty} c_n x_0^n = 0$: ovvero, i termini della successione $c_n x_0^n$ sono definitivamente minori di 1. Consideriamo ora

$$|c_n x^n| = \left| c_n x_0^n \frac{x^n}{x_0^n} \right| = |c_n x_0^n| \left| \frac{x^n}{x_0^n} \right| \le \left| \left( \frac{x}{x_0} \right)^n \right|$$

□

**Definizione 64.** Si definisce *polinomio trigonometrico* una somma parziale

$$f(x) = \frac{a_0}{2} + \sum_{n=1}^{N} (a_n \cos(nx) + b_n \sin(nx))$$

con $x$ reale e $a_1, \ldots, a_N$, $b_1, \ldots, b_N$ complessi.

Studiamo ora alcune caratteristiche delle serie di Fourier e delle loro somme.

- È palese che $f(x)$ sia una funzione periodica, di periodo $2\pi$ (è la somma di funzioni periodiche con periodo $\frac{2\pi}{n}$, con $n = 1, \ldots, N$. Di conseguenza ha periodo $2\pi$).
- Esiste una semplice condizione di convergenza per la serie; infatti, è possibile applicare il Criterio di Weierstrass:

  $$\left| a_n \cos(nx) + b_n \sin(nx) \right| \le \left| a_n \cos(nx) \right| + \left| b_n \sin(nx) \right| \le |a_n| + |b_n|$$

  Poiché la condizione è verificata per ogni $n$, se le serie $\sum_{n=1}^\infty |a_n|$ e $\sum_{n=1}^\infty |b_n|$ convergono per il criterio del confronto converge anche la serie di Fourier. In particolare, converge totalmente e quindi uniformemente; di conseguenza, essendo le $f_n$ delle funzioni continue anche la somma della serie è continua.
- Ammettendo che la serie di Fourier presa in esame sia una serie uniformemente convergente, si ha che $f \in \mathscr{C}(\mathbb{R}) \subset \mathscr{R}(\mathbb{R})$. Moltiplichiamo ambo i membri della serie per $\cos(mx)$

  $$f(x)\cos(mx) = \frac{a_0}{2}\cos(mx) + \sum_{n=1}^\infty (a_n \cos(nx)\cos(mx) + b_n \sin(nx)\cos(mx))$$

  e integriamo nell'intervallo $[-\pi, \pi]$

  $$\int_{-\pi}^{\pi} f(x)\cos(mx)\ dx =$$
  $$\int_{-\pi}^{\pi} \frac{a_0}{2}\cos(mx)\ dx + \int_{-\pi}^{\pi} \sum_{n=1}^\infty (a_n \cos(nx)\cos(mx) + b_n \sin(nx)\cos(mx))\ dx$$

  Per l'uniforme convergenza della serie di Fourier otteniamo

  $$\int_{-\pi}^{\pi} f(x)\cos(mx)\ dx = \int_{-\pi}^{\pi} \frac{a_0}{2}\cos(mx)\ dx +$$
  $$\sum_{n=1}^\infty \left( \int_{-\pi}^{\pi} a_n \cos(nx)\cos(mx)\ dx + \int_{-\pi}^{\pi} b_n \sin(nx)\cos(mx)\ dx \right)$$

Consideriamo gli integrali contenuti nell'indice di sommatoria; si ha[13] che

$$\int_{-\pi}^{\pi} a_n \cos(nx)\cos(mx)\ dx = \pi\delta_{nm}$$
$$\int_{-\pi}^{\pi} a_n \sin(nx)\cos(mx)\ dx = 0$$

Consideriamo il caso $m = 0$; in questo caso si ha

$$\int_{-\pi}^{\pi} f(x)\ dx = \int_{-\pi}^{\pi} \frac{a_0}{2}\ dx = a_0\pi$$

Per $m \neq 0$ invece

$$a_m = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos(mx)\ dx$$

Moltiplicando ambo i membri per $\sin(mx)$ si ottiene invece che

$$b_m = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin(mx)\ dx$$

$a_n$ e $b_n$ vengono detti *coefficienti di Fourier* di $f$.

- Si noti che la restrizione al caso in cui il periodo $\tau$ è $2\pi$ non è limitativa: infatti, si consideri una serie di Fourier definita nel modo seguente

$$\frac{a_0}{2} + \sum_{n=1}^{N}\left(a_n \cos\left(\frac{2\pi n}{\tau}y\right) + b_n \sin\left(\frac{2\pi n}{\tau}y\right)\right)$$

In questo caso la somma della serie $g(y)$ (se quest'ultima è convergente) risulta essere una funzione di periodo $\tau$, con $\tau$ generico. Si può facilmente osservare come da questo caso ci si possa ricondurre al caso sopra riportato operando un cambio di variabile, ed in particolare ponendo

$$x = \frac{2\pi}{\tau}y \qquad f(x) = f\left(\frac{2\pi}{\tau}y\right) = g(y)$$

Se, data una certa funzione $f$, siamo in grado di calcolare i coefficienti di Fourier $a_n$ e $b_n$, allora siamo in grado di scrivere la serie di Fourier ad essa associata, e scriveremo

$$f(x) \sim \frac{a_0}{2} + \sum_{n=1}^{N}\left(a_n \cos\left(\frac{2\pi n}{\tau}x\right) + b_n \sin\left(\frac{2\pi n}{\tau}x\right)\right)$$

Tuttavia, il poter calcolare i coefficienti non implica nulla sulla convergenza della serie e sul fatto che la somma sia la funzione $f$. A titolo di esempio, si consideri la funzione $\{x\} := x - [x]$, ovvero la mantissa di $x$. È una funzione periodica di periodo 1. Possiamo pertanto calcolarne i coefficienti di Fourier:

$$a_0 = 2\int_0^1 x\ dx = 1$$
$$a_n = 2\int_0^1 x\cos(2\pi nx)\ dx = \frac{1}{n\pi}\int_0^1 \sin(2\pi nx)\ dx = \frac{1}{2n^2\pi^2}(\cos(2\pi nx)|_1^0) = 0$$
$$b_n = 2\int_0^1 x\sin(2\pi nx)\ dx = \frac{-1}{n\pi}\left(1 - \int_0^1 \cos(2\pi nx)\ dx\right) = \frac{-1}{n\pi}$$

Di conseguenza $\{x\} \sim \frac{1}{2} + \sum_{n=1}^{\infty} \frac{-1}{n\pi}\sin(2\pi nx)$. Si può facilmente notare che la funzione $f$ non è la somma della serie: infatti per $x = 0$ la serie di Fourier vale 0.5, mentre $\{0\} = 0$.

[13] Si sono utilizzate le formule di Werner: $\sin\alpha\cos\beta = \frac{1}{2}(\sin(\alpha+\beta) + \sin(\alpha-\beta))$, $\cos\alpha\cos\beta = \frac{1}{2}(\cos(\alpha+\beta) + \cos(\alpha-\beta))$ e $\sin\alpha\cos\beta = \frac{1}{2}(\cos(\alpha-\beta) - \cos(\alpha+\beta))$.

Consideriamo ora il prolungamento periodico (diciamolo $f$) della restrizione all'intervallo $[-\pi, \pi]$ della funzione $g\colon \mathbb{R} \to \mathbb{R}$ che manda $x$ in $x^2$. Poiché la funzione considerata è una funzione pari, l'unico contributo nella serie di Fourier associata alla funzione deriverà dal coseno. Pertanto

$$\begin{aligned}
a_0 =& \frac{1}{\pi}\int_{-\pi}^{\pi} x^2\ dx = \frac{1}{\pi}\frac{2\pi^3}{3} = \frac{2\pi^2}{3}\\
a_n =& \frac{1}{\pi}\int_{-\pi}^{\pi} x^2\cos(nx)\ dx = \frac{-2}{\pi n}\left(\int_{-\pi}^{\pi} x\sin(nx)\ dx\right) =\\
& \frac{2}{\pi n^2}\left(x\cos(nx)|_{-\pi}^{\pi} - \int_{-\pi}^{\pi}\cos(nx)\ dx\right) = \frac{2}{\pi n^2}2\pi(-1)^n = \frac{4}{n^2}(-1)^n
\end{aligned}$$

Si ha quindi che $f(x) \sim \frac{\pi^2}{3} + 4\sum_{n=1}^{\infty}\frac{(-1)^n}{n^2}\cos(nx)$. In questo caso vale l'uguaglianza: infatti la serie è uniformemente convergente. La serie $\sum_{n=1}^{\infty}|a_n|$ è infatti convergente, e di conseguenza la serie di Fourier è totalmente convergente.

**Definizione 65.** Sia $\{\phi_n\}$ una successione di funzioni (a valori reali o complessi) su $[a, b]$ tali che

$$\int_a^b \phi_n(x)\overline{\phi_m(x)}\ dx = 0\ (n \neq m) \qquad \int_a^b \left|\phi_n(x)\right|^2\ dx = 1\ \forall n$$

Diremo che $\{\phi_n\}$ è un *sistema ortonormale di funzioni su* $[a, b]$.

Si considerino ad esempio

$$\phi_0(x) = \frac{1}{\sqrt{2\pi}} \qquad \phi_{2m}(x) = \frac{\cos(mx)}{\sqrt{\pi}} \qquad \phi_{2m-1}(x) = \frac{\sin(mx)}{\sqrt{\pi}}\ m = 1, \ldots, n$$

Queste funzioni costituiscono un sistema ortonormale di funzioni. Dato un sistema ortonormale di funzioni su $[a, b]$, i coefficienti di Fourier ottenuti nel modo sopra indicato vengono detti i coefficienti di Fourier di $f$ relativamente a $\{\phi_n\}$. Consideriamo ora delle funzioni $2\pi$-periodiche integrabili secondo Riemann e le serie di Fourier ad esse associate. Vale il seguente Teorema.

**Teorema 56.** *Sia $\{\phi_n\}$ un sistema ortonormale di funzioni su $\mathbb{R}$, sia $f$ una funzione, sia $S_N$ il polinomio trigonometrico di grado $N$ che corrisponde alla somma parziale della serie di Fourier associata a $f$ e sia $\sigma_N$ un generico polinomio trigonometrico di grado $N$. Valgono le seguenti affermazioni:*

1. *Al variare di $\sigma_N$ tra tutti i polinomi trigonometrici di grado $N$, in corrispondenza di $S_N$ l'approssimazione in media quadratica di $f$ è la migliore (i.e. lo scarto quadratico medio è minimo)*
$$\int_{-\pi}^{\pi}\left|f(x) - S_N(x)\right|^2\ dx \leq \int_{-\pi}^{\pi}\left|f(x) - \sigma_N(x)\right|^2\ dx$$
2. $$\int_{-\pi}^{\pi}\left|f(x) - S_N(x)\right|^2\ dx = \int_{-\pi}^{\pi}\left|f(x)\right|^2\ dx - \pi\left(\frac{a_0}{2} + \sum_{n=1}^{N}(a_n^2 + b_n^2)\right)$$
3. $$\int_{-\pi}^{\pi}\left|f(x)\right|^2\ dx \geq \pi\left(\frac{a_0}{2} + \sum_{n=1}^{\infty}(a_n^2 + b_n^2)\right)$$

*Dimostrazione.* Consideriamo $\int_{-\pi}^{\pi}\left|f(x) - \sigma_N(x)\right|^2\ dx$. Vale che

$$\int_{-\pi}^{\pi}\left|f(x) - \sigma_N(x)\right|^2\ dx = \langle f - \sigma_N, f - \sigma_N\rangle = \langle f, f\rangle - 2\langle f, \sigma_N\rangle + \langle\sigma_N, \sigma_n\rangle$$

$$\int_{-\pi}^{\pi}\left|f(x)\right|^2\ dx - 2\int_{-\pi}^{\pi} f(x)\sum_{n=1}^{N}\gamma_n\phi_n(x)\ dx + \int_{-\pi}^{\pi}\sum_{n=1}^{N}\gamma_n\phi_n(x)\sum_{k=1}^{N}\gamma_k\phi_k(x)\ dx$$

Inoltre, per le Definizioni precedenti si ha che

$$\int_{-\pi}^{\pi} f(x) \sum_{n=1}^{N} \gamma_n \phi_n(x)\ dx = \sum_{n=1}^{N} \gamma_n \int_{-\pi}^{\pi} f(x)\phi_n(x)\ dx = \sum_{n=1}^{N} c_n \gamma_n$$

$$\int_{-\pi}^{\pi} \sum_{n=1}^{N} \gamma_n \phi_n(x) \sum_{k=1}^{N} \gamma_k \phi_k(x)\ dx = \sum_{n=1}^{N}\sum_{k=1}^{N} \gamma_n \gamma_k \int_{-\pi}^{\pi} \phi_n(x)\phi_k(x)\ dx =$$

$$\sum_{n=1}^{N}\sum_{k=1}^{N} \gamma_n \gamma_k \delta_{nk} = \sum_{n=1}^{N} \gamma_n^2$$

Sostituendo, si ha che

$$0 \leq \int_{-\pi}^{\pi} \left| f(x) - \sigma_N(x) \right|^2\ dx = \int_{-\pi}^{\pi} \left| f(x) \right|^2\ dx - 2\sum_{n=1}^{N} c_n \gamma_n + \sum_{n=1}^{N} \gamma_n^2$$

Il termine a destra può essere riscritto nel modo seguente

$$0 \leq \int_{-\pi}^{\pi} \left| f(x) - \sigma_N(x) \right|^2\ dx = \int_{-\pi}^{\pi} \left| f(x) \right|^2\ dx - \sum_{n=1}^{N} c_n^2 + \sum_{n=1}^{N} |\gamma_n - c_n|^2$$

Il valore minimo lo si raggiunge per $\gamma_n = c_n$, e abbiamo così il primo asserto. Magheggiando un pò si ottiene, per $\gamma_n = c_n$, che

$$\int_{-\pi}^{\pi} \left| f(x) \right|^2\ dx \geq \sum_{n=1}^{N} c_n^2$$

Abbiamo quindi anche il secondo asserto. Per il $\lim_{N\to\infty}$ abbiamo anche il terzo asserto (detto *Disuguaglianza di Bessel*). □

Vogliamo ora occuparci della convergenza della serie di Fourier associata ad $f$ ad $f$.

**Teorema 57** (Lemma di Riemann-Lebesgue)**.** *Sia $f\colon [a,b] \to \mathbb{R}$ una funzione Riemann-integrabile. Vale che*

$$\lim_{\alpha\to\infty} \int_a^b f(t)\cos(\alpha t)\ dt = \lim_{\alpha\to\infty} \int_a^b f(t)\sin(\alpha t)\ dt = 0$$

**Definizione 66.** Diremo che $f\colon [a,b] \to \mathbb{R}$ è *continua a tratti* se è continua in $[a,b]$ a meno di un numero finito di punti in cui esistono (finiti) i limiti destro e sinistro (diciamoli $f(x+)$ e $f(x-)$). Se, data una funzione $f\colon \mathbb{R} \to \mathbb{R}$, è verificata la condizione precedente per ogni intervallo finito, allora $f$ è continua a tratti.

Indicheremo con $\mathscr{P}_\tau$ l'insieme delle funzioni definite su $\mathbb{R}$ a valori reali $\tau$-periodiche e continue a tratti.

**Definizione 67** (Criterio di Dirichlet)**.** Data una funzione $f\colon [a,b] \to \mathbb{R}$, diremo che $f$ soddisfa la condizione di Dirichlet in $x_0 \in (a,b)$ se è valida una delle seguenti condizioni:

1. $f$ è derivabile in $x_0$;
2. $f$ è continua in $x_0$ ed esistono finiti $f(x_0+)$ e $f(x_0-)$;
3. $f$ presenta una discontinuità di salto in $x_0$ ed esistono (finiti)

$$f'^*(x_0+) = \lim_{x\to x_0^+} \frac{f(x) - f(x+)}{x - x_0} \qquad f'^*(x_0-) = \lim_{x\to x_0^-} \frac{f(x) - f(x-)}{x - x_0}$$

Alla luce di questo Teorema possiamo riprendere il caso della funzione mantissa $\{x\}$: infatti, computando $f(0+)$ e $f(0-)$ si ottiene che $s(0) = \frac{f(0+)+f(0-)}{2} = \frac{0+1}{2}$, compatibilmente a quanto si osserva nella serie di Fourier associata $s(0) = \frac{1}{2} + \sum_{n=1}^{\infty} \frac{-1}{n\pi}\sin(n0) = \frac{1}{2}$.

**Teorema 58** (di Dirichlet)**.** *Sia $f \in \mathscr{P}_{2\pi}$. Supponiamo che l'intervallo di periodicità $[-\pi, \pi]$ possa essere suddiviso in un numero finito di sottointervalli in cui $f$ è monotona. Allora la serie di Fourier associata ad $f$ converge a $s(x) = \frac{f(x+)+f(x-)}{2}$.*

Questo Teorema assume particolare rilevanza nel caso di funzioni che presentano punti di non derivabilità particolari, come $f(x) = \sqrt{|x|}$.

**Teorema 59** (Formula di Dirichlet)**.** *Definiamo il* Nucleo di Dirichlet*: $\forall n \in \mathbb{N}$ vale che*

$$\frac{1}{2} + \cos(x) + \cos(2x) + \cdots + \cos(nx) = \frac{\sin\left(\left(n + \frac{1}{2}\right)x\right)}{2\sin\left(\frac{x}{2}\right)} = D_n(x)$$

*Sia $f\colon \mathbb{R} \to \mathbb{R}$. Se $S_N(x)$ è la somma parziale della serie di Fourier associata a $f$, allora*

$$S_N(x) = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)D_N(t)\ dt = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)\frac{\sin\left(\left(n + \frac{1}{2}\right)t\right)}{2\sin\left(\frac{t}{2}\right)}\ dt$$

*Dimostrazione.*

- Dimostriamo che vale l'uguaglianza per il Nucleo di Dirichlet. Procediamo per induzione su $n$. La condizione è banalmente verificata per $n = 0$; supponiamola vera per $n - 1$ e verifichiamo che sia verificata per $n$.

$$\frac{1}{2} + \cos(x) + \cos(2x) + \cdots + \cos((n-1)x) + \cos(nx) = \frac{\sin\left(n - \frac{1}{2}\right)x}{2\sin\left(\frac{x}{2}\right)} + \cos(nx)$$

$$\frac{\sin\left(n - \frac{1}{2}\right)x}{2\sin\left(\frac{x}{2}\right)} + \cos(nx) = \frac{\sin\left(n - \frac{1}{2}\right)x + \cos(nx)2\sin\left(\frac{x}{2}\right)}{2\sin\left(\frac{x}{2}\right)}$$

Applicando le forumle di Werner si ottiene che

$$\sin\left(n - \frac{1}{2}\right)x + \cos(nx)2\sin\left(\frac{x}{2}\right) = \sin\left(n - \frac{1}{2}\right)x + \sin\left(n + \frac{1}{2}\right)x + \sin\left(n - \frac{1}{2}\right)x$$

- Consideriamo ora il polinomio trigonometrico di grado $N$ ottenuto dalla serie di Fourier associata a $f$

$$S_N(x) = \frac{a_0}{2} + \sum_{n=1}^{N}\Big(a_n\cos(nx) + b_n\sin(nx)\Big)$$

Sostituendo i valori effettivi per i coefficienti della serie, si ha che

$$\frac{1}{2\pi}\int_{-\pi}^{\pi} f(t)\ dt + \frac{1}{\pi}\sum_{n=1}^{N}\left(\int_{-\pi}^{\pi} f(t)\cos(nt)\ dt\cos(nx) + \int_{-\pi}^{\pi} f(t)\sin(nt)\ dt\sin(nx)\right) =$$
$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left[\frac{1}{2} + \sum_{n=1}^{N}\Big(\cos(nt)\cos(nx) + \sin(nt)\sin(nx)\Big)\right]\ dt =$$
$$\frac{1}{\pi}\int_{-\pi}^{\pi} f(t)\left[\frac{1}{2} + \sum_{n=1}^{N}\cos(n(t-x))\right]\ dt = \frac{1}{\pi}\int_{-\pi}^{\pi} f(t)D_N(t-x)\ dt = \frac{1}{\pi}\int_{-\pi}^{\pi} f(s-t)D_N(s)\ ds$$

L'asserto è quindi dimostrato.

□

**Teorema 60** (Criterio per la convergenza puntuale)**.** *Sia data una funzione $f \in \mathscr{P}_{2\pi}$. La serie di Fourier associata a $f$ converge in ogni punti $x \in \mathbb{R}$ in cui è verificata la condizione di Dirichlet e vale che*

$$s(x) = \frac{a_0}{2} + \sum_{n=1}^{\infty} \Big( a_n \cos(nx) + b_n \sin(nx) \Big) = \frac{f(x+) + f(x-)}{2}$$

*Dimostrazione.* Consideriamo il Nucleo di Dirichlet: si ha che

$$\int_0^{\pi} \frac{1}{2} + \sum_{n=1}^{N} \cos(nx)\ dx = \int_0^{\pi} D_N(x)\ dx = \frac{\pi}{2} = \int_{-\pi}^{0} D_N(x)\ dx$$

in quanto l'integrale da 0 a $\pi$ di $\cos(nx)$ è sempre nullo. Consideriamo ora $S_N(x) - \frac{f(x+)+f(x-)}{2}$: abbiamo che

$$\begin{aligned}
&S_N(x) - \frac{1}{\pi}\int_0^{\pi} f(x+)D_N(t)\ dt - \frac{1}{\pi}\int_{-\pi}^{0} f(x-)D_N(t)\ dt = 0 \\
&\frac{1}{\pi}\int_{-\pi}^{\pi} f(x+t)D_N(t)\ dt - \frac{1}{\pi}\int_0^{\pi} f(x+)D_N(t)\ dt - \frac{1}{\pi}\int_{-\pi}^{0} f(x-)D_N(t)\ dt = 0 \\
&\frac{1}{\pi}\int_0^{\pi} f(x+t) - f(x+)D_N(t)\ dt + \frac{1}{\pi}\int_{-\pi}^{0} f(x+t) - f(x-)D_N(t)\ dt = 0
\end{aligned}$$

Definiamo ora la funzione ausiliaria

$$F(t) = \begin{cases} \frac{f(x+t)-f(x-)}{2\sin\left(\frac{t}{2}\right)} & -\pi \le t < 0 \\ 0 & t = 0 \\ \frac{f(x+t)-f(x+)}{2\sin\left(\frac{t}{2}\right)} & 0 < t \le \pi \end{cases}$$

Questa funzione è una funzione continua in $\mathbb{R} \setminus \{0\}$; verifichiamo cosa accade in 0.

$$\begin{aligned}
\lim_{t\to 0^+} F(t) &= \frac{f(x+t) - f(x+)}{2\sin\left(\frac{t}{2}\right)} \cong \frac{f(x+t) - f(x+)}{t} = f'^*(x+) \\
\lim_{t\to 0^-} F(t) &= \frac{f(x+t) - f(x-)}{2\sin\left(\frac{t}{2}\right)} \cong \frac{f(x+t) - f(x-)}{t} = f'^*(x-)
\end{aligned}$$

Sebbene la funzione non sia continua in 0, i limiti destro e sinistro esistono e sono finiti: è valida la condizione di Dirichlet e la funzione $F$ è integrabile secondo Riemann. Consideriamo ora

$$\int_{-\pi}^{\pi} F(t)\sin\left(\left(N + \frac{1}{2}\right)t\right)\ dt =$$
$$\int_{-\pi}^{\pi} F(t)\sin(Nt)\cos\left(\frac{t}{2}\right)\ dt + \int_{-\pi}^{\pi} F(t)\cos(Nt)\sin\left(\frac{t}{2}\right)\ dt$$

Per il Lemma di Riemann-Lebesgue tendono entrambi a 0 per $N \to \infty$. Si ha quindi che

$$\lim_{N\to\infty} S_N(x) - \frac{1}{2}(f(x+) + f(x-)) = 0$$

e abbiamo l'asserto. □

**Teorema 61** (Criterio per la convergenza uniforme)**.** *Sia $f\colon \mathbb{R} \to \mathbb{R}$, $f \in \mathscr{P}_{2\pi}$ una funzione continua con derivata continua a tratti. La serie di Fourier associata a $f$ converge uniformemente in $\mathbb{R}$.*

*Dimostrazione.* Per il Criterio di Weierstrass, affinché la serie di Fourier associata ad una data funzione converga uniformemente è sufficiente che le serie $\sum_n^\infty |a_n|$ e $\sum_n^\infty |b_n|$ convergano. Per ipotesi sappiamo che $f'$ è continua a tratti; $f'$ è quindi integrabile seconodo Riemann in $\mathbb{R}$. Inoltre, $f' \in \mathscr{P}_{2\pi}$: pertanto possiamo calcolare la serie di Fourier ad essa associata. I coefficienti risultano essere

$$a'_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f'(x)\cos(nx)\ dx = \frac{1}{\pi}\left[f(x)\cos(nx)|_{-\pi}^{\pi} + n\int_{-\pi}^{\pi} f(x)\sin(nx)\ dx\right] = \frac{1}{\pi}nb_n$$
$$b'_n = \frac{1}{\pi}\int_{-\pi}^{\pi} f'(x)\sin(nx)\ dx = \frac{1}{\pi}\left[f(x)\sin(nx)|_{-\pi}^{\pi} - n\int_{-\pi}^{\pi} f(x)\cos(nx)\ dx\right] = \frac{-1}{\pi}na_n$$

Sfruttiamo ora la disuguaglianza di Bessel

$$\int_{-\pi}^{\pi} \left|f'(x)\right|^2\ dx \geq \frac{a_0^2}{2} + \sum_{n=1}^{\infty}(a_n'^2 + b_n'^2) = \frac{a_0^2}{2} + \sum_{n=1}^{\infty}(n^2a_n^2 + n^2b_n^2)$$

Poiché $f' \in \mathscr{R}(\mathbb{R})$, l'integrale a sinistra del segno di disuguaglianza esiste finito; pertanto le serie $\sum_n^\infty n^2|a_n|^2$ e $\sum_n^\infty n^2|b_n|^2$ convergono. Consideriamo ora la disuguaglianza $2ab \leq a^2 + b^2$, ponendo $a = n|a_n|$ e $b = \frac{1}{n}$: otteniamo che

$$|a_n| = n|a_n|\,\frac{1}{n} \leq (n^2|a_n|^2 + \frac{1}{n^2})$$

Consideriamo le serie associate ai termini che entrano nella disuguaglianza: si ha che

$$\sum_{n=1}^{\infty}|a_n| \leq \sum_{n=1}^{\infty}\left(n^2|a_n|^2 + \frac{1}{n^2}\right) = \sum_{n=1}^{\infty} n^2|a_n|^2 + \sum_{n=1}^{\infty}\frac{1}{n^2}$$

Entrambe le serie a destra del segno di disuguaglianza sono convergenti; pertanto anche la serie $\sum_n^\infty |a_n|$ è convergente. Un ragionamento analogo può essere fatto considerando la serie $\sum_n^\infty |b_n|$: pertanto per il Criterio di Weierstrass la serie di Fourier associata a $f$ converge uniformemente a $f$. □

# 8 Equazioni differenziali

Sia la coppia $(\mathrm{End}(\mathbb{R}^n), \|.\|_{Eucl})$ uno spazio di Banach, e sia $A \in \mathrm{End}(\mathbb{R}^n)$ un'applicazione lineare. Come possiamo dare significato all'espressione $e^A$? Consideriamo innanzitutto, date due applicazioni lineari $A$ e $B$, $(AB)_{ij}^2 = \left(\sum_{k=1}^n a_{ik}b_{kj}\right)^2$. Possiamo esprimere il termine a destra del segno di uguaglianza come $\langle \mathbf{A_i}, \mathbf{B_j}\rangle^2$, ove $\mathbf{A_i} = \{a_{i1}, \ldots, a_{in}\}$ e $\mathbf{B_i} = \{b_{1j}, \ldots, b_{nj}\}$. Possiamo ora applicare la disuguaglianza di Cauchy-Schwarz: otteniamo così che $\langle \mathbf{A_i}, \mathbf{B_j}\rangle^2 \leq \|\mathbf{A_i}\|^2\|\mathbf{B_j}\|^2$. Consideriamo ora la $\|AB\|_{Eucl}^2 = \sum_{i,j=1}^n (AB)_{ij}^2 \leq \sum_{i,j}^n \|\mathbf{A_i}\|^2\|\mathbf{B_j}\|^2$. Sviluppando la somma del termine a destra del segno di disuguaglianza si ottiene che $\|AB\|_{Eucl}^2 \leq \|A\|_{Eucl}^2\|B\|_{Eucl}^2$. Generalizzando il risultato, si può affermare che $\|A^k\|_{Eucl} \leq \|A\|_{Eucl}^k$. Sia ora $\|A\|_{Eucl}$ un certo numero appartenente a $\mathbb{R}$. Possiamo utilizzare lo sviluppo di Mc Laurin dell'esponenziale per scrivere $e^{\|A\|} = \sum_{k=0}^\infty \frac{\|A\|^k}{k!}$, ove la serie è una serie di Cauchy. Definiamo allora la funzione esponenziale di una matrice quadrata (che chiameremo *esponenziale di matrice*) come

$$e^A := \sum_{k=0}^{\infty}\frac{A^k}{k!}$$

Poiché $\|A\|^k$ maggiora la $\|A^k\|$, per il teorema del confronto la serie sopra indicata converge sempre (e quindi la matrice esponenziale è sempre ben definita). In particolare, il risultato è una matrice $e^A \in M_n(\mathbb{R})$.
Si consideri a titolo di esempio la matrice $A = \left(\begin{smallmatrix} 0 & -\alpha \\ \alpha & 0 \end{smallmatrix}\right)$: si ha che

$$A^0 = I_n \qquad A^1 = A \qquad A^2 = \left(\begin{smallmatrix} -\alpha^2 & 0 \\ 0 & -\alpha^2 \end{smallmatrix}\right) \qquad A^3 = \left(\begin{smallmatrix} 0 & \alpha^3 \\ -\alpha^3 & 0 \end{smallmatrix}\right)$$

Sommando i vari termini si ottiene che

$$e^A = \begin{pmatrix} 1 - \frac{\alpha^2}{2} + \dots & -\alpha + \frac{\alpha^3}{6} + \dots \\ \alpha - \frac{\alpha^3}{6} \dots & 1 - \frac{\alpha^2}{2} + \dots \end{pmatrix} = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

Sia $\gamma\colon \mathbb{R} \to \mathrm{End}(\mathbb{R}^n)$ una funzione differenziabile tale che $t \mapsto \gamma(t) = e^{tA}$, con $A \in \mathrm{End}(\mathbb{R}^n)$. Consideriamo $d\gamma(t)h = \gamma'(t)h$. Com'è definito? La funzione $\gamma$ definita precedentemente può essere vista come una serie di potenze $\gamma(t) = \sum_{k=0}^{\infty} c_k t^k$; di conseguenza, $\gamma'(t) = \sum_{k=1}^{\infty} k c_k t^{k-1}$. In questo caso si ha che

$$\alpha = \lim_{k\to\infty} \sup \frac{\|A\|^{k+1}}{(k+1)!} \frac{k!}{\|A\|^k} = \lim_{k\to\infty} \frac{\|A\|}{k+1} = 0 \qquad \rho = \infty$$

$$\gamma(t) = \sum_{k=0}^{\infty} \frac{A^k}{k!} t^k \qquad \gamma'(t) = \sum_{k=1}^{\infty} \frac{A^k}{(k-1)!} t^{k-1} = A \sum_{j=0}^{\infty} \frac{A^j}{j!} t^j = Ae^{tA}$$

Consideriamo ora l'uguaglianza

$$\gamma'(t) = A\gamma(t)$$

Questa è un'*equazione differenziale ordinaria scritta in forma*, la cui soluzione è una funzione $\gamma\colon \mathbb{R} \to \mathrm{End}(\mathbb{R}^n)$ a valori in endomorfismi. Possiamo inoltre costruire un cosiddetto *problema di Cauchy*, imponendo una condizione iniziale $\gamma(0)$. Sappiamo già che una soluzione dell'equazione differenziale è $\gamma(t) = e^{tA}$. Ma è l'unica soluzione? Consideriamo la funzione $\tilde{\gamma}\colon \mathbb{R} \to \mathrm{End}(\mathbb{R}^n)$, e supponiamo che sia un'altra soluzione dell'equazione differenziale sopra riportata.

**Definizione 68.** Un'*equazione differenziale ordinaria di ordine n* (ordine di derivazione della funzione incognita, ndr) è un'equazione del tipo

$$F(t, y(t), y'(t), \dots, y^{(n)}(t)) = 0$$

ove $F$ è una funzione $F\colon U \subseteq \mathbb{R}^{n+2} \to \mathbb{R}$. Un'equazione differenziale si dice espressa in forma normale quando è nella forma

$$y^{(n)}(t) = f(t, y(t), y'(t), \dots, y^{(n-1)}(t))$$

- Se la funzione $F$ non dipende da $t$, l'equazione differenziale si dice *autonoma*;
- Se $F$ è un polinomio di primo grado in $y, y', \dots, y^{(n)}$ l'equazione differenziale viene detta *lineare*

  $$\alpha_n y^{(n)}(t) + \dots + \alpha_1 y'(t) + \alpha_0 y(t) = b(t)$$

  Se $b(t) = 0$, l'equazione differenziale si dice *omogenea*.

**Definizione 69.** Si dice *soluzione* (o *integrale*) di $y^{(n)}(t) = f(t, y(t), y'(t), \dots, y^{(n-1)}(t))$ una funzione $\phi \in \mathscr{C}(I \subseteq \mathbb{R})$ tale che $(t, \phi(t), \phi'(t), \dots, \phi^{(n-1)}(t)) \in D$ (ove con $D$ si indica il dominio di $f$) e che $\phi^{(n)}(t) = f(t, \phi(t), \phi'(t), \dots, \phi^{(n-1)}(t))$ per ogni $t \in I$.

È d'uopo notare che si richiede che la funzione sia $\mathscr{C}^n$ solamente in un certo intervallo $I$, e che l'uguaglianza valga in quell'intervallo: quindi la proprietà di essere soluzione di un'equazione differenziale è una proprietà locale.

**Definizione 70.** Chiameremo *sistema di equazioni differenziali in forma normale del primo ordine* un sistema

$$\begin{cases} \dot{y}_1 = f_1(t, y_1, \dots, y_n) \\ \vdots \\ \dot{y}_n = f_n(t, y_1, \dots, y_n) \end{cases}$$

ove le funzioni $f_i\colon D \to \mathbb{R}$ accoppiano le equazioni all'interno del sistema.

È possibile riformulare il sistema sopra riportato nel modo seguente. Si considerino la funzione $\mathbf{y}\colon \mathbb{R} \to \mathbb{R}^n$ tale che $\mathbf{y}(t) = (y_1(t), \dots, y_n(t))$ e la funzione $\mathbf{F}\colon \mathbb{R} \to \mathbb{R}^n$ tale che $\mathbf{F}(t, \mathbf{y}(t)) = (f_1(t), \dots, f_n(t))$. Possiamo riscrivere il sistema precedente nella forma

$$\dot{\mathbf{y}}(t) = \mathbf{F}(t, \mathbf{y}(t))$$

I sistemi di equazioni differenziali del primo ordine rivestono particolare importanza perché, data un'equazione differenziale di ordine qualsiasi, è possibile ricondursi ad un sistema di equazioni differenziali del primo ordine operando le seguenti sostituizioni

$$\begin{cases} y = y_1 \\ y' = y_2 \\ \vdots \\ y^{(n-1)} = y_n \end{cases} \qquad \begin{cases} \dot{y}_1 = y_2 \\ \dot{y}_2 = y_3 \\ \vdots \\ \dot{y}_n = f(t, y_1, \dots, y_n) \end{cases}$$

**Definizione 71.** Chiameremo *problema di Cauchy* il sistema

$$\begin{cases} y^{(n)}(t) = f(t, y(t), y'(t), \dots, y^{(n-1)}) \\ y(\tau) = \epsilon_0 \\ y'(\tau) = \epsilon_1 \\ \vdots \\ y^{(n-1)}(\tau) = \epsilon_{n-1} \end{cases}$$

ove $y^{(n)}(t) = f(t, y(t), y'(t), \dots, y^{(n-1)})$ rappresenta l'equazione differenziale di ordine $n$ in forma normale, e le restanti equazioni del sistema costituiscono $n-1$ condizioni iniziali del tipo $y^{(i)}(\tau) = \epsilon_i$, $i = 0, 1, \dots, n-1$, ove $\tau$ indica l'*istante iniziale*. Risolvere un problema di Cauchy significa trovare una soluzione $\phi(t)$ dell'equazione differenziale definita in $U(\tau)$ per cui valgano le condizioni iniziali. Ciò consente di estrapolare, nella classe di infinite funzioni che costituiscono una soluzione dell'equazione differenziale, una soluzione particolare.

**Teorema 62.** *Sia data un'equazione differenziale ordinaria del primo ordine in forma normale $\dot{y}(t) = P(t)y + Q(t)$, con $P, Q \in \mathscr{C}^0(I)$. La soluzione generale (una funzione parametrica che al variare del parametro spazia tutta la classe di funzioni soluzione dell'equazione) dell'equazione differenziale è una funzione $y\colon I \to \mathbb{R}$ tale che*

$$y(t) = e^{\int P(t)\ dt}\left(c + \int Q(t) e^{-\int P(t)\ dt}\ dt\right)$$

*Dimostrazione.* Si consideri

$$\frac{d}{dt}\left(e^{-\int P(t)\ dt} y(t)\right) = -P(t)y(t)e^{-\int P(t)\ dt} + \dot{y}(t)e^{-\int P(t)\ dt} = e^{-\int P(t)\ dt}\big[\dot{y}(t) - P(t)y(t)\big]$$

Come si può notare dall'equazione differenziale, $\dot{y}(t) = P(t)y(t) + Q(t)$. Si ha quindi che

$$\frac{d}{dt}\left(e^{-\int P(t)\ dt} y(t)\right) = Q(t)e^{-\int P(t)\ dt}$$

Integrando ambo i membri e moltiplicando per $e^{\int P(t)\ dt}$ si ha l'asserto. □

Si consideri, a titolo d'esempio, il seguente problema di Cauchy.

$$\begin{cases} \dot{y} = \frac{y}{t} + 3t^3 \\ y(-1) = 2 \end{cases}$$

Scegliamo di prendere come dominio di $y(t)$ l'intervallo $(-\infty, 0)$ (contente l'istante iniziale), intervallo in cui $\dot{y}(t)$ è continua. Applicando la formula precedentemente ottenuta considerando $P(t) = \frac{1}{t}$ e $Q(t) = 3t^3$, si ha

$$y(t) = e^{\int \frac{1}{t}\ dt}\left(c + \int 3t^3 e^{-\int \frac{1}{t}\ dt}\ dt\right) = t\left(c + \int 3t^2\ dt\right) = t^4 + ct$$

Sostituendo le condizioni iniziali si ha che $c = 1$.

**Teorema 63.** *Sia data un'equazione differenziale ordinaria del primo ordine esatta*[14] $\dot{y}(t) = -\frac{P(t,y)}{Q(t,y)}$*, con* $P, Q \in \mathscr{C}^0(A \subseteq \mathbb{R}^2)$ *e* $Q \neq 0$ *in* $A$*. Esiste una funzione* $F = F(t, y)\colon U(\tau, \epsilon) \to \mathbb{R}$ *tale che* $D_t F = P$*,* $D_y F = Q$ *e* $D_t Q = D_y P$ *data da*

$$F(t, y) = \int_\tau^t P(s, y)\ ds + \int_\epsilon^y Q(\tau, s)\ ds$$

*In particolare, la soluzione dell'equazione differenziale esatta sopra riportata è data in forma implicita da* $F(t, y) = c$*,* $c \in \mathbb{R}$*.*

*Dimostrazione.* Dimostriamo l'implicazione diretta: se una funzione $\phi(t)$ è soluzione dell'equazione differenziale esatta, allora è definita implicitamente da $F(t, y) = c$. Consideriamo

$$\frac{d}{dt}F(t, y(t)) = \frac{d}{dt}F(t, \phi(t)) = \frac{\partial}{\partial t}F(t, \phi(t)) + \frac{\partial}{\partial y}F(t, \phi(t))\frac{d}{dt}\phi(t) = P(t, \phi(t)) + Q(t, \phi(t))\dot{\phi}(t) = 0$$

Abbiamo quindi la prima implicazione. È ora necessario dimostrare l'implicazione inversa. Sia dato un punto $(t_0, y_0) \in A$ tale che $F(t_0, y_0) = c$. Per il Teorema della Funzione Implicita di Dini l'equazione $F(t, y) = c$ definisce una funzione $\phi(t)$ tale che $\phi(t_0) = y_0$ e $\dot{\phi}(t) = -\frac{D_t F}{D_y F}$. Abbiamo quindi l'asserto. □

Si consideri il seguente problema di Cauchy

$$\begin{cases} \dot{y} = -\frac{t+y}{t-3y} \\ y(0) = 1 \end{cases}$$

In questo caso consideriamo $P(t, y) = t + y$ e $Q(t, y) = t - 3y$. Quest'ultima funzione è diversa da 0 per $t < 3y \wedge t > 3y$. Poiché il punto $(0, 1)$ appartiene al secondo intervallo, quello sarà l'iniziale macrointervallo di definizione della funzione. La soluzione è una funzione $y\colon U(0, 1) \to \mathbb{R}$ che in forma implicita è data da

$$F(t, y) = \int_0^t s + y\ ds - \int_1^y 3s\ ds = \frac{t^2}{2} + ty - \frac{3}{2}y^2 + \frac{3}{2}$$

Consideriamo ora $F(t, y) = 0 = t^2 + 2ty - 3y^2 + 3$.

**Teorema 64.** *Sia data un'equazione differenziale ordinaria del primo ordine a variabili separabili* $\dot{y}(t) = f(t)g(y)$*, con* $f \in \mathscr{C}^0(I \subseteq \mathbb{R})$*,* $g \in \mathscr{C}^0(J \subseteq \mathbb{R})$*. Poniamo* $P(t, y) = f(t)$ *e* $Q(t, y) = -\frac{1}{g(y)}$*; la risoluzione è così analoga al caso precedente*

$$F(t, y) = \int f(t)\ dt - \int \frac{1}{g(y)}\ dy = c$$

Si consideri ad esempio l'equazione differenziale $y' = 2t\sqrt{1 - y(t)^2}$. Utilizzando la formula precedente, si ottiene che

$$F(t, y) = \int 2t\ dt - \int \frac{1}{\sqrt{1.y(t)^2}}\ dy = t^2 - \arcsin(y) = c$$

Si ha quindi che $y(t) = \sin(t^2 + c)$.

[14] $dF$ è un differenziale esatto: $dF = \frac{\partial F}{\partial t}dt + \frac{\partial F}{\partial y}dy$.

**Teorema 65.** *Sia data un'equazione di Bernoulli $\dot{y} = P(t)y + Q(t)y^\alpha$, con $\alpha \neq 0, 1$ e $P, Q \in \mathscr{C}^0(I)$. Una soluzione banale di questa equazione differenziale è la funzione $y(t) = 0$. Per l'altra, si proceda nel modo seguente*

$$y^{-\alpha}\dot{y} = P(t)y^{1-\alpha} + Q(t)$$

*Definiamo una funzione ausiliaria $z(t) := y^{1-\alpha}$, per cui vale che $z' = (1-\alpha)y^\alpha y'$. Sostituendo si ha*

$$z' = (1-\alpha)P(t)z + (1-\alpha)Q(t)$$

*Ci siamo quindi ricondotti ad un'ODE lineare del primo ordine. Si ottiene così la soluzione $z(t)$; vale infine che $y(t) = z(t)^{\frac{1}{1-\alpha}}$.*

Si consideri a titolo di esempio il problema di Cauchy

$$\begin{cases} \dot{y} = -ty + t^3y^2 \\ y(1) = 1 \end{cases}$$

In questo caso si ponga $z(t) = y^{-1}$; si ha che $z' = tz - t^3$ e che $z(1) = 1$. Risolviamo l'equazione differenziale ordinaria lineare del primo ordine

$$z(t) = e^{\int t\ dt}\left(c - \int t^3 e^{-\int t\ dt}\ dt\right) = e^{\frac{t^2}{2}}\left(c + (t^2+2)e^{-\frac{t^2}{2}}\right) = ce^{\frac{t^2}{2}} + t^2 + 2$$

Sostituendo le condizioni iniziali, si ha che $c = -\frac{2}{\sqrt{e}}$, e che $y(t) = \frac{1}{t^2+2-2\exp(\frac{t^2}{2}-\frac{1}{2})}$.

**Teorema 66.** *Sia data un'equazione differenziale omogenea $y' = f(\frac{y}{t})$, con $f \in \mathscr{C}^0(I)$. Possiamo procedere come prima per sostituzione, ponendo $z(t) = \frac{y}{t}$ e quindi $y' = z'(t)t + z(t)$. Sostituendo, si ha che*

$$z' = \frac{z - f(z)}{t}$$

*In questo caso possiamo ricondurci ad un'equazione differenziale ordinaria esatta, ponendo $P(t,z) = f(z) - z$ e $Q(t,z) = t$.*

Si consideri l'equazione differenziale $y' = \frac{t^3+y^3}{ty^2} = (\frac{t}{y})^2 + \frac{y}{t}$. Sostituendo, si ottiene $z' = \frac{z^{-2}-z+z}{t} = \frac{1}{tz^2}$. Possiamo quindi risolverla considerandola un'equazione differenziabile a variabili separabili

$$\int z^2\ dz = \int \frac{1}{t}\ dt + c$$

Otteniamo così che $z(t) = \sqrt[3]{3\log(t)+c}$, e di conseguenza $y(t) = t\sqrt[3]{3\log(t)+c}$.

**Teorema 67.** *Sia data un'equazione di Clairault $y = ty' + \phi(y')$, con $\phi \in \mathscr{C}(I \subseteq \mathbb{R})$. Poniamo $p(t) = y'$: otteniamo così che*

$$y = tp + \phi(p)$$

*Derivando otteniamo che*

$$y' = p = p + tp' + \phi'(p)p'$$

*Si ha quindi che $p'(t + \phi'(p)) = 0$. Si può quindi procedere in due direzioni diverse. Abbiamo che $p' = 0$; di conseguenza $p(t) = c$, e quindi $y(t) = ct + \phi(c)$, equazione che rappresenta un fascio di rette. Tale fascio è detto integrale generale. Abbiamo poi una soluzione particolare, dovuta al secondo termine. Abbiamo infatti $t = -\phi'(p)$, e sostituendo si ha che $y = \phi(p) - t\phi'(p)$.*

Si consideri l'equazione differenziale $y = ty' + (y')^2$. In questo caso l'integrale generale è $y = ct + c^2$, mentre la soluzione particolare è data, in forma parametrica, da

$$\begin{cases} t = -2p \\ y = -2p^2 + p^2 = -p^2 \end{cases}$$

La suddetta equazione parametrica altro non è se non una parabola la cui equazione è $y(t) = -\frac{t^2}{4}$. Dopo aver analizzato il "bestiario" delle equazioni differenziali, è necessario verificare l'esistenza e l'unicità della soluzione. Si considerino ad esempio i seguenti casi.

- Sia data un'equazione differenziale $\dot{y} = f(t)$, con $f$ che presenta una discontinuità di salto nell'istante iniziale $\tau$. Nessun intorno di $\tau$ ammette una primitiva.
- Si consideri il seguente esempio (Peano). Sia dato il seguente problema di Cauchy

$$\begin{cases} \dot{y} = \sqrt[3]{t} \\ y(0) = 0 \end{cases}$$

La funzione $f$ è una funzione $\mathscr{C}^0(\mathbb{R})$. Tuttavia, non esiste un'unica soluzione. Infatti

$$\phi_1(t) = 0$$

$$\phi_2(t) = \begin{cases} \left(\frac{2}{3}t\right)^{1.5} & t > 0 \\ 0 & t \leq 0 \end{cases}$$

In questo caso infatti $f'(t) = \frac{1}{3\sqrt[3]{y^2}}$; la derivata non è limitata.

**Definizione 72.** Sia $\mathbf{F}\colon A \subset \mathbb{R} \times \mathbb{R}^n \to \mathbb{R}^n$. Diremo che $\mathbf{F}$ è *lipschitziana* in $A$ rispetto a $\mathbf{y}$ e uniformemente in $t$ se esiste $l \in \mathbb{R}$ per la quale vale

$$\|\mathbf{F}(t, \mathbf{y}) - \mathbf{F}(t, \mathbf{z})\| \leq l\|\mathbf{y} - \mathbf{z}\|$$

per ogni $\mathbf{y}, \mathbf{z} \in \mathbb{R}^n$ ($l$ è indipendente da $t$). Diremo altresì che $\mathbf{F}$ è localmente lipschitziana in A rispetto a $\mathbf{y}$ e uniformemente in $t$ se $\forall(\bar{t}, \bar{\mathbf{y}}) \in A$ esiste un intorno $U(\bar{t}, \bar{\mathbf{y}})$ per cui vale la definizione precedente. Logicamente, essendo $\mathbf{F} = (f_1, \ldots, f_n)$, ciascuna proprietà vale per ogni funzione $f_j$.

È bene notare che una funzione localmente lipschitziana non è sempre continua: si consideri ad esempio la funzione

$$f(t, y) = a(t)b(y)$$

con $a(t)$ una funzione localmente limitata ma non continua e $b(y)$ una funzione lipschitziana. Per costruzione, $\forall(t_0, y_0)\ \exists M_0(t_0) : a(t) < M\ \forall t \in U(t_0)$. Consideriamo ora

$$\left|f(t, y') - f(t, y)\right| = \left|a(t)(b(y') - b(y))\right| = \left|a(t)\right|\left|b(y') - b(y)\right|$$

Poiché per quanto detto antecedentemente, $\forall(t_0, y_0)\ \exists\, U(t_0) : \left|a(t)\right|\left|b(y') - b(y)\right| \leq M_0\left|b(y') - b(y)\right|$. Inoltre, poiché $b$ è lipschitziana, vale che $M_0\left|b(y') - b(y)\right| \leq M_0L\left|y' - y\right|$. Vale quindi che $\forall(t_0, y_0)\ \exists\ U(t_0, y_0) : \left|f(t, y') - f(t, y)\right| \leq M_0L\left|y' - y\right|$. $f$ è quindi localmente lipschitziana, ma non continua.

**Teorema 68** (Condizione sufficiente per la lipschitzianità locale)**.** *Se* $\mathbf{F}$ *e* $\partial_i f_j$ *sono continue in* $A$ *allora* $\mathbf{F}$ *è localmente lipschitziana in* $A$ *rispetto a* $\mathbf{y}$ *uniformemente in* $t$*.*

Consideriamo ad esempio la funzione

$$f(t, y) = \frac{|y|^\alpha}{1 + t^2} \qquad \alpha > 0$$

La funzione è palesemente continua (è composizione di funzioni continue). Le derivate parziali sono $\partial_t f = \frac{-2t|y|^\alpha}{(1+t^2)^2}$ e $\partial_y f = \frac{\alpha|y|^{(\alpha-1)}}{1+t^2}$. Se $\alpha \in (0, 1)$ $f$ non è lipschitziana in $\mathbb{R}^2$. Se $\alpha = 1$ entrambe le derivate sono continue e quindi $f$ è lipschitiziana in $\mathbb{R}^2$. Infatti

$$\left|f(t, y) - f(t, z)\right| = \frac{1}{1 + t^2}\left||y| - |z|\right| \leq \frac{1}{1 + t^2}|y - z| \leq |y - z|$$

Consideriamo ora il caso $\alpha > 1$. Si ha

$$\left|f(t, y) - f(t, z)\right| = \frac{1}{1 + t^2}\left||y|^\alpha - |z|^\alpha\right| \leq \left||y|^\alpha - |z|^\alpha\right| = \left|\alpha \int_{|z|}^{|y|} t^{\alpha-1}\ dt\right| \leq \left|\alpha \int_{|z|}^{|y|} \left|t^{\alpha-1}\right|\ dt\right|$$

Poiché $\left|t^{\alpha-1}\right|$ è una funzione continua, è limitata in qualsiasi compatto. Per Weierstrass ammette quindi un massimo globale. Si ha quindi

$$\left|f(t,y) - f(t,z)\right| \leq \max_K(x^{\alpha-1})\left\||y| - |z|\right| \leq L|y-z|$$

Si ha quindi che $f$ è localmente lipschitziana in $\mathbb{R}^2$ per $\alpha > 1$.

**Teorema 69** (Esistenza e unicità in piccolo di una soluzione)**.** *Sia* $\mathbf{F}\colon A \subseteq \mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$ *una funzione tale che* $\mathbf{F} \in \mathscr{C}^0(A)$ *(la soluzione è quindi una funzione di classe* $\mathscr{C}^1(I_\delta)$*) e tale che* $\mathbf{F}$ *sia localmente lipschitziana in* $A$*. Allora per ogni punto* $(\tau,\xi) \in A$ *esiste un intorno chiuso* $I_\delta$ *ti* $\tau$ *nel quale esiste una soluzione al problema di Cauchy*

$$\begin{cases} \dot{\mathbf{y}} = \mathbf{F}(t,\mathbf{y}) \\ \mathbf{y}(\tau) = \xi \end{cases}$$

*Tale soluzione è unica.*

*Dimostrazione.* La dimostrazione si svolge in tre passaggi fondamentali:

1. Trasformiamo il problema di Cauchy in un'equazione integrale.

   **Teorema 70.** *Siano valide le condizioni del Teorema di esistenza e unicità. Se* $\phi \in \mathscr{C}^1(I_\delta)$ *è soluzione del problema di Cauchy, allora soddisfa l'equazione integrale di Volterra*

   $$\mathbf{y}(t) = \xi + \int_\tau^t \mathbf{F}(s,\mathbf{y}(s))\ ds$$

   *Dimostrazione.* Sia $\phi\colon \mathbb{R} \to \mathbb{R}^n$ una soluzione del problema di Cauchy. Vale che $\phi'(t) = \mathbf{F}(s,\phi(s))$. Consideriamo ora l'equazione di Volterra[15]. Si ha che

   $$\phi(t) = \xi + \int_\tau^t \mathbf{F}(s,\phi(s))\ ds = \xi + \int_\tau^t \phi'(s)\ ds = \xi + \phi(t) - \phi(\tau) = \phi(t)$$

   □

   **Teorema 71.** *Sia* $\mathbf{f}\colon [a,b] \to \mathbb{R}^n$*,* $\mathbf{f} \in \mathscr{C}^0([a,b])$*. Vale la disuguaglianza*

   $$\|\int_a^b \mathbf{f}(t)\ dt\| \leq \left|\int_a^b \|\mathbf{f}(t)\|\ dt\right|$$

   *Dimostrazione.* Poniamo $\mathbf{v} := \int_a^b \mathbf{f}(t)\ dt$. Consideriamo ora

   $$\|\mathbf{v}\|^2 = \left|\langle\mathbf{v},\mathbf{v}\rangle\right| = \left|\langle\mathbf{v}, \int_a^b \mathbf{f}(t)\ dt\rangle\right| = \left|\sum_{i=1}^n v_i \int_a^b f_i(t)\ dt\right| = \left|\int_a^b \sum_{i=1}^n v_i f_i(t)\ dt\right| =$$
   $$\left|\int_a^b \langle\mathbf{v},\mathbf{f}(t)\rangle\ dt\right| \leq \left|\int_a^b \|\mathbf{v}\|\|\mathbf{f}(t)\|\right| = \|\mathbf{v}\|\left|\int_a^b \|\mathbf{f}\|(t)\ dt\right|$$

   Poiché per $\mathbf{v} = \mathbf{0}$ l'asserto è banalmente verificato, possiamo porci nel caso $\mathbf{v} \neq \mathbf{0}$. Dividendo ambo i membri per $\|\mathbf{v}\|$ e risostituendo $\mathbf{v}$ abbiamo l'asserto. □

   Consideriamo ora la funzione $f[]\colon D \to C$ che, data una funzione $\mathbf{y}(t) \in D$, la manda in $f[\mathbf{y}](t) = \xi + \int_\tau^t \mathbf{F}(s,\mathbf{y}(s))\ ds$. La funzione $f[\mathbf{y}]$ manterrà certe proprietà della funzione $\mathbf{y}$; per esempio, se $\mathbf{y} \in \mathscr{C}^0(I_\delta,\mathbb{R}^n)$, anche $f[\mathbf{y}]$ sarà continua.

[15] Data una funzione vettoriale $\mathbf{f}\colon \mathbb{R} \to \mathbb{R}^n$ tale che $\mathbf{f} = (f_i,\dots,f_n)$, vale che $\int_a^b \mathbf{f}(t)\ dt = \left(\int_a^b f_1(t)\ dt, \dots, \int_a^b f_n(t)\ dt\right)$

2. Si interpreta la soluzione dell'equazione integrale come un problema di punto fisso. Data la funzione $f[]$, vogliamo trovarne un insieme di definizione $D$ tale per cui $f[]$ è un endomorfismo. Suddetto insieme non può essere $\mathscr{C}^0(I_\delta, \mathbb{R}^n)$ perché, avendo le funzioni $\mathbf{y}$ grafico in $A$, non è detto che la funzione trasformata abbia grafico in $A$. Pertanto, consideriamo il cilindro[16] (un insieme chiuso e limitato e quindi, per il Teorema di Heine-Borel, un compatto in $\mathbb{R}^{n+1}$)

$$\Gamma \subset A \subseteq \mathbb{R}^{n+1} \qquad \Gamma = \{I_a \times \overline{B(\xi, b)}\} \qquad I_a = [\tau - a, \tau + a]$$

e consideriamo il sottoinsieme metrico $Y_\delta := \{\phi \in \mathscr{C}^0(I_\delta, \mathbb{R}^n) : \|\phi(t) - \xi\| \leq b \ \forall t \in I_\delta\}$. Poiché la norma è una funzione continua, la controimmagine di un insieme chiuso è un chiuso: di conseguenza, essendo $0 \leq \|\phi(t) - \xi\| \leq b$ un chiuso, l'insieme $Y_\delta$ è un chiuso in $\mathscr{C}^0(I_\delta, \mathbb{R}^n)$. Vale il seguente Teorema.

**Teorema 72.** *Sia $(X, d)$ uno spazio metrico completo. Sia $Y \subseteq X$; vale che $(Y, d|_y)$ è uno spazio metrico completo se e solo se $Y$ è chiuso.*

Si ha quindi che la coppia $(Y_\delta, d_\infty)$ è uno spazio metrico completo. Consideriamo ora la funzione $f[]$; vale che è un endomorfismo. Per verificarlo, controlliamo, data $\phi \in Y_\delta$, che $\|f[\phi](t) - \xi\| \leq b$.

$$\|f[\phi](t) - \xi\| = \|\int_\tau^t \mathbf{F}(s, \phi(s))\ ds\| \leq \left|\int_\tau^t \|\mathbf{F}(s, \phi(s))\|\ ds\right|$$

Poiché $\mathbf{F}(s, \phi(s))$ è una funzione continua, per il Teorema di Weierstrass ammette massimo assoluto nel compatto $\Gamma$. Si ha quindi che

$$\left|\int_\tau^t \|\mathbf{F}(s, \phi(s))\|\ ds\right| \leq \left|\int_\tau^t \max_{(t,\mathbf{y})\in\Gamma} \|\mathbf{F}(s, \phi(s))\|\ ds\right| = M|t - \tau|$$

Affinché $f[]$ sia un endomorfismo, deve valere che $M|t - \tau| \leq b$; poiché $t \in [\tau - \delta, \tau + \delta]$, vale che $\delta \leq \frac{b}{M}$. Abbiamo così ottenuto una prima stima di $\delta$. Vogliamo ora verificare che $f[]$ sia una contrazione, ovvero che $d_\infty(f[\phi](t), f[\psi](t)) \leq d_\infty(\phi, \psi)$. Andando a sostituire l'espressione per il termine a destra, si ha

$$\sup_{t\in I_\delta} \|f[\phi](t) - f[\psi](t)\| \overset{\mathscr{C}^0}{=} \max_{t\in I_\delta} \|f[\phi](t) - f[\psi](t)\| = \max_{t\in I_\delta} \|\int_\tau^t \mathbf{F}(s, \phi(s)) - \mathbf{F}(s, \psi(s))\ ds\| \leq$$
$$\max_{t\in I_\delta} \left|\int_\tau^t \|\mathbf{F}(s, \phi(s)) - \mathbf{F}(s, \psi(s))\|\ ds\right|$$

Poiché $\mathbf{F}$ è localmente lipschitziana in $A$, vale che $\|\mathbf{F}(s, \phi(s)) - \mathbf{F}(s, \psi(s))\| \leq L\|\phi(s) - \psi(s)\|$ per ogni $(s, \mathbf{v}) \in \Gamma$. Si ha quindi che

$$\max_{t\in I_\delta} \left|\int_\tau^t \|\mathbf{F}(s, \phi(s)) - \mathbf{F}(s, \psi(s))\|\ ds\right| \leq \max_{t\in I_\delta} \left|\int_\tau^t L\|\phi(s) - \psi(s)\|\ ds\right| \leq$$
$$\left|\int_\tau^t L d_\infty(\phi(s), \psi(s))\ ds\right| = L|t - \tau|\, d_\infty(\phi(s), \psi(s))$$

Affinché $f[]$ sia una contrazione, deve valere che $L\delta < 1$; quindi $\delta < \frac{1}{L}$. Ponendo $\delta = \min a, \frac{b}{M}, \frac{1}{L}$ abbiamo quindi che $f[]$ è una contrazione da uno spazio metrico completo a valori in uno spazio metrico completo.

[16] Viene sovente detto *cilindro di sicurezza*, ad indicare il suo ruolo fondamentale nell'evitare che la funzione diverga o simili (?).

3. È quindi possibile applicare il Teorema di Banach-Caccioppoli; $f[]$ ha quindi un solo punto fisso. Vale quindi che esiste un'unica funzione $\bar{\phi}$ tale per cui

$$\bar{\phi}(t) = f[\bar{\phi}](t) = \xi + \int_{\tau}^{t} \mathbf{F}(s, \bar{\phi}(s))\; ds$$

   Per il Teorema precedente, tale funzione è soluzione del Problema di Cauchy.

□

È d'uopo notare che il Teorema 70 fornisce delle condizioni sufficienti; è infatti possibile trovare un'equazione differenziale $\dot{\mathbf{y}} = \mathbf{F}(t, \mathbf{y})$ che, pur non soddisfando le condizioni del Teorema, ammette soluzioni. Si consideri, ad esempio, l'equazione differenziale

$$\dot{y} = f(y) \qquad f(y) = \begin{cases} y \log|y|\,, & y \neq 0 \\ 0, & y = 0 \end{cases}$$

La funzione è $\mathscr{C}^0(\mathbb{R})$; essa non è tuttavia localmente lipschitziana, in quanto la derivata non è limitata in un intorno di 0. Si possono tuttavia trovare delle soluzioni: quelle immediate sono le funzioni costanti $\phi(t) = 0$ e $\phi(t) = \pm 1$. Queste rette individuano delle porzioni di spazio in $\mathbb{R}^2$ in cui sono confinate le altre soluzioni, che non posso intersecare le rette sopra riportate per l'unicità della soluzione in $\mathbb{R} \setminus 0$. Risolvendo l'equazione differenziale, si ottiene che

$$y = \pm e^{\pm e^{t+c}}$$

In particolare, per $y \in (-\infty, -1)$ $y = -e^{e^{t+c}}$, per $y \in (-1, 0)$ $y = -e^{-e^{t+c}}$, per $y \in (0, 1)$ $y = e^{-e^{t+c}}$ e per $y \in (1, \infty)$ $y = e^{e^{t+c}}$.
Consideriamo il problema di Cauchy

$$\begin{cases} \dot{\mathbf{y}} = \mathbf{F}(t, \mathbf{y}) \\ \mathbf{y}(\tau) = \xi \end{cases}$$

La sua soluzione è una funzione definita in $I_\delta = [\tau - \delta, \tau + \delta]$, con $\delta = \min\{a, \frac{b}{M}\}$, a grafico in $\Gamma$. Consideriamo ora $\tau_1 = \tau + \delta$, $\mathbf{y}(\tau_1) = \xi_1$ e il problema di Cauchy

$$\begin{cases} \dot{\mathbf{y}} = \mathbf{F}(t, \mathbf{y}) \\ \mathbf{y}(\tau_1) = \xi_1 \end{cases}$$

Possiamo trovare una soluzione $\phi_1$ definita in un intorno $I_{\delta_1} = [\tau_1 - \delta_1, \tau_1 + \delta_1]$, con $\delta_1 = \min\{a_1, \frac{b_1}{M}\}$. Poiché $\delta_1$ è positivo per costruzione, $I_\delta \cap I_{\delta_1} \neq \emptyset$; di conseguenza nell'intersezione $\phi$ e $\phi_1$ coincidono per l'unicità della soluzione. A questo punto reitero il procedimento, ponendo $\tau_2 = \tau + \delta + \delta_1$ e $\mathbf{y}(\tau_2) = \xi_2$ etc. Questa procedura iterativa identifica un metodo per estendere le soluzioni locale di un Problema di Cauchy tramite un prolungamento della funzione, detto *prolungamento massimale*. Non è tuttavia sempre possibile applicare questa procedura; si consideri infatti il problema di Cauchy

$$\begin{cases} \dot{y} = y^2 \\ y(0) = 1 \end{cases}$$

la cui soluzione è $y(t) = \frac{1}{1-t}$. L'intervallo massimale (ovvero il massimo intervallo in cui l'equazione differenziale presenta una soluzione, non più prolungabile) è $(-\infty, 1)$.

**Teorema 73** (Cauchy globale)**.** *Sia $S = (\tau_1, \tau_2) \times \mathbb{R}^n$ e sia $\mathbf{F} \colon \overline{S} \to \mathbb{R}^n$. Se per $\mathbf{F}$ valgono le condizioni per l'esistenza della soluzione del problema di Cauchy in piccolo e se esistono $k_1 \geq 0$, $k_2 \geq 0$ tali che $\|\mathbf{F}(\tau, \mathbf{y})\| \leq k_1 + k_2 \|\mathbf{y}\|$ $\forall (t, \mathbf{y}) \in S$ allora per ogni $(\tau, \xi) \in S$ esiste una funzione $\phi(t)$ soluzione del problema di Cauchy in $[\tau_1, \tau_2]$.*

*Dimostrazione.* Data la coppia $(\tau, \xi)$ di condizioni iniziali costruiamo il cilindro $\Gamma := \{(t, \mathbf{y}) \in \mathbb{R} \times \mathbb{R}^n : \|\mathbf{y} - \xi\| \leq b \wedge |t - \tau| \leq a\}$ scegliendo $a$ in modo che $\Gamma \subset \overline{S}$. Consideriamo la coppia $(t, \mathbf{y}) \in \Gamma$: deve valere che $\|\mathbf{y} - \xi\| \leq b$. Vale quindi che $\|\mathbf{y}\| \leq \|\xi\| + b$. Consideriamo ora la seconda condizione. Per ogni coppia $(\tau, \xi) \in \Gamma$ vale che

$$\|\mathbf{F}(\tau, \mathbf{y})\| \leq k_1 + k_2\|\mathbf{y}\| \leq k_1 + k_2\|\xi\| + k_2 b$$

Per la continuità di $f$ e la compattezza di $\Gamma$ possiamo scrivere

$$\frac{\max_\Gamma \|\mathbf{F}(t, \mathbf{y})\|}{b} \leq \frac{k_1 + k_2\|\xi\|}{b} + k_2$$

Ponendo $b = k_1 + k_2\|\xi\|$ otteniamo che $\frac{b}{M} \geq \frac{1}{1+k_2}$, e quindi $\delta := \min\{a, \frac{b}{M}\} \geq \max\{a, \frac{1}{1+k_2}\}$. Focalizzaimoci su questo ultimo termine: abbiamo ottenuto un $\delta_0$ che non dipende più dalle condizioni iniziali $(\tau, \xi)$. Posso quindi, con la procedura iterativa sopra presentata, trovare una funzione soluzione dell'equazione procedendo per $\delta_0$ fino a coprire tutto l'intervallo $[\tau_1, \tau_2]$. $\square$

Sebbene la condizione di sottolinearità (o coercività) possa sembrare limitativa, in realtà le seguenti condizioni sufficienti per la sottolinearità dimostrano il contrario:

1. $\mathbf{F}$ è limitata in $\overline{S}$.

2. $\mathbf{F}(t, \mathbf{0})$ è limitata e $\mathbf{F}$ è lipschitziana in $\overline{S}$ rispetto a $\mathbf{y}$ e uniformemente in $t$. Cosideriamo infatti
$$\|\mathbf{F}(t, \mathbf{y}) - \mathbf{F}(t, \mathbf{z})\| \leq L\|\mathbf{y} - \mathbf{z}\|$$
Ponendo $\mathbf{z} = \mathbf{0}$, otteniamo che
$$\|\mathbf{F}(t, \mathbf{y})\| - \|\mathbf{F}(t, \mathbf{0})\| \leq \|\mathbf{F}(t, \mathbf{y}) - \mathbf{F}(t, \mathbf{0})\| \leq L\|\mathbf{y}\|$$
Vale quindi che
$$\|\mathbf{F}(t, \mathbf{y})\| \leq \|\mathbf{F}(t, \mathbf{0})\| + L\|\mathbf{y}\|$$
Ponendo $\|\mathbf{F}(t, \mathbf{0})\| = k_1$ e $L = k_2$ otteniamo la condizione di sottolinearità.

3. $\mathbf{F}(t, \mathbf{0})$ è limitata, $\partial_i f_j$ sono continue e limitate in $I$. In questo modo, sfruttando il Teorema del valor medio, è possibile procedere in modo analogo a quanto detto sopra.

Consideriamo a titolo di esempio il sistema differenziale lineare del primo ordine non autonomo non omogeneo

$$\begin{cases} \dot{y_1} = a_{11}y_1 + \cdots + a_{1n}y_n + b_1(t) \\ \vdots \\ \dot{y_n} = a_{n1}y_1 + \cdots + a_{nn}y_n + b_n(t) \end{cases}$$

con $a_{ji}(t), b_j(t) \in \mathscr{C}^0([\tau_1, \tau_2])$; esso può essere riscritto, ponendo $\mathbf{y} = (y_1, \dots, y_n)^t$, $\mathbf{b} = (b_1(t), \dots, b_n(t))^t$ e $A = \begin{pmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \\ a_{n1} & \dots & a_{nn} \end{pmatrix}$, nella forma matriciale

$$\dot{\mathbf{y}} = A\mathbf{y} + \mathbf{b} = \mathbf{F}(t, \mathbf{y})$$

Consideriamo $\frac{\partial f_j}{\partial y_i} = a_{ji}(t)$; essendo per ipotesi una funzione continua, $\mathbf{F} \in \mathscr{C}^1([\tau_1, \tau_2])$. Inoltre, poiché le funzioni $a_{ji}(t)$ sono continute in $[\tau_1, \tau_2]$ è possibile definire $L_{ji} := \max_{t \in [\tau_1, \tau_2]} |a_{ji}(t)|$ e $L_0 := \max_{j,i=1,\dots,n} |L_{ji}|$. Vale quindi che $\|\frac{\partial f_j}{\partial y_i}(t, \mathbf{y})\| \leq L_0 \; \forall (t, \mathbf{y}) \in \overline{S}$. Abbiamo quindi verificato la sottolinearità di $\mathbf{F}$ in $[\tau_1, \tau_2]$; è quindi possibile applicare il Teorema precedente, che ci assicura l'esistenza di una soluzione del sistema lineare di equazioni differenziali sopra riportato.
Esaminiamo ora le condizioni per cui si può prolungare la soluzione di un Problema di Cauchy sino alla frontiera del dominio di definizione di $\mathbf{F}$.

**Teorema 74.** *Sia $\mathbf{F}\colon A \subseteq \mathbb{R}^{n+1} \to \mathbb{R}^n$; supponiamo soddisfatte le condizioni del Teorema di esistenza e unicità in piccolo della soluzione del Problema di Cauchy*

$$\begin{cases} \dot{\mathbf{y}} = \mathbf{F}(t, \mathbf{y}) \\ \mathbf{y}(\tau) = \xi \end{cases}$$

*Sia $A_0 \subseteq A$ un compatto tale che $(\tau, \xi) \in A_0$ e $J_l = (T_{\min}, T_{\max})$ l'intervallo massimale di esistenza della soluzione $\phi$. Allora il grafico di $\phi$ esce definitivamente da $A_0$.*

*Dimostrazione.* Poiché $A$ è un aperto di $\mathbb{R}^{n+1}$ e $A_0 \subset A$ è un compatto, $\forall (t, \mathbf{y}) \in A_0\ d((t, \mathbf{y}), \partial A) > 0$. In particolare, definiamo $d = \inf_{(t,\mathbf{y}) \in A_0} d((t, y), \partial A)$. Definiamo l'insieme $K := \{(t, \mathbf{y}) \in A : d((t, \mathbf{y}), A_0) \leq \frac{d}{2}\}$, generalmente detto *collarino di raggio* $\frac{d}{2}$. Poiché la distanza è una funzione continua e l'immagine di $K$ è un insieme chiuso e limitato in $\mathbb{R}$, $K$ è un compatto. Vale che $A_0 \subseteq K \subseteq A$. Poniamo ora $a^2 + b^2 \leq \frac{d^2}{4}$, in modo che $\Gamma := \{(t, \mathbf{y}) \in A : |t - \tau| \leq a \wedge \|\mathbf{y} - \xi\| \leq b\} \subset K$ per ogni $(\tau, \xi) \in A_0$. Vogliamo ora determinare l'intervallo in cui esiste... Ricordando la definizione della semilarghezza dell'intervallo, $\delta = \min\{a, \frac{b}{M}\}$, consideriamo la determinazione di $M$: vale infatti che $M = \max_\Gamma \|\mathbf{F}(t, \mathbf{y}\| \leq \max_K \|\mathbf{F}(t, \mathbf{y})\|$. In particolare, poniamo $M$ uguale all'ultimo termine. Sulla base delle determinazioni fatte sino ad ora, possiamo dire che $\delta$ è indipendente dalla scelta di $(\tau, \xi)$. Quindi, partendo da un qualsiasi punto $(\tau, \xi) \in A_0$, siamo in grado di costruire un intervallo $I_\delta = [\tau - \delta, \tau + \delta]$ e, estendolo a partire da uno degli estremi, di raggiungere ed uscire da $A_0$ in un numero finito di passi. La traiettoria della soluzione massimale tende quindi al bordo dell'aperto di definizione (l'intervallo $J_l$). $\square$

**Definizione 73.** Sia data una funzione $g\colon A \subseteq \mathbb{R}^n \to \mathbb{R}$ tale che $g(x_1, \ldots, x_n) = 0$. Chiameremo $f\colon B \subseteq \mathbb{R}^{n-1} \to \mathbb{R}$ *funzione implicita* di $g$ una funzione tale che $\mathrm{Graf}(f) = \{(\mathbf{x}, y) \in \mathbb{R}^n | y = f(\mathbf{x})\}$ sia un sottoinsieme di $A$ e $g(x_1, x_2, \ldots, x_{n-1}, f(\mathbf{x})) = 0\ \forall \mathbf{x} \in B$.

**Teorema 75** (del Dini). *Sia $g\colon A \subseteq \mathbb{R}^2 \to \mathbb{R}$, e supponiamo che*

1. *$\partial_y g, g \in \mathscr{C}^0(A)$;*
2. *Nel punto $(x_0, y_0) \in A$ $g(x_0, y_0) = 0$ e $\partial_y g(x_0, y_0) \neq 0$.*

*Allora esistono un intorno $U(x_0)$ e un'unica funzione $f\colon U(x_0) \to \mathbb{R}$ continua in $U(x_0)$ tale che $y_0 = f(x_0)$ e $g(x, f(x)) = 0\ \forall x \in U(x_0)$. Se inoltre $\partial_x g$ è continua in $A$ (quindi $g \in \mathscr{C}^1(A)$) allora $f \in \mathscr{C}^1(U(x_0))$ e vale che*

$$f'(x) = -\frac{\partial_x g(x, f(x))}{\partial_y g(x, f(x))}\ \forall x \in U(x_0)$$

*Dimostrazione.* Supponiamo, senza perdere in generalità, che $\partial_y g(x_0, y_0) > 0$. Allora per il teorema della permanenza del segno esiste un rettangolo $W = [x_0 - a, x_0 + a] \times [y_0 - b, y_0 + b]$ in cui $\partial_y g$ è positiva. Fissiamo $x_0$, e definiamo una funzione $f\colon [y_0 - b, y_0 + b] \to \mathbb{R}$ che, dato un elemento $y \in [y_0 - b, y_0 + b]$ lo manda in $g(x_0, y)$. Dato che $f(x_0, y_0) = 0$ e poiché $g$ è monotona, vale che $g(x_0, y_0 - b) < 0 \wedge g(x_0, y_0 + b) > 0$. Considero ora le mappe $x \mapsto g(x, y_0 - b)$ e $x \mapsto g(x, y_0 + b)$: sono funzioni continue in $A$. Per il teorema della permanenza del segno, esiste un $\delta > 0$ tale che per ogni $x \in [x_0 - \delta, x_0 + \delta]$ $g(x, y_0 - b) < 0 \wedge g(x, y_0 + b) > 0$. Fissato un $\bar{x} \in (x_0 - \delta, x_0 + \delta)$, consideriamo la mappa che, dato un $y \in [y_0 - b, y_0 + b]$, manda $y$ in $g(\bar{x}, y)$. La funzione è continua in $[y_0 - b, y_0 + b]$ e sappiamo che $g(\bar{x}, y_0 - b) < 0 \wedge g(\bar{x}, y_0 + b) > 0$. Applicando il Teorema di Bolzano, sappiamo che esiste un punto $\bar{y} \in [y_0 - b, y_0 + b]$ tale che $g(\bar{x}, \bar{y}) = 0$. Poiché $g$ è monotona in $[y_0 - b, y_0 + b]$, tale punto è unico. Abbiamo quindi individuato una relazione univoca $f\colon [x_0 - \delta, x_0 + \delta] \to [y_0 - b, y_0 + b]$ che manda $\bar{x} \in [x_0 - \delta, x_0 + \delta]$ in $f(\bar{x}) = \bar{y}$. Per dimostrare la continuità di $f$, consideriamo ora $0 < \epsilon < b$ e ripetiamo il procedimento sopra riportato: troviamo una $f^*$ e un $\delta_\epsilon$ che definisce un intervallo $(x_0 - \delta_\epsilon, x_0 + \delta_\epsilon) \subseteq (x_0 - \delta, x_0 + \delta)$. Vale che $y_0 = f^*(x_0)$, $g(x_0, f^*(x_0)) = 0$ e $g(x, f^*(x)) = 0\ \forall x \in (x_0 - \delta_\epsilon, x_0 + \delta_\epsilon)$, ove $f^*(x) \in (y_0 - \epsilon, y_0 + \epsilon)$. Possiamo quindi ritrovare la definizione di continuità: $\left|f^*(x) - y_0\right| \leq \epsilon\ \forall x \in (x_0 - \delta_\epsilon, x_0 + \delta_\epsilon)$.
Dimostriamo ora la seconda parte dell'asserto. Poiché $g$ è $\mathscr{C}^1(A)$, è anche $\mathscr{C}^1(W)$. Consideriamo

ora il segmento $[(x,y),(\bar{x},\bar{y})] \subset U$: esiste un punto $(\xi,\eta) \in [(x,y),(\bar{x},\bar{y})]$ per cui si ha che $g(x,y) - g(\bar{x},\bar{y}) = \langle \nabla g(\xi,\eta),(x-\bar{x},y-\bar{y})\rangle$. Poiché $x,\bar{x} \in U$ vale che $y = f(x)$ e $\bar{y} = f(\bar{x})$. Si ha quindi che

$$g(x,f(x)) - g(\bar{x},f(\bar{x})) = \langle \nabla g(\xi,\eta),(x-\bar{x},f(x)-f(\bar{x}))\rangle = \partial_x g(\xi,\eta)(x-\bar{x}) + \partial_y g(\xi,\eta)(f(x)-f(\bar{x}))$$

Ricordando che, appartenendo i punti $(x,y)$ e $(\bar{x},\bar{y})$, essi costituiscono degli zeri di $g$, si ha che

$$\frac{f(x)-f(\bar{x})}{x-\bar{x}} = -\frac{\partial_x g(\xi,\eta)}{\partial_y g(\xi,\eta)}$$

Passando al limite

$$\lim_{x\to\bar{x}} \frac{f(x)-f(\bar{x})}{x-\bar{x}} = -\frac{\partial_x g(\bar{x},f(\bar{x}))}{\partial_y g(\bar{x},f(\bar{x}))}$$

L'asserto è dimostrato. □

Consideriamo ora la funzione $g(x,y) = x^2 - y^2 - 1 = 0$. Supponiamo che la funzione implicita sia una funzione $f(x)$: verifichiamo le ipotesi del Teorema del Dini. Vale che $\partial_y g(x,y) = -2y$, che è diversa da 0 per ogni $y \in \mathbb{R} \setminus 0$. In tutti questi punti posso definire la funzione implicita $f(x)\colon U(x_0) \to \mathbb{R}\ x \mapsto \sqrt{x^2-1}$ (oppure $x \mapsto -\sqrt{x^2-1}$, a seconda del segno di $y_0$). In 0 tuttavia tale funzione non è definita: tale problema può essere imputato alla non unicità della funzione, in quanto in un qualsiasi intorno dei punti $(\pm 1, 0)$ esistono due possibili funzioni implicite. Qualora volessimo trovare una funzione $f(y)$ invece il problema si pone in $x = 0$, in quanto $\partial_x g(x,y) = 2x$. Sull'onda del secondo asserto del Teorema di Dini, possiamo affermare che, se la funzione $g$ è $\mathscr{C}^2(A)$, allora $f \in \mathscr{C}^2(U)$, e vale che (derivando il risultato precedente $\langle \nabla g(x,f(x)), \phi'(x)\rangle = 0\ \forall x \in U(x_0)$ e applicando il Teorema di Schwarz)

$$\partial^2_{xx} g(x,f(x)) + 2\partial^2_{yx} g(x,f(x)) f'(x) + \partial^2_{yy} g(x,f(x)) f'^2(x) + \partial_y g(x,f(x)) f''(x) = 0$$
$$f''(x) = -\frac{\partial_{xx} g(x,f(x)) + 2\partial_{yx} g(x,f(x)) f'(x) + \partial_{yy} g(x,f(x)) f'^2(x)}{\partial_y g(x,f(x))}$$

È quindi possibile utilizzare le informazioni ottenute in questo modo per conoscere la forma approssimata della funzione $f$ in un intorno del punto $(x_0)$ attraverso uno sviluppo di Taylor centrato in $x_0$; logicamente, qualora la funzione fosse $\mathscr{C}^\infty$ si avrebbe accesso a tutti i coefficienti della serie e si avrebbe conoscenza completa della funzione.
Consideriamo ad esempio la funzione $g(x,y) = 2xe^y + y + 1 = 0$, e calcoliamo la funzione implicita $f(x)\colon U(0) \to \mathbb{R}$. Innanzitutto verifichiamo che il Dini sia applicabile. La funzione è $\mathscr{C}^\infty(\mathbb{R})$ in quanto composizione e combinazione lineare di funzioni $\mathscr{C}^\infty$, e $\partial_y g(x_0,y_0) = 2xe^y + 1 = 1 \neq 0$. Inoltre $g(0,-1) = 0$. Applichiamo quindi il Teorema del Dini: abbiamo che

$$f(x) = -1 - \frac{2}{e}x + \frac{4}{e^2}x^2 + o(x^2)$$

**Teorema 76** (generalizzazione del Dini)**.** *Sia $A \subseteq \mathbb{R}^{n+1}$ e sia $g\colon A \to \mathbb{R}$. Detto $X_0 = (x_1,\ldots,x_n)$, se valgono le ipotesi*

1. $\partial_y g, \in \mathscr{C}^0(A)$;
2. $(X_0,y_0) \in A : g(X_0,y_0) = 0 \wedge \partial_y g(X_0,y_0) \neq 0$

*allora esistono un intorno $U(X_0) \subseteq \mathbb{R}^n$ e un'unica funzione $f\colon U \to \mathbb{R}$ tale che $f \in \mathscr{C}^0(U)$, $f(X_0) = y_0$ e $g(X,f(X)) = 0\ \forall X \in U(X_0)$.*
*Se inoltre $\partial_{x_i} g \in \mathscr{C}^0(A)$ $i = 1,\ldots,n$ allora $f \in \mathscr{C}^1(U)$ e vale che*

$$\partial_{x_i} f = -\frac{\partial_{x_i} g(X,f(X))}{\partial_y g(X,f(X)}\ \forall X \in U(X_0)$$

Anche in questo caso è applicabile il risultato precedente; in particolare si ha che

$$\frac{\partial^2 f}{\partial x_k \partial x_i} = -\,\partial^2_{x_k x_i} g(X, f(X)) + \partial^2_{y x_i} g(X, f(X)) \partial_{x_k} f(X) + \partial^2_{x_k y} g(X, f(X)) \partial_{x_i} f(X) + \\ \partial^2_{yy} g(X, f(X)) \partial_{x_k} f(X) \partial_{x_i} f(X))(\partial_y g(X, f(X)))^{-1}$$

Consideriamo a titolo di esempio la funzione $g(x, y, z) = z^3 - (x + 2y)z - 2 = 0$, e calcoliamo la funzione $f(x, y)$ in un intorno $U(x_0, y_0)$ con $(x_0, y_0) = (-1, 0)$. In questo caso, affinché la condizione di azzeramento della funzione $g$ sia rispettata, deve valere che $z = 1$. Verifichiamo ora che $\partial_z g(x_0, y_0, z_0) = 4 \neq 0$. Poiché $g$ è un polinomio, è una funzione $\mathscr{C}^\infty(\mathbb{R})$; è pertanto possibile applicare il Teorema del Dini. Abbiamo che

$$f(x, y) = 1 + \frac{1}{4}(x + 1) + \frac{1}{2}y + \frac{1}{64}(x + 1)^2 + \frac{1}{16}(x + 1)y + \frac{1}{16}y^2 + o(\|(x, y)\|^2)$$

Enunciamo ora un Teorema che fornisce delle condizioni per cui localmente esiste una funzione inversa che abbia la stessa regolarità di $\mathbf{f}$.

**Definizione 74.** Siano $A, B \subseteq \mathbb{R}^n$ due aperti e sia $\mathbf{f}\colon A \to B$. Diremo che $\mathbf{f}$ è un *diffeomorfismo* di classe $\mathscr{C}^k$ se

1. $\mathbf{f}$ è biiettiva tra $A$ e $B$;
2. $\mathbf{f}$ è di classe $\mathscr{C}^k(A)$;
3. $\mathbf{f}^{-1}\colon B \to A$ è di classe $\mathscr{C}^k(B)$.

Supponiamo che sia data una funzione $\mathbf{f}\colon A \to \mathbb{R}^n$. Diremo che $\mathbf{f}$ è un diffeomorfismo locale in di classe $\mathscr{C}^k$ in $\mathbf{x_0} \in A$ se

1. esiste un intorno $U(\mathbf{x_0})$ tale che $\mathbf{f}(U)$ sia aperto in $\mathbb{R}^n$;
2. $\mathbf{f}|_U$ è biiettiva di classe $\mathscr{C}^k$;
3. $\mathbf{f}^{-1}\colon \mathbf{f}(U) \to U$ è di classe $\mathscr{C}^k$

**Teorema 77** (Dell'inversione locale)**.**