Analisi Matematica III

Appunti delle lezioni tenute dal Prof. A. Fonda

Università di Trieste, CdL Matematica, a.a. 2009/2010

1 La derivata direzionale

In questa sezione, E sarà un sottoinsieme aperto di $\mathbb{R}^N$, $\boldsymbol{x}_0$ un punto di E e $f : E \to \mathbb{R}$ una funzione. Vogliamo estendere il concetto di derivata già introdotto nel caso $N = 1$. Iniziamo con il fissare una "direzione", ossia un vettore $\boldsymbol{v} \in \mathbb{R}^N$ tale che $\|\boldsymbol{v}\| = 1$ (detto anche "versore"). Chiamiamo, se esiste, "derivata direzionale" di f in $\boldsymbol{x}_0$ nella direzione $\boldsymbol{v}$ il seguente limite

$$\lim_{t\to 0} \frac{f(\boldsymbol{x}_0 + t\boldsymbol{v}) - f(\boldsymbol{x}_0)}{t}\,,$$

che verrà indicato con il simbolo

$$\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0)\,.$$

Se $\boldsymbol{v}$ coincide con un elemento $\boldsymbol{e}_k$ della base canonica $(\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_N)$ di $\mathbb{R}^N$, la derivata direzionale si chiamerà "derivata parziale" $k-$esima di f in $\boldsymbol{x}_0$ e si indicherà con

$$\frac{\partial f}{\partial x_k}(\boldsymbol{x}_0)\,.$$

Se $\boldsymbol{x}_0 = (x_1^0, x_2^0, \dots, x_N^0)$, si ha quindi:

$$\begin{aligned}\frac{\partial f}{\partial x_k}(\boldsymbol{x}_0) &= \lim_{t\to 0} \frac{f(\boldsymbol{x}_0 + t\boldsymbol{e}_k) - f(\boldsymbol{x}_0)}{t} \\ &= \lim_{t\to 0} \frac{f(x_1^0, x_2^0, \dots, x_k^0 + t, \dots, x_N^0) - f(x_1^0, x_2^0, \dots, x_k^0, \dots, x_N^0)}{t}\,,\end{aligned}$$

per cui si usa parlare di "derivata rispetto alla $k-$esima variabile".

Esistono delle funzioni che, pur avendo derivate direzionali in tutte le possibili direzioni, non sono continue. Ad esempio, la funzione $f : \mathbb{R}^2 \to \mathbb{R}$ definita da

$$f(x,y) = \begin{cases} \dfrac{x^4 y^2}{(x^4 + y^2)^2} & \text{se } (x,y) \neq (0,0)\,, \\[2ex] 0 & \text{se } (x,y) = (0,0)\,, \end{cases}$$

ha tutte le derivate direzionali nulle in $\boldsymbol{x}_0 = (0,0)$, ma non è continua in tale punto, come si vede considerando la restrizione alla parabola $\{(x,y) \in \mathbb{R}^2 : y = x^2\}$. Questo fatto ci porta a cercare una generalizzazione più appropriata del concetto di derivata.

2 Il differenziale di una funzione a valori scalari

Definizione. *Diremo che la funzione f è "differenziabile" in $\boldsymbol{x}_0$ se esiste una applicazione lineare $\ell : \mathbb{R}^N \to \mathbb{R}$ per cui si possa scrivere*

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \ell(\boldsymbol{x} - \boldsymbol{x}_0) + r(\boldsymbol{x})\,,$$

dove r è una funzione tale che

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{r(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0\,.$$

Se f è differenziabile in $\boldsymbol{x}_0$, l'applicazione lineare ℓ si chiama "differenziale" di f in $\boldsymbol{x}_0$ e si indica con il simbolo

$$df(\boldsymbol{x}_0)\,.$$

Teorema. *Se f è differenziabile in $\boldsymbol{x}_0$, allora f è continua in $\boldsymbol{x}_0$.*

Dimostrazione. Sappiamo che l'applicazione $\ell = df(\boldsymbol{x}_0)$, essendo lineare, è continua e $\ell(\boldsymbol{0}) = 0$. Ne segue che

$$\begin{aligned}
\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} f(\boldsymbol{x}) &= \lim_{\boldsymbol{x} \to \boldsymbol{x}_0} [f(\boldsymbol{x}_0) + \ell(\boldsymbol{x} - \boldsymbol{x}_0) + r(\boldsymbol{x})] \\
&= f(\boldsymbol{x}_0) + \ell(\boldsymbol{0}) + \lim_{\boldsymbol{x} \to \boldsymbol{x}_0} r(\boldsymbol{x}) \\
&= f(\boldsymbol{x}_0) + \lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{r(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} \lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \|\boldsymbol{x} - \boldsymbol{x}_0\| \\
&= f(\boldsymbol{x}_0)\,.
\end{aligned}$$

■

Seguendo un'abitudine consolidata per le applicazioni lineari, si usa spesso scrivere $df(\boldsymbol{x}_0)\boldsymbol{h}$ invece di $df(\boldsymbol{x}_0)(\boldsymbol{h})$.

Teorema. *Se f è differenziabile in $\boldsymbol{x}_0$, allora esistono tutte le derivate direzionali di f in $\boldsymbol{x}_0$: per ogni direzione $\boldsymbol{v} \in \mathbb{R}^N$ si ha*

$$\frac{\partial f}{\partial \boldsymbol{v}}(\boldsymbol{x}_0) = df(\boldsymbol{x}_0)\boldsymbol{v}\,.$$

Dimostrazione. Usando la definizione di differenziale, abbiamo

$$\begin{aligned}\lim_{t\to 0}\frac{f(\boldsymbol{x}_0+t\boldsymbol{v})-f(\boldsymbol{x}_0)}{t} &= \lim_{t\to 0}\frac{df(\boldsymbol{x}_0)(t\boldsymbol{v})+r(\boldsymbol{x}_0+t\boldsymbol{v})}{t}\\ &= \lim_{t\to 0}\frac{t\,df(\boldsymbol{x}_0)\boldsymbol{v}+r(\boldsymbol{x}_0+t\boldsymbol{v})}{t}\\ &= df(\boldsymbol{x}_0)\boldsymbol{v}+\lim_{t\to 0}\frac{r(\boldsymbol{x}_0+t\boldsymbol{v})}{t}\,;\end{aligned}$$

d'altra parte, essendo $\|\boldsymbol{v}\|=1$, si ha

$$\lim_{t\to 0}\left|\frac{r(\boldsymbol{x}_0+t\boldsymbol{v})}{t}\right| = \lim_{\boldsymbol{x}\to\boldsymbol{x}_0}\frac{|r(\boldsymbol{x})|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|}=0\,,$$

da cui la tesi. ∎

In particolare, se $\boldsymbol{v}$ coincide con un elemento $\boldsymbol{e}_k$ della base canonica $(\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_N)$, si ha:

$$\frac{\partial f}{\partial x_k}(\boldsymbol{x}_0) = df(\boldsymbol{x}_0)\boldsymbol{e}_k\,.$$

Scrivendo il vettore $\boldsymbol{h}\in\mathbb{R}^N$ come $\boldsymbol{h}=h_1\boldsymbol{e}_1+h_2\boldsymbol{e}_2+\ldots+h_N\boldsymbol{e}_N$, abbiamo

$$\begin{aligned}df(\boldsymbol{x}_0)\boldsymbol{h} &= h_1 df(\boldsymbol{x}_0)\boldsymbol{e}_1+h_2 df(\boldsymbol{x}_0)\boldsymbol{e}_2+\ldots+h_N df(\boldsymbol{x}_0)\boldsymbol{e}_N\\ &= h_1\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)+h_2\frac{\partial f}{\partial x_2}(\boldsymbol{x}_0)+\ldots+h_N\frac{\partial f}{\partial x_N}(\boldsymbol{x}_0)\,,\end{aligned}$$

ossia

$$df(\boldsymbol{x}_0)\boldsymbol{h}=\sum_{k=1}^N\frac{\partial f}{\partial x_k}(\boldsymbol{x}_0)h_k\,.$$

Introducendo il vettore "gradiente" di f in $\boldsymbol{x}_0$

$$\nabla f(\boldsymbol{x}_0)=\left(\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0),\frac{\partial f}{\partial x_2}(\boldsymbol{x}_0),\ldots,\frac{\partial f}{\partial x_N}(\boldsymbol{x}_0)\right)\,,$$

si può scrivere

$$df(\boldsymbol{x}_0)\boldsymbol{h}=\nabla f(\boldsymbol{x}_0)\cdot\boldsymbol{h}\,.$$

3 Funzioni di classe C^1

Il seguente risultato è noto come "teorema del differenziale totale".

Teorema. *Se f possiede le derivate parziali in un intorno di $\boldsymbol{x}_0$ ed esse sono continue in $\boldsymbol{x}_0$, allora f è differenziabile in $\boldsymbol{x}_0$.*

Dimostrazione. Supporremo per semplicità di notazioni $N=2$. Definiamo l'applicazione lineare $\ell:\mathbb{R}^2\to\mathbb{R}$ che ad ogni vettore $\boldsymbol{h}=(h_1,h_2)$ associa

$$\ell(\boldsymbol{h})=\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)h_1+\frac{\partial f}{\partial x_2}(\boldsymbol{x}_0)h_2\,.$$

Vedremo che ℓ è proprio il differenziale di f in $\boldsymbol{x}_0$. Intanto, è lineare, come si vede immediatamente. Inoltre, scrivendo $\boldsymbol{x}_0 = (x_1^0, x_2^0)$ e $\boldsymbol{x} = (x_1, x_2)$, per il teorema di Lagrange si ha

$$\begin{aligned} f(\boldsymbol{x}) - f(\boldsymbol{x}_0) &= (f(x_1, x_2) - f(x_1^0, x_2)) + (f(x_1^0, x_2) - f(x_1^0, x_2^0)) \\ &= \frac{\partial f}{\partial x_1}(\xi_1, x_2)(x_1 - x_1^0) + \frac{\partial f}{\partial x_2}(x_1^0, \xi_2)(x_2 - x_2^0)\,, \end{aligned}$$

per un certo $\xi_1 \in]x_1^0, x_1[$ e un certo $\xi_2 \in]x_2^0, x_2[$. Quindi,

$$\begin{aligned} r(\boldsymbol{x}) &= f(\boldsymbol{x}) - f(\boldsymbol{x}_0) - \ell(\boldsymbol{x} - \boldsymbol{x}_0) \\ &= \left[\frac{\partial f}{\partial x_1}(\xi_1, x_2) - \frac{\partial f}{\partial x_1}(x_1^0, x_2^0)\right](x_1 - x_1^0) + \\ &\quad + \left[\frac{\partial f}{\partial x_2}(x_1^0, \xi_2) - \frac{\partial f}{\partial x_2}(x_1^0, x_2^0)\right](x_2 - x_2^0)\,, \end{aligned}$$

ed essendo $|x_1 - x_1^0| \le \|\boldsymbol{x} - \boldsymbol{x}_0\|$ e $|x_2 - x_2^0| \le \|\boldsymbol{x} - \boldsymbol{x}_0\|$,

$$\frac{|r(\boldsymbol{x})|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} \le \left|\frac{\partial f}{\partial x_1}(\xi_1, x_2) - \frac{\partial f}{\partial x_1}(x_1^0, x_2^0)\right| + \left|\frac{\partial f}{\partial x_2}(x_1^0, \xi_2) - \frac{\partial f}{\partial x_2}(x_1^0, x_2^0)\right|\,.$$

Facendo tendere $\boldsymbol{x}$ a $\boldsymbol{x}_0$, si ha che $(\xi_1, x_2) \to (x_1^0, x_2^0)$ e $(x_1^0, \xi_2) \to (x_1^0, x_2^0)$ per cui, essendo $\frac{\partial f}{\partial x_1}$ e $\frac{\partial f}{\partial x_2}$ continue in $\boldsymbol{x}_0 = (x_1^0, x_2^0)$, si ha

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{|r(\boldsymbol{x})|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0\,,$$

da cui la tesi. ■

Diremo che la funzione f è di classe $\mathcal{C}^1$ su E se f possiede le derivate parziali ed esse sono continue su tutto E. Dal teorema precedente segue che una funzione di classe $\mathcal{C}^1$ è "differenziabile su E", ossia in ogni punto di E.

4 Derivate parziali successive

Supponiamo, per semplicità, $N = 2$. Consideriamo E, un insieme aperto di $\mathbb{R}^2$ e una funzione $f : E \to \mathbb{R}$ che abbia le derivate parziali $\frac{\partial f}{\partial x_1}$, $\frac{\partial f}{\partial x_2}$ in tutti i punti di E. Se esse posseggono a loro volta derivate parziali in un punto $\boldsymbol{x}_0$, queste si dicono "derivate parziali seconde" della f in $\boldsymbol{x}_0$ e si denotano con i simboli

$$\frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_1}\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)\,, \quad \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_2}\frac{\partial f}{\partial x_1}(\boldsymbol{x}_0)\,,$$
$$\frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_1}\frac{\partial f}{\partial x_2}(\boldsymbol{x}_0)\,, \quad \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}_0) = \frac{\partial}{\partial x_2}\frac{\partial f}{\partial x_2}(\boldsymbol{x}_0)\,.$$

Teorema (di Schwarz). *Se esistono le derivate parziali seconde* $\frac{\partial^2 f}{\partial x_2 \partial x_1}$, $\frac{\partial^2 f}{\partial x_1 \partial x_2}$ *in un intorno di* $\boldsymbol{x}_0$ *ed esse sono continue in* $\boldsymbol{x}_0$*, allora*

$$\frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}_0) = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}_0)\,.$$

Dimostrazione. [1] Sia $\rho > 0$ tale che $B(\boldsymbol{x}_0, \rho) \subseteq E$. Scriviamo $\boldsymbol{x}_0 = (x_1^0, x_2^0)$ e prendiamo un $\boldsymbol{x} = (x_1, x_2) \in B(\boldsymbol{x}_0, \rho)$ tale che $x_1 \neq x_1^0$ e $x_2 \neq x_2^0$. Possiamo allora definire

$$g(x_1, x_2) = \frac{f(x_1, x_2) - f(x_1, x_2^0)}{x_2 - x_2^0}\,, \quad h(x_1, x_2) = \frac{f(x_1, x_2) - f(x_1^0, x_2)}{x_1 - x_1^0}\,.$$

Si verifica che vale l'uguaglianza

$$\frac{g(x_1, x_2) - g(x_1^0, x_2)}{x_1 - x_1^0} = \frac{h(x_1, x_2) - h(x_1, x_2^0)}{x_2 - x_2^0}\,.$$

Per il teorema di Lagrange, esiste un $\xi_1 \in]x_1^0, x_1[$ tale che

$$\frac{g(x_1, x_2) - g(x_1^0, x_2)}{x_1 - x_1^0} = \frac{\partial g}{\partial x_1}(\xi_1, x_2) = \frac{\frac{\partial f}{\partial x_1}(\xi_1, x_2) - \frac{\partial f}{\partial x_1}(\xi_1, x_2^0)}{x_2 - x_2^0}\,,$$

ed esiste un $\xi_2 \in]x_2^0, x_2[$ tale che

$$\frac{h(x_1, x_2) - h(x_1, x_2^0)}{x_2 - x_2^0} = \frac{\partial h}{\partial x_2}(x_1, \xi_2) = \frac{\frac{\partial f}{\partial x_2}(x_1, \xi_2) - \frac{\partial f}{\partial x_2}(x_1^0, \xi_2)}{x_1 - x_1^0}\,.$$

Di nuovo per il teorema di Lagrange, esiste un $\eta_2 \in]x_2^0, x_2[$ tale che

$$\frac{\frac{\partial f}{\partial x_1}(\xi_1, x_2) - \frac{\partial f}{\partial x_1}(\xi_1, x_2^0)}{x_2 - x_2^0} = \frac{\partial^2 f}{\partial x_2 \partial x_1}(\xi_1, \eta_2)\,,$$

ed esiste un $\eta_1 \in]x_1^0, x_1[$ tale che

$$\frac{\frac{\partial f}{\partial x_2}(x_1, \xi_2) - \frac{\partial f}{\partial x_2}(x_1^0, \xi_2)}{x_1 - x_1^0} = \frac{\partial^2 f}{\partial x_1 \partial x_2}(\eta_1, \xi_2)\,.$$

Quindi,

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

Facendo tendere $\boldsymbol{x} = (x_1, x_2)$ a $\boldsymbol{x}_0 = (x_1^0, x_2^0)$, si ha che sia (ξ_1, η_2) che (η_1, ξ_2) tendono a $\boldsymbol{x}_0$, e per la continuità delle derivate seconde miste si ha la tesi. ■

Diremo che la funzione f è di classe $\mathcal{C}^2$ su E se f possiede tutte le derivate parziali seconde ed esse sono continue su tutto E. Dal teorema precedente segue che se una funzione di classe $\mathcal{C}^2$, le derivate parziali "miste" sono uguali.

[1] Dimostrazione solo accennata a lezione.

È utile definire la "matrice hessiana" di f nel punto $\boldsymbol{x}_0$:

$$Hf(\boldsymbol{x}_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(\boldsymbol{x}_0) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(\boldsymbol{x}_0) \\ \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(\boldsymbol{x}_0) & \frac{\partial^2 f}{\partial x_2^2}(\boldsymbol{x}_0) \end{pmatrix};$$

se f è di classe $\mathcal{C}^2$, si tratta di una matrice simmetrica.

Quanto sopra si può estendere senza difficoltà alle funzioni di N variabili, con N qualunque. Se f è di classe $\mathcal{C}^2$, la matrice hessiana risulta allora una matrice simmetrica del tipo $N \times N$.

Procedendo per induzione, si possono definire le derivate parziali $n-$esime di una funzione. Si dice che la funzione f è di classe $\mathcal{C}^n$ su E se f possiede tutte le derivate parziali $n-$esime ed esse sono continue su tutto E.

5 La formula di Taylor

Supponiamo ora che $f : E \to \mathbb{R}$ sia una funzione di classe $\mathcal{C}^{n+1}$, per un certo $n \geq 1$.

Consideriamo come sopra, per semplicità, il caso $N = 2$. Introduciamo le seguenti notazioni:

$$D_{x_1} = \frac{\partial}{\partial x_1}, \quad D_{x_2} = \frac{\partial}{\partial x_2},$$

$$D_{x_1}^2 = \frac{\partial^2}{\partial x_1^2}, \quad D_{x_1} D_{x_2} = \frac{\partial^2}{\partial x_1 \partial x_2}, \quad D_{x_2}^2 = \frac{\partial^2}{\partial x_2^2},$$

e così via, per le derivate parziali successive. Si noti che, per un vettore $\boldsymbol{h} = (h_1, h_2) \in \mathbb{R}^2$, si ha

$$df(\boldsymbol{x}_0)\boldsymbol{h} = h_1 D_{x_1} f(\boldsymbol{x}_0) + h_2 D_{x_2} f(\boldsymbol{x}_0),$$

che risulterà conveniente scrivere

$$df(\boldsymbol{x}_0)\boldsymbol{h} = [h_1 D_{x_1} + h_2 D_{x_2}] f(\boldsymbol{x}_0).$$

In questo modo, possiamo pensare che f viene trasformata dall'operatore $[h_1 D_{x_1} + h_2 D_{x_2}]$ nella nuova funzione $[h_1 D_{x_1} + h_2 D_{x_2}] f = h_1 D_{x_1} f + h_2 D_{x_2} f$.

Dati due punti $\boldsymbol{x}_0$ e $\boldsymbol{x}$ in $\mathbb{R}^N$, si definisce il "segmento" che li congiunge:

$$[\boldsymbol{x}_0, \boldsymbol{x}] = \{\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0) : t \in [0, 1]\};$$

analogamente, scriveremo

$$]\boldsymbol{x}_0, \boldsymbol{x}[= \{\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0) : t \in]0, 1[\}.$$

Supponiamo ora che $[\boldsymbol{x}_0, \boldsymbol{x}]$ sia un segmento contenuto in E e consideriamo la funzione $\phi : [0, 1] \to \mathbb{R}$ definita da

$$\phi(t) = f(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0)).$$

Dimostriamo che ϕ è derivabile $n+1$ volte su $[0,1]$. Per $t \in [0,1]$, essendo f differenziabile in $\boldsymbol{u}_0 = \boldsymbol{x}_0 + t(\boldsymbol{x}-\boldsymbol{x}_0)$, si ha

$$f(\boldsymbol{u}) = f(\boldsymbol{u}_0) + df(\boldsymbol{u}_0)(\boldsymbol{u}-\boldsymbol{u}_0) + r(\boldsymbol{u})\,,$$

con

$$\lim_{\boldsymbol{u}\to\boldsymbol{u}_0} \frac{r(\boldsymbol{u})}{\|\boldsymbol{u}-\boldsymbol{u}_0\|} = 0\,.$$

Quindi,

$$\begin{aligned}\lim_{s\to t}\frac{\phi(s)-\phi(t)}{s-t} &= \lim_{s\to t}\frac{f(\boldsymbol{x}_0+s(\boldsymbol{x}-\boldsymbol{x}_0)) - f(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))}{s-t}\\ &= \lim_{s\to t}\frac{df(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))((s-t)(\boldsymbol{x}-\boldsymbol{x}_0)) + r(\boldsymbol{x}_0+s(\boldsymbol{x}-\boldsymbol{x}_0))}{s-t}\\ &= df(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))(\boldsymbol{x}-\boldsymbol{x}_0) + \lim_{s\to t}\frac{r(\boldsymbol{x}_0+s(\boldsymbol{x}-\boldsymbol{x}_0))}{s-t}\,,\end{aligned}$$

ed essendo

$$\lim_{s\to t}\left|\frac{r(\boldsymbol{x}_0+s(\boldsymbol{x}-\boldsymbol{x}_0))}{s-t}\right| = \lim_{\boldsymbol{u}\to\boldsymbol{u}_0}\frac{|r(\boldsymbol{u})|}{\|\boldsymbol{u}-\boldsymbol{u}_0\|}\|\boldsymbol{x}-\boldsymbol{x}_0\| = 0\,,$$

si ha

$$\phi'(t) = \lim_{s\to t}\frac{\phi(s)-\phi(t)}{s-t} = df(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))(\boldsymbol{x}-\boldsymbol{x}_0)\,.$$

Con le nuove notazioni, ponendo $\boldsymbol{x}-\boldsymbol{x}_0 = \boldsymbol{h} = (h_1,h_2)$, abbiamo

$$\phi'(t) = [h_1 D_{x_1} + h_2 D_{x_2}]f(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0)) = g(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))\,,$$

dove g è la nuova funzione $[h_1 D_{x_1} + h_2 D_{x_2}]f$. Possiamo allora iterare il procedimento, e calcolare la derivata seconda di ϕ:

$$\begin{aligned}\phi''(t) &= [h_1 D_{x_1} + h_2 D_{x_2}]g(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))\\ &= [h_1 D_{x_1} + h_2 D_{x_2}][h_1 D_{x_1} + h_2 D_{x_2}]f(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))\,.\end{aligned}$$

Per brevità, scriveremo

$$\phi''(t) = [h_1 D_{x_1} + h_2 D_{x_2}]^2 f(\boldsymbol{x}_0+t(\boldsymbol{x}-\boldsymbol{x}_0))\,.$$

Notiamo che, usando la linearità delle derivate parziali e l'uguaglianza delle derivate miste (teorema di Schwarz), si ha

$$\begin{aligned}[h_1 D_{x_1} + h_2 D_{x_2}]^2 f &= h_1^2 D_{x_1}^2 f + 2h_1h_2 D_{x_1}D_{x_2} f + h_2^2 D_{x_2}^2 f\\ &= [h_1^2 D_{x_1}^2 + 2h_1h_2 D_{x_1}D_{x_2} + h_2^2 D_{x_2}^2]f\,.\end{aligned}$$

Osserviamo che l'espressione

$$[h_1 D_{x_1} + h_2 D_{x_2}]^2 = [h_1^2 D_{x_1}^2 + 2h_1h_2 D_{x_1}D_{x_2} + h_2^2 D_{x_2}^2]$$

si ottiene formalmente come il quadrato di un binomio. Procedendo in questo modo, si può dimostrare per induzione che, per $k = 1, 2, \ldots, n+1$, la formula della derivata $k-$esima di ϕ è

$$\phi^{(k)}(t) = [h_1 D_{x_1} + h_2 D_{x_2}]^k f(\boldsymbol{x}_0 + t(\boldsymbol{x} - \boldsymbol{x}_0))\,,$$

e che, usando formalmente la formula del binomio di Newton, si ha

$$[h_1 D_{x_1} + h_2 D_{x_2}]^k = \left[\sum_{j=0}^{k} \binom{k}{j} h_1^{k-j} h_2^j D_{x_1}^{k-j} D_{x_2}^j\right]$$

(in questa formula, i simboli $D_{x_1}^0$ e $D_{x_2}^0$ vanno interpretati come l'operatore identità).

Per poter scrivere agevolmente la formula di Taylor, introduciamo la notazione

$$d^k f(\boldsymbol{x}_0)\boldsymbol{h}^k = [h_1 D_{x_1} + h_2 D_{x_2}]^k f(\boldsymbol{x}_0)\,.$$

Teorema. *Sia $f : E \to \mathbb{R}$ di classe $\mathcal{C}^{n+1}$ e $[\boldsymbol{x}_0, \boldsymbol{x}]$ un segmento contenuto in E. Allora esiste un $\boldsymbol{\xi} \in]\boldsymbol{x}_0, \boldsymbol{x}[$ tale che*

$$f(\boldsymbol{x}) = p_n(\boldsymbol{x}) + r_n(\boldsymbol{x})\,,$$

dove

$$p_n(\boldsymbol{x}) = f(\boldsymbol{x}_0) + df(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0) + \frac{1}{2!}d^2 f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^2 + \ldots + \frac{1}{n!}d^n f(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)^n$$

è il "polinomio di Taylor di grado n associato alla funzione f nel punto $\boldsymbol{x}_0$" e

$$r_n(\boldsymbol{x}) = \frac{1}{(n+1)!} d^{n+1} f(\boldsymbol{\xi})(\boldsymbol{x} - \boldsymbol{x}_0)^{n+1}$$

è il "resto di Lagrange".

<u>Dimostrazione</u>. Per la formula di Taylor applicata alla funzione ϕ, si ha

$$\phi(t) = \phi(0) + \phi'(0)t + \frac{1}{2!}\phi''(0)t^2 + \ldots + \frac{1}{n!}\phi^{(n)}(0)t^n + \frac{1}{(n+1)!}\phi^{(n+1)}(\xi)t^{n+1}\,,$$

per un certo $\xi \in]0, t[$. La formula cercata si ottiene prendendo $t = 1$ e sostituendo i valori delle derivate di ϕ trovati sopra. ∎

Il polinomio di Taylor si può anche scrivere nella forma compatta

$$p_n(\boldsymbol{x}) = \sum_{k=0}^{n} \frac{1}{k!} d^k f(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)^k\,,$$

con la convenzione che $d^0 f(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)^0$, il primo addendo della somma, sia $f(\boldsymbol{x}_0)$. Si ha quindi

$$\begin{aligned} p_n(\boldsymbol{x}) &= \sum_{k=0}^{n} \frac{1}{k!}[(x_1 - x_1^0)D_{x_1} + (x_2 - x_2^0)D_{x_2}]^k f(\boldsymbol{x}_0) \\ &= \sum_{k=0}^{n} \frac{1}{k!} \left(\sum_{j=0}^{k} \binom{k}{j} \frac{\partial^k f}{\partial^{k-j} x_1 \partial^j x_2}(\boldsymbol{x}_0)\ (x_1 - x_1^0)^{k-j}(x_2 - x_2^0)^j \right)\,. \end{aligned}$$

Può essere utile la seguente espressione per il polinomio di secondo grado:

$$p_2(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \nabla f(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) + \frac{1}{2}\Big(Hf(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)\Big) \cdot (\boldsymbol{x} - \boldsymbol{x}_0)\,.$$

Il teorema sopra dimostrato resta valido per qualsiasi dimensione N, pur di interpretare correttamente le notazioni: ad esempio, per un vettore $\boldsymbol{h} = (h_1, h_2, \ldots, h_N)$, si dovrà leggere

$$d^k f(\boldsymbol{x}_0)\boldsymbol{h}^k = [h_1 D_{x_1} + h_2 D_{x_2} + \ldots + h_N D_{x_N}]^k f(\boldsymbol{x}_0)\,.$$

6 Il differenziale di una funzione a valori vettoriali

Sia E un sottoinsieme aperto di $\mathbb{R}^N$, $\boldsymbol{x}_0$ un punto di E e $f : E \to \mathbb{R}^M$ una funzione.

Definizione. *Diremo che la funzione f è "differenziabile" in $\boldsymbol{x}_0$ se esiste una applicazione lineare $\ell : \mathbb{R}^N \to \mathbb{R}^M$ per cui si possa scrivere*

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \ell(\boldsymbol{x} - \boldsymbol{x}_0) + r(\boldsymbol{x})\,,$$

dove r è una funzione tale che

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{r(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = \boldsymbol{0}\,.$$

Se f è differenziabile in $\boldsymbol{x}_0$, l'applicazione lineare ℓ si chiama "differenziale" di f in $\boldsymbol{x}_0$ e si indica con il simbolo

$$df(\boldsymbol{x}_0)\,.$$

Siano $f_1, f_2, \ldots, f_M$ le componenti di f, per cui

$$f(\boldsymbol{x}) = (f_1(\boldsymbol{x}), f_2(\boldsymbol{x}), \ldots, f_M(\boldsymbol{x}))\,.$$

Teorema. *La funzione f è differenziabile in $\boldsymbol{x}_0$ se e solo se lo sono tutte le sue componenti. In tal caso, per ogni vettore $\boldsymbol{h} \in \mathbb{R}^N$ si ha*

$$df(\boldsymbol{x}_0)\boldsymbol{h} = (df_1(\boldsymbol{x}_0)\boldsymbol{h}, df_2(\boldsymbol{x}_0)\boldsymbol{h}, \ldots, df_M(\boldsymbol{x}_0)\boldsymbol{h})\,.$$

Dimostrazione. Considerando le componenti nell'equazione

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + \ell(\boldsymbol{x} - \boldsymbol{x}_0) + r(\boldsymbol{x})\,,$$

possiamo scrivere

$$f_j(\boldsymbol{x}) = f_j(\boldsymbol{x}_0) + \ell_j(\boldsymbol{x} - \boldsymbol{x}_0) + r_j(\boldsymbol{x})\,,$$

con $j = 1, 2, \ldots, M$, e sappiamo che

$$\lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{r(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = \boldsymbol{0} \quad\Longleftrightarrow\quad \lim_{\boldsymbol{x}\to\boldsymbol{x}_0} \frac{r_j(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0 \;\text{ per ogni } j = 1, 2, \ldots, M\,,$$

da cui la tesi. ■

Il teorema precedente permette di ricondurre lo studio del differenziale di una funzione a valori vettoriali a quello delle sue componenti, che sono funzioni a valori scalari.

È utile considerare la matrice associata all'applicazione lineare $\ell = df(\boldsymbol{x}_0)$, data da

$$\begin{pmatrix} \ell_1(\boldsymbol{e}_1) & \ell_1(\boldsymbol{e}_2) & \dots & \ell_1(\boldsymbol{e}_N) \\ \ell_2(\boldsymbol{e}_1) & \ell_2(\boldsymbol{e}_2) & \dots & \ell_2(\boldsymbol{e}_N) \\ \vdots & \vdots & & \vdots \\ \ell_M(\boldsymbol{e}_1) & \ell_M(\boldsymbol{e}_2) & \dots & \ell_M(\boldsymbol{e}_N) \end{pmatrix},$$

dove $\boldsymbol{e}_1, \boldsymbol{e}_2, \dots \boldsymbol{e}_N$ sono i vettori della base canonica di $\mathbb{R}^N$. Tale matrice si chiama "matrice jacobiana" associata alla funzione f nel punto $\boldsymbol{x}_0$ e si denota con $Jf(\boldsymbol{x}_0)$. Ricordando che

$$\frac{\partial f_j}{\partial x_k}(\boldsymbol{x}_0) = df_j(\boldsymbol{x}_0)\boldsymbol{e}_k\,,$$

con $j = 1, 2, \dots, M$ e $k = 1, 2, \dots, N$, si ottiene la matrice

$$Jf(\boldsymbol{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}_0) & \frac{\partial f_1}{\partial x_2}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_1}{\partial x_N}(\boldsymbol{x}_0) \\ \frac{\partial f_2}{\partial x_1}(\boldsymbol{x}_0) & \frac{\partial f_2}{\partial x_2}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_2}{\partial x_N}(\boldsymbol{x}_0) \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_M}{\partial x_1}(\boldsymbol{x}_0) & \frac{\partial f_M}{\partial x_2}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_M}{\partial x_N}(\boldsymbol{x}_0) \end{pmatrix}.$$

Studiamo ora la differenziabilità di una funzione composta.

Teorema. *Se $f : E \to \mathbb{R}^M$ è differenziabile in $\boldsymbol{x}_0$, E' è un aperto di $\mathbb{R}^M$ contenente $f(E)$ e $g : E' \to \mathbb{R}^L$ è differenziabile in $f(\boldsymbol{x}_0)$, allora $g \circ f$ è differenziabile in $\boldsymbol{x}_0$, e si ha*

$$d(g \circ f)(\boldsymbol{x}_0) = dg(f(\boldsymbol{x}_0)) \circ df(\boldsymbol{x}_0)\,.$$

Dimostrazione. [2] Ponendo $\boldsymbol{y}_0 = f(\boldsymbol{x}_0)$, si ha

$$f(\boldsymbol{x}) = f(\boldsymbol{x}_0) + df(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + r_1(\boldsymbol{x})\,, \quad g(\boldsymbol{y}) = g(\boldsymbol{y}_0) + dg(\boldsymbol{y}_0)(\boldsymbol{y} - \boldsymbol{y}_0) + r_2(\boldsymbol{y})\,,$$

con

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{r_1(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = \boldsymbol{0}\,, \quad \lim_{\boldsymbol{y} \to \boldsymbol{y}_0} \frac{r_2(\boldsymbol{y})}{\|\boldsymbol{y} - \boldsymbol{y}_0\|} = \boldsymbol{0}\,.$$

Introduciamo la funzione $R_2 : E' \to \mathbb{R}^L$ così definita:

$$R_2(\boldsymbol{y}) = \begin{cases} \dfrac{r_2(\boldsymbol{y})}{\|\boldsymbol{y} - \boldsymbol{y}_0\|} & \text{se } \boldsymbol{y} \neq \boldsymbol{y}_0\,, \\ \boldsymbol{0} & \text{se } \boldsymbol{y} = \boldsymbol{y}_0\,. \end{cases}$$

[2] Dimostrazione non svolta a lezione.

Si noti che R_2 è continua in $\boldsymbol{y}_0$. Allora

$$\begin{aligned} g(f(\boldsymbol{x})) &= g(f(\boldsymbol{x}_0)) + dg(f(\boldsymbol{x}_0))[f(\boldsymbol{x}) - f(\boldsymbol{x}_0)] + r_2(f(\boldsymbol{x})) \\ &= g(f(\boldsymbol{x}_0)) + dg(f(\boldsymbol{x}_0))[df(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + r_1(\boldsymbol{x})] + r_2(f(\boldsymbol{x})) \\ &= g(f(\boldsymbol{x}_0)) + [dg(f(\boldsymbol{x}_0)) \circ df(\boldsymbol{x}_0)](\boldsymbol{x} - \boldsymbol{x}_0) + r_3(\boldsymbol{x}) \,, \end{aligned}$$

dove

$$\begin{aligned} r_3(\boldsymbol{x}) &= dg(f(\boldsymbol{x}_0))(r_1(\boldsymbol{x})) + r_2(f(\boldsymbol{x})) \\ &= dg(f(\boldsymbol{x}_0))(r_1(\boldsymbol{x})) + \|f(\boldsymbol{x}) - f(\boldsymbol{x}_0)\| R_2(f(\boldsymbol{x})) \\ &= dg(f(\boldsymbol{x}_0))(r_1(\boldsymbol{x})) + \|df(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0) + r_1(\boldsymbol{x})\| R_2(f(\boldsymbol{x})) \,. \end{aligned}$$

Quindi,

$$\begin{aligned} \frac{\|r_3(\boldsymbol{x})\|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} \leq & \left\| dg(f(\boldsymbol{x}_0)) \left(\frac{r_1(\boldsymbol{x})}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} \right) \right\| + \\ & + \left(\left\| df(\boldsymbol{x}_0) \left(\frac{\boldsymbol{x} - \boldsymbol{x}_0}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} \right) \right\| + \frac{\|r_1(\boldsymbol{x})\|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} \right) \|R_2(f(\boldsymbol{x}))\| \,. \end{aligned}$$

Se $\boldsymbol{x} \to \boldsymbol{x}_0$, il primo addendo tende a 0, poiché $dg(f(\boldsymbol{x}_0))$ è continua; f è continua in $\boldsymbol{x}_0$ e R_2 è continua in $\boldsymbol{y}_0 = f(\boldsymbol{x}_0)$ con $R_2(\boldsymbol{y}_0)) = \mathbf{0}$, per cui $\|R_2(f(\boldsymbol{x}))\|$ tende a 0; $df(\boldsymbol{x}_0)$, essendo continua, è limitata sull'insieme compatto $\bar{B}(\mathbf{0}, 1)$. Quindi, si ha che

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{\|r_3(\boldsymbol{x})\|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0 \,.$$

Ne segue che $g \circ f$ è differenziabile in x_0 con differenziale $dg(f(\boldsymbol{x}_0)) \circ df(\boldsymbol{x}_0)$. ■

Come noto, la matrice associata alla composizione di due applicazioni lineari è il prodotto delle due matrici corrispondenti. Dal teorema precedente abbiamo quindi la seguente formula per le matrici jacobiane:

$$J(g \circ f)(\boldsymbol{x}_0) = Jg(f(\boldsymbol{x}_0)) \cdot Jf(\boldsymbol{x}_0) \,,$$

ossia

$$\begin{pmatrix} \frac{\partial (g\circ f)_1}{\partial x_1}(\boldsymbol{x}_0) & \cdots & \frac{\partial (g\circ f)_1}{\partial x_N}(\boldsymbol{x}_0) \\ \vdots & \cdots & \vdots \\ \frac{\partial (g\circ f)_L}{\partial x_1}(\boldsymbol{x}_0) & \cdots & \frac{\partial (g\circ f)_L}{\partial x_N}(\boldsymbol{x}_0) \end{pmatrix} = \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(f(\boldsymbol{x}_0)) & \cdots & \frac{\partial g_1}{\partial y_M}(f(\boldsymbol{x}_0)) \\ \vdots & \cdots & \vdots \\ \frac{\partial g_L}{\partial y_1}(f(\boldsymbol{x}_0)) & \cdots & \frac{\partial g_L}{\partial y_M}(f(\boldsymbol{x}_0)) \end{pmatrix} \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_1}{\partial x_N}(\boldsymbol{x}_0) \\ \vdots & \cdots & \vdots \\ \frac{\partial f_M}{\partial x_1}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_M}{\partial x_N}(\boldsymbol{x}_0) \end{pmatrix} .$$

Ne segue la formula per le derivate parziali:

$$\begin{aligned} \frac{\partial (g \circ f)_i}{\partial x_k}(\boldsymbol{x}_0) &= \frac{\partial g_i}{\partial y_1}(f(\boldsymbol{x}_0)) \frac{\partial f_1}{\partial x_k}(\boldsymbol{x}_0) + \frac{\partial g_i}{\partial y_2}(f(\boldsymbol{x}_0)) \frac{\partial f_2}{\partial x_k}(\boldsymbol{x}_0) + \ldots + \frac{\partial g_i}{\partial y_M}(f(\boldsymbol{x}_0)) \frac{\partial f_M}{\partial x_k}(\boldsymbol{x}_0) \\ &= \sum_{k=1}^{M} \frac{\partial g_i}{\partial y_j}(f(\boldsymbol{x}_0)) \frac{\partial f_j}{\partial x_k}(\boldsymbol{x}_0) \,, \end{aligned}$$

dove $i = 1, 2, \ldots, L$ e $k = 1, 2, \ldots, N$.

7 Il teorema della funzione implicita

Teorema. *Siano* $\Omega \subseteq \mathbb{R} \times \mathbb{R}$ *un aperto,* $g : \Omega \to \mathbb{R}$ *una funzione di classe* C^1 *e* (x_0, y_0) *un punto di* Ω *per cui si abbia:*

$$g(x_0, y_0) = 0 \qquad \frac{\partial g}{\partial y}(x_0, y_0) \neq 0 \,.$$

Allora esistono un intorno aperto U *di* x_0*, un intorno aperto* V *di* y_0 *e una funzione* $\eta : U \to V$ *di classe* C^1 *tali che* $U \times V \subseteq \Omega$ *e, presi* $x \in U$ *e* $y \in V$*, si ha:*

$$g(x, y) = 0 \quad \Longleftrightarrow \quad y = \eta(x) \,.$$

Inoltre, la funzione η *è di classe* C^1 *e vale la formula*

$$\eta'(x) = -\frac{\frac{\partial g}{\partial x}(x, \eta(x))}{\frac{\partial g}{\partial y}(x, \eta(x))} \,.$$

La funzione η risulta definita "implicitamente" dall'equazione $g(x, y) = 0$; il suo grafico è l'insieme

$$Gr(\eta) = \{(x, y) \in U \times V : g(x, y) = 0\} \,.$$

Dimostrazione. Supponiamo ad esempio $\frac{\partial g}{\partial y}(x_0, y_0) > 0$. Per la proprietà di permanenza del segno, esiste un $\delta > 0$ tale che, se $|x - x_0| \leq \delta$ e $|y - y_0| \leq \delta$, allora $\frac{\partial g}{\partial y}(x, y) > 0$. Quindi, per ogni $x \in [x_0 - \delta, x_0 + \delta]$, la funzione $g(x, \cdot)$ è strettamente crescente su $[y_0 - \delta, y_0 + \delta]$. Essendo $g(x_0, y_0) = 0$, avremo che

$$g(x_0, y_0 - \delta) < 0 < g(x_0, y_0 + \delta) \,.$$

Per la permanenza del segno, esiste un $\delta' > 0$ tale che, se $x \in [x_0 - \delta', x_0 + \delta']$, allora

$$g(x, y_0 - \delta) < 0 < g(x, y_0 + \delta) \,.$$

Definiamo $U =]x_0 - \delta', x_0 + \delta'[$ e $V =]y_0 - \delta, y_0 + \delta[$. Quindi, per ogni $x \in U$, siccome $g(x, \cdot)$ è strettamente crescente, esiste uno ed un solo $y \in]y_0 - \delta, y_0 + \delta[$ per cui $g(x, y) = 0$; chiamo $\eta(x)$ tale y. Resta così definita una funzione $\eta : U \to V$ tale che, presi $x \in U$ e $y \in V$, si ha:

$$g(x, y) = 0 \quad \Longleftrightarrow \quad y = \eta(x) \,.$$

Per vedere che η è continua, fissiamo ora un $\bar{x} \in U$ e dimostriamo la continuità in $\bar{x}$. Preso un $x \in U$ e considerata la funzione $\gamma : [0, 1] \to \mathbb{R} \times \mathbb{R}$ definita da

$$\gamma(t) = (\bar{x} + t(x - \bar{x}), \eta(\bar{x}) + t(\eta(x) - \eta(\bar{x}))),$$

applicando il teorema di Lagrange alla funzione $g \circ \gamma$ si ha che esiste un $\xi \in]0, 1[$ per cui

$$g(x, \eta(x)) - g(\bar{x}, \eta(\bar{x})) = \frac{\partial g}{\partial x}(\gamma(\xi))(x - \bar{x}) + \frac{\partial g}{\partial y}(\gamma(\xi))(\eta(x) - \eta(\bar{x})) \,.$$

Essendo $g(x,\eta(x)) = g(\bar{x},\eta(\bar{x})) = 0$, si ha che

$$|\eta(x) - \eta(\bar{x})| = \left| \frac{\frac{\partial g}{\partial x}(\gamma(\xi))}{\frac{\partial g}{\partial y}(\gamma(\xi))} \right| |x - \bar{x}| \, .$$

Siccome le derivate parziali di g sono continue e $\frac{\partial g}{\partial y}$ è non nulla sul compatto $\overline{U} \times \overline{V}$, si ha che $|\frac{\partial g}{\partial x}(\gamma(\xi))(\frac{\partial g}{\partial y}(\gamma(\xi)))^{-1}|$ è limitato superiormente e ne segue la continuità di η in $\bar{x}$. Resta da vedere la derivabilità: procedendo come sopra si ha che

$$\frac{\eta(x) - \eta(\bar{x})}{x - \bar{x}} = -\frac{\frac{\partial g}{\partial x}(\gamma(\xi))}{\frac{\partial g}{\partial y}(\gamma(\xi))} \, ,$$

con $\gamma(\xi)$ appartenente al segmento che congiunge $(\bar{x},\eta(\bar{x}))$ con $(x,\eta(x))$. Se x tende a $\bar{x}$, si ha che $\gamma(\xi)$ tende a $(\bar{x},\eta(\bar{x}))$ e quindi

$$\eta'(\bar{x}) = \lim_{x \to \bar{x}} \frac{\eta(x) - \eta(\bar{x})}{x - \bar{x}} = -\frac{\frac{\partial g}{\partial x}(\bar{x},\eta(\bar{x}))}{\frac{\partial g}{\partial y}(\bar{x},\eta(\bar{x}))} \, .$$

Ne segue che η è di classe C^1. ■

8 Il teorema della funzione implicita - caso generale

Vediamo come si generalizza il teorema della funzione implicita. Consideremo un insieme aperto Ω di $\mathbb{R}^M \times \mathbb{R}^N$ e una funzione $g : \Omega \to \mathbb{R}^N$, di classe C^1. Quindi, g ha N componenti

$$g(\boldsymbol{x},\boldsymbol{y}) = (g_1(\boldsymbol{x},\boldsymbol{y}), \dots, g_N(\boldsymbol{x},\boldsymbol{y})) \, .$$

Qui $\boldsymbol{x} = (x_1, \dots, x_M) \in \mathbb{R}^M$ e $\boldsymbol{y} = (y_1, \dots, y_N) \in \mathbb{R}^N$. Useremo la seguente notazione per le matrici jacobiane:

$$\frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x},\boldsymbol{y}) = \begin{pmatrix} \frac{\partial g_1}{\partial x_1}(\boldsymbol{x},\boldsymbol{y}) & \cdots & \frac{\partial g_1}{\partial x_M}(\boldsymbol{x},\boldsymbol{y}) \\ \vdots & \dots & \vdots \\ \frac{\partial g_N}{\partial x_1}(\boldsymbol{x},\boldsymbol{y}) & \cdots & \frac{\partial g_N}{\partial x_M}(\boldsymbol{x},\boldsymbol{y}) \end{pmatrix}, \quad \frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{x},\boldsymbol{y}) = \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(\boldsymbol{x},\boldsymbol{y}) & \cdots & \frac{\partial g_1}{\partial y_N}(\boldsymbol{x},\boldsymbol{y}) \\ \vdots & \dots & \vdots \\ \frac{\partial g_N}{\partial y_1}(\boldsymbol{x},\boldsymbol{y}) & \cdots & \frac{\partial g_N}{\partial y_N}(\boldsymbol{x},\boldsymbol{y}) \end{pmatrix}.$$

Teorema. *Siano $\Omega \subseteq \mathbb{R}^M \times \mathbb{R}^N$ un aperto, $g : \Omega \to \mathbb{R}^N$ una funzione di classe C^1 e $(\boldsymbol{x}_0,\boldsymbol{y}_0)$ un punto di Ω per cui si abbia:*

$$g(\boldsymbol{x}_0,\boldsymbol{y}_0) = 0 \qquad \det \frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{x}_0,\boldsymbol{y}_0) \neq 0 \, .$$

Allora esistono un intorno aperto U di $\boldsymbol{x}_0$, un intorno aperto V di $\boldsymbol{y}_0$ e una funzione $\eta : U \to V$ di classe C^1 tali che $U \times V \subseteq \Omega$ e, presi $\boldsymbol{x} \in U$ e $\boldsymbol{y} \in V$, si ha:

$$g(\boldsymbol{x},\boldsymbol{y}) = 0 \quad \Longleftrightarrow \quad \boldsymbol{y} = \eta(\boldsymbol{x}) \, .$$

Inoltre, la funzione η è di classe C^1 e vale la formula

$$J\eta(\boldsymbol{x}) = -\left(\frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{x}, \eta(\boldsymbol{x}))\right)^{-1} \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}, \eta(\boldsymbol{x}))\,.$$

<u>Dimostrazione</u>. [3] Faremo la dimostrazione per induzione su N.

Nel caso $N = 1$ e $M \geq 2$, si procede in modo del tutto analogo a quanto già fatto nel caso $M = 1$. Basterà prendere, al posto dell'intervallo $[x_0-\delta, x_0+\delta]$, la palla chiusa $\bar{B}(\boldsymbol{x}_0, \delta)$, e similmente per gli intorni aperti di $\boldsymbol{x}_0$, per dimostrare l'esistenza e la continuità della funzione η. Resta da vedere la derivabilità: considerato $\bar{\boldsymbol{x}} = (\bar{x}_1, \dots, \bar{x}_M)$, prendiamo ora $\boldsymbol{x} = (\bar{x}_1+h, \dots, \bar{x}_M)$; procedendo come in precedenza, si ha che

$$\frac{\eta(\bar{x}_1 + h, \dots, \bar{x}_M) - \eta(\bar{x}_1, \dots, \bar{x}_M)}{h} = -\frac{\frac{\partial g}{\partial x_1}(\gamma(\xi))}{\frac{\partial g}{\partial y}(\gamma(\xi))}\,,$$

con $\gamma(\xi)$ appartenente al segmento che congiunge $(\bar{\boldsymbol{x}}, \eta(\bar{\boldsymbol{x}}))$ con $(\boldsymbol{x}, \eta(\boldsymbol{x}))$. Se h tende a 0, si ha che $\gamma(\xi)$ tende a $(\bar{\boldsymbol{x}}, \eta(\bar{\boldsymbol{x}}))$ e quindi

$$\frac{\partial \eta}{\partial x_1}(\bar{\boldsymbol{x}}) = \lim_{h\to 0} \frac{\eta(\bar{x}_1 + h, \dots, \bar{x}_M) - \eta(\bar{x}_1, \dots, \bar{x}_M)}{h} = -\frac{\frac{\partial g}{\partial x_1}(\bar{\boldsymbol{x}}, \eta(\bar{\boldsymbol{x}}))}{\frac{\partial g}{\partial y}(\bar{\boldsymbol{x}}, \eta(\bar{\boldsymbol{x}}))}\,.$$

Analogamente si calcolano le derivate parziali rispetto a $x_2, \dots, x_M$, per cui si vede che η è di classe C^1 e

$$J\eta(\boldsymbol{x}) = -\frac{1}{\frac{\partial g}{\partial y}(\boldsymbol{x}, \eta(\boldsymbol{x}))} \frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}, \eta(\boldsymbol{x}))\,.$$

Supponiamo ora l'enunciato valido fino a $N - 1$, per un certo $N \geq 2$ (e $M \geq 1$ qualsiasi) e dimostriamo che vale anche per N. Useremo la notazione

$$\tilde{\boldsymbol{y}}_1 = (y_1, \dots, y_{N-1})\,,$$

per cui scriveremo $y = (\tilde{\boldsymbol{y}}_1, y_N)$. Siccome

$$\det \begin{pmatrix} \frac{\partial g_1}{\partial y_1}(\boldsymbol{x}_0, \boldsymbol{y}_0) & \cdots & \frac{\partial g_1}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0) \\ \vdots & \cdots & \vdots \\ \frac{\partial g_N}{\partial y_1}(\boldsymbol{x}_0, \boldsymbol{y}_0) & \cdots & \frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0) \end{pmatrix} \neq 0\,,$$

almeno uno degli elementi dell'ultima colonna è non nullo. Possiamo supporre senza perdita di generalità, eventualmente permutando le righe, che sia $\frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0) \neq 0$. Scrivendo $\boldsymbol{y}_0 = (\tilde{\boldsymbol{y}}_1^0, y_N^0)$, con $\tilde{\boldsymbol{y}}_1^0 = (y_1^0, \dots, y_{N-1}^0)$, sarà

$$g_N(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0, y_N^0) = 0\,, \qquad \frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0, y_N^0) \neq 0\,.$$

[3] Dimostrazione non vista a lezione.

Allora (caso unidimensionale) esistono un intorno aperto U_1 di $(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0)$, un intorno aperto V_N di y_N^0 e una funzione $\eta_1 : U_1 \to V_N$ di classe C^1 tali che $U_1 \times V_N \subseteq \Omega$, per cui si abbia: se $(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \in U_1$ e $y_N \in V_N$,

$$g_N(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, y_N) = 0 \quad \Longleftrightarrow \quad y_N = \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1)\,,$$

e

$$J\eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) = -\frac{1}{\frac{\partial g_N}{\partial y_N}(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1)))}\;\frac{\partial g_N}{\partial(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1)}(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1))\,.$$

Possiamo supporre U_1 della forma $\tilde{U} \times \tilde{V}_1$, con $\tilde{U}$ intorno aperto di $\boldsymbol{x}_0$ e $\tilde{V}_1$ intorno aperto di $\tilde{\boldsymbol{y}}_1^0$. Definiamo la funzione $\phi : \tilde{U} \times \tilde{V}_1 \to \mathbb{R}^{N-1}$, ponendo

$$\phi(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) = (g_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1)), \ldots, g_{N-1}(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1)))\,.$$

Per brevità, scriveremo

$$g_{(1,\ldots,N-1)}(\boldsymbol{x}, \boldsymbol{y}) = (g_1(\boldsymbol{x}, \boldsymbol{y}), \ldots, g_{N-1}(\boldsymbol{x}, \boldsymbol{y}))\,.$$

Notiamo che $\phi(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) = 0$ e che, essendo $\eta_1(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) = y_N^0$,

$$\frac{\partial \phi}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) = \frac{\partial g_{(1,\ldots,N-1)}}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \boldsymbol{y}_0) + \frac{\partial g_{(1,\ldots,N-1)}}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0)\frac{\partial \eta_1}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0)\,. \qquad (*)$$

Inoltre, siccome $g_N(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1)) = 0$, per ogni $(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \in U_1$, differenziando si ha:

$$0 = \frac{\partial g_N}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \boldsymbol{y}_0) + \frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0)\frac{\partial \eta_1}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0)\,. \qquad (**)$$

Scriviamo

$$\det \frac{\partial \phi}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) = \frac{1}{\frac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0)} \det \left(\begin{array}{c|c} \dfrac{\partial \phi}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) & \dfrac{\partial g_{(1,\ldots,N-1)}}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0) \\ \hline 0 & \dfrac{\partial g_N}{\partial y_N}(\boldsymbol{x}_0, \boldsymbol{y}_0) \end{array}\right),$$

avendo usato la notazione di matrice suddivisa a blocchi. Sostituendo le due uguaglianze $(*)$, $(**)$ e usando le proprietà dei determinanti, si ha:

$$\begin{aligned}
&\det \left(\begin{array}{c|c} \dfrac{\partial \phi}{\partial \tilde{y}_1}(x_0, \tilde{y}_1^0) & \dfrac{\partial g_{(1,\ldots,N-1)}}{\partial y_N}(x_0, y_0) \\ \hline 0 & \dfrac{\partial g_N}{\partial y_N}(x_0, y_0) \end{array}\right) = \\
&= \det \left(\begin{array}{c|c} \dfrac{\partial g_{(1,\ldots,N-1)}}{\partial \tilde{y}_1}(x_0, y_0) + \dfrac{\partial g_{(1,\ldots,N-1)}}{\partial y_N}(x_0, y_0)\dfrac{\partial \eta_1}{\partial \tilde{y}_1}(x_0, \tilde{y}_1^0) & \dfrac{\partial g_{(1,\ldots,N-1)}}{\partial y_N}(x_0, y_0) \\ \hline \dfrac{\partial g_N}{\partial \tilde{y}_1}(x_0, y_0) + \dfrac{\partial g_N}{\partial y_N}(x_0, y_0)\dfrac{\partial \eta_1}{\partial \tilde{y}_1}(x_0, \tilde{y}_1^0) & \dfrac{\partial g_N}{\partial y_N}(x_0, y_0) \end{array}\right) \\
&= \det \left(\frac{\partial g}{\partial \tilde{y}_1}(x_0, y_0) + \frac{\partial g}{\partial y_N}(x_0, y_0)\frac{\partial \eta_1}{\partial \tilde{y}_1}(x_0, \tilde{y}_1^0) \;\middle|\; \frac{\partial g}{\partial y_N}(x_0, y_0) \right) \\
&= \det \left(\frac{\partial g}{\partial \tilde{y}_1}(x_0, y_0) \;\middle|\; \frac{\partial g}{\partial y_N}(x_0, y_0) \right) = \det \frac{\partial g}{\partial y}(x_0, y_0) \neq 0\,.
\end{aligned}$$

Abbiamo quindi

$$\phi(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) = 0\,, \qquad \det \frac{\partial \phi}{\partial \tilde{\boldsymbol{y}}_1}(\boldsymbol{x}_0, \tilde{\boldsymbol{y}}_1^0) \neq 0\,.$$

Per l'ipotesi induttiva, esistono un intorno aperto U di $\boldsymbol{x}_0$, un intorno aperto V_1 di $\tilde{\boldsymbol{y}}_1^0$ e una funzione $\eta_2 : U \to V_1$ di classe C^1 tali che $U \times V_1 \subseteq \tilde{U} \times \tilde{V}_1$, per cui si abbia: per ogni $\boldsymbol{x} \in U$ e $\tilde{\boldsymbol{y}}_1 \in V_1$,

$$\phi(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) = 0 \quad \Longleftrightarrow \quad \tilde{\boldsymbol{y}}_1 = \eta_2(\boldsymbol{x})\,.$$

In conclusione, per $\boldsymbol{x} \in U$ e $\boldsymbol{y} = (\tilde{\boldsymbol{y}}_1, y_N) \in V_1 \times V_2$, si ha:

$$\begin{aligned}
g(\boldsymbol{x}, \boldsymbol{y}) = 0 \quad & \Longleftrightarrow \quad \begin{cases} g_{(1,\dots,N-1)}(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, y_N) = 0 \\ g_N(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, y_N) = 0 \end{cases} \\
& \Longleftrightarrow \quad \begin{cases} g_{(1,\dots,N-1)}(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1, y_N) = 0 \\ y_N = \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \end{cases} \\
& \Longleftrightarrow \quad \begin{cases} \phi(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) = 0 \\ y_N = \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \end{cases} \\
& \Longleftrightarrow \quad \begin{cases} \tilde{\boldsymbol{y}}_1 = \eta_2(\boldsymbol{x}) \\ y_N = \eta_1(\boldsymbol{x}, \tilde{\boldsymbol{y}}_1) \end{cases} \\
& \Longleftrightarrow \quad \boldsymbol{y} = (\eta_2(\boldsymbol{x}), \eta_1(\boldsymbol{x}, \eta_2(\boldsymbol{x})))\,.
\end{aligned}$$

Ponendo $V = V_1 \times V_2$, resta pertanto definita la funzione $\eta : U \to V$:

$$\eta(\boldsymbol{x}) = (\eta_2(\boldsymbol{x}), \eta_1(\boldsymbol{x}, \eta_2(\boldsymbol{x})))\,.$$

Tale funzione è di classe C^1, siccome lo sono sia η_1 che η_2. Siccome $g(\boldsymbol{x}, \eta(\boldsymbol{x})) = 0$ per ogni $\boldsymbol{x} \in U$, se ne deduce che che

$$\frac{\partial g}{\partial \boldsymbol{x}}(\boldsymbol{x}, \eta(\boldsymbol{x})) + \frac{\partial g}{\partial \boldsymbol{y}}(\boldsymbol{x}, \eta(\boldsymbol{x}))J\eta(\boldsymbol{x}) = 0\,,$$

da cui la formula per $J\eta(\boldsymbol{x})$. ■

9 $M-$superfici

Indichiamo con I un rettangolo di $\mathbb{R}^M$, dove $1 \leq M \leq N$. Avremo quindi:

$$I = [a_1, b_1] \times [a_2, b_2] \times \ldots \times [a_M, b_M]\,.$$

Definizione. *Chiameremo "$M-$superficie" in $\mathbb{R}^N$ una funzione* [4] $\sigma : I \to \mathbb{R}^N$ *di classe C^1. Se $M = 1$, σ si dirà anche "curva"; se $M = 2$, si dirà semplicemente "superficie". L'insieme $\sigma(I)$ è detto "supporto" della $M-$superficie σ. Diremo che la $M-$superficie σ è "regolare" se, per ogni $\boldsymbol{u}$ interno ad I, la matrice jacobiana $J\sigma(\boldsymbol{u})$ ha rango M.*

[4] Le derivate parziali di σ devono essere continue su tutto I e nei punti di frontiera vanno intese, se necessario, come derivate destre o sinistre. Equivalentemente, si potrebbe estendere σ ad una funzione di classe C^1 definita su un aperto contenente I.

Consideriamo da vicino il caso $N = 3$. Una curva in $\mathbb{R}^3$ è una funzione $\sigma : [a,b] \to \mathbb{R}^3$, $\sigma = (\sigma_1, \sigma_2, \sigma_3)$. La curva è regolare se, per ogni $t \in]a,b[$, il "vettore derivata" (o "vettore velocità") $\sigma'(t) = (\sigma_1'(t), \sigma_2'(t), \sigma_3'(t))$ è non nullo. È individuata una retta, passante per il punto $\sigma(t)$, avente la direzione di $\sigma'(t)$; essa è detta "retta tangente" alla curva nel punto $\sigma(t)$. Si definisce il seguente "versore tangente":

$$\tau_\sigma(t) = \frac{\sigma'(t)}{\|\sigma'(t)\|} .$$

Una superficie in $\mathbb{R}^3$ è una funzione $\sigma : [a_1, b_1] \times [a_2, b_2] \to \mathbb{R}^3$. La superficie è regolare se, per ogni $(u,v) \in]a_1, b_1[\times]a_2, b_2[$, i vettori $\frac{\partial\sigma}{\partial u}(u,v)$, $\frac{\partial\sigma}{\partial v}(u,v)$ sono linearmente indipendenti. È individuato un piano, contenete il punto $\sigma(u,v)$, parallelo al piano generato dai due vettori $\frac{\partial\sigma}{\partial u}(u,v)$ e $\frac{\partial\sigma}{\partial v}(u,v)$; esso è detto "piano tangente" alla superficie nel punto $\sigma(u,v)$. Si definisce il seguente "versore normale":

$$\nu_\sigma(u,v) = \frac{\frac{\partial\sigma}{\partial u}(u,v) \times \frac{\partial\sigma}{\partial v}(u,v)}{||\frac{\partial\sigma}{\partial u}(u,v) \times \frac{\partial\sigma}{\partial v}(u,v)||} .$$

Esempi. 1. La superficie $\sigma : [0,\pi] \times [0,\pi] \to \mathbb{R}^3$ definita da

$$\sigma(u,v) = (R \sin u \cos v, R \sin u \sin v, R \cos u)$$

ha come supporto la semisfera

$$\{(x,y,z) : x^2 + y^2 + z^2 = R^2, y \geq 0\} .$$

Calcolando

$$\frac{\partial\sigma}{\partial u}(u,v) = (R \cos u \cos v, R \cos u \sin v, -R \sin u) ,$$
$$\frac{\partial\sigma}{\partial v}(u,v) = (-R \sin u \sin v, R \sin u \cos v, 0) ,$$

vediamo che

$$\frac{\partial\sigma}{\partial u}(u,v) \times \frac{\partial\sigma}{\partial v}(u,v) = (\sin^2 u \cos v, \sin^2 u \sin v, \sin u \cos u) .$$

Essendo

$$\left\| \frac{\partial\sigma}{\partial u}(u,v) \times \frac{\partial\sigma}{\partial v}(u,v) \right\| = \sin u ,$$

si tratta di una superficie regolare, e si ha:

$$\nu_\sigma(u,v) = (\sin u \cos v, \sin u \sin v, \cos u) .$$

2. La superficie $\sigma : [0, 2\pi] \times [0, 2\pi] \to \mathbb{R}^3$ definita da

$$\sigma(u,v) = ((R + r \cos u) \cos v, (R + r \cos u) \sin v, r \sin u)$$

dove $0 < r < R$, ha come supporto l'anello toroidale o "toro"

$$\{(x,y,z) : (\sqrt{x^2 + y^2} - R)^2 + z^2 = r^2\} .$$

Si può verificare che anche in questo caso si tratta di una superficie regolare.

3. La superficie $\sigma : [r, R] \times [0, 2\pi] \to \mathbb{R}^3$, con $0 \leq r < R$, data da

$$\sigma(u, v) = (u \cos v, u \sin v, 0)\,,$$

ha come supporto l'insieme

$$\{(x, y, z) : r^2 \leq x^2 + y^2 \leq R^2,\ z = 0\}\,,$$

che è un cerchio se $r = 0$, una corona circolare se $r > 0$. Anche in questo caso si tratta di una superficie regolare.

4. La superficie $\sigma : [r, R] \times [0, 2\pi] \to \mathbb{R}^3$, con $0 < r < R$, definita da

$$\begin{aligned}\sigma(u, v) = \Bigg(&\left(\frac{r+R}{2} + \left(u - \frac{r+R}{2}\right)\cos\left(\frac{v}{2}\right)\right)\cos v, \\ &\left(\frac{r+R}{2} + \left(u - \frac{r+R}{2}\right)\cos\left(\frac{v}{2}\right)\right)\sin v, \\ &\left(u - \frac{r+R}{2}\right)\sin\left(\frac{v}{2}\right)\Bigg)\,,\end{aligned}$$

ha come supporto un nastro di Möbius. Si può verificare anche in questo caso che si tratta di una superficie regolare.

10 $M-$parametrizzazioni locali

In questa sezione supporremo $1 \leq M < N$. Identificando $\mathbb{R}^N$ con $\mathbb{R}^M \times \mathbb{R}^{N-M}$, ogni vettore $\boldsymbol{p} = (p_1, \ldots, p_N)$ di $\mathbb{R}^N$ si scriverà nella forma $\boldsymbol{p} = (\hat{\boldsymbol{p}}, \tilde{\boldsymbol{p}})$, con $\hat{\boldsymbol{p}} = (p_1, \ldots, p_M)$ e $\tilde{\boldsymbol{p}} = (p_{M+1}, \ldots, p_N)$.

Useremo inoltre la seguente notazione: dato $\hat{\boldsymbol{p}} = (p_1, \ldots, p_M) \in \mathbb{R}^M$ e $r > 0$,

$$B[\hat{\boldsymbol{p}}, r] = [p_1 - r, p_1 + r] \times \ldots \times [p_M - r, p_M + r] \subseteq \mathbb{R}^M\,.$$

Per semplicità, scriveremo $B[r]$ invece di $B[\mathbf{0}, r]$.

Teorema. *Siano Ω un sottoinsieme aperto di $\mathbb{R}^N$, $\boldsymbol{p}_0$ un punto di Ω e $g : \Omega \to \mathbb{R}^{N-M}$ una funzione di classe C^1, tale che $g(\boldsymbol{p}_0) = 0$ e $Jg(\boldsymbol{p}_0)$ abbia rango $N - M$. Allora esistono un intorno U di $\boldsymbol{p}_0$ e una $M-$superficie regolare e iniettiva $\sigma : B[r] \to \mathbb{R}^N$, per un certo $r > 0$, tali che $\sigma(\mathbf{0}) = \boldsymbol{p}_0$ e*

$$\{\boldsymbol{p} \in U : g(\boldsymbol{p}) = 0\} = \sigma(B[r])\,.$$

<u>Dimostrazione</u>. Supponiamo, per esempio, che sia invertibile

$$\frac{\partial g}{\partial \tilde{\boldsymbol{p}}}(\boldsymbol{p}_0) = \begin{pmatrix} \frac{\partial g_1}{\partial p_{M+1}}(\boldsymbol{p}_0) & \cdots & \frac{\partial g_1}{\partial p_N}(\boldsymbol{p}_0) \\ \vdots & \cdots & \vdots \\ \frac{\partial g_{N-M}}{\partial p_{M+1}}(\boldsymbol{p}_0) & \cdots & \frac{\partial g_{N-M}}{\partial p_N}(\boldsymbol{p}_0) \end{pmatrix}.$$

Per il teorema della funzione implicita, esistono un intorno aperto $\hat{U}$ di $\hat{\boldsymbol{p}}_0$, un intorno aperto $\tilde{U}$ di $\tilde{\boldsymbol{p}}_0$ e una funzione $\eta : \hat{U} \to \tilde{U}$, tali che $\hat{U} \times \tilde{U} \subseteq \Omega$ e, se $\hat{\boldsymbol{p}} \in \hat{U}$ e $\tilde{\boldsymbol{p}} \in \tilde{U}$, si ha:

$$g(\hat{\boldsymbol{p}}, \tilde{\boldsymbol{p}}) = 0 \quad \Longleftrightarrow \quad \tilde{\boldsymbol{p}} = \eta(\hat{\boldsymbol{p}}) .$$

Preso $r > 0$ tale che $B[\hat{\boldsymbol{p}}_0, r] \subseteq \hat{U}$, sia $U = B[\hat{\boldsymbol{p}}_0, r] \times \tilde{U}$ e $\sigma : B[r] \to \mathbb{R}^N$ definita da $\sigma(\boldsymbol{u}) = (\boldsymbol{u} + \hat{\boldsymbol{p}}_0, \eta(\boldsymbol{u} + \hat{\boldsymbol{p}}_0))$. Si verifica che la matrice jacobiana $J\sigma(\boldsymbol{u})$ ha come sottomatrice la matrice $M \times M$ identità, per cui σ è regolare. Si vede facilmente che σ è iniettiva, in quanto lo è la prima componente $\boldsymbol{u} \mapsto \boldsymbol{u} + \hat{\boldsymbol{p}}_0$. Inoltre, se $\boldsymbol{p} = (\hat{\boldsymbol{p}}, \tilde{\boldsymbol{p}}) \in U$,

$$g(\hat{\boldsymbol{p}}, \tilde{\boldsymbol{p}}) = 0 \quad \Leftrightarrow \quad \tilde{\boldsymbol{p}} = \eta(\hat{\boldsymbol{p}}) \quad \Leftrightarrow \quad (\hat{\boldsymbol{p}}, \tilde{\boldsymbol{p}}) = \sigma(\hat{\boldsymbol{p}} - \hat{\boldsymbol{p}}_0) ,$$

da cui la tesi. Nel caso in cui la sottomatrice considerata non sia invertibile, basterà operare degli scambi nelle colonne della matrice $Jg(\boldsymbol{p}_0)$ per ricondursi alla situazione precedente. ■

La $M-$superficie σ individuata dal teorema precedente è detta "$M-$parametrizzazione locale".

Vediamo tre casi di particolare interesse. Iniziamo con una curva in $\mathbb{R}^2$ (caso $M = 1$, $N = 2$).

Corollario 1. Siano Ω un sottoinsieme aperto di $\mathbb{R}^2$, (x_0, y_0) un punto di Ω e $g : \Omega \to \mathbb{R}$ una funzione di classe C^1, tale che $g(x_0, y_0) = 0$ e $\nabla g(x_0, y_0) \neq 0$. Allora esistono un intorno U di (x_0, y_0) e una curva regolare e iniettiva $\sigma : [-r, r] \to \mathbb{R}^2$, per un certo $r > 0$, tali che $\sigma(0) = (x_0, y_0)$ e

$$\{(x, y) \in U : g(x, y) = 0\} = \sigma([-r, r]) .$$

Vediamo ora il caso di una superficie in $\mathbb{R}^3$ (caso $M = 2$, $N = 3$).

Corollario 2. Siano Ω un sottoinsieme aperto di $\mathbb{R}^3$, (x_0, y_0, z_0) un punto di Ω e $g : \Omega \to \mathbb{R}$ una funzione di classe C^1, tale che $g(x_0, y_0, z_0) = 0$ e $\nabla g(x_0, y_0, z_0) \neq 0$. Allora esistono un intorno U di (x_0, y_0, z_0) e una superficie regolare e iniettiva $\sigma : [-r, r] \times [-r, r] \to \mathbb{R}^3$, per un certo $r > 0$, tali che $\sigma(0, 0) = (x_0, y_0, z_0)$ e

$$\{(x, y, z) \in U : g(x, y, z) = 0\} = \sigma([-r, r] \times [-r, r]) .$$

Infine, vediamo il caso di una curva in $\mathbb{R}^3$ (caso $M = 1$, $N = 3$).

Corollario 3. Siano Ω un sottoinsieme aperto di $\mathbb{R}^3$, (x_0, y_0, z_0) un punto di Ω e $g_1, g_2 : \Omega \to \mathbb{R}$ due funzioni di classe C^1, tali che

$$g_1(x_0, y_0, z_0) = g_2(x_0, y_0, z_0) = 0 ,$$

e

$$\nabla g_1(x_0, y_0, z_0) \times \nabla g_2(x_0, y_0, z_0) \neq 0 .$$

Allora esistono un intorno U di (x_0, y_0, z_0) e una curva regolare e iniettiva $\sigma : [-r, r] \to \mathbb{R}^3$, per un certo $r > 0$, tali che $\sigma(0) = (x_0, y_0, z_0)$ e

$$\{(x, y, z) \in U : g_1(x, y, z) = g_2(x, y, z) = 0\} = \sigma([-r, r]) .$$

11 Moltiplicatori di Lagrange

Teorema. Siano Ω un aperto di $\mathbb{R}^N$, $\boldsymbol{p}_0$ un punto di Ω, $g : \Omega \to \mathbb{R}^{N-M}$ una funzione di classe C^1 tale che $g(\boldsymbol{p}_0) = 0$ e $Jg(\boldsymbol{p}_0)$ abbia rango $N - M$, e $f : \Omega \to \mathbb{R}$ una funzione differenziabile in $\boldsymbol{p}_0$. Posto

$$S = \{\boldsymbol{p} \in \Omega : g(\boldsymbol{p}) = 0\} ,$$

se $\boldsymbol{p}_0$ è un punto di minimo o massimo relativo per $f|_S$, allora esistono $(N-M)$ numeri reali $\lambda_1, \ldots, \lambda_{N-M}$ tali che

$$\nabla f(\boldsymbol{p}_0) = \sum_{j=1}^{N-M} \lambda_j \nabla g_j(\boldsymbol{p}_0) .$$

I numeri $\lambda_1, \ldots, \lambda_{N-M}$ si chiamano **moltiplicatori di Lagrange**.

<u>Dimostrazione</u>. Per il teorema precedente, esistono un intorno U di $\boldsymbol{p}_0$, un $r > 0$ e una $M-$superficie regolare $\sigma : B[\mathbf{0}, r] \to \mathbb{R}^N$ tali che $\sigma(\mathbf{0}) = \boldsymbol{p}_0$ e

$$S \cap U = \sigma(B[r]) .$$

Considerata la funzione $F(\boldsymbol{y}) = f(\sigma(\boldsymbol{y}))$, si ha che $\mathbf{0}$ è un punto di minimo o massimo relativo per F. Quindi, $\nabla F(\mathbf{0}) = \mathbf{0}$, per cui

$$0 = JF(\mathbf{0}) = Jf(\boldsymbol{p}_0) J\sigma(\mathbf{0}) .$$

Ne segue che $\nabla f(\boldsymbol{p}_0)$ è ortogonale a $\frac{\partial \sigma}{\partial y_1}(\mathbf{0}), \ldots, \frac{\partial \sigma}{\partial y_M}(\mathbf{0})$. Inoltre, essendo per ogni $y \in B[r]$

$$g(\sigma(\boldsymbol{y})) = 0 ,$$

si ha che

$$Jg(\boldsymbol{p}_0) J\sigma(\mathbf{0}) = 0 .$$

Quindi anche i vettori $\nabla g_1(\boldsymbol{p}_0), \ldots, \nabla g_{N-M}(\boldsymbol{p}_0)$ sono tutti ortogonali a $\frac{\partial \sigma}{\partial y_1}(\mathbf{0})$, $\ldots$, $\frac{\partial \sigma}{\partial y_M}(\mathbf{0})$.

Essendo $J\sigma(\mathbf{0})$ di rango M, lo spazio vettoriale $\mathcal{T}$ generato da $\frac{\partial \sigma}{\partial y_1}(\mathbf{0})$, $\ldots$, $\frac{\partial \sigma}{\partial y_M}(\mathbf{0})$ ha dimensione M. Quindi lo spazio ortogonale $\mathcal{T}^\perp$ ha dimesnione $N - M$. Siccome, come abbiamo visto,

$$\nabla f(\boldsymbol{p}_0), \nabla g_1(\boldsymbol{p}_0), \ldots, \nabla g_{N-M}(\boldsymbol{p}_0) \in \mathcal{T}^\perp ,$$

essi devono essere linearmente dipendenti. Essendo i vettori $\nabla g_1(\boldsymbol{p}_0)$, $\ldots$, $\nabla g_{N-M}(\boldsymbol{p}_0)$ linearmente indipendenti, ne segue che $\nabla f(\boldsymbol{p}_0)$ si deve poter esprimere come combinazione lineare dei $\nabla g_j(\boldsymbol{p}_0)$. ■

Vediamo anche qui tre casi particolari interessanti.

Corollario 1. Siano Ω un aperto di $\mathbb{R}^2$, (x_0, y_0) un punto di Ω, $g : \Omega \to \mathbb{R}$ una funzione di classe C^1 tale che

$$g(x_0, y_0) = 0 \quad \text{e} \quad \nabla g(x_0, y_0) \neq 0,$$

e $f : \Omega \to \mathbb{R}$ una funzione differenziabile in (x_0, y_0). Posto

$$S = \{(x, y) \in \Omega : g(x, y) = 0\} ,$$

se (x_0, y_0) è un punto di minimo o massimo relativo per $f|_S$, allora esiste un numero reale λ tale che

$$\nabla f(x_0, y_0) = \lambda \nabla g(x_0, y_0) .$$

Corollario 2. Siano Ω un aperto di $\mathbb{R}^3$, (x_0, y_0, z_0) un punto di Ω, $g : \Omega \to \mathbb{R}$ una funzione di classe C^1 tale che

$$g(x_0, y_0, z_0) = 0 \quad \text{e} \quad \nabla g(x_0, y_0, z_0) \neq 0,$$

e $f : \Omega \to \mathbb{R}$ una funzione differenziabile in (x_0, y_0, z_0). Posto

$$S = \{(x, y, z) \in \Omega : g(x, y, z) = 0\} ,$$

se (x_0, y_0, z_0) è un punto di minimo o massimo relativo per $f|_S$, allora esiste un numero reale λ tale che

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0) .$$

Corollario 3. Siano Ω un aperto di $\mathbb{R}^3$, (x_0, y_0, z_0) un punto di Ω, $g_1, g_2 : \Omega \to \mathbb{R}$ due funzioni di classe C^1 tali che

$$g_1(x_0, y_0, z_0) = g_2(x_0, y_0, z_0) = 0 \quad \text{e} \quad \nabla g_1(x_0, y_0, z_0) \times \nabla g_2(x_0, y_0, z_0) \neq 0 ,$$

e $f : \Omega \to \mathbb{R}$ una funzione differenziabile in (x_0, y_0, z_0). Posto

$$S = \{(x, y, z) \in U : g_1(x, y, z) = 0,\ g_2(x, y, z) = 0\} ,$$

se (x_0, y_0, z_0) è un punto di minimo o massimo relativo per $f|_S$, allora esistono due numeri reali λ_1, λ_2 tali che

$$\nabla f(x_0, y_0, z_0) = \lambda_1 \nabla g_1(x_0, y_0, z_0) + \lambda_2 \nabla g_2(x_0, y_0, z_0) .$$

12 Convergenza uniforme

Siano E de E' due spazi metrici. Indichiamo con $\mathcal{B}(E, E')$ l'insieme delle funzioni $f : E \to E'$ limitate. [5] Definiamo in $\mathcal{B}(E, E')$ una distanza:

$$d_{\mathcal{B}}(f, g) = \sup\{d(f(x), g(x)) : x \in E\} .$$

[5] Si ricorda che $f : E \to E'$ è limitata se la sua immagine $f(E)$ è limitata, ossia è contenuta in una palla di E'.

Verifichiamo che si tratta effettivamente di una distanza. Prese due funzioni f, g in $\mathcal{B}(E, E')$, è evidente che $d_{\mathcal{B}}(f, g) \geq 0$. Inoltre, se $d_{\mathcal{B}}(f, g) = 0$, allora $f(x) = g(x)$ per ogni $x \in E$, quindi $f = g$. La simmetria $d_{\mathcal{B}}(f, g) = d_{\mathcal{B}}(g, f)$ è anch'essa di immediata verifica. Vediamo la disuguaglianza triangolare: presa una terza funzione h, per ogni $x \in E$ si ha

$$d(f(x), h(x)) \leq d(f(x), g(x)) + d(g(x), h(x)) \leq d_{\mathcal{B}}(f, g) + d_{\mathcal{B}}(g, h)\,,$$

per cui

$$d_{\mathcal{B}}(f, h) = \sup\{d(f(x), h(x)) : x \in E\} \leq d_{\mathcal{B}}(f, g) + d_{\mathcal{B}}(g, h)\,.$$

Nel seguito, $\mathcal{B}(E, E')$ sarà sempre considerato uno spazio metrico con la distanza $d_{\mathcal{B}}$ sopra introdotta.

Teorema. Se E' è completo, allora anche $\mathcal{B}(E, E')$ è completo.

Dimostrazione. Sia $(f_n)_n$ una successione di Cauchy in $\mathcal{B}(E, E')$. Essendo, per ogni $x \in E$,

$$d(f_m(x), f_n(x)) \leq d_{\mathcal{B}}(f_m, f_n)\,,$$

si ha che $(f_n(x))_n$ è una successione di Cauchy in E', per ogni $x \in E$. Essendo E' completo, la successione $(f_n(x))_n$ ha limite in E'. Indicheremo tale limite con $f(x)$. Resta così definita una funzione $f : E \to E'$. Vediamo che $f_n \to f$ in $\mathcal{B}(E, E')$. Fissiamo $\varepsilon > 0$. Essendo $(f_n)_n$ di Cauchy, esiste un $\bar{n} \in \mathbb{N}$ tale che

$$[\, m \geq \bar{n} \text{ e } n \geq \bar{n}\,] \quad \Rightarrow \quad d_{\mathcal{B}}(f_n, f_m) < \varepsilon \quad \Rightarrow \quad \forall x \in E \quad d(f_n(x), f_m(x)) < \varepsilon\,.$$

Passando al limite, se $m \to \infty$ si ha

$$n \geq \bar{n} \quad \Rightarrow \quad \forall x \in E \quad d(f_n(x), f(x)) \leq \varepsilon\,.$$

Ne segue che $f \in \mathcal{B}(E, E')$ e

$$n \geq \bar{n} \;\Rightarrow\; \quad d_{\mathcal{B}}(f_n, f) \leq \varepsilon\,.$$

■

Se $(f_n)_n$ è una successione e $f_n \to f$ in $\mathcal{B}(E, E')$, diremo che $(f_n)_n$ "converge uniformemente" a f. In tal caso, si ha che

$$\forall \varepsilon > 0 \quad \exists \bar{n} \in \mathbb{N} : \quad \forall x \in E \quad n \geq \bar{n} \;\Rightarrow\; d(f_n(x), f(x)) < \varepsilon\,.$$

Teorema. L'insieme

$$\{f \in \mathcal{B}(E, E') : f \text{ è continua su } E\}$$

è chiuso in $\mathcal{B}(E, E')$.

Dimostrazione. Sia $(f_n)_n$ una successione di funzioni continue in $\mathcal{B}(E,E')$ tale che $f_n \to f$, per una certa funzione $f \in \mathcal{B}(E,E')$. Si tratta di dimostrare che tale f è continua. Prendiamo un punto x_0 in E e dimostriamo che f è continua in x_0. Fissiamo un $\varepsilon > 0$. Esiste un $\bar{n} \in \mathbb{N}$ tale che

$$n \geq \bar{n} \quad \Rightarrow \quad d_{\mathcal{B}}(f_n, f) < \frac{1}{3}\varepsilon \, .$$

Siccome

$$d(f(x), f(x_0)) \leq d(f(x), f_n(x)) + d(f_n(x), f_n(x_0)) + d(f_n(x_0), f(x_0)) \, ,$$

per $n = \bar{n}$ si ha che

$$d(f(x), f(x_0)) < \frac{2}{3}\varepsilon + d(f_{\bar{n}}(x), f_{\bar{n}}(x_0)) \, .$$

Usando la continuità di $f_{\bar{n}}$, troviamo un $\delta > 0$ tale che

$$d(x, x_0) < \delta \quad \Rightarrow \quad d(f_{\bar{n}}(x), f_{\bar{n}}(x_0)) < \frac{1}{3}\varepsilon \, .$$

Ne segue che

$$d(x, x_0) < \delta \quad \Rightarrow \quad d(f(x), f(x_0)) < \varepsilon \, .$$

■

Indichiamo con $\mathcal{C}(E,E')$ l'insieme delle funzioni $f : E \to E'$ continue. Abbiamo dimostrato che $\mathcal{C}(E,E') \cap \mathcal{B}(E,E')$ è chiuso in $\mathcal{B}(E,E')$. Pertanto, se una successione di funzioni continue e limitate converge uniformemente a una funzione f, allora f è anch'essa continua. Possiamo allora scrivere

$$\lim_n \left(\lim_{x \to x_0} f_n(x) \right) = \lim_{x \to x_0} \left(\lim_n f_n(x) \right) ,$$

per ogni $x_0 \in E$.

Ci occuperemo ora del caso in cui E sia compatto.

Teorema. Se E è compatto e $f \in \mathcal{C}(E,E')$, allora $f(E)$ è compatto.

Dimostrazione. Sia $(y_n)_n$ una successione in $f(E)$. In corrispondenza, possiamo trovare una successione $(x_n)_n$ in E tale che $f(x_n) = y_n$. Essendo E compatto, esiste una sottosuccessione $(x_{n_k})_k$ che ha un limite $\bar{x} \in E$. Essendo $y_{n_k} = f(x_{n_k})$, per la continuità di f,

$$\lim_k f(x_{n_k}) = f(\lim_k x_{n_k}) = f(\bar{x}) \, .$$

Quindi, la sottosuccessione $(y_{n_k})_k$ ha limite $f(\bar{x}) \in f(E)$. ■

Siccome ogni insieme compatto è limitato, dal teorema precedente segue che, se E è compatto, allora $\mathcal{C}(E,E') \subseteq \mathcal{B}(E,E')$.

Corollario. Se E è compatto ed E' è completo, allora $\mathcal{C}(E,E')$ è completo.

Un tipico esempio ne è il caso $E = [a,b] \subseteq \mathbb{R}$ ed $E' = \mathbb{R}$.

13 Due teoremi di passaggio al limite

Dimostreremo ora il seguente teorema di "passaggio al limite sotto il segno di integrale".

Teorema. *Sia $(f_n)_n$ una successione in $\mathcal{C}([a,b],\mathbb{R})$ che converge uniformemente. Allora*

$$\lim_n \int_a^b f_n(t)\,dt = \int_a^b \lim_n f_n(t)\,dt\,.$$

<u>Dimostrazione</u>. Consideriamo la funzione

$$\Psi : \mathcal{C}([a,b],\mathbb{R}) \to \mathcal{C}([a,b],\mathbb{R})$$

così definita:

$$[\Psi(f)](x) = \int_a^x f(t)\,dt\,.$$

Prese $f, g \in \mathcal{C}([a,b],\mathbb{R})$, per ogni $x \in [a,b]$ si ha

$$\begin{aligned} |[\Psi(f)](x) - [\Psi(g)](x)| &= \Big|\int_a^x (f-g)\Big| \le \int_a^x |f-g| \le \int_a^b |f-g| \\ &\le (b-a) d_{\mathcal{B}}(f,g)\,, \end{aligned}$$

da cui segue che

$$d_{\mathcal{B}}(\Psi(f),\Psi(g)) \le (b-a) d_{\mathcal{B}}(f,g)\,.$$

Vediamo così che Ψ è continua.

Sia ora $\lim_n f_n = f$. Per quanto visto sopra, essendo la convergenza uniforme, si ha che $f \in \mathcal{C}([a,b],\mathbb{R})$. Inoltre, essendo Ψ continua, si ha che $\lim_n \Psi(f_n) = \Psi(f)$. Ne segue che

$$\lim_n [\Psi(f_n)](x) = [\Psi(f)](x)\,,$$

per ogni $x \in [a,b]$. In particolare, prendendo $x = b$,

$$\lim_n \int_a^b f_n(t)\,dt = \int_a^b f(t)\,dt\,,$$

che è quanto si voleva dimostrare. ■

Vediamo ora un teorema di "passaggio al limite nella derivata".

Teorema. *Sia $(f_n)_n$ una successione in $\mathcal{C}^1([a,b],\mathbb{R})$ tale che $\lim_n f_n(a) = f(a)$ e $\lim_n f_n' = g$, uniformemente. Allora $f \in \mathcal{C}^1([a,b],\mathbb{R})$ e $f' = g$. Potremo quindi scrivere:*

$$\lim_n \Big(\frac{d}{dt} f_n(t)\Big) = \frac{d}{dt}\Big(\lim_n f_n(t)\Big)\,.$$

Dimostrazione. Per ogni n, si ha:

$$f_n(x) = f_n(a) + \int_a^x f_n'(t)\, dt\,.$$

Siccome $\lim_n f_n(a) = f(a)$ e $(f_n')_n$ converge uniformemente a g, possiamo passare al limite sotto il segno di integrale, ottenendo

$$f(x) = f(a) + \lim_n \int_a^x f_n'(t)\, dt = f(a) + \int_a^x \lim_n f_n'(t)\, dt = f(a) + \int_a^x g(t)\, dt\,.$$

Ne segue che f è derivabile e $f'(x) = g(x)$, per ogni $x \in I$. ■

14 Spazi di Banach

Se E' è uno spazio vettoriale normato, con norma $\|\cdot\|$ e distanza $d(x,x') = \|x - x'\|$, si può definire, per ogni $f \in \mathcal{B}(E,E')$,

$$\|f\|_{\mathcal{B}} = \sup\{\|f(x)\| : x \in E\}\,.$$

Si può verificare che si tratta di una norma, poiché verifica le seguenti proprietà:

$$\begin{aligned}
&a)\ \|f\|_{\mathcal{B}} \geq 0\,,\\
&b)\ \|f\|_{\mathcal{B}} = 0 \quad \Leftrightarrow \quad f = 0\,,\\
&c)\ \|\lambda f\|_{\mathcal{B}} = |\lambda|\, \|f\|_{\mathcal{B}} \qquad (\text{qui } \lambda \text{ è un numero reale})\,,\\
&d)\ \|f + g\|_{\mathcal{B}} \leq \|f\|_{\mathcal{B}} + \|g\|_{\mathcal{B}}\,.
\end{aligned}$$

Si vede inoltre che

$$d_{\mathcal{B}}(f,g) = \|f - g\|_{\mathcal{B}}\,.$$

Definizione. *Si chiama "spazio di Banach" uno spazio vettoriale normato e completo.*

Abbiamo immediatamente i seguenti corollari.

Corollario. Se E' è uno spazio di Banach, allora $\mathcal{B}(E,E')$ è uno spazio di Banach.

Corollario. Se E è compatto ed E' è uno spazio di Banach, allora $\mathcal{C}(E,E')$ è uno spazio di Banach.

Dato un intervallo $I = [a,b]$, indichiamo con il simbolo $\mathcal{C}^m(I)$ l'insieme delle funzioni $f : I \to \mathbb{R}$ di classe C^m, su cui definiremo ora una norma $\|\cdot\|_m$. Se $m = 0$, si intende che $\mathcal{C}^0(I) = \mathcal{C}(I,\mathbb{R})$, e la norma $\|\cdot\|_0$ coinciderà con la norma $\|\cdot\|_{\mathcal{B}}$ già introdotta. Se $m = 1$, poniamo

$$\|f\|_1 = \|f\|_{\mathcal{B}} + \|f'\|_{\mathcal{B}}\,.$$

Si può verificare che si tratta effettivamente di una norma. Vogliamo dimostrare che, con questa norma, $\mathcal{C}^1(I)$ è uno spazio di Banach. A questo scopo, dimostriamo dapprima il seguente risultato.

Dimostriamo ora che $\mathcal{C}^1(I)$ è uno spazio di Banach. Sia $(f_n)_n$ una successione di Cauchy in $\mathcal{C}^1(I)$. Allora $(f_n)_n$ e $(f'_n)_n$ sono di Cauchy in $\mathcal{C}(I, \mathbb{R})$, per cui convergono uniformemente. Siano $f = \lim_n f_n$ e $g = \lim_n f'_n$. Per il teorema della sezione precedente, si ha che $f \in \mathcal{C}^1(I)$ e $f' = g$.

In generale, si definisce su $\mathcal{C}^m(I)$ la norma

$$\|f\|_m = \sum_{j=1}^{m} \|f^{(j)}\|_{\mathcal{B}} .$$

Si dimostra per induzione che tutti questi sono spazi di Banach.

15 Serie in spazi di Banach

Sia V uno spazio vettoriale normato. Indicheremo con $\| \cdot \|$ la sua norma. V risulta quindi uno spazio metrico, con la distanza definita, come al solito, da $d(x, x') = \|x - x'\|$. Data una successione $(a_k)_k$ in V chiameremo "serie" (ad essa associata) la successione $(s_n)_n$ così definita:

$$\begin{aligned}
&s_0 = a_0 \, , \\
&s_1 = a_0 + a_1 \, , \\
&s_2 = a_0 + a_1 + a_2 \, , \\
&\;\dots \\
&s_n = a_0 + a_1 + a_2 + \ldots + a_n \, , \\
&\;\dots
\end{aligned}$$

L'elemento a_k si dice "termine $k-$esimo", mentre $s_n = \sum_{k=0}^{n} a_k$ si dice "somma parziale $n-$esima" della serie. Nel caso in cui esiste il limite di $(s_n)_n$ in V, si dice che la serie "converge". In tal caso, il limite $S = \lim_n s_n$ si dice "somma della serie" e si scrive

$$S = \lim_n \left(\sum_{k=0}^{n} a_k \right) = \sum_{k=0}^{\infty} a_k \, ,$$

o talvolta anche

$$S = a_0 + a_1 + a_2 + \ldots + a_n + \ldots$$

Con un abuso di notazioni, si usa spesso indicare la serie stessa con i simboli

$$\sum_{k=0}^{\infty} a_k \, , \qquad \text{oppure} \qquad a_0 + a_1 + a_2 + \ldots + a_n + \ldots .$$

Talvolta, per brevità, scriveremo semplicemente $\sum_k a_k$.

Raramente si ha la fortuna di riuscire a calcolare la somma di una serie, se questa esiste. Molto spesso, ci si accontenta di stabilire se la serie converge o meno.

Teorema. *Se una serie* $\sum_k a_k$ *converge, allora*

$$\lim_n a_n = 0\,.$$

<u>Dimostrazione</u>. Sia $\lim_n s_n = S \in V$. Allora anche $\lim_n s_{n-1} = S$, da cui

$$\lim_n a_n = \lim_n (s_n - s_{n-1}) = \lim_n s_n - \lim_n s_{n-1} = S - S = 0\,.$$

■

Teorema. *Supponiamo che le due serie* $\sum_k a_k$ *e* $\sum_k b_k$ *convergano e abbiano somma* A *e* B*, rispettivamente. Allora converge anche la serie* $\sum_k (a_k + b_k)$ *e la sua somma vale* $A+B$*. Inoltre, per ogni scalare fissato* α*, converge anche la serie* $\sum_k (\alpha a_k)$ *e la sua somma vale* αA*. Scriveremo brevemente*

$$\sum_{k=0}^{\infty} (a_k + b_k) = \sum_{k=0}^{\infty} a_k + \sum_{k=0}^{\infty} b_k\,, \qquad \sum_{k=0}^{\infty} (\alpha a_k) = \alpha \sum_{k=0}^{\infty} a_k\,.$$

<u>Dimostrazione</u>. Siano $s_n = \sum_{k=0}^n a_k$ e $s'_n = \sum_{k=0}^n b_k$. Allora

$$s_n + s'_n = \sum_{k=0}^{n} (a_k + b_k)\,, \qquad \alpha s_n = \sum_{k=0}^{n} (\alpha a_k)\,,$$

da cui segue la tesi. ■

Vediamo come si adatta alle serie il **criterio di Cauchy**.

Teorema. *Se* V *è uno spazio di Banach, la serie* $\sum_k a_k$ *converge se e solo se*

$$\forall \varepsilon > 0 \quad \exists \bar{n} : \quad n > m \geq \bar{n} \quad \Rightarrow \quad \left\| \sum_{k=m+1}^{n} a_k \right\| < \varepsilon\,.$$

<u>Dimostrazione</u>. Essendo V completo, abbiamo che la successione $(s_n)_n$ ha limite finito se e solo se è di Cauchy, ossia

$$\forall \varepsilon > 0 \quad \exists \bar{n} : \quad [\, m \geq \bar{n} \text{ e } n \geq \bar{n}\,] \quad \Rightarrow \quad \|s_n - s_m\| < \varepsilon\,.$$

Non essendo restrittivo supporre $m < n$, sostituendo $s_n = \sum_{k=0}^n a_k$ e $s_m = \sum_{k=0}^m a_k$ si ha la tesi. ■

Dimostreremo il seguente **criterio della convergenza totale**.

Teorema. *Se* V *è uno spazio di Banach e la serie a termini reali* $\sum_k \|a_k\|$ *converge, allora anche la serie* $\sum_k a_k$ *converge.*

In tal caso, si dice che la serie $\sum_k a_k$ "converge totalmente".

Dimostrazione. Supponiamo che la serie $\sum_k \|a_k\|$ converga. Fissiamo un $\varepsilon > 0$. Per il criterio di Cauchy, esiste un $\bar{n} \in \mathbb{N}$ tale che

$$n > m \geq \bar{n} \quad \Rightarrow \quad \sum_{k=m+1}^{n} \|a_k\| < \varepsilon\,.$$

Essendo

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

si ha che

$$n > m \geq \bar{n} \quad \Rightarrow \quad \left\| \sum_{k=m+1}^{n} a_k \right\| < \varepsilon\,.$$

Quindi, $(s_n)_n$ è di Cauchy. Siccome V è completo, si ha che $(s_n)_n$ ha limite in V, per cui la serie $\sum_k a_k$ converge. ■

La convergenza totale di una serie in V ci riporta allo studio di una serie a termini reali positivi. Vediamo di richiamare alcuni criteri di convergenza visti per le serie a termini reali. [6] È molto utile il seguente **criterio del confronto**.

Teorema. *Siano $\sum_k a_k$ una serie nello spazio di Banach V e $\sum_k b_k$ una serie a termini reali, tali che*

$$\exists \bar{k} : \quad k \geq \bar{k} \quad \Rightarrow \quad \|a_k\| \leq b_k\,.$$

Se la serie $\sum_k b_k$ converge, allora converge anche la serie $\sum_k a_k$.

Dimostrazione. Poniamo

$$s_n = \|a_0\| + \|a_1\| + \|a_2\| + \ldots + \|a_n\|\,, \qquad s'_n = b_0 + b_1 + b_2 + \ldots + b_n\,.$$

Se la serie $\sum_k b_k$ converge, esiste il limite

$$\lim_n s'_n = S' \in \mathbb{R}\,.$$

Usiamo ora la seguente considerazione generale:

Se modifichiamo solo un numero finito di termini di una serie, otteniamo una nuova serie con la seguente proprietà: essa converge se e solo se converge la serie di partenza.

Possiamo quindi modificare un numero finito di termini nelle due serie e supporre che sia $\|a_k\| \leq b_k$, per ogni k. Allora anche

$$s_n \leq s'_n\,,$$

per ogni n. Inoltre, la successione $(s_n)_n$ è crescente e pertanto ha limite. Da quanto sopra, $\lim_n s_n \leq S'$, per cui la serie $\sum_k \|a_k\|$ converge. Quindi, la serie $\sum_k a_k$ converge totalmente. ■

[6] Delle note che seguono si è solo accennato a lezione.

Vediamo il **criterio della radice**.

Corollario 1. *Sia* $\sum_k a_k$ *una serie nello spazio di Banach* V *per cui si abbia*

$$\limsup_k \sqrt[k]{\|a_k\|} < 1\,.$$

Allora, la serie converge.

Dimostrazione. sia $L = \limsup_k \sqrt[k]{\|a_k\|}$. Fissiamo un $\alpha \in]L, 1[$. Allora esiste un $\bar{k}$ tale che

$$k \geq \bar{k} \quad \Rightarrow \quad \sqrt[k]{\|a_k\|} \leq \alpha\,,$$

ossia

$$k \geq \bar{k} \quad \Rightarrow \quad \|a_k\| \leq \alpha^k\,,$$

La conclusione segue per confronto con la serie geometrica di base α, che converge, essendo $0 < \alpha < 1$. ■

Ricordiamo ora il seguente fatto generale.

Proposizione. *Data una successione* $(\alpha_k)_k$ *di numeri reali positivi, si ha sempre*

$$\liminf_k \frac{\alpha_{k+1}}{\alpha_k} \leq \liminf_k \sqrt[k]{\alpha_k} \leq \limsup_k \sqrt[k]{\alpha_k} \leq \limsup_k \frac{\alpha_{k+1}}{\alpha_k}\,.$$

Il seguente corollario ci fornisce il cosiddetto **criterio del rapporto**.

Corollario 2. *Sia* $\sum_k a_k$ *una serie nello spazio di Banach* V*, a termini non nulli, per cui si abbia*

$$\limsup_k \frac{\|a_{k+1}\|}{\|a_k\|} < 1\,.$$

Allora, la serie converge.

16 Serie a termini complessi e serie di funzioni

Se V è l'insieme $\mathbb{C}$ dei numeri complessi, lo si identifica con $\mathbb{R}^2$, che è dotato della norma euclidea. Quindi, se $z = a + ib$, la sua norma coincide con il suo modulo complesso $|z| = \sqrt{a^2 + b^2}$.

Data una serie a termini complessi $a_k = x_k + iy_k$, possiamo scrivere la sua somma parziale in questo modo:

$$s_n = \sum_{k=0}^{n} (x_k + iy_k) = \sum_{k=0}^{n} x_k + i \sum_{k=0}^{n} y_k = \sigma_n + i\tau_n$$

dove σ_n e τ_n sono la parte reale e la parte immaginaria di s_n. Abbiamo quindi una successione $(\sigma_n, \tau_n)_n$ in $\mathbb{R}^2$. Ricordando che una tale successione ha limite in $\mathbb{R}^2$ se e solo se le sue componenti hanno limite in $\mathbb{R}$, avremo il risultato seguente.

Teorema. *Se* $a_k = x_k + iy_k$, *con* x_k *e* y_k *reali, la serie* $\sum_k a_k$ *converge se e solo se convergono le serie* $\sum_k x_k$ *e* $\sum_k y_k$. *In tal caso, la somma della serie è data da*

$$\sum_{k=0}^{\infty} a_k = \sum_{k=0}^{\infty} x_k + i \sum_{k=0}^{\infty} y_k \,.$$

Esempio. Consideriamo la serie

$$\sum_{k=0}^{\infty} \frac{i^k}{k+1} = 1 + \frac{i}{2} - \frac{1}{3} - \frac{i}{4} + \frac{1}{5} + \frac{i}{6} - \frac{1}{7} - \frac{i}{8} + \ldots + \frac{i^n}{n+1} + \ldots \,.$$

La serie delle parti reali è

$$1 - \frac{1}{3} + \frac{1}{5} - \frac{1}{7} + \ldots \,,$$

mentre quella delle parti immaginarie è

$$\frac{1}{2} - \frac{1}{4} + \frac{1}{6} - \frac{1}{8} + \ldots \,.$$

Entrambe convergono, per il criterio di Leibniz sulle serie reali a segni alternati. Quindi anche la serie di partenza converge.

Nel caso delle serie a termini complessi, invece di “convergenza totale” si preferisce parlare di “convergenza assoluta”, come nel caso delle serie a termini reali. Nel caso della serie dell'esempio precedente, si noti che la serie dei moduli non converge:

$$\sum_{k=0}^{\infty} \left| \frac{i^k}{k+1} \right| = \sum_{k=0}^{\infty} \frac{1}{k+1} \,;$$

è la serie armonica, che diverge a $+\infty$. Quindi, in questo caso la serie data non converge assolutamente.

Supponiamo ora che sia $V = \mathcal{C}(E, E')$, dove E è uno spazio metrico ed E' è uno spazio di Banach. Ricordiamo che, in tal caso, V è anch'esso uno spazio di Banach, con la norma $\| \cdot \|_{\mathcal{B}}$. I termini di una serie in V sono delle funzioni continue, che indicheremo con $f_k : E \to E'$; avremo quindi la serie

$$\sum_{k=0}^{\infty} f_k \,.$$

Se la serie converge in V e ha per somma una funzione $f : E \to E'$, si dice che la serie “converge uniformemente” a f. Ricordiamo che ciò significa, in simboli,

$$\forall \varepsilon > 0 \quad \exists \bar{n} \in \mathbb{N} : \quad \forall x \in E \quad n \geq \bar{n} \;\Rightarrow\; \left\| \sum_{k=0}^{n} f_k(x) - f(x) \right\| < \varepsilon \,.$$

Se questo accade per una certa funzione $f : E \to E'$, allora necessariamente anche f deve essere continua.

Esempio. Sia $E = [a,b] \subseteq \mathbb{R}$ ed $E' = \mathbb{R}$. Consideriamo la serie

$$\sum_{k=1}^{\infty} \frac{1}{k^2} \sin(e^{3kx-1} + \arctan(x^2 + \sqrt{k}\,))\,.$$

Studiamo la serie delle norme in $\mathcal{C}([a,b])$: si ha

$$\sup\left\{\left|\frac{1}{k^2} \sin(e^{3kx-1} + \arctan(x^2 + \sqrt{k}\,))\right| : x \in [a,b]\right\} \leq \frac{1}{k^2}\,,$$

e la serie a termini reali $\sum_k \frac{1}{k^2}$ converge. Essendo $[a,b]$ compatto, $\mathcal{C}([a,b],\mathbb{R})$ è uno spazio di Banach. Pertanto, la serie data converge totalmente.

17 Serie di potenze

Un esempio importante di serie di funzioni si ha prendendo $E = \mathbb{C}$, $E' = \mathbb{C}$ e le funzioni $f_k : \mathbb{C} \to \mathbb{C}$ definite da $f_k(x) = a_k x^k$, per un certo $a_k \in \mathbb{C}$. Si ottiene così la "serie di potenze"

$$(SP) \qquad \sum_{k=0}^{\infty} a_k x^k\,,$$

dove gli a_k e x sono numeri complessi.

Teorema. *Sia*

$$L = \limsup_k \sqrt[k]{|a_k|}\,.$$

Allora

(i) se $L = +\infty$, la serie converge solo se $x = 0$;

(ii) se $L = 0$, la serie converge per ogni $x \in \mathbb{C}$;

(iii) se $L \in\,]0,+\infty[$, la serie $\begin{cases} \text{converge se } |x| < \frac{1}{L}\,, \\ \text{non converge se } |x| > \frac{1}{L}\,. \end{cases}$

Dimostrazione. Se $L = +\infty$ e $x \neq 0$, allora $\sqrt[k]{|a_k|} > \frac{1}{|x|}$ per infiniti k, quindi

$$|a_k x^k| = (\sqrt[k]{|a_k|}|x|)^k > 1\,, \qquad \text{per infiniti } k\,.$$

Se la serie convergesse, dovrebbe essere $\lim_k a_k x^k = 0$, mentre così non è. Pertanto, se $L = +\infty$, la serie converge solo se $x = 0$.

Se $L = 0$, prendendo un qualsiasi $x \in \mathbb{C}$, si ha

$$\limsup_k \sqrt[k]{|a_k x^k|} = |x| \limsup_k \sqrt[k]{|a_k|} = 0\,,$$

e la serie converge assolutamente, per il criterio della radice.

Supponiamo ora $L \in \,]0,+\infty[\,$. Se $|x| < \frac{1}{L}$, allora

$$\limsup_k \sqrt[k]{|a_k x^k|} = |x| \limsup_k \sqrt[k]{|a_k|} = |x|\, L < 1\,,$$

e la serie converge, per il criterio della radice. Se invece $|x| > \frac{1}{L}$, ossia $L > \frac{1}{|x|}$, allora $\sqrt[k]{|a_k|} > \frac{1}{|x|}$ per infiniti k, quindi

$$|a_k x^k| = (\sqrt[k]{|a_k|}|x|)^k > 1\,, \qquad \text{per infiniti } k\,.$$

Se la serie convergesse, dovrebbe essere $\lim_k a_k x^k = 0$, mentre così non è. ■

Corollario. *Se esiste il limite*

$$L = \lim_k \frac{|a_{k+1}|}{|a_k|}\,,$$

si ha la stessa conclusione del teorema precedente.

Dimostrazione. In questo caso si ha che $\lim_k \sqrt[k]{|a_k|} = \lim_k \frac{|a_{k+1}|}{|a_k|}$. ■

Si definisce il "raggio di convergenza" r della serie (SP):

$$r = \begin{cases} 0 & \text{se } L = +\infty\,, \\[2ex] +\infty & \text{se } L = 0\,, \\[2ex] \dfrac{1}{L} & \text{se } L \in \,]0,+\infty[\,. \end{cases}$$

Se $r > 0$, diremo che $B(0,r)$ è il suo "cerchio di convergenza". Si tratta di un cerchio, privato della circonferenza esterna. Nel caso in cui $r = +\infty$, poniamo $B(0,r) = \mathbb{C}$.

Abbiamo dimostrato che la serie di potenze converge per ogni $x \in B(0,r)$. Purtroppo, tale convergenza potrebbe non essere uniforme. Vediamo ora che si ha la convergenza uniforme se ci si restringe a un cerchio più piccolo.

Teorema. *Se $r > 0$, allora, per ogni $\rho \in \,]0,r[\,$, la serie di potenze (SP) converge totalmente, e quindi uniformemente, su $\bar{B}(0,\rho)$.*

Dimostrazione. Sia $\rho \in \,]0,r[\,$. Allora

$$\sup\{|a_k x^k| : |x| \le \rho\} \le |a_k|\rho^k$$

e siccome

$$\limsup_k \sqrt[k]{|a_k|\rho^k} = \rho \limsup_k \sqrt[k]{|a_k|} < 1\,,$$

la serie $\sum_k |a_k|\rho^k$ converge, per il criterio della radice. Quindi, la serie converge totalmente su $\bar{B}(0,\rho)$. ■

Corollario. Se $r > 0$ e $f : B(0,r) \to \mathbb{C}$ è la somma della serie,

$$f(x) = \sum_{k=0}^{\infty} a_k x^k\,,$$

allora f è continua in $B(0,r)$.

Dimostrazione. Dal teorema precedente, per ogni $\rho \in]0,r[$ la convergenza è uniforme su $\bar{B}(0,\rho)$, per cui f è continua su $\bar{B}(0,\rho)$. Per l'arbitrarietà di $\rho \in]0,r[$ si vede allora che f è continua su tutto $B(0,r)$. ∎

18 Serie di Taylor (a termini reali)

Consideriamo ancora la serie di potenze

$$(SP) \qquad \sum_{k=0}^{\infty} a_k x^k\,,$$

dove però stavolta tutti gli a_k e anche x sono numeri reali. In questo caso, prendiamo $E = \mathbb{R}$, $E' = \mathbb{R}$ e le funzioni $f_k : \mathbb{R} \to \mathbb{R}$ definite da $f_k(x) = a_k x^k$. Il cerchio di convergenza $B(0,r)$, in questo caso, si riduce all'intervallo $]-r,r[$. Se $r = +\infty$, è tutto $\mathbb{R}$.

Siccome le $f_k : \mathbb{R} \to \mathbb{R}$ sono funzioni derivabili, ci chiediamo se, nel caso in cui il raggio di convergenza sia non nullo, la funzione che si ottiene come somma della serie di potenze (SP) sia essa stessa derivabile. A tal scopo, abbiamo bisogno di un risultato preliminare.

Teorema. *Data una serie $\sum_k f_k$, i cui termini $f_k : [a,b] \to \mathbb{R}$ sono funzioni di classe C^1, se essa converge uniformemente con $\sum_k f_k = f$ e la serie delle derivate converge uniformemente con $\sum_k f_k' = g$, allora f è di classe C^1 con derivata $f' = g$.*

Dimostrazione. Consideriamo le successioni delle somme parziali $s_n = \sum_{k=0}^{n} f_k$ e $s_n' = \sum_{k=0}^{n} f_k'$. Quindi $(s_n)_n$ è una successione in $\mathcal{C}^1([a,b])$ e si ha $\lim_n s_n = f$ e $\lim_n s_n' = g$, uniformemente. Come abbiamo visto in precedenza, ne segue che $f \in \mathcal{C}^1([a,b])$ e $f' = g$. ∎

Iterando il ragionamento, possiamo generalizzare il teorema precedente.

Teorema. *Data una serie $\sum_k f_k$, i cui termini $f_k : [a,b] \to \mathbb{R}$ sono funzioni di classe C^m, se le seguenti serie convergono uniformemente con le rispettive somme*

$$\sum_k f_k = f\,, \quad \sum_k f_k' = g_1\,, \quad \sum_k f_k'' = g_2\,,\ \dots\,,\ \sum_k f_k^{(m)} = g_m\,,$$

allora f è di classe C^m e si ha:

$$f' = g_1\,, \quad f'' = g_2\,,\ \dots\,,\ f^{(m)} = g_m\,.$$

Possiamo ora affrontare il problema della derivabilità della funzione somma della serie di potenze (SP).

Teorema. *Sia $r > 0$ il raggio di convergenza della serie di potenze (SP) a termini reali, e sia*

$$\sum_{k=0}^{\infty} a_k x^k = f(x)\,, \qquad \textit{per ogni } x \in \,]-r,r[\,.$$

La serie

$$\sum_{k=1}^{\infty} k a_k x^{k-1}$$

ha come raggio di convergenza lo stesso r. Inoltre, la funzione $f :]-r,r[\to \mathbb{R}$ è derivabile e si ha:

$$f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}\,, \qquad \text{per ogni } x \in \,]-r,r[\,.$$

<u>Dimostrazione</u>. Siccome

$$\limsup_k \sqrt[k]{|ka_k|} = \lim_k \sqrt[k]{k}\ \limsup_k \sqrt[k]{|a_k|} = \limsup_k \sqrt[k]{|a_k|}\,,$$

si vede che il raggio di convergenza è lo stesso r, e possiamo definire $g(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$, per $x \in \,]-r,r[\,$. Fissato $\rho \in]0,r[$ sappiamo che si ha la convergenza uniforme su $[-\rho,\rho]$. Ponendo $f_k(x) = a_k x^k$, abbiamo che $f = \sum_k f_k$ e $g = \sum_k f_k'$, uniformemente entrambe su $[-\rho,\rho]$. Per un teorema precedente, abbiamo che f è derivabile e $f'(x) = g(x)$, per ogni $x \in [-\rho,\rho]$. Ne segue la conclusione, essendo $\rho \in]0,r[$ arbitrario. ■

Iterando il procedimento, si dimostra la generalizzazione seguente.

Teorema. *Sia $r > 0$ il raggio di convergenza della serie di potenze (SP) a termini reali, e sia*

$$\sum_{k=0}^{\infty} a_k x^k = f(x)\,, \qquad \textit{per ogni } x \in \,]-r,r[\,.$$

Le serie

$$\sum_{k=1}^{\infty} k a_k x^{k-1}\,, \quad \sum_{k=2}^{\infty} k(k-1) a_k x^{k-2}\,, \quad \sum_{k=3}^{\infty} k(k-1)(k-2) a_k x^{k-3}\,, \quad \dots \quad ,$$

$$\sum_{k=m}^{\infty} k(k-1)(k-2)\cdots(k-m+1) a_k x^{k-m}\,, \quad \dots$$

hanno tutte come raggio di convergenza lo stesso r. Inoltre, la funzione $f :]-r,r[\to \mathbb{R}$ è derivabile infinite volte e, per ogni $x \in \,]-r,r[\,$, si ha:

$$f^{(j)}(x) = \sum_{k=j}^{\infty} k(k-1)(k-2)\cdots(k-j+1) a_k x^{k-j}\,.$$

A questo punto, se usiamo la formula ora scritta prendendo $x = 0$, otteniamo

$$f^{(j)}(0) = j!\, a_j\,,$$

per ogni $j \in \mathbb{N}$ (ricordo che $f^{(0)} = f$). Quindi,

$$f(x) = \sum_{k=0}^{\infty} a_k x^k = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(0) x^k\,.$$

Quest'ultima è la "serie di Taylor" associata alla funzione f nel punto $x_0 = 0$. Abbiamo così dimostrato che una serie di potenze con raggio di convergenza $r > 0$ definisce una funzione f, che è analitica su $]-r, r[$.

Ridordiamo, ad esempio, che la funzione esponenziale è analitica su tutto $\mathbb{R}$, e si ha

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!}\,.$$

Notiamo infine che tutta la teoria può essere svolta più in generale per le serie del tipo

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k\,,$$

semplicemente operando la sostituzione $u = x - x_0$. Il cerchio di convergenza sarà del tipo $B(x_0, r)$, e se la serie è a termini reali e $f(x)$ ne è la somma, la serie di Taylor è la seguente:

$$f(x) = \sum_{k=0}^{\infty} \frac{1}{k!} f^{(k)}(x_0)(x - x_0)^k\,.$$

19 La funzione esponenziale complessa

Date due serie $\sum_{k=0}^{\infty} a_k$ e $\sum_{k=0}^{\infty} b_k$, si definisce la serie "prodotto alla Cauchy" in questo modo:

$$\sum_{k=0}^{\infty} \left(\sum_{j=0}^{k} a_{k-j} b_j \right)\,.$$

Enunciamo il seguente importante risultato.

Teorema (di Mertens). *Se la serie $\sum_{k=0}^{\infty} a_k$ converge e ha somma A, la serie $\sum_{k=0}^{\infty} b_k$ converge e ha somma B e almeno una delle due converge assolutamente, allora la serie prodotto alla Cauchy converge e ha somma AB.*

<u>Dimostrazione</u>. [7] Supponiamo, per fissare le idee, che sia la serie $\sum_{k=0}^{\infty} a_k$ a convergere assolutamente. Indichiamo con $c_k = \sum_{j=0}^{k} a_{k-j} b_j$, il termine

[7] Dimostrazione non svolta a lezione.

$k-$esimo della serie prodotto. Siano $s_n = \sum_{k=0}^n a_k$, $s'_n = \sum_{k=0}^n b_k$ e $s''_n = \sum_{k=0}^n c_k$. Sia inoltre $r'_n = B - s'_n$. Allora

$$\begin{aligned} s''_n &= a_0b_0 + (a_1b_0 + a_0b_1) + \ldots + (a_nb_0 + a_{n-1}b_1 + \ldots + a_1b_{n-1} + a_0b_n) \\ &= a_0s'_n + a_1s'_{n-1} + \ldots + a_{n-1}s'_1 + a_ns'_0 \\ &= a_0(B - r'_n) + a_1(B - r'_{n-1}) + \ldots + a_{n-1}(B - r'_1) + a_n(B - r'_0) \\ &= s_nB - (a_0r'_n + a_1r'_{n-1} + \ldots + a_{n-1}r'_1 + a_nr'_0)\,. \end{aligned}$$

Siccome $\lim_n s_nB = AB$, la tesi sarà dimostrata se

$$\lim_n\,(a_0r'_n + a_1r'_{n-1} + \ldots + a_{n-1}r'_1 + a_nr'_0) = 0\,.$$

Fissiamo $\varepsilon > 0$. Essendo $\lim_n r'_n = 0$, esiste un $\bar n_1$ tale che

$$n \geq \bar n_1 \quad \Rightarrow \quad |r'_n| < \varepsilon\,.$$

Poniamo inoltre $\bar R = \max\{|r'_n| : n \in \mathbb{N}\}$. Per ipotesi, la serie $\sum_{k=0}^\infty |a_k|$ converge: sia $\bar A$ la sua somma. Per il criterio di Cauchy esiste un $\bar n_2$ tale che $\bar n_2 \geq \bar n_1$ e

$$n \geq \bar n_2 \quad \Rightarrow \quad |a_{n-\bar n_1+1}| + |a_{n-\bar n_1+2}| + \ldots + |a_n| < \varepsilon\,.$$

Allora, se $n \geq \bar n_2$,

$$\begin{aligned} |a_0r'_n +\, & a_1r'_{n-1} + \ldots + a_{n-1}r'_1 + a_nr'_0| \leq \\ &\leq |a_0|\,|r'_n| + \ldots + |a_{n-\bar n_1}|\,|r'_{\bar n_1}| + |a_{n-\bar n_1+1}|\,|r'_{\bar n_1-1}| + \ldots + |a_n|\,|r'_0| \\ &\leq \varepsilon(|a_0| + \ldots + |a_{n-\bar n_1}|) + \bar R(|a_{n-\bar n_1+1}| + \ldots + |a_n|) \\ &\leq \varepsilon\bar A + \bar R\varepsilon \\ &= (\bar A + \bar R)\varepsilon\,, \end{aligned}$$

il che completa la dimostrazione. ■

Per $z \in \mathbb{C}$, consideriamo la serie

$$\sum_{k=0}^\infty \frac{z^k}{k!}\,.$$

Si tratta di una serie di potenze. Vediamo che converge assolutamente: si ha che

$$\sum_{k=0}^\infty \left|\frac{z^k}{k!}\right| = \sum_{k=0}^\infty \frac{|z|^k}{k!} = e^{|z|}\,.$$

È pertanto possibile definire una funzione $f : \mathbb{C} \to \mathbb{C}$ in questo modo:

$$f(z) = \sum_{k=0}^\infty \frac{z^k}{k!}\,.$$

Teorema. *Per ogni z_1, z_2 si ha*

$$f(z_1 + z_2) = f(z_1)f(z_2)\,.$$

Dimostrazione. Le serie $\sum_{k=0}^{\infty} \frac{z_1^k}{k!}$ e $\sum_{k=0}^{\infty} \frac{z_2^k}{k!}$ convergono assolutamente e hanno per somma $f(z_1)$ e $f(z_2)$, rispettivamente. Per il teorema di Mertens, la serie prodotto alla Cauchy converge e ha per somma $f(z_1)f(z_2)$. Ma il prodotto alla Cauchy è la serie

$$\sum_{k=0}^{\infty} \left(\sum_{j=0}^{k} \frac{z_1^{k-j}}{(k-j)!} \frac{z_2^j}{j!} \right) = \sum_{k=0}^{\infty} \frac{1}{k!} \left(\sum_{j=0}^{k} \binom{k}{j} z_1^{k-j} z_2^j \right) = \sum_{k=0}^{\infty} \frac{(z_1 + z_2)^k}{k!}\,,$$

che ha per somma $f(z_1 + z_2)$. Da qui la tesi. ■

Questa proprietà ci porta a chiamare la funzione f "esponenziale complessa": invece di $f(z)$ scriveremo $\exp(z)$ o anche e^z:

$$e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}\,.$$

La proprietà dimostrata sopra si può allora scrivere in questo modo: se z_1 e z_2 sono due numeri complessi,

$$e^{z_1+z_2} = e^{z_1} e^{z_2}\,.$$

Sia ora $z = a + ib$. Allora

$$e^{a+ib} = e^a e^{ib}\,;$$

vediamo quest'ultimo:

$$\begin{aligned} e^{ib} &= 1 + ib + \frac{(ib)^2}{2!} + \frac{(ib)^3}{3!} + \frac{(ib)^4}{4!} + \frac{(ib)^5}{5!} + \frac{(ib)^6}{6!} + \frac{(ib)^7}{7!} + \dots \\ &= 1 + ib - \frac{b^2}{2!} - i\frac{b^3}{3!} + \frac{b^4}{4!} + i\frac{b^5}{5!} - \frac{b^6}{6!} - i\frac{b^7}{7!} + \dots \\ &= \left(1 - \frac{b^2}{2!} + \frac{b^4}{4!} - \frac{b^6}{6!} + \dots\right) + i\left(b - \frac{b^3}{3!} + \frac{b^5}{5!} - \frac{b^7}{7!} + \dots\right) \\ &= \cos b + i \sin b\,. \end{aligned}$$

Abbiamo così la **formula di Eulero**

$$e^{a+ib} = e^a(\cos b + i \sin b)\,.$$

Si possono allora verificare le seguenti uguaglianze: per ogni $t \in \mathbb{R}$,

$$\cos t = \frac{e^{it} + e^{-it}}{2}\,, \quad \sin t = \frac{e^{it} - e^{-it}}{2i}\,.$$

Queste formule possono essere utilizzate, ad esempio, per estendere anche le funzioni trigonometriche al campo complesso. Anche le funzioni iperboliche, definite da

$$\cosh(x) = \frac{e^x + e^{-x}}{2}, \qquad \sinh(x) = \frac{e^x - e^{-x}}{2},$$

si possono estendere con le stesse formule al campo complesso. Si noti che

$$\cos t = \cosh(it), \qquad \sin t = -i \sinh(it).$$

A questo punto risulteranno finalmente spiegate le similitudini incontrate tra le funzioni trigonometriche e quelle iperboliche.

Il fatto che, per ogni $z \in \mathbb{C}$,

$$e^{z+2\pi i} = e^z e^{2\pi i} = e^z,$$

si può interpretare dicendo che la funzione esponenziale complessa è periodica di periodo $2\pi i$. Questo fatto compromette la possibile definizione di una funzione "logaritmo" nel campo complesso: dato $z \in \mathbb{C}$, con $z \neq 0$, l'equazione

$$e^u = z,$$

vista la periodicità della funzione esponenziale, presenta molteplici soluzioni. Precisamente, se scriviamo $z = \rho(\cos\theta + i \sin\theta)$, il numero complesso $u = x + iy$ ne è soluzione se e solo se

$$e^x(\cos y + i \sin y) = \rho(\cos\theta + i \sin\theta),$$

ossia

$$x = \ln\rho, \quad y = \theta + 2\pi k,$$

con $k \in \mathbb{Z}$. Pertanto,

$$e^u = z \quad \Leftrightarrow \quad u \in \{\ln|z| + i(Arg(z) + 2\pi k) : k \in \mathbb{Z}\}.$$

Talvolta si interpreta il "logaritmo complesso" come una "funzione multivoca" che assume in questo caso infiniti valori, riservando il nome di "logaritmo principale" al particolare valore ottenuto scegliendo $k = 0$.

Ad esempio, il logaritmo complesso del numero i assume tutti i valori dell'insieme

$$\left\{ i\left(\frac{\pi}{2} + 2\pi k\right) : k \in \mathbb{Z} \right\}.$$

Il logaritmo principale di i vale pertanto $\frac{\pi}{2} i$.

20 Serie di Fourier

Studieremo ora i seguenti "polinomi trigonometrici":

$$c_0 + \sum_{k=1}^{n} \left(a_k \cos\left(\frac{2\pi k}{T} t\right) + b_k \sin\left(\frac{2\pi k}{T} t\right) \right),$$

dove c_0, a_k e b_k sono dei numeri reali. Si tratta di funzioni $T-$periodiche. Ci interessa il problema dell'eventuale convergenza della serie che si ottiene facendo tendere n all'infinito.

Teorema. *Supponiamo che la serie converga uniformemente su $[0,T]$ a una certa funzione f. Allora f si può estendere per T-periodicità a una funzione continua su tutto $\mathbb{R}$, e si ha*

$$f(t) = c_0 + \sum_{k=1}^{\infty} \left(a_k \cos\left(\frac{2\pi k}{T} t\right) + b_k \sin\left(\frac{2\pi k}{T} t\right)\right),$$

uniformemente per ogni $t \in \mathbb{R}$. Inoltre, i coefficienti c_0, a_k e b_k sono necessariamente i seguenti:

$$c_0 = \frac{1}{T}\int_0^T f(t)\,dt\,,$$

$$a_k = \frac{2}{T}\int_0^T f(t)\cos\left(\frac{2\pi k}{T}t\right)\,dt\,, \quad \text{e} \quad b_k = \frac{2}{T}\int_0^T f(t)\sin\left(\frac{2\pi k}{T}t\right)\,dt\,.$$

<u>Dimostrazione</u>. La prima parte dell'enunciato segue dal fatto che tutte le funzioni nella serie sono continue e T-periodiche. Vediamo la seconda parte. Usando la convergenza uniforme, abbiamo che

$$\begin{aligned}\int_0^T f(t)\,dt &= \int_0^T \left(c_0 + \sum_{k=1}^{\infty} \left(a_k \cos\left(\frac{2\pi k}{T} t\right) + b_k \sin\left(\frac{2\pi k}{T} t\right)\right)\right) dt \\ &= c_0 T + \sum_{k=1}^{\infty} \int_0^T \left(a_k \cos\left(\frac{2\pi k}{T} t\right) + b_k \sin\left(\frac{2\pi k}{T} t\right)\right) dt = c_0 T\,,\end{aligned}$$

da cui la formula per c_0. Analogamente, per un intero $j \geq 1$,

$$\begin{aligned}&\int_0^T f(t)\cos\left(\frac{2\pi j}{T}t\right)\,dt = \\ &= \int_0^T \left(c_0 + \sum_{k=1}^{\infty} \left(a_k \cos\left(\frac{2\pi k}{T} t\right) + b_k \sin\left(\frac{2\pi k}{T} t\right)\right)\right)\cos\left(\frac{2\pi j}{T}t\right) dt \\ &= \sum_{k=1}^{\infty} \int_0^T \left(a_k \cos\left(\frac{2\pi k}{T} t\right) + b_k \sin\left(\frac{2\pi k}{T} t\right)\right)\cos\left(\frac{2\pi j}{T}t\right) dt\,.\end{aligned}$$

D'altra parte, usando per due volte la formula di integrazione per parti, si può verificare che, se $k \neq j$,

$$\int_0^T \sin\left(\frac{2\pi k}{T}t\right)\cos\left(\frac{2\pi j}{T}t\right)\,dt = 0\,, \quad \int_0^T \cos\left(\frac{2\pi k}{T}t\right)\cos\left(\frac{2\pi j}{T}t\right)\,dt = 0\,,$$

mentre, se $k = j$,

$$\int_0^T \sin\left(\frac{2\pi j}{T}t\right)\cos\left(\frac{2\pi j}{T}t\right)\,dt = 0\,, \quad \int_0^T \cos^2\left(\frac{2\pi j}{T}t\right)\,dt = \frac{T}{2}\,.$$

Quindi,

$$\int_0^T f(t)\cos\left(\frac{2\pi j}{T}t\right)\,dt = \frac{T}{2}a_j\,,$$

da cui la formula per a_j. Analogamente si vede che

$$\int_0^T f(t) \sin\left(\frac{2\pi j}{T} t\right) \, dt = \frac{T}{2} b_j \,,$$

da cui la formula per b_j. ■

Data una funzione continua $f : [0,T] \to \mathbb{R}$, si definiscono, per $k \in \mathbb{N}$, i "coefficienti di Fourier"

$$a_k = \frac{2}{T} \int_0^T f(t) \cos\left(\frac{2\pi k}{T} t\right) \, dt \,, \quad b_k = \frac{2}{T} \int_0^T f(t) \sin\left(\frac{2\pi k}{T} t\right) \, dt \,,$$

e la relativa "serie di Fourier"

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos\left(\frac{2\pi k}{T} t\right) + b_k \sin\left(\frac{2\pi k}{T} t\right) \right) \,.$$

Si pone immediatamente il problema della convergenza di tale serie. Purtroppo, anche se f è defnita su tutto $\mathbb{R}$ ed è T-periodica, la sua continuità non è sufficiente a garantirla. Come dimostrato da Du Bois-Raymond, ci sono delle funzioni continue e T-periodiche la cui serie di Fourier diverge per alcuni $t \in [0,T]$. In altri termini, se consideriamo le funzioni

$$f_n(t) = \frac{a_0}{2} + \sum_{k=1}^{n} \left(a_k \cos\left(\frac{2\pi k}{T} t\right) + b_k \sin\left(\frac{2\pi k}{T} t\right) \right) \,,$$

potrebbe essere che, per alcuni $t \in [0,T]$, la successione $(f_n(t))_n$ non abbia limite finito.

21 Il teorema di Fejer

Consideriamo allora le "medie"

$$\sigma_n(t) = \frac{1}{n+1} \left[f_0(t) + f_1(t) + \ldots + f_n(t) \right] \,.$$

Si potrebbe sperare che, pur non essendoci il limite di $(f_n(t))_n$, esista almeno il limite della successione $(\sigma_n(t))_n$. È quanto afferma il teorema seguente.

Teorema (di Fejer). *Se $f : \mathbb{R} \to \mathbb{R}$ è continua e T-periodica, allora*

$$\lim_n \sigma_n(t) = f(t) \,,$$

uniformemente in $t \in \mathbb{R}$.

Per semplificare le notazioni, introduciamo i seguenti coefficienti (complessi) di Fourier: per $k \in \mathbb{Z}$,

$$c_k = \frac{1}{T}\int_0^T f(t)e^{-i\frac{2\pi k}{T}t}\,dt\,.$$

Ponendo $b_0 = 0$, sii vede allora che

$$c_k = \begin{cases} \frac{1}{2}(a_{-k} + ib_{-k}) & \text{se } k < 0, \\ \\ \frac{1}{2}(a_k - ib_k) & \text{se } k \geq 0, \end{cases}$$

per cui

$$f_n(t) = \sum_{k=-n}^{n} c_k e^{i\frac{2\pi k}{T}t}$$

e

$$\begin{aligned}
\sigma_n(t) = \frac{1}{n+1}\Big[&c_0e^0 + (c_{-1}e^{-i\frac{2\pi}{T}t} + c_0e^0 + c_1e^{i\frac{2\pi}{T}t}) + \ldots + \\
&+(c_{-n}e^{-i\frac{2\pi n}{T}t} + \ldots + c_{-1}e^{-i\frac{2\pi}{T}t} + c_0e^0 + c_1e^{i\frac{2\pi}{T}t} + \ldots + c_ne^{i\frac{2\pi n}{T}t})\Big] \\
= \frac{1}{n+1}\Big[&(n+1)c_0e^0 + n(c_{-1}e^{-i\frac{2\pi}{T}t} + c_1e^{i\frac{2\pi}{T}t}) + \ldots + 1(c_{-n}e^{-i\frac{2\pi n}{T}t} + c_ne^{i\frac{2\pi n}{T}t})\Big] \\
= \frac{1}{n+1}&\sum_{k=-n}^{n}(n+1-|k|)c_ke^{i\frac{2\pi k}{T}t}\,.
\end{aligned}$$

Dimostrazione del Teorema di Fejer.[8] Per semplicità, considereremo il caso $T = 2\pi$. Questa ipotesi non sarà comunque restrittiva, potendocisi sempre ricondurre con un cambiamento di variabile. Abbiamo quindi

$$c_k = \frac{1}{2\pi}\int_0^{2\pi} f(t)e^{-ikt}\,dt\,, \quad f_n(t) = \sum_{k=-n}^{n} c_ke^{ikt}\,,$$

e

$$\begin{aligned}
\sigma_n(t) &= \frac{1}{n+1}\sum_{k=-n}^{n}(n+1-|k|)c_ke^{ikt} \\
&= \frac{1}{n+1}\sum_{k=-n}^{n}(n+1-|k|)\left(\frac{1}{2\pi}\int_0^{2\pi} f(s)e^{-iks}\,ds\right)e^{ikt} \\
&= \frac{1}{2\pi}\int_0^{2\pi} f(s)K_n(t-s)\,ds\,,
\end{aligned}$$

avendo posto

$$K_n(x) = \sum_{k=-n}^{n}\frac{n+1-|k|}{n+1}e^{ikx}\,.$$

[8] Dimostrazione non svolta a lezione.

Notiamo che, se $x = 0$ (o un multiplo intero di 2π),

$$\begin{aligned} K_n(0) &= \frac{1 + 2 + \ldots + n + (n+1) + n + \ldots + 2 + 1}{n+1} \\ &= \frac{n(n+1) + (n+1)}{n+1} = n + 1\,, \end{aligned}$$

mentre, se x non è un multiplo intero di 2π,

$$\begin{aligned} K_n(x) &= \frac{1}{n+1}\big[e^{-inx} + 2e^{-i(n-1)x} + \ldots + ne^{-ix} + (n+1)e^0 + \\ &\qquad +ne^{ix} + \ldots + 2e^{i(n-1)x} + e^{inx}\big] \\ &= \frac{1}{n+1}\big[e^{-inx}(1 + 2e^{ix} + \ldots + ne^{i(n-1)x}) + (n+1)e^0 + \\ &\qquad +e^{inx}(1 + 2e^{-ix} + \ldots + ne^{-i(n-1)x})\big]\ . \end{aligned}$$

Essendo, per $\alpha \neq 1$,

$$\sum_{k=1}^{n} k\alpha^{k-1} = \frac{1 - (n+1)\alpha^n + n\alpha^{n+1}}{(1-\alpha)^2}$$

(lo si può dimostrare per induzione), ponendo alternativamente $\alpha = e^{ix}$ oppure $\alpha = e^{-ix}$ otteniamo:

$$\begin{aligned} K_n(x) &= \frac{1}{n+1}\left(e^{-inx}\,\frac{1 - (n+1)e^{inx} + ne^{i(n+1)x}}{(1-e^{ix})^2} + (n+1)e^0 + \right. \\ &\qquad \left. +e^{inx}\,\frac{1 - (n+1)e^{-inx} + ne^{-i(n+1)x}}{(1-e^{-ix})^2}\right) \\ &= \frac{1}{n+1}\frac{e^{-inx} - 2e^{ix} + e^{i(n+2)x}}{(1-e^{ix})^2} = \frac{e^{-inx}}{n+1}\left(\frac{1 - e^{i(n+1)x}}{1 - e^{ix}}\right)^2 \\ &= \frac{1}{n+1}\left(\frac{e^{-\frac{i(n+1)x}{2}} - e^{\frac{i(n+1)x}{2}}}{e^{-\frac{ix}{2}} - e^{\frac{ix}{2}}}\right)^2 = \frac{1}{n+1}\left(\frac{\sin\left(\frac{(n+1)x}{2}\right)}{\sin\left(\frac{x}{2}\right)}\right)^2\ . \end{aligned}$$

Osserviamo che $K_n(x) \geq 0$ per ogni x. Inoltre, fissato un $\delta \in]0, 2\pi[$, abbiamo che, se $x \in]\delta, 2\pi - \delta[$,

$$K_n(x) \leq \frac{1}{n+1}\left(\frac{1}{\sin\left(\frac{x}{2}\right)}\right)^2 \leq \frac{1}{n+1}\left(\frac{1}{\sin\left(\frac{\delta}{2}\right)}\right)^2,$$

per cui $K_n(x)$ tende a 0 uniformemente in $]\delta, 2\pi - \delta[$. Osserviamo inoltre che

$$\int_0^{2\pi} K_n(x)\,dx = \sum_{k=-n}^{n} \frac{n+1-|k|}{n+1}\int_0^{2\pi} e^{ikx}\,dx = 2\pi\,,$$

indipendentemente da n. Ritornando a $\sigma_n(t)$, si ha:

$$\begin{aligned}\sigma_n(t) &= \frac{1}{2\pi}\int_0^{2\pi} f(s)K_n(t-s)\,ds \\ &= \frac{1}{2\pi}\int_t^{t+2\pi} f(t-x)K_n(x)\,dx \\ &= \frac{1}{2\pi}\int_0^{2\pi} f(t-x)K_n(x)\,dx\,.\end{aligned}$$

Dimostriamo che $\sigma_n(t)$ converge a $f(t)$ uniformemente. Fissiamo $\varepsilon > 0$. Essendo f continua e T-periodica, esiste un $M > 0$ tale che

$$|f(t)| \leq M\,, \quad \text{per ogni } t \in \mathbb{R}\,,$$

e f è uniformemente continua, per cui esiste un $\delta > 0$ tale che

$$|t - t'| < \delta \quad \Rightarrow \quad |f(t) - f(t')| < \frac{\varepsilon}{4}\,.$$

Inoltre, esiste un $\bar{n}$ tale che

$$n \geq \bar{n} \quad \Rightarrow \quad \forall x \in [\delta, 2\pi - \delta], \quad |K_n(x)| \leq \frac{\varepsilon}{4M}\,,$$

e quindi:

$$\begin{aligned}&|\sigma_n(t) - f(t)| = \left|\frac{1}{2\pi}\int_0^{2\pi} f(t-x)K_n(x)\,dx - \frac{1}{2\pi}\int_0^{2\pi} f(t)K_n(x)\,dx\right| \\ &= \frac{1}{2\pi}\left|\int_0^{2\pi}(f(t-x) - f(t))K_n(x)\,dx\right| \\ &= \frac{1}{2\pi}\left|\int_0^{\delta}\ldots\,dx + \int_\delta^{2\pi-\delta}\ldots\,dx + \int_{2\pi-\delta}^{2\pi}\ldots\,dx\right| \\ &\leq \frac{1}{2\pi}\left(\frac{\varepsilon}{4}\int_0^{\delta} K_n(x)\,dx + \frac{\varepsilon}{4M}\int_\delta^{2\pi-\delta}(|f(t-x)| + |f(t)|)\,dx + \frac{\varepsilon}{4}\int_\delta^{2\pi-\delta} K_n(x)\,dx\right) \\ &\leq \frac{\varepsilon}{4} + \frac{\varepsilon}{4M}2M + \frac{\varepsilon}{4} = \varepsilon\,.\end{aligned}$$

La dimostrazione è così completa. ∎

Ne segue immediatamente il seguente importante

Corollario. *Siano $f, \tilde{f} : \mathbb{R} \to \mathbb{R}$ due funzioni continue e T-periodiche. Se per i rispettivi coefficienti di Fourier c_k e $\tilde{c}_k$ si ha che $c_k = \tilde{c}_k$ per ogni k, allora f e $\tilde{f}$ coincidono.*

<u>Dimostrazione</u>. Con le ovvie notazioni, avremo che $\sigma_n(t) = \tilde{\sigma}_n(t)$, per ogni n, per cui

$$f(t) - \tilde{f}(t) = \lim_{n\to\infty}(\sigma_n(t) - \tilde{\sigma}_n(t)) = 0\,,$$

per ogni $t \in \mathbb{R}$. ∎

22 Convergenza della serie di Fourier

Ci interessiamo ora alla convergenza delle funzioni f_n. Useremo la notazione

$$\sum_{k=-\infty}^{\infty} \alpha_k = \lim_{n\to\infty} \left(\sum_{k=-n}^{n} \alpha_k \right) .$$

Teorema. *Se $f : \mathbb{R} \to \mathbb{R}$ è continua e T-periodica e la serie $\sum_{k=-\infty}^{\infty} |c_k|$ converge, allora*

$$\lim_{n\to\infty} f_n(t) = f(t) \,,$$

uniformemente in $t \in \mathbb{R}$.

Scriveremo quindi:

$$f(t) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \left(a_k \cos\left(\frac{2\pi k}{T} t\right) + b_k \sin\left(\frac{2\pi k}{T} t\right)\right) \;;$$

oppure

$$f(t) = \sum_{k=-\infty}^{\infty} c_k e^{i\frac{2\pi k}{T} t} \,.$$

Dimostrazione. Essendo

$$\left| c_k e^{i\frac{2\pi k}{T} t} \right| \le |c_k| \,,$$

siccome la serie $\sum_{k=-\infty}^{\infty} |c_k|$ converge, per il criterio di Weierstrass si ha che le f_n convergono uniformemente ad una certa funzione continua $\tilde{f}$. D'altra parte, per tale funzione si ha:

$$\begin{aligned} \tilde{c}_k &= \frac{1}{T} \int_0^T \tilde{f}(t) e^{-i\frac{2\pi k}{T} t}\, dt = \frac{1}{T} \int_0^T \left(\lim_{n\to\infty} f_n(t) \right) e^{-i\frac{2\pi k}{T} t}\, dt \\ &= \lim_{n\to\infty} \frac{1}{T} \int_0^T f_n(t) e^{-i\frac{2\pi k}{T} t}\, dt = \lim_{n\to\infty} \frac{1}{T} \int_0^T \sum_{j=-n}^{n} c_j e^{i\frac{2\pi j}{T} t} e^{-i\frac{2\pi k}{T} t}\, dt \\ &= \lim_{n\to\infty} \sum_{j=-n}^{n} \frac{1}{T} \int_0^T c_j e^{i\frac{2\pi (j-k)}{T} t}\, dt = \lim_{n\to\infty} \frac{1}{T} \int_0^T c_k\, dt = c_k \,, \end{aligned}$$

per ogni k. Per il corollario al teorema di Fejer, f e $\tilde{f}$ coincidono, per cui la dimostrazione è completa. ∎

Corollario. *Se $f : \mathbb{R} \to \mathbb{R}$ è di classe C^2 ed è T-periodica, allora*

$$\lim_{n\to\infty} f_n(t) = f(t) \,,$$

uniformemente in $t \in \mathbb{R}$.

<u>Dimostrazione</u>. Integrando per parti due volte,

$$\begin{aligned}
c_k &= \frac{1}{T}\int_0^T f(t)e^{-i\frac{2\pi k}{T}t}\,dt \\
&= \frac{1}{T}\left(\left[f(t)\frac{T}{-i2\pi k}e^{-i\frac{2\pi k}{T}t}\right]_0^T + \frac{T}{i2\pi k}\int_0^T f'(t)e^{-i\frac{2\pi k}{T}t}\,dt\right) \\
&= \frac{1}{i2\pi k}\int_0^T f'(t)e^{-i\frac{2\pi k}{T}t}\,dt \\
&= \frac{1}{i2\pi k}\left(\left[f'(t)\frac{T}{-i2\pi k}e^{-i\frac{2\pi k}{T}t}\right]_0^T + \frac{T}{i2\pi k}\int_0^T f''(t)e^{-i\frac{2\pi k}{T}t}\,dt\right) \\
&= \frac{T}{(i2\pi k)^2}\int_0^T f''(t)e^{-i\frac{2\pi k}{T}t}\,dt\,.
\end{aligned}$$

Essendo f'' continua, esiste un $M>0$ per cui

$$|f''(t)|\leq M\,,\quad \text{per ogni } t\in[0,T]\,.$$

Quindi,

$$|c_k| = \frac{T}{(2\pi k)^2}\left|\int_0^T f''(t)e^{-i\frac{2\pi k}{T}t}\,dt\right| \leq \frac{T}{(2\pi k)^2}\int_0^T |f''(t)e^{-i\frac{2\pi k}{T}t}|\,dt \leq \frac{T^2M}{(2\pi k)^2}\,,$$

per cui la serie $\sum_{k=-\infty}^{\infty}|c_k|$ converge, e il risultato segue dal teorema precedente.

■

Nota. I risultati sopra dimostrati si estendono senza difficoltà alle funzioni a valori complessi anziché reali.

Concludiamo con il seguente enunciato. Qui f potrebbe non essere continua.

Teorema. *Supponiamo che esistano un numero finito di punti $t_0, t_1, t_2, \ldots, t_N$, con*

$$0 = t_0 < t_1 < t_2 < \ldots < t_N = T\,,$$

per cui f risulti di classe C^1 su ogni intervallo $]t_{j-1}, t_j[$, con $j = 1, 2, \ldots, N$. Nei punti t_j la funzione f potrebbe non essere continua, oppure, pur essendo continua, potrebbe non essere derivabile. Supponiamo comunque che, per ogni $t \in \{t_0, t_1, t_2, \ldots, t_N\}$, esistano e siano finiti i seguenti limiti:

$$\lim_{s\to t^-} f(s)\,,\quad \lim_{s\to t^+} f(s)\,,\quad \lim_{s\to t^-} f'(s)\,,\quad \lim_{s\to t^+} f'(s)\,.$$

Allora la serie di Fourier associata a f converge per ogni $t\in\mathbb{R}$, e si ha

$$\frac{a_0}{2} + \sum_{k=1}^{\infty}\left(a_k\cos\left(\frac{2\pi k}{T}t\right) + b_k\sin\left(\frac{2\pi k}{T}t\right)\right) = \frac{1}{2}\Big(\lim_{s\to t^-} f(s) + \lim_{s\to t^+} f(s)\Big)\,.$$

Si noti che, se f è continua in un certo t, allora

$$\frac{1}{2}\Big(\lim_{s\to t^-} f(s) + \lim_{s\to t^+} f(s)\Big) = f(t)\,.$$

Per la dimostrazione di questo teorema, si possono consultare, ad esempio, i libri di E. Giusti o di G. Prodi consigliati per questo corso.

23 Equazioni differenziali: il problema di Cauchy

Consideriamo equazioni del tipo

$$u' = f(t,u)\,.$$

Qui f è una funzione continua definita su un sottoinsieme aperto A di $\mathbb{R}\times\mathbb{R}^N$, a valori in $\mathbb{R}^N$. Si tratta di trovare una funzione u, definita su un intervallo I di $\mathbb{R}$, da determinarsi, per cui si abbia

$$u'(t) = f(t,u(t))\,,$$

per ogni $t \in I$. Ricordiamo che

$$u'(t) = \lim_{s\to t} \frac{u(s)-u(t)}{s-t} \in \mathbb{R}^N$$

è il "vettore derivata" di u in t. Spesso, pensando ai modelli della meccanica, si parla di "vettore velocità" di u "al tempo t".

Nelle applicazioni ci si trova spesso ad affrontare la ricerca di una soluzione dell'equazione differenziale che soddisfi una "condizione iniziale" del tipo

$$u(t_0) = u_0\,.$$

Avendo in mente i modelli della meccanica, si dice che la "posizione" di u "al tempo iniziale t_0" è u_0.

Ci si trova così a dover risolvere il "problema di Cauchy"

$$\begin{cases} u' = f(t,u) \\ u(t_0) = u_0 \end{cases}$$

dove $f : A \subseteq \mathbb{R}\times\mathbb{R}^N \to \mathbb{R}^N$ è una funzione continua, e $(t_0,u_0)\in A$. Si tratta di trovare un intervallo I, contenente t_0, e una funzione derivabile $u : I \to \mathbb{R}^N$, per cui si abbia che $u'(t) = f(t,u(t))$, per ogni $t\in I$, e $u(t_0) = u_0$.

Teorema. *Il problema di Cauchy è equivalente a trovare una funzione continua $u : I \to \mathbb{R}^N$ per cui si abbia*

$$u(t) = u_0 + \int_{t_0}^{t} f(s,u(s))\,ds\,,$$

per ogni $t\in I$, dove I è un intervallo contenente t_0.[9]

[9] L'integrale di una funzione a valori vettoriali $g : [a,b] \to \mathbb{R}^N$, con

$$g(t) = (g_1(t),\dots,g_N(t))\,,$$

Dimostrazione. Dobbiamo dimostrare una doppia implicazione. Se u è una soluzione del problema di Cauchy, allora

$$u(t) - u(t_0) = \int_{t_0}^{t} u'(s)\, ds = \int_{t_0}^{t} f(s, u(s))\, ds\,,$$

il che dimostra una delle due implicazioni, essendo $u(t_0) = u_0$.

Viceversa, se u è una funzione continua per cui si abbia

$$u(t) = u_0 + \int_{t_0}^{t} f(s, u(s))\, ds\,,$$

si ha che $u(t_0) = u_0$, e la funzione $s \mapsto f(s, u(s))$ è continua. Quindi u è derivabile, e si ha

$$u'(t) = \frac{d}{dt} \int_{t_0}^{t} f(s, u(s))\, ds = f(t, u(t))\,.$$

■

Affronteremo ora il problema dell'esistenza e dell'unicità della soluzione per un problema di Cauchy, definita in un opportuno intorno di t_0. A tal scopo, aggiungeremo un'ipotesi sulla funzione f.

Siano $r_1 > 0$ e $r_2 > 0$ tali che il "rettangolo"

$$\mathcal{R} = [t_0 - r_1, t_0 + r_1] \times \bar{B}(u_0, r_2)$$

risulti contenuto in A, e sia

$$M = \max\{\|f(t, u)\| : (t, u) \in \mathcal{R}\}\,.$$

Supponiamo che esista una costante $L \geq 0$ tale che

$$\|f(t, v) - f(t, w)\| \leq L\|v - w\|\,,$$

per ogni $t \in [t_0 - r_1, t_0 + r_1]$ e $v, w \in \bar{B}(u_0, r_2)$. Infine, sia $r_0 > 0$ tale che

$$r_0 < \min\left\{r_1, \frac{r_2}{M}, \frac{1}{L}\right\}\,.$$

Consideriamo l'insieme X, costituito dalle funzioni continue $v : I_0 \to \mathbb{R}^N$, con $I_0 = [t_0 - r_0, t_0 + r_0]$, tali che $v(t) \in \bar{B}(u_0, r_2)$, per ogni $t \in I_0$. Sotto queste ipotesi, vogliamo dimostrare che il problema di Cauchy ha una ed una sola soluzione $u \in X$.

è definito da

$$\int_a^b g(t)\, dt = \left(\int_a^b g_1(t)\, dt, \ldots, \int_a^b g_N(t)\, dt\right)\,.$$

Esso è quindi un vettore di $\mathbb{R}^N$. Se $a < b$, risulterà utile la seguente disuguaglianza:

$$\left\|\int_a^b g(t)\, dt\right\| \leq \int_a^b \|g(t)\|\, dt\,.$$

Sia F la funzione che associa ad ogni $v \in X$ la funzione $F(v) : I_0 \to \mathbb{R}^N$ definita da

$$[F(v)](t) = u_0 + \int_{t_0}^{t} f(s, v(s))\, ds\,.$$

Dimostriamo che $F(v) \in X$. Infatti, $F(v)$ è una funzione continua, e si ha

$$\|[F(v)](t) - u_0\| = \left\| \int_{t_0}^{t} f(s, v(s))\, ds \right\| \leq \left| \int_{t_0}^{t} \|f(s, v(s))\|\, ds \right| \leq M|t - t_0| < r_2\,.$$

Notiamo inoltre che, prese due funzioni v, w in X, si ha

$$\begin{aligned}
\|F(v) - F(w)\|_0 &= \sup\{\|[F(v)](t) - [F(w)](t)\| : t \in I_0\} \\
&= \sup\left\{ \left\| \int_{t_0}^{t} (f(s, v(s)) - f(s, w(s)))\, ds \right\| : t \in I_0 \right\} \\
&\leq \sup\left\{ \left| \int_{t_0}^{t} \|f(s, v(s)) - f(s, w(s))\|\, ds \right| : t \in I_0 \right\} \\
&\leq \sup\left\{ \left| \int_{t_0}^{t} L\|v(s) - w(s)\|\, ds \right| : t \in I_0 \right\} \\
&\leq L r_0 \|v - w\|_0\,.
\end{aligned}$$

Ponendo $\alpha = L r_0$, si ha quindi che $\alpha < 1$ e

$$\|F(v) - F(w)\|_0 \leq \alpha \|v - w\|_0\,.$$

Osserviamo che, se $u \in X$ soddisfa l'equazione

$$F(u) = u\,,$$

allora u è una soluzione del nostro problema. Si tratta quindi di trovare un "punto fisso" della funzione $F : X \to X$. Notiamo che X è un sottoinsieme chiuso di $C(I_0, \mathbb{R}^N)$, per cui X è uno spazio metrico completo.

Dato uno spazio metrico E, diremo che una funzione $F : E \to E$ è una "contrazione" se, per un certo $\alpha < 1$, si ha

$$d(F(v), F(w)) \leq \alpha\, d(v, w)\,,$$

per ogni $v, w \in E$.

Teorema. *Se E è uno spazio metrico completo e $F : E \to E$ è una contrazione, allora esiste un unico $x \in E$ tale che $F(x) = x$. Inoltre, scegliendo $x_0 \in E$ arbitrariamente, la successione $(x_n)_n$ definita da*

$$x_{n+1} = F(x_n)$$

è tale che $\lim_n x_n = x$.

Ricordiamo che, se $F(x) = x$, si dice che x è un "punto fisso" di F.

Dimostrazione. Consideriamo la successione $(x_n)_n$ definita come nell'enunciato, con $x_0 \in E$ arbitrario. Dimostriamo che è una successione di Cauchy. Osserviamo che, se m e n sono due numeri naturali, con $m < n$, si ha

$$d(x_m, x_n) \leq \sum_{k=m}^{n-1} d(x_k, x_{k+1}) \,.$$

Dimostriamo per induzione che, per ogni $k \in \mathbb{N}$, vale la seguente proposizione:

$$(P_k) \qquad d(x_k, x_{k+1}) \leq \alpha^k d(x_0, x_1) \,.$$

Infatti, se $k = 0$ si ha chiaramente l'uguaglianza, per cui (P_0) è vera. Supponiamo ora vera (P_k), per un certo $k \in \mathbb{N}$; allora

$$d(x_{k+1}, x_{k+2}) = d(F(x_k), F(x_{k+1})) \leq \alpha\, d(x_k, x_{k+1}) \leq \alpha^{k+1} d(x_0, x_1) \,,$$

per cui è vera anche (P_{k+1}).

Usando (P_k), abbiamo quindi

$$d(x_m, x_n) \leq \sum_{k=m}^{n-1} \alpha^k d(x_0, x_1) \leq \alpha^m d(x_0, x_1) \sum_{k=0}^{n-1-m} \alpha^k.$$

Siccome $\alpha \in [0, 1[$, la serie geometrica di base α converge e ha somma $\frac{1}{1-\alpha}$, per cui

$$d(x_m, x_n) \leq \alpha^m \frac{d(x_0, x_1)}{1 - \alpha} \,.$$

Fissato $\varepsilon > 0$, siccome $\alpha \in [0, 1[$, esiste un $\bar{n} \in \mathbb{N}$ tale che

$$m \geq \bar{n} \quad \Rightarrow \quad \alpha^m \frac{d(x_0, x_1)}{1 - \alpha} < \varepsilon \,.$$

Ne segue che

$$n > m \geq \bar{n} \quad \Rightarrow \quad d(x_m, x_n) < \varepsilon \,,$$

per cui $(x_n)_n$ è di Cauchy. Siccome E è completo, esiste il limite di $(x_n)_n$. Sia esso un certo $x \in E$:

$$\lim_n x_n = x \,.$$

Allora, essendo F continua,

$$F(x) = F(\lim_n x_n) = \lim_n F(x_n) = \lim_n x_{n+1} = x \,,$$

il che dimostra che x è un punto fisso di F.

Resta da dimostrare che il punto fisso è unico. Supponiamo che x' sia anch'esso un punto fisso di F. Allora

$$d(x, x') = d(F(x), F(x')) \leq \alpha\, d(x, x') \,,$$

e siccome $\alpha < 1$, deve essere $x = x'$. ∎

Abbiamo quindi il seguente corollario.

Corollario. *Siano* $A \subseteq \mathbb{R} \times \mathbb{R}^N$ *un insieme aperto,* (t_0, u_0) *un punto di* A *e* $f : A \to \mathbb{R}^N$ *una funzione continua. Siano* $r_1 > 0$ *e* $r_2 > 0$ *tali che*

$$[t_0 - r_1, t_0 + r_1] \times \bar{B}(u_0, r_2) \subseteq A\,,$$

e sia

$$M = \max\{\|f(t,u)\| : t \in [t_0 - r_1, t_0 + r_1]\,,\ u \in \bar{B}(u_0, r_2)\}\,.$$

Supponiamo che esista una costante $L \geq 0$ *tale che*

$$\|f(t,v) - f(t,w)\| \leq L\|v - w\|\,,$$

per ogni $t \in [t_0 - r_1, t_0 + r_1]$ *e ogni* $v, w \in \bar{B}(u_0, r_2)$*. Sia* $r_0 > 0$ *tale che*

$$r_0 < \min\left\{r_1, \frac{r_2}{M}, \frac{1}{L}\right\}\,.$$

Allora il problema di Cauchy

$$\begin{cases} u' = f(t,u) \\ u(t_0) = u_0 \end{cases}$$

ha una ed una sola soluzione $u : [t_0 - r_0, t_0 + r_0] \to \mathbb{R}^N$*, i cui valori sono in* $\bar{B}(u_0, r_2)$*. Inoltre, presa una qualsiasi funzione continua* $x_0 : [t_0 - r_0, t_0 + r_0] \to \mathbb{R}^N$*, con valori in* $\bar{B}(u_0, r_2)$*, la successione di funzioni* $(x_n)_n$*, definita da*

$$x_{n+1}(t) = u_0 + \int_{t_0}^{t} f(s, x_n(s))\, ds\,,$$

è tale che $\lim_n x_n = u$*, uniformemente su* $[t_0 - r_0, t_0 + r_0]$*.*