

Dispense di
MATEMATICA II

M. Romeo

Febbraio 2013

Indice

1	Funzioni vettoriali e curve nello spazio	1
1.1	Limiti, continuità e derivate di vettori	1
1.2	Curve in forma parametrica	3
1.3	Terna intrinseca	8
2	Funzioni di più variabili	11
2.1	Limiti e funzioni continue	11
2.2	Derivate parziali	13
2.3	Funzioni differenziabili	15
2.4	Funzioni vettoriali e funzioni composte	20
2.5	Derivate successive e formula di Taylor	23
2.6	Massimi e minimi relativi	26
3	Spazi metrici	33
3.1	Metrica	33
3.2	Convergenza uniforme	35
3.3	Teorema delle contrazioni	37
4	Equazioni differenziali	39
4.1	Equazioni del primo ordine	39
4.2	Teorema di esistenza e unicità	46
4.3	Equazioni differenziali lineari	48
4.4	Equazioni omogenee a coefficienti costanti	52
4.5	Equazioni non omogenee	55
4.6	Sistemi di equazioni differenziali lineari	59
4.7	Alcune applicazioni delle equazioni differenziali	66
4.7.1	Probabilità di collisione	66
4.7.2	Moto di un punto soggetto a forza resistente	67
4.7.3	Oscillatore armonico forzato	68
5	Superfici regolari	71
5.1	Equazioni di superfici regolari	71
5.2	Curva su una superficie regolare	73
5.3	Teorema delle funzioni implicite	76
5.4	Superfici in forma implicita	79

6	Integrazione delle funzioni di più variabili	81
6.1	Insiemi misurabili	81
6.2	Integrali doppi	83
6.3	Formule di riduzione per gli integrali doppi	86
6.4	Cambiamento di variabili negli integrali doppi	90
6.5	Calcolo di aree di superfici	97
6.6	Integrali di superficie	99
6.7	Integrali tripli	100
6.8	Applicazioni degli integrali doppi e tripli	105
6.9	Derivazione sotto il segno di integrale	110
7	Integrali impropri	113
7.1	Integrali estesi ad intervalli non limitati	113
7.2	Integrali di funzioni non limitate	120
7.3	Integrali impropri di funzioni di più variabili	123
7.4	Esempi di calcolo di integrali impropri	125
8	Serie numeriche	127
8.1	Definizioni e operazioni sulle serie	127
8.2	Criteri di convergenza	130
8.3	Serie alternanti e serie assolutamente convergenti	134
9	Serie di funzioni	137
9.1	Convergenza uniforme	137
9.2	Serie di potenze	141
9.3	Serie di Taylor	145

Capitolo 1

Funzioni vettoriali e curve nello spazio

1.1 Limiti, continuità e derivate di vettori

Una legge che ad ogni numero reale t nel dominio $A \subseteq \mathbb{R}$ faccia corrispondere un vettore di V_3 , viene detta *funzione a valori vettoriali* o *funzione vettoriale* di variabile reale in V_3 . Indichiamo con $\mathbf{v}(t)$ una tale funzione e prendiamo una qualunque base $\{\mathbf{e}_i\}$ in V_3 . Per ogni $t \in A$ possiamo scrivere

$$\mathbf{v}(t) = \sum_{i=1}^3 v_i(t)\mathbf{e}_i.$$

Poiché le componenti $v_i(t)$, ($i = 1, 2, 3$) individuano univocamente, per ogni t , il vettore $\mathbf{v}(t)$ nella base scelta, esse, come funzioni reali della variabile reale t nel dominio A , costituiranno le *componenti* della funzione vettoriale stessa in quella base.

Vogliamo estendere alle funzioni vettoriali i concetti di limite, di funzione continua e di derivata. A tale scopo consideriamo la funzione vettoriale $\mathbf{v}(t)$, definita nel dominio $A \subseteq \mathbb{R}$, a valori in V_3 , ed il vettore $\mathbf{w} \in V_3$. Preso un punto t_0 di accumulazione per A diremo che la funzione $\mathbf{v}(t)$ converge al vettore \mathbf{w} per t che tende a t_0 , se è possibile determinare un intorno di t_0 in A , escluso al più t_0 in modo tale che per valori di t in tale intorno, la norma della differenza $\mathbf{v}(t) - \mathbf{w}$ sia più piccola di un qualunque numero positivo prefissato. In altri termini scriveremo

$$\lim_{t \rightarrow t_0} \mathbf{v}(t) = \mathbf{w},$$

se, $\forall \varepsilon > 0, \exists \delta > 0 : 0 < |t - t_0| < \delta \Rightarrow \|\mathbf{v}(t) - \mathbf{w}\| < \varepsilon$. Dette w_i , ($i = 1, 2, 3$) le componenti di \mathbf{w} nella base $\{\mathbf{e}_i\}$, si ha

$$\|\mathbf{v}(t) - \mathbf{w}\| = \sqrt{(v_1(t) - w_1)^2 + (v_2(t) - w_2)^2 + (v_3(t) - w_3)^2}.$$

Ne segue che la condizione $\|\mathbf{v}(t) - \mathbf{w}\| < \varepsilon$ implica che siano soddisfatte contemporaneamente le tre condizioni

$$|v_i(t) - w_i| < \varepsilon, \quad i = 1, 2, 3.$$

Viceversa, queste ultime condizioni implicano che $\|\mathbf{v}(t) - \mathbf{w}\| < \sqrt{3}\varepsilon$, ovvero, che la norma della differenza $\mathbf{v}(t) - \mathbf{w}$ si può rendere minore di un qualunque numero positivo prefissato. Concludiamo allora che la funzione vettoriale $\mathbf{v}(t)$ converge, per $t \rightarrow t_0$, al vettore \mathbf{w} se e solo se le funzioni componenti $v_i(t)$ convergono rispettivamente a w_i , ($i = 1, 2, 3$), ovvero

$$\lim_{t \rightarrow t_0} \mathbf{v}(t) = \mathbf{w} \iff \lim_{t \rightarrow t_0} v_i(t) = w_i, \quad i = 1, 2, 3.$$

Questo risultato ci permette di ricondurre il calcolo del limite di una funzione vettoriale al calcolo del limite delle sue componenti. Di conseguenza, tutti i teoremi e le proprietà sui limiti di funzione si estendono immediatamente alle funzioni vettoriali. In particolare questo vale anche per il limite del prodotto scalare tra due funzioni vettoriali ed il limite del prodotto vettoriale tra due funzioni vettoriali.

Una funzione vettoriale $\mathbf{v}(t)$ definita nel dominio $A \subseteq \mathbb{R}$ si dice continua in un punto t_0 di accumulazione per A se risulta

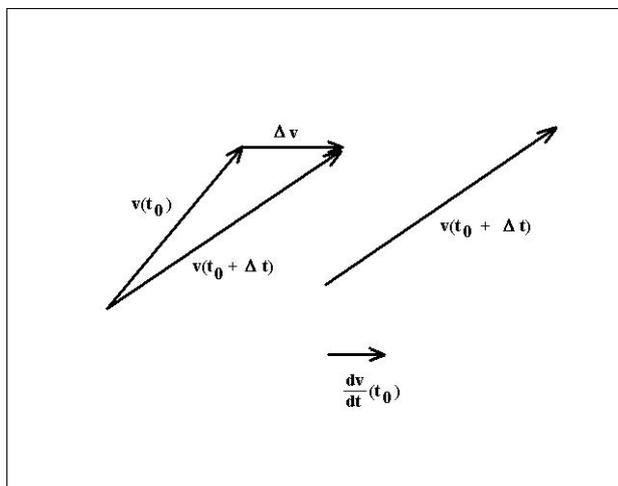
$$\lim_{t \rightarrow t_0} \mathbf{v}(t) = \mathbf{v}(t_0).$$

In forza del risultato precedente sul limite di una funzione vettoriale, possiamo dire che la continuità di $\mathbf{v}(t)$ in t_0 è equivalente alla continuità delle sue funzioni componenti $v_i(t)$ in t_0 .

Proprietà analoghe a quelle delle funzioni continue valgono per le funzioni vettoriali. Per esempio, la somma di più funzioni vettoriali continue è una funzione vettoriale continua. Il prodotto scalare tra due funzioni vettoriali continue è una funzione continua ed il prodotto vettoriale tra due funzioni vettoriali continue è una funzione vettoriale continua.

Preso un punto t_0 in A ed un incremento Δt della variabile t in modo che $t_0 + \Delta t \in A$, consideriamo i vettori $\mathbf{v}(t_0 + \Delta t)$ e $\mathbf{v}(t_0)$. Denotato con $\Delta \mathbf{v} = \mathbf{v}(t_0 + \Delta t) - \mathbf{v}(t_0)$ l'incremento della funzione vettoriale relativa al punto t_0 e all'incremento Δt , chiamiamo *rapporto incrementale* il vettore

$$\frac{\Delta \mathbf{v}}{\Delta t} = \frac{\mathbf{v}(t_0 + \Delta t) - \mathbf{v}(t_0)}{\Delta t}.$$



Il rapporto incrementale risulta una funzione vettoriale di t . Se esiste ed è finito, il limite di questa funzione per $\Delta t \rightarrow 0$ si dice derivata della funzione vettoriale $\mathbf{v}(t)$ in t_0 e si scrive

$$\lim_{\Delta t \rightarrow 0} \frac{\mathbf{v}(t_0 + \Delta t) - \mathbf{v}(t_0)}{\Delta t} = \frac{d\mathbf{v}}{dt}(t_0).$$

Talvolta useremo notazioni analoghe a quelle relative alle funzioni, ovvero $D\mathbf{v}(t_0)$, oppure $\mathbf{v}'(t_0)$. Considerata una base $\{\mathbf{e}_i\}$ e usando il risultato precedente sui limiti di funzioni vettoriali, si può

scrivere

$$\begin{aligned} \frac{d\mathbf{v}}{dt}(t_0) &= \lim_{\Delta t \rightarrow 0} \sum_{i=1}^3 \frac{v_i(t_0 + \Delta t) - v_i(t_0)}{\Delta t} \mathbf{e}_i \\ &= \sum_{i=1}^3 \lim_{\Delta t \rightarrow 0} \frac{v_i(t_0 + \Delta t) - v_i(t_0)}{\Delta t} \mathbf{e}_i = \sum_{i=1}^3 \frac{dv_i}{dt}(t_0) \mathbf{e}_i. \end{aligned}$$

In altri termini, le componenti della derivata di una funzione vettoriale rispetto ad una data base sono le derivate delle funzioni componenti di quel vettore rispetto a quella base.

Anche per le funzioni vettoriali derivabili valgono regole e teoremi analoghi a quelli delle funzioni reali ordinarie derivabili. Ci limitiamo ad osservare che una funzione vettoriale derivabile risulta continua. Estenderemo poi alle funzioni vettoriali la notazione per gli insiemi di funzioni con derivate continue. Per esempio, la notazione $\mathbf{v}(t) \in C^k(A)$ indica che la funzione vettoriale $\mathbf{v}(t)$ ha derivate continue fino all'ordine k nel dominio A , ovvero, le sue componenti sono funzioni appartenenti a $C^k(A)$. Si possono poi dimostrare facilmente le seguenti regole di derivazione

$$\begin{aligned} \frac{d}{dt}[\mathbf{u}(t) \cdot \mathbf{v}(t)] &= \frac{d}{dt} \mathbf{u}(t) \cdot \mathbf{v}(t) + \mathbf{u}(t) \cdot \frac{d}{dt} \mathbf{v}(t) \\ \frac{d}{dt}[\mathbf{u}(t) \times \mathbf{v}(t)] &= \frac{d}{dt} \mathbf{u}(t) \times \mathbf{v}(t) + \mathbf{u}(t) \times \frac{d}{dt} \mathbf{v}(t) \\ \frac{d}{dx} \mathbf{u}(t(x)) &= \frac{d}{dt} \mathbf{u}(t(x)) t'(x). \end{aligned}$$

1.2 Curve in forma parametrica

Le equazioni in forma parametrica della retta passante per P_0 e diretta secondo il vettore \mathbf{v} ,

$$\begin{cases} x = x_0 + tv_x \\ y = y_0 + tv_y \\ z = z_0 + tv_z, \end{cases}$$

sono un esempio di curva nello spazio espressa mediante funzioni di un parametro reale. Più in generale, data una curva \mathcal{C} nello spazio ed un riferimento cartesiano con origine in O , supponiamo che sia possibile stabilire una corrispondenza biunivoca tra i punti P di \mathcal{C} ed i valori di un parametro t in un intervallo reale $[a, b]$. Si può esprimere allora il vettore \vec{OP} come funzione del parametro t , ovvero si può introdurre una funzione vettoriale $\hat{\mathbf{x}}(t)$ tale che

$$\vec{OP} = (O, \hat{\mathbf{x}}(t)).$$

Le componenti del vettore posizione $\mathbf{x} = \hat{\mathbf{x}}(t)$ sono le coordinate del punto P nel riferimento adottato e risultano funzioni di t definite in $[a, b]$. In altri termini, si può scrivere

$$\begin{cases} x = \hat{x}(t) \\ y = \hat{y}(t) \\ z = \hat{z}(t) \end{cases} \quad t \in [a, b],$$

che rappresentano le equazioni parametriche della curva \mathcal{C} .

 • Esempi

1) Le equazioni dell'elica cilindrica di raggio R , asse z e passo p , si possono scrivere in forma parametrica usando come parametro la coordinata cilindrica ψ . Si ha infatti

$$\begin{cases} x = R \cos \psi \\ y = R \sin \psi \\ z = \frac{p}{2\pi} \psi, \end{cases} \quad \psi \in \mathbb{R}.$$

2) Analogamente le equazioni di una circonferenza di raggio R e centro nell'origine, sul piano $\psi = \pi/2$, in coordinate sferiche, si possono esprimere in forma parametrica usando come parametro la coordinata sferica θ . Si ha

$$\begin{cases} x = 0 \\ y = R \sin \theta \\ z = R \cos \theta, \end{cases} \quad \theta \in [0, 2\pi)$$

Sia $\mathbf{x} = \hat{\mathbf{x}}(t)$ l'equazione parametrica di una curva \mathcal{C} definita in $[a, b]$. Supponiamo che $\hat{\mathbf{x}}(t) \in C^1([a, b])$. Supponiamo inoltre che $\|\mathbf{dx}/dt\| \neq 0, \forall t \in [a, b]$ e che $\nexists t_1, t_2$ con $t_1 \neq t_2 : \hat{\mathbf{x}}(t_1) = \hat{\mathbf{x}}(t_2)$ in tutto $[a, b]$. Sotto queste condizioni si dice che $\mathbf{x} = \hat{\mathbf{x}}(t)$ è una rappresentazione parametrica *regolare* della curva \mathcal{C} . Una curva che ammette almeno una rappresentazione parametrica regolare si dice *regolare*. Dalla definizione segue che nelle equazioni parametriche di una curva regolare le funzioni $\hat{x}(t), \hat{y}(t), \hat{z}(t)$ hanno derivate prime continue in $[a, b]$ che non si annullano mai contemporaneamente in $[a, b]$ e, al variare di t in $[a, b]$, la curva non passa per un dato punto più di una volta.

 • Esempio

I due sistemi

$$\begin{cases} x = x_0 + at \\ y = y_0 + bt \\ z = z_0 + ct \end{cases}, \quad \begin{cases} x = x_0 + at^3 \\ y = y_0 + bt^3 \\ z = z_0 + ct^3 \end{cases}, \quad t \in \mathbb{R},$$

rappresentano la stessa retta, in quanto, eliminando il parametro t , da ambedue i sistemi si ottengono le medesime equazioni cartesiane

$$\frac{x - x_0}{a} = \frac{y - y_0}{b} = \frac{z - z_0}{c}.$$

Tuttavia solo il primo sistema è una rappresentazione regolare della retta. Infatti, per il secondo sistema si ha $\|\mathbf{x}'(t)\| = 0$ per $t = 0$.

Di una curva regolare è possibile definire una lunghezza. A questo scopo consideriamo una partizione P_n dell'intervallo $[a, b]$ in n parti prendendo $a = t_0 < t_1 < \dots < t_n = b$ e siano $\mathbf{x}(t_0), \mathbf{x}(t_1), \dots, \mathbf{x}(t_n)$

i corrispondenti vettori posizione dei punti sulla curva \mathcal{C} . Congiungendo gli estremi di questi vettori otterremo una spezzata di lunghezza

$$\begin{aligned} L(P_n) &= \sum_{i=1}^n \|\mathbf{x}(t_i) - \mathbf{x}(t_{i-1})\| \\ &= \sum_{i=1}^n \sqrt{[x(t_i) - x(t_{i-1})]^2 + [y(t_i) - y(t_{i-1})]^2 + [z(t_i) - z(t_{i-1})]^2}. \end{aligned}$$

Poiché le funzioni $x(t), y(t), z(t)$ sono derivabili, applicando a ciascuna di esse il teorema di Lagrange si potranno trovare tre punti ξ_i, η_i, ζ_i nell'intervallo (t_{i-1}, t_i) in modo tale che

$$\begin{aligned} x(t_i) - x(t_{i-1}) &= x'(\xi_i)(t_i - t_{i-1}), \\ y(t_i) - y(t_{i-1}) &= y'(\eta_i)(t_i - t_{i-1}), \\ z(t_i) - z(t_{i-1}) &= z'(\zeta_i)(t_i - t_{i-1}), \end{aligned}$$

Otteniamo così

$$L(P_n) = \sum_{i=1}^n (t_i - t_{i-1}) \sqrt{[x'(\xi_i)]^2 + [y'(\eta_i)]^2 + [z'(\zeta_i)]^2}.$$

Posto

$$\begin{aligned} \bar{x}'_i &= \inf\{|x'(t)|, t \in [t_{i-1}, t_i]\}, & \bar{X}'_i &= \sup\{|x'(t)|, t \in [t_{i-1}, t_i]\} \\ \bar{y}'_i &= \inf\{|y'(t)|, t \in [t_{i-1}, t_i]\}, & \bar{Y}'_i &= \sup\{|y'(t)|, t \in [t_{i-1}, t_i]\} \\ \bar{z}'_i &= \inf\{|z'(t)|, t \in [t_{i-1}, t_i]\}, & \bar{Z}'_i &= \sup\{|z'(t)|, t \in [t_{i-1}, t_i]\} \end{aligned}$$

si ha

$$L_m(P_n) \leq L(P_n) \leq L_M(P_n),$$

dove

$$\begin{aligned} L_m(P_n) &= \sum_{i=1}^n (t_i - t_{i-1}) \sqrt{[\bar{x}'_i]^2 + [\bar{y}'_i]^2 + [\bar{z}'_i]^2} \\ L_M(P_n) &= \sum_{i=1}^n (t_i - t_{i-1}) \sqrt{[\bar{X}'_i]^2 + [\bar{Y}'_i]^2 + [\bar{Z}'_i]^2}. \end{aligned}$$

Diremo allora che la curva \mathcal{C} in $[a, b]$ è *rettificabile* se

$$\sup_{P_n} [L_m(P_n)] = \inf_{P_n} [L_M(P_n)] = l$$

e chiameremo l la lunghezza di \mathcal{C} . Vale il seguente risultato che non dimostriamo.

Teorema 1.1 *Una curva $\mathbf{x} = \hat{\mathbf{x}}(t)$, regolare in $[a, b]$, è rettificabile.*

L'analisi precedente ci permette di effettuare il calcolo della lunghezza di una curva.

Teorema 1.2 *Data la curva $\mathbf{x} = \hat{\mathbf{x}}(t)$, regolare in $[a, b]$, la sua lunghezza è data da*

$$l = \int_a^b \sqrt{[\hat{x}'(t)]^2 + [\hat{y}'(t)]^2 + [\hat{z}'(t)]^2} dt.$$

Dim. Consideriamo la funzione

$$f(t) = \sqrt{[\hat{x}'(t)]^2 + [\hat{y}'(t)]^2 + [\hat{z}'(t)]^2}.$$

In base alla definizione delle quantità $L_m(P_n)$ e $L_M(P_n)$ e indicate con $s(f, P_n)$ e $S(f, P_n)$ rispettivamente le somme integrali inferiore e superiore della funzione f relative alla partizione P_n , valgono le seguenti disuguaglianze

$$L_m(P_n) \leq s(f, P_n) \leq S(f, P_n) \leq L_M(P_n).$$

Poiché la curva è regolare, $f(t)$ risulta continua, e quindi integrabile in $[a, b]$, ovvero

$$\sup_{P_n} [s(f, P_n)] = \inf_{P_n} [S(f, P_n)] = \int_a^b \sqrt{[\hat{x}'(t)]^2 + [\hat{y}'(t)]^2 + [\hat{z}'(t)]^2} dt.$$

D'altra parte, la curva risulta rettificabile quindi

$$\sup_{P_n} [L_m(P_n)] = \inf_{P_n} [L_M(P_n)] = l,$$

e le precedenti disuguaglianze ci permettono di concludere

$$l = \int_a^b \sqrt{[\hat{x}'(t)]^2 + [\hat{y}'(t)]^2 + [\hat{z}'(t)]^2} dt = \int_a^b \left\| \frac{d\hat{\mathbf{x}}}{dt} \right\| dt.$$

• **Esempio**

Vogliamo calcolare la lunghezza dell'arco di curva di equazioni parametriche

$$\begin{cases} x = t \cos t \\ y = t \sin t \\ z = \sqrt{3} t \end{cases}$$

relativo all'intervallo $0 \leq t \leq 1$.

Si può verificare che tratta di una curva regolare. Si ha

$$\begin{cases} x' = \cos t - t \sin t \\ y' = \sin t + t \cos t \\ z' = \sqrt{3} \end{cases}$$

da cui

$$\left\| \frac{d\mathbf{x}}{dt} \right\| = \sqrt{4 + t^2}.$$

Si ottiene quindi

$$l = \int_0^1 \sqrt{4 + t^2} dt.$$

Questo integrale si può risolvere mediante la sostituzione $t = 2 \sinh \eta$, ottenendo

$$l = 4 \int_0^{\sinh^{-1} \frac{1}{2}} \cosh^2 \eta d\eta = 2 \sinh^{-1} \frac{1}{2} + \frac{\sqrt{5}}{2}.$$

Supponiamo ora di prendere un punto P_0 corrispondente al vettore \mathbf{x}_0 su una curva regolare \mathcal{C} e di aver scelto il parametro t nelle sue equazioni parametriche in modo tale che $\mathbf{x}(0) = \mathbf{x}_0$. Preso un generico punto P di \mathcal{C} corrispondente al generico valore t del parametro, definiamo *ascissa curvilinea* s su \mathcal{C} la quantità

$$s(t) = \pm \int_0^t \left\| \frac{d\mathbf{x}}{d\tau} \right\| d\tau,$$

dove vale il segno positivo se si vuole che s cresca al crescere di t , mentre vale il segno negativo se si richiede che s decresca al crescere di t . In valore assoluto, l'ascissa curvilinea misura la lunghezza dell'arco di curva da \mathbf{x}_0 a $\mathbf{x}(t)$. Nel seguito, se non specificato diversamente, assumeremo che s cresca al crescere del parametro t . Per la regolarità della curva, le derivate delle componenti di $\mathbf{x}(t)$ non si annullano mai contemporaneamente, ovvero $\|d\mathbf{x}/dt\| \neq 0$. L'ascissa curvilinea risulta quindi una funzione continua e strettamente monotona di t . Essa è dunque invertibile, con inversa $t(s)$ continua. Ne segue che si possono scrivere le equazioni parametriche della curva \mathcal{C} in funzione di s come $\mathbf{x} = \hat{\mathbf{x}}(t(s))$. Quest'ultima risulta sempre essere una rappresentazione regolare in quanto, dalla definizione stessa di ascissa curvilinea

$$\left\| \frac{d\mathbf{x}}{ds} \right\| = \left\| \frac{d\mathbf{x}}{dt} \right\| \left| \frac{dt}{ds} \right| = \left\| \frac{d\mathbf{x}}{dt} \right\| \frac{1}{|s'|} = 1 \neq 0.$$

Inoltre, poiché $s(t)$ è biunivoca, si avrà $s(t_1) \neq s(t_2)$ per $t_1 \neq t_2$, quindi $\mathbf{x}(s_1) \neq \mathbf{x}(s_2)$ se $s_1 \neq s_2$ ($s_1 = s(t_1)$, $s_2 = s(t_2)$).

- Esempi

1) Consideriamo l'equazione di una retta in forma parametrica

$$\begin{cases} x = x_0 + tv_x \\ y = y_0 + tv_y \\ z = z_0 + tv_z, \end{cases}$$

Dalla definizione di ascissa curvilinea abbiamo

$$s(t) = \int_0^t \sqrt{v_x^2 + v_y^2 + v_z^2} d\tau = \|\mathbf{v}\| t.$$

L'equazione parametrica della medesima retta in funzione della sua ascissa curvilinea è

$$\begin{cases} x = x_0 + \frac{v_x}{\|\mathbf{v}\|} s \\ y = y_0 + \frac{v_y}{\|\mathbf{v}\|} s \\ z = z_0 + \frac{v_z}{\|\mathbf{v}\|} s. \end{cases}$$

In tal caso si verifica subito che la lunghezza del segmento $\overline{P_0P}$ è proprio $|s|$.

2) Consideriamo l'equazione della circonferenza di raggio R e centro nell'origine, sul piano $z = 0$ di un riferimento cartesiano. Le equazioni parametriche si possono scrivere nella forma

$$\begin{cases} x = R \cos \psi \\ y = R \sin \psi \end{cases} \quad \psi \in [0, 2\pi).$$

In tal caso si ha

$$s(\psi) = \int_0^\psi R \sqrt{\sin^2 \tau + \cos^2 \tau} d\tau = R\psi.$$

che, come è noto, corrisponde, in valore assoluto, alla lunghezza dell'arco di circonferenza di apertura ψ .

3) Consideriamo le equazioni parametriche dell'elica cilindrica di raggio R , asse z e passo p . Abbiamo visto che si può scrivere

$$\begin{cases} x = R \cos \psi \\ y = R \sin \psi \\ z = \frac{p}{2\pi} \psi, \end{cases} \quad \psi \in \mathbb{R}.$$

Calcoliamo l'ascissa curvilinea. Si ha

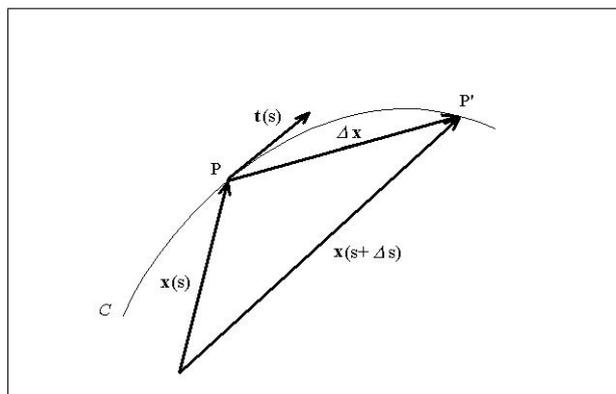
$$s(\psi) = \int_0^\psi \sqrt{R^2 \sin^2 \tau + R^2 \cos^2 \tau + \left(\frac{p}{2\pi}\right)^2} d\tau = \sqrt{R^2 + \left(\frac{p}{2\pi}\right)^2} \psi.$$

1.3 Terna intrinseca

Siano P e P' due punti di una curva regolare \mathcal{C} , corrispondenti a valori s e $s + \Delta s$ dell'ascissa curvilinea. Essi saranno individuati dai vettori $\mathbf{x}(s)$ e $\mathbf{x}(s + \Delta s)$. Al tendere di Δs a zero il vettore

$$\frac{\mathbf{x}(s + \Delta s) - \mathbf{x}(s)}{\Delta s},$$

tende ad assumere la direzione della tangente alla curva nel punto P di ascissa s .



Definiamo *versore tangente* alla curva in P il vettore

$$\mathbf{t}(s) = \frac{d\mathbf{x}}{ds} = \lim_{\Delta s \rightarrow 0} \frac{\mathbf{x}(s + \Delta s) - \mathbf{x}(s)}{\Delta s}.$$

Per verificare che si tratta effettivamente di un versore, osserviamo che, da quanto detto sull'ascissa curvilinea,

$$\left\| \frac{d\mathbf{x}}{ds} \right\| = 1, \quad \implies \quad \|\mathbf{t}(s)\| = 1.$$

Supponiamo ora che la funzione $\mathbf{x}(t)$ sia derivabile almeno due volte. Derivando ambo i membri dell'identità

$$\mathbf{t}(s) \cdot \mathbf{t}(s) = 1,$$

ricaviamo

$$2\mathbf{t}(s) \cdot \frac{d\mathbf{t}}{ds} = 0.$$

Se escludiamo il caso in cui la curva sia una retta, cioè supponiamo $d\mathbf{t}/ds \neq \mathbf{0}$, la condizione precedente implica che il vettore $d\mathbf{t}/ds$ risulta ortogonale al versore tangente alla curva in P . Chiameremo *normale* il versore \mathbf{n} di $d\mathbf{t}/ds$, ovvero, porremo

$$\frac{d\mathbf{t}}{ds} = c\mathbf{n},$$

dove la quantità $c = \|d\mathbf{t}/ds\|$ è detta *curvatura di flessione* della curva. Il suo reciproco $\rho = 1/c$ si dice *raggio di curvatura*.

Sempre nell'ipotesi che la curva non sia una retta, si definisce versore *binormale* il vettore

$$\mathbf{b} = \mathbf{t} \times \mathbf{n}.$$

I tre versori \mathbf{t} , \mathbf{n} , \mathbf{b} risultano mutuamente ortogonali e costituiscono la cosiddetta *terna intrinseca* della curva in P . Ovviamente, al variare del punto P su \mathcal{C} , varierà l'orientazione della terna intrinseca.

• Esempi

1) Calcoliamo la terna intrinseca e il raggio di curvatura di una circonferenza di raggio R centro nell'origine, nel piano $z = 0$ di un riferimento cartesiano. Abbiamo

$$\begin{cases} x = R \cos \psi \\ y = R \sin \psi, \end{cases} \quad s = R\psi \quad \implies \quad \begin{cases} x = R \cos(s/R) \\ y = R \sin(s/R). \end{cases}$$

Il versore tangente è dato da

$$\mathbf{t} = \frac{d\mathbf{x}}{ds} = -\sin(s/R)\mathbf{e}_1 + \cos(s/R)\mathbf{e}_2 = -\sin \psi \mathbf{e}_1 + \cos \psi \mathbf{e}_2,$$

da cui

$$\frac{d\mathbf{t}}{ds} = -\frac{1}{R} \cos(s/R)\mathbf{e}_1 - \frac{1}{R} \sin(s/R)\mathbf{e}_2.$$

confrontando con la definizione di \mathbf{n} abbiamo

$$\mathbf{n} = -\cos \psi \mathbf{e}_1 - \sin \psi \mathbf{e}_2, \quad \rho = R.$$

Infine, il versore binormale è

$$\mathbf{b} = \mathbf{t} \times \mathbf{n} = (-\sin \psi \mathbf{e}_1 + \cos \psi \mathbf{e}_2) \times (-\cos \psi \mathbf{e}_1 - \sin \psi \mathbf{e}_2) = \mathbf{e}_3.$$

2) Calcoliamo la terna intrinseca e la curvatura di flessione dell'elica cilindrica di asse z , raggio R e passo p . Dai risultati già ottenuti abbiamo

$$\begin{cases} x = R \cos \psi \\ y = R \sin \psi \\ z = \frac{p}{2\pi} \psi, \end{cases} \quad s(\psi) = \sqrt{R^2 + \left(\frac{p}{2\pi}\right)^2} \psi =: \alpha \psi \quad \Rightarrow \quad \begin{cases} x = R \cos(s/\alpha) \\ y = R \sin(s/\alpha) \\ z = \frac{p}{2\pi} \frac{s}{\alpha}. \end{cases}$$

Si ricava così,

$$\mathbf{t} = -\frac{R}{\alpha} \sin(s/\alpha) \mathbf{e}_1 + \frac{R}{\alpha} \cos(s/\alpha) \mathbf{e}_2 + \frac{p}{2\pi\alpha} \mathbf{e}_3 = \frac{R}{\alpha} (-\sin \psi \mathbf{e}_1 + \cos \psi \mathbf{e}_2 + \frac{p}{2\pi R} \mathbf{e}_3).$$

Derivando ulteriormente si ottiene

$$\frac{d\mathbf{t}}{ds} = -\frac{R}{\alpha^2} \cos \psi \mathbf{e}_1 - \frac{R}{\alpha^2} \sin \psi \mathbf{e}_2,$$

da cui si ricava

$$\mathbf{n} = -\cos \psi \mathbf{e}_1 - \sin \psi \mathbf{e}_2, \quad c = \frac{1}{\rho} = \frac{4\pi^2 R}{p^2 + 4\pi^2 R^2}.$$

Infine

$$\begin{aligned} \mathbf{b} &= \frac{R}{\alpha} (-\sin \psi \mathbf{e}_1 + \cos \psi \mathbf{e}_2 + \frac{p}{2\pi R} \mathbf{e}_3) \times (-\cos \psi \mathbf{e}_1 - \sin \psi \mathbf{e}_2) \\ &= \left(R^2 + \frac{p^2}{4\pi^2}\right)^{-1/2} \left[\frac{p}{2\pi} (\sin \psi \mathbf{e}_1 - \cos \psi \mathbf{e}_2) + R \mathbf{e}_3\right]. \end{aligned}$$

Osserviamo che il calcolo dei versori della terna intrinseca può essere effettuato anche senza conoscere l'ascissa curvilinea. Infatti, data la rappresentazione regolare $\mathbf{x} = \hat{\mathbf{x}}(s(t))$ della curva \mathcal{C} , si ha

$$\begin{aligned} \mathbf{t} &= \frac{d\mathbf{x}}{ds} = \frac{d\mathbf{x}}{dt} \frac{dt}{ds} = \frac{\mathbf{x}'}{s'} = \frac{\mathbf{x}'}{\|\mathbf{x}'\|}, \\ \mathbf{n} &= \frac{\frac{d\mathbf{t}}{ds}}{\left\|\frac{d\mathbf{t}}{ds}\right\|} = \frac{\mathbf{t}'/s'}{\|\mathbf{t}'/s'\|} = \frac{\mathbf{t}'}{\|\mathbf{t}'\|}. \end{aligned}$$

Ne segue anche

$$c = \frac{1}{\rho} = \frac{\|\mathbf{t}'\|}{\|\mathbf{x}'\|}.$$

- Esempio

Calcoliamo il versore tangente della elica conica di equazioni parametriche

$$\begin{cases} x = R\psi \cos \psi \\ y = R\psi \sin \psi \\ z = \frac{p}{2\pi} \psi \end{cases} \quad \psi \in \mathbb{R}.$$

Si ha

$$\mathbf{x}' = (R \cos \psi - R\psi \sin \psi) \mathbf{e}_1 + (R \sin \psi + R\psi \cos \psi) \mathbf{e}_2 + \frac{p}{2\pi} \mathbf{e}_3,$$

da cui

$$\mathbf{t} = \frac{\mathbf{x}'}{\|\mathbf{x}'\|} = \frac{R[(\cos \psi - \psi \sin \psi) \mathbf{e}_1 + (\sin \psi + \psi \cos \psi) \mathbf{e}_2 + \frac{p}{2\pi R} \mathbf{e}_3]}{\sqrt{R^2(1 + \psi^2) + \left(\frac{p}{2\pi}\right)^2}}.$$

Capitolo 2

Funzioni di più variabili

2.1 Limiti e funzioni continue

Consideriamo l'insieme \mathbb{R}^n delle n -uple ordinate di numeri reali. Introdotta la usuale struttura euclidea in \mathbb{R}^n , sia $\{\mathbf{e}_i\}$, ($i = 1, \dots, n$) la base ortonormale canonica. Detto $\mathbf{x} = (x_1, x_2, \dots, x_n)$ un vettore di \mathbb{R}^n , la sua norma sarà $\|\mathbf{x}\| = \sqrt{x_1^2 + x_2^2 + \dots + x_n^2}$. In analogia con la notazione dei sottoinsiemi di \mathbb{R} , diamo le seguenti definizioni.

Si dice *intorno sferico* di centro \mathbf{x}_0 e raggio r ($r > 0$) l'insieme $B_r(\mathbf{x}_0)$, definito da

$$B_r(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{R}^n : \|\mathbf{x} - \mathbf{x}_0\| < r\}.$$

Nel caso in cui si escluda da questo il solo punto \mathbf{x}_0 scriveremo

$$B_r^0(\mathbf{x}_0) = B_r(\mathbf{x}_0) \setminus \{\mathbf{x}_0\}.$$

Diremo che l'insieme $A \subset \mathbb{R}^n$ è un *aperto* di \mathbb{R}^n se per ogni $\mathbf{x} \in A$ è possibile trovare un intorno $B_r(\mathbf{x})$ interamente contenuto in A . Diremo poi che l'insieme C è *chiuso* se è il complementare di un aperto. Chiameremo punto di accumulazione per $A \subset \mathbb{R}^n$ un punto \mathbf{x}_0 tale che, preso un qualunque intorno sferico $B_r(\mathbf{x}_0)$, esiste almeno un punto di A distinto da \mathbf{x}_0 appartenente a $B_r(\mathbf{x}_0)$. Diremo che un insieme $A \subset \mathbb{R}^n$ è *limitato*, se esiste un $k \in \mathbb{R}^+$ tale che

$$\|\mathbf{x}_1 - \mathbf{x}_2\| \leq k, \quad \forall \mathbf{x}_1, \mathbf{x}_2 \in A.$$

Si dice poi *diametro* di un insieme limitato $A \subset \mathbb{R}^n$, la quantità

$$\sup_{\mathbf{x}_1, \mathbf{x}_2 \in A} \|\mathbf{x}_1 - \mathbf{x}_2\|.$$

Consideriamo una funzione f definita in un sottoinsieme A di \mathbb{R}^n a valori in \mathbb{R} e sia \mathbf{x}_0 un punto di accumulazione per A . Diremo che la funzione $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ tende al limite l per \mathbf{x} che tende a \mathbf{x}_0 , e scriveremo $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = l$ se, $\forall \varepsilon > 0$, $\exists \delta > 0 : \mathbf{x} \in B_\delta^0(\mathbf{x}_0) \cap A \Rightarrow |f(\mathbf{x}) - l| < \varepsilon$. Analogamente diremo che $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = +\infty (-\infty)$ se $\forall k > 0$, $\exists \delta > 0 : \mathbf{x} \in B_\delta^0(\mathbf{x}_0) \cap A \Rightarrow f(\mathbf{x}) > k (< -k)$.

- Esempi

1) Data la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, definita da

$$f(\mathbf{x}) = x - 2y, \quad \mathbf{x} = (x, y) \in \mathbb{R}^2,$$

verificare che $\lim_{\mathbf{x} \rightarrow \mathbf{0}} f(\mathbf{x}) = 0$. Poiché si ha

$$|x - 2y| \leq |x| + 2|y| \leq 3\sqrt{x^2 + y^2} = 3\|\mathbf{x} - \mathbf{0}\|,$$

ne segue che, fissato un $\varepsilon > 0$, scegliendo $\delta = \varepsilon/3$, si ha

$$|x - 2y| < \varepsilon, \quad \forall \mathbf{x} \in B_\delta(\mathbf{0}),$$

ed il limite è verificato.

2) Data la funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, definita da

$$f(\mathbf{x}) = \frac{\sin(x - 2y)}{x - 2y},$$

verificare che $\lim_{\mathbf{x} \rightarrow \mathbf{0}} f(\mathbf{x}) = 1$. La funzione è definita in $A = \mathbb{R}^2 \setminus \{\mathbf{x} \in \mathbb{R}^2 : x - 2y = 0\}$, cioè in tutti i punti del piano \mathbb{R}^2 esclusa la retta di equazione $x - 2y = 0$. Sappiamo che $\lim_{z \rightarrow 0} \frac{\sin z}{z} = 1$, ovvero,

$$\forall \varepsilon > 0, \exists \sigma > 0 : 0 < |z| < \sigma \Rightarrow \left| \frac{\sin z}{z} - 1 \right| < \varepsilon.$$

In corrispondenza di σ , per il limite dell'esempio precedente, possiamo trovare un $\delta > 0$ tale che

$$\|\mathbf{x} - \mathbf{0}\| < \delta \Rightarrow |x - 2y| < \sigma.$$

Mettendo insieme i due risultati, otteniamo

$$\left| \frac{\sin(x - 2y)}{x - 2y} - 1 \right| < \varepsilon, \quad \forall \mathbf{x} \in B_\delta^0(\mathbf{0}) \cap A.$$

3) Data la funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, definita da

$$f(\mathbf{x}) = \frac{x^3 y}{x^6 + y^2},$$

vogliamo studiare il $\lim_{\mathbf{x} \rightarrow \mathbf{0}} f(\mathbf{x})$. La funzione è definita in $A = \mathbb{R}^2 \setminus \{\mathbf{0}\}$. Osserviamo che, presa la restrizione della f in $A' = \{\mathbf{x} \in \mathbb{R}^2 : y - mx = 0\}$, essendo $m \in \mathbb{R}$, si ha

$$f(\mathbf{x})|_{A'} = \frac{mx^4}{x^2(x^4 + m^2)} = \frac{mx^2}{x^4 + m^2}.$$

Se facciamo tendere $\mathbf{x} \rightarrow \mathbf{0}$, in A' , otteniamo

$$\lim_{\mathbf{x} \rightarrow \mathbf{0}} f(\mathbf{x})|_{A'} = 0.$$

D'altra parte, se prendiamo la restrizione della funzione f in $A'' = \{\mathbf{x} \in \mathbb{R}^2 : y - x^3 = 0\}$, otteniamo

$$f(\mathbf{x})|_{A''} = \frac{x^6}{x^6 + x^6} = \frac{1}{2} \Rightarrow \lim_{\mathbf{x} \rightarrow \mathbf{0}} f(\mathbf{x})|_{A''} = \frac{1}{2}.$$

In definitiva, comunque scegliamo un intorno sferico di $\mathbf{x} = \mathbf{0}$, in questo intorno la funzione assumerà sempre tutti i valori dell'intervallo $[0, 1/2]$. La funzione dunque non ammette limite per $\mathbf{x} \rightarrow \mathbf{0}$.

Sia f definita in un dominio $A \subset \mathbb{R}^n$ non limitato. Diciamo che la funzione f tende a l al tendere di \mathbf{x} all'infinito e scriveremo $\lim_{\mathbf{x} \rightarrow \infty} f(\mathbf{x}) = l$, se, $\forall \varepsilon > 0, \exists \delta > 0 : \forall \mathbf{x} \in A, \mathbf{x} \notin B_\delta(\mathbf{0}) \Rightarrow |f(\mathbf{x}) - l| < \varepsilon$. Analogamente si definisce il limite infinito per $\mathbf{x} \rightarrow \infty$.

• Esempio

Data la funzione $f : A \subset \mathbb{R}^3 \rightarrow \mathbb{R}$, definita da

$$f(\mathbf{x}) = \frac{x + y - z}{x^2 + y^2 + z^2}, \quad \mathbf{x} = (x, y, z) \in \mathbb{R}^3 \setminus \{\mathbf{0}\},$$

verifichiamo che $\lim_{\mathbf{x} \rightarrow \infty} f(\mathbf{x}) = 0$. Osserviamo che si ha

$$\left| \frac{x + y - z}{x^2 + y^2 + z^2} \right| \leq \frac{|x| + |y| + |z|}{\|\mathbf{x}\|^2} \leq \frac{3\|\mathbf{x}\|}{\|\mathbf{x}\|^2} = \frac{3}{\|\mathbf{x}\|}.$$

Allora, fissato un $\varepsilon > 0$, avremo $|f(\mathbf{x})| < \varepsilon$, purchè

$$\frac{3}{\|\mathbf{x}\|} < \varepsilon, \quad \Rightarrow \quad \|\mathbf{x}\| > \frac{3}{\varepsilon}.$$

Scelto $\delta = 3/\varepsilon$, otteniamo

$$|f(\mathbf{x})| < \varepsilon \quad \text{quando} \quad \mathbf{x} \in A, \mathbf{x} \notin B_\delta(\mathbf{0}).$$

Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e \mathbf{x}_0 un punto di accumulazione di A . Diremo che f è *continua* in \mathbf{x}_0 se esiste il $\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x})$ e si ha

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0).$$

Si dice che la funzione f è continua in A se è continua in ogni punto di A . Valgono inoltre proprietà analoghe a quelle delle funzioni di una variabile reale. In particolare, la somma ed il prodotto di funzioni continue sono funzioni continue. Valgono inoltre i seguenti teoremi.

Teorema 2.1 *Se f è continua in $\mathbf{x}_0 \in \mathbb{R}^n$ e si ha $f(\mathbf{x}_0) > 0 (< 0)$, allora esiste un intorno sferico di \mathbf{x}_0 in ogni punto del quale $f(\mathbf{x}) > 0 (< 0)$.*

Teorema 2.2 *(di Weierstrass) Se f è continua in un insieme $A \subset \mathbb{R}^n$ chiuso e limitato, allora f è limitata (cioè la sua immagine in \mathbb{R} è un insieme limitato) ed ammette un massimo e un minimo.*

2.2 Derivate parziali

Sia f definita in un aperto $A \subset \mathbb{R}^n$ e sia \mathbf{x}_0 un punto di accumulazione per A . Consideriamo un vettore \mathbf{v} di \mathbb{R}^n . Chiameremo *retta* passante per \mathbf{x}_0 e di direzione \mathbf{v} , il luogo dei punti $r = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \mathbf{x}_0 + \mathbf{v}t, t \in \mathbb{R}\}$. Consideriamo la restrizione di f ad r ,

$$f(\mathbf{x})|_r = f(\mathbf{x}_0 + \mathbf{v}t).$$

Tale restrizione è una funzione della sola variabile reale t . Se questa restrizione è derivabile in $t = 0$, ovvero se esiste il

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

lo chiameremo *derivata direzionale* di f nella direzione \mathbf{v} , in \mathbf{x}_0 e lo indicheremo con $D_{\mathbf{v}}f(\mathbf{x}_0)$. Se \mathbf{v} coincide con uno dei versori della base ortonormale canonica, \mathbf{e}_k , porremo

$$\frac{\partial f}{\partial x_k}(\mathbf{x}_0) := D_{\mathbf{e}_k}f(\mathbf{x}_0),$$

che chiameremo *derivata parziale* di f rispetto a x_k , in \mathbf{x}_0 . Poiché risulta

$$\frac{\partial f}{\partial x_k}(\mathbf{x}_0) = \lim_{t \rightarrow 0} \frac{1}{t} [f(x_1^0, \dots, x_k^0 + t, \dots, x_n^0) - f(x_1^0, \dots, x_n^0)],$$

la derivata parziale di una funzione rispetto alla variabile x_k è la derivata della funzione che si ottiene da f fissando tutte le altre variabili e facendo variare solo la k -esima.

• Esempi

1) Calcolare la derivata direzionale della funzione

$$f(x_1, x_2, x_3, x_4) = \ln(x_1^2 + x_2^2) + x_3^2 - x_4^2,$$

nella direzione $\mathbf{v} = \frac{1}{2}(1, 1, -1, 1)$ nel punto $\mathbf{x}_0 = (1, 0, 0, 1)$.

Secondo la definizione data avremo

$$D_{\mathbf{v}}f(\mathbf{x}_0) = \lim_{t \rightarrow 0} \frac{\ln \left[\left(1 + \frac{1}{2}t\right)^2 + \left(\frac{1}{2}t\right)^2 \right] + \left(-\frac{1}{2}t\right)^2 - \left(1 + \frac{1}{2}t\right)^2 - (\ln 1 - 1)}{t} = \lim_{t \rightarrow 0} \frac{\ln \left(1 + t + \frac{1}{2}t^2\right) - t}{t}$$

da cui

$$D_{\mathbf{v}}f(\mathbf{x}_0) = \lim_{t \rightarrow 0} \frac{\ln \left(1 + t + \frac{1}{2}t^2\right)}{t} - 1 = 0$$

2) Data la funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, definita da

$$f(\mathbf{x}) = x \ln(xy^2),$$

si ha $A = \{(x, y) : x > 0, y \neq 0\}$. Calcoliamo le sue derivate parziali.

$$\begin{aligned} \frac{\partial f}{\partial x} &= \ln(xy^2) + x \frac{1}{xy^2} y^2 = 1 + \ln(xy^2), \\ \frac{\partial f}{\partial y} &= x \frac{1}{xy^2} 2xy = \frac{2x}{y}, \end{aligned}$$

che risultano definite in tutto l'aperto A .

3) Data la funzione $f : A \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ definita da

$$f(\mathbf{x}) = xyz - \sin(xz) + ye^{x^2z},$$

si ha $A = \mathbb{R}^3$. Calcoliamo le derivate parziali nel punto $\mathbf{x}_0 = (1, 2, \pi)$.

$$\begin{aligned} \frac{\partial f}{\partial x} &= yz - z \cos(xz) + 2xyz e^{x^2z}, \\ \frac{\partial f}{\partial y} &= xz + e^{x^2z} \\ \frac{\partial f}{\partial z} &= xy - x \cos(xz) + x^2 y e^{x^2z}, \end{aligned}$$

che risultano definite in A . Per $\mathbf{x} = \mathbf{x}_0$ si ha

$$\frac{\partial f}{\partial x}(\mathbf{x}_0) = \pi(3 + 4e^\pi), \quad \frac{\partial f}{\partial y}(\mathbf{x}_0) = \pi + e^\pi, \quad \frac{\partial f}{\partial z}(\mathbf{x}_0) = 3 + 2e^\pi.$$

4) Data la funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da

$$f(\mathbf{x}) = 2\sqrt{xy^3} - \sqrt{x}\sqrt{y},$$

si ha $A = \{(x, y) : x \geq 0, y \geq 0\}$. Calcoliamo le derivate parziali.

$$\begin{aligned} \frac{\partial f}{\partial x} &= \frac{1}{2\sqrt{x}}(2\sqrt{y^3} - \sqrt{y}) \\ \frac{\partial f}{\partial y} &= \sqrt{x} \left(3\sqrt{y} - \frac{1}{4} \frac{1}{\sqrt{y^3}} \right) \end{aligned}$$

che risultano definite rispettivamente in $A \setminus \{x = 0\}$ e $A \setminus \{y = 0\}$.

2.3 Funzioni differenziabili

Sia f definita in $A \subset \mathbb{R}^n$ e sia $\mathbf{x}_0 \in A$, un punto di accumulazione. Diremo che f è *differenziabile* in \mathbf{x}_0 se esiste una applicazione lineare $L_{\mathbf{x}_0}$ da \mathbb{R}^n in \mathbb{R} tale che

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{f(\mathbf{x}_0 + \mathbf{h}) - f(\mathbf{x}_0) - L_{\mathbf{x}_0}(\mathbf{h})}{\|\mathbf{h}\|} = 0.$$

con $\mathbf{h} \in \mathbb{R}^n$. Se f è differenziabile in ogni punto di A , diremo che è differenziabile in A .

Dimostriamo che una funzione differenziabile in \mathbf{x}_0 ammette derivata direzionale in qualunque direzione individuata da un versore \mathbf{v} e quindi ammette derivate parziali. Posto $\mathbf{h} = \mathbf{v}t$ si ha infatti, per la linearità di $L_{\mathbf{x}_0}$,

$$0 = \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + \mathbf{v}t) - f(\mathbf{x}_0) - L_{\mathbf{x}_0}(\mathbf{v}t)}{\|\mathbf{v}\||t|} = \lim_{t \rightarrow 0} \left[\frac{f(\mathbf{x}_0 + \mathbf{v}t) - f(\mathbf{x}_0)}{|t|} - \frac{tL_{\mathbf{x}_0}(\mathbf{v})}{|t|} \right]$$

da cui otteniamo

$$L_{\mathbf{x}_0}(\mathbf{v}) = \lim_{t \rightarrow 0} \frac{f(\mathbf{x}_0 + \mathbf{v}t) - f(\mathbf{x}_0)}{t} \quad \Rightarrow \quad D_{\mathbf{v}}f(\mathbf{x}_0) = L_{\mathbf{x}_0}(\mathbf{v}).$$

In particolare, se $\mathbf{v} = \mathbf{e}_k$, otteniamo

$$L_{\mathbf{x}_0}(\mathbf{e}_k) = \frac{\partial f}{\partial x_k}(\mathbf{x}_0).$$

L'applicazione $L_{\mathbf{x}_0}$ si dice *differenziale* di f in \mathbf{x}_0 e la si indica più spesso con $df(\mathbf{x}_0)$. I risultati precedenti si scriveranno allora nella forma

$$D_{\mathbf{v}}f(\mathbf{x}_0) = df(\mathbf{x}_0)(\mathbf{v}), \quad \frac{\partial f}{\partial x_k}(\mathbf{x}_0) = df(\mathbf{x}_0)(\mathbf{e}_k),$$

e indicano che l'applicazione di $df(\mathbf{x}_0)$ ad un vettore $\mathbf{h} \in \mathbb{R}^n$ equivale alla norma di \mathbf{h} per la derivata direzionale di f in \mathbf{x}_0 . Notiamo che, essendo $\mathbf{v} = \sum_{i=1}^n v_i \mathbf{e}_i$, risulta, per la linearità di $df(\mathbf{x}_0)$,

$$df(\mathbf{x}_0)(\mathbf{v}) = df(\mathbf{x}_0) \sum_{i=1}^n v_i \mathbf{e}_i = \sum_{i=1}^n v_i df(\mathbf{x}_0)(\mathbf{e}_i) = \sum_{i=1}^n v_i \frac{\partial f}{\partial x_i}(\mathbf{x}_0).$$

Se f è una funzione differenziabile in \mathbf{x}_0 , si dice *gradiente* di f in \mathbf{x}_0 , il vettore

$$\nabla f(\mathbf{x}_0) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0) \mathbf{e}_i.$$

Il differenziale di f si può esprimere allora come l'applicazione di un prodotto scalare,

$$df(\mathbf{x}_0)(\mathbf{v}) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v}, \quad \forall \mathbf{v} \in \mathbb{R}^n.$$

Consideriamo in particolare le funzioni f_i così definite in \mathbb{R}^n ,

$$f_1(\mathbf{x}) = x_1, \quad f_2(\mathbf{x}) = x_2, \quad \dots, \quad f_n(\mathbf{x}) = x_n.$$

Poiché ogni funzione f_i è differenziabile,

$$df_i(\mathbf{x}_0)(\mathbf{v}) = \nabla f_i(\mathbf{x}_0) \cdot \mathbf{v} = \sum_{j=1}^n \frac{\partial f_i}{\partial x_j}(\mathbf{x}_0) v_j = \sum_{j=1}^n \delta_{ij} v_j = v_i, \quad (i = 1, \dots, n)$$

Ne segue che, posto in questo caso $df_i = dx_i$, e risultando i dx_i indipendenti da \mathbf{x}_0 ,

$$dx_i(\mathbf{v}) = v_i.$$

Possiamo allora riscrivere il differenziale di una qualunque funzione f in termini della n -upla di applicazioni lineari dx_i , ($i = 1, \dots, n$). Posto $d\mathbf{x} = (dx_1, dx_2, \dots, dx_n)$, avremo

$$df(\mathbf{x}_0)(\mathbf{v}) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0) dx_i(\mathbf{v}) = \nabla f(\mathbf{x}_0) \cdot d\mathbf{x}(\mathbf{v}), \quad \forall \mathbf{v} \in \mathbb{R}^n.$$

Osservazione. Nel caso $n = 1$, la definizione di funzione differenziabile si riduce a quella di derivabilità di una funzione di una variabile reale. In tal caso si ha infatti, per $h \in \mathbb{R}$

$$\lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0) - \frac{df}{dx}(x_0)h}{|h|} = 0.$$

da cui segue

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{f(x_0 + h) - f(x_0)}{h}.$$

• Esempi

1) Rifare il calcolo della derivata direzionale dell'esempio 1) del paragrafo precedente utilizzando il gradiente della funzione.

La funzione è

$$f(x_1, x_2, x_3, x_4) = \ln(x_1^2 + x_2^2) + x_3^2 - x_4^2.$$

Poiché si ha

$$\frac{\partial f}{\partial x_1} = \frac{2x_1}{x_1^2 + x_2^2}, \quad \frac{\partial f}{\partial x_2} = \frac{2x_2}{x_1^2 + x_2^2}, \quad \frac{\partial f}{\partial x_3} = 2x_3, \quad \frac{\partial f}{\partial x_4} = -2x_4,$$

il gradiente in $\mathbf{x}_0 = (1, 0, 0, 1)$ è dato da

$$\nabla f(\mathbf{x}_0) = (2, 0, 0, -2)$$

Ne segue

$$D_{\mathbf{v}}f(\mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v} = \frac{1}{2}(2, 0, 0, -2)(1, 1, -1, 1)^T = 0$$

2) Dimostrare che, dato il vettore costante $\bar{\mathbf{x}}$, il gradiente della funzione $f(\mathbf{x}) = \mathbf{x} \cdot \bar{\mathbf{x}}$, definita in \mathbb{R}^n è dato da $\bar{\mathbf{x}}$.

Dalla definizione di differenziale in \mathbf{x}_0 abbiamo

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} = 0.$$

Nel nostro caso otteniamo

$$\frac{|\mathbf{x} \cdot \bar{\mathbf{x}} - \mathbf{x}_0 \cdot \bar{\mathbf{x}} - \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)|}{\|\mathbf{x} - \mathbf{x}_0\|} = \frac{|(\mathbf{x} - \mathbf{x}_0) \cdot [\bar{\mathbf{x}} - \nabla f(\mathbf{x}_0)]|}{\|\mathbf{x} - \mathbf{x}_0\|}$$

Posto $\mathbf{x} - \mathbf{x}_0 = \mathbf{v}t$, dove \mathbf{v} è un arbitrario versore e traducendo il limite per $\mathbf{x} \rightarrow \mathbf{x}_0$ nel limite per $t \rightarrow 0$, otteniamo

$$0 = \lim_{t \rightarrow 0} \frac{|t|\mathbf{v} \cdot [\bar{\mathbf{x}} - \nabla f(\mathbf{x}_0)]|}{\|\mathbf{v}\||t|} = |\mathbf{v} \cdot [\bar{\mathbf{x}} - \nabla f(\mathbf{x}_0)]|.$$

Quest'ultimo risultato vale per ogni scelta del versore $\mathbf{v} \in \mathbb{R}^n$ e per ogni punto \mathbf{x}_0 . Ne segue dunque la tesi.

3) Dimostrare che il gradiente di $f(\mathbf{x}) = \mathbf{x} \cdot \mathbf{A}\mathbf{x}$, dove \mathbf{A} è un operatore lineare indipendente da \mathbf{x} in \mathbb{R}^n , è dato da $\mathbf{A}\mathbf{x} + \mathbf{A}^T\mathbf{x}$.

Sempre per la definizione di differenziale, scelto un $\mathbf{x}_0 \in \mathbb{R}^n$, in questo caso abbiamo

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{\mathbf{x} \cdot \mathbf{A}\mathbf{x} - \mathbf{x}_0 \cdot \mathbf{A}\mathbf{x}_0 - \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} = 0.$$

D'altra parte, si può scrivere

$$\begin{aligned} & \frac{|\mathbf{x} \cdot \mathbf{A}\mathbf{x} - \mathbf{x}_0 \cdot \mathbf{A}\mathbf{x}_0 - \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)|}{\|\mathbf{x} - \mathbf{x}_0\|} \\ &= \frac{|\mathbf{x} \cdot \mathbf{A}\mathbf{x} - \mathbf{x}_0 \cdot \mathbf{A}\mathbf{x} + \mathbf{x}_0 \cdot \mathbf{A}\mathbf{x} - \mathbf{x}_0 \cdot \mathbf{A}\mathbf{x}_0 - \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)|}{\|\mathbf{x} - \mathbf{x}_0\|} \\ &= \frac{|(\mathbf{x} - \mathbf{x}_0) \cdot [\mathbf{A}\mathbf{x} + \mathbf{A}^T\mathbf{x}_0 - \nabla f(\mathbf{x}_0)]|}{\|\mathbf{x} - \mathbf{x}_0\|} \end{aligned}$$

Come nel caso precedente poniamo $\mathbf{x} - \mathbf{x}_0 = \mathbf{v}t$ e passiamo al limite per $t \rightarrow 0$. Otteniamo

$$0 = \lim_{t \rightarrow 0} \frac{|t|\mathbf{v} \cdot [\mathbf{A}(\mathbf{x}_0 + \mathbf{v}t) + \mathbf{A}^T\mathbf{x}_0 - \nabla f(\mathbf{x}_0)]|}{\|\mathbf{v}\||t|} = |\mathbf{v} \cdot [\mathbf{A}\mathbf{x}_0 + \mathbf{A}^T\mathbf{x}_0 - \nabla f(\mathbf{x}_0)]|.$$

Questo risultato vale per ogni versore \mathbf{v} e per ogni punto \mathbf{x}_0 . Ne segue il risultato richiesto.

Dimostriamo il seguente importante risultato.

Teorema 2.3 *Ogni funzione differenziabile in \mathbf{x}_0 è continua in \mathbf{x}_0 .*

Dim. Calcoliamo la differenza $f(\mathbf{x}) - f(\mathbf{x}_0)$. Si ha

$$f(\mathbf{x}) - f(\mathbf{x}_0) = \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} \|\mathbf{x} - \mathbf{x}_0\| + df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0).$$

Facendo uso della disuguaglianza triangolare, della disuguaglianza di Schwarz e osservando che $df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)$, otteniamo

$$\begin{aligned} |f(\mathbf{x}) - f(\mathbf{x}_0)| &\leq \left| \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} \|\mathbf{x} - \mathbf{x}_0\| \right| + |\nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0)| \\ &\leq \left| \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} \right| \|\mathbf{x} - \mathbf{x}_0\| + \|\nabla f(\mathbf{x}_0)\| \|\mathbf{x} - \mathbf{x}_0\| \end{aligned}$$

Infine, passando al limite per $\mathbf{x} \rightarrow \mathbf{x}_0$ e sfruttando la differenziabilità, ricaviamo

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} |f(\mathbf{x}) - f(\mathbf{x}_0)| \leq 0 + \|\nabla f(\mathbf{x}_0)\| \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \|\mathbf{x} - \mathbf{x}_0\| = 0,$$

da cui segue la continuità di f .

Teorema 2.4 *Condizione sufficiente affinché la funzione f sia differenziabile in \mathbf{x}_0 è che esista un intorno sferico $B_r(\mathbf{x}_0)$ in cui f ammette derivate parziali e che queste siano continue in \mathbf{x}_0 .*

Dim. Dimostriamo il teorema nel caso più semplice, quello delle funzioni di due variabili. Sia f definita in $A \subset \mathbb{R}^2$ e sia $\mathbf{x}_0 \in A$, $\mathbf{x} \in B_r(\mathbf{x}_0) \cap A$. Posto

$$f(\mathbf{x}) - f(\mathbf{x}_0) = f(x, y) - f(x_0, y_0) = f(x, y) - f(x_0, y) + f(x_0, y) - f(x_0, y_0),$$

supporremo che la funzione abbia derivate parziali continue in \mathbf{x}_0 . Applicando il teorema del valor medio alle due differenze $f(x, y) - f(x_0, y)$ e $f(x_0, y) - f(x_0, y_0)$, rispettivamente negli intorni I_x di x_0 e I_y di y_0 (con $I_x \times I_y \subset B_r(\mathbf{x}_0)$), otteniamo

$$\begin{aligned} f(x, y) - f(x_0, y) &= \frac{\partial f}{\partial x}(\xi_1, y)(x - x_0), \quad \xi_1 \in I_x, \\ f(x_0, y) - f(x_0, y_0) &= \frac{\partial f}{\partial y}(x_0, \xi_2)(y - y_0), \quad \xi_2 \in I_y. \end{aligned}$$

Poniamo ora

$$N(\mathbf{x}, \mathbf{x}_0) = f(\mathbf{x}) - f(\mathbf{x}_0) - \frac{\partial f}{\partial x}(\mathbf{x}_0)(x - x_0) - \frac{\partial f}{\partial y}(\mathbf{x}_0)(y - y_0),$$

e sostituiamo i risultati precedenti.

$$N(\mathbf{x}, \mathbf{x}_0) = \left[\frac{\partial f}{\partial x}(\xi_1, y) - \frac{\partial f}{\partial x}(\mathbf{x}_0) \right] (x - x_0) + \left[\frac{\partial f}{\partial y}(x_0, \xi_2) - \frac{\partial f}{\partial y}(\mathbf{x}_0) \right] (y - y_0).$$

Si ottiene così, in $I_x \times I_y$,

$$\begin{aligned} |N(\mathbf{x}, \mathbf{x}_0)| &\leq \left| \frac{\partial f}{\partial x}(\xi_1, y) - \frac{\partial f}{\partial x}(\mathbf{x}_0) \right| |x - x_0| + \left| \frac{\partial f}{\partial y}(x_0, \xi_2) - \frac{\partial f}{\partial y}(\mathbf{x}_0) \right| |y - y_0| \\ &\leq \left\{ \left| \frac{\partial f}{\partial x}(\xi_1, y) - \frac{\partial f}{\partial x}(\mathbf{x}_0) \right| + \left| \frac{\partial f}{\partial y}(x_0, \xi_2) - \frac{\partial f}{\partial y}(\mathbf{x}_0) \right| \right\} \|\mathbf{x} - \mathbf{x}_0\|. \end{aligned}$$

Per la continuità delle derivate parziali, passando al limite per $\mathbf{x} \rightarrow \mathbf{x}_0$ si arriva a

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{|N(\mathbf{x}, \mathbf{x}_0)|}{\|\mathbf{x} - \mathbf{x}_0\|} = 0.$$

Ne segue la differenziabilità di f in \mathbf{x}_0 .

La differenziabilità di una funzione di due variabili $z = f(x, y)$ in un punto (x_0, y_0) , consente di definire il piano tangente al *grafico* della funzione in $(x_0, y_0, f(x_0, y_0))$. A tale scopo denotiamo con $\{(\mathbf{x}, f(\mathbf{x})) \subset \mathbb{R}^3, \mathbf{x} \in A\}$, il *grafico* della funzione f . Il punto $(x_0, y_0, z_0 = f(x_0, y_0))$ apparterrà a tale grafico. Consideriamo la funzione reale, di due variabili,

$$\phi(\mathbf{x}) = f(\mathbf{x}_0) + df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0).$$

Posto $z = \phi(\mathbf{x})$, si può scrivere

$$z - z_0 = \nabla f(\mathbf{x}_0) \cdot (\mathbf{x} - \mathbf{x}_0) = \frac{\partial f}{\partial x}(\mathbf{x}_0)(x - x_0) + \frac{\partial f}{\partial y}(\mathbf{x}_0)(y - y_0).$$

Quest'ultima equazione rappresenta un piano in \mathbb{R}^3 . Tale piano passa per il punto $(\mathbf{x}_0, f(\mathbf{x}_0))$ ed il versore \mathbf{n} , ad esso normale, può essere ricavato dal gradiente di f in \mathbf{x}_0 . Si ha infatti

$$\mathbf{n} = \frac{\frac{\partial f}{\partial x}(\mathbf{x}_0)\mathbf{e}_1 + \frac{\partial f}{\partial y}(\mathbf{x}_0)\mathbf{e}_2 - \mathbf{e}_3}{\sqrt{\left[\frac{\partial f}{\partial x}(\mathbf{x}_0)\right]^2 + \left[\frac{\partial f}{\partial y}(\mathbf{x}_0)\right]^2 + 1}}$$

Notiamo che, in forza della differenziabilità di f in \mathbf{x}_0 ,

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{f(\mathbf{x}) - \phi(\mathbf{x})}{\|\mathbf{x} - \mathbf{x}_0\|} = \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{f(\mathbf{x}) - f(\mathbf{x}_0) - df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|} = 0.$$

Ne segue che la differenza $f(\mathbf{x}) - \phi(\mathbf{x})$ è un infinitesimo di ordine superiore a $\|\mathbf{x} - \mathbf{x}_0\|$ quando $\mathbf{x} \rightarrow \mathbf{x}_0$. In analogia al caso della tangente al grafico di una funzione di una variabile reale, questa proprietà corrisponde alla condizione di tangenza del piano $z = \phi(\mathbf{x})$ al grafico di f in \mathbf{x}_0 .

• Esempio

Verificare che la funzione di due variabili

$$f(x, y) = x^2 + \sin y$$

è differenziabile in $\mathbf{x}_0 = (1, 0)$ e determinare l'equazione del piano tangente in $(\mathbf{x}_0, f(\mathbf{x}_0))$. La funzione è definita in tutto \mathbb{R}^2 e si ha

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = \cos y.$$

Le derivate parziali sono funzioni continue in tutto \mathbb{R}^2 . Quindi la funzione f è differenziabile in tutto \mathbb{R}^2 . Poiché si ha

$$\frac{\partial f}{\partial x}(\mathbf{x}_0) = 2, \quad \frac{\partial f}{\partial y}(\mathbf{x}_0) = 1,$$

ed essendo $f(\mathbf{x}_0) = 1$, si ottiene l'equazione del piano tangente,

$$z - 1 = 2(x - 1) + y.$$

2.4 Funzioni vettoriali e funzioni composte

Consideriamo una funzione $\mathbf{f} : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ che ad n -uple di numeri reali fa corrispondere m -uple di numeri reali. È possibile estendere i concetti di limite e di continuità per tali funzioni, in modo perfettamente analogo a quanto fatto per le funzioni vettoriali di una variabile reale. In particolare, denotate con f_1, f_2, \dots, f_m , le componenti di \mathbf{f} rispetto alla base canonica in \mathbb{R}^m , si dimostra facilmente che

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \mathbf{f}(\mathbf{x}) = \mathbf{l} \quad \text{se e solo se} \quad \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} f_i(\mathbf{x}) = l_i, \quad (i = 1, \dots, m),$$

essendo l_i le componenti di \mathbf{l} nella stessa base. Da ciò segue anche che la continuità di \mathbf{f} in un punto $\mathbf{x}_0 \in A$, è equivalente alla continuità delle sue componenti $f_i : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ nello stesso punto \mathbf{x}_0 .

Estendiamo ora il concetto di differenziabilità alle funzioni vettoriali di più variabili.

Si dice che $\mathbf{f} : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ è differenziabile in $\mathbf{x}_0 \in A$ se esiste una applicazione lineare $L_{\mathbf{x}_0} : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tale che

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{\mathbf{f}(\mathbf{x}_0 + \mathbf{h}) - \mathbf{f}(\mathbf{x}_0) - L_{\mathbf{x}_0}(\mathbf{h})}{\|\mathbf{h}\|} = \mathbf{0},$$

dove $\mathbf{h} \in \mathbb{R}^n$ e dove il limite è ovviamente inteso nel senso della norma di \mathbb{R}^m . Come prima chiameremo $L_{\mathbf{x}_0}$ il differenziale di \mathbf{f} in \mathbf{x}_0 e lo denoteremo con $d\mathbf{f}(\mathbf{x}_0)$. In base alle premesse fatte, il limite precedente è equivalente agli m limiti

$$\lim_{\mathbf{h} \rightarrow \mathbf{0}} \frac{f_i(\mathbf{x}_0 + \mathbf{h}) - f_i(\mathbf{x}_0) - df_i(\mathbf{x}_0)(\mathbf{h})}{\|\mathbf{h}\|} = 0, \quad i = 1, \dots, m.$$

Quindi, una funzione $\mathbf{f} : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ è differenziabile se e solo se lo sono le sue funzioni componenti. Preso un generico versore $\mathbf{v} \in \mathbb{R}^n$ si ha

$$df_i(\mathbf{x}_0)(\mathbf{v}) = \nabla f_i(\mathbf{x}_0) \cdot \mathbf{v} = \sum_{k=1}^n \frac{\partial f_i}{\partial x_k}(\mathbf{x}_0) v_k.$$

La matrice $m \times n$ di elementi

$$J_{ik} = \frac{\partial f_i}{\partial x_k}$$

si dice matrice *jacobiana* della funzione \mathbf{f} . Le righe di questa matrice sono costituite dalle componenti dei gradienti delle f_i .

- Esempio

Calcolare la matrice jacobiana della funzione $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^3$, data da

$$f_1(\mathbf{x}) = 2x - 3y$$

$$f_2(\mathbf{x}) = \sin xy$$

$$f_3(\mathbf{x}) = 2 \cos x - 3 \sin y,$$

nel punto $\mathbf{x}_0 = (0, 0)$. Le componenti di \mathbf{f} sono continue con derivate continue. Si ha

$$\begin{aligned} \frac{\partial f_1}{\partial x} &= 2, & \frac{\partial f_2}{\partial x} &= y \cos xy, & \frac{\partial f_3}{\partial x} &= -2 \sin x \\ \frac{\partial f_1}{\partial y} &= -3, & \frac{\partial f_2}{\partial y} &= x \cos xy, & \frac{\partial f_3}{\partial y} &= -3 \cos y. \end{aligned}$$

Per $\mathbf{x} = \mathbf{x}_0$ si ottiene,

$$J = \begin{pmatrix} 2 & -3 \\ 0 & 0 \\ 0 & -3 \end{pmatrix}.$$

Consideriamo ora due funzioni $\mathbf{f} : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$, e $\mathbf{g} : B \subset \mathbb{R}^m \rightarrow \mathbb{R}^p$, con $\mathbf{f}(A) \subset B$ e sia $\mathbf{F} = \mathbf{g} \circ \mathbf{f}$ la funzione composta $\mathbf{F} : A \subset \mathbb{R}^n \rightarrow \mathbb{R}^p$, definita da

$$\mathbf{F}(\mathbf{x}) = \mathbf{g}[\mathbf{f}(\mathbf{x})], \quad \mathbf{x} \in A.$$

Enunciamo, senza dimostrare, i seguenti risultati.

Teorema 2.5 *Sia \mathbf{f} continua in $\mathbf{x}_0 \in A$ e \mathbf{g} continua in $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$. Allora la funzione composta $\mathbf{F} = \mathbf{g} \circ \mathbf{f}$ è continua in \mathbf{x}_0 .*

Teorema 2.6 *Sia \mathbf{f} differenziabile in $\mathbf{x}_0 \in A$ e \mathbf{g} differenziabile in $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$. Allora la funzione composta $\mathbf{F} = \mathbf{g} \circ \mathbf{f}$ è differenziabile in \mathbf{x}_0 e si ha*

$$d\mathbf{F}(\mathbf{x}_0) = d\mathbf{g}(\mathbf{y}_0) \circ d\mathbf{f}(\mathbf{x}_0).$$

Presi $\mathbf{v} \in \mathbb{R}^n$ e $\mathbf{w} \in \mathbb{R}^m$, si ha, per $i = 1, \dots, m$ e $j = 1, \dots, p$,

$$df_i(\mathbf{x}_0)(\mathbf{v}) = \sum_{k=1}^n \frac{\partial f_i}{\partial x_k}(\mathbf{x}_0)v_k, \quad dg_j(\mathbf{y}_0)(\mathbf{w}) = \sum_{h=1}^m \frac{\partial g_j}{\partial y_h}(\mathbf{y}_0)w_h.$$

Dal teorema 2.6 si ottiene

$$dF_q(\mathbf{x}_0)(\mathbf{v}) = \sum_{h=1}^m \frac{\partial g_q}{\partial y_h}(\mathbf{y}_0) \sum_{k=1}^n \frac{\partial f_h}{\partial x_k}(\mathbf{x}_0)v_k = \sum_{k=1}^n \left(\sum_{h=1}^m \frac{\partial g_q}{\partial y_h}(\mathbf{y}_0) \frac{\partial f_h}{\partial x_k}(\mathbf{x}_0) \right) v_k.$$

Ne segue che la matrice jacobiana della funzione composta \mathbf{F} è data dal prodotto righe per colonne delle due matrici jacobiane di \mathbf{g} e di \mathbf{f} , ovvero

$$J_{\mathbf{F}}(\mathbf{x}_0) = J_{\mathbf{g}}(\mathbf{y}_0)J_{\mathbf{f}}(\mathbf{x}_0) \quad J_{qk}^{\mathbf{F}} = \sum_{h=1}^m J_{qh}^{\mathbf{g}}J_{hk}^{\mathbf{f}}.$$

Consideriamo il caso particolare di $\mathbf{f} : A \subset \mathbb{R} \rightarrow \mathbb{R}^3$ e $g : B \subset \mathbb{R}^3 \rightarrow \mathbb{R}$. Si ha

$$F(t) = g(f_1(t), f_2(t), f_3(t)), \quad t \in A.$$

Calcoliamo la derivata di F in $t_0 \in A$. Posto $g = g(x_1, x_2, x_3)$ si ha

$$F'(t_0)dt = dF(t_0) = \left(\sum_{k=1}^3 \frac{\partial g}{\partial x_k}(\mathbf{f}(t_0)) \frac{df_k}{dt}(t_0) \right) dt.$$

da cui,

$$F'(t_0) = \frac{\partial g}{\partial x_1}(\mathbf{f}(t_0))f'_1(t_0) + \frac{\partial g}{\partial x_2}(\mathbf{f}(t_0))f'_2(t_0) + \frac{\partial g}{\partial x_3}(\mathbf{f}(t_0))f'_3(t_0).$$

• Esempi

1) Calcolare la derivata prima della seguente funzione composta F mediante la formula precedente e in modo diretto, dopo aver ricavato esplicitamente $F(t)$,

$$F(t) = (g \circ \mathbf{f})(t), \quad \begin{cases} f_1(t) = e^t \\ f_2(t) = \sin t \\ f_3(t) = \arctan t, \end{cases} \quad g(x, y, z) = zy - y \sin x.$$

Si ha

$$\frac{\partial g}{\partial x} = -y \cos x, \quad \frac{\partial g}{\partial y} = z - \sin x, \quad \frac{\partial g}{\partial z} = y.$$

Calcolando queste derivate parziali per $x = e^t, y = \sin t, z = \arctan t$, e sostituendo nella formula

$$F'(t) = \frac{\partial g}{\partial x} f_1'(t) + \frac{\partial g}{\partial y} f_2'(t) + \frac{\partial g}{\partial z} f_3'(t),$$

si ottiene

$$F'(t) = \sin t \left(\frac{1}{1+t^2} - e^t \cos e^t \right) + \cos t (\arctan t - \sin e^t).$$

Calcolando direttamente $F(t)$, si ha

$$F(t) = g(f_1(t), f_2(t), f_3(t)) = \arctan t \sin t - \sin t \sin e^t,$$

da cui, derivando, si ottiene il risultato precedente.

2) Calcolare la matrice jacobiana della funzione composta $\mathbf{F} = \mathbf{g} \circ \mathbf{f}$, dove $\mathbf{f} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ e $\mathbf{g} : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ sono date da

$$\begin{cases} f_1(x, y) = x \\ f_2(x, y) = xy, \end{cases} \quad \begin{cases} g_1(x, y) = xe^y \\ g_2(x, y) = ye^x \end{cases}.$$

Si ha

$$\begin{aligned} \frac{\partial f_1}{\partial x} &= 1, & \frac{\partial f_2}{\partial x} &= y, & \frac{\partial f_1}{\partial y} &= 0, & \frac{\partial f_2}{\partial y} &= x, \\ \frac{\partial g_1}{\partial x} &= e^y, & \frac{\partial g_2}{\partial x} &= ye^x, & \frac{\partial g_1}{\partial y} &= xe^y, & \frac{\partial g_2}{\partial y} &= e^x. \end{aligned}$$

da cui

$$\frac{\partial g_1}{\partial x}(x, xy) = e^{xy}, \quad \frac{\partial g_2}{\partial x}(x, xy) = xye^x, \quad \frac{\partial g_1}{\partial y}(x, xy) = xe^{xy}, \quad \frac{\partial g_2}{\partial y}(x, xy) = e^x.$$

Ne segue

$$J_{\mathbf{F}} = J_{\mathbf{g}} J_{\mathbf{f}} = \begin{pmatrix} e^{xy} & xe^{xy} \\ xye^x & e^x \end{pmatrix} \begin{pmatrix} 1 & 0 \\ y & x \end{pmatrix} = \begin{pmatrix} (1+xy)e^{xy} & x^2e^{xy} \\ ye^x(1+x) & xe^x \end{pmatrix}.$$

3) Calcolare la matrice jacobiana delle funzioni composte $\mathbf{F}_1 = \mathbf{g} \circ \mathbf{f}$ e $F_2 = \mathbf{f} \circ \mathbf{g}$ dove $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ e $\mathbf{g} : \mathbb{R} \rightarrow \mathbb{R}^2$ sono date da

$$f(x, y) = x^2 + 2y, \quad \begin{cases} g_1(t) = te^t \\ g_2(t) = \sin 2t \end{cases}.$$

Le derivate sono

$$\frac{\partial f}{\partial x} = 2x, \quad \frac{\partial f}{\partial y} = 2, \quad g'_1(t) = (1+t)e^t, \quad g'_2(t) = 2 \cos 2t.$$

Nel primo caso si ha

$$g'_1(x^2 + 2y) = (1 + x^2 + 2y)e^{x^2 + 2y}, \quad g'_2(x^2 + 2y) = 2 \cos(2x^2 + 4y),$$

e di conseguenza

$$\begin{aligned} J_{F_1} &= J_g J_f = \begin{pmatrix} (1 + x^2 + 2y)e^{x^2 + 2y} \\ 2 \cos(2x^2 + 4y) \end{pmatrix} \begin{pmatrix} 2x & 2 \end{pmatrix} \\ &= \begin{pmatrix} 2x(1 + x^2 + 2y)e^{x^2 + 2y} & 2(1 + x^2 + 2y)e^{x^2 + 2y} \\ 4x \cos(2x^2 + 4y) & 4 \cos(2x^2 + 4y) \end{pmatrix}. \end{aligned}$$

Nel secondo caso si ha

$$\frac{\partial f}{\partial x}(te^t, \sin 2t) = 2te^t, \quad \frac{\partial f}{\partial y}(te^t, \sin 2t) = 2,$$

da cui

$$J_{F_2} = J_f J_g = \begin{pmatrix} 2te^t & 2 \end{pmatrix} \begin{pmatrix} (1+t)e^t \\ 2 \cos 2t \end{pmatrix} = 2te^{2t}(1+t) + 4 \cos 2t = F'_2(t).$$

2.5 Derivate successive e formula di Taylor

Le derivate parziali prime di una funzione $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ possono essere considerate funzioni nel dominio B di derivabilità. Se tali funzioni sono derivabili, si potranno calcolare le derivate parziali seconde, e così via si definiscono le derivate parziali terze, quarte, etc. In generale la derivata parziale seconda

$$\frac{\partial^2 f}{\partial x_i \partial x_j},$$

rispetto alle variabili x_i, x_j può risultare diversa dalla derivata parziale rispetto a x_j, x_i , cioè quella che si ottiene invertendo l'ordine di derivazione. Per esempio la funzione

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

ha derivate parziali seconde in $x = 0, y = 0$ date da

$$\left. \frac{\partial^2 f}{\partial x \partial y} \right|_{\mathbf{x}=\mathbf{0}} = -1, \quad \left. \frac{\partial^2 f}{\partial y \partial x} \right|_{\mathbf{x}=\mathbf{0}} = 1.$$

Vale il seguente risultato, che non dimostreremo.

Teorema 2.7 (di Schwarz). *Data una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ con A aperto, e dato un punto $(x_0, y_0) \in A$, se le derivate parziali seconde*

$$\frac{\partial^2 f}{\partial x \partial y}, \quad \frac{\partial^2 f}{\partial y \partial x},$$

sono definite in un intorno sferico $B_r[(x_0, y_0)]$ e continue in (x_0, y_0) allora esse coincidono in (x_0, y_0) .

Il teorema di Schwarz si può generalizzare a funzioni di n variabili di classe C^k con k arbitrario. Nell'ambito di validità di questo teorema possiamo calcolare le derivate parziali di ordine superiore al secondo, senza preoccuparci dell'ordine in cui vengono effettuate le derivate stesse.

Consideriamo una funzione $f \in C^k(A)$, essendo A un aperto di \mathbb{R}^n . Preso $\mathbf{x}_0 \in A$, sia $\mathbf{w} \in \mathbb{R}^n$ tale che $B_{\|\mathbf{w}\|}(\mathbf{x}_0) \subset A$. Consideriamo la restrizione F della funzione f al "segmento" $\{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} = \mathbf{x}_0 + t\mathbf{w}, t \in [0, 1]\}$, ovvero

$$F(t) = f(\mathbf{x}_0 + t\mathbf{w}).$$

Questa può essere considerata una funzione composta di t tramite la f e la $\mathbf{x}(t) = \mathbf{x}_0 + t\mathbf{w}$. Essendo $f \in C^k(A)$, sarà $F \in C^k([0, 1])$. Sviluppiamo allora la funzione $F(t)$ nel suo polinomio di Taylor attorno a $t = 0$. Si ha

$$F(t) = F(0) + F'(0)t + \frac{F''(0)}{2}t^2 + \dots + \frac{F^{(k)}(0)}{k!}t^k + R_k(t, 0).$$

In particolare, per $t = 1$, abbiamo

$$F(1) = F(0) + F'(0) + \frac{F''(0)}{2} + \dots + \frac{F^{(k)}(0)}{k!} + R_k(1, 0).$$

Osserviamo ora che

$$\begin{aligned} F(0) &= f(\mathbf{x}_0), \\ F'(0) &= \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0)w_i, \\ F''(0) &= \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)w_i w_j, \\ &\vdots \\ F^{(k)}(0) &= \sum_{i_1, i_2, \dots, i_k=1}^n \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}}(\mathbf{x}_0)w_{i_1} w_{i_2} \dots w_{i_k}. \end{aligned}$$

Posto $\mathbf{x} = \mathbf{x}_0 + \mathbf{w}$, avremo $F(1) = f(\mathbf{x})$ e $\mathbf{w} = \mathbf{x} - \mathbf{x}_0$. Sostituendo nello sviluppo di F otteniamo

$$\begin{aligned} f(\mathbf{x}) &= f(\mathbf{x}_0) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0)(x_i - x_i^0) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)(x_i - x_i^0)(x_j - x_j^0) \\ &+ \dots + \frac{1}{k!} \sum_{i_1, i_2, \dots, i_k=1}^n \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}}(\mathbf{x}_0)(x_{i_1} - x_{i_1}^0)(x_{i_2} - x_{i_2}^0) \dots (x_{i_k} - x_{i_k}^0) \\ &+ R_k(\mathbf{x}, \mathbf{x}_0). \end{aligned}$$

Questa espressione rappresenta la formula di Taylor per funzioni di più variabili. Il resto $R_k(\mathbf{x}, \mathbf{x}_0)$ della formula di Taylor si può scrivere nella forma di Lagrange, purchè $f \in C^{k+1}$. Poiché si ha

$$R_k(1, 0) = \frac{F^{(k+1)}(\tau)}{(k+1)!}, \quad \text{con } \tau \in [0, 1],$$

si ottiene

$$R_k(\mathbf{x}, \mathbf{x}_0) = \frac{1}{(k+1)!} \times \sum_{i_1, i_2, \dots, i_{k+1}=1}^n \frac{\partial^{(k+1)} f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_{k+1}}}(\mathbf{x}_0 + \tau \mathbf{w})(x_{i_1} - x_{i_1}^0)(x_{i_2} - x_{i_2}^0) \dots (x_{i_{k+1}} - x_{i_{k+1}}^0).$$

Osserviamo che, per ogni $i = 1, \dots, n$, si ha

$$|x_i - x_i^0| \leq \|\mathbf{x} - \mathbf{x}_0\|.$$

Ciò comporta che

$$|R_k(\mathbf{x}, \mathbf{x}_0)| \leq M_k \|\mathbf{x} - \mathbf{x}_0\|^{k+1},$$

dove

$$M_k = \frac{1}{(k+1)!} \max_{\tau \in [0,1]} \left\{ \sum_{i_1, i_2, \dots, i_{k+1}=1}^n \left| \frac{\partial^{(k+1)} f}{\partial x_{i_1} \dots \partial x_{i_{k+1}}}(\mathbf{x}_0 + \mathbf{w}\tau) \right| \right\}$$

Ne segue che $R_k(\mathbf{x}, \mathbf{x}_0)$ è un infinitesimo di ordine superiore rispetto a $\|\mathbf{x} - \mathbf{x}_0\|^k$, ovvero

$$\lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{R_k(\mathbf{x}, \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|^k} = 0.$$

• Esempio

Scrivere la formula di Taylor per la funzione $f(x, y) = ye^{xy}$ attorno al punto $(0, 0)$ fino al terzo ordine. Esplicitamente si ha

$$f(x, y) = f(0, 0) + \frac{\partial f}{\partial x}(0, 0)x + \frac{\partial f}{\partial y}(0, 0)y + \frac{1}{2} \left[\frac{\partial^2 f}{\partial x^2}(0, 0)x^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(0, 0)xy + \frac{\partial^2 f}{\partial y^2}(0, 0)y^2 \right] + \frac{1}{6} \left[\frac{\partial^3 f}{\partial x^3}(0, 0)x^3 + 3 \frac{\partial^3 f}{\partial x^2 \partial y}(0, 0)x^2y + 3 \frac{\partial^3 f}{\partial x \partial y^2}(0, 0)xy^2 + \frac{\partial^3 f}{\partial y^3}(0, 0)y^3 \right] + R_3(\mathbf{x}, \mathbf{0}).$$

Le derivate parziali che compaiono nella formula sono date da

$$\begin{aligned} \frac{\partial f}{\partial x} &= y^2 e^{xy}, & \frac{\partial f}{\partial y} &= (1 + xy)e^{xy}, \\ \frac{\partial^2 f}{\partial x^2} &= y^3 e^{xy}, & \frac{\partial^2 f}{\partial x \partial y} &= (2y + y^2x)e^{xy}, & \frac{\partial^2 f}{\partial y^2} &= (2x + x^2y)e^{xy}, \\ \frac{\partial^3 f}{\partial x^3} &= y^4 e^{xy}, & \frac{\partial^3 f}{\partial x^2 \partial y} &= (3y^2 + y^3x)e^{xy}, & \frac{\partial^3 f}{\partial x \partial y^2} &= (2 + 4xy + x^2y^2)e^{xy}, \\ \frac{\partial^3 f}{\partial y^3} &= (3x^2 + x^3y)e^{xy} \end{aligned}$$

Dopo aver valutato queste derivate in $(0, 0)$, otteniamo

$$f(x, y) = y + xy^2 + R_3(\mathbf{x}, \mathbf{0}).$$

Nel seguito sarà utile avere una espressione di tipo vettoriale per la formula di Taylor con $k = 2$. Sia $\mathbf{x}_0 \in A$ e sia $f \in C^2$. Per $k = 2$ possiamo scrivere

$$f(\mathbf{x}) = f(\mathbf{x}_0) + \sum_{i=1}^n \frac{\partial f}{\partial x_i}(\mathbf{x}_0)(x_i - x_i^0) + \frac{1}{2} \sum_{i,j=1}^n \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0)(x_i - x_i^0)(x_j - x_j^0) + R_2(\mathbf{x}, \mathbf{x}_0).$$

Introduciamo la matrice $H(\mathbf{x}_0)$ i cui elementi sono dati da

$$h_{ij}(\mathbf{x}_0) = \frac{\partial^2 f}{\partial x_i \partial x_j}(\mathbf{x}_0).$$

Si tratta di una matrice $n \times n$, detta *hessiana* della funzione f in \mathbf{x}_0 . Per il teorema di Schwarz essa è simmetrica. Facendo uso di questa matrice la formula di Taylor si può anche scrivere

$$f(\mathbf{x}) = f(\mathbf{x}_0) + df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot \mathbf{H}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + R_2(\mathbf{x}, \mathbf{x}_0).$$

dove $\mathbf{H}(\mathbf{x}_0)$ è un operatore lineare simmetrico in \mathbb{R}^n a cui è associata la matrice hessiana. Ricordiamo infine che R_2 è un infinitesimo di ordine superiore rispetto a $\|\mathbf{x} - \mathbf{x}_0\|^2$.

2.6 Massimi e minimi relativi

Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ e sia $\mathbf{x}_0 \in A$. Si dice che \mathbf{x}_0 è un punto di *massimo relativo* (*minimo relativo*) per f se esiste un intorno sferico $B_\delta(\mathbf{x}_0)$ tale che $\forall \mathbf{x} \in B_\delta(\mathbf{x}_0) \cap A$ si abbia

$$f(\mathbf{x}) \leq f(\mathbf{x}_0) \quad (f(\mathbf{x}) \geq f(\mathbf{x}_0)).$$

Teorema 2.8 *Data $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ differenziabile in un punto \mathbf{x}_0 interno ad A , se \mathbf{x}_0 è un punto di massimo o di minimo relativo, si ha*

$$df(\mathbf{x}_0) = 0.$$

Dim. Dobbiamo dimostrare che sotto le ipotesi date, si ha $df(\mathbf{x}_0)(\mathbf{v}) = 0$, $\forall \mathbf{v} \in \mathbb{R}^n$. Supponiamo, per fissare le idee, che \mathbf{x}_0 sia un punto di massimo relativo per f . Allora $f(\mathbf{x}) \leq f(\mathbf{x}_0)$ per ogni $\mathbf{x} \in B_\delta(\mathbf{x}_0) \cap A$. Preso un generico versore $\mathbf{v} \in \mathbb{R}^n$, consideriamo la funzione

$$F(t) = f(\mathbf{x}_0 + t\mathbf{v}),$$

dove $t \in \mathbb{R}$ è tale che $(\mathbf{x}_0 + t\mathbf{v}) \in B_\delta(\mathbf{x}_0) \cap A$. Poiché questa funzione è una restrizione di f definita in un intorno contenente \mathbf{x}_0 , si avrà $f(\mathbf{x}_0 + t\mathbf{v}) \leq f(\mathbf{x}_0)$, quindi $F(t)$ è una funzione di variabile reale che ammette un massimo relativo in $t = 0$. Ne segue che $F'(0) = 0$. Essendo

$$F'(0) = \sum_k \frac{\partial f}{\partial x_k}(\mathbf{x}_0)v_k = df(\mathbf{x}_0)(\mathbf{v}),$$

si avrà $df(\mathbf{x}_0)(\mathbf{v}) = 0$ per ogni scelta di \mathbf{v} in \mathbb{R}^n . La dimostrazione è identica nel caso si supponga che \mathbf{x}_0 sia un punto di minimo relativo.

Coerentemente con la terminologia usata nel caso delle funzioni di una variabile reale, chiameremo punti di *stazionarietà* per f , tutti i punti per i quali $df = 0$. Osserviamo che l'annullarsi del differenziale comporta l'annullarsi del gradiente di f in \mathbf{x}_0 e quindi di tutte le derivate parziali prime di f in \mathbf{x}_0 . Infatti, essendo

$$df(\mathbf{x}_0)(\mathbf{v}) = \nabla f(\mathbf{x}_0) \cdot \mathbf{v},$$

si ha

$$df(\mathbf{x}_0)(\mathbf{v}) = 0, \quad \forall \mathbf{v} \in \mathbb{R}^n \quad \Leftrightarrow \quad \nabla f(\mathbf{x}_0) = \mathbf{0}, \quad \Leftrightarrow \quad \frac{\partial f}{\partial x_i}(\mathbf{x}_0) = 0, \quad (i = 1, \dots, n).$$

Teorema 2.9 *Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ una funzione di classe C^2 e sia \mathbf{x}_0 un punto di stazionarietà per f , interno ad A . Condizione necessaria affinché \mathbf{x}_0 sia un punto di minimo relativo (massimo relativo) è che la matrice hessiana in \mathbf{x}_0 sia semidefinita positiva (semidefinita negativa).*

Dim. Essendo $f \in C^2$, scriviamo la formula di Taylor per $k = 2$.

$$f(\mathbf{x}) = f(\mathbf{x}_0) + df(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot \mathbf{H}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + R_2(\mathbf{x}, \mathbf{x}_0).$$

Poiché \mathbf{x}_0 è un punto di stazionarietà, si ha $df(\mathbf{x}_0) = 0$, e di conseguenza, posto $\mathbf{x} = \mathbf{x}_0 + t\mathbf{v}$, dove \mathbf{v} è un generico versore in \mathbb{R}^n , avremo

$$f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) = \frac{1}{2}t^2\mathbf{v} \cdot \mathbf{H}\mathbf{v} + R_2(\mathbf{x}_0 + t\mathbf{v}, \mathbf{x}_0),$$

dove abbiamo fatto uso della linearità dell'operatore \mathbf{H} . Supponiamo ora che \mathbf{x}_0 sia un punto di minimo relativo. Esisterà un intorno di $t = 0$ in cui $f(\mathbf{x}_0 + t\mathbf{v}) - f(\mathbf{x}_0) \geq 0$ e quindi, in quell'intorno

$$\mathbf{v} \cdot \mathbf{H}\mathbf{v} + 2t^{-2}R_2(\mathbf{x}_0 + t\mathbf{v}, \mathbf{x}_0) \geq 0.$$

D'altra parte, si ha

$$0 = \lim_{\mathbf{x} \rightarrow \mathbf{x}_0} \frac{R_2(\mathbf{x}, \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|^2} = \lim_{t \rightarrow 0} \frac{R_2(\mathbf{x}_0 + t\mathbf{v}, \mathbf{x}_0)}{t^2},$$

quindi, essendo $\mathbf{v} \cdot \mathbf{H}\mathbf{v}$ indipendente da t , si ha

$$0 \leq \lim_{t \rightarrow 0} [\mathbf{v} \cdot \mathbf{H}\mathbf{v} + 2t^{-2}R_2(\mathbf{x}_0 + t\mathbf{v}, \mathbf{x}_0)] = \mathbf{v} \cdot \mathbf{H}\mathbf{v}.$$

Questo risultato vale per ogni scelta di \mathbf{v} quindi \mathbf{H} è semidefinito positivo, cioè la matrice hessiana in \mathbf{x}_0 è semidefinita positiva. In modo perfettamente analogo si dimostra che se il punto \mathbf{x}_0 è di massimo relativo, allora la matrice hessiana in \mathbf{x}_0 deve essere semidefinita negativa.

Il teorema precedente fornisce solo una condizione necessaria affinché un punto di stazionarietà sia un massimo (o un minimo) relativo. La condizione non è sufficiente. A questo proposito consideriamo il seguente esempio.

Data la funzione

$$f(x, y) = x^2 - y^4,$$

cerchiamo i suoi punti di stazionarietà imponendo l'annullarsi del differenziale, ovvero l'annullarsi delle sue derivate parziali prime.

$$\frac{\partial f}{\partial x} = 2x = 0, \quad \frac{\partial f}{\partial y} = -4y^3 = 0,$$

da cui ricaviamo $\mathbf{x}_0 = (0, 0)$. La matrice hessiana in tale punto è

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

e risulta semidefinita positiva, in quanto $\mathbf{v} \cdot \mathbf{H}(\mathbf{0})\mathbf{v} = 2v_1^2 \geq 0$. La funzione, tuttavia, non possiede un minimo in $(0, 0)$ perché risulta positiva lungo l'asse delle x e negativa lungo l'asse delle y , in un qualunque intorno di $(0, 0)$.

Teorema 2.10 Sia $f : A \subset \mathbb{R}^n \rightarrow \mathbb{R}$ di classe C^2 e sia \mathbf{x}_0 un punto di stazionarietà per f , interno ad A . Condizione sufficiente affinché \mathbf{x}_0 sia un punto di minimo relativo (massimo relativo) è che la matrice hessiana in \mathbf{x}_0 sia definita positiva (definita negativa).

Dim. Nell'ipotesi $f \in C^2$, ed essendo \mathbf{x}_0 di stazionarietà, possiamo scrivere

$$f(\mathbf{x}) - f(\mathbf{x}_0) = \frac{1}{2}(\mathbf{x} - \mathbf{x}_0) \cdot \mathbf{H}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + R_2(\mathbf{x}, \mathbf{x}_0),$$

in un intorno di \mathbf{x}_0 contenuto in A . Poiché $H(\mathbf{x}_0)$ è una matrice simmetrica, essa ammetterà n autovettori ortonormali $\mathbf{e}_1, \mathbf{e}_2, \dots, \mathbf{e}_n$. Questi autovettori formeranno una base in \mathbb{R}^n e potremo scrivere

$$\mathbf{x} - \mathbf{x}_0 = \sum_{i=1}^n c_i \mathbf{e}_i.$$

si ottiene così,

$$\mathbf{H}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = \mathbf{H}(\mathbf{x}_0) \sum_{i=1}^n c_i \mathbf{e}_i = \sum_{i=1}^n c_i \lambda^{(i)} \mathbf{e}_i,$$

essendo $\lambda^{(i)}$ gli autovalori corrispondenti. Moltiplicando scalarmente a sinistra per $(\mathbf{x} - \mathbf{x}_0)$, otteniamo

$$(\mathbf{x} - \mathbf{x}_0) \cdot \mathbf{H}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = \sum_{i,j=1}^n c_i c_j \lambda^{(i)} \mathbf{e}_j \cdot \mathbf{e}_i = \sum_{i=1}^n \lambda^{(i)} c_i^2.$$

Supponiamo ora che $\mathbf{H}(\mathbf{x}_0)$ sia definito positivo. Allora tutti i suoi autovalori sono positivi. Quindi, preso il più piccolo tra questi, chiamiamolo $\bar{\lambda}$, avremo

$$(\mathbf{x} - \mathbf{x}_0) \cdot \mathbf{H}(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) = \sum_{i=1}^n \lambda^{(i)} c_i^2 \geq \bar{\lambda} \sum_{i=1}^n c_i^2 = \bar{\lambda} \|\mathbf{x} - \mathbf{x}_0\|^2.$$

Si ottiene così

$$f(\mathbf{x}) - f(\mathbf{x}_0) \geq \frac{1}{2} \bar{\lambda} \|\mathbf{x} - \mathbf{x}_0\|^2 + R_2(\mathbf{x}, \mathbf{x}_0).$$

Ricordando che R_2 è un infinitesimo di ordine superiore rispetto a $\|\mathbf{x} - \mathbf{x}_0\|^2$, fissato un $\epsilon \leq \frac{1}{2} \bar{\lambda}$, esisterà un intorno $B_\delta(\mathbf{x}_0)$ in cui

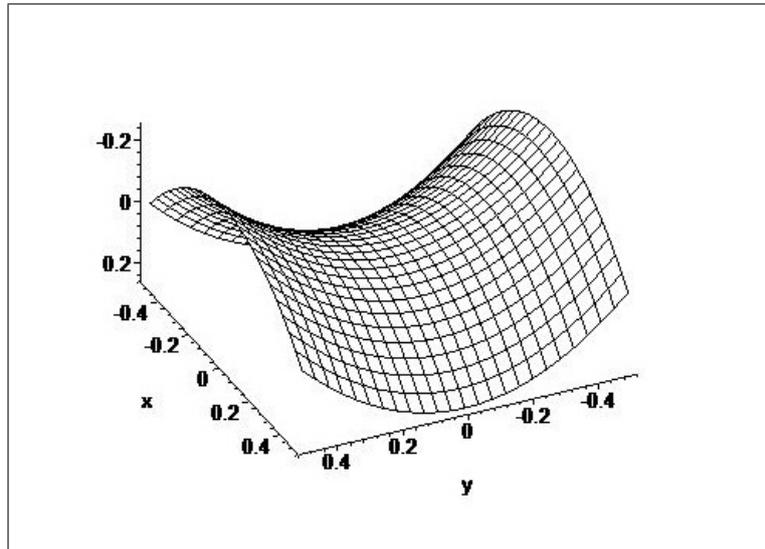
$$-\frac{1}{2} \bar{\lambda} \leq -\epsilon < \frac{R_2(\mathbf{x}, \mathbf{x}_0)}{\|\mathbf{x} - \mathbf{x}_0\|^2} < \epsilon \leq \frac{1}{2} \bar{\lambda},$$

e quindi $f(\mathbf{x}) - f(\mathbf{x}_0) \geq 0$, $\forall \mathbf{x} \in B_\delta(\mathbf{x}_0) \cap A$. In modo perfettamente analogo si procede per dimostrare che se nel punto di stazionarietà \mathbf{x}_0 di f la matrice hessiana è definita negativa, allora \mathbf{x}_0 è un punto di massimo relativo per f .

I risultati precedenti forniscono un metodo di ricerca dei massimi e dei minimi relativi di una funzione di più variabili. In pratica, bisogna cercare i punti di stazionarietà della funzione e, in tali punti, calcolare la matrice hessiana. Se questa matrice risulta definita abbiamo sicuramente un punto di massimo o di minimo relativo. Se la matrice hessiana non è nè definita nè semidefinita, cioè, come si usa dire, è *non definita*, il punto non è nè di massimo nè di minimo. Nel caso di funzioni di due variabili, un punto in cui la matrice hessiana non è definita si dice punto di *sella*. Si consideri ad esempio la funzione $f(x, y) = x^2 - y^2$. Essa ha un unico punto di stazionarietà in $(0, 0)$ e la sua matrice hessiana in questo punto è

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

che non è definita. Il punto $(0,0)$ è allora un punto di sella. Esso è caratterizzato dal fatto che la restrizione della F lungo la retta $y = 0$ ha un minimo in $(0,0)$ mentre la restrizione della F lungo la retta $x = 0$ ha un massimo in $(0,0)$ (vedi il grafico della funzione, in figura).



Per stabilire se la matrice hessiana è definita, basta calcolarne gli autovalori. Se questi sono tutti positivi o tutti negativi, la matrice sarà definita rispettivamente, positiva o negativa. Se gli autovalori sono tutti non negativi o tutti non positivi la matrice sarà semidefinita rispettivamente, positiva o negativa. In tutti gli altri casi l'hessiana sarà non definita.

- Esempi

1) Determinare gli eventuali punti di massimo e di minimo relativo della funzione

$$f(x, y) = x^2 + y^3 - xy.$$

Determiniamo i punti di stazionarietà, imponendo l'annullarsi delle derivate parziali prime. Si ha

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2x - y = 0, \\ \frac{\partial f}{\partial y} &= 3y^2 - x = 0, \end{aligned}$$

da cui si ottengono i due punti di stazionarietà $\mathbf{x}_1 = (0,0)$, $\mathbf{x}_2 = (1/12, 1/6)$. La matrice hessiana è

$$H(\mathbf{x}) = \begin{pmatrix} 2 & -1 \\ -1 & 6y \end{pmatrix}.$$

si ha, allora

$$H(\mathbf{x}_1) = \begin{pmatrix} 2 & -1 \\ -1 & 0 \end{pmatrix}, \quad H(\mathbf{x}_2) = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

Determiniamo gli autovalori di $H(\mathbf{x}_1)$. Dall'equazione caratteristica

$$\det(H(\mathbf{x}_1) - \lambda I) = \lambda^2 - 2\lambda - 1 = 0$$

ricaviamo $\lambda_1 = 1 + \sqrt{2}$, $\lambda_2 = 1 - \sqrt{2}$. Poiché $\lambda_1 > 0$, $\lambda_2 < 0$, ne segue che la matrice $H(\mathbf{x}_1)$ è non definita, quindi \mathbf{x}_1 è un punto di sella. Determiniamo gli autovalori di $H(\mathbf{x}_2)$. Dall'equazione caratteristica

$$\det(H(\mathbf{x}_2) - \lambda I) = \lambda^2 - 3\lambda + 1 = 0,$$

otteniamo,

$$\lambda_1 = \frac{3 + \sqrt{5}}{2}, \quad \lambda_2 = \frac{3 - \sqrt{5}}{2}.$$

Ambedue gli autovalori sono positivi quindi la matrice hessiana risulta definita positiva. Il punto \mathbf{x}_2 è allora un punto di minimo relativo per f .

2) Determinare gli eventuali punti di massimo e di minimo relativo della funzione

$$f(x, y, z) = x^2 + 2y^2 + z^2 + xy - xz.$$

Impongo le condizioni di stazionarietà

$$\begin{aligned} \frac{\partial f}{\partial x} &= 2x + y - z = 0, \\ \frac{\partial f}{\partial y} &= 4y + x = 0, \\ \frac{\partial f}{\partial z} &= 2z - x = 0, \end{aligned}$$

che ammettono la sola soluzione $\mathbf{x} = (0, 0, 0)$. Si ha

$$H(\mathbf{0}) = \begin{pmatrix} 2 & 1 & -1 \\ 1 & 4 & 0 \\ -1 & 0 & 2 \end{pmatrix}$$

da cui si ricava

$$\det(H(\mathbf{0}) - \lambda I) = 0 \quad \Rightarrow \quad -\lambda^3 + 8\lambda^2 - 18\lambda + 10 = 0.$$

Questa equazione non ammette radici non positive quindi le tre radici saranno necessariamente positive e la funzione avrà in $(0, 0, 0)$ un punto di minimo relativo.

3) Determinare gli eventuali punti di massimo o di minimo della seguente funzione

$$f(x, y) = 3x^2 - y^2 + 2xy - x.$$

La condizione di stazionarietà implica

$$\begin{cases} 6x + 2y - 1 = 0 \\ -2y + 2x = 0 \end{cases}$$

da cui si ricava l'unico punto di stazionarietà

$$\mathbf{x}_0 = \left(\frac{1}{8}, \frac{1}{8} \right).$$

La matrice hessiana ha elementi costanti e vale

$$H = \begin{pmatrix} 6 & 2 \\ 2 & -2 \end{pmatrix}.$$

I suoi autovalori sono

$$\lambda_1 = 2 - 2\sqrt{5}, \quad \lambda_2 = 2 + 2\sqrt{5}.$$

L'hessiana è dunque non definita ed il punto trovato è un punto di sella.

Capitolo 3

Spazi metrici

3.1 Metrica

Sia W un insieme e u, v due suoi elementi. Si definisce *distanza* o *metrica* su W una funzione $d(u, v)$ che ad ogni coppia di elementi di W fa corrispondere un numero reale tale che

- a) $d(u, v) \geq 0$, con $d(u, v) = 0$ se e solo se $u = v$,
- b) $d(u, v) = d(v, u)$,
- c) $d(u, v) \leq d(u, w) + d(w, v)$, $\forall u, v, w \in W$.

Un insieme dotato di una distanza così definita si dice *spazio metrico*.

- Esempi

1) Dato lo spazio vettoriale \mathbb{R} , definiamo la distanza tra due numeri reali x, y come

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

È immediato verificare che questa distanza è una metrica in \mathbb{R} . Lo stesso vale per un qualunque intervallo $(a, b) \subset \mathbb{R}$.

2) Dato lo spazio vettoriale euclideo E_n , definiamo

$$d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|, \quad \forall \mathbf{u}, \mathbf{v} \in E_n.$$

La distanza così introdotta in E_n è una metrica infatti si ha

$$\begin{aligned} \|\mathbf{u} - \mathbf{v}\| &\geq 0, & \text{con } \|\mathbf{u} - \mathbf{v}\| = 0, & \text{ se e solo se } \mathbf{u} - \mathbf{v} = \mathbf{0}. \\ \|\mathbf{u} - \mathbf{v}\| &= \sqrt{(\mathbf{u} - \mathbf{v}) \cdot (\mathbf{u} - \mathbf{v})} = \sqrt{(\mathbf{v} - \mathbf{u}) \cdot (\mathbf{v} - \mathbf{u})} = \|\mathbf{v} - \mathbf{u}\|, \\ \|\mathbf{u} - \mathbf{v}\| &= \|\mathbf{u} - \mathbf{w} + \mathbf{w} - \mathbf{v}\| \leq \|\mathbf{u} - \mathbf{w}\| + \|\mathbf{w} - \mathbf{v}\| \quad (\text{per la diseg. triang.}), \end{aligned}$$

e lo spazio E_n assume la struttura di spazio metrico.

3) Sia $C([a, b])$ lo spazio delle funzioni reali di variabile reale continue in $[a, b]$. Date due funzioni $f, g \in C([a, b])$, definiamo

$$d(f, g) = \sup_{x \in [a, b]} |f(x) - g(x)|.$$

Questa definizione ha senso in quanto le funzioni di $C([a, b])$ sono limitate in $[a, b]$. Sono inoltre soddisfatti i requisiti della metrica, in quanto,

$$\begin{aligned} \sup_{x \in [a, b]} |f(x) - g(x)| &\geq 0, \quad (= 0 \text{ se e solo se } f(x) = g(x) \text{ identicamente}), \\ \sup_{x \in [a, b]} |f(x) - g(x)| &= \sup_{x \in [a, b]} |g(x) - f(x)|, \\ \sup_{x \in [a, b]} |f(x) - g(x)| &= \sup_{x \in [a, b]} |f(x) - h(x) + h(x) - g(x)| \\ &\leq \sup_{x \in [a, b]} |f(x) - h(x)| + \sup_{x \in [a, b]} |h(x) - g(x)|, \end{aligned}$$

quindi lo spazio $C([a, b])$ assume, con la definizione data, la struttura di spazio metrico. Questa metrica si dice anche metrica della *convergenza uniforme*. Tale denominazione sarà chiarita in seguito.

4) Sia \mathbb{C}^n l'insieme delle n -uple ordinate di numeri complessi $\mathbf{z} = (z_1, z_2, \dots, z_n)$. Poniamo

$$d(\mathbf{z}, \mathbf{z}') = \sqrt{\sum_{i=1}^n |z_i - z'_i|^2}.$$

Con questa distanza \mathbb{C}^n assume la struttura di spazio metrico. La verifica delle proprietà a)-c) segue dalle proprietà dei moduli dei numeri complessi ed è immediata.

Nel seguito chiameremo semplicemente *punti* gli elementi di uno spazio metrico.

Consideriamo uno spazio metrico W ed una successione di punti $\{y_n\}_{n \in \mathbb{N}}$ di W . Diremo che $\{y_n\}$ converge al punto $y \in W$ e scriveremo $\lim_{n \rightarrow \infty} y_n = y$, se, fissato un $\varepsilon > 0$, esiste un indice $\nu \in \mathbb{N}$ tale che per $n > \nu$ si abbia

$$d(y_n, y) < \varepsilon,$$

essendo d la metrica di W . In altri termini, la convergenza di una successione di punti in W è equivalente alla convergenza della successione numerica

$$\eta_n = d(y_n, y),$$

a zero.

Si dice che una successione $\{y_n\}$ di punti di uno spazio metrico è *di Cauchy* (o *fondamentale*) se, fissato un $\varepsilon > 0$, esiste un indice $\nu \in \mathbb{N}$ tale che, per ogni coppia di indici $n, m > \nu$ si abbia

$$d(y_n, y_m) < \varepsilon.$$

Osserviamo subito che ogni successione convergente in uno spazio metrico W è di Cauchy. Infatti, dalle proprietà della metrica e dall'ipotesi di convergenza, abbiamo

$$d(y_n, y_m) \leq d(y_n, y) + d(y, y_m) = d(y_n, y) + d(y_m, y) < 2\varepsilon,$$

per ogni coppia di indici $n, m > \nu$. Viceversa è possibile avere successioni fondamentali che non sono convergenti. Per esempio, l'insieme dei numeri razionali \mathbb{Q} con la metrica

$$d(x, y) = |x - y|, \quad \forall x, y \in \mathbb{Q}$$

è uno spazio metrico. Consideriamo la successione in \mathbb{Q}

$$y_n = \left(1 + \frac{1}{n}\right)^n.$$

Essa risulta di Cauchy, in quanto, posto $m = n + p$, si ha

$$\begin{aligned} \left(1 + \frac{1}{m}\right)^m - \left(1 + \frac{1}{n}\right)^n &= \left(1 + \frac{1}{n+p}\right)^{n+p} - \left(1 + \frac{1}{n}\right)^n < \\ \left(1 + \frac{1}{n}\right)^{n+p} - \left(1 + \frac{1}{n}\right)^n &= \left(1 + \frac{1}{n}\right)^n \left[\left(1 + \frac{1}{n}\right)^p - 1\right] < 3 \left[\left(1 + \frac{1}{n}\right)^p - 1\right]. \end{aligned}$$

Fissato un $\varepsilon > 0$, e comunque preso p , si può certamente trovare un ν tale che per ogni $n > \nu$ l'ultimo termine della disuguaglianza precedente venga maggiorato da ε . Ne segue che

$$\left|\left(1 + \frac{1}{m}\right)^m - \left(1 + \frac{1}{n}\right)^n\right| < \varepsilon, \quad \forall n, m > \nu,$$

e quindi la successione data è fondamentale. D'altra parte sappiamo che questa successione numerica converge al numero e che non appartiene a \mathbb{Q} . La successione non è dunque convergente in \mathbb{Q} .

3.2 Convergenza uniforme

Uno spazio metrico W si dice *completo* se ogni successione di Cauchy di punti di W è convergente.

Introdotta la metrica $d(x, y) = |x - y|$ in \mathbb{R} consideriamo un intervallo chiuso $[a, b]$. Dimostriamo che $[a, b]$ è uno spazio metrico completo. A tale scopo consideriamo una successione $\{x_n\}$ di punti di $[a, b]$ che sia di Cauchy. Tale successione sarà ovviamente limitata ed ammetterà quindi una sottosuccessione $\{x_{n_k}\}_{k \in \mathbb{N}}$ convergente a un punto \bar{x} di $[a, b]$, ovvero esisterà un $k_1 \in \mathbb{N}$ tale che $|x_{n_k} - \bar{x}| < \frac{\varepsilon}{2}$ per $k > k_1$. D'altra parte, essendo $\{x_n\}$ di Cauchy, avremo che per $n, m > \nu$, $|x_n - x_m| < \frac{\varepsilon}{2}$. Detto k_2 un indice tale che $\nu < n_{k_2}$, prendiamo $\bar{k} \geq \max\{k_1, k_2\}$ e avremo

$$|x_n - \bar{x}| = |x_n - x_{n_{\bar{k}}} + x_{n_{\bar{k}}} - \bar{x}| \leq |x_n - x_{n_{\bar{k}}}| + |x_{n_{\bar{k}}} - \bar{x}| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon$$

per $n > n_{\bar{k}}$. Otteniamo così la convergenza di $\{x_n\}$ a \bar{x} .

Teorema 3.1 *L'insieme delle funzioni continue $f : [a, b] \rightarrow [c, d]$, con la metrica $d(f, g) = \max_{x \in [a, b]} |f(x) - g(x)|$, è uno spazio metrico completo.*

Dim. Introduciamo la metrica $d(x_1, x_2) = |x_1 - x_2|$ negli intervalli $[a, b]$ e $[c, d]$. Consideriamo una successione di Cauchy $\{f_n\}$ di funzioni continue $f_n : [a, b] \rightarrow [c, d]$ e dimostriamo che f_n converge ad una funzione continua $f : [a, b] \rightarrow [c, d]$. Fissato $\bar{x} \in [a, b]$, avremo

$$d(f_n(\bar{x}), f_m(\bar{x})) = |f_n(\bar{x}) - f_m(\bar{x})| \leq d(f_n, f_m) < \varepsilon, \quad n, m > \nu,$$

e quindi la successione numerica $\{f_n(\bar{x})\}$ è di Cauchy. Poiché $[c, d]$ è completo, tale successione sarà convergente ad un certo numero di $[c, d]$, che indicheremo con $f(\bar{x})$. Poiché questo risultato vale per ogni scelta di $\bar{x} \in [a, b]$, rimane così definita una funzione $f(x)$ da $[a, b]$ in $[c, d]$ che è il limite di $\{f_n(x)\}$. Facciamo vedere che f è continua. Preso $x_0 \in [a, b]$, per la proprietà c) della metrica avremo

$$|f(x) - f(x_0)| \leq d(f(x), f(x_0)) \leq d(f(x), f_n(x)) + d(f_n(x), f_n(x_0)) + d(f_n(x_0), f(x_0)), \quad \forall x \in [a, b].$$

Per quanto detto finora, fissato $\varepsilon > 0$, esisterà un indice ν tale che, per $n > \nu$ e per ogni $x \in [a, b]$,

$$d(f(x), f_n(x)) < \frac{\varepsilon}{3}, \quad d(f(x_0), f_n(x_0)) < \frac{\varepsilon}{3}.$$

Scelto allora $\bar{n} > \nu$, possiamo scrivere

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

Ma $f_{\bar{n}}(x)$ è continua in x_0 , quindi esisterà un intorno $I_\delta(x_0)$ tale che $|f_{\bar{n}}(x) - f_{\bar{n}}(x_0)| < \frac{\varepsilon}{3}$, o, equivalentemente, $x \in I_\delta(x_0) \Rightarrow d(f_{\bar{n}}(x), f_{\bar{n}}(x_0)) < \frac{\varepsilon}{3}$. In conclusione avremo

$$|f(x) - f(x_0)| \leq d(f(x), f(x_0)) < \varepsilon, \quad \forall x \in I_\delta(x_0),$$

ovvero la continuità di $f(x)$.

La convergenza di una successione di funzioni $f_n(x)$ alla funzione $f(x)$, secondo la metrica $d(f, g) = \sup_{x \in [a, b]} |f(x) - g(x)|$, si dice *convergenza uniforme*. Questo tipo di convergenza differisce dalla semplice convergenza *puntuale* di una successione di funzioni. Per chiarire questo concetto osserviamo che, fissato un $\bar{x} \in [a, b]$, la disuguaglianza

$$|f_n(\bar{x}) - f(\bar{x})| < \varepsilon, \quad \text{per } n > \nu,$$

comporta la convergenza della successione numerica $f_n(\bar{x})$ a $f(\bar{x})$. Il numero ν dipenderà da \bar{x} . Se allora accade che

$$\forall x \in [a, b], \quad |f_n(x) - f(x)| < \varepsilon, \quad \text{per } n > \nu(x),$$

avremo la convergenza *puntuale* di f_n a f nell'intervallo $[a, b]$. La convergenza uniforme richiede invece che, per $n > \nu$, sia soddisfatta la disuguaglianza (che non è più condizionata dal valore di x in $[a, b]$),

$$d(f_n, f) = \sup_{x \in [a, b]} |f_n(x) - f(x)| < \varepsilon.$$

Poiché si ha

$$|f_n(x) - f(x)| \leq \sup_{x \in [a, b]} |f_n(x) - f(x)| = d(f_n, f),$$

la convergenza uniforme implica la convergenza puntuale in $[a, b]$. Tuttavia una successione di funzioni può convergere puntualmente in un intervallo ma non essere, in quello stesso intervallo, uniformemente convergente.

- Esempio

Consideriamo la successione di funzioni

$$f_n(x) = n^\alpha x e^{-nx}, \quad x \in [0, 1], \quad \alpha \in \mathbb{R}.$$

Tale successione converge puntualmente a $f(x) = 0$ in $[0, 1]$ in quanto

$$\lim_{n \rightarrow \infty} n^\alpha \bar{x} e^{-n\bar{x}} = 0, \quad \forall \bar{x} \in [0, 1].$$

Consideriamo ora la convergenza uniforme e calcoliamo

$$d(f_n, 0) = \sup_{x \in [0, 1]} |n^\alpha x e^{-nx}| = \max_{x \in [0, 1]} (n^\alpha x e^{-nx}).$$

Considerata come funzione di $x \in [0, 1]$, f_n ha un massimo assoluto in $x = \frac{1}{n}$, da cui

$$d(f_n, 0) = \frac{n^{\alpha-1}}{e}.$$

Se $\alpha < 1$ la successione di funzioni convergerà uniformemente a 0, mentre, se $\alpha \geq 1$, $f_n(x)$ non convergerà uniformemente.

3.3 Teorema delle contrazioni

Sia W uno spazio metrico. Consideriamo una funzione $F : W \rightarrow W$ tale che

$$d(F(u), F(v)) \leq \alpha d(u, v), \quad \forall u, v \in W,$$

con $\alpha \in (0, 1)$. Ogni funzione F che soddisfa questa proprietà si dice una *contrazione* in W .

Teorema 3.2 (*delle contrazioni*). *Dato uno spazio metrico completo W ed una contrazione $F : W \rightarrow W$, esiste un unico punto $\bar{u} \in W$, detto punto fisso di F , tale che*

$$F(\bar{u}) = \bar{u}.$$

Dim. Sia $u_0 \in W$. Costruiamo la successione di punti in W

$$u_1 = F(u_0), \quad u_2 = F(u_1), \quad \dots \quad u_n = F(u_{n-1}), \quad \dots$$

Poiché, per ipotesi, F è una contrazione, avremo

$$d(u_{n+1}, u_n) = d(F(u_n), F(u_{n-1})) \leq \alpha d(u_n, u_{n-1}).$$

Da questa, per induzione, otteniamo

$$d(u_{n+1}, u_n) \leq \alpha^n d(u_1, u_0).$$

Prendiamo un $m > n$. Applicando $m - n$ volte la disuguaglianza triangolare, abbiamo

$$\begin{aligned} d(u_{m+1}, u_n) &\leq d(u_{m+1}, u_m) + d(u_m, u_{m-1}) + \dots + d(u_{n+1}, u_n) \\ &\leq (\alpha^m + \alpha^{m-1} + \dots + \alpha^n) d(u_1, u_0) \\ &= \alpha^n (1 + \alpha + \dots + \alpha^{m-n}) d(u_1, u_0) = \alpha^n \frac{1 - \alpha^{m-n+1}}{1 - \alpha} d(u_1, u_0) \\ &< \frac{\alpha^n}{1 - \alpha} d(u_1, u_0). \end{aligned}$$

Per n sufficientemente grande l'ultimo termine di questa disuguaglianza può essere reso piccolo a piacere, quindi la successione $\{u_n\}$ è di Cauchy. Poiché W è completo, essa è anche convergente e chiameremo \bar{u} il suo limite.

Dimostriamo ora che \bar{u} è un punto fisso per F . Posto $F(\bar{u}) = u^*$, consideriamo la distanza $d(\bar{u}, u^*)$. Si ha

$$d(\bar{u}, u^*) \leq d(\bar{u}, u_n) + d(u_n, u_{n+1}) + d(u_{n+1}, u^*).$$

Poiché $\{u_n\}$ converge a \bar{u} , fissato un $\varepsilon > 0$, si può trovare un $\nu_1 \in \mathbb{N}$ tale che, per $n > \nu_1$

$$d(\bar{u}, u_n) < \frac{\varepsilon}{3}.$$

D'altra parte, essendo $\{u_n\}$ di Cauchy, esisterà un $\nu_2 \in \mathbb{N}$ tale che, per $n > \nu_2$,

$$d(u_n, u_{n+1}) < \frac{\varepsilon}{3},$$

ed inoltre, per $n > \nu_1$,

$$d(u_{n+1}, u^*) = d(F(u_n), F(\bar{u})) \leq \alpha d(u_n, \bar{u}) < \frac{\varepsilon}{3} \alpha < \frac{\varepsilon}{3}.$$

Ne segue che, per $n > \max\{\nu_1, \nu_2\}$,

$$d(\bar{u}, u^*) < \varepsilon.$$

Per l'arbitrarietà di ε si ha allora $\bar{u} = u^*$.

Dimostriamo infine che il punto fisso è unico. Se infatti esistesse un altro punto \hat{u} tale che $F(\hat{u}) = \hat{u}$, allora si avrebbe

$$d(\bar{u}, \hat{u}) = d(F(\bar{u}), F(\hat{u})) \leq \alpha d(\bar{u}, \hat{u}) \quad \Rightarrow \quad d(\bar{u}, \hat{u})(1 - \alpha) \leq 0.$$

Ma $\alpha < 1$ per ipotesi, quindi deve essere

$$d(\bar{u}, \hat{u}) = 0, \quad \Rightarrow \quad \hat{u} = \bar{u}.$$

Capitolo 4

Equazioni differenziali

4.1 Equazioni del primo ordine

Una equazione differenziale di ordine n per la funzione $y(t)$ è una equazione in cui compaiono le derivate, al più di ordine n , della funzione $y(t)$. In generale essa ha la forma

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

Qualora sia possibile esplicitare la derivata di ordine massimo come funzione di t , di $y(t)$ e di tutte le derivate di ordine minore di n , l'equazione differenziale si può porre nella forma

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

detta *forma normale*.

- Esempio

L'equazione

$$(y^{(4)} - 1)(y^{(5)} - t) + 2e^t y' = 0$$

è una equazione differenziale di ordine 5. Essa si può porre in forma normale. Infatti è possibile dividere l'equazione per $(y^{(4)} - 1)$, in quanto, se fosse $y^{(4)} = 1$, integrando quattro volte, si avrebbe $y = \frac{1}{24}t^4 + \frac{1}{6}c_1 t^3 + \frac{1}{2}c_2 t^2 + c_3 t + c_4$. Sostituendo si otterrebbe l'eguaglianza

$$\left(\frac{1}{3}t^3 + c_1 t^2 + 2c_2 t + 2c_3\right) e^t = 0,$$

la quale non è identicamente soddisfatta per alcuna scelta delle costanti c_1, c_2, c_3 . Si può scrivere allora

$$y^{(5)} = t - \frac{2e^t y'}{y^{(4)} - 1},$$

che è la forma normale dell'equazione di partenza.

Una equazione del primo ordine in forma normale si scrive

$$y' = f(t, y)$$

Nel caso più semplice, in cui f non dipende da y , l'equazione si risolve immediatamente tramite una integrazione, in quanto y risulta essere una primitiva di f , e si ha

$$y' = f(t), \quad \Rightarrow \quad y = \int f(t) dt.$$

La funzione y sarà individuata in modo univoco se si conosce il suo valore per un determinato valore di t , per esempio t_0 , e si ha

$$y(t) = y(t_0) + \int_{t_0}^t f(\tau) d\tau.$$

L'insieme di tutte le soluzioni dell'equazione differenziale $y' = f(t, y)$ si dice *integrale generale* mentre ogni soluzione soddisfacente la condizione $y(t_0) = y_0$ con y_0 assegnato, si dice un *integrale particolare*. Più avanti, sotto opportune ipotesi, dimostreremo l'esistenza e l'unicità della soluzione dell'equazione $y' = f(t, y)$ una volta assegnata la condizione $y(t_0) = y_0$. Nella pratica però, spesso non si riesce a trovare l'integrale generale o una soluzione particolare di una equazione differenziale. Esistono inoltre metodi sistematici di soluzione solo nei casi più semplici come quello delle equazioni lineari, che saranno studiati in seguito.

Uno dei casi in cui è possibile integrare una equazione del primo ordine in forma normale è quello in cui la funzione $f(t, y)$ può essere fattorizzata come $f(t, y) = f_1(t)f_2(y)$. In questo caso si dice che l'equazione

$$y' = f_1(t)f_2(y),$$

è a *variabili separabili*. Se si conosce una primitiva F_2 di $1/f_2(y)$ ed una primitiva F_1 di $f_1(t)$, allora l'equazione precedente può essere così integrata

$$\begin{aligned} \frac{y'}{f_2(y)} = f_1(t), \quad \Rightarrow \quad \frac{dy}{f_2(y)} = f_1(t) dt \quad \Rightarrow \quad \int \frac{dy}{f_2(y)} = \int f_1(t) dt \\ \Rightarrow \quad F_2(y) = F_1(t) + c. \end{aligned}$$

L'ultima equazione rappresenta, in forma implicita, l'integrale generale della equazione di partenza. Se la funzione F_2 è invertibile, allora si ottiene l'integrale generale in forma esplicita

$$y(t) = F_2^{-1}[F_1(t) + c].$$

Assegnato $y(0) = y_0$, si ottiene una soluzione particolare con

$$c = F_2(y_0) - F_1(0).$$

- Esempi

1) Trovare l'integrale generale dell'equazione $y' = y/t$. Separando le variabili si ottiene

$$\frac{dy}{y} = \frac{dt}{t} \quad \Rightarrow \quad \int \frac{dy}{y} = \int \frac{dt}{t},$$

da cui si ricava

$$\ln |y| = \ln |t| + c \quad \Rightarrow \quad |y| = k|t|,$$

dove $k = e^c$. Questa soluzione si può sempre scrivere nella forma

$$y = k_1 t, \quad k_1 \neq 0.$$

Osserviamo che nel separare le variabili abbiamo diviso per y , perdendo così la soluzione particolare $y = 0$. Possiamo inglobare questa soluzione nell'integrale generale trovato, eliminando la restrizione $k_1 \neq 0$.

2) Trovare la soluzione particolare dell'equazione

$$y' = (1 + y^2) \sin t,$$

soddisfacente alla *condizione iniziale* $y(0) = 1$. L'equazione è a variabili separabili. Possiamo dividere ambo i membri per $(1 + y^2)$ che è sempre diverso da zero, ottenendo

$$\frac{dy}{1 + y^2} = \sin t \, dt \quad \Rightarrow \quad \arctan y = -\cos t + c,$$

da cui ricaviamo, esplicitamente, l'integrale generale, $y(t) = \tan(-\cos t + c)$. Imponendo la condizione iniziale si ha

$$\arctan 1 = -\cos 0 + c \quad \Rightarrow \quad c = \frac{\pi}{4} + 1.$$

La soluzione richiesta è dunque,

$$y(t) = \tan\left(1 + \frac{\pi}{4} - \cos t\right).$$

Certe equazioni possono essere ricondotte nella forma a variabili separabili. Per esempio l'equazione

$$y' = f\left(\frac{y}{t}\right)$$

detta di tipo *omogeneo*, può essere risolta mediante la sostituzione $y = tz$. In tal modo si ha $y' = z + tz'$ e l'equazione diventa una equazione a variabili separabili per la funzione $z(t)$,

$$z't = f(z) - z, \quad \Rightarrow \quad \frac{dz}{f(z) - z} = \frac{dt}{t}.$$

L'integrale generale si può allora esprimere in forma implicita come

$$\int \frac{dz}{f(z) - z} = \ln |t| + c.$$

Un altro esempio è quello delle equazioni nella forma

$$y' = f(at + by).$$

In tal caso infatti, effettuando la sostituzione $z = at + by$, si ha $y' = (z' - a)/b$ e quindi

$$\frac{dz}{bf(z) + a} = dt \quad \Rightarrow \quad t = \int \frac{dz}{bf(z) + a}.$$

- Esempi

1) Determinare l'integrale generale della equazione

$$y' = \frac{y^2 + ty + t^2}{t^2}.$$

Si tratta di una equazione omogenea in quanto il secondo membro si può scrivere come $z^2 + z + 1$, essendo $z = y/t$. Effettuando la sostituzione e integrando si ottiene

$$\frac{z'}{z^2 + 1} = \frac{1}{t} \Rightarrow \arctan z = \ln |t| + c,$$

da cui

$$z(t) = \tan(\ln |t| + c) \Rightarrow y(t) = t \tan(\ln |t| + c).$$

2) Determinare l'integrale generale dell'equazione

$$y' = \frac{1}{t - y} + 1.$$

Effettuiamo la sostituzione $t - y = z$. Otteniamo

$$z' = -\frac{1}{z} \Rightarrow z dz = -dt,$$

da cui, integrando, si ricava l'integrale generale

$$z^2 = -2t + c \Rightarrow (t - y)^2 = -2t + c.$$

Una equazione del primo ordine in forma normale si dice *lineare* se $f(t, y)$ è una funzione lineare di y cioè se

$$f(t, y) = q(t) + p(t)y.$$

Supporremo qui che le funzioni $p(t)$ e $q(t)$ siano continue in un dominio $D \subseteq \mathbb{R}$. L'equazione lineare

$$y' - p(t)y = q(t),$$

si dice *omogenea* se $q(t)$ è identicamente nulla. In quest'ultimo caso l'equazione è a variabili separabili e si ha

$$\frac{dy}{y} = p(t) dt, \Rightarrow \ln |y| = \int p(t) dt + c = P(t) + c,$$

dove con $P(t)$ abbiamo indicato una primitiva di $p(t)$. Si ricava così l'integrale generale dell'equazione omogenea,

$$\bar{y}(t) = ke^{P(t)},$$

dove $k \in \mathbb{R}$. Osserviamo in particolare che la soluzione $\bar{y} = 0$, persa nella separazione di variabili, si ottiene per $k = 0$.

L'integrazione dell'equazione non omogenea si può effettuare mediante il cosiddetto metodo della *variazione della costante*, una volta noto l'integrale generale dell'equazione omogenea corrispondente, detta *equazione omogenea associata*. A questo scopo cerchiamo una soluzione particolare nella forma

$$\hat{y}(t) = k(t)e^{P(t)},$$

dove $k(t)$ è una funzione continua tale che $k(t_0) = 0$, con $t_0 \in D$. Verifichiamo che tale soluzione esiste e calcoliamo $k(t)$. Sostituendo nell'equazione di partenza otteniamo infatti

$$[k'(t) + k(t)p(t)]e^{P(t)} - p(t)k(t)e^{P(t)} = q(t),$$

da cui

$$k'(t) = q(t)e^{-P(t)}.$$

Integrando sull'intervallo $[t_0, t] \subseteq D$ e ricordando che $k(t_0) = 0$, otteniamo

$$k(t) = \int_{t_0}^t q(\tau)e^{-P(\tau)} d\tau.$$

Abbiamo così dimostrato che

$$\hat{y}(t) = e^{P(t)} \int_{t_0}^t q(\tau)e^{-P(\tau)} d\tau$$

è una soluzione particolare dell'equazione non omogenea. Consideriamo ora una qualsiasi altra soluzione $\tilde{y}(t)$ dell'equazione non omogenea. Poiché si ha contemporaneamente

$$\begin{aligned}\tilde{y}' &= p(t)\tilde{y} + q(t) \\ \hat{y}' &= p(t)\hat{y} + q(t).\end{aligned}$$

sottraendo membro a membro queste due eguaglianze, otteniamo

$$\tilde{y}' - \hat{y}' = p(t)(\tilde{y} - \hat{y}).$$

Ne segue che $\tilde{y} - \hat{y}$ è soluzione dell'equazione omogenea associata, cioè esiste una costante k per cui $\tilde{y} - \hat{y} = ke^{P(t)}$. Concludiamo allora che tutte le soluzioni dell'equazione non omogenea hanno la forma

$$y(t) = ke^{P(t)} + \hat{y} = ke^{P(t)} + e^{P(t)} \int_{t_0}^t q(\tau)e^{-P(\tau)} d\tau.$$

In altri termini, l'integrale generale dell'equazione non omogenea è dato dalla somma della soluzione particolare $\hat{y}(t)$ e dell'integrale generale dell'equazione omogenea associata. Più avanti generalizzeremo questa conclusione al caso di equazioni lineari di qualsiasi ordine.

- Esempi

1) Determinare l'integrale generale dell'equazione lineare

$$y' + y \tan t = \frac{1}{\sin t}.$$

Ricaviamo prima la soluzione dell'omogenea associata

$$y' = -y \tan t \quad \Rightarrow \quad \frac{y'}{y} = -\tan t,$$

da cui, integrando,

$$\ln |y| = \ln |\cos t| + c \Rightarrow \bar{y} = k \cos t, \quad k \in \mathbb{R}.$$

Facciamo variare la costante ponendo $k = k(t)$. Abbiamo

$$y = k(t) \cos t \Rightarrow y' = k' \cos t - k \sin t.$$

Sostituendo nell'equazione non omogenea ricaviamo

$$k'(t) = \frac{1}{\sin t \cos t},$$

da cui integrando,

$$k(t) = -\ln |\cos t| + \ln |\sin t| + k_1 = -\ln |\cot t| + k_1.$$

Ne segue l'integrale generale della non omogenea

$$y(t) = k_1 \cos t - \cos t \ln |\cot t|.$$

2) Determinare l'integrale generale dell'equazione lineare

$$y' + 2ty = t + t^3.$$

Dalla omogenea associata si ricava $y = ke^{-t^2}$ con $k \in \mathbb{R}$. Facendo variare la costante si ha $y' = (k' - 2kt)e^{-t^2}$. Sostituendo si ottiene

$$k' = (t + t^3)e^{t^2},$$

la cui integrazione fornisce

$$k(t) = \frac{1}{2}t^2e^{t^2} + k_1.$$

Si ottiene così l'integrale generale della non omogenea

$$y = k_1e^{-t^2} + \frac{1}{2}t^2.$$

Alcune equazioni differenziali del primo ordine in forma normale possono essere ridotte mediante una opportuna sostituzione ad equazioni lineari. Consideriamo per esempio l'*equazione di Bernoulli*

$$y' + p(t)y = f(t)y^n, \quad \text{con } n \neq 1.$$

Se dividiamo per y^n (osserviamo che $y = 0$ è comunque soluzione), abbiamo

$$y'y^{-n} + p(t)y^{1-n} = f(t).$$

Posto $z = y^{1-n}$, otteniamo $z' = (1-n)y'y^{-n}$ da cui, sostituendo troviamo una equazione lineare per z ,

$$\frac{1}{1-n}z' + p(t)z = f(t).$$

Consideriamo l'equazione di Riccati

$$y' + p(t)y + q(t)y^2 = f(t).$$

È possibile trasformare questa equazione in una equazione di Bernoulli e quindi in una equazione lineare. Per raggiungere tale scopo è necessario però conoscere una soluzione particolare $y_1(t)$ dell'equazione di Riccati. Posto $y = y_1 + z$, si ha

$$y_1' + z' + p(t)(y_1 + z) + q(t)(y_1 + z)^2 = f(t),$$

e sfruttando il fatto che y_1 è una soluzione, ci si riduce all'equazione

$$z' + [p(t) + 2q(t)y_1]z + q(t)z^2 = 0,$$

che è una equazione di Bernoulli.

Oltre a singole equazioni differenziali studieremo i sistemi di equazioni differenziali, le cui incognite sono n -uple di funzioni $y_i(t)$, ($i = 1, \dots, n$). Un sistema di equazioni differenziali del primo ordine in forma normale può essere scritto come

$$\begin{aligned} y_1' &= f_1(t, y_1, y_2, \dots, y_n) \\ y_2' &= f_2(t, y_1, y_2, \dots, y_n) \\ &\vdots \\ y_n' &= f_n(t, y_1, y_2, \dots, y_n) \end{aligned}$$

L'integrale generale del sistema è costituito da tutte le n -uple $y_i(t)$, ($i = 1, \dots, n$) che soddisfano tutte le equazioni del sistema. Si può dare una forma più compatta alle equazioni precedenti introducendo una funzione vettoriale $\mathbf{y}(t)$ di componenti $y_i(t)$, ($i = 1, \dots, n$) a valori in \mathbb{R}^n ed una funzione vettoriale di $n + 1$ variabili $\mathbf{f} : \mathbb{R}^{n+1} \rightarrow \mathbb{R}^n$. Scriveremo allora

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y}).$$

È importante osservare che se sappiamo risolvere un sistema di equazioni differenziali del primo ordine in forma normale, possiamo anche risolvere equazioni o sistemi di equazioni in forma normale di ordine superiore al primo. Consideriamo il seguente esempio. L'equazione

$$y^{(4)} + 2y''y'e^y = t(1 + y''),$$

è del quarto ordine e può essere espressa in forma normale. Poniamo

$$y =: u_1, \quad y' = u_1' =: u_2, \quad y'' = u_2' =: u_3, \quad y''' = u_3' =: u_4.$$

L'equazione data si può scrivere nella forma di un sistema per u_1, u_2, u_3, u_4 , cioè

$$\begin{aligned} u_1' &= u_2 \\ u_2' &= u_3 \\ u_3' &= u_4 \\ u_4' &= -2u_3u_2e^{u_1} + t(1 + u_3), \end{aligned}$$

ovvero nella forma

$$\mathbf{u}' = \mathbf{f}(t, \mathbf{u}),$$

dove

$$\mathbf{u} = (u_1, u_2, u_3, u_4), \quad \mathbf{f} = (f_1, f_2, f_3, f_4),$$

con

$$\begin{aligned} f_1(t, u_1, u_2, u_3, u_4) &= u_2, & f_2(t, u_1, u_2, u_3, u_4) &= u_3, \\ f_3(t, u_1, u_2, u_3, u_4) &= u_4, & f_4(t, u_1, u_2, u_3, u_4) &= -2u_3u_2e^{u_1} + t(1 + u_3). \end{aligned}$$

In generale, una equazione di ordine k in forma normale si può trasformare in un sistema di k equazioni del primo ordine in forma normale.

4.2 Teorema di esistenza e unicità

Consideriamo un sistema di equazioni differenziali del primo ordine in forma normale

$$\mathbf{y}' = \mathbf{f}(t, \mathbf{y}),$$

e supponiamo di avere assegnato il valore della funzione vettoriale $\mathbf{y}(t)$ per $t = t_0$, ovvero $\mathbf{y}(t_0) = \mathbf{y}^0$. Si dice *problema di Cauchy* il sistema

$$\begin{cases} \mathbf{y}' = \mathbf{f}(t, \mathbf{y}) \\ \mathbf{y}(t_0) = \mathbf{y}^0, \end{cases}$$

formato dall'equazione differenziale vettoriale per la funzione $\mathbf{y}(t)$, in forma normale e dalla condizione iniziale imposta su tale funzione. Sotto opportune ipotesi di regolarità della funzione vettoriale \mathbf{f} e della funzione $\mathbf{y}(t)$ si può dimostrare che esiste una ed una sola soluzione al problema di Cauchy.

Consideriamo per il momento il caso di una sola equazione differenziale per la funzione $y(t)$ e scriviamo il problema di Cauchy nella forma

$$\begin{cases} y' = f(t, y) \\ y(t_0) = y_0. \end{cases}$$

Supponiamo che la funzione $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ sia continua e integriamo tra t_0 e t ambo i membri dell'equazione differenziale. La funzione $y(t)$ risulterà soddisfare l'equazione integrale

$$y(t) = y_0 + \int_{t_0}^t f(\tau, y(\tau)) \, d\tau.$$

D'altra parte, se $y(t)$ è una funzione continua con derivata continua, che sia soluzione di questa equazione integrale, allora risulta $y(t_0) = y_0$ e $y' = f(t, y(t))$, cioè essa è soluzione del problema di Cauchy. In sintesi possiamo dire che una funzione $y(t)$ con derivata continua è soluzione del problema di Cauchy se e solo se soddisfa l'equazione integrale precedente.

Cerchiamo soluzioni $y(t) \in C^1(D)$, essendo $D = \{t \in \mathbb{R} : |t - t_0| \leq \delta\}$. Poiché tali funzioni sono limitate e $y(t_0) = y_0$, esse assumeranno valori in un dominio B esprimibile come

$$B = \{y \in \mathbb{R} : |y - y_0| \leq b\},$$

e di conseguenza l'insieme delle possibili soluzioni del problema di Cauchy sarà

$$W = \{y(t) \in C^1(D) \mid y : D \rightarrow B\}.$$

Nell'ipotesi che f sia continua in $D \times B$ essa sarà limitata, cioè

$$|f(t, y)| \leq M, \quad \forall (t, y) \in D \times B.$$

Diremo poi che f soddisfa la *condizione di Lipschitz* se esiste un $\eta > 0$ tale che

$$|f(t, y_1) - f(t, y_2)| \leq \eta |y_1 - y_2|, \quad \forall t \in D, \quad \forall y_1, y_2 \in B.$$

Tenuto conto di tutte queste osservazioni, possiamo dimostrare il seguente risultato.

Teorema 4.1 *Esiste una ed una sola soluzione $y \in W$ al problema di Cauchy*

$$\begin{cases} y' = f(t, y) \\ y(t_0) = y_0, \end{cases}$$

con f continua e soddisfacente la condizione di Lipschitz.

Dim. Nell'insieme $C^1(D)$ introduciamo la metrica

$$d(y_1, y_2) = \sup_{t \in D} |y_1(t) - y_2(t)| = \max_{t \in D} |y_1(t) - y_2(t)|.$$

Per il teorema 3.1 il sottoinsieme W con la suddetta metrica è uno spazio metrico completo. Consideriamo la funzione F definita in W da

$$F(y)(t) = y_0 + \int_{t_0}^t f(\tau, y(\tau)) \, d\tau.$$

Poiché f è continua in D , per ogni scelta di y in W , la funzione $F(y)(t)$ risulta avere derivata prima continua. Inoltre, scelto $\delta \leq b/M$, si ha

$$\left| \int_{t_0}^t f(\tau, y(\tau)) \, d\tau \right| \leq \left| \int_{t_0}^t |f(\tau, y(\tau))| \, d\tau \right| \leq M|t - t_0| \leq M\delta \leq b,$$

quindi le funzioni $F(y)(t)$, ottenute scegliendo y in W , hanno valori in B . Riassumendo, fissato b e scelto $\delta \leq b/M$, si ha $F : W \rightarrow W$.

Dimostriamo ora che, precisato ulteriormente δ , e cioè scelto $\delta < \min\{1/\eta, b/M\}$, F è una contrazione. Infatti, tenuto conto della condizione di Lipschitz, $\forall y_1, y_2 \in W$, si ha

$$\begin{aligned} |F(y_1) - F(y_2)| &= \left| \int_{t_0}^t [f(\tau, y_1(\tau)) - f(\tau, y_2(\tau))] \, d\tau \right| \leq \\ &\left| \int_{t_0}^t |f(\tau, y_1(\tau)) - f(\tau, y_2(\tau))| \, d\tau \right| \leq \eta \left| \int_{t_0}^t |y_1(\tau) - y_2(\tau)| \, d\tau \right| \leq \\ &\eta |t - t_0| \max_{t \in D} |y_1(t) - y_2(t)| \leq \eta \delta \max_{t \in D} |y_1(t) - y_2(t)|, \end{aligned}$$

da cui otteniamo

$$d(F(y_1), F(y_2)) = \max_{t \in D} |F(y_1) - F(y_2)| \leq \eta \delta \max_{t \in D} |y_1(t) - y_2(t)| = \alpha d(y_1, y_2),$$

dove $\alpha = \eta \delta < 1$. Per il teorema delle contrazioni esiste allora un unico punto fisso \bar{y} dello spazio metrico completo W tale che $F(\bar{y}) = \bar{y}$, ovvero una funzione $\bar{y}(t) \in W$ tale che

$$\bar{y}(t) = y_0 + \int_{t_0}^t f(\tau, \bar{y}(\tau)) \, d\tau.$$

Dato che quest'ultima equazione è equivalente al problema di Cauchy, perveniamo alla tesi.

Il teorema di esistenza e unicità si può estendere senza difficoltà al caso di un sistema di equazioni differenziali del primo ordine in forma normale,

$$\begin{cases} \mathbf{y}' = \mathbf{f}(t, \mathbf{y}) \\ \mathbf{y}(t_0) = \mathbf{y}^0. \end{cases}$$

In questo caso si può ripetere la dimostrazione precedente lavorando nello spazio delle funzioni vettoriali con derivate continue, definite in D , a valori in \mathbb{R}^n , con la metrica

$$d(\mathbf{y}_1, \mathbf{y}_2) = \max_{t \in D} \|\mathbf{y}_1(t) - \mathbf{y}_2(t)\|,$$

dove $\|\cdot\|$ indica l'usuale norma in \mathbb{R}^n . La funzione vettoriale \mathbf{f} si suppone continua in $D \times \mathbb{R}^n$ e la condizione di Lipschitz si scrive nella forma

$$\|\mathbf{f}(t, \mathbf{y}_1) - \mathbf{f}(t, \mathbf{y}_2)\| \leq \eta \|\mathbf{y}_1 - \mathbf{y}_2\|, \quad \forall t \in D, \quad \forall \mathbf{y}_1, \mathbf{y}_2 \in B,$$

dove $B = \{\mathbf{y} \in \mathbb{R}^n : \|\mathbf{y} - \mathbf{y}^0\| \leq b\}$.

OSSERVAZIONE. Si può dimostrare che la condizione di Lipschitz è soddisfatta se la funzione $f(t, y)$ ha derivata parziale rispetto a y continua. Ciò vale anche nel caso di un sistema di equazioni, nell'ipotesi cioè che la funzione \mathbf{f} abbia derivate parziali rispetto a ciascuna delle componenti di \mathbf{y} , continue.

4.3 Equazioni differenziali lineari

Una equazione differenziale lineare di ordine n ha la forma

$$a_0(t)y^{(n)} + a_1(t)y^{(n-1)} + \dots + a_{n-1}(t)y' + a_n(t)y = \phi(t).$$

Se $\phi(t)$ è identicamente nulla, l'equazione si dice *omogenea*.

Supponiamo che la funzione $a_0(t)$ non si annulli mai in un dominio $D \subset \mathbb{R}$ e dividiamo ambo i membri dell'equazione data per $a_0(t)$. Allora si ottiene l'equazione

$$y^{(n)} + p_1(t)y^{(n-1)} + \dots + p_n(t)y = f(t),$$

che è riducibile a forma normale, in cui cioè, è possibile esplicitare la derivata di ordine massimo. Come sappiamo questa equazione è equivalente ad un sistema di n equazioni differenziali del primo ordine in forma normale. Se si assegnano gli n dati iniziali

$$y(t_0) = y_0, \quad y'(t_0) = y'_0, \quad \dots \quad y^{(n-1)}(t_0) = y_0^{(n-1)},$$

nell'ipotesi di continuità delle funzioni $p_i(t)$, ($i = 1, \dots, n$) e della $f(t)$ in D , l'osservazione alla fine del paragrafo precedente implica che siano soddisfatte le ipotesi del teorema di esistenza e unicità.

Sia $C^{(n)}(D)$ lo spazio delle funzioni derivabili n volte in D con derivata n -esima continua. Consideriamo l'applicazione $L : C^{(n)}(D) \rightarrow C(D)$ definita da

$$L(y) = y^{(n)} + p_1(t)y^{(n-1)} + \dots + p_n(t)y.$$

Per ogni coppia di funzioni $y_1(t)$ e $y_2(t)$ in $C^{(n)}(D)$ e per ogni scalare c si ha

$$L(y_1 + y_2) = L(y_1) + L(y_2), \quad L(cy_1) = cL(y_1).$$

L'applicazione risulta dunque lineare. L'equazione differenziale si può scrivere nella forma

$$L(y) = f.$$

Notiamo inoltre che se y_1 e y_2 sono due soluzioni particolari di questa equazione, la loro differenza $y_1 - y_2$ è soluzione dell'equazione lineare omogenea $L(y) = 0$, infatti, per la linearità di L ,

$$L(y_1 - y_2) = L(y_1) - L(y_2) = f - f = 0.$$

Se allora si conosce una soluzione particolare \hat{y} dell'equazione non omogenea, ogni altra soluzione di questa equazione si otterrà aggiungendo a \hat{y} una soluzione della omogenea $L(y) = 0$ che chiameremo equazione *omogenea associata*. In altri termini tutte le soluzioni dell'equazione non omogenea $L(y) = f$ sono date da

$$y = \hat{y} + \bar{y} \quad \text{con} \quad \bar{y} \in \text{Ker}(L).$$

Studiamo allora l'equazione lineare omogenea $L(y) = 0$.

Teorema 4.2 *L'insieme delle soluzioni dell'equazione lineare omogenea di ordine n*

$$L(y) = 0,$$

costituisce uno spazio vettoriale di dimensione n .

Dim. La tesi equivale a dire che $\dim[\text{Ker}(L)] = n$. Innanzitutto osserviamo che $\text{Ker}(L)$ è un sottospazio vettoriale di $C^{(n)}(D)$. Facciamo vedere quindi che esistono n elementi di $\text{Ker}(L)$ linearmente indipendenti e che tali elementi generano $\text{Ker}(L)$. Per raggiungere il primo obiettivo consideriamo i seguenti n problemi di Cauchy

$$\begin{aligned} 1) \quad & L(y) = 0, & \{y^{(k-1)}(t_0) = \delta_{1k}, k = 1, \dots, n\} \\ 2) \quad & L(y) = 0, & \{y^{(k-1)}(t_0) = \delta_{2k}, k = 1, \dots, n\} \\ & \vdots & \\ n) \quad & L(y) = 0, & \{y^{(k-1)}(t_0) = \delta_{nk}, k = 1, \dots, n\} \end{aligned}$$

dove con δ_{ij} ($i, j = 1, \dots, n$) si intende la delta di Kronecker. Ciascuno di questi problemi ammetterà una soluzione y_i , ($i = 1, \dots, n$), che sarà un elemento di $\text{Ker}(L)$ coincidente con una soluzione particolare di $L(y) = 0$. Costruiamo ora una combinazione lineare di queste n soluzioni,

$$y = \sum_{i=1}^n c_i y_i.$$

Per la linearità di L , essa sarà ancora soluzione dell'equazione omogenea. Dalle condizioni iniziali sulle y_i , segue che la funzione y è l'elemento di $\text{Ker}(L)$ soddisfacente le condizioni

$$y^{(k-1)}(t_0) = c_k, \quad k = 1, \dots, n.$$

La soluzione identicamente nulla è ancora un elemento di $\text{Ker}(L)$ e si ricava dalla combinazione precedente scegliendo $c_1 = c_2 = \dots = c_n = 0$. Poiché la soluzione nulla è unica, la combinazione $c_1 y_1 + c_2 y_2 + \dots + c_n y_n$ si annulla se e solo se si annullano tutti i coefficienti c_i . In altri termini, gli

elementi $y_i \in \text{Ker}(L)$, ($i = 1, \dots, n$) sono linearmente indipendenti. Dimostriamo ora che essi generano $\text{Ker}(L)$. Sia z una qualunque soluzione particolare di $L(y) = 0$, soddisfacente le condizioni iniziali

$$z^{(k-1)}(t_0) = \eta_k, \quad k = 1, \dots, n.$$

Si verifica subito che essa è esprimibile come

$$z(t) = \sum_{k=1}^n c_k y_k(t),$$

purchè si scelga $c_k = \eta_k$, ($k = 1, 2, \dots, n$). Infatti la combinazione $z = \sum_{k=1}^n \eta_k y_k$ è proprio l'elemento di $\text{Ker}(L)$ corrispondente alle condizioni iniziali imposte. Ne segue che ogni elemento di $\text{Ker}(L)$ è dato da una combinazione lineare delle y_i . In definitiva le n soluzioni y_i costituiscono una base di $\text{Ker}(L)$ e quindi $\dim[\text{Ker}(L)] = n$.

Chiameremo *sistema fondamentale* di soluzioni di una equazione lineare omogenea di ordine n , un insieme arbitrario di n soluzioni particolari linearmente indipendenti di questa equazione. L'integrale generale sarà dato dalle combinazioni lineari degli elementi di un sistema fondamentale.

Siano $y_1(t), y_2(t), \dots, y_n(t)$, n funzioni $C^n(D)$. Consideriamo il determinante

$$W(t) = \begin{vmatrix} y_1 & y_2 & \dots & y_n \\ y_1' & y_2' & \dots & y_n' \\ \vdots & \vdots & \dots & \vdots \\ y_1^{(n-1)} & y_2^{(n-1)} & \dots & y_n^{(n-1)} \end{vmatrix},$$

detto *determinante di Wronsky* (o *wronskiano*) dell'insieme di funzioni $y_k(t)$, $k = 1, 2, \dots, n$. Dimostriamo che se $\{y_k(t)\}$, ($t \in D$) è un sistema fondamentale di soluzioni dell'equazione differenziale omogenea di ordine n , $L(y) = 0$, il suo wronskiano è diverso da zero in tutto D . Supponiamo per assurdo che esista un $t_0 \in D$ tale che $W(t_0) = 0$ e consideriamo il sistema algebrico lineare nelle incognite $\alpha_1, \alpha_2, \dots, \alpha_n$,

$$\begin{aligned} \sum_{k=1}^n y_k(t_0) \alpha_k &= 0 \\ \sum_{k=1}^n y_k'(t_0) \alpha_k &= 0 \\ \vdots & \\ \sum_{k=1}^n y_k^{(n-1)}(t_0) \alpha_k &= 0. \end{aligned}$$

Il determinante della matrice dei coefficienti di questo sistema è proprio $W(t_0)$ e di conseguenza il sistema ammette soluzioni non banali. Ne segue che la combinazione lineare

$$\sum_{k=1}^n y_k(t) \alpha_k,$$

di cui il sistema precedente rappresenta le condizioni iniziali, è una soluzione particolare di $L(y) = 0$ soddisfacente condizioni iniziali tutte nulle. Per l'unicità della soluzione, tale combinazione deve essere

la funzione identicamente nulla. In altri termini si ha

$$\sum_{k=1}^n y_k(t) \alpha_k = 0, \quad \forall t \in D,$$

con α_k non tutte nulle. Ma ciò contraddice l'indipendenza lineare delle funzioni $y_k(t)$. Per l'arbitrarietà di t_0 concludiamo che $W(t) \neq 0, \forall t \in D$.

• Esempio

Consideriamo l'equazione lineare omogenea

$$y''' - y' = 0.$$

Si tratta di una equazione del terzo ordine. Essa ammette come sistema fondamentale una qualunque terna di soluzioni particolari linearmente indipendenti. Notiamo che le seguenti funzioni

$$y_1 = 1, \quad y_2 = e^t, \quad y_3 = \cosh t, \quad y_4 = \sinh t,$$

sono tutte soluzioni particolari dell'equazione data. Di queste però solo tre sono linearmente indipendenti. Per esempio, le terne

$$\{y_1, y_2, y_3\} \quad \{y_1, y_2, y_4\} \quad \{y_1, y_3, y_4\},$$

sono sistemi fondamentali dell'equazione data. L'integrale generale si può scrivere, per esempio, come

$$y = c_1 + c_2 e^t + c_3 \sinh t.$$

Osserviamo che in questo caso

$$W(t) = \begin{vmatrix} 1 & e^t & \sinh t \\ 0 & e^t & \cosh t \\ 0 & e^t & \sinh t \end{vmatrix} = -1,$$

ovvero il wronskiano è diverso da zero per ogni t .

Ritorniamo al problema dell'equazione lineare non omogenea di ordine n ,

$$L(y) = f.$$

Riassumendo i risultati ottenuti fin qui, possiamo concludere che, data una sua soluzione particolare \hat{y} , l'equazione non omogenea ha la soluzione generale

$$y = \sum_{i=1}^n c_i y_i + \hat{y}.$$

dove $\{y_1, y_2, \dots, y_n\}$ è un sistema fondamentale dell'equazione omogenea associata.

4.4 Equazioni omogenee a coefficienti costanti

Consideriamo l'equazione lineare omogenea di ordine n , $L(y) = 0$ in cui, al posto delle funzioni $p_i(t)$ ($i = 1, \dots, n$) abbiamo n costanti a_i , ($i = 1, \dots, n$),

$$L(y) = y^{(n)} + a_1 y^{(n-1)} + \dots + a_n y = 0.$$

In questo caso, l'applicazione $L : C^{(n)} \rightarrow C$ si può scrivere sotto forma di composizione di applicazioni. Consideriamo per esempio il caso di una equazione omogenea del secondo ordine

$$y'' + a_1 y' + a_2 y = 0.$$

e introduciamo l'applicazione di derivazione $D : V^{(n)} \rightarrow C$ definita da $D(y) = y'$. L'equazione precedente si può scrivere allora nella forma

$$(D - \lambda_1)(D - \lambda_2)y = 0,$$

dove λ_1, λ_2 sono le radici dell'equazione $\lambda^2 + a_1 \lambda + a_2 = 0$. Più in generale si ha

$$L(y) = y^{(n)} + a_1 y^{(n-1)} + \dots + a_n y = (D - \lambda_1)(D - \lambda_2) \dots (D - \lambda_n)y,$$

dove $\lambda_1, \lambda_2, \dots, \lambda_n$ sono le radici dell'equazione, detta *equazione caratteristica*,

$$\lambda^n + a_1 \lambda^{n-1} + \dots + a_n = 0.$$

Le applicazioni $D - \lambda_i$ sono lineari ed inoltre le loro composizioni sono commutative. Infatti

$$\begin{aligned} (D - \lambda_i)(D - \lambda_j)y &= y'' - (\lambda_i + \lambda_j)y' + \lambda_i \lambda_j y, \\ (D - \lambda_j)(D - \lambda_i)y &= y'' - (\lambda_j + \lambda_i)y' + \lambda_j \lambda_i y. \end{aligned}$$

A questo punto osserviamo che il nucleo dell'applicazione L contiene tutti gli elementi dei nuclei delle applicazioni $D - \lambda_i$ e, di conseguenza, le soluzioni delle singole equazioni del primo ordine

$$(D - \lambda_1)y = 0, \quad (D - \lambda_2)y = 0, \quad \dots \quad (D - \lambda_n)y = 0,$$

sono soluzioni particolari dell'equazione $L(y) = 0$. Poiché la soluzione dell'equazione $(D - \lambda_i)y = 0$ è data da

$$c_i e^{\lambda_i t},$$

se l'equazione caratteristica ammette n radici distinte, le funzioni $e^{\lambda_i t}$, ($i = 1, \dots, n$) saranno linearmente indipendenti e formeranno quindi un sistema fondamentale per l'equazione omogenea. In altri termini l'equazione $L(y) = 0$ avrà l'integrale generale

$$y(t) = \sum_{i=1}^n c_i e^{\lambda_i t}.$$

- Esempi

1) Determinare l'integrale generale dell'equazione

$$y'' - 3y' + 2y = 0.$$

L'equazione caratteristica è

$$\lambda^2 - 3\lambda + 2 = 0,$$

da cui si ottiene $\lambda_1 = 1$, $\lambda_2 = 2$. Allora l'integrale generale sarà

$$y = c_1 e^t + c_2 e^{2t}.$$

2) Determinare l'integrale generale dell'equazione

$$y''' - y' = 0.$$

Si ha l'equazione caratteristica $\lambda^3 - \lambda = 0$ da cui si ricavano le radici $\lambda_1 = 0$, $\lambda_2 = -1$, $\lambda_3 = 1$. L'integrale generale sarà

$$y = c_1 + c_2 e^{-t} + c_3 e^t.$$

L'equazione caratteristica può ammettere radici complesse. In tal caso, se ammette la radice $\lambda = \alpha + i\beta$, ammetterà anche la radice coniugata $\bar{\lambda} = \alpha - i\beta$. Usando la formula di Eulero, le due soluzioni particolari corrispondenti, si possono porre nella forma

$$\begin{aligned} y_\lambda &= e^{\lambda t} = e^{\alpha t} (\cos \beta t + i \sin \beta t), \\ y_{\bar{\lambda}} &= e^{\bar{\lambda} t} = e^{\alpha t} (\cos \beta t - i \sin \beta t), \end{aligned}$$

Possiamo sempre sostituire queste soluzioni con una coppia di loro combinazioni lineari. Volendo ottenere soluzioni reali, scriveremo

$$\begin{aligned} y_1 &= \frac{1}{2}(e^{\lambda t} + e^{\bar{\lambda} t}) = e^{\alpha t} \cos \beta t, \\ y_2 &= \frac{1}{2i}(e^{\lambda t} - e^{\bar{\lambda} t}) = e^{\alpha t} \sin \beta t, \end{aligned}$$

Quindi, ad ogni combinazione lineare $c_1 e^{\lambda t} + c_2 e^{\bar{\lambda} t}$ presente nella soluzione di $L(y) = 0$ possiamo sostituire la combinazione $e^{\alpha t}(c_1 \cos \beta t + c_2 \sin \beta t)$.

- Esempio

Risolvere l'equazione dell'oscillatore armonico smorzato,

$$s'' + 2hs' + \omega^2 s = 0, \quad (h > 0).$$

L'equazione caratteristica è $\lambda^2 + 2h\lambda + \omega^2 = 0$ da cui si ottengono le radici

$$\lambda_1 = -h + \sqrt{h^2 - \omega^2}, \quad \lambda_2 = -h - \sqrt{h^2 - \omega^2}.$$

Se $h > \omega$, le due radici sono reali e, posto $\sqrt{h^2 - \omega^2} = \nu$, si ha l'integrale generale

$$s(t) = e^{-ht}(c_1 e^{\nu t} + c_2 e^{-\nu t}).$$

Se $h < \omega$, le due radici sono complesse e posto $\sqrt{\omega^2 - h^2} = \mu$ si ha

$$\lambda_1 = -h + i\mu, \quad \lambda_2 = -h - i\mu,$$

da cui segue l'integrale generale

$$s(t) = c_1 e^{-ht+i\mu t} + c_2 e^{-ht-i\mu t} = e^{-ht}(\alpha_1 \cos \mu t + \alpha_2 \sin \mu t).$$

Supponiamo ora che le radici dell'equazione caratteristica non siano tutte distinte ma ce ne sia una, chiamiamola genericamente λ , di molteplicità m . Allora le soluzioni particolari corrispondenti dell'equazione $L(y) = 0$ sono soluzioni dell'equazione

$$(D - \lambda)^m y = 0.$$

Osserviamo che vale la seguente identità, facilmente dimostrabile per induzione sull'indice m ,

$$(D - \lambda)^m t^p e^{\lambda t} = p(p-1)(p-2)\dots(p-m+1)t^{p-m} e^{\lambda t}.$$

In forza di questa identità, per ogni $p \leq m-1$, la funzione $t^p e^{\lambda t}$ è soluzione dell'equazione $(D - \lambda)^m y = 0$. Poiché le funzioni

$$e^{\lambda t}, t e^{\lambda t}, \dots, t^{m-1} e^{\lambda t},$$

sono linearmente indipendenti, la soluzione generale dell'equazione $L(y) = 0$ sarà data in questo caso da

$$y(t) = \sum_{i=1}^{n-m} \hat{c}_i e^{\hat{\lambda}_i t} + \sum_{i=1}^m c_i t^{i-1} e^{\lambda t},$$

dove $\hat{\lambda}_i$ sono le rimanenti soluzioni distinte dell'equazione caratteristica.

- Esempi

1) Determinare l'integrale generale dell'equazione

$$y'' - 2y' + y = 0.$$

Dall'equazione caratteristica $\lambda^2 - 2\lambda + 1 = 0$ ricaviamo la radice doppia $\lambda = 1$. Allora la soluzione è

$$y(t) = c_1 e^t + c_2 t e^t.$$

2) Determinare l'integrale generale dell'equazione

$$y^{(4)} + 2y'' + y = 0.$$

Dall'equazione caratteristica $\lambda^4 + 2\lambda^2 + 1 = 0$ otteniamo le due radici doppie $\lambda_1 = i$, $\lambda_2 = -i$. Come sistema fondamentale avremo le seguenti funzioni

$$e^{it}, e^{-it}, t e^{it}, t e^{-it}.$$

In termini di funzioni reali, si ottiene il sistema equivalente

$$\cos t, \sin t, t \cos t, t \sin t,$$

da cui segue l'integrale generale

$$y(t) = (c_1 + c_2 t) \cos t + (c_3 + c_4 t) \sin t.$$

4.5 Equazioni non omogenee

Consideriamo ora una equazione lineare di ordine n non omogenea,

$$L(y) := y^{(n)} + p_1(t)y^{(n-1)} + \dots + p_n(t)y = f(t).$$

Sappiamo che l'integrale generale di questa equazione è dato dalla somma di una sua soluzione particolare e della soluzione dell'omogenea associata. Dobbiamo quindi trovare delle procedure per determinare una soluzione particolare dell'equazione $L(y) = f$.

Osserviamo innanzitutto che se il termine non omogeneo è dato dalla somma di più funzioni, cioè se $f(t) = \sum_{q=1}^r f_q(t)$, vale il seguente *principio di sovrapposizione* delle soluzioni, ovvero,

Se \hat{y}_q è una soluzione particolare dell'equazione $L(y) = f_q$ ($q = 1, \dots, r$), allora $\hat{y} = \sum_{q=1}^r \hat{y}_q$ è una soluzione particolare dell'equazione $L(y) = f = \sum_{q=1}^r f_q$.

Questo risultato segue dalla linearità dell'equazione in quanto si ha

$$L(\hat{y}) = L\left(\sum_{q=1}^r \hat{y}_q\right) = \sum_{q=1}^r L(\hat{y}_q) = \sum_{q=1}^r f_q = f.$$

Il principio di sovrapposizione consente di ricavare un integrale particolare di una equazione lineare non omogenea come somma di integrali particolari di equazioni contenenti singoli termini non omogenei in cui può essere decomposto il termine non omogeneo dell'equazione di partenza.

Detto questo, un metodo generale per la ricerca di un integrale particolare di una equazione lineare è quello della variazione delle costanti, già introdotto nel caso delle equazioni lineari del primo ordine. Ci limiteremo qui ad illustrare questo metodo nel caso di equazioni del secondo ordine in quanto, nella pratica, per equazioni di ordine superiore esso risulta poco maneggevole.

Supponiamo di aver determinato la soluzione

$$y = c_1 y_1 + c_2 y_2,$$

dell'omogenea associata all'equazione non omogenea

$$y'' + p_1(t)y' + p_2(t)y = f(t),$$

dove y_1 e y_2 sono funzioni linearmente indipendenti definite in un dominio D . Cerchiamo una soluzione particolare di questa equazione nella forma

$$\hat{y}(t) = c_1(t)y_1(t) + c_2(t)y_2(t).$$

Derivando otteniamo

$$\hat{y}' = c_1' y_1 + c_2' y_2 + c_1 y_1' + c_2 y_2'.$$

Poiché abbiamo a disposizione le due funzioni $c_1(t)$ e $c_2(t)$ per soddisfare una sola equazione, possiamo imporre una condizione supplementare su tali funzioni. Si può porre ad esempio

$$c_1' y_1 + c_2' y_2 = 0,$$

e di conseguenza, una ulteriore derivazione della \hat{y} fornisce

$$\hat{y}'' = c_1' y_1' + c_2' y_2' + c_1 y_1'' + c_2 y_2''.$$

Sostituendo nell'equazione di partenza, si ottiene

$$c_1' y_1' + c_2' y_2' + c_1 y_1'' + c_2 y_2'' + p_1(c_1 y_1' + c_2 y_2') + p_2(c_1 y_1 + c_2 y_2) = f.$$

Poiché y_1 e y_2 sono soluzioni dell'omogenea associata, da quest'ultima si ricava,

$$c_1' y_1' + c_2' y_2' = f.$$

Affinché \hat{y} sia una soluzione, le derivate delle funzioni c_1 e c_2 devono dunque soddisfare il sistema

$$\begin{cases} c_1' y_1 + c_2' y_2 = 0 \\ c_1' y_1' + c_2' y_2' = f. \end{cases}$$

Il determinante dei coefficienti di questo sistema è il wronskiano del sistema fondamentale di soluzioni. Esso, come sappiamo, è diverso da zero in tutto D , quindi il sistema ammette una sola soluzione data da

$$c_1'(t) = \frac{y_2(t)f(t)}{y_2(t)y_1'(t) - y_1(t)y_2'(t)}, \quad c_2'(t) = \frac{y_1(t)f(t)}{y_1(t)y_2'(t) - y_2(t)y_1'(t)}.$$

Integrando queste ultime e sostituendo nell'integrale generale della non omogenea, si ottiene

$$y(t) = c_1 y_1(t) + c_2 y_2(t) + y_1(t) \int \frac{y_2 f}{y_2 y_1' - y_1 y_2'} dt + y_2(t) \int \frac{y_1 f}{y_1 y_2' - y_2 y_1'} dt.$$

• Esempio

Determinare l'integrale generale dell'equazione

$$y'' + y = \frac{1}{\cos t}.$$

L'equazione omogenea associata ammette la soluzione

$$y = c_1 \cos t + c_2 \sin t.$$

Facendo variare c_1 e c_2 , prendendo cioè $c_1 = c_1(t)$ e $c_2 = c_2(t)$, abbiamo

$$y' = c_1' \cos t + c_2' \sin t - c_1 \sin t + c_2 \cos t.$$

Imponendo la condizione

$$c_1' \cos t + c_2' \sin t = 0,$$

e sostituendo nell'equazione non omogenea abbiamo

$$-c_1' \sin t + c_2' \cos t = \frac{1}{\cos t},$$

da cui si ricava il sistema

$$\begin{cases} c_1' \cos t + c_2' \sin t = 0 \\ -c_1' \sin t + c_2' \cos t = \frac{1}{\cos t}, \end{cases}$$

che ammette la soluzione $c_1' = -\tan t$, $c_2' = 1$. Una integrazione fornisce $c_1 = \ln |\cos t| + \alpha_1$, $c_2 = t + \alpha_2$, e dunque una soluzione particolare della non omogenea è

$$\hat{y}(t) = \ln |\cos t| \cos t + t \sin t.$$

L'integrale generale sarà allora,

$$y(t) = (c_1 + \ln |\cos t|) \cos t + (c_2 + t) \sin t.$$

In alcuni casi, la forma del termine non omogeneo $f(t)$ e l'uso del principio di sovrapposizione delle soluzioni suggeriscono una soluzione particolare dell'equazione non omogenea. Consideriamo i casi più elementari.

1) Supponiamo che $f(t) = e^{at}P_n(t)$ dove $a \in \mathbb{R}$ e dove P_n è un polinomio di grado n in t . Allora se a non è radice dell'equazione caratteristica dell'omogenea associata, una soluzione particolare \hat{y} è data da

$$\hat{y}(t) = e^{at}Q_n(t),$$

dove Q_n è un polinomio di grado n in t a coefficienti indeterminati.

Se invece a è una radice dell'equazione caratteristica dell'omogenea associata, si pone

$$\hat{y}(t) = t^p e^{at}Q_n(t),$$

dove p è la molteplicità della radice a .

2) Supponiamo che $f(t) = e^{at}[P_n(t) \cos bt + Q_m(t) \sin bt]$ con $a, b \in \mathbb{R}$ e P_n e Q_m polinomi di grado n ed m in t . Allora se $a \pm ib$ non è una coppia di radici dell'equazione caratteristica dell'omogenea associata, una soluzione particolare ha la forma

$$\hat{y}(t) = e^{at}[R_q(t) \cos bt + S_q(t) \sin bt],$$

dove R_q e S_q sono polinomi di grado $q = \max\{n, m\}$.

Se invece $a \pm ib$ è una coppia di soluzioni dell'equazione caratteristica dell'omogenea associata, si pone

$$\hat{y}(t) = t^p e^{at}[R_q(t) \cos bt + S_q(t) \sin bt],$$

dove p è la molteplicità delle coppie di radici $a \pm ib$.

- Esempi

1) Determinare l'integrale generale dell'equazione

$$y''' - y' = (2t + 1)e^{2t}.$$

L'equazione omogenea associata $y''' - y' = 0$ ammette l'integrale generale

$$\bar{y} = c_1 + c_2 e^t + c_3 e^{-t}.$$

Cerchiamo una soluzione particolare dell'equazione non omogenea nella forma

$$\hat{y} = (At + B)e^{2t}.$$

Derivando otteniamo

$$\hat{y}' = (A + 2B + 2At)e^{2t}, \quad \hat{y}'' = (4A + 4B + 4At)e^{2t}, \quad \hat{y}''' = (12A + 8B + 8At)e^{2t}.$$

Sostituendo si trova

$$11A + 6B + 6At = 2t + 1 \quad \Rightarrow \quad A = \frac{1}{3}, \quad B = -\frac{4}{9}.$$

L'integrale generale sarà dunque

$$y(t) = \bar{y}(t) + \hat{y}(t) = c_1 + c_2 e^t + c_3 e^{-t} + \left(\frac{1}{3}t - \frac{4}{9}\right)e^{2t}.$$

2) Determinare l'integrale generale dell'equazione

$$y'' - y = (t + 1)e^t.$$

L'integrale generale dell'omogenea associata $y'' - y = 0$ è dato da

$$\bar{y} = c_1 e^t + c_2 e^{-t}.$$

Poiché l'esponenziale del termine non omogeneo compare nella soluzione della omogenea associata dobbiamo cercare una soluzione della non omogenea nella forma di un polinomio moltiplicato per te^t , ovvero

$$\hat{y} = (At + B)te^t.$$

Derivando otteniamo

$$\hat{y}' = [B + (2A + B)t + At^2]e^t, \quad \hat{y}'' = [2A + 2B + (4A + B)t + At^2]e^t.$$

Sostituendo ricaviamo

$$2A + 2B + 4At = t + 1, \quad \Rightarrow \quad A = \frac{1}{4}, \quad B = \frac{1}{4}.$$

L'integrale generale sarà allora

$$y(t) = \bar{y}(t) + \hat{y}(t) = c_1 e^t + c_2 e^{-t} + \frac{1}{4}(t + 1)te^t.$$

3) Determinare l'integrale generale della equazione

$$y'' + 4y = t \sin 2t.$$

L'omogenea associata $y'' + 4y = 0$ ha l'integrale

$$\bar{y} = c_1 \sin 2t + c_2 \cos 2t.$$

Anche in questo caso il termine non omogeneo compare nella soluzione della omogenea associata. Cercheremo allora la soluzione particolare della non omogenea nella forma

$$\hat{y} = (At + C)t \sin 2t + (Bt + D)t \cos 2t.$$

Derivando otteniamo

$$\begin{aligned} \hat{y}' &= [C + 2(A - D)t - 2Bt^2] \sin 2t + [D + 2(B + C)t + 2At^2] \cos 2t. \\ \hat{y}'' &= 4 \left[\frac{A}{2} - D - (C + 2B)t - At^2 \right] \sin 2t + 4 \left[\frac{B}{2} + C - (D - 2A)t - Bt^2 \right] \cos 2t. \end{aligned}$$

Sostituendo ricaviamo

$$\begin{aligned} 2(A - D - 4Bt) \sin 2t + 2(B + C + 4At) \cos 2t &= t \sin 2t \\ \Rightarrow \quad A = D = 0, \quad -B = C = \frac{1}{8}. \end{aligned}$$

4.6 Sistemi di equazioni differenziali lineari

Un sistema di equazioni differenziali del primo ordine $\mathbf{y}' = \mathbf{f}(t, \mathbf{y})$ si dice lineare se \mathbf{f} è una funzione lineare di \mathbf{y} ovvero se

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{b},$$

dove \mathbf{A} è un operatore lineare in \mathbb{R}^n , in generale dipendente da $t \in D$, che agisce sulla funzione vettoriale $\mathbf{y} : D \rightarrow \mathbb{R}^n$ a componenti con derivate prime continue e dove \mathbf{b} è una funzione vettoriale assegnata di $t \in D$. Il sistema si può scrivere anche nella forma

$$L(\mathbf{y}) = \mathbf{b},$$

dove L è l'applicazione definita da $D(\mathbf{y}) - \mathbf{A}\mathbf{y}$, essendo D l'applicazione di derivazione sulla funzione vettoriale \mathbf{y} . Per la linearità di quest'ultima, l'applicazione L risulta lineare. Ne segue che, come nel caso delle equazioni lineari, se $\bar{\mathbf{y}}$ rappresenta l'integrale del sistema omogeneo associato $L(\mathbf{y}) = \mathbf{0}$, e se $\hat{\mathbf{y}}$ è una soluzione particolare del sistema non omogeneo, l'integrale generale del sistema non omogeneo è dato da

$$\mathbf{y} = \bar{\mathbf{y}} + \hat{\mathbf{y}}.$$

Teorema 4.3 *L'integrale generale del sistema omogeneo $L(\mathbf{y}) = \mathbf{0}$ è dato da tutte e sole le combinazioni lineari di n soluzioni particolari linearmente indipendenti $\mathbf{y}_{(k)}$ ($k = 1, \dots, n$).*

Dim. Consideriamo i seguenti n problemi di Cauchy

$$L(\mathbf{y}) = \mathbf{0}, \quad \mathbf{y}(t_0) = \mathbf{e}_k, \quad k = 1, \dots, n,$$

dove \mathbf{e}_k sono i versori della base canonica in \mathbb{R}^n , ovvero $\mathbf{e}_1 = (1, 0, 0, \dots, 0)$, $\mathbf{e}_2 = (0, 1, 0, \dots, 0)$, ..., $\mathbf{e}_n = (0, 0, \dots, 1)$. Chiamiamo $\mathbf{y}_{(k)}(t)$ le soluzioni di questi problemi per t appartenente ad un intorno di t_0 . Esse costituiscono n soluzioni particolari del sistema omogeneo. Si verifica subito che la combinazione lineare

$$\mathbf{y} = \sum_{k=1}^n c_k \mathbf{y}_{(k)},$$

è soluzione del problema di Cauchy

$$L(\mathbf{y}) = \mathbf{0}, \quad \mathbf{y}(t_0) = \sum_{k=1}^n c_k \mathbf{e}_k.$$

Affinché tale soluzione sia quella identicamente nulla è sufficiente e necessario che tutti i coefficienti c_k siano nulli. Ne segue che le funzioni $\mathbf{y}_{(k)}$ sono linearmente indipendenti.

D'altra parte, una qualunque soluzione dell'equazione $L(\mathbf{y}) = \mathbf{0}$ avente condizioni iniziali $\mathbf{y}(t_0) = \mathbf{y}^{(0)}$ si può scrivere nella forma

$$\mathbf{y} = \sum_{k=1}^n \alpha_k \mathbf{y}_{(k)}.$$

Infatti, basta scegliere $\alpha_k = y_k(t_0)$ affinché valga

$$\mathbf{y}(t_0) = \sum_{k=1}^n y_k(t_0) \mathbf{y}_{(k)}(t_0) = \sum_{k=1}^n y_k(t_0) \mathbf{e}_k = \mathbf{y}^{(0)}.$$

che corrisponde alla condizione iniziale data.

OSSERVAZIONE. Poiché una equazione differenziale lineare di ordine n in forma normale è sempre riconducibile a un sistema di n equazioni lineari del primo ordine in n funzioni incognite, il teorema 4.2 può essere ricavato dal teorema 4.3. Vale la pena esaminare in dettaglio questa implicazione.

Se nell'equazione lineare omogenea

$$u^{(n)} + p_1(t)u^{(n-1)} + \dots + p_{n-1}(t)u = 0,$$

si pone $u = y_1$, $u' = y_2$, ..., $u^{(n-1)} = y_n$, si ottiene il sistema equivalente

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix}' = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \vdots & \vdots & \vdots & \dots & 0 & 1 \\ -p_{n-1} & -p_{n-2} & -p_{n-3} & \dots & -p_1 & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{n-1} \\ y_n \end{pmatrix}.$$

Dette $\mathbf{y}_{(k)}(t)$, $t \in D$, ($k = 1, 2, \dots, n$), n soluzioni linearmente indipendenti di tale sistema, soddisfacenti le condizioni iniziali $\mathbf{y}_{(k)}(t_0) = \mathbf{y}_{(k)}^0$, ad esse corrispondono le n soluzioni $u_k = y_{(k)1}$, ($k = 1, 2, \dots, n$) dell'equazione di partenza, soddisfacenti le condizioni

$$u_k(t_0) = y_{(k)1}^0, \quad u_k'(t_0) = y_{(k)2}^0, \quad \dots \quad u_k^{(n-1)}(t_0) = y_{(k)n}^0, \quad k = 1, 2, \dots, n.$$

Osserviamo che il wronskiano $W(t)$ delle u_k non è altro che il determinante della matrice che ha per colonne i vettori $\mathbf{y}_{(k)}(t)$. L'indipendenza lineare delle $\mathbf{y}_{(k)}$ comporta quindi che W sia diverso da zero in tutto D e di conseguenza le soluzioni u_k risultano linearmente indipendenti.

Infine, ogni soluzione u soddisfacente le condizioni iniziali

$$u^{(k-1)}(t_0) = \eta_k, \quad k = 1, 2, \dots, n,$$

è la corrispondente dell'unica soluzione del sistema di equazioni differenziali nelle y_i , soddisfacente le condizioni iniziali

$$y_i(t_0) = \eta_i, \quad i = 1, 2, \dots, n.$$

Poiché quest'ultima è una combinazione lineare delle n soluzioni $\mathbf{y}_{(k)}$, la soluzione u sarà una combinazione lineare delle u_k .

Consideriamo ora il caso di un sistema di n equazioni differenziali lineari a coefficienti costanti

$$\mathbf{y}' = \mathbf{A}\mathbf{y} + \mathbf{b}$$

dove \mathbf{A} è un operatore in \mathbb{R}^n indipendente da t . Ovviamente, anche in questo caso la soluzione si trova aggiungendo all'integrale generale del sistema omogeneo associato una soluzione particolare di quello non omogeneo. Studiamo dunque in dettaglio l'equazione omogenea.

In forma matriciale il sistema di equazioni differenziali si scrive

$$Y' = AY,$$

dove Y è il vettore colonna delle componenti del vettore $\mathbf{y}(t)$ in una base data e dove A è la matrice corrispondente all'operatore \mathbf{A} .

Se la matrice A è diagonalizzabile, il sistema può essere trasformato in n equazioni differenziali del primo ordine nelle singole componenti $z_i(t)$ di una funzione incognita $\mathbf{z}(t)$, legate alle componenti y_i dalla trasformazione che diagonalizza A . Più precisamente supponiamo che A ammetta n autovalori

distinti λ_i , ($i = 1, 2, \dots, n$), cui corrispondono gli n vettori linearmente indipendenti $X^{(i)}$. Costruita la matrice invertibile Q che ha per colonne gli autovettori $X^{(i)}$, operiamo la trasformazione che diagonalizza A ,

$$Q^{-1}AQ = \Lambda = \text{diag}(\lambda_1, \lambda_2, \dots, \lambda_n), \quad Y = QZ,$$

dove Z è il vettore colonna formato dalle componenti del vettore \mathbf{y} nella nuova base. Poiché Q è una matrice di costanti, si ha $Y' = QZ'$ e il sistema di partenza diviene

$$QZ' = AQZ \implies Z' = Q^{-1}AQZ = \Lambda Z.$$

Si ottiene così, esplicitamente

$$\begin{aligned} z_1'(t) &= \lambda_1 z_1(t), \\ z_2'(t) &= \lambda_2 z_2(t), \\ &\vdots \\ z_n'(t) &= \lambda_n z_n(t). \end{aligned}$$

Le equazioni così ottenute si dicono *disaccoppiate* perché ciascuna può essere risolta indipendentemente dalle altre. Risolvendole si ha

$$z_i(t) = c_i e^{\lambda_i t}, \quad i = 1, 2, \dots, n.$$

Da queste soluzioni si ricavano poi le $y_i(t)$ mediante la trasformazione $Y = QZ$, ovvero

$$y_i(t) = \sum_{j=1}^n Q_{ij} c_j e^{\lambda_j t} = \sum_{j=1}^n X_i^{(j)} c_j e^{\lambda_j t}.$$

In altri termini le $y_i(t)$ sono combinazioni lineari delle funzioni $e^{\lambda_j t}$. Dato che gli autovalori λ_i non dipendono dalla base scelta e osservando che i vettori $X_i^{(j)} c_j$ altro non sono che le n -uple di componenti degli autovettori $\mathbf{k}^{(j)}$ di \mathbf{A} nella base data, la soluzione del sistema di partenza si scrive come

$$\mathbf{y}(t) = \sum_{i=1}^n \mathbf{k}^{(i)} e^{\lambda_i t}.$$

Le singole funzioni $\mathbf{k}^{(i)} e^{\lambda_i t}$ si dicono anche *autofunzioni* del sistema di equazioni differenziali $\mathbf{y}' = \mathbf{A}\mathbf{y}$.

• Esempio

Risolvere il sistema di equazioni differenziali lineari omogenee

$$\begin{cases} y_1' = y_1 + 2y_2 \\ y_2' = 4y_1 + 3y_2 \end{cases}$$

Si ha

$$A = \begin{pmatrix} 1 & 2 \\ 4 & 3 \end{pmatrix} \implies \det(A - \lambda I) = \lambda^2 - 4\lambda - 5.$$

Si ottiene così $\lambda_1 = -1$, $\lambda_2 = 5$. Gli autovettori corrispondenti sono

$$K^{(1)} = c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix}, \quad K^{(2)} = c_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix}.$$

L'integrale generale del sistema sarà

$$Y = c_1 \begin{pmatrix} 1 \\ -1 \end{pmatrix} e^{-t} + c_2 \begin{pmatrix} 1 \\ 2 \end{pmatrix} e^{5t} \implies \begin{cases} y_1 = c_1 e^{-t} + c_2 e^{5t} \\ y_2 = -c_1 e^{-t} + 2c_2 e^{5t} \end{cases}.$$

Se il polinomio caratteristico ammette una radice complessa, esso ammetterà anche la sua coniugata e i corrispondenti autovettori saranno coniugati. Infatti, essendo A reale, se $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$, ne segue $\mathbf{A}\bar{\mathbf{x}} = \bar{\lambda}\bar{\mathbf{x}}$. Posto $\lambda = a + ib$, le corrispondenti autofunzioni saranno

$$\mathbf{y} = \mathbf{x}e^{(a+ib)t}, \quad \bar{\mathbf{y}} = \bar{\mathbf{x}}e^{(a-ib)t}.$$

queste ultime si possono sempre esprimere in termini reali, sostituendole con le combinazioni

$$\begin{aligned} \frac{1}{2}(\mathbf{y} + \bar{\mathbf{y}}) &= \Re(\mathbf{y}) = e^{at}(\mathbf{x}_a \cos bt - \mathbf{x}_b \sin bt), \\ \frac{1}{2i}(\mathbf{y} - \bar{\mathbf{y}}) &= \Im(\mathbf{y}) = e^{at}(\mathbf{x}_a \sin bt + \mathbf{x}_b \cos bt), \end{aligned}$$

dove si è posto $\mathbf{x} = \mathbf{x}_a + i\mathbf{x}_b$.

• Esempio

Determinare l'integrale generale del sistema di equazioni omogenee

$$\begin{cases} y_1' = y_1 - 5y_2 \\ y_2' = 2y_1 - y_2 \end{cases}$$

Si ha

$$A = \begin{pmatrix} 1 & -5 \\ 2 & -1 \end{pmatrix} \Rightarrow \lambda_1 = 3i, \quad \lambda_2 = -3i.$$

I corrispondenti autovettori sono

$$X^{(1)} = \begin{pmatrix} 1 + 3i \\ 2 \end{pmatrix}, \quad X^{(2)} = \begin{pmatrix} 1 - 3i \\ 2 \end{pmatrix}.$$

Si ottengono così le due soluzioni reali

$$\begin{aligned} Y_A &= \begin{pmatrix} 1 \\ 2 \end{pmatrix} \cos 3t - \begin{pmatrix} 3 \\ 0 \end{pmatrix} \sin 3t \\ Y_B &= \begin{pmatrix} 1 \\ 2 \end{pmatrix} \sin 3t + \begin{pmatrix} 3 \\ 0 \end{pmatrix} \cos 3t, \end{aligned}$$

Prendendo una loro combinazione lineare otteniamo l'integrale generale

$$\begin{aligned} y_1 &= (c_1 + 3c_2) \cos 3t - (3c_1 - c_2) \sin 3t \\ y_2 &= 2c_1 \cos 3t + 2c_2 \sin 3t. \end{aligned}$$

Il procedimento indicato per la soluzione dei sistemi omogenei non cambia se la matrice A ammette autovalori con molteplicità maggiore di 1, purchè sia sempre diagonalizzabile. Se, per esempio, l'autovalore λ ha molteplicità m e a tale autovalore corrispondono esattamente m autovettori $\hat{X}^{(j)}$, ($j = 1, 2, \dots, m$) linearmente indipendenti, si ricava

$$y_i(t) = \sum_{j=1}^m \hat{X}_i^{(j)} \hat{c}_j e^{\lambda t} + \sum_{j=1}^{n-m} X_i^{(j)} c_j e^{\lambda_j t},$$

da cui

$$\mathbf{y}(t) = e^{\lambda t} \sum_{j=1}^m \hat{\mathbf{k}}^{(j)} + \sum_{j=1}^{n-m} \mathbf{k}^{(j)} e^{\lambda_j t},$$

dove $\mathbf{k}^{(j)}$ sono gli autovettori corrispondenti agli autovalori λ_j con molteplicità 1.

• Esempio

Determinare l'integrale generale del sistema di equazioni differenziali

$$\begin{aligned} y_1' &= 6y_1 + 2y_2 + 2y_3 \\ y_2' &= 2y_1 + 3y_2 - 4y_3 \\ y_3' &= 2y_1 - 4y_2 + 3y_3 \end{aligned}$$

Si ha

$$A = \begin{pmatrix} 6 & 2 & 2 \\ 2 & 3 & -4 \\ 2 & -4 & 3 \end{pmatrix} \implies \lambda_1 = 7 \text{ (2volte)}, \quad \lambda_2 = -2$$

Poiché la matrice A è simmetrica, essa ammette 3 autovettori linearmente indipendenti. In particolare si ha, per $\lambda = \lambda_1$,

$$X^{(1)} = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}, \quad X^{(2)} = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix}$$

mentre, per $\lambda = \lambda_2$,

$$X^{(3)} = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}.$$

La soluzione del sistema sarà allora

$$\mathbf{y} = \left[c_1 \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right] e^{7t} + c_3 \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix} e^{-2t}.$$

Supponiamo ora che la matrice A non ammetta n autovettori linearmente indipendenti e non sia quindi diagonalizzabile. Dall'algebra sappiamo che in questo caso è sempre possibile ridurre A alla forma canonica di Jordan mediante una opportuna trasformazione di base. Indicata con Q la matrice di trasformazione, costruita con n autovettori generalizzati di A , si ha

$$Q^{-1}AQ = J = \text{diag}(J_1, \dots, J_p), \quad Y = QZ,$$

dove J_i ($i = 1, 2, \dots, p$) sono blocchi di Jordan di ordine n_i e dove $\sum_{i=1}^p n_i = n$. Il sistema lineare $Y' = AY$ diventa

$$QZ' = AQZ \implies Z' = Q^{-1}AQZ = JZ.$$

Quest'ultimo sistema si spezza in p sottosistemi disaccoppiati

$$Z_i' = J_i Z_i, \quad i = 1, 2, \dots, p,$$

ciascuno nelle n_i -uple di funzioni incognite $\{z_{i1}(t), z_{i2}(t), \dots, z_{in_i}(t)\}$. Esplicitamente, fissato i si ha

$$\begin{aligned} z'_{i1}(t) &= \lambda_i z_{i1}(t) + z_{i2}(t), \\ z'_{i2}(t) &= \lambda_i z_{i2}(t) + z_{i3}(t), \\ &\vdots \\ z'_{in_i-1}(t) &= \lambda_i z_{in_i-1}(t) + z_{in_i}(t), \\ z'_{in_i}(t) &= \lambda_i z_{in_i}(t). \end{aligned}$$

Dall'ultima equazione ricaviamo

$$z_{in_i}(t) = c_{n_i} e^{\lambda_i t},$$

e sostituendo nella precedente otteniamo l'equazione lineare non omogenea

$$z'_{in_i-1}(t) = \lambda_i z_{in_i-1}(t) + c_{n_i} e^{\lambda_i t},$$

la cui soluzione è

$$z_{in_i-1}(t) = (c_{n_i-1} + c_{n_i} t) e^{\lambda_i t}.$$

Sostituendo ancora nell'equazione precedente e continuando così fino a risolvere la prima equazione, si ottiene

$$\begin{aligned} z_{in_i-2}(t) &= (c_{n_i-2} + c_{n_i-1} t + \frac{1}{2} c_{n_i} t^2) e^{\lambda_i t}, \\ &\vdots \\ z_{i1}(t) &= (c_1 + c_2 t + \frac{1}{2} c_3 t^2 + \dots + \frac{1}{(n_i-1)!} c_{n_i} t^{n_i-1}) e^{\lambda_i t}. \end{aligned}$$

In generale si può dunque scrivere la soluzione dell'intero sistema nella forma

$$z_{ij}(t) = e^{\lambda_i t} P_{n_i-j}(t), \quad j = 1, \dots, n_i, \quad i = 1, \dots, p,$$

dove $P_{n_i-j}(t)$ è un polinomio in t di grado $n_i - j$. Osservando che $Z = (z_{11}, \dots, z_{1n_1}, z_{21}, \dots, z_{2n_2}, \dots, z_{p1}, \dots, z_{pn_p})^T$ si ottiene infine la soluzione del sistema di partenza dalla relazione $Y = QZ$.

• Esempio

Determinare l'integrale generale del sistema differenziale lineare omogeneo

$$\begin{aligned} y'_1 &= -y_1 + 2y_3, \\ y'_2 &= 2y_1 + y_2 - 2y_3, \\ y'_3 &= -2y_1 + 3y_3. \end{aligned}$$

Si ha $Y' = AY$ con

$$A = \begin{pmatrix} -1 & 0 & 2 \\ 2 & 1 & -2 \\ -2 & 0 & 3 \end{pmatrix},$$

da cui si ricava $\lambda_1 = 1$ (3 volte). All'unico autovalore corrispondono i due autovettori linearmente indipendenti

$$X^{(1)} = (1, 0, 1)^T, \quad X^{(2)} = (1, 1, 1)^T.$$

La matrice non è diagonalizzabile e quindi cerchiamo gli autovettori generalizzati di A relativi a λ_1 . Posto

$$(A - I)X \neq 0, \quad (A - I)^2 X = 0,$$

troviamo un primo autovettore generalizzato,

$$X = (0, 1, 1)^T,$$

da cui si ottiene il secondo autovettore generalizzato

$$\tilde{X} = (A - I)X = (2, -2, 2)^T.$$

Abbiamo così due autovettori generalizzati X, \tilde{X} linearmente indipendenti e possiamo scegliere come terzo autovettore, $X^{(1)}$. Si ottiene così

$$Q = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 1 & 1 & 2 \end{pmatrix}, \quad Q^{-1} = \begin{pmatrix} 2 & 1 & -1 \\ -1 & 0 & 1 \\ -1/2 & -1/2 & 1/2 \end{pmatrix},$$

e quindi

$$J = Q^{-1}AQ = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}.$$

Il sistema $Z' = JZ$ è dato esplicitamente da

$$\begin{aligned} z_1'(t) &= z_1(t) \\ z_2'(t) &= z_2(t) \\ z_3'(t) &= z_3(t) + z_2(t), \end{aligned}$$

da cui ricaviamo

$$z_1(t) = c_1 e^t, \quad z_2(t) = c_2 e^t, \quad z_3(t) = c_3 e^t + c_2 t e^t.$$

Ritornando alle funzioni incognite $y_i(t)$, abbiamo

$$Y = QZ = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 1 & 1 & 2 \end{pmatrix} \begin{pmatrix} c_1 e^t \\ c_2 e^t \\ (c_3 + c_2 t) e^t \end{pmatrix},$$

ovvero,

$$\begin{aligned} y_1(t) &= (c_1 + 2c_3 + 2c_2 t) e^t \\ y_2(t) &= (c_2 - 2c_3 - 2c_2 t) e^t \\ y_3(t) &= (c_1 + c_2 + 2c_3 + 2c_2 t) e^t. \end{aligned}$$

Nel caso di un sistema lineare non omogeneo si può cercare una soluzione particolare facendo ricorso sempre al metodo della variazione delle costanti. Illustriamo il metodo in questo caso mediante un esempio.

- Esempio

Determinare l'integrale generale del sistema non omogeneo

$$\begin{cases} y_1' = y_2 \\ y_2' = -y_1 + \frac{1}{\cos t}. \end{cases}$$

Il sistema omogeneo associato ammette l'integrale generale

$$\begin{aligned} y_1 &= c_1 \cos t + c_2 \sin t \\ y_2 &= -c_1 \sin t + c_2 \cos t. \end{aligned}$$

Facendo variare le costanti abbiamo

$$\begin{aligned} y_1 &= c_1(t) \cos t + c_2(t) \sin t \\ y_2 &= -c_1(t) \sin t + c_2(t) \cos t, \end{aligned}$$

e sostituendo nel sistema di partenza otteniamo

$$\begin{aligned} c_1' \cos t + c_2' \sin t &= 0 \\ -c_1' \sin t + c_2' \cos t &= \frac{1}{\cos t}. \end{aligned}$$

La soluzione di questo sistema è

$$c_1' = -\tan t, \quad c_2' = 1,$$

da cui si ricava

$$c_1 = \ln |\cos t| + \alpha_1, \quad c_2 = t + \alpha_2.$$

L'integrale generale del sistema non omogeneo risulta

$$\begin{aligned} y_1 &= \alpha_1 \cos t + \alpha_2 \sin t + \cos t \ln |\cos t| + t \sin t \\ y_2 &= -\alpha_1 \sin t + \alpha_2 \cos t - \sin t \ln |\cos t| + t \cos t. \end{aligned}$$

4.7 Alcune applicazioni delle equazioni differenziali

Consideriamo qui alcuni esempi di sistemi fisici il cui studio conduce alla risoluzione di un problema di Cauchy.

4.7.1 Probabilità di collisione

Consideriamo uno strato materiale a facce piane di spessore L e area A . Sia n il numero di molecole del materiale per unità di volume e supponiamo di poter rappresentare ciascuna molecola come una sfera di raggio R . Lo strato viene bombardato con molecole-proiettile anch'esse di forma sferica e raggio r . Una molecola-proiettile urterà una molecola del bersaglio se si avvicinerà a quest'ultima ad una distanza minore di $R + r$. La sezione d'urto di collisione per una molecola sarà allora $\sigma = \pi(R + r)^2$. Detta $x \in [0, L]$ la coordinata trasversale allo strato lungo cui incidono le molecole-proiettile, il numero

di molecole presenti nell'elemento di strato compreso tra x_0 e $x_0 + dx$ sarà dato da $nAdx$, mentre la sezione d'urto relativa allo stesso elemento sarà $S = \sigma nAdx$. Se si divide tale sezione d'urto per l'area A si trova la probabilità P che, su N proiettili incidenti, un numero pari a $|dN|$ urti le molecole del suddetto elemento di strato. In altri termini si ha

$$\left| \frac{dN}{N} \right| = P = \frac{\sigma nAdx}{A} = \sigma ndx.$$

Se N_0 è il numero di molecole incidenti sulla superficie dello strato (cioè per $x = 0$), il numero di molecole $N(x)$ che riescono a giungere alla profondità x senza collidere soddisferà il seguente problema di Cauchy del primo ordine,

$$\begin{cases} \frac{dN(x)}{dx} = -n\sigma N(x), \\ N(0) = N_0. \end{cases}$$

dove il segno meno è dovuto al fatto che dN rappresenta la perdita di proiettili che attraversano lo strato senza collidere, ed è quindi una quantità negativa. Integrando l'equazione differenziale si ha

$$\ln N(x) = -n\sigma x \implies N(x) = k e^{-n\sigma x},$$

e imponendo la condizione iniziale, si ottiene

$$N(x) = N_0 e^{-n\sigma x}.$$

Il numero di molecole-proiettile che attraversa uno spessore x ($0 < x < L$) decresce quindi esponenzialmente con x . La quantità

$$l = \frac{1}{n\sigma}$$

ha le dimensioni di una lunghezza ed è una misura della distanza percorsa in media da un proiettile senza collidere. Per tale motivo essa viene chiamata *libero cammino medio* del processo di collisione. In particolare il numero di proiettili che emergono dallo strato senza aver subito collisioni sarà

$$N = N_0 e^{-L/l}.$$

Nei problemi che seguono le funzioni incognite saranno funzioni f del tempo ed indicheremo con \dot{f}, \ddot{f}, \dots , le derivate prima, seconda, etc., di f rispetto a t .

4.7.2 Moto di un punto soggetto a forza resistente

Sia \mathbf{x} la posizione nello spazio di un punto materiale di massa m e sia $\mathbf{v} = \dot{\mathbf{x}}$ la sua velocità. Supponiamo che il punto sia soggetto ad una forza resistente $\mathbf{F}_R = -h\|\mathbf{v}\|\mathbf{v}$ dove $h \in \mathbb{R}^{++}$, e alla forza peso $m\mathbf{g}$. Assumiamo inoltre che all'istante iniziale $t = 0$ si abbia $\mathbf{x}(0) = \mathbf{x}^0$ e $\mathbf{v}(0) = \dot{\mathbf{x}}(0) = \mathbf{0}$. Vogliamo determinare il moto del punto.

L'equazione del moto $m\mathbf{a} = \mathbf{F}$, si scrive

$$m\ddot{\mathbf{x}} = -h\|\dot{\mathbf{x}}\|\dot{\mathbf{x}} + m\mathbf{g}.$$

Dette x_1, x_2, x_3 le componenti di \mathbf{x} rispetto ad una terna cartesiana ortogonale di versori $\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3$ e prendendo $\mathbf{g} = -g\mathbf{e}_3$, con g accelerazione di gravità, si ottiene il sistema di equazioni differenziali

$$\begin{cases} m\ddot{x}_1 = -h\|\dot{\mathbf{x}}\|\dot{x}_1 \\ m\ddot{x}_2 = -h\|\dot{\mathbf{x}}\|\dot{x}_2 \\ m\ddot{x}_3 = -h\|\dot{\mathbf{x}}\|\dot{x}_3 - mg, \end{cases}$$

con le condizioni iniziali $x_1(0) = x_1^0, x_2(0) = x_2^0, x_3(0) = x_3^0, \dot{x}_1(0) = \dot{x}_2(0) = \dot{x}_3(0) = 0$. Notiamo che le prime due equazioni ammettono certamente la soluzione $x_1(t) = x_1^0, x_2(t) = x_2^0$ essendo verificate le condizioni iniziali. Per l'unicità della soluzione, il punto materiale manterrà, durante il moto, sempre le stesse coordinate sul piano x_1, x_2 . Il moto avverrà dunque lungo l'asse x_3 , secondo l'equazione

$$m\ddot{x}_3 = -h|\dot{x}_3|\dot{x}_3 - mg.$$

Da questa equazione, facendo uso delle condizioni iniziali, si deduce che $\ddot{x}_3(0) = -g, \dot{x}_3(0) = 0$. Allora, se x_3 è continua, essa ha un massimo in $t = 0$ e quindi, in un intorno $I = (0, \delta)$ si avrà $\dot{x}_3 < 0$. Ne segue che in I l'equazione diventa

$$m\ddot{x}_3 = h\dot{x}_3^2 - mg.$$

Dividendo per m e ponendo $u = \dot{x}_3, \alpha = h/m$, si ottiene

$$\begin{cases} \dot{u} - \alpha u^2 = -g \\ u(0) = 0. \end{cases}$$

Risolviamo l'equazione di questo problema separando le variabili

$$-\int \frac{du}{g - \alpha u^2} = t,$$

da cui,

$$\sqrt{\alpha g} t = \tanh^{-1} \left(-\sqrt{\frac{\alpha}{g}} u \right) + c.$$

Imponendo la condizione iniziale $u(0) = 0$, si ha $c = 0$ e ricavando u ,

$$u(t) = \dot{x}_3(t) = -\sqrt{\frac{g}{\alpha}} \tanh(\sqrt{\alpha g} t).$$

Poiché da questa soluzione \dot{x}_3 risulta minore di zero per ogni t , la soluzione del problema può essere estesa per qualunque $t > 0$. Integrando ulteriormente e tenendo conto della condizione $x_3(0) = x_3^0$, abbiamo

$$x_3(t) = x_3^0 - \frac{1}{\alpha} \ln[\cosh(\sqrt{\alpha g} t)].$$

Osserviamo che il moto del punto è tale che per $t \rightarrow \infty$ si ha $\dot{x}_3 \rightarrow -\sqrt{\frac{g}{\alpha}}$, cioè, asintoticamente, il punto assume velocità costante pari a $\sqrt{\frac{mg}{h}}$.

4.7.3 Oscillatore armonico forzato

Supponiamo che un punto materiale di massa m possa scorrere lungo una guida orizzontale priva di attrito. Detta x l'ascissa del punto sulla guida, rispetto ad una origine prefissata O , applichiamo al punto una forza elastica $-kx$ che tende a richiamarlo verso O ed una forza esterna periodica $F(t) = F \cos(\gamma t + \beta)$. Determinare il moto del punto sapendo che $x(0) = x_0$ e che $\dot{x}(0) = v_0$.

L'equazione del moto è

$$m\ddot{x} = -kx + F \cos(\gamma t + \beta).$$

Dividendo per m e ponendo $k/m = \omega^2, F/m = f$, si ottiene il seguente problema di Cauchy

$$\begin{cases} \ddot{x} + \omega^2 x = f \cos(\gamma t + \beta) \\ x(0) = x_0, \dot{x}(0) = v_0. \end{cases}$$

L'omogenea associata $\ddot{x} + \omega^2 x = 0$ ammette l'integrale generale

$$\bar{x}(t) = c_1 \cos \omega t + c_2 \sin \omega t.$$

Supposto $\gamma \neq \omega$, cerchiamo ora una soluzione particolare dell'equazione non omogenea nella forma

$$\hat{x}(t) = b \cos(\gamma t + \beta).$$

Sostituendo nell'equazione si ottiene

$$-b\gamma^2 \cos(\gamma t + \beta) + \omega^2 b \cos(\gamma t + \beta) = f \cos(\gamma t + \beta),$$

che vale identicamente se

$$b = \frac{f}{\omega^2 - \gamma^2}.$$

In tal caso otteniamo l'integrale generale nella forma

$$x(t) = c_1 \cos \omega t + c_2 \sin \omega t + \frac{f}{\omega^2 - \gamma^2} \cos(\gamma t + \beta).$$

Derivando si ha anche

$$\dot{x}(t) = -\omega c_1 \sin \omega t + \omega c_2 \cos \omega t - \frac{f\gamma}{\omega^2 - \gamma^2} \sin(\gamma t + \beta).$$

Imponendo le condizioni iniziali su x e \dot{x} si ha

$$\begin{aligned} x_0 &= c_1 + \frac{f}{\omega^2 - \gamma^2} \cos \beta, \\ v_0 &= \omega c_2 - \frac{f\gamma}{\omega^2 - \gamma^2} \sin \beta. \end{aligned}$$

da cui si ottiene

$$c_1 = x_0 - \frac{f}{\omega^2 - \gamma^2} \cos \beta, \quad c_2 = \frac{1}{\omega} \left(v_0 + \frac{f\gamma}{\omega^2 - \gamma^2} \sin \beta \right).$$

La soluzione del problema sarà dunque

$$x(t) = \left(x_0 - \frac{f \cos \beta}{\omega^2 - \gamma^2} \right) \cos \omega t + \left(\frac{v_0}{\omega} + \frac{(f\gamma/\omega) \sin \beta}{\omega^2 - \gamma^2} \right) \sin \omega t + \frac{f \cos(\gamma t + \beta)}{\omega^2 - \gamma^2}.$$

Nel caso in cui $\gamma = \omega$ la soluzione particolare precedentemente trovata non vale più e si deve cercare un integrale particolare nella forma

$$\hat{x}(t) = at \sin(\omega t + \beta).$$

Sostituendo si ha

$$2a\omega \cos(\omega t + \beta) = f \cos(\omega t + \beta) \quad \Rightarrow \quad a = \frac{f}{2\omega}.$$

L'integrale generale sarà allora

$$x(t) = c_1 \cos \omega t + c_2 \sin \omega t + \frac{f}{2\omega} t \sin(\omega t + \beta).$$

Sfruttando le condizioni iniziali si ricava

$$x_0 = c_1, \quad v_0 = c_2 \omega + \frac{f}{2\omega} \sin \beta,$$

da cui

$$c_1 = x_0, \quad c_2 = \frac{1}{\omega} \left(v_0 - \frac{f}{2\omega} \sin \beta \right).$$

La soluzione del problema sarà dunque

$$x(t) = x_0 \cos \omega t + \left(\frac{v_0}{\omega} - \frac{f \sin \beta}{2\omega^2} \right) \sin \omega t + \frac{f}{2\omega} t \sin(\omega t + \beta).$$

Osserviamo che, nel caso $\gamma = \omega$ la soluzione particolare \hat{x} rappresenta una oscillazione di frequenza ω con ampiezza crescente in modo lineare e quindi illimitata per $t \rightarrow \infty$.

Nel caso in cui la differenza tra γ e ω sia "relativamente piccola", cioè $|\omega - \gamma| \ll \omega$, allora la soluzione \hat{x} può essere approssimata in modo da descrivere il fenomeno dei *battimenti*. A questo scopo poniamo $\omega - \gamma = \nu$ dove $|\nu| \ll \omega$ e riscriviamo la soluzione del problema nella forma

$$x(t) = x_0 \cos \omega t + \frac{v_0}{\omega} \sin \omega t + x_1(t),$$

dove $x_1(t)$ è la parte di soluzione dipendente da γ , ovvero

$$x_1 = \left[\frac{f}{\omega^2 - \gamma^2} (-\cos \omega t \cos \beta) + \frac{f \gamma / \omega}{\omega^2 - \gamma^2} \sin \omega t \sin \beta + \frac{f \cos(\gamma t + \beta)}{\omega^2 - \gamma^2} \right].$$

Si ha

$$\begin{aligned} \omega^2 - \gamma^2 &= (\omega + \gamma)(\omega - \gamma) = (2\omega - \nu)\nu \simeq 2\omega\nu, \\ \frac{\gamma/\omega}{\omega^2 - \gamma^2} &= \frac{\omega - \nu}{\omega(2\omega - \nu)\nu} \simeq \frac{1}{2\omega\nu}. \end{aligned}$$

La funzione $x_1(t)$ può essere approssimata come

$$\begin{aligned} x_1(t) &\simeq -\frac{f}{2\omega\nu} [\cos(\omega t + \beta) - \cos(\omega t - \nu t + \beta)] = \\ &= -\frac{f}{2\omega\nu} \left[\cos(\omega t + \beta) 2 \sin^2 \frac{\nu}{2} t - \sin(\omega t + \beta) 2 \sin \frac{\nu}{2} t \cos \frac{\nu}{2} t \right] = \\ &= \frac{f}{\omega\nu} \sin \frac{\nu t}{2} \sin \left[\left(\omega - \frac{\nu}{2} \right) t + \beta \right]. \end{aligned}$$

Quest'ultima rappresenta una oscillazione di frequenza $\omega - \nu/2$ con ampiezza modulata dal fattore $\sin(\nu/2)t$, e quindi, con frequenza di modulazione $\nu/2 \ll \omega$. La soluzione $x_1(t)$ descrive i cosiddetti battimenti che si sovrappongono alla parte puramente armonica della soluzione.

Capitolo 5

Superfici regolari

5.1 Equazioni di superfici regolari

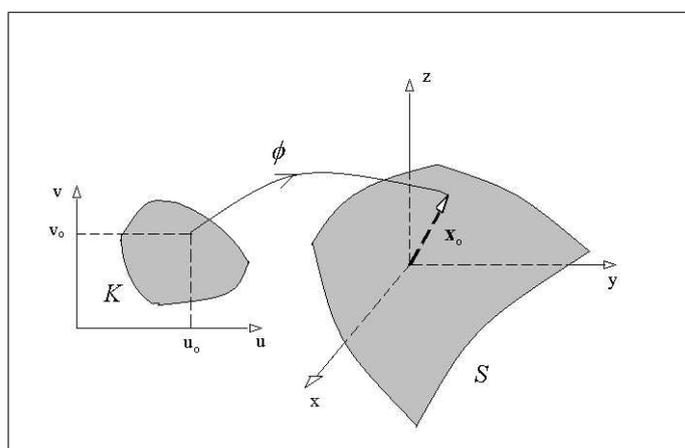
Nel seguito saranno utili le seguenti definizioni. Diremo *frontiera* ∂A di un insieme chiuso $A \subset \mathbb{R}^n$ l'insieme dei punti di A tali che in un loro qualunque intorno cadano punti di A e del suo complementare in \mathbb{R}^n . Indicheremo con $\overset{\circ}{A}$ l'insieme dei punti interni di A , ovvero l'insieme $A \setminus \partial A$. Diremo poi che un insieme $A \subset \mathbb{R}^n$ è *connesso* se non esiste una coppia di sottoinsiemi A_1, A_2 , chiusi e non vuoti, di \mathbb{R}^n tali che $A = A_1 \cup A_2$ e $A_1 \cap A_2 = \emptyset$.

Sia $K \subset \mathbb{R}^2$ un insieme chiuso, limitato e connesso e sia $\Phi : K \rightarrow \mathbb{R}^3$ una funzione vettoriale di componenti

$$\phi_1(u, v), \quad \phi_2(u, v), \quad \phi_3(u, v), \quad (u, v) \in K.$$

Consideriamo il luogo S dei punti di \mathbb{R}^3 ottenuti come immagine di ϕ , ovvero il luogo di equazioni parametriche

$$\begin{cases} x = \phi_1(u, v) \\ y = \phi_2(u, v) \\ z = \phi_3(u, v) \end{cases} \quad (u, v) \in K.$$



Diremo che S è una superficie regolare se si verificano le seguenti condizioni:

- a) $\Phi \in C^1(K)$,
- b) Esiste una corrispondenza biunivoca tra $\overset{\circ}{K}$ e i punti di S .
- c) $\text{rank}(J_\phi) = 2, \quad \forall (u, v) \in \overset{\circ}{K}$,

dove J_ϕ è la matrice jacobiana di ϕ , data da

$$J_\phi = \begin{pmatrix} \frac{\partial \phi_1}{\partial u} & \frac{\partial \phi_1}{\partial v} \\ \frac{\partial \phi_2}{\partial u} & \frac{\partial \phi_2}{\partial v} \\ \frac{\partial \phi_3}{\partial u} & \frac{\partial \phi_3}{\partial v} \end{pmatrix}.$$

La condizione c) si può anche tradurre nella indipendenza lineare dei due vettori

$$\Phi_u = \left(\frac{\partial \phi_1}{\partial u}, \frac{\partial \phi_2}{\partial u}, \frac{\partial \phi_3}{\partial u} \right), \quad \Phi_v = \left(\frac{\partial \phi_1}{\partial v}, \frac{\partial \phi_2}{\partial v}, \frac{\partial \phi_3}{\partial v} \right),$$

per ogni coppia (u, v) interna a K , ovvero, nella condizione

$$\|\Phi_u \times \Phi_v\| \neq 0, \quad \forall (u, v) \in \overset{\circ}{K}.$$

Consideriamo una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ dove A è chiuso, limitato e connesso e denotiamo con $G = \{(x, y, f(x, y)) \in \mathbb{R}^3 : (x, y) \in A\}$ il grafico di f . Tale grafico si può considerare una superficie S di equazioni parametriche

$$\begin{cases} x = u \\ y = v \\ z = f(u, v), \end{cases} \quad (u, v) \in A.$$

Dimostriamo che, se f è continua insieme alle sue derivate prime, allora la superficie S è regolare. In questo caso abbiamo

$$\phi_1 = u, \quad \phi_2 = v, \quad \phi_3 = f(u, v).$$

Poiché $f \in C^1(A)$, ne segue che la funzione vettoriale ϕ è in $C^1(A)$. Inoltre, preso un punto $(\bar{u}, \bar{v}) \in \overset{\circ}{A}$, ad esso corrisponde un solo punto di \mathbb{R}^3

$$\bar{x} = \bar{u}, \quad \bar{y} = \bar{v}, \quad \bar{z} = f(\bar{u}, \bar{v}).$$

Viceversa, ogni punto $\bar{\mathbf{x}} = (\bar{x}, \bar{y}, \bar{z})$ di S è individuato da una unica coppia (\bar{u}, \bar{v}) tale che $\bar{u} = \bar{x}$, $\bar{v} = \bar{y}$.

Ne segue la corrispondenza biunivoca tra S e $\overset{\circ}{A}$. Infine si ha

$$\Phi_u = \left(1, 0, \frac{\partial f}{\partial u} \right), \quad \Phi_v = \left(0, 1, \frac{\partial f}{\partial v} \right),$$

da cui risulta

$$\|\Phi_u \times \Phi_v\| = \sqrt{1 + \left(\frac{\partial f}{\partial u} \right)^2 + \left(\frac{\partial f}{\partial v} \right)^2} \geq 1, \quad \forall (u, v) \in \overset{\circ}{A}.$$

Ne segue che S è regolare.

Abbiamo visto nel paragrafo 2.3 che, data una funzione $f : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$, continua con le sue derivate prime in A , il suo grafico ammette in ogni punto un piano tangente. Come vedremo questa proprietà è comune a tutte le superfici regolari di cui sia data una equazione parametrica $\mathbf{x} = \Phi(u, v)$.

- Esempio

Consideriamo la superficie S di equazioni parametriche

$$\begin{cases} x = a \cos u \cos v \\ y = b \sin u \cos v \\ z = c \sin v \end{cases} \quad u \in [0, 2\pi], \quad v \in [-\pi/2, \pi/2], \quad a, b, c \in \mathbb{R}.$$

Verifichiamo la regolarità di S . Si ha $K = [0, 2\pi] \times [-\pi/2, \pi/2]$, che è chiuso e connesso. Le tre funzioni ϕ_1, ϕ_2, ϕ_3 sono certamente continue insieme alle loro derivate di qualsiasi ordine. Notiamo che la condizione $b)$ della definizione di regolarità non è soddisfatta nei punti della frontiera di K aventi $u = 0, u = 2\pi$. Ambedue i valori infatti corrispondono agli stessi punti di S . Tuttavia la condizione è sempre soddisfatta nei punti interni di K . Per quanto riguarda la condizione $c)$, si ha

$$\Phi_u = (-a \sin u \cos v, b \cos u \cos v, 0), \quad \Phi_v = (-a \cos u \sin v, -b \sin u \sin v, c \cos v),$$

da cui

$$\|\Phi_u \times \Phi_v\| = |\cos v| \sqrt{(bc \cos u \cos v)^2 + (ac \sin u \cos v)^2 + (ab \sin v)^2}.$$

La condizione $c)$ non è soddisfatta sui punti della frontiera di K per i quali $v = \pm\pi/2$, ma è sicuramente verificata nei punti interni.

5.2 Curva su una superficie regolare

Consideriamo ora una superficie regolare S definita dalla equazione

$$\mathbf{x} = \Phi(u, v), \quad \text{con } (u, v) \in K,$$

e sia $\mathcal{C} : [a, b] \rightarrow K$ una curva regolare fatta di punti interni a K , di equazioni parametriche

$$\begin{cases} u = \hat{u}(t) \\ v = \hat{v}(t) \end{cases} \quad t \in [a, b].$$

Consideriamo la funzione composta $\Psi(t) = (\Phi \circ \mathcal{C})(t) = \Phi(\hat{u}(t), \hat{v}(t))$. Preso $t_0 \in (a, b)$, poniamo $\hat{u}(t_0) = u_0, \hat{v}(t_0) = v_0$ e $\Phi(u_0, v_0) = \mathbf{x}_0$. Dimostriamo che $\Psi : [a, b] \rightarrow \mathbb{R}^3$ è una curva regolare passante per \mathbf{x}_0 . Osserviamo innanzitutto che $\Psi(t_0) = \Phi(u_0, v_0) = \mathbf{x}_0$ e quindi che \mathbf{x}_0 è un punto della curva $\mathbf{x} = \Psi(t)$. Inoltre si ha

$$\begin{cases} x = \phi_1(\hat{u}(t), \hat{v}(t)) \\ y = \phi_2(\hat{u}(t), \hat{v}(t)) \\ z = \phi_3(\hat{u}(t), \hat{v}(t)) \end{cases} \quad t \in [a, b].$$

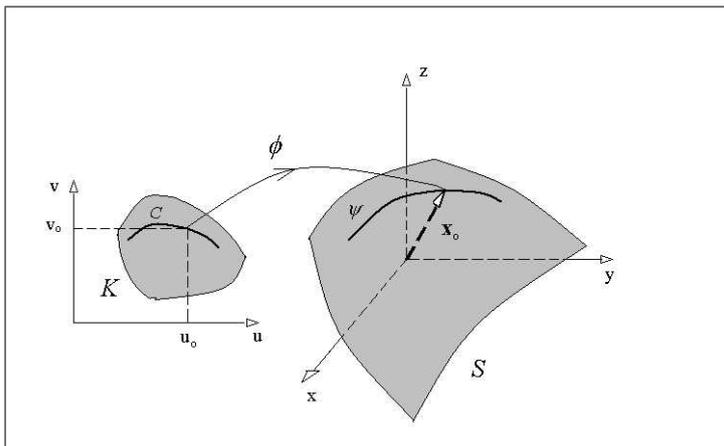
Per la regolarità di \mathcal{C} , le funzioni \hat{u}, \hat{v} sono di classe C^1 e le loro derivate non si annullano mai contemporaneamente in $[a, b]$. Inoltre, a valori di t distinti, corrispondono sempre valori distinti delle coppie (u, v) . Poiché S è regolare, le funzioni ϕ_1, ϕ_2, ϕ_3 sono di classe C^1 . Per il teorema sulla continuità delle funzioni composte, anche Ψ sarà continua con derivate continue. Le funzioni ϕ_1, ϕ_2, ϕ_3 sono poi in corrispondenza biunivoca con u e v e, a valori distinti di t corrispondono valori distinti di $\Psi(t)$.

Infine consideriamo la derivata di Ψ rispetto a t . Si ha $\Psi'(t) = \frac{d}{dt}\Phi(\hat{u}(t), \hat{v}(t))$. Per le proprietà di derivazione delle funzioni composte,

$$\begin{aligned}\frac{d\psi_1}{dt} &= \frac{\partial\phi_1}{\partial u}\hat{u}' + \frac{\partial\phi_1}{\partial v}\hat{v}', \\ \frac{d\psi_2}{dt} &= \frac{\partial\phi_2}{\partial u}\hat{u}' + \frac{\partial\phi_2}{\partial v}\hat{v}', \\ \frac{d\psi_3}{dt} &= \frac{\partial\phi_3}{\partial u}\hat{u}' + \frac{\partial\phi_3}{\partial v}\hat{v}'.\end{aligned}$$

Queste ultime si possono riguardare come un sistema di equazioni lineari nelle \hat{u}' , \hat{v}' . Se fosse $\|d\Psi/dt\| = 0$, in qualche punto di $[a, b]$, allora tale sistema sarebbe, in tale punto, omogeneo. Poiché la matrice dei coefficienti è lo jacobiano di Φ , che per ipotesi ha rango 2, l'unica soluzione sarebbe $\hat{u}' = 0, \hat{v}' = 0$. Ma per la regolarità di \mathcal{C} , ciò non è possibile, quindi deve essere necessariamente

$$\left\| \frac{d\Psi}{dt} \right\| \neq 0, \quad \forall t \in [a, b].$$



In definitiva, la curva Ψ è una curva regolare tracciata sulla superficie S . Essa ammetterà un vettore tangente in \mathbf{x}_0 , dato da

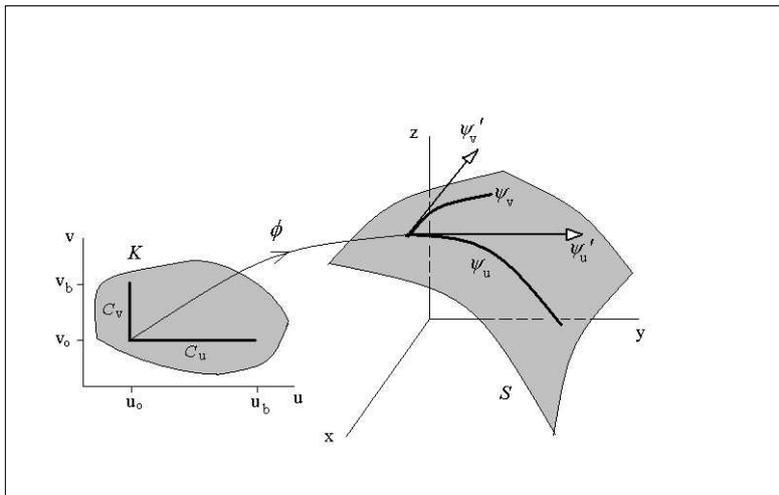
$$\Psi'(t_0) = \Phi_u(u_0, v_0)\hat{u}'(t_0) + \Phi_v(u_0, v_0)\hat{v}'(t_0).$$

Tale vettore risulta dunque combinazione lineare dei vettori Φ_u , Φ_v , i quali dipendono da S ma non da \mathcal{C} . Facciamo vedere che i vettori $\Phi_u(u_0, v_0)$ e $\Phi_v(u_0, v_0)$ individuano il piano tangente alla superficie in \mathbf{x}_0 . A tale scopo consideriamo le due curve \mathcal{C}_u e \mathcal{C}_v da $[a, b]$ in K ,

$$\mathcal{C}_u : \begin{cases} u = \hat{u}(t), \\ v = v_0, \end{cases} \quad \mathcal{C}_v : \begin{cases} u = u_0, \\ v = \hat{v}(t), \end{cases} \quad t \in [a, b].$$

A tali curve corrispondono due curve regolari $\Psi_u : [a, b] \rightarrow \mathbb{R}^3$ e $\Psi_v : [a, b] \rightarrow \mathbb{R}^3$, tracciate su S . I vettori tangenti alle due curve in \mathbf{x}_0 sono dati da

$$\Psi'_u(t_0) = \Phi_u(u_0, v_0)\hat{u}'(t_0), \quad \Psi'_v(t_0) = \Phi_v(u_0, v_0)\hat{v}'(t_0),$$



e risultano rispettivamente paralleli ai vettori Φ_u e Φ_v in \mathbf{x}_0 . Poiché Ψ'_u e Ψ'_v sono tangenti alla superficie in \mathbf{x}_0 , i vettori Φ_u e Φ_v apparterranno al piano tangente a S in \mathbf{x}_0 . Ma per la regolarità di S tali vettori sono linearmente indipendenti quindi essi costituiscono una base nel piano tangente. Il vettore

$$\mathbf{N}(\mathbf{x}_0) = \Phi_u(u_0, v_0) \times \Phi_v(u_0, v_0).$$

risulta normale alla superficie in \mathbf{x}_0 in quanto perpendicolare al piano individuato da Φ_u e Φ_v . L'equazione del piano tangente in \mathbf{x}_0 si potrà scrivere allora come

$$(\mathbf{x} - \mathbf{x}_0) \cdot \Phi_u(u_0, v_0) \times \Phi_v(u_0, v_0) = 0,$$

ovvero

$$(x - x_0)N_x + (y - y_0)N_y + (z - z_0)N_z = 0,$$

essendo x_0, y_0, z_0 le coordinate del punto \mathbf{x}_0 .

• Esempio

Determinare la retta normale alla superficie, studiata nell'esempio precedente,

$$\begin{cases} x = a \cos u \cos v \\ y = b \sin u \cos v \\ z = c \sin v \end{cases} \quad u \in [0, 2\pi], \quad v \in [-\pi/2, \pi/2], \quad a, b, c \in \mathbb{R},$$

nel punto $x_0 = a/2, y_0 = b/2, z_0 = c\sqrt{2}/2$.

Il punto corrisponde a $u_0 = \pi/4, v_0 = \pi/4$. Si ha

$$\Phi_u(u_0, v_0) \times \Phi_v(u_0, v_0) = bc \frac{\sqrt{2}}{4} \mathbf{e}_1 + ac \frac{\sqrt{2}}{4} \mathbf{e}_2 + \frac{ab}{2} \mathbf{e}_3.$$

da cui si ottiene la retta normale

$$\frac{x - a/2}{bc\sqrt{2}} = \frac{y - b/2}{ac\sqrt{2}} = \frac{z - c\sqrt{2}/2}{2ab}.$$

5.3 Teorema delle funzioni implicite

Spesso una curva piana o una superficie non sono date mediante equazioni parametriche ma, rispettivamente, attraverso equazioni cartesiane del tipo

$$F(x, y) = 0, \quad G(x, y, z) = 0.$$

Si pone allora il problema di stabilire sotto quali condizioni sulle funzioni F o G esse risultano regolari. Da quanto detto sulle superfici regolari, sappiamo che se è possibile riscrivere $G(x, y, z) = 0$ nella forma esplicita $z = g(x, y)$, allora semplici condizioni di regolarità sulla g sono sufficienti a garantire la regolarità della superficie. Nel seguito vogliamo trovare le condizioni per le quali è possibile determinare una forma esplicita della equazione di una curva o di una superficie, espresse in forma cartesiana.

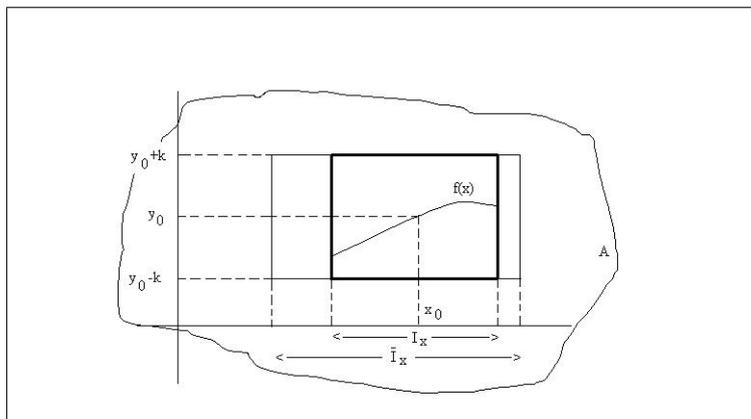
Teorema 5.1 (delle funzioni implicite). *Sia $F : A \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ di classe C^1 e sia $\mathbf{x}_0 = (x_0, y_0) \in A$ tale che $F(x_0, y_0) = 0$, $\frac{\partial F}{\partial y}(x_0, y_0) \neq 0$. Allora è possibile trovare un intorno I_x di x_0 ed un intorno I_y di y_0 tali che esista $f : I_x \rightarrow I_y$ soddisfacente la condizione*

$$F(x, f(x)) = 0.$$

Inoltre la funzione $f(x)$ risulta avere derivata continua data da

$$f'(x) = -\frac{\frac{\partial F}{\partial x}(x, f(x))}{\frac{\partial F}{\partial y}(x, f(x))}.$$

Dim. Per fissare le idee supponiamo che $\frac{\partial F}{\partial y}(x_0, y_0) > 0$. Dato che $\partial F/\partial y$ è continua in A , per il teorema della permanenza del segno, esisterà un intorno rettangolare $I = \bar{I}_x \times I_y$ di (x_0, y_0) , contenuto in A in cui $\partial F/\partial y > 0$. Fissato \bar{x} in \bar{I}_x , la funzione $F(\bar{x}, y)$ risulta dunque crescente in I_y . Ciò vale in particolare se $\bar{x} = x_0$, ed essendo $F(x_0, y_0) = 0$, posto $I_y = [y_0 - k, y_0 + k]$, avremo $F(x_0, y_0 + k) > 0$ e $F(x_0, y_0 - k) < 0$.



Consideriamo ora le funzioni di x , $F(x, y_0 - k)$ e $F(x, y_0 + k)$ che risultano continue in \bar{I}_x . Applicando anche qui il teorema della permanenza del segno, è possibile trovare un intorno $I_x \subset \bar{I}_x$ tale che

$$F(x, y_0 - k) < 0, \quad F(x, y_0 + k) > 0, \quad \forall x \in I_x.$$

Ma allora, per ogni $x \in I_x$, la funzione $F(x, y)$ è strettamente crescente e passa da un valore negativo (in $y_0 - k$) ad un valore positivo (in $y_0 + k$). Esisterà dunque, per ciascun $x \in I_x$ una ed una sola

$y \in I_y$ tale che $F(x, y) = 0$. In altri termini, esisterà una funzione $f : I_x \rightarrow I_y$ tale che $F(x, f(x)) = 0$, identicamente in I_x .

Dimostriamo ora che f è continua. Presi due punti x, x_1 in I_x , possiamo scrivere

$$F(x, f(x)) - F(x_1, f(x_1)) = 0.$$

D'altra parte, si ha

$$F(x, f(x)) - F(x_1, f(x_1)) = F(x, f(x)) - F(x, f(x_1)) + F(x, f(x_1)) - F(x_1, f(x_1)).$$

Utilizzando la derivabilità di F rispetto a x e a y , applichiamo il teorema del valor medio negli intervalli di estremi x, x_1 e $f(x), f(x_1)$. Otteniamo

$$\frac{\partial F}{\partial x}(\xi, f(x_1))(x - x_1) + \frac{\partial F}{\partial y}(x, \eta)(f(x) - f(x_1)) = 0,$$

dove (ξ, η) è un punto interno al rettangolo $[x, x_1] \times [f(x), f(x_1)]$, contenuto in $I_x \times I_y$. Poiché $\frac{\partial F}{\partial y} > 0$ in tutto $I_x \times I_y$, si ha

$$f(x) - f(x_1) = -\frac{\frac{\partial F}{\partial x}(\xi, f(x_1))}{\frac{\partial F}{\partial y}(x, \eta)}(x - x_1),$$

da cui, prendendo il valore assoluto e osservando che $\partial F/\partial x$ è limitata e che $\partial F/\partial y > 0$ in $I_x \times I_y$, si ottiene

$$|f(x) - f(x_1)| = \left| \frac{\partial F/\partial x(\xi, f(x_1))}{\partial F/\partial y(x, \eta)} \right| |x - x_1| \leq \left| \frac{\max_{I_x \times I_y}(\partial F/\partial x)}{\min_{I_x \times I_y}(\partial F/\partial y)} \right| |x - x_1|.$$

Prendendo il limite per $x \rightarrow x_1$ si ricava

$$\lim_{x \rightarrow x_1} |f(x) - f(x_1)| = 0,$$

e quindi, la continuità di $f(x)$ in x_1 . Ciò vale per qualunque scelta di x_1 in I_x , quindi la funzione f risulta continua in I_x .

Dalla relazione precedente si ha anche

$$\frac{f(x) - f(x_1)}{x - x_1} = -\frac{\partial F/\partial x(\xi, f(x_1))}{\partial F/\partial y(x, \eta)}.$$

Passando al limite per $x \rightarrow x_1$, si ottiene, per la continuità di f e delle derivate parziali di F ,

$$f'(x_1) = -\frac{\partial F/\partial x(x_1, f(x_1))}{\partial F/\partial y(x_1, f(x_1))},$$

e questo completa la dimostrazione del teorema.

Questo teorema si può estendere facilmente a funzioni di tre o più variabili. Nel caso di tre variabili si ha il seguente risultato.

Teorema 5.2 Sia $F : A \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ di classe C^1 e sia $\mathbf{x}_0 = (x_0, y_0, z_0)$ un punto di A tale che $F(\mathbf{x}_0) = 0$ e $\frac{\partial F}{\partial z}(\mathbf{x}_0) \neq 0$. Allora è possibile trovare un intorno K del punto (x_0, y_0) ed un intorno I_z di z_0 , con $K \times I_z \subset A$, in modo tale che esista una funzione $f : K \rightarrow I_z$ soddisfacente la condizione

$$F(x, y, f(x, y)) = 0,$$

identicamente in K . Inoltre la funzione f risulta avere derivate parziali continue in K , date da

$$\frac{\partial f}{\partial x}(x, y) = -\frac{\frac{\partial F}{\partial x}(x, y, f(x, y))}{\frac{\partial F}{\partial z}(x, y, f(x, y))}, \quad \frac{\partial f}{\partial y}(x, y) = -\frac{\frac{\partial F}{\partial y}(x, y, f(x, y))}{\frac{\partial F}{\partial z}(x, y, f(x, y))}.$$

 • Esempio

Calcolare le derivate parziali della funzione $f(x, y)$, definita implicitamente da $F(x, y, z) = \exp(x^2 + y^2 - z^2) - \frac{z}{5}$ nel punto $(x_0, y_0) = (3, 4)$.

Si ha $F(3, 4, z) = \exp(25 - z^2) - \frac{z}{5} = 0$ per $z = z_0 = 5$. Inoltre

$$\frac{\partial F}{\partial z} = -2z \exp(x^2 + y^2 - z^2) - \frac{1}{5} \implies \frac{\partial F}{\partial z}(x_0, y_0, z_0) = -\frac{51}{5} \neq 0.$$

È dunque possibile applicare il teorema delle funzioni implicite nel punto $(3, 4, 5)$. In un intorno di $(3, 4)$ esiste allora una funzione $f(x, y)$ continua insieme alle sue derivate parziali prime e si ha

$$\frac{\partial f}{\partial x}(x, y) = \frac{2x \exp(x^2 + y^2 - f(x, y)^2)}{2f(x, y) \exp(x^2 + y^2 - f(x, y)^2) + \frac{1}{5}}, \quad \frac{\partial f}{\partial y}(x, y) = \frac{2y \exp(x^2 + y^2 - f(x, y)^2)}{2f(x, y) \exp(x^2 + y^2 - f(x, y)^2) + \frac{1}{5}},$$

da cui

$$\frac{\partial f}{\partial x}(3, 4) = \frac{30}{51}, \quad \frac{\partial f}{\partial y}(3, 4) = \frac{40}{51}.$$

I teoremi delle funzioni implicite fin qui esaminati sono casi particolari di un teorema generale applicabile a funzioni vettoriali di più variabili. Diamo qui l'enunciato di questo importante teorema senza riportarne la dimostrazione.

Teorema 5.3 *Sia $\mathbf{F} : A \subseteq \mathbb{R}^n \rightarrow \mathbb{R}^p$ una funzione vettoriale di classe C^1 nell'aperto A , con $p < n$. Sia $\mathbf{x}_0 \in A$ tale che $\mathbf{F}(\mathbf{x}_0) = \mathbf{0}$ e detta $J_{\mathbf{F}}$ la matrice jacobiana di \mathbf{F} sia*

$$\text{rank}(J_{\mathbf{F}}) = p, \quad \text{in } A.$$

Allora, posto

$$\mathbf{x} = (y_1, y_2, \dots, y_{n-p}, z_1, z_2, \dots, z_p) =: (\mathbf{y}, \mathbf{z}), \quad \text{con } \mathbf{y} \in \mathbb{R}^{n-p}, \quad \mathbf{z} \in \mathbb{R}^p,$$

$$J_{ij}^{(p)} = \frac{\partial F_i}{\partial z_j}, \quad J_{hk}^{(n-p)} = \frac{\partial F_h}{\partial y_k},$$

esiste un intorno $B(\mathbf{x}_0)$ ed una funzione vettoriale $\mathbf{f} : H \rightarrow K$, con $H \subseteq \mathbb{R}^{n-p}$, $K \subseteq \mathbb{R}^p$ e $H \times K \subseteq B(\mathbf{x}_0)$, tale che

$$\mathbf{F}(\mathbf{y}, \mathbf{f}(\mathbf{y})) = \mathbf{0}, \quad \forall \mathbf{y} \in H.$$

Inoltre \mathbf{f} è continua e differenziabile in H e si ha

$$J_{\mathbf{f}}(\mathbf{y}) = [J^{(p)}(\mathbf{y}, \mathbf{f}(\mathbf{y}))]^{-1} J^{(n-p)}(\mathbf{y}, \mathbf{f}(\mathbf{y})).$$

Anche in questo caso, si dice che la funzione vettoriale \mathbf{F} definisce implicitamente la funzione vettoriale \mathbf{f} . In altri termini, la tesi di questo teorema stabilisce che p delle n variabili indipendenti possono essere espresse come funzioni delle restanti $n - p$ variabili, ovvero

$$\begin{aligned} z_1 &= f_1(y_1, y_2, \dots, y_{n-p}) \\ z_2 &= f_2(y_1, y_2, \dots, y_{n-p}) \\ &\vdots \\ z_p &= f_p(y_1, y_2, \dots, y_{n-p}) \end{aligned}$$

Tali funzioni risultano continue e differenziabili in H e le loro derivate parziali sono date da

$$\frac{\partial f_i}{\partial y_j} = \sum_{k=1}^p (J^{(p)})_{ik}^{-1} \frac{\partial F_k}{\partial y_j}$$

5.4 Superfici in forma implicita

Consideriamo una funzione $G(x, y, z)$ continua insieme alle sue derivate parziali prime in un sottoinsieme A , connesso, di \mathbb{R}^3 . La funzione risulterà differenziabile in ogni punto di A . Sia $\mathbf{x}_0 \in A$ e supponiamo che $\nabla G(\mathbf{x}_0) \neq 0$. Posto allora

$$G(\mathbf{x}_0) = \alpha,$$

consideriamo la funzione

$$F(\mathbf{x}) = G(\mathbf{x}) - \alpha.$$

Si ha $F(\mathbf{x}_0) = 0$, ed inoltre

$$\nabla F(\mathbf{x}_0) = \nabla G(\mathbf{x}_0) \neq 0.$$

Ne segue che almeno una delle derivate parziali prime di F in \mathbf{x}_0 è diversa da zero, per esempio $\partial F/\partial z$. Sono allora soddisfatte le ipotesi del teorema delle funzioni implicite e si ha, in un intorno $B(x_0, y_0) \times I_z$ del punto \mathbf{x}_0 ,

$$z = f(x, y).$$

Analogamente, se fosse $\partial F/\partial x \neq 0$, oppure $\partial F/\partial y \neq 0$, si otterrebbero rispettivamente le identità

$$x = g(y, z), \quad \text{in } B(y_0, z_0) \times I_x, \quad \text{oppure} \quad y = h(x, z), \quad \text{in } B(x_0, z_0) \times I_y.$$

Ritornando al caso in cui $\partial F/\partial z \neq 0$, l'equazione $z = f(x, y)$ rappresenta una superficie regolare definita in $B(x_0, y_0)$. Infatti, sempre per il teorema delle funzioni implicite, la funzione f ammette derivate parziali prime continue in $B(x_0, y_0)$. Da tutto ciò segue che l'equazione

$$F(x, y, z) = 0$$

rappresenta in $B(x_0, y_0) \times I_z$ una superficie regolare e, di conseguenza, $G(\mathbf{x}) = \alpha$ è l'equazione di una superficie regolare in quello stesso dominio.

Riassumendo, abbiamo dimostrato che, data una funzione $G(x, y, z)$ di classe C^1 in A , connesso di \mathbb{R}^3 , l'equazione $G(x, y, z) = \alpha$ rappresenta una superficie regolare in tutti i punti in cui $\nabla G \neq 0$.

Facciamo vedere, infine, che ∇G è normale alla superficie $G(\mathbf{x}) = \alpha$. Dalla condizione di regolarità possiamo sempre dire che esiste un K connesso ed una funzione vettoriale $\Phi(u, v)$ tali che

$$G(\phi_1(u, v), \phi_2(u, v), \phi_3(u, v)) = \alpha,$$

identicamente in K . Preso allora un generico punto \mathbf{x}_0 tale che $\nabla G(\mathbf{x}_0) \neq 0$, in tale punto si avrà

$$\begin{aligned} \frac{\partial G}{\partial u} &= \frac{\partial G}{\partial x} \frac{\partial \phi_1}{\partial u} + \frac{\partial G}{\partial y} \frac{\partial \phi_2}{\partial u} + \frac{\partial G}{\partial z} \frac{\partial \phi_3}{\partial u} = 0 \\ \frac{\partial G}{\partial v} &= \frac{\partial G}{\partial x} \frac{\partial \phi_1}{\partial v} + \frac{\partial G}{\partial y} \frac{\partial \phi_2}{\partial v} + \frac{\partial G}{\partial z} \frac{\partial \phi_3}{\partial v} = 0, \end{aligned}$$

ovvero

$$\nabla G(\mathbf{x}_0) \cdot \Phi_u(u_0, v_0) = 0, \quad \nabla G(\mathbf{x}_0) \cdot \Phi_v(u_0, v_0) = 0.$$

Ma i vettori Φ_u, Φ_v definiscono il piano tangente alla superficie quindi ∇G è ortogonale al piano tangente in \mathbf{x}_0 .

- Esempio

Calcolare la retta normale ed il piano tangente alla superficie di equazione implicita

$$3xyz - z^3 - a^3 = 0,$$

nel punto $x_0 = 0, y_0 = a$, essendo $a \neq 0$. Si ha $G(\mathbf{x}) = 3xyz - z^3 - a^3$, da cui

$$\nabla G(\mathbf{x}) = 3yz\mathbf{e}_1 + 3xz\mathbf{e}_2 + (3xy - 3z^2)\mathbf{e}_3,$$

ed essendo $\mathbf{x}_0 = (0, a, -a)$, si ottiene

$$\nabla G(\mathbf{x}_0) = -3a^2\mathbf{e}_1 - 3a^2\mathbf{e}_3.$$

L'equazione della retta normale alla superficie in \mathbf{x}_0 sarà

$$\begin{cases} x = -3a^2t \\ y = a \\ z = -a - 3a^2t \end{cases} .$$

L'equazione del piano tangente sarà

$$x + z + a = 0.$$

Capitolo 6

Integrazione delle funzioni di più variabili

6.1 Insiemi misurabili

Consideriamo il rettangolo $R = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, c \leq y \leq d\}$ e denotiamo con $\text{mis}(R) = (b - a)(d - c)$, la sua area. Chiameremo *plurirettangolo* \bar{R} l'insieme dato dall'unione di un numero finito di rettangoli $R_i, (i = 1, \dots, n)$, privi di punti interni comuni e definiremo $\text{mis}(\bar{R}) = \sum_{i=1}^n \text{mis}(R_i)$. Sia D un insieme limitato di \mathbb{R}^2 e denotiamo con \mathcal{R}_D l'insieme di tutti i plurirettangoli contenuti in D e con \mathcal{R}'_D l'insieme di tutti i plurirettangoli contenuti in D . Diremo che D è *misurabile* se si verifica uno dei seguenti casi.

- D è privo di punti interni e

$$\inf_{\bar{R} \in \mathcal{R}_D} \{\text{mis}(\bar{R})\} = 0.$$

Allora D è un insieme di misura nulla ovvero $\text{mis}(D) = 0$.

- D è dotato di punti interni e risulta

$$\inf_{\bar{R} \in \mathcal{R}_D} \{\text{mis}(\bar{R})\} = \sup_{\bar{R}' \in \mathcal{R}'_D} \{\text{mis}(\bar{R}')\} = A.$$

Allora $\text{mis}(D) = A$.

Se l'insieme D non è limitato esso si dirà misurabile se risultano misurabili le sue intersezioni con un qualunque cerchio. L'estremo superiore delle misure di queste intersezioni sarà la misura di D .

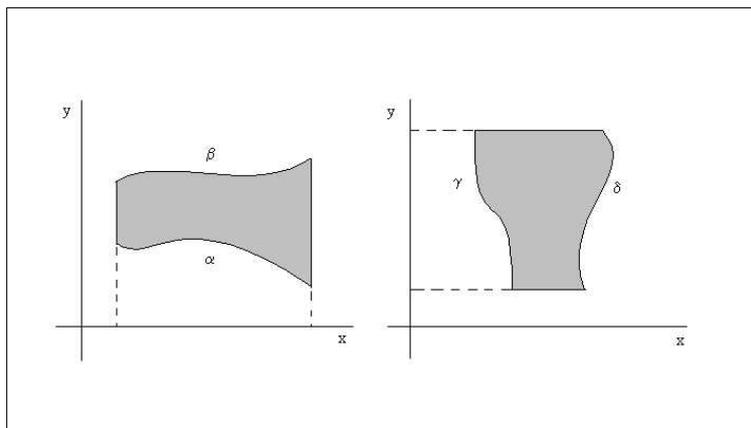
Semplici esempi di insiemi misurabili in \mathbb{R}^2 sono i *domini normali*.

Siano $\alpha(x)$ e $\beta(x)$ due funzioni integrabili in $[a, b]$ e tali che $\alpha(x) \leq \beta(x), \forall x \in [a, b]$. Consideriamo l'insieme $D = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \alpha(x) \leq y \leq \beta(x)\}$. Si dice che D è un *dominio normale rispetto all'asse x*. Verifichiamo che D è un insieme misurabile. Infatti, per ogni partizione P di $[a, b]$ si può costruire l'insieme $\mathcal{R}_D(P)$ dei plurirettangoli R_P contenenti D e l'insieme $\mathcal{R}'_D(P)$ dei plurirettangoli R'_P contenuti in D tali che

$$\text{mis}(R_P) = S(\beta, P) - s(\alpha, P), \quad \text{mis}(R'_P) = s(\beta, P) - S(\alpha, P).$$

dove S e s indicano rispettivamente le somme integrali superiore e inferiore. Si ha

$$\begin{aligned} \inf_P \text{mis}(R_P) &= \inf_P [S(\beta, P) - s(\alpha, P)] = \inf_P S(\beta, P) - \sup_P s(\alpha, P) \\ \sup_P \text{mis}(R'_P) &= \sup_P [s(\beta, P) - S(\alpha, P)] = \sup_P s(\beta, P) - \inf_P S(\alpha, P). \end{aligned}$$



Ma per l'integrabilità di α e di β si ha $\inf_P S(\beta, P) = \sup_P s(\beta, P)$ e $\sup_P s(\alpha, P) = \inf_P S(\alpha, P)$, e quindi

$$\inf_P \text{mis}(R_P) = \sup_P \text{mis}(R'_P).$$

Ne segue la misurabilità di D . Si ha inoltre

$$\text{mis}(D) = \int_a^b [\beta(x) - \alpha(x)] dx.$$

Notiamo che se fosse $\alpha = \beta = f$ in $[a, b]$, D coinciderebbe con il grafico di f . In tal caso si avrebbe un insieme privo di punti interni e tale che

$$\inf_P \text{mis}(R_P) = 0.$$

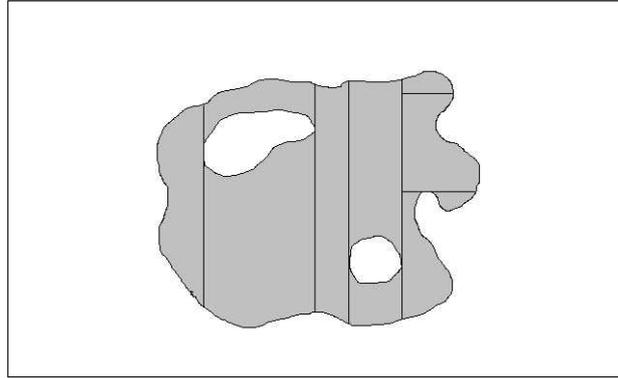
Ciò significa che il grafico di una funzione integrabile ha misura nulla in \mathbb{R}^2 .

Analogamente, date due funzioni integrabili $\gamma(y), \delta(y)$ definite in $[c, d]$, con $\gamma(y) \leq \delta(y), \forall y \in [c, d]$, l'insieme $D = \{(x, y) \in \mathbb{R}^2 : \gamma(y) \leq x \leq \delta(y), c \leq y \leq d\}$, si dice *dominio normale rispetto all'asse y* e risulta un insieme misurabile, con

$$\text{mis}(D) = \int_c^d [\delta(y) - \gamma(y)] dy.$$

Se le funzioni $\alpha, \beta, \gamma, \delta$ degli esempi precedenti sono anche derivabili con derivate prime continue, allora i loro grafici saranno curve regolari ed i domini D si diranno *domini normali regolari*. Tutti gli insiemi piani delimitati da curve regolari a tratti si possono decomporre nella unione di domini normali regolari privi di punti interni in comune, e quindi risultano misurabili (vedi figura).

Concetti analoghi possono essere introdotti in \mathbb{R}^3 . Sia R il parallelepipedo definito da $\{(x, y, z) \in \mathbb{R}^3 : a \leq x \leq a', b \leq y \leq b', c \leq z \leq c'\}$. Denoteremo con $\text{mis}(R) = (a' - a)(b' - b)(c' - c)$ il suo volume. Diremo *pluriparallelepipedo* \bar{R} l'insieme dato dall'unione di un numero finito di parallelepipedi $R_i (i = 1, \dots, n)$ privi di punti interni comuni e definiamo $\text{mis}(\bar{R}) = \sum_{i=1}^n \text{mis}(R_i)$. Sia $T \subset \mathbb{R}^3$, limitato. Denotati con \mathcal{R}_T l'insieme di tutti i pluriparallelepipedo contenenti T e con \mathcal{R}'_T l'insieme di tutti i pluriparallelepipedo contenuti in T , diremo che T è *misurabile* (o *cubabile*) se si verifica uno dei seguenti casi.



- T è privo di punti interni e si ha

$$\inf_{\bar{R} \in \mathcal{R}_T} \text{mis}(\bar{R}) = 0.$$

Allora $\text{mis}(T) = 0$.

- L'estremo superiore delle misure dei pluriparallelepipedi di \mathcal{R}'_T coincide con l'estremo inferiore delle misure dei pluriparallelepipedi di \mathcal{R}_T . In tal caso

$$\text{mis}(T) = \inf_{\bar{R} \in \mathcal{R}_T} \text{mis}(\bar{R}) = \sup_{\bar{R}' \in \mathcal{R}'_T} \text{mis}(\bar{R}').$$

Un insieme $T \subset \mathbb{R}^3$ non limitato si dice misurabile se sono misurabili le sue intersezioni con una qualunque sfera. L'estremo superiore delle misure di tali intersezioni sarà la misura di T .

Dato l'insieme $D \subset \mathbb{R}^2$ e il numero $h \in \mathbb{R}^+$, diremo cilindro (in senso generalizzato) l'insieme $T \subset \mathbb{R}^3$ definito da $T = D \times [0, h]$. Si dimostra facilmente che se D è limitato e misurabile allora T è misurabile e si ha

$$\text{mis}(T) = \text{mis}(D) \cdot h.$$

Si dice *pluricilindro* l'unione di un numero finito di cilindri privi di punti interni in comune. La sua misura sarà la somma delle misure dei singoli cilindri.

Sia $f : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}^+$. L'insieme $T = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D, 0 \leq z \leq f(x, y)\}$ si dice *cilindroide* relativo a D . Si può dimostrare che se D è limitato e misurabile e se f è continua in D , allora il cilindroide T è misurabile.

6.2 Integrali doppi

Sia $D \in \mathbb{R}^2$ un insieme misurabile e sia $f(x, y)$ una funzione limitata e non negativa in D . Effettuiamo una partizione P di D in un numero finito di sottoinsiemi D_i misurabili tali che $D = \cup_{i=1}^n D_i$ e $\overset{\circ}{D}_i \cap \overset{\circ}{D}_j = \emptyset$ per $i \neq j$. Siano m_i e M_i rispettivamente l'estremo inferiore e l'estremo superiore di f in D_i . Consideriamo le *somme integrali*

$$s(f, P) = \sum_{i=1}^n m_i \text{mis}(D_i), \quad S(f, P) = \sum_{i=1}^n M_i \text{mis}(D_i),$$

che rappresentano le misure di due pluricilindri, il primo contenuto ed il secondo contenente il cilindroide di f relativo a D . Poiché $m_i \leq M_i, (i = 1, \dots, n)$, si ha, per ogni partizione P ,

$$s(f, P) \leq S(f, P).$$

Inoltre, se si considera un raffinamento P' della partizione P , ad esempio in modo tale che gli insiemi D_i abbiano diametro inferiore a quello degli insiemi nella partizione P , si ha

$$s(f, P') \geq s(f, P), \quad S(f, P') \leq S(f, P).$$

Al variare di tutte le possibili partizioni, si avrà

$$\sup_P [s(f, P)] \leq \inf_P [S(f, P)].$$

Si dice allora che $f(x, y)$ è *integrabile* in D se accade che

$$\sup_P [s(f, P)] = \inf_P [S(f, P)] = I_D,$$

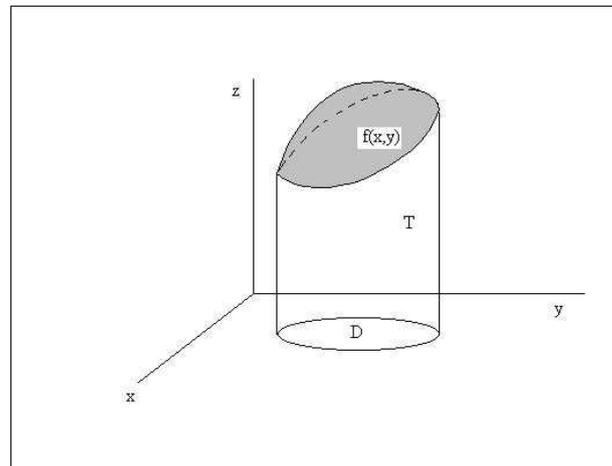
e si pone

$$I_D = \iint_D f(x, y) \, dx dy,$$

detto *integrale doppio* di f esteso al dominio D . In forma più sintetica si scrive anche

$$I_D = \int_D f(\mathbf{x}) \, d\mathbf{x},$$

dove $\mathbf{x} = (x, y)$.



In base a questa definizione e al risultato del paragrafo precedente sulla misurabilità del cilindroide $T = \{(x, y, z) \in \mathbb{R}^3 : (x, y) \in D, 0 \leq z \leq f(x, y)\}$, l'integrale doppio della funzione $f(x, y)$ continua in D assume il significato geometrico della misura di T , ovvero del volume del cilindroide di sezione D ,

$$\text{mis}(T) = \int_D f(\mathbf{x}) \, d\mathbf{x}.$$

Notiamo che, se $f(x, y) = 1, \forall (x, y) \in D$, allora si ha $m_i = M_i = 1$ e $s(f, P) = S(f, P) = \sum_{i=1}^n \text{mis}(D_i) = \text{mis}(D)$. Ne segue

$$\int_D dx = \text{mis}(D).$$

Se f ha segno variabile in D , si può sempre considerare la funzione ausiliaria $g(x, y) = f(x, y) + k$ ($k > 0$) dove, essendo f limitata, k è tale che $|f(x, y)| < k$. In tal caso $g(x, y)$ è non negativa in D e si ha

$$\begin{aligned} \sup_P [s(g, P)] - \inf_P [S(g, P)] &= \sup_P [s(f, P)] + k \text{mis}(D) - \inf_P [S(f, P)] - k \text{mis}(D) \\ &= \sup_P [s(f, P)] - \inf_P [S(f, P)]. \end{aligned}$$

Quindi, se g è integrabile lo sarà anche f e si avrà

$$\int_D f(\mathbf{x}) d\mathbf{x} = \int_D g(\mathbf{x}) d\mathbf{x} - k \text{mis}(D).$$

Si può osservare che la precedente definizione di funzione integrabile è formalmente simile a quella data per funzioni di una sola variabile. Questa analogia consente di stabilire il seguente teorema di integrabilità, analogo a quello dimostrato per le funzioni di una variabile.

Teorema 6.1 *Data $f : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ con D limitato e misurabile, condizione necessaria e sufficiente affinché essa sia integrabile in D è che, fissato un $\varepsilon > 0$, esista una partizione \bar{P} di D tale che*

$$S(f, \bar{P}) - s(f, \bar{P}) < \varepsilon.$$

La dimostrazione di questo teorema è identica a quella effettuata per funzioni di una sola variabile, in quanto fa riferimento ai soli concetti di partizione e di somme integrali e non verrà quindi ripetuta.

Il seguente teorema, che non dimostreremo, costituisce una importante condizione sufficiente per l'integrabilità.

Teorema 6.2 *Sia data $f : D \subset \mathbb{R}^2 \rightarrow \mathbb{R}$ con D limitato e misurabile. Se f è continua in D , eccettuato al più un insieme U_0 di misura nulla, allora f è integrabile in D .*

Come corollario a questo teorema segue anche che se f e g sono due funzioni continue, definite in un insieme limitato e misurabile e risulta

$$f(x, y) = g(x, y), \quad \forall (x, y) \in D \setminus U_0,$$

allora le due funzioni sono integrabili e i loro integrali estesi a D coincidono.

Gli integrali doppi soddisfano proprietà analoghe a quelle degli integrali delle funzioni di una variabile. Le più importanti sono le seguenti.

1) Date due funzioni f e g integrabili in D e dati $c_1, c_2 \in \mathbb{R}$ si ha

$$\iint_D [c_1 f(x, y) + c_2 g(x, y)] dx dy = c_1 \iint_D f(x, y) dx dy + c_2 \iint_D g(x, y) dx dy.$$

2) Siano D_1 e D_2 due sottoinsiemi misurabili di D tali che $D_1 \cup D_2 = D$ e $\overset{\circ}{D}_1 \cap \overset{\circ}{D}_2 = \emptyset$. Per ogni f

integrabile in D si ha

$$\iint_D f(x, y) \, dx dy = \iint_{D_1} f(x, y) \, dx dy + \iint_{D_2} f(x, y) \, dx dy.$$

3) Se f e g sono due funzioni integrabili in D e tali che $f(x, y) \geq g(x, y), \forall (x, y) \in D$, si ha

$$\iint_D f(x, y) \, dx dy \geq \iint_D g(x, y) \, dx dy.$$

4) Dalla proprietà precedente segue che, essendo $-|f(\mathbf{x})| \leq f(\mathbf{x}) \leq |f(\mathbf{x})|$, si ha

$$-\int_D |f(\mathbf{x})| \, d\mathbf{x} \leq \int_D f(\mathbf{x}) \, d\mathbf{x} \leq \int_D |f(\mathbf{x})| \, d\mathbf{x},$$

da cui

$$\left| \int_D f(\mathbf{x}) \, d\mathbf{x} \right| \leq \int_D |f(\mathbf{x})| \, d\mathbf{x}.$$

5) (Teorema della media). Se f è continua in un insieme misurabile chiuso e limitato D , esiste un punto $\bar{\mathbf{x}} \in D$ tale che

$$\int_D f(\mathbf{x}) \, d\mathbf{x} = f(\bar{\mathbf{x}}) \, \text{mis}(D).$$

La dimostrazione di questa proprietà è simile al caso delle funzioni di una variabile. Nel nostro caso, detti m ed M rispettivamente il minimo e il massimo di f in D , scriviamo la disuguaglianza

$$m \leq f(\mathbf{x}) \leq M, \quad \forall \mathbf{x} \in D.$$

Integrando su D e tenendo conto delle proprietà 1) e 3) si ha

$$m \, \text{mis}(D) \leq \int_D f(\mathbf{x}) \, d\mathbf{x} \leq M \, \text{mis}(D),$$

da cui

$$m \leq \frac{1}{\text{mis}(D)} \int_D f(\mathbf{x}) \, d\mathbf{x} \leq M.$$

Per la sua continuità, f assumerà in D tutti i valori compresi tra m ed M e quindi esisterà un $\bar{\mathbf{x}}$ tale che

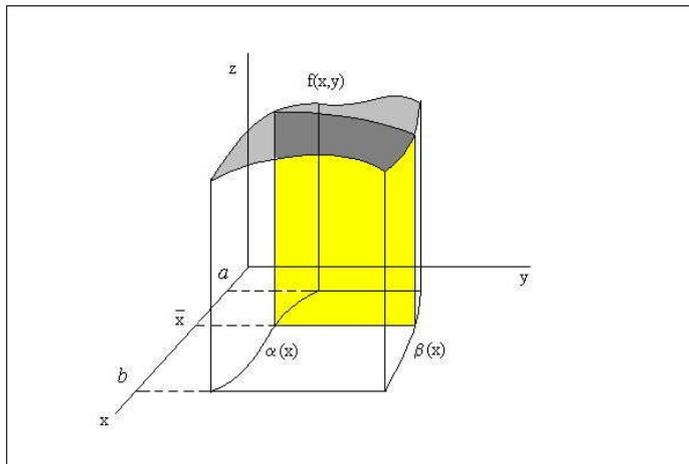
$$f(\bar{\mathbf{x}}) = \frac{1}{\text{mis}(D)} \int_D f(\mathbf{x}) \, d\mathbf{x}.$$

6.3 Formule di riduzione per gli integrali doppi

Consideriamo nel piano cartesiano un dominio normale rispetto all'asse x , dato da $D = \{(x, y) \in \mathbb{R}^2 : a \leq x \leq b, \alpha(x) \leq y \leq \beta(x)\}$. Sia $f(x, y)$ continua in D e quindi integrabile. Fissato $\bar{x} \in [a, b]$ consideriamo la quantità

$$\int_{\alpha(\bar{x})}^{\beta(\bar{x})} f(\bar{x}, y) \, dy.$$

Poiché $f(x, y)$ è continua, la sua restrizione per $x = \bar{x}$ è una funzione continua di y . Inoltre, per ogni $x \in [a, b]$ si ha $\alpha(x) \leq \beta(x)$ e quindi $\alpha(\bar{x}) \leq \beta(\bar{x})$. Queste osservazioni danno senso all'integrale



precedente, comunque si scelga \bar{x} nell'intervallo $[a, b]$. Esso rappresenta geometricamente l'area della sezione del cilindroide T di f relativo a D , sul piano $x = \bar{x}$.

Definiamo ora la funzione

$$G(x) = \int_{\alpha(x)}^{\beta(x)} f(x, y) dy.$$

Si può dimostrare che, se $\alpha(x)$ e $\beta(x)$ sono continue in $[a, b]$, allora $G(x)$ è continua in $[a, b]$. Essa risulta quindi integrabile e si ha

$$\int_a^b G(x) dx = \sup_P s(G, P) = \inf_P S(G, P),$$

dove $s(G, P)$ e $S(G, P)$ sono rispettivamente le somme integrali inferiore e superiore di G per una data partizione P dell'intervallo $[a, b]$. Si può osservare che s e S rappresentano le misure di due pluricilindroidi T_s e T_S . In generale T_s non è contenuto in T , nè T_S contiene T . Sussiste tuttavia la disegualianza

$$s(G, P) \leq \text{mis}(T) \leq S(G, P), \quad \forall P.$$

Allora dall'integrabilità di G si dovrà necessariamente avere

$$\int_a^b G(x) dx = \text{mis}(T).$$

Ne segue la formula di riduzione per gli integrali doppi estesi ad un dominio normale rispetto all'asse x

$$\iint_D f(x, y) dx dy = \int_a^b dx \int_{\alpha(x)}^{\beta(x)} f(x, y) dy.$$

In modo analogo, considerato un dominio normale rispetto all'asse y , sotto le corrispondenti ipotesi, si ricava

$$\iint_D f(x, y) dx dy = \int_c^d dy \int_{\gamma(y)}^{\delta(y)} f(x, y) dx.$$

Se nelle formule precedenti si pone $f(x, y) = 1, \forall (x, y) \in D$, si ottiene il risultato già noto

$$\text{mis}(D) = \iint_D dx dy = \begin{cases} \int_a^b [\beta(x) - \alpha(x)] dx, & \text{se } D \text{ è normale rispetto a } x \\ \int_c^d [\delta(y) - \gamma(y)] dy, & \text{se } D \text{ è normale rispetto a } y \end{cases}$$

• Esempi

1) Consideriamo il dominio D compreso tra l'arco di ellisse di equazione $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ del primo quadrante e gli assi x e y . Vogliamo calcolare

$$\iint_D x \, dx \, dy.$$

Il dominio D risulta normale sia rispetto all'asse x che rispetto all'asse y . Come dominio normale rispetto all'asse x esso è dato da

$$D : \begin{cases} 0 \leq x \leq a \\ 0 \leq y \leq \frac{b}{a} \sqrt{a^2 - x^2}. \end{cases}$$

Mediante la formula di riduzione, otteniamo

$$\begin{aligned} \iint_D x \, dx \, dy &= \int_0^a dx \int_0^{\frac{b}{a} \sqrt{a^2 - x^2}} x \, dy = \int_0^a x \, dx \int_0^{\frac{b}{a} \sqrt{a^2 - x^2}} dy \\ &= \int_0^a \frac{b}{a} \sqrt{a^2 - x^2} x \, dx = -\frac{b}{2a} \int_0^a -2x \sqrt{a^2 - x^2} \, dx = \frac{b}{3} a^2. \end{aligned}$$

2) Calcoliamo l'integrale doppio

$$\iint_D (2x - y + 3) \, dx \, dy,$$

dove D è il dominio compreso tra la parabola $y = x^2$ e la retta $y = x$. Anche in questo caso il dominio può essere visto come normale rispetto all'asse delle x o all'asse delle y . Considerato come dominio normale rispetto all'asse x si ha

$$D : \begin{cases} 0 \leq x \leq 1 \\ x^2 \leq y \leq x. \end{cases}$$

Si ottiene così

$$\begin{aligned} \iint_D (2x - y + 3) \, dx \, dy &= \int_0^1 dx \int_{x^2}^x (2x - y + 3) \, dy = \int_0^1 \left[2xy - \frac{y^2}{2} + 3y \right]_{y=x^2}^{y=x} dx \\ &= \int_0^1 \left(2x^2 - \frac{x^2}{2} + 3x - 2x^3 + \frac{x^4}{2} - 3x^2 \right) dx \\ &= \int_0^1 \left(3x - \frac{3}{2}x^2 - 2x^3 + \frac{x^4}{2} \right) dx = \frac{3}{5}. \end{aligned}$$

Considerato come dominio normale rispetto all'asse delle y , si ha

$$D : \begin{cases} 0 \leq y \leq 1 \\ y \leq x \leq \sqrt{y}. \end{cases}$$

Ne segue che

$$\begin{aligned} \iint_D (2x - y + 3) \, dx \, dy &= \int_0^1 dy \int_y^{\sqrt{y}} (2x - y + 3) \, dx = \int_0^1 [x^2 - xy + 3x]_{x=y}^{x=\sqrt{y}} dy \\ &= \int_0^1 (3\sqrt{y} - 2y - \sqrt{y^3}) \, dy = \frac{3}{5}. \end{aligned}$$

3) Vogliamo calcolare il volume della regione T dello spazio delimitata dal paraboloide di equazione $z = x^2 + y^2$ ed il piano $z = h^2$. Tale volume è dato da

$$\iiint_D (h^2 - x^2 - y^2) \, dx \, dy,$$

dove D è il cerchio di raggio h e centro nell'origine del piano x, y . Considerando D come dominio normale rispetto all'asse x si ha

$$D: \begin{cases} -h \leq x \leq h \\ -\sqrt{h^2 - x^2} \leq y \leq \sqrt{h^2 - x^2}. \end{cases}$$

Si ottiene allora

$$\begin{aligned} \iint_D (h^2 - x^2 - y^2) \, dx \, dy &= \int_{-h}^h dx \int_{-\sqrt{h^2 - x^2}}^{\sqrt{h^2 - x^2}} (h^2 - x^2 - y^2) \, dy \\ &= \int_{-h}^h \left[(h^2 - x^2)y - \frac{y^3}{3} \right]_{-\sqrt{h^2 - x^2}}^{\sqrt{h^2 - x^2}} dx \\ &= \int_{-h}^h \frac{4}{3} (h^2 - x^2) \sqrt{h^2 - x^2} \, dx \end{aligned}$$

Quest'ultimo integrale può essere risolto mediante la sostituzione $x = h \sin t$. Si ottiene

$$\int_{-h}^h \frac{4}{3} (h^2 - x^2) \sqrt{h^2 - x^2} \, dx = \frac{4}{3} h^4 \int_{-\pi/2}^{\pi/2} \cos^4 t \, dt = \frac{h^4}{2} \pi.$$

4) Calcoliamo l'integrale doppio

$$\iint_D |y - x^2| \, dx \, dy, \quad \text{dove } D: \begin{cases} -1 \leq x \leq 1 \\ -1 \leq y \leq 1 \end{cases}.$$

A causa del valore assoluto, dobbiamo distinguere due casi. In D si ha

$$|y - x^2| = \begin{cases} y - x^2 & (x, y) \in D_1: \begin{cases} -1 \leq x \leq 1 \\ x^2 \leq y \leq 1 \end{cases} \\ x^2 - y & (x, y) \in D_2: \begin{cases} -1 \leq x \leq 1 \\ -1 \leq y \leq x^2 \end{cases} \end{cases}$$

Utilizzando la proprietà 2) si ottiene

$$\iint_D |y - x^2| \, dx \, dy = \iint_{D_1} (y - x^2) \, dx \, dy + \iint_{D_2} (x^2 - y) \, dx \, dy.$$

Si ha

$$\begin{aligned} \iint_D |y - x^2| \, dx \, dy &= \int_{-1}^1 dx \int_{x^2}^1 (y - x^2) \, dy + \int_{-1}^1 dx \int_{-1}^{x^2} (x^2 - y) \, dy \\ &= \int_{-1}^1 \left[\frac{y^2}{2} - x^2 y \right]_{x^2}^1 dx + \int_{-1}^1 \left[x^2 y - \frac{y^2}{2} \right]_{-1}^{x^2} dx \\ &= \int_{-1}^1 (x^4 + 1) \, dx = \frac{12}{5}. \end{aligned}$$

5) Calcolare l'integrale doppio

$$\iint_D e^{x+y} \, dx \, dy,$$

dove D è la regione di piano compresa tra i due quadrati

$$D_1 : \begin{cases} -1 \leq x \leq 1 \\ -1 \leq y \leq 1 \end{cases} \quad D_2 : \begin{cases} -2 \leq x \leq 2 \\ -2 \leq y \leq 2 \end{cases}$$

Tenuto conto della proprietà 2) si ha

$$\iint_{D_2} e^{x+y} \, dx \, dy = \iint_{D_1} e^{x+y} \, dx \, dy + \iint_D e^{x+y} \, dx \, dy.$$

Otteniamo così,

$$\begin{aligned} \iint_D e^{x+y} \, dx \, dy &= \iint_{D_2} e^{x+y} \, dx \, dy - \iint_{D_1} e^{x+y} \, dx \, dy \\ &= \int_{-2}^2 dx \int_{-2}^2 e^{x+y} \, dy - \int_{-1}^1 dx \int_{-1}^1 e^{x+y} \, dy \\ &= \int_{-2}^2 e^x \, dx \int_{-2}^2 e^y \, dy - \int_{-1}^1 e^x \, dx \int_{-1}^1 e^y \, dy \\ &= (e^2 - e^{-2})^2 - (e - e^{-1})^2 = 2 \cosh 4 - 2 \cosh 2. \end{aligned}$$

6.4 Cambiamento di variabili negli integrali doppi

Sia $K \subset \mathbb{R}^2$ un dominio limitato con frontiera regolare a tratti e sia $\Phi : K \rightarrow \mathbb{R}^2$ una funzione di componenti $\phi_1(u, v), \phi_2(u, v)$, con $(u, v) \in K$, tale che

$$\Phi \in C^1(K), \quad \det J(u, v) \neq 0, \quad \forall (u, v) \in K,$$

essendo J la matrice jacobiana di Φ . Detta D l'immagine di Φ in \mathbb{R}^2 , poniamo

$$\begin{cases} x = \phi_1(u, v) \\ y = \phi_2(u, v), \end{cases} \quad (u, v) \in K, \quad (x, y) \in D,$$

e supponiamo che queste equazioni mettano in corrispondenza biunivoca ∂K con ∂D . Allora si può dimostrare che le equazioni precedenti stabiliscono una corrispondenza biunivoca tra tutto K e tutto D .

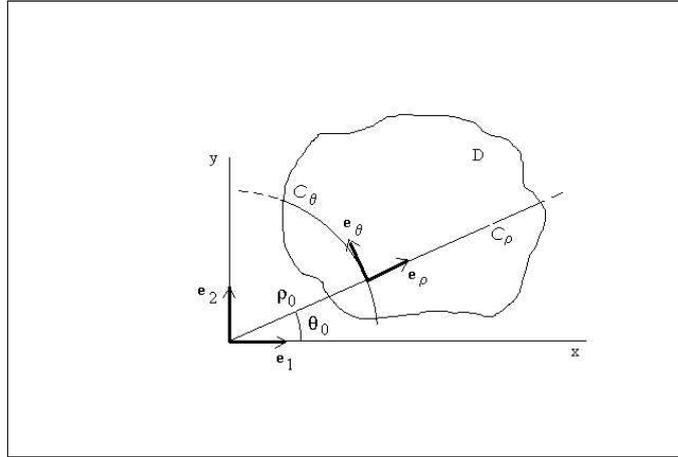
Se x e y sono coordinate cartesiane in D , le funzioni ϕ_1 e ϕ_2 individuano nuove coordinate u, v in \mathbb{R}^2 , che diremo *coordinate curvilinee*. Preso un punto $(u_0, v_0) \in K$ le due curve

$$\mathcal{C}_u : \begin{cases} x = \phi_1(u, v_0) \\ y = \phi_2(u, v_0), \end{cases} \quad \mathcal{C}_v : \begin{cases} x = \phi_1(u_0, v) \\ y = \phi_2(u_0, v), \end{cases} \quad (u, v) \in K,$$

hanno il grafico in D e risultano regolari. Inoltre, i versori tangenti a ciascuna curva in (u_0, v_0) ,

$$\mathbf{e}_u = \frac{\partial \mathbf{x}}{\partial u} \Big|_0 / \left\| \frac{\partial \mathbf{x}}{\partial u} \Big|_0 \right\|, \quad \mathbf{e}_v = \frac{\partial \mathbf{x}}{\partial v} \Big|_0 / \left\| \frac{\partial \mathbf{x}}{\partial v} \Big|_0 \right\|,$$

risultano linearmente indipendenti a causa della ipotesi $\det J \neq 0$. Le curve \mathcal{C}_u e \mathcal{C}_v si dicono *linee coordinate* ed $\mathbf{e}_u, \mathbf{e}_v$, versori del sistema di coordinate u, v . Porremo inoltre $|\det J| = |J|$ e lo chiameremo *jacobiano della trasformazione* dalle coordinate u, v alle coordinate x, y .



Un esempio già noto di coordinate curvilinee nel piano è quello delle coordinate polari, definite da

$$\begin{cases} x = \rho \cos \theta \\ y = \rho \sin \theta, \end{cases} \quad K = \{(\rho, \theta) \in \mathbb{R}^2 : \rho > 0, \theta \in [0, 2\pi[\}.$$

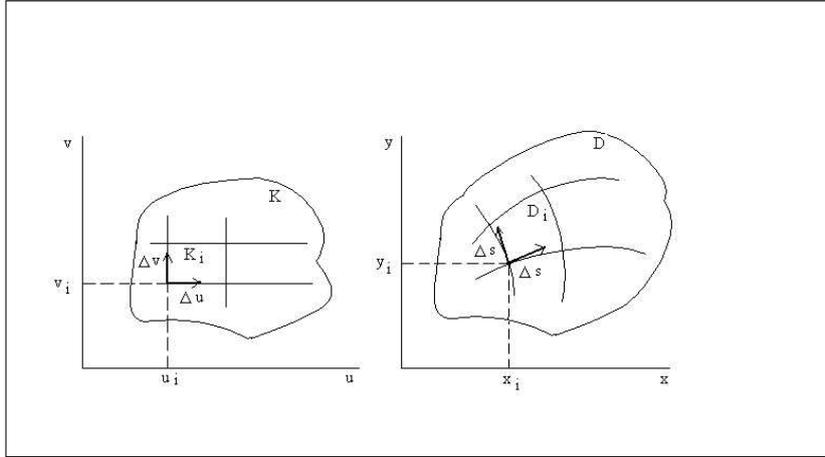
In tal caso, preso un punto $(\rho_0, \theta_0) \in K$, le curve \mathcal{C}_ρ e \mathcal{C}_θ , passanti per (ρ_0, θ_0) sono rispettivamente un segmento di retta per l'origine ed un arco di circonferenza di centro l'origine e raggio ρ_0 . Si ha inoltre

$$|J| = \begin{vmatrix} \frac{\partial x}{\partial \rho} & \frac{\partial x}{\partial \theta} \\ \frac{\partial y}{\partial \rho} & \frac{\partial y}{\partial \theta} \end{vmatrix} = \begin{vmatrix} \cos \theta & -\rho \sin \theta \\ \sin \theta & \rho \cos \theta \end{vmatrix} = \rho > 0, \quad \begin{cases} \mathbf{e}_\rho = \cos \theta \mathbf{e}_1 + \sin \theta \mathbf{e}_2 \\ \mathbf{e}_\theta = -\sin \theta \mathbf{e}_1 + \cos \theta \mathbf{e}_2. \end{cases}$$

Ritornando al caso generale, poiché K è delimitato da una curva regolare a tratti, esso risulta misurabile. La corrispondenza tra ∂K e ∂D e la condizione $\Phi \in C^1(K)$, implicano che anche ∂D è regolare a tratti e quindi D risulta misurabile.

Vogliamo trovare una formula per esprimere la misura di D tramite le nuove coordinate in K . A tale scopo consideriamo l'insieme \mathcal{R}_K dei plurirettangoli contenenti K e l'insieme \mathcal{R}'_K dei plurirettangoli contenuti in K , ottenuti mediante una partizione P . Fissato un plurirettangolo di \mathcal{R}_K consideriamo il suo i -esimo rettangolo K_i e sia (u_i, v_i) un suo spigolo (vedi figura). Tramite la trasformazione $\mathbf{x} = \Phi(u, v)$, a K_i corrisponderà un quadrilatero curvilineo D_i di vertice (x_i, y_i) corrispondente a (u_i, v_i) . Dette Δu e Δv le misure dei lati di K_i si avrà

$$\text{mis}(K_i) = \Delta u \Delta v.$$



I lati corrispondenti di D_i saranno dati dagli incrementi delle ascisse curvilinee su \mathcal{C}_u e \mathcal{C}_v passanti per (u_i, v_i) , ovvero

$$\Delta s_u = \int_{u_i}^{u_i + \Delta u} \left\| \frac{\partial \mathbf{x}}{\partial u} \right\| du, \quad \Delta s_v = \int_{v_i}^{v_i + \Delta v} \left\| \frac{\partial \mathbf{x}}{\partial v} \right\| dv,$$

e la misura di D_i si potrà pensare come l'area di un parallelogramma curvilineo, il cui valore approssimato è dato dalla norma del prodotto vettoriale tra i due vettori tangenti ai lati curvilinei in (x_i, y_i) , di lunghezza pari agli incrementi Δs_u e Δs_v . A meno di infinitesimi di ordine superiore a Δu e Δv possiamo confondere gli incrementi Δs_u e Δs_v delle ascisse curvilinee s_u e s_v con i loro differenziali ottenendo

$$\Delta s_u = \left\| \frac{\partial \mathbf{x}}{\partial u} \right\|_{(u_i, v_i)} \Delta u, \quad \Delta s_v = \left\| \frac{\partial \mathbf{x}}{\partial v} \right\|_{(u_i, v_i)} \Delta v,$$

e di conseguenza ricaviamo

$$\begin{aligned} \text{mis}(D_i) &= \|\Delta s_u \mathbf{e}_u \times \Delta s_v \mathbf{e}_v\| = \left\| \frac{\partial \mathbf{x}}{\partial u} \Big|_{(u_i, v_i)} \times \frac{\partial \mathbf{x}}{\partial v} \Big|_{(u_i, v_i)} \right\| \Delta u \Delta v \\ &= |J(u_i, v_i)| \Delta u \Delta v. \end{aligned}$$

Poiché $\Delta u \Delta v = \text{mis}(K_i)$, avremo

$$\text{mis}(D_i) = \text{mis}(K_i) |J(u_i, v_i)|,$$

da cui

$$\sum_i \text{mis}(D_i) = \sum_i \text{mis}(K_i) |J(u_i, v_i)|.$$

D'altra parte, per ogni partizione P si può scrivere

$$s(|J|, P) \leq \sum_i \text{mis}(K_i) |J(u_i, v_i)|.$$

Per l'integrabilità di $|J|$ in K segue che

$$\inf_P \sum_i \text{mis}(D_i) = \inf_P \sum_i \text{mis}(K_i) |J(u_i, v_i)| \geq \sup_P s(|J|, P) = \iint_K |J(u, v)| du dv.$$

Analogamente, considerando un plurirettangolo di \mathcal{R}'_K contenuto in K e detti K'_i i suoi rettangoli di spigoli (u'_i, v'_i) si ottiene

$$\sum_i \text{mis}(D'_i) = \sum_i \text{mis}(K'_i) |J(u'_i, v'_i)| \leq S(|J|, P),$$

e, sempre per l'integrabilità di $|J|$, si ha

$$\sup_P \sum_i \text{mis}(D'_i) \leq \inf_P S(|J|, P) = \iint_K |J(u, v)| \, du \, dv.$$

Ma per la misurabilità di D ,

$$\inf_P \sum_i \text{mis}(D_i) = \sup_P \sum_i \text{mis}(D'_i),$$

da cui concludiamo

$$\text{mis}(D) = \iint_K |J(u, v)| \, du \, dv.$$

Sapendo che $\text{mis}(D) = \iint_D dx \, dy$, possiamo enunciare il seguente risultato.

Teorema 6.3 *Sia K un dominio limitato con frontiera regolare a tratti e sia $\mathbf{x} = \Phi(u, v)$ una trasformazione di coordinate con $\Phi \in C^1(K)$, a valori in D , con $|J| \neq 0$ in K e tale da garantire una corrispondenza biunivoca tra ∂K e ∂D . Allora*

$$\iint_D dx \, dy = \iint_K |J(u, v)| \, du \, dv.$$

• Esempio

Calcoliamo l'area della superficie della regione D del primo quadrante del piano x, y delimitato dalle due iperboli $xy = a^2$ e $xy = b^2$ con $0 < a < b$ e dalle rette $y = \alpha x$ e $y = \beta x$, con $0 < \alpha < \beta$ (vedi figura). L'area della superficie D è data da $\text{mis}(D) = \iint_D dx \, dy$.

Consideriamo la trasformazione di coordinate

$$xy = u, \quad \frac{y}{x} = v.$$

Il dominio corrispondente è $K = \{(u, v) \in \mathbb{R}^2 : a^2 \leq u \leq b^2, \alpha \leq v \leq \beta\}$. Si può scrivere

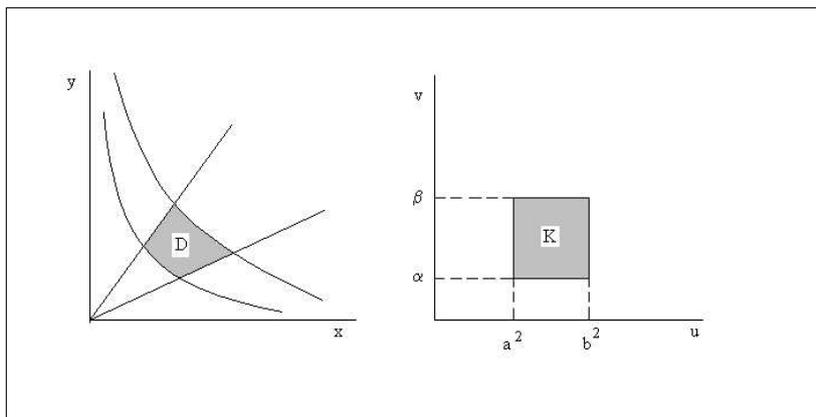
$$\begin{cases} x = \sqrt{\frac{u}{v}} \\ y = \sqrt{uv} \end{cases} \quad (u, v) \in K,$$

e si ha

$$|J| = \begin{vmatrix} \frac{1}{2\sqrt{uv}} & -\frac{\sqrt{u}}{2v\sqrt{v}} \\ \frac{1}{2}\sqrt{\frac{v}{u}} & \frac{1}{2}\sqrt{\frac{u}{v}} \end{vmatrix} = \frac{1}{2v} > 0.$$

Questa trasformazione soddisfa tutte le ipotesi del teorema 6.3. Si ottiene allora

$$\iint_D dx \, dy = \iint_K |J(u, v)| \, du \, dv = \int_{a^2}^{b^2} du \int_{\alpha}^{\beta} \frac{1}{2v} \, dv = \frac{b^2 - a^2}{2} \ln \frac{\beta}{\alpha}.$$



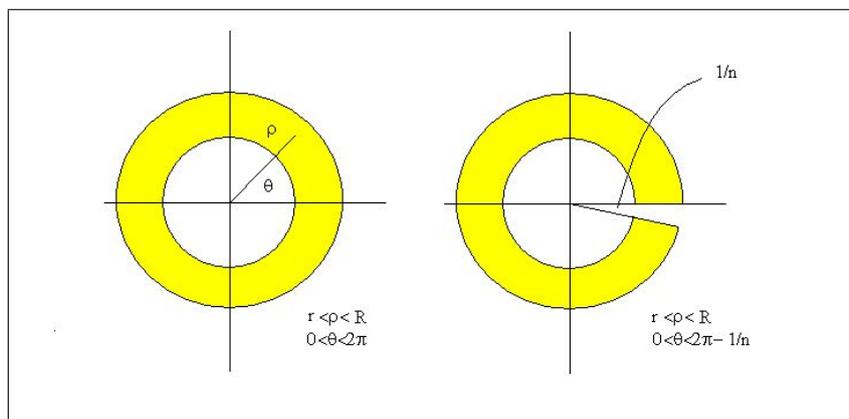
La tesi del teorema 6.3 sussiste anche nel caso in cui le ipotesi non siano soddisfatte in un sottoinsieme di misura nulla di D . Consideriamo ad esempio il dominio costituito dalla corona circolare D di centro l'origine, raggio interno r e raggio esterno R ,

$$D = \{(x, y) \in \mathbb{R}^2 : r^2 \leq x^2 + y^2 \leq R^2\}.$$

Trasformando in coordinate polari si ottiene il dominio

$$K = \{(\rho, \theta) \in \mathbb{R}^2 : r \leq \rho \leq R, \theta \in [0, 2\pi]\}.$$

La frontiera ∂K non è in corrispondenza biunivoca con ∂D in quanto ai valori $\theta = 0$ e $r < \rho < R$ corrispondono i punti interni di D dati da $r < x < R, y = 0$.



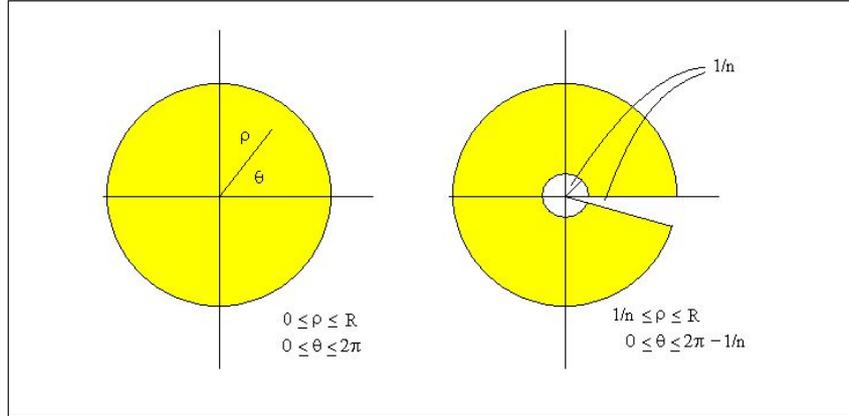
Come altro esempio consideriamo il cerchio D di centro l'origine e raggio R

$$D = \{(x, y) \in \mathbb{R}^2 : 0 \leq x^2 + y^2 \leq R^2\}.$$

Trasformando in coordinate polari, si ha

$$K = \{(\rho, \theta) \in \mathbb{R}^2 : 0 \leq \rho \leq R, \theta \in [0, 2\pi]\},$$

e, per $\rho = 0$, si ha $|J| = 0$.



In ambedue i casi si può pensare D come il limite di una successione di domini per i quali sussistono le ipotesi del teorema 6.3. In particolare, nel caso della corona circolare consideriamo la successione $\{K_n\}$, dove

$$K_n = \left\{ (\rho, \theta) \in \mathbb{R}^2 : r \leq \rho \leq R, \theta \in \left[0, 2\pi - \frac{1}{n} \right] \right\}.$$

Si ha

$$\begin{aligned} \text{mis}(D) &= \lim_{n \rightarrow \infty} \iint_{K_n} \rho \, d\rho d\theta = \lim_{n \rightarrow \infty} \int_0^{2\pi - \frac{1}{n}} d\theta \int_r^R \rho \, d\rho \\ &= \lim_{n \rightarrow \infty} \frac{1}{2} \left(2\pi - \frac{1}{n} \right) (R^2 - r^2) = \pi(R^2 - r^2). \end{aligned}$$

Nel caso del cerchio consideriamo la successione $\{K_n\}$, con

$$K_n = \left\{ (\rho, \theta) \in \mathbb{R}^2 : \frac{1}{n} \leq \rho \leq R, \theta \in \left[0, 2\pi - \frac{1}{n} \right] \right\},$$

e si ha

$$\begin{aligned} \text{mis}(D) &= \lim_{n \rightarrow \infty} \iint_{K_n} \rho \, d\rho d\theta = \lim_{n \rightarrow \infty} \int_0^{2\pi - \frac{1}{n}} d\theta \int_{\frac{1}{n}}^R \rho \, d\rho \\ &= \lim_{n \rightarrow \infty} \frac{1}{2} \left(2\pi - \frac{1}{n} \right) \left(R^2 - \frac{1}{n^2} \right) = \pi R^2. \end{aligned}$$

La conseguenza di questi risultati è che il teorema 6.3 continua ad essere valido se le ipotesi sono soddisfatte per una coppia di successioni di domini $\{K_n\}$ e $\{D_n\}$ e se accade che $\lim_{n \rightarrow \infty} \text{mis}(K_n) = \text{mis}(K)$ e $\lim_{n \rightarrow \infty} \text{mis}(D_n) = \text{mis}(D)$.

Con una dimostrazione analoga a quella del teorema 6.3 si può provare la formula di trasformazione degli integrali doppi per cambiamento di variabili. Sussiste il seguente risultato.

Teorema 6.4 *Sotto le stesse ipotesi del teorema 6.3, data la funzione $f(x, y)$, continua in D , si ha*

$$\iint_D f(x, y) \, dx dy = \iint_K f[\phi_1(u, v), \phi_2(u, v)] |J(u, v)| \, du dv.$$

Dim. Applicando il teorema della media al risultato del teorema 6.3 si ha, per un qualunque dominio misurabile D ,

$$\text{mis}(D) = \iint_K |J(u, v)| \, dudv = |J(u^*, v^*)| \text{mis}(K).$$

dove (u^*, v^*) è un opportuno punto di K . Effettuata allora una partizione di K in sottodomini K_i , e denotati con D_i i corrispondenti domini in D , si può scrivere

$$\text{mis}(D_i) = |J(u_i^*, v_i^*)| \text{mis}(K_i).$$

Supponiamo ora che la funzione f sia non negativa in tutto D . Moltiplicando ambo i membri di questa relazione per $m_i = \inf_{D_i} f(x, y) = \inf_{K_i} f(\phi_1(u, v), \phi_2(u, v))$ e sommando su i si ottiene

$$\begin{aligned} \sum_i m_i \text{mis}(D_i) &= \sum_i \inf_{K_i} f(\phi_1(u, v), \phi_2(u, v)) |J| \text{mis}(K_i) \geq \sum_i \inf_{K_i} [f(\phi_1(u, v), \phi_2(u, v)) |J|] \text{mis}(K_i) \\ &= s(f|J, P). \end{aligned}$$

Analogamente, moltiplicando per $M_i = \sup_{D_i} f(x, y) = \sup_{K_i} f(\phi_1(u, v), \phi_2(u, v))$ e sommando su i si ottiene

$$\begin{aligned} \sum_i M_i \text{mis}(D_i) &= \sum_i \sup_{K_i} f(\phi_1(u, v), \phi_2(u, v)) |J| \text{mis}(K_i) \leq \sum_i \sup_{K_i} [f(\phi_1(u, v), \phi_2(u, v)) |J|] \text{mis}(K_i) \\ &= S(f|J, P). \end{aligned}$$

Prendendo rispettivamente il sup e l'inf dei membri estremi di queste due disuguaglianze al variare della partizione e tenuto conto della integrabilità di $f(x, y)$ in D e di $f[\phi_1(u, v), \phi_2(u, v)]|J(u, v)|$ in K , avremo

$$\begin{aligned} \iint_D f(x, y) \, dxdy &\geq \iint_K f[\phi_1(u, v), \phi_2(u, v)] |J(u, v)| \, dudv, \\ \iint_D f(x, y) \, dxdy &\leq \iint_K f[\phi_1(u, v), \phi_2(u, v)] |J(u, v)| \, dudv, \end{aligned}$$

da cui, necessariamente, segue la tesi. Se la funzione f non fosse non negativa in tutto D , la dimostrazione si si potrebbe ripetere considerando la funzione $g(x, y) = f(x, y) + h$ dove $h = |\inf_D f(x, y)|$, che risulta non negatva in D . Si verifica facilmente la tesi anche in questo caso.

• Esempio

Calcolare l'integrale doppio

$$\iint_D \frac{1}{\sqrt{1+x^2+y^2}} \, dxdy, \quad \text{con } D = \{(x, y) \in \mathbb{R}^+ \times \mathbb{R}^+ : x^2 + y^2 \leq 1\}.$$

Passando a coordinate polari otteniamo

$$K = \left\{ (\rho, \theta) \in \mathbb{R}^2 : 0 \leq \rho \leq 1, 0 \leq \theta \leq \frac{\pi}{2} \right\}.$$

Poiché $|J| = \rho$, si ha $|J| = 0$ per $\rho = 0$. Anche nel caso del teorema 6.4 vale però l'osservazione fatta a proposito della validità del teorema 6.3 e la formula di calcolo per cambiamento di coordinate vale ancora. Si ottiene

$$\iint_D \frac{dxdy}{\sqrt{1+x^2+y^2}} = \iint_K \frac{\rho}{\sqrt{1+\rho^2}} \, d\rho d\theta = \int_0^{\frac{\pi}{2}} d\theta \int_0^1 \frac{\rho \, d\rho}{\sqrt{1+\rho^2}} = \frac{\pi}{2}(\sqrt{2}-1).$$

6.5 Calcolo di aree di superfici

Consideriamo una superficie regolare S , definita da

$$S : \begin{cases} x = \phi_1(u, v) \\ y = \phi_2(u, v) \\ z = \phi_3(u, v), \end{cases} \quad (u, v) \in K,$$

dove $K \subset \mathbb{R}^2$ è chiuso, limitato e connesso e $\Phi \in C^1(K)$. Preso un punto $(u_0, v_0) \in K$, sia \mathbf{x}_0 il punto corrispondente di S . Sappiamo che, fissato $v = v_0$ e facendo variare u le equazioni precedenti descrivono una curva regolare \mathcal{C}_u passante per \mathbf{x}_0 , il cui versore tangente in \mathbf{x}_0 è dato da

$$\mathbf{e}_u = \Phi_u|_{(u_0, v_0)} / \|\Phi_u|_{(u_0, v_0)}\|.$$

Analogamente, fissato $u = u_0$ e facendo variare v si ottiene una curva regolare \mathcal{C}_v per \mathbf{x}_0 , il cui versore tangente è

$$\mathbf{e}_v = \Phi_v|_{(u_0, v_0)} / \|\Phi_v|_{(u_0, v_0)}\|.$$

Per la regolarità di S i parametri u e v costituiscono un sistema di coordinate curvilinee sulla superficie S . In modo analogo a quanto visto per il cambiamento di variabili negli integrali doppi, possiamo fare una partizione P della superficie S in un numero finito di porzioni S_i ($i = 1, \dots, n$), tutte contenute in S e delimitate da archi di curve \mathcal{C}_u e \mathcal{C}_v . Se indichiamo con Δs_u e Δs_v le lunghezze di una coppia di tali archi, l'area approssimata del parallelogramma curvilineo che approssima la superficie della porzione S_i sarà data da

$$A_S(S_i) = \|\Delta s_u \mathbf{e}_u \times \Delta s_v \mathbf{e}_v\| = \|\Phi_u|_{(u_i, v_i)} \times \Phi_v|_{(u_i, v_i)}\| \Delta u \Delta v,$$

essendo Δu e Δv gli incrementi delle variabili in K . Sommando i contributi di tutte le porzioni di superficie, al variare della partizione P su S , otteniamo un insieme di aree che ammette come estremo superiore l'area della superficie di S . In altri termini otteniamo

$$A_S(S) = \sup_P \sum_i A_S(S_i) = \sup_P \sum_i \|\Phi_u \times \Phi_v\| \text{mis}(K_i) = \iint_K \|\Phi_u \times \Phi_v\| \, du \, dv.$$

Se la superficie è data in forma esplicita $z = f(x, y)$ dove f è una funzione continua con derivate parziali continue in un dominio $A \subset \mathbb{R}^2$ chiuso, limitato e connesso, allora, ponendo

$$S : \begin{cases} x = u \\ y = v \\ z = f(u, v), \end{cases} \quad (u, v) \in A.$$

Si ottiene

$$\|\Phi_u \times \Phi_v\| = \sqrt{1 + \left(\frac{\partial f}{\partial u}\right)^2 + \left(\frac{\partial f}{\partial v}\right)^2},$$

e di conseguenza la formula per il calcolo della superficie diventa

$$A_S(S) = \iint_A \sqrt{1 + \left(\frac{\partial f}{\partial x}\right)^2 + \left(\frac{\partial f}{\partial y}\right)^2} \, dx \, dy.$$

• Esempi

1) Calcolare l'area della superficie di una sfera di raggio R .

Le equazioni parametriche della sfera di raggio R sono

$$S \begin{cases} x = R \sin \theta \cos \psi \\ y = R \sin \theta \sin \psi \\ z = R \cos \theta, \end{cases} \quad (\theta, \psi) \in K,$$

dove $K = [0, \pi] \times [0, 2\pi[$. Si ottiene

$$\begin{aligned} \Phi_\theta &= (R \cos \theta \cos \psi, R \cos \theta \sin \psi, -R \sin \theta), \\ \Phi_\psi &= (-R \sin \theta \sin \psi, R \sin \theta \cos \psi, 0), \end{aligned}$$

da cui

$$\|\Phi_\theta \times \Phi_\psi\| = R^2 \sin \theta.$$

Si ricava così,

$$A_S(S) = \int_0^{2\pi} d\psi \int_0^\pi R^2 \sin \theta d\theta = 2\pi R^2 [-\cos \theta]_0^\pi = 4\pi R^2.$$

OSSERVAZIONE. Si deve notare che θ, ψ perdono il requisito di coordinate curvilinee su S per $\theta = 0, \pi$ a causa della non biunivocità della corrispondenza coi punti dell'asse z . Tuttavia, se si considerano i domini $K_n = [\frac{1}{n}, \pi - \frac{1}{n}] \times [0, 2\pi[$, con $n \in \mathbb{N}$ arbitrario, le coordinate θ, ψ sono effettive coordinate curvilinee in S ed inoltre, detta A_n l'area della porzione di S ottenuta per $(\theta, \psi) \in K_n$, si ha $\lim_{n \rightarrow \infty} A_n = A_S(S)$. Ciò giustifica la validità della formula del calcolo dell'area di una superficie anche in tal caso.

2) Calcolare l'area della superficie del cilindro di equazione $y^2 + z^2 = R^2$ contenuta nella regione di spazio individuata da $z \geq 0, |x| \leq y, y \geq 0$.

In questo caso possiamo esprimere la superficie in forma esplicita scrivendo

$$z = f(x, y) = \sqrt{R^2 - y^2}, \quad (x, y) \in A, \quad A = \{(x, y) \in \mathbb{R}^2 : 0 \leq y \leq R, |x| \leq y\}.$$

Poiché si ha

$$\frac{\partial f}{\partial x} = 0, \quad \frac{\partial f}{\partial y} = \frac{-y}{\sqrt{R^2 - y^2}},$$

si ottiene

$$\begin{aligned} \iint_A \sqrt{1 + \frac{y^2}{R^2 - y^2}} dx dy &= \int_0^R \left(\int_{-y}^y \frac{R}{\sqrt{R^2 - y^2}} dx \right) dy \\ &= \int_0^R \frac{2Ry}{\sqrt{R^2 - y^2}} dy = \left[-2R\sqrt{R^2 - y^2} \right]_0^R = 2R^2. \end{aligned}$$

6.6 Integrali di superficie

Sia S una superficie regolare definita da $\Phi : K \subset \mathbb{R}^2 \rightarrow \mathbb{R}^3$ con K limitato, chiuso e connesso e sia $f(\mathbf{x})$ una funzione definita in un dominio $T \subset \mathbb{R}^3$ contenente S , continua in T . Facciamo una partizione di K in un numero finito di domini K_i , non aventi punti interni in comune. A questi K_i corrisponderanno altrettante superfici S_i , prive di punti interni in comune, la cui area può essere approssimata a

$$A_S(S_i) = \|\Phi_u|_{(u_i, v_i)} \times \Phi_v|_{(u_i, v_i)}\| \Delta u \Delta v.$$

Siano m_i e M_i rispettivamente l'estremo inferiore e l'estremo superiore della restrizione di f in S_i . Se al variare della partizione P su S si ha

$$\sup_P \sum_i m_i A_S(S_i) = \inf_P \sum_i M_i A_S(S_i) = I_S,$$

allora si pone

$$I_S = \int_S f(\mathbf{x}) \, dS,$$

che si chiama *integrale di superficie* della funzione f . Sotto le condizioni date, l'integrale di superficie di f esiste e si ha

$$\int_S f(\mathbf{x}) \, dS = \iint_K f[\phi_1(u, v), \phi_2(u, v), \phi_3(u, v)] \|\Phi_u \times \Phi_v\| \, du \, dv.$$

Nel caso in cui la superficie ammetta la rappresentazione cartesiana $z = \phi(x, y)$, si ha

$$\int_S f(\mathbf{x}) \, dS = \iint_A f[x, y, \phi(x, y)] \sqrt{1 + \left(\frac{\partial \phi}{\partial x}\right)^2 + \left(\frac{\partial \phi}{\partial y}\right)^2} \, dx \, dy,$$

dove A è un dominio limitato e connesso.

- Esempio

Data la funzione $f(\mathbf{x}) = x^2 + y^2 + z^2$ e la porzione S di paraboloidi di equazione $z = \frac{1}{2}(x^2 + y^2)$ con $0 \leq z \leq 1$, calcolare l'integrale di superficie di f su S .

Si ha

$$\int_S f(\mathbf{x}) \, dS = \iint_A \left[x^2 + y^2 + \frac{1}{4}(x^2 + y^2)^2 \right] \sqrt{1 + x^2 + y^2} \, dx \, dy,$$

dove $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 2\}$. Passando in coordinate polari si ottiene

$$\int_S f(\mathbf{x}) \, dS = \int_0^{2\pi} d\theta \int_0^{\sqrt{2}} \left(\rho^2 + \frac{1}{4}\rho^4 \right) \sqrt{1 + \rho^2} \rho \, d\rho.$$

Mediante la sostituzione $\sqrt{1 + \rho^2} = t$, si ha $\rho^2 = t^2 - 1$, $\rho \, d\rho = t \, dt$, da cui

$$\int_S f(\mathbf{x}) \, dS = 2\pi \int_1^{\sqrt{3}} \left[t^2 - 1 + \frac{1}{4}(t^2 - 1)^2 \right] t^2 \, dt = \frac{2\pi}{35} (39\sqrt{3} + 4).$$

6.7 Integrali tripli

La definizione di integrale triplo si può introdurre in modo analogo al caso dell'integrale doppio. Si tratta di estendere a domini tridimensionali e a funzioni in \mathbb{R}^3 quanto è stato detto circa gli integrali di funzioni di due variabili in domini di \mathbb{R}^2 .

Sia $T \subset \mathbb{R}^3$ un insieme cubabile ed $f : T \rightarrow \mathbb{R}$ una funzione limitata in T . Effettuando una partizione P di T in un numero finito di insiemi T_i (per esempio dei parallelepipedi) tali che $T = \cup_{i=1}^n T_i$, $\overset{\circ}{T}_i \cap \overset{\circ}{T}_j = \emptyset$, per $i \neq j$, denotiamo con m_i e M_i rispettivamente l'estremo inferiore e l'estremo superiore di $f(\mathbf{x})$ in T_i . Definiamo le somme integrali inferiore e superiore,

$$s(f, P) = \sum_{i=1}^n m_i \text{mis}(T_i), \quad S(f, P) = \sum_{i=1}^n M_i \text{mis}(T_i).$$

Per ogni partizione P si ha $s(f, P) \leq S(f, P)$ ed ogni raffinamento P' di P è tale che

$$s(f, P') \geq s(f, P), \quad S(f, P') \leq S(f, P).$$

Al variare di tutte le possibili partizioni di T si avrà

$$\sup_P [s(f, P)] \leq \inf_P [S(f, P)].$$

Si dice che $f(\mathbf{x})$ è *integrabile* in T se accade che

$$\sup_P [s(f, P)] = \inf_P [S(f, P)] = I_T,$$

e si pone

$$I_T = \iiint_T f(x, y, z) \, dx \, dy \, dz, \quad \text{o anche} \quad I_T = \int_T f(\mathbf{x}) \, d\mathbf{x}.$$

Diversamente dal caso degli integrali doppi, l'integrale I_T non ha, in generale, alcun significato geometrico. Tuttavia, nel caso particolare in cui $f(\mathbf{x}) = 1$ identicamente in T , si ha $m_i = M_i = 1$, per ogni T_i e, in base alla definizione di insieme cubabile, si ha

$$\int_T d\mathbf{x} = \text{mis}(T).$$

In analogia con il caso degli integrali doppi, vale la seguente condizione di integrabilità.

Teorema 6.5 *Sia $f : T \subset \mathbb{R}^3 \rightarrow \mathbb{R}$ con T limitato e misurabile e sia f continua in T eccettuato al più un insieme V_0 di misura nulla, allora f è integrabile in T .*

Si ha anche, come corollario, che se f e g sono continue in T limitato e misurabile e se

$$f(\mathbf{x}) = g(\mathbf{x}), \quad \forall \mathbf{x} \in T \setminus V_0,$$

allora le due funzioni sono integrabili in T e i loro integrali coincidono.

Per gli integrali tripli si possono provare proprietà analoghe a quelle date per gli integrali doppi. Valgono, in particolare, con le dovute modifiche, le proprietà 1) – 5) del paragrafo 6.2.

Per quanto riguarda il calcolo, gli integrali tripli ammettono formule di riduzione analoghe a quelle degli integrali doppi. Dato un insieme D limitato e misurabile e due funzioni $\phi_1(x, y)$ e $\phi_2(x, y)$ continue in D e tali che $\phi_1(x, y) \leq \phi_2(x, y) \forall (x, y) \in D$, l'insieme

$$T = \{\mathbf{x} \in \mathbb{R}^3 : (x, y) \in D, \phi_1(x, y) \leq z \leq \phi_2(x, y)\},$$

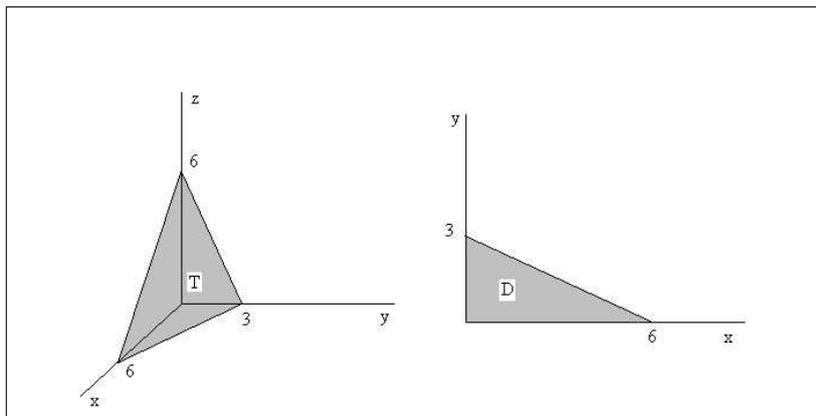
si dice *dominio normale* rispetto al piano x, y . Da quanto detto alla fine del paragrafo 6.1, essendo T la differenza tra due cilindroidi relativi a D , esso risulterà misurabile. Per domini di questo tipo vale la seguente formula di riduzione

$$\iiint_T f(x, y, z) dx dy dz = \iint_D \left(\int_{\phi_1(x, y)}^{\phi_2(x, y)} f(x, y, z) dz \right) dx dy.$$

In questo modo, dopo una prima integrazione su z ci si riduce al calcolo di un integrale doppio. Naturalmente valgono analoghe formule di riduzione per domini normali rispetto agli altri piani coordinati. Decomponendo un generico dominio T misurabile in domini normali, l'uso delle formule di riduzione e della proprietà di additività rispetto al dominio, permette il calcolo di un integrale triplo su T .

• Esempi

- 1) Calcolare il volume del tetraedro T delimitato dai tre piani coordinati e dal piano di equazione $x + 2y + z - 6 = 0$.



Considerato come dominio normale rispetto al piano x, y , si ha

$$T = \{\mathbf{x} \in \mathbb{R}^3 : (x, y) \in D, 0 \leq z \leq 6 - x - 2y\},$$

$$D = \{\mathbf{x} \in \mathbb{R}^2 : 0 \leq x \leq 6, 0 \leq y \leq 3 - x/2\}.$$

Poiché, in questo caso dobbiamo porre $f(\mathbf{x}) = 1$, identicamente, si ha

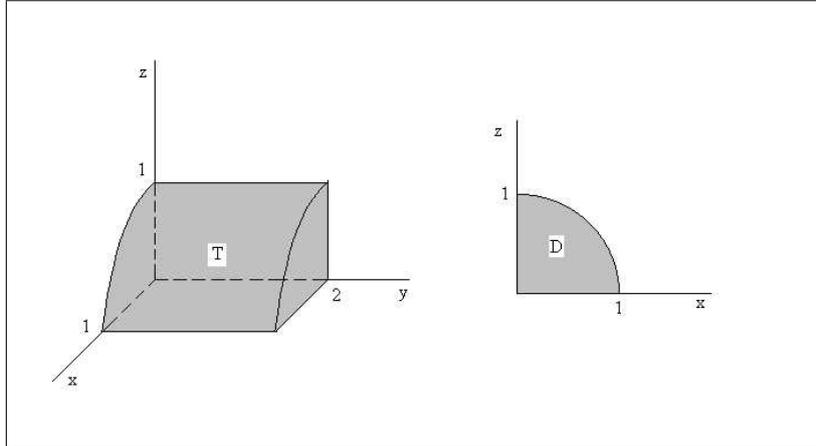
$$\begin{aligned} \text{mis}(T) &= \iint_D \left(\int_0^{6-x-2y} dz \right) dx dy = \int_0^6 \left[\int_0^{3-\frac{x}{2}} \left(\int_0^{6-x-2y} dz \right) dy \right] dx \\ &= \int_0^6 \left[\int_0^{3-\frac{x}{2}} (6-x-2y) dy \right] dx = \int_0^6 [6y - xy - y^2]_0^{3-\frac{x}{2}} dx \\ &= \frac{1}{4} \int_0^6 (6-x)^2 dx = \frac{1}{4} \left[-\frac{(6-x)^3}{3} \right]_0^6 = 18, \end{aligned}$$

in accordo con il risultato ottenibile dalla geometria elementare.

2) Calcolare l'integrale triplo

$$\iiint_T y \, dx \, dy \, dz,$$

dove T è la regione di $\mathbb{R}^+ \times \mathbb{R}^+ \times \mathbb{R}^+$ delimitata dal cilindro di equazione $x^2 + z^2 = 1$, dai piani coordinati e dal piano $y = 2$.



Si può pensare a T come a un dominio normale rispetto al piano x, z , ovvero

$$T = \{\mathbf{x} \in \mathbb{R}^3 : (x, z) \in D, 0 \leq y \leq 2\},$$

$$D = \{(x, z) \in \mathbb{R}^+ \times \mathbb{R}^+ : x^2 + z^2 \leq 1\}$$

Si ottiene così,

$$\iiint_T y \, dx \, dy \, dz = \iint_D \left(\int_0^2 y \, dy \right) dx \, dz = \iint_D 2 \, dx \, dz = 2 \text{mis}(D) = \frac{\pi}{2}.$$

Così come per gli integrali doppi, si possono ricavare formule per il cambiamento di variabili negli integrali tripli. Senza ripetere le dimostrazioni già fatte, riportiamo i risultati essenziali ai fini del calcolo degli integrali tripli.

Osserviamo innanzitutto che le equazioni

$$\begin{cases} x = \phi_1(u, v, w) \\ y = \phi_2(u, v, w) \\ z = \phi_3(u, v, w), \end{cases} \quad (u, v, w) \in K, \quad (x, y, z) \in T,$$

definiscono un sistema di coordinate curvilinee in T se le funzioni $\phi_1, \phi_2, \phi_3 \in C^1(K)$ stabiliscono una corrispondenza biunivoca tra i punti di K e quelli di T e se $\det J \neq 0, \forall (u, v, w) \in K$, essendo J la matrice jacobiana della trasformazione di coordinate,

$$J = \begin{pmatrix} \frac{\partial \phi_1}{\partial u} & \frac{\partial \phi_1}{\partial v} & \frac{\partial \phi_1}{\partial w} \\ \frac{\partial \phi_2}{\partial u} & \frac{\partial \phi_2}{\partial v} & \frac{\partial \phi_2}{\partial w} \\ \frac{\partial \phi_3}{\partial u} & \frac{\partial \phi_3}{\partial v} & \frac{\partial \phi_3}{\partial w} \end{pmatrix}.$$

Fissato un punto $(u_0, v_0, w_0) \in K$ chiameremo linee coordinate passanti per $\mathbf{x}_0 = \Phi(u_0, v_0, w_0)$ le curve regolari

$$\mathcal{C}_u : \begin{cases} x = \phi_1(u, v_0, w_0) \\ y = \phi_2(u, v_0, w_0) \\ z = \phi_3(u, v_0, w_0), \end{cases} \quad \mathcal{C}_v : \begin{cases} x = \phi_1(u_0, v, w_0) \\ y = \phi_2(u_0, v, w_0) \\ z = \phi_3(u_0, v, w_0), \end{cases} \quad \mathcal{C}_w : \begin{cases} x = \phi_1(u_0, v_0, w) \\ y = \phi_2(u_0, v_0, w) \\ z = \phi_3(u_0, v_0, w), \end{cases}$$

e denoteremo con $\mathbf{e}_u, \mathbf{e}_v, \mathbf{e}_w$ i loro versori tangenti in \mathbf{x}_0 ,

$$\mathbf{e}_u = \frac{\partial \mathbf{x}}{\partial u} \Big|_0 / \left\| \frac{\partial \mathbf{x}}{\partial u} \Big|_0 \right\|, \quad \mathbf{e}_v = \frac{\partial \mathbf{x}}{\partial v} \Big|_0 / \left\| \frac{\partial \mathbf{x}}{\partial v} \Big|_0 \right\|, \quad \mathbf{e}_w = \frac{\partial \mathbf{x}}{\partial w} \Big|_0 / \left\| \frac{\partial \mathbf{x}}{\partial w} \Big|_0 \right\|.$$

Le coordinate sferiche e le coordinate cilindriche, sono esempi di coordinate curvilinee in \mathbb{R}^3 . Nel primo caso si ha

$$\begin{cases} x = r \sin \theta \cos \psi \\ y = r \sin \theta \sin \psi \\ z = r \cos \theta, \end{cases} \quad K = \mathbb{R}^{++} \times (0, \pi) \times [0, 2\pi),$$

$$|J| = \begin{vmatrix} \sin \theta \cos \psi & r \cos \theta \cos \psi & -r \sin \theta \sin \psi \\ \sin \theta \sin \psi & r \cos \theta \sin \psi & r \sin \theta \cos \psi \\ \cos \theta & -r \sin \theta & 0 \end{vmatrix} = r^2 \sin \theta,$$

$$\begin{aligned} \mathbf{e}_r &= (\sin \theta \cos \psi, \sin \theta \sin \psi, \cos \theta), \\ \mathbf{e}_\theta &= (\cos \theta \cos \psi, \cos \theta \sin \psi, -\sin \theta), \\ \mathbf{e}_\psi &= (-\sin \psi, \cos \psi, 0). \end{aligned}$$

Le curve \mathcal{C}_r sono archi di rette per l'origine, le curve \mathcal{C}_θ sono archi di circonferenze di centro l'origine e con z asse diametrale, le curve \mathcal{C}_ψ sono archi di circonferenze su piani paralleli al piano x, y e con centro sull'asse z .

Nel secondo caso si ha

$$\begin{cases} x = r \cos \psi \\ y = r \sin \psi \\ z = z, \end{cases} \quad K = \mathbb{R}^{++} \times [0, 2\pi) \times \mathbb{R},$$

$$|J| = \begin{vmatrix} \cos \psi & -r \sin \psi & 0 \\ \sin \psi & r \cos \psi & 0 \\ 0 & 0 & 1 \end{vmatrix} = r$$

$$\begin{aligned} \mathbf{e}_r &= (\cos \psi, \sin \psi, 0) \\ \mathbf{e}_\psi &= (-\sin \psi, \cos \psi, 0) \\ \mathbf{e}_z &= (0, 0, 1). \end{aligned}$$

Le curve \mathcal{C}_r sono archi di rette che intersecano ortogonalmente l'asse z , le curve \mathcal{C}_ψ sono archi di circonferenze su piani paralleli al piano x, y e con centro sull'asse z , le curve \mathcal{C}_z sono archi di rette parallele all'asse z .

Dato un dominio misurabile $T \subset \mathbb{R}^3$, introdotte coordinate curvilinee $(u, v, w) \in K$ mediante la trasformazione $\mathbf{x} = \Phi(u, v, w)$, ($\mathbf{x} \in T$), e denotato con $|J|$ il determinante della matrice jacobiana, si dimostra che

$$\text{mis}(T) = \iiint_K |J(u, v, w)| \, dudvdw,$$

da cui segue la formula

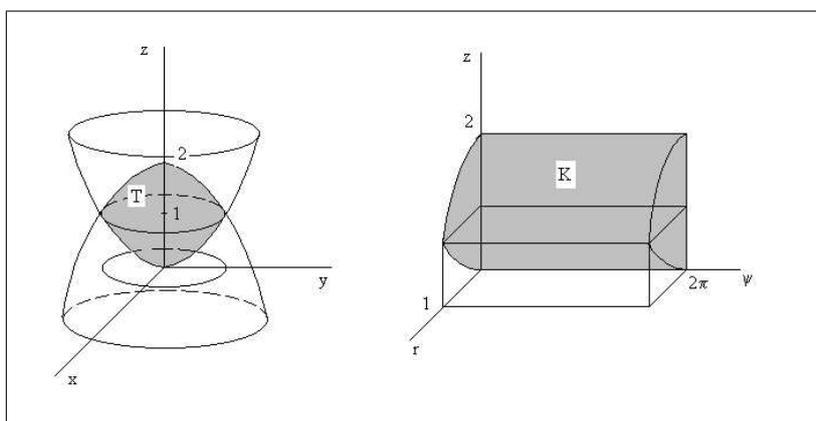
$$\int_T d\mathbf{x} = \iiint_K |J(u, v, w)| du dv dw.$$

Inoltre, data una funzione $f(\mathbf{x})$ continua nell'insieme misurabile $T \subset \mathbb{R}^3$, si ha la seguente formula di trasformazione degli integrali tripli per cambiamento di variabili

$$\int_T f(\mathbf{x}) d\mathbf{x} = \iiint_K f[\phi_1(u, v, w), \phi_2(u, v, w), \phi_3(u, v, w)] |J(u, v, w)| du dv dw.$$

• Esempi

- 1) Calcolare il volume del solido delimitato dai due paraboloidi di equazioni $z = x^2 + y^2$ e $z = 2 - x^2 - y^2$.



Indicato con T il dominio occupato dal solido, si ha

$$T = \{\mathbf{x} \in \mathbb{R}^3 : x^2 + y^2 \leq z \leq 2 - x^2 - y^2\}.$$

La limitazione su z implica che $x^2 + y^2 \leq 1$. Utilizzando coordinate cilindriche si ottiene il dominio trasformato K , dato da

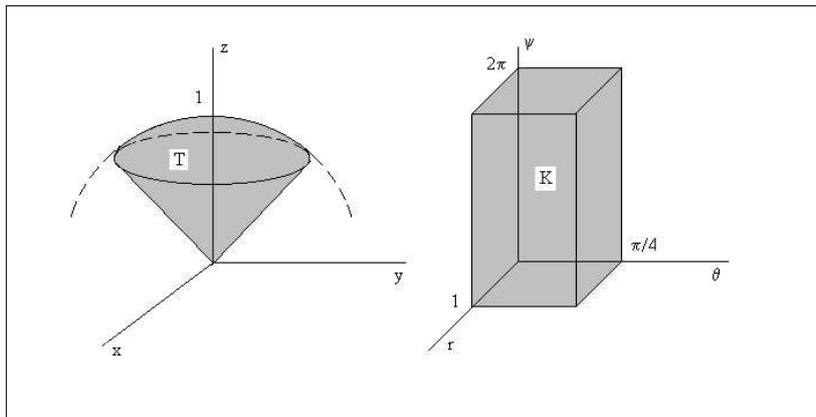
$$K = \{(r, \psi, z) \in \mathbb{R}^3 : 0 \leq r \leq 1, 0 \leq \psi \leq 2\pi, r^2 \leq z \leq 2 - r^2\}$$

Considerando quest'ultimo come un dominio normale rispetto al piano r, ψ , essendo $|J| = r$, si ottiene

$$\begin{aligned} \text{mis}(T) &= \iiint_K r dr d\psi dz \\ &= \int_0^{2\pi} d\psi \int_0^1 \left(\int_{r^2}^{2-r^2} r dz \right) dr = 2\pi \int_0^1 (2r - 2r^3) dr = \pi. \end{aligned}$$

- 2) Calcolare l'integrale triplo

$$\iiint_T (x^2 + yz) dx dy dz,$$



dove T è la regione di \mathbb{R}^3 delimitata dalla sfera di centro l'origine e raggio 1 e dalla falda superiore del cono di equazione $x^2 + y^2 - z^2 = 0$.

Utilizzando coordinate sferiche il dominio T viene trasformato nel dominio K dato da

$$K = \{(r, \theta, \psi) \in \mathbb{R}^3 : 0 \leq r \leq 1, 0 \leq \theta \leq \pi/4, 0 \leq \psi \leq 2\pi\}$$

Tenuto conto che $|J| = r^2 \sin \theta$, si ottiene così

$$\begin{aligned} & \iiint_T (x^2 + yz) \, dx \, dy \, dz \\ &= \int_0^{\pi/4} \left(\int_0^{2\pi} \left(\int_0^1 (r^2 \sin^2 \theta \cos^2 \psi + r^2 \sin \theta \cos \theta \sin \psi) r^2 \sin \theta \, dr \right) d\psi \right) d\theta \\ &= \int_0^{\pi/4} \left(\int_0^{2\pi} \left(\frac{1}{5} \sin^3 \theta \cos^2 \psi + \frac{1}{5} \sin^2 \theta \cos \theta \sin \psi \right) d\psi \right) d\theta \\ &= \int_0^{\pi/4} \frac{\pi}{5} \sin^3 \theta \, d\theta = \frac{\pi}{15} \left(2 - \frac{5\sqrt{2}}{4} \right). \end{aligned}$$

6.8 Applicazioni degli integrali doppi e tripli

Una importante applicazione degli integrali di funzioni di più variabili nasce dal più semplice modello matematico dei corpi fisici. Dato un corpo \mathcal{B} assumeremo che esso occupi una regione di spazio $T \subset \mathbb{R}^3$ chiusa, limitata e misurabile. Nel dominio T definiamo una funzione positiva e continua $\mu : T \rightarrow \mathbb{R}$ detta *densità di massa*. Definiamo *massa* di \mathcal{B} la quantità

$$m(\mathcal{B}) = \int_T \mu(\mathbf{x}) \, d\mathbf{x}.$$

Tale definizione soddisfa le proprietà elementari del concetto fisico di massa di un corpo, quali la positività e l'additività.

Si definisce *centro di massa* o *baricentro* di \mathcal{B} il punto $\mathbf{x}_G \in \mathbb{R}^3$, dato da

$$\mathbf{x}_G(\mathcal{B}) = \frac{\int_T \mu(\mathbf{x}) \mathbf{x} \, d\mathbf{x}}{\int_T \mu(\mathbf{x}) \, d\mathbf{x}} = \frac{1}{m(\mathcal{B})} \int_T \mu(\mathbf{x}) \mathbf{x} \, d\mathbf{x},$$

dove, in componenti, la notazione precedente ha il significato seguente,

$$\begin{aligned} x_G &= \frac{1}{m} \iiint_T \mu(x, y, z) x \, dx dy dz, \\ y_G &= \frac{1}{m} \iiint_T \mu(x, y, z) y \, dx dy dz, \\ z_G &= \frac{1}{m} \iiint_T \mu(x, y, z) z \, dx dy dz. \end{aligned}$$

Se una delle tre dimensioni spaziali del corpo ha misura trascurabile rispetto alle altre, si può definire una funzione *densità superficiale di massa* $\mu : S \rightarrow \mathbb{R}$ dove S è la superficie occupata dal corpo. La massa di \mathcal{B} sarà data allora dall'integrale di superficie

$$m(\mathcal{B}) = \int_S \mu(\mathbf{x}) \, dS.$$

Nel caso piano, detto D il dominio occupato dal corpo, si avrà $\mu : D \rightarrow \mathbb{R}$ e

$$m(\mathcal{B}) = \iint_D \mu(x, y) \, dx dy.$$

Il baricentro di un corpo piano sarà dato da

$$x_G = \frac{1}{m} \iint_D \mu(x, y) x \, dx dy, \quad y_G = \frac{1}{m} \iint_D \mu(x, y) y \, dx dy.$$

Se poi il corpo è omogeneo, la densità sarà una costante e la massa del corpo sarà data dalla misura della regione di spazio o di superficie moltiplicata per tale costante. In questo caso il baricentro è dato da

$$\mathbf{x}_G(\mathcal{B}) = \frac{1}{\text{mis}(T)} \iiint_T \mathbf{x} \, dx dy dz, \quad \mathbf{x}_G(\mathcal{B}) = \frac{1}{\text{mis}(D)} \iint_D \mathbf{x} \, dx dy$$

rispettivamente nei casi tridimensionale e bidimensionale.

L'additività rispetto al dominio di integrazione implica che se il corpo \mathcal{B} viene suddiviso in n parti \mathcal{B}_i , $i = 1, \dots, n$, e se di ciascuna parte viene calcolato il baricentro $\mathbf{x}_G^{(i)}$, allora il baricentro dell'intero corpo \mathcal{B} sarà dato da

$$\mathbf{x}_G = \frac{1}{m(\mathcal{B})} \iiint_T \mu \mathbf{x} \, dx dy dz = \frac{1}{m(\mathcal{B})} \sum_{i=1}^n \iiint_{T_i} \mu \mathbf{x} \, dx dy dz = \frac{1}{m(\mathcal{B})} \sum_{i=1}^n m(\mathcal{B}_i) \mathbf{x}_G^{(i)}.$$

• Esempi

1) Calcolare la massa e il baricentro del cilindro $T = \{\mathbf{x} \in \mathbb{R}^3 : x^2 + y^2 \leq R^2, 0 \leq z \leq z_0\}$ di densità $\mu(\mathbf{x}) = \mu_0 \frac{z_0}{z+z_0}$ in T .

Dalla definizione, facendo uso di coordinate cilindriche, si ha

$$m = \int_T \mu(\mathbf{x}) \, d\mathbf{x} = \mu_0 \int_0^{2\pi} d\psi \int_0^R r \, dr \int_0^{z_0} \frac{z_0}{z+z_0} \, dz = \mu_0 R^2 z_0 \pi \ln 2.$$

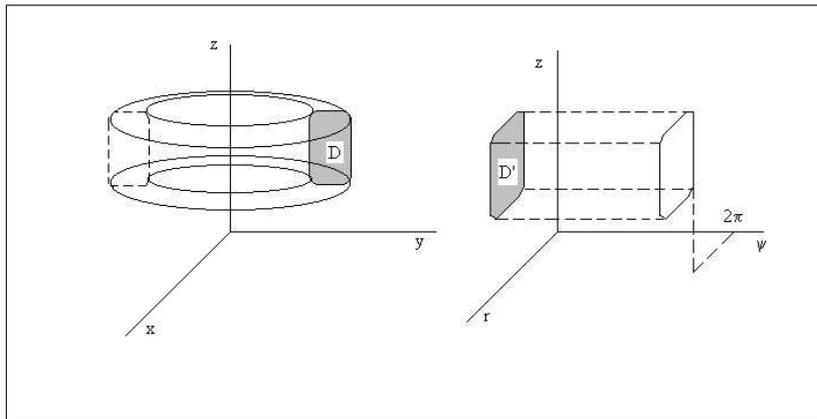
Per la simmetria del problema si ha $x_G = y_G = 0$, mentre

$$\begin{aligned} z_G &= \frac{1}{m} \iiint_T \mu(x, y, z) z \, dx dy dz = \frac{1}{R^2 z_0 \pi \ln 2} \int_0^{2\pi} d\psi \int_0^R r \, dr \int_0^{z_0} \frac{z z_0}{z + z_0} dz \\ &= \frac{1}{\ln 2} [z - z_0 \ln |z + z_0|]_0^{z_0} = z_0 \frac{1 - \ln 2}{\ln 2}. \end{aligned}$$

2) Calcolare il baricentro di una lamina omogenea a forma di semicerchio D di raggio R . Consideriamo un riferimento sulla lamina piana con origine nel centro del semicerchio e asse x lungo il suo diametro. Notando che $x_G = 0$ e usando coordinate polari, avremo

$$y_G = \frac{1}{\text{mis}(D)} \iint_D y \, dx dy = \frac{2}{\pi R^2} \int_0^\pi \sin \theta \, d\theta \int_0^R r^2 \, dr = \frac{4R}{3\pi}.$$

Il concetto di baricentro può essere sfruttato per il calcolo del volume dei solidi di rotazione. Consideriamo un dominio D limitato e misurabile nel piano y, z di un riferimento nello spazio tridimensionale e assumiamo $y \geq 0$, $\forall (y, z) \in D$. Facendo ruotare D attorno all'asse z di un angolo pari a 2π radianti, si genera un solido di rotazione T . Il dominio T si può trasformare tramite coordinate cilindriche in



$$K = \{(r, \psi, z) \in \mathbb{R}^3 : (r, z) \in D', 0 \leq \psi \leq 2\pi\}.$$

Questo dominio risulta normale rispetto al piano r, z e si ha dunque

$$\text{mis}(T) = \iiint_T dx dy dz = \int_0^{2\pi} d\psi \iint_{D'} r \, dr dz.$$

Poiché la coordinata r in D' coincide con y in D , si ottiene

$$\iint_{D'} r \, dr dz = \iint_D y \, dy dz.$$

In base alla definizione di baricentro di un corpo piano omogeneo, si ha

$$\text{mis}(T) = \int_0^{2\pi} d\psi \iint_D y \, dy dz = 2\pi \text{mis}(D) y_G.$$

Questo risultato è anche conosciuto come *secondo teorema di Guldino*.

- Esempio

Calcolare il volume di un toro T ottenuto ruotando la circonferenza di centro $C = (0, y_0, 0)$ e raggio $R (\leq y_0)$.

Il dominio D che genera il toro ha misura pari a πR^2 e baricentro nel centro C della circonferenza. Di conseguenza l'applicazione diretta del primo teorema di Guldino fornisce

$$\text{mis}(T) = 2\pi(\pi R^2)y_0 = 2\pi^2 R^2 y_0.$$

Consideriamo un corpo \mathcal{B} che occupa un dominio T nello spazio riferito ad un sistema di assi cartesiani x, y, z . Detta $\mu(x, y, z)$ la densità di massa del corpo, si definisce *momento di inerzia* di \mathcal{B} rispetto all'asse x , la quantità

$$I_x(\mathcal{B}) = \iiint_T \mu(x, y, z)(y^2 + z^2) dx dy dz.$$

Analogamente, si definiscono i momenti di inerzia di \mathcal{B} rispetto all'asse y e all'asse z , le quantità

$$I_y(\mathcal{B}) = \iiint_T \mu(x, y, z)(x^2 + z^2) dx dy dz, \quad I_z(\mathcal{B}) = \iiint_T \mu(x, y, z)(x^2 + y^2) dx dy dz$$

Più in generale, considerata una retta a nello spazio, si definisce momento di inerzia del corpo \mathcal{B} rispetto all'asse a la quantità

$$I_a(\mathcal{B}) = \int_T \mu(\mathbf{x}) d^2(\mathbf{x}, a) d\mathbf{x},$$

dove $d(\mathbf{x}, a)$ è la funzione distanza del generico punto $\mathbf{x} \in T$ dalla retta a . I_a viene anche detto *momento d'inerzia assiale* di \mathcal{B} rispetto ad a . La definizione precedente vale anche nel caso di un corpo bidimensionale che occupa un dominio $D \in \mathbb{R}^2$, essendo a una retta del piano o dello spazio.

- Esempio

Calcolare il momento d'inerzia assiale di una lamina omogenea di massa m a forma di cerchio D di raggio R e centro in $(0, 0)$ rispetto all'asse di equazione $y = x - a$.

La densità (costante) della lamina sarà $\mu_0 = \frac{m}{\pi R^2}$. Poiché la distanza del generico punto del piano dalla retta è

$$d = \frac{|x - y - a|}{\sqrt{2}},$$

utilizzando coordinate polari, si ottiene,

$$\begin{aligned} I_a &= \iint_D \mu_0 \frac{(x - y - a)^2}{2} dx dy = \frac{m}{\pi R^2} \int_0^{2\pi} d\theta \int_0^R \frac{r}{2} (r \cos \theta - r \sin \theta - a)^2 dr \\ &= \frac{m}{2\pi R^2} \int_0^{2\pi} \left[\frac{R^4}{4} + a^2 \frac{R^2}{2} - \frac{1}{2} R^4 \sin \theta \cos \theta + 2a \frac{R^3}{3} (\sin \theta - \cos \theta) \right] d\theta \\ &= \frac{m}{2\pi R^2} \int_0^{2\pi} \left[\frac{R^4}{4} + a^2 \frac{R^2}{2} \right] d\theta = \frac{m}{4} (R^2 + 2a^2). \end{aligned}$$

Una tra le più importanti proprietà del momento d'inerzia assiale di un corpo è data dal *teorema di Huygens* secondo il quale, date due rette parallele a e a' , distanti tra loro d , e tali che a' passi per il baricentro del corpo \mathcal{B} si ha

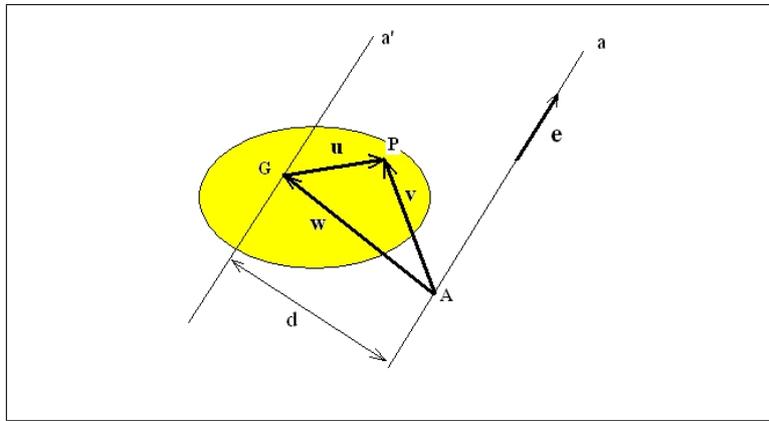
$$I_a(\mathcal{B}) = I_{a'}(\mathcal{B}) + m(\mathcal{B}) d^2.$$

Per dimostrare questo risultato consideriamo un punto A sull'asse a e denotiamo con \mathbf{x}_A e \mathbf{x}_G rispettivamente i vettori posizione di A e di G e poniamo

$$\mathbf{u} = \mathbf{x} - \mathbf{x}_G, \quad \mathbf{v} = \mathbf{x} - \mathbf{x}_A, \quad \mathbf{w} = \mathbf{x}_G - \mathbf{x}_A,$$

dove \mathbf{x} è il vettore posizione del generico punto P di \mathcal{B} . Indicato poi con \mathbf{e} il versore dei due assi a e a' , osserviamo che

$$d(\mathbf{x}, a) = \|\mathbf{v} \times \mathbf{e}\|, \quad d(\mathbf{x}, a') = \|\mathbf{u} \times \mathbf{e}\|, \quad d = \|\mathbf{w} \times \mathbf{e}\|.$$



Poiché $\mathbf{v} = \mathbf{u} + \mathbf{w}$, si ha

$$\begin{aligned} d^2(\mathbf{x}, a) &= [(\mathbf{u} + \mathbf{w}) \times \mathbf{e}] \cdot [(\mathbf{u} + \mathbf{w}) \times \mathbf{e}] \\ &= (\mathbf{u} \times \mathbf{e}) \cdot (\mathbf{u} \times \mathbf{e}) + (\mathbf{w} \times \mathbf{e}) \cdot (\mathbf{w} \times \mathbf{e}) + 2(\mathbf{u} \times \mathbf{e}) \cdot (\mathbf{w} \times \mathbf{e}) \\ &= d^2(\mathbf{x}, a') + d^2 + 2\mathbf{u} \cdot [\mathbf{e} \times (\mathbf{w} \times \mathbf{e})] \end{aligned}$$

dove, nell'ultimo termine si è sfruttata una proprietà del prodotto misto tra vettori. Per le proprietà degli integrali si ottiene allora

$$\begin{aligned} &\int_T \mu(\mathbf{x}) d^2(\mathbf{x}, a) d\mathbf{x} \\ &= \int_T \mu(\mathbf{x}) d^2(\mathbf{x}, a') d\mathbf{x} + d^2 \int_T \mu(\mathbf{x}) d\mathbf{x} + 2[\mathbf{e} \times (\mathbf{w} \times \mathbf{e})] \cdot \int_T \mu(\mathbf{x}) \mathbf{u} d\mathbf{x} \end{aligned}$$

Il secondo integrale a secondo membro non è altro che la massa del corpo \mathcal{B} mentre, facendo uso della definizione di baricentro e di massa, dall'ultimo integrale a secondo membro otteniamo

$$\int_T \mu(\mathbf{x}) \mathbf{u} d\mathbf{x} = \int_T \mu(\mathbf{x}) (\mathbf{x} - \mathbf{x}_G) d\mathbf{x} = m(\mathcal{B}) \mathbf{x}_G - \mathbf{x}_G \int_T \mu(\mathbf{x}) d\mathbf{x} = \mathbf{0}.$$

In definitiva, in base alla definizione di momento d'inerzia assiale si ricava la tesi del teorema. Dal teorema di Huygens si deduce anche che il momento d'inerzia assiale di un corpo rispetto ad una retta di data direzione è minimo quando la retta passa per il baricentro del corpo.

6.9 Derivazione sotto il segno di integrale

Sia $f(x, y)$ una funzione continua in $D = [a, b] \times [c, d]$. Allora la funzione della sola x

$$\int_c^d f(x, y) \, dy$$

risulta continua in $[a, b]$. Infatti, poiché $f(x, \bar{y})$ è continua per ogni $\bar{y} \in [c, d]$, fissato un $\varepsilon > 0$ e scelto un $x_0 \in [a, b]$, si ha $|f(x, \bar{y}) - f(x_0, \bar{y})| < \varepsilon$ per ogni x tale che $|x - x_0| < \delta(\varepsilon, \bar{y})$. Per le proprietà degli integrali, si ha

$$H(x, x_0) = \left| \int_c^d f(x, y) \, dy - \int_c^d f(x_0, y) \, dy \right| \leq \left| \int_c^d |f(x, y) - f(x_0, y)| \, dy \right|.$$

Allora, considerato $\delta_\varepsilon = \min_{\bar{y} \in [c, d]} \delta(\varepsilon, \bar{y})$, si ottiene

$$H(x, x_0) < (d - c)\varepsilon, \quad \forall x : |x - x_0| < \delta_\varepsilon.$$

In altri termini si ha $\lim_{x \rightarrow x_0} H(x, x_0) = 0$, ovvero, la continuità di $\int_c^d f(x, y) \, dy$.

Fatta questa premessa, vogliamo dimostrare che, se la funzione f ammette anche derivata parziale rispetto a x continua in D , allora sussiste la seguente formula di *derivazione sotto il segno di integrale*,

$$\frac{d}{dx} \int_c^d f(x, y) \, dy = \int_c^d \frac{\partial f}{\partial x}(x, y) \, dy, \quad \forall x \in [a, b].$$

A tale scopo osserviamo che, se si considera f come funzione della sola x si ha

$$\int_a^x \frac{\partial f}{\partial x}(\tau, y) \, d\tau = f(\tau, y)|_a^x = f(x, y) - f(a, y),$$

da cui

$$f(x, y) = \int_a^x \frac{\partial f}{\partial x}(\tau, y) \, d\tau + f(a, y).$$

Integrando rispetto a y nell'intervallo $[c, d]$, abbiamo

$$\int_c^d f(x, y) \, dy = \int_c^d \int_a^x \frac{\partial f}{\partial x}(\tau, y) \, d\tau \, dy + \int_c^d f(a, y) \, dy.$$

Ma nell'integrale doppio a secondo membro si può scambiare l'ordine di integrazione e quindi si ha

$$\int_c^d f(x, y) \, dy = \int_a^x \left(\int_c^d \frac{\partial f}{\partial x}(\tau, y) \, dy \right) d\tau + \int_c^d f(a, y) \, dy.$$

Per la continuità di $\partial f / \partial x$, la funzione di τ

$$\int_c^d \frac{\partial f}{\partial x}(\tau, y) \, dy,$$

risulta continua e quindi, dalla equazione precedente, derivando rispetto a x e applicando il teorema fondamentale del calcolo integrale, si ottiene

$$\frac{d}{dx} \int_c^d f(x, y) \, dy = \int_c^d \frac{\partial f}{\partial x}(x, y) \, dy,$$

che è quanto volevamo dimostrare.

Consideriamo ora la funzione

$$F(x, y, z) = \int_y^z f(x, \tau) d\tau,$$

dove f è continua con derivata $\partial f/\partial x$ continua. Facendo ancora uso del teorema fondamentale del calcolo integrale, si ottiene

$$\frac{\partial F}{\partial y}(x, y, z) = -f(x, y), \quad \frac{\partial F}{\partial z}(x, y, z) = f(x, z).$$

Applichiamo questi risultati al caso in cui $y = \alpha(x)$, $z = \beta(x)$ sono funzioni derivabili e riguardiamo la funzione

$$F(x, \alpha(x), \beta(x)) = \int_{\alpha(x)}^{\beta(x)} f(x, \tau) d\tau,$$

come funzione composta. Derivando rispetto a x e facendo uso della regola di derivazione delle funzioni composte, otteniamo,

$$\begin{aligned} \frac{d}{dx} F[x, \alpha(x), \beta(x)] &= \frac{\partial F}{\partial x}[x, \alpha(x), \beta(x)] \\ &+ \frac{\partial F}{\partial y}[x, \alpha(x), \beta(x)]\alpha'(x) + \frac{\partial F}{\partial z}[x, \alpha(x), \beta(x)]\beta'(x), \end{aligned}$$

ovvero

$$\frac{d}{dx} \int_{\alpha(x)}^{\beta(x)} f(x, \tau) d\tau = \int_{\alpha(x)}^{\beta(x)} \frac{\partial f}{\partial x}(x, \tau) d\tau + \beta'(x)f[x, \beta(x)] - \alpha'(x)f[x, \alpha(x)].$$

Quest'ultima formula generalizza la precedente regola di derivazione sotto il segno di integrale.

• Esempio

Calcolare la seguente derivata di integrale,

$$\frac{d}{dx} \int_{\sqrt{x}}^{x+3} (2x^2 - y) dy,$$

prima in modo diretto e poi usando la regola di derivazione sotto segno di integrale.

Nel primo modo si ha

$$\frac{d}{dx} \left\{ \left[2x^2 y - \frac{y^2}{2} \right]_{\sqrt{x}}^{x+3} \right\} = \frac{d}{dx} \left[2x^3 - 2x^2\sqrt{x} + \frac{11}{2}x^2 - \frac{5}{2}x - \frac{9}{2} \right] = 6x^2 - 5x\sqrt{x} + 11x - \frac{5}{2}.$$

Nel secondo caso, osserviamo che sono soddisfatte le condizioni per l'applicabilità della regola di derivazione sotto segno di integrale e otteniamo

$$\begin{aligned} \frac{d}{dx} \int_{\sqrt{x}}^{x+3} (2x^2 - y) dy &= \int_{\sqrt{x}}^{x+3} 4x dy + 2x^2 - x - 3 - \frac{1}{2\sqrt{x}}(2x^2 - \sqrt{x}) \\ &= 4x(x+3 - \sqrt{x}) + 2x^2 - x - 3 - x\sqrt{x} + \frac{1}{2} = 6x^2 - 5x\sqrt{x} + 11x - \frac{5}{2}. \end{aligned}$$

Capitolo 7

Integrali impropri

La teoria degli integrali definiti per le funzioni di una variabile era basata sull'ipotesi che l'intervallo di integrazione fosse finito e che la funzione da integrare fosse limitata. In alcuni casi è possibile definire integrali estesi a domini non limitati oppure integrali di funzioni non limitate.

7.1 Integrali estesi ad intervalli non limitati

Sia $f(x)$ una funzione definita nell'intervallo $[a, +\infty)$ e supponiamo che per ogni $b > a$ essa risulti integrabile in $[a, b]$. Vogliamo dare senso alla scrittura

$$\int_a^{+\infty} f(x) dx,$$

detto *integrale improprio* di f in $[a, +\infty)$. Introdotta la funzione $I(b) = \int_a^b f(x) dx$ consideriamo il limite di $I(b)$ per $b \rightarrow +\infty$. Se tale limite esiste ed è finito, porremo

$$\int_a^{+\infty} f(x) dx = \lim_{b \rightarrow +\infty} \int_a^b f(x) dx,$$

e diremo che l'integrale improprio di f converge. Se il suddetto limite non esiste, o non è finito, diremo che l'integrale non converge (o diverge).

- Esempi

1) Dato l'integrale improprio

$$I = \int_0^{+\infty} \frac{1}{1+x^2} dx,$$

si ha

$$\int_0^b \frac{1}{1+x^2} dx = \arctan b,$$

da cui

$$\lim_{b \rightarrow +\infty} \int_0^b \frac{1}{1+x^2} dx = \frac{\pi}{2},$$

quindi l'integrale improprio I converge.

2) Dato l'integrale improprio

$$I = \int_1^{+\infty} \frac{x^2+1}{x^3} dx,$$

si ha

$$\int_1^b \frac{x^2 + 1}{x^3} dx = \ln b - \frac{1}{2b^2} + \frac{1}{2}.$$

Ne segue che

$$\lim_{b \rightarrow +\infty} \int_1^b \frac{x^2 + 1}{x^3} dx = +\infty.$$

quindi I è divergente.

3) Dato l'integrale improprio

$$I = \int_0^{+\infty} \cos x dx,$$

si ha

$$\int_0^b \cos x dx = \sin b,$$

da cui

$$\lim_{b \rightarrow +\infty} \int_0^b \cos x dx = \lim_{b \rightarrow +\infty} \sin b.$$

Ma quest'ultimo limite non esiste quindi I non è convergente.

4) Consideriamo l'integrale improprio

$$I = \int_a^{+\infty} \frac{1}{x^\alpha} dx,$$

con $a \in \mathbb{R}^{++}$ e studiamo il

$$\lim_{b \rightarrow +\infty} \int_a^b \frac{1}{x^\alpha} dx,$$

al variare di α in \mathbb{R} . Osserviamo innanzitutto che per $\alpha = 1$ l'integrale diverge. Consideriamo allora $\alpha \neq 1$. Si ha

$$\int_a^b \frac{1}{x^\alpha} dx = \frac{b^{1-\alpha}}{1-\alpha} - \frac{a^{1-\alpha}}{1-\alpha},$$

da cui,

$$\lim_{b \rightarrow +\infty} \int_a^b \frac{1}{x^\alpha} dx = \begin{cases} \frac{a^{1-\alpha}}{\alpha-1}, & \text{se } \alpha > 1 \\ +\infty, & \text{se } \alpha < 1. \end{cases}$$

Ne segue che I converge per $\alpha > 1$ e diverge per $\alpha \leq 1$.

5) Sia $a > 1$ e $\alpha \in \mathbb{R}$. Studiamo il comportamento dell'integrale improprio

$$I = \int_a^{+\infty} \frac{1}{x(\ln x)^\alpha} dx.$$

Se $\alpha \neq 1$ abbiamo

$$I = \lim_{b \rightarrow +\infty} \left[\frac{(\ln x)^{1-\alpha}}{1-\alpha} \right]_a^b = \begin{cases} \frac{(\ln a)^{1-\alpha}}{\alpha-1} & \alpha > 1 \\ +\infty & \alpha < 1 \end{cases}$$

Se invece $\alpha = 1$ si ottiene

$$I = \lim_{b \rightarrow +\infty} [\ln(\ln x)]_a^b = +\infty.$$

Concludiamo quindi

$$I = \begin{cases} \frac{(\ln a)^{1-\alpha}}{\alpha-1} & \alpha > 1 \\ +\infty & \alpha \leq 1 \end{cases}$$

Dalla definizione di integrale improprio segue che se i due integrali $\int_a^{+\infty} f(x) dx$ e $\int_a^{+\infty} g(x) dx$ convergono, allora, per ogni coppia di numeri λ e μ , l'integrale improprio

$$\int_a^{+\infty} [\lambda f(x) + \mu g(x)] dx,$$

converge. La dimostrazione di questa proprietà segue dall'osservare che

$$\int_a^b [\lambda f(x) + \mu g(x)] dx = \lambda \int_a^b f(x) dx + \mu \int_a^b g(x) dx.$$

Poiché i due integrali a secondo membro ammettono limite per $b \rightarrow +\infty$, anche il primo membro ammetterà limite e si avrà

$$\int_a^{+\infty} [\lambda f(x) + \mu g(x)] dx = \lambda \int_a^{+\infty} f(x) dx + \mu \int_a^{+\infty} g(x) dx.$$

Per gli integrali impropri vale anche la regola di integrazione per parti. Se f e g sono funzioni con derivate continue in $[a, +\infty)$, sussiste la formula

$$\int_a^{+\infty} f(x)g'(x) dx = [f(x)g(x)]_a^{+\infty} - \int_a^{+\infty} f'(x)g(x) dx,$$

purchè due dei tre termini presenti in questa formula abbiano senso.

• Esempio

Calcolare, se risulta convergente, il seguente integrale improprio

$$I_n = \int_0^{+\infty} x^n e^{-x} dx,$$

essendo $n \in \mathbb{N}$. Mediante una integrazione per parti si ha

$$\int_0^b x^n e^{-x} dx = [-x^n e^{-x}]_0^b + n \int_0^b x^{n-1} e^{-x} dx.$$

Passando al limite per $b \rightarrow +\infty$ si ottiene

$$I_n = nI_{n-1}.$$

Riapplicando successive integrazioni per parti ricaviamo

$$I_n = n(n-1)(n-2)\dots 2 \cdot 1 I_0 = n! I_0.$$

L'integrale I_n risulta allora convergente se lo è l'integrale I_0 . Poiché

$$I_0 = \lim_{b \rightarrow +\infty} \int_0^b e^{-x} dx = \lim_{b \rightarrow +\infty} [-e^{-x}]_0^b = 1,$$

concludiamo che I_n è convergente e vale

$$I_n = n!.$$

Non sempre è possibile stabilire se un integrale improprio è convergente o meno, facendo uso della definizione. È tuttavia possibile dimostrare alcuni criteri sufficienti per la convergenza degli integrali impropri. Considereremo qui i più importanti.

Teorema 7.1 (*Criterio del confronto*) Siano $f(x)$ e $g(x)$ due funzioni integrabili in $[a, b]$, per qualunque $b > a$ e supponiamo che $0 \leq f(x) \leq g(x)$ per ogni $x \geq a$. Allora, se $\int_a^{+\infty} g(x) dx$ converge, anche $\int_a^{+\infty} f(x) dx$ converge, mentre se $\int_a^{+\infty} f(x) dx$ diverge, allora diverge anche $\int_a^{+\infty} g(x) dx$.

Dim. Supponiamo che $\int_a^{+\infty} g(x) dx$ converga. Posto $I(b) = \int_a^b f(x) dx$, $J(b) = \int_a^b g(x) dx$, da una nota proprietà degli integrali, per $b > a$ si ha

$$0 \leq I(b) \leq J(b).$$

Inoltre, essendo $f(x) \geq 0$ e $g(x) \geq 0$ per $x \geq a$, le funzioni $I(b)$ e $J(b)$ risultano non decrescenti in $[a, +\infty)$. D'altra parte, poiché $J(b)$ ammette limite finito per $b \rightarrow +\infty$, dalla disuguaglianza precedente risulta che $I(b)$ è limitata superiormente. Ne segue che $I(b)$ converge per $b \rightarrow +\infty$.

La seconda parte del teorema segue immediatamente dalla prima. Infatti, se per ipotesi $\int_a^{+\infty} f(x) dx$ diverge, $\int_a^{+\infty} g(x) dx$ non può convergere senza contraddire l'ipotesi.

• Esempio

Studiare la convergenza dell'integrale

$$I = \int_0^{+\infty} \frac{\tanh x}{1 + x^2 + \cos^2 x} dx.$$

Poiché si ha

$$0 \leq \frac{\tanh x}{1 + x^2 + \cos^2 x} \leq \frac{1}{1 + x^2}, \quad \forall x \in [0, +\infty),$$

ed essendo $\int_0^{+\infty} \frac{1}{1+x^2} dx$ convergente, in forza del teorema 7.1, si ottiene la convergenza di I .

Teorema 7.2 (*Criterio del confronto asintotico*) Siano $f(x)$ e $g(x)$ integrabili in $[a, b]$, per qualunque b e non negative in $[a, +\infty)$. Supponiamo che esista un \bar{x} tale che, per $x > \bar{x}$ sia $g(x) > 0$ e supponiamo inoltre che esista il limite

$$\lim_{x \rightarrow +\infty} \frac{f(x)}{g(x)}.$$

Allora se tale limite è finito e diverso da zero, la convergenza (o la divergenza) di $\int_a^{+\infty} f(x) dx$ implica la convergenza (o la divergenza) di $\int_a^{+\infty} g(x) dx$ e viceversa. Se il limite è nullo la convergenza dell'integrale di g implica la convergenza di quello di f mentre la divergenza dell'integrale di f implica la divergenza di quello di g . Infine se il limite è infinito, la divergenza dell'integrale di g implica la divergenza di quello di f e la convergenza dell'integrale di f implica la convergenza dell'integrale di g .

Dim. Se il limite del rapporto f/g è $k > 0$, dalla definizione di limite segue che, $\forall \varepsilon > 0 \exists x_\varepsilon : k - \varepsilon < \frac{f(x)}{g(x)} < k + \varepsilon, \forall x > x_\varepsilon$. Preso $x > x_1 = \max(x_\varepsilon, \bar{x})$ si ha $g(x)$ strettamente positiva, quindi scegliendo $\varepsilon < k$, si può scrivere

$$0 < g(x)(k - \varepsilon) < f(x) < g(x)(k + \varepsilon).$$

Se $\int_{x_1}^{+\infty} g(x) dx$ converge, convergerà anche $\int_{x_1}^{+\infty} (k + \varepsilon)g(x) dx$ e, per il teorema 7.1, converge anche $\int_{x_1}^{+\infty} f(x) dx$. Viceversa, se $\int_{x_1}^{+\infty} f(x) dx$ converge, allora $\int_{x_1}^{+\infty} (k - \varepsilon)g(x) dx$ converge e quindi $\int_{x_1}^{+\infty} g(x) dx$ converge. Analogamente, la divergenza di uno dei due integrali implica la divergenza dell'altro.

Se il limite del rapporto f/g fosse nullo, allora si avrebbe

$$0 \leq \frac{f(x)}{g(x)} < \varepsilon \quad \text{per } x > x_\varepsilon,$$

da cui, se $x > \max(x_\varepsilon, \bar{x}) = x_1$, si ha $0 \leq f(x) < \varepsilon g(x)$. Utilizzando ancora il criterio del confronto, concludiamo che la convergenza di $\int_{x_1}^{+\infty} g(x) dx$ implica la convergenza di $\int_{x_1}^{+\infty} f(x) dx$ mentre la divergenza di $\int_{x_1}^{+\infty} f(x) dx$ implica la divergenza di $\int_{x_1}^{+\infty} g(x) dx$.

Se poi il limite di f/g fosse $+\infty$, esisterebbe un M tale che, per $x > x_M$,

$$f(x) > M g(x).$$

Ancora il teorema del confronto ci permette di concludere che la convergenza di $\int_{x_M}^{+\infty} f(x) dx$ implica la convergenza di $\int_{x_M}^{+\infty} g(x) dx$ e la divergenza di $\int_{x_M}^{+\infty} g(x) dx$ implica la divergenza di $\int_{x_M}^{+\infty} f(x) dx$.

Notiamo infine che se, per $x_1 > a$, l'integrale $\int_{x_1}^{+\infty} f(x) dx$ converge (diverge), allora anche $\int_a^{+\infty} f(x) dx$ converge (diverge) in quanto i due integrali differiscono per un numero finito (lo stesso vale ovviamente per $x_M > a$).

- Esempio

Stabilire la natura dell'integrale improprio

$$\int_1^{+\infty} \frac{2x^2 + 1}{x^3 + 3x + 4} dx.$$

Osserviamo che l'integrando è un infinitesimo dello stesso ordine di $\frac{1}{x}$ per $x \rightarrow +\infty$, e si ha

$$\lim_{x \rightarrow +\infty} \frac{\frac{2x^2+1}{x^3+3x+4}}{\frac{1}{x}} = 2 > 0.$$

Dal teorema 7.2 segue che l'integrale di partenza si comporta come l'integrale improprio $\int_1^{+\infty} \frac{1}{x} dx$, il quale è divergente.

Dal teorema del confronto si può ottenere una condizione necessaria per la convergenza di cui omettiamo la dimostrazione. Essa stabilisce che se $f(x)$ ammette limite per $x \rightarrow +\infty$, allora vale l'implicazione

$$\int_a^{+\infty} f(x) dx \text{ convergente} \implies \lim_{x \rightarrow +\infty} f(x) = 0.$$

Sempre facendo uso del teorema 7.1, si può provare facilmente il seguente risultato.

Teorema 7.3 Sia $f(x)$ integrabile in $[a, b]$ per qualunque $b > a > 0$. Supponiamo che esista un $\bar{x} > a$ tale che $\forall x > \bar{x}$,

$$0 \leq f(x) \leq \frac{M}{x^\alpha},$$

essendo $M > 0$ e $\alpha > 1$. Allora $\int_a^{+\infty} f(x) dx$ converge. Se invece, per $x > \bar{x}$ si ha

$$f(x) \geq \frac{M}{x},$$

essendo sempre $M > 0$, allora $\int_a^{+\infty} f(x) dx$ diverge.

Dim. La dimostrazione segue direttamente dal teorema del confronto e da quanto visto sull'integrale improprio $\int_a^{+\infty} \frac{1}{x^\alpha} dx$ per $\alpha > 1$ e per $\alpha = 1$.

Osserviamo ancora che i risultati precedenti ci permettono di stabilire la convergenza o la divergenza dell'integrale $\int_a^{+\infty} f(x) dx$, studiando il comportamento asintotico di f . In particolare, se f non è un infinitesimo per $x \rightarrow \infty$, l'integrale non potrà convergere, mentre se $f(x)$ è un infinitesimo per $x \rightarrow \infty$ allora la convergenza dipenderà dall'ordine di infinitesimo. Se questo è maggiore di 1 avremo convergenza, se invece è minore o uguale a 1 l'integrale divergerà.

• Esempi

1) Studiare la natura dell'integrale

$$I = \int_3^{+\infty} \frac{x-2}{x^3+x^2+2x+5} dx.$$

Poiché si può scrivere

$$\frac{x-2}{x^3+x^2+2x+5} = \frac{x-2}{x^2(x-2)+3x^2+2x+5} < \frac{1}{x^2},$$

per il teorema 7.3 si ha la convergenza di I .

2) Stabilire la natura dell'integrale improprio

$$\int_1^{+\infty} \left(1 - \cos \frac{1}{x}\right) dx.$$

La funzione integranda è un infinitesimo per $x \rightarrow +\infty$. Calcolandone l'ordine troviamo che

$$\lim_{x \rightarrow +\infty} \frac{1 - \cos \frac{1}{x}}{\frac{1}{x^\alpha}} = \frac{1}{2}, \quad \text{per } \alpha = 2.$$

Ne segue, per $x > \bar{x} > 1$,

$$\frac{1}{2} - \varepsilon < x^2 \left(1 - \cos \frac{1}{x}\right) < \frac{1}{2} + \varepsilon \quad \implies \quad \left(1 - \cos \frac{1}{x}\right) < \frac{\frac{1}{2} + \varepsilon}{x^2}.$$

Per il teorema 7.3, otteniamo quindi la convergenza dell'integrale.

3) Stabilire la natura dell'integrale

$$\int_0^{+\infty} \left(\frac{\pi}{2} - \arctan x\right) dx.$$

Anche in questo caso la funzione integranda è un infinitesimo per $x \rightarrow +\infty$. Poiché si verifica facilmente che

$$\lim_{x \rightarrow +\infty} \frac{\frac{\pi}{2} - \arctan x}{\frac{1}{x}} = 1,$$

ricaviamo, per $x > \bar{x}$,

$$\frac{1}{x}(1 - \varepsilon) < \frac{\pi}{2} - \arctan x < \frac{1}{x}(1 + \varepsilon)$$

da cui, sempre per il teorema 7.3, deduciamo la divergenza dell'integrale improprio.

Si dice che l'integrale improprio $\int_a^{+\infty} f(x) dx$ è *assolutamente convergente* se è convergente l'integrale $\int_a^{+\infty} |f(x)| dx$. Si dimostra che ogni integrale improprio assolutamente convergente è convergente mentre, in generale, non vale il viceversa.

• Esempi

1) Verificare che l'integrale

$$I = \int_1^{+\infty} \frac{\cos x}{x^2} dx,$$

è assolutamente convergente.

Poiché si ha $|\frac{\cos x}{x^2}| \leq \frac{1}{x^2}$, per ogni $x \in \mathbb{R}$, mediante il teorema 7.3 concludiamo che I è assolutamente convergente.

2) Verificare che l'integrale improprio

$$I = \int_1^{+\infty} \frac{\sin x}{x} dx,$$

è convergente ma non è assolutamente convergente.

Mediante una integrazione per parti si ha

$$\int_1^{+\infty} \frac{\sin x}{x} dx = \left[-\frac{\cos x}{x} \right]_1^{+\infty} - \int_1^{+\infty} \frac{\cos x}{x^2} dx = \cos 1 - \int_1^{+\infty} \frac{\cos x}{x^2} dx.$$

Poiché l'integrale all'ultimo membro è assolutamente convergente, esso sarà anche convergente e quindi I è convergente. Si può verificare in modo simile che $\int_1^{+\infty} \frac{\cos x}{x} dx$ converge. Consideriamo ora l'integrale

$$\bar{I} = \int_1^{+\infty} \frac{|\sin x|}{x} dx.$$

Poiché si ha $|\sin x| \geq \sin^2 x = \frac{1 - \cos 2x}{2}$, si ottiene

$$\int_1^b \frac{|\sin x|}{x} dx \geq \int_1^b \frac{1 - \cos 2x}{2x} dx = \int_1^b \frac{1}{2x} dx - \int_1^b \frac{\cos 2x}{2x} dx.$$

Prendendo il limite per $b \rightarrow +\infty$ si riconosce che il primo integrale a secondo membro diverge mentre il secondo converge. Quindi \bar{I} diverge.

7.2 Integrali di funzioni non limitate

Consideriamo una funzione $f(x)$ definita in $[a, b]$ e integrabile in $[a, b - \epsilon]$ essendo ϵ un numero positivo piccolo a piacere. Supponiamo che $f(x)$ sia non limitata in $(b - \epsilon, b)$. Chiameremo

$$I = \int_a^b f(x) dx,$$

integrale improprio di $f(x)$ in $[a, b]$. Introdotta la funzione di ϵ

$$I(\epsilon) = \int_a^{b-\epsilon} f(x) dx,$$

diremo che l'integrale improprio I è convergente se esiste ed è finito il

$$\lim_{\epsilon \rightarrow 0^+} I(\epsilon) = \lim_{\epsilon \rightarrow 0^+} \int_a^{b-\epsilon} f(x) dx.$$

Se tale limite non è finito, diremo che I è divergente, se il limite non esiste diremo che l'integrale non converge. Analogamente, se $f(x)$ definita in $(a, b]$ è integrabile in $[a + \epsilon, b]$ e non limitata in $(a, a + \epsilon)$, diremo che I è convergente se esiste ed è finito il

$$\lim_{\epsilon \rightarrow 0^+} \int_{a+\epsilon}^b f(x) dx,$$

altrimenti diremo che I è divergente. Si può anche avere il caso in cui $f(x)$ sia definita in $[a, b]$ escluso un punto interno c e sia $f(x)$ integrabile in $[a, c - \epsilon]$ e in $[c + \epsilon, b]$, mentre risulti non limitata in $(c - \epsilon, c + \epsilon)$. In tal caso si pone

$$\int_a^b f(x) dx = \lim_{\epsilon \rightarrow 0^+} \left[\int_a^{c-\epsilon} f(x) dx + \int_{c+\epsilon}^b f(x) dx \right],$$

qualora il limite esista e sia finito.

- Esempi

1) Consideriamo l'integrale improprio

$$\int_1^2 \frac{1}{\sqrt{x-1}} dx.$$

Posto $I(\epsilon) = \int_{1+\epsilon}^2 \frac{1}{\sqrt{x-1}} dx$, si ha $I(\epsilon) = [2\sqrt{x-1}]_{1+\epsilon}^2 = 2 - 2\sqrt{\epsilon}$, da cui, passando al limite per $\epsilon \rightarrow 0^+$, otteniamo

$$\int_1^2 \frac{1}{\sqrt{x-1}} dx = 2.$$

Osserviamo che questo integrale improprio può essere anche ricondotto ad un integrale ordinario mediante la sostituzione $\sqrt{x-1} = t$. Si ha infatti,

$$\int_1^2 \frac{1}{\sqrt{x-1}} dx = \int_0^1 \frac{2t}{t} dt = \int_0^1 2 dt = 2.$$

2) Studiare la convergenza dell'integrale improprio

$$I = \int_0^b \frac{1}{x^\alpha} dx,$$

al variare di $\alpha \in \mathbb{R}$. Dalla definizione, per $\alpha \neq 1$, abbiamo

$$\int_0^b \frac{1}{x^\alpha} dx = \lim_{\epsilon \rightarrow 0^+} \int_\epsilon^b \frac{1}{x^\alpha} dx = \lim_{\epsilon \rightarrow 0^+} \left[\frac{x^{1-\alpha}}{1-\alpha} \right]_\epsilon^b = \lim_{\epsilon \rightarrow 0^+} \left[\frac{b^{1-\alpha}}{1-\alpha} - \frac{\epsilon^{1-\alpha}}{1-\alpha} \right].$$

Tale limite è finito se $\alpha < 1$ mentre è infinito se $\alpha > 1$. Se $\alpha = 1$ si ha

$$\int_\epsilon^b \frac{1}{x} dx = [\ln x]_\epsilon^b = \ln b - \ln \epsilon.$$

Per $\epsilon \rightarrow 0^+$ si ottiene un limite infinito. Concludiamo che I è convergente per $\alpha < 1$ e divergente per $\alpha \geq 1$.

3) Sia $0 < b < 1$ e $\alpha \in \mathbb{R}$. Studiamo l'integrale improprio

$$I = \int_0^b \frac{1}{x |\ln x|^\alpha} dx.$$

Per $\alpha \neq 1$ abbiamo

$$I = \lim_{\epsilon \rightarrow 0^+} \left[\frac{(\ln \frac{1}{x})^{1-\alpha}}{\alpha-1} \right]_\epsilon^b = \begin{cases} \frac{(\ln \frac{1}{b})^{1-\alpha}}{\alpha-1} & \alpha > 1 \\ +\infty & \alpha < 1 \end{cases}$$

Per $\alpha = 1$ si ha poi

$$I = \lim_{\epsilon \rightarrow 0^+} \left[-\ln \left(\ln \frac{1}{x} \right) \right]_\epsilon^b = +\infty.$$

Concludiamo quindi

$$I = \begin{cases} \frac{(\ln \frac{1}{b})^{1-\alpha}}{\alpha-1} & \alpha > 1 \\ +\infty & \alpha \leq 1 \end{cases}$$

Come nel caso degli integrali estesi ad intervalli non limitati, si possono ricavare alcune condizioni sufficienti per la convergenza (o la divergenza) degli integrali di funzioni non limitate. Enunceremo tali risultati senza dimostrarli.

Teorema 7.4 (Criterio del confronto) *Siano $f(x)$ e $g(x)$ integrabili in $[a, b - \epsilon]$ e non limitate in $(b - \epsilon, b)$ e tali che $0 \leq f(x) \leq g(x)$ in $[a, b)$. Allora la convergenza di $\int_a^b g(x) dx$ implica la convergenza di $\int_a^b f(x) dx$ e la divergenza di $\int_a^b f(x) dx$ implica la divergenza di $\int_a^b g(x) dx$.*

Teorema 7.5 (Criterio del confronto asintotico) Siano $f(x)$ e $g(x)$ positive e integrabili in $[a, b - \epsilon]$, non limitate in $(b - \epsilon, b)$. Supponiamo inoltre che esista il limite

$$\lim_{x \rightarrow b^-} \frac{f(x)}{g(x)}.$$

Allora se tale limite è finito e diverso da zero, la convergenza (o la divergenza) di $\int_a^b f(x) dx$ implica la convergenza (o la divergenza) di $\int_a^b g(x) dx$ e viceversa. Se il limite è nullo la convergenza dell'integrale di g implica la convergenza di quello di f mentre la divergenza dell'integrale di f implica la divergenza di quello di g . Infine se il limite è infinito, la divergenza dell'integrale di g implica la divergenza di quello di f e la convergenza dell'integrale di f implica la convergenza dell'integrale di g .

Si dice che l'integrale improprio $\int_a^b f(x) dx$ è assolutamente convergente se l'integrale $\int_a^b |f(x)| dx$ è convergente. Si dimostra anche in questo caso che se un integrale improprio è assolutamente convergente allora è convergente.

Teorema 7.6 Sia $f(x)$ continua in $[a, b)$ e non limitata in $(b - \epsilon, b)$. Se esiste un $M > 0$ ed un $\alpha < 1$ tali che

$$|f(x)| \leq \frac{M}{(b-x)^\alpha}, \quad \forall x \in (b - \epsilon, b),$$

allora l'integrale $\int_a^b f(x) dx$ è assolutamente convergente. Se invece esiste un $M > 0$ ed un $\alpha \geq 1$ tali che

$$|f(x)| \geq \frac{M}{(b-x)^\alpha}, \quad \forall x \in (b - \epsilon, b),$$

allora l'integrale $\int_a^b f(x) dx$ è divergente.

• Esempi

1) Studiare la convergenza dell'integrale improprio

$$\int_0^1 \ln x dx.$$

Poiché il limite

$$\lim_{x \rightarrow 0^+} \left| \frac{\ln x}{\frac{1}{x^\alpha}} \right| = \lim_{x \rightarrow 0^+} \left| \frac{\frac{1}{x}}{-\alpha x^{-\alpha-1}} \right| = \lim_{x \rightarrow 0^+} \left| \frac{x^\alpha}{\alpha} \right|,$$

risulta nullo per $0 < \alpha < 1$, concludiamo, per il confronto asintotico, che l'integrale dato è assolutamente convergente.

2) Studiare la convergenza dell'integrale improprio

$$\int_0^1 \ln(\tan x) dx.$$

Osserviamo che in un intorno destro dello zero la tangente è maggiore del suo argomento. Quindi per $0 < \epsilon < \pi/4$,

$$|\ln(\tan x)| < |\ln x| \quad \forall x \in (0, \epsilon)$$

In base all'esempio precedente, tramite il criterio del confronto, concludiamo che l'integrale dato è assolutamente convergente.

3) Stabilire la natura del seguente integrale improprio (di Fresnel),

$$I = \int_0^{+\infty} \sin(x^2) dx.$$

Osserviamo che mediante la sostituzione $x^2 = t$ questo integrale può essere trasformato in un integrale improprio di una funzione non limitata. Si ottiene infatti $dx = \frac{1}{2\sqrt{t}} dt$,

$$I = \int_0^{+\infty} \frac{\sin t}{2\sqrt{t}} dt$$

Studiamo questo integrale. Si ha

$$I = \int_0^1 \frac{\sin t}{2\sqrt{t}} dt + \int_1^{\infty} \frac{\sin t}{2\sqrt{t}} dt.$$

Poiché

$$\left| \frac{\sin t}{2\sqrt{t}} \right| \leq \frac{1}{2\sqrt{t}},$$

il primo integrale risulta assolutamente convergente. Si ha poi, mediante una integrazione per parti,

$$\int_1^{\infty} \frac{\sin t}{2\sqrt{t}} dt = \left[\frac{-\cos t}{2\sqrt{t}} \right]_1^{\infty} - \frac{1}{4} \int_1^{\infty} \frac{\cos t}{t^{3/2}} dt = \frac{1}{2} \cos 1 - \frac{1}{4} \int_1^{\infty} \frac{\cos t}{t^{3/2}} dt.$$

D'altra parte,

$$\left| \frac{\cos t}{t^{3/2}} \right| \leq \frac{1}{t^{3/2}},$$

quindi il secondo integrale è convergente. L'integrale di Fresnel è quindi convergente. Un risultato analogo vale per l'integrale

$$\int_0^{+\infty} \cos(x^2) dx.$$

7.3 Integrali impropri di funzioni di più variabili

Il concetto di integrazione in domini non limitati o di funzioni non limitate si può estendere al caso di funzioni di più variabili. Per semplicità consideriamo qui il caso delle funzioni di due variabili. Sia $D \subseteq \mathbb{R}^2$ un dominio non necessariamente limitato ed $f(x, y)$ una funzione non necessariamente limitata in D . Vogliamo dare un senso alla scrittura

$$\iint_D f(x, y) dx dy,$$

che chiameremo *integrale improprio* di f in D . Supponiamo che esista una successione di domini contenuti in D , limitati e misurabili $\{D_n\}$ tali che $D_n \subset D_{n+1}$ per ogni $n \in \mathbb{N}$. Supponiamo inoltre che $\lim_{n \rightarrow \infty} \text{mis}(D_n) = \text{mis}(D)$, essendo $\text{mis}(D)$ l'estremo superiore delle intersezioni di D con un qualunque cerchio. Se $f(x, y)$ è integrabile in D_n per qualsiasi n , e se esiste ed è finito il

$$\lim_{n \rightarrow \infty} \iint_{D_n} f(x, y) dx dy,$$

indipendentemente dalla scelta della successione $\{D_n\}$, allora diremo che l'integrale improprio di f in D è convergente e scriveremo

$$\iint_D f(x, y) \, dx dy = \lim_{n \rightarrow \infty} \iint_{D_n} f(x, y) \, dx dy$$

• Esempi

1) Stabilire se l'integrale improprio

$$\iint_D \frac{1}{\sqrt[3]{x^2 + y^2}} \, dx dy, \quad \text{con } D = \{(x, y) \in \mathbb{R}^2 : 0 \leq x^2 + y^2 \leq 1\}$$

è convergente.

La funzione f non è limitata in un intorno dell'origine. Consideriamo allora la successione di domini $\{D_n\}$ con

$$D_n = \{(x, y) \in \mathbb{R}^2 : \frac{1}{n^2} \leq x^2 + y^2 \leq 1\}.$$

In D_n la funzione f è integrabile e si ha

$$\iint_{D_n} \frac{1}{\sqrt[3]{x^2 + y^2}} \, dx dy = \int_0^{2\pi} \left(\int_{\frac{1}{n}}^1 \frac{1}{\rho^{2/3}} \rho \, d\rho \right) d\theta = \frac{3}{2}\pi \left(1 - \frac{1}{n^{4/3}} \right).$$

passando al limite, si trova che l'integrale improprio converge e si ha

$$\iint_D \frac{1}{\sqrt[3]{x^2 + y^2}} \, dx dy = \lim_{n \rightarrow \infty} \iint_{D_n} \frac{1}{\sqrt[3]{x^2 + y^2}} \, dx dy = \frac{3}{2}\pi.$$

2) Stabilire se l'integrale improprio

$$I = \iint_{\mathbb{R}^2} \frac{1}{(x^2 + y^2 + 1)^2} \, dx dy,$$

converge.

Il dominio non è limitato e ha misura infinita. Consideriamo la successione di domini limitati $\{D_n\}$ dove

$$D_n = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq n^2\}.$$

Si ha

$$\iint_{D_n} \frac{1}{(x^2 + y^2 + 1)^2} \, dx dy = \int_0^{2\pi} \left(\int_0^n \frac{\rho \, d\rho}{(\rho^2 + 1)^2} \right) d\theta = \pi \left(1 - \frac{1}{n^2 + 1} \right).$$

Passando al limite per $n \rightarrow \infty$ si ottiene la convergenza e il risultato $I = \pi$.

7.4 Esempi di calcolo di integrali impropri

Consideriamo qui alcuni esempi notevoli di integrali impropri il cui calcolo richiede tecniche particolari.

1) Calcolare l'integrale

$$I = \int_0^{+\infty} e^{-x^2} dx.$$

Osserviamo che si può sempre scrivere

$$I^2 = \int_0^{+\infty} e^{-x^2} dx \int_0^{+\infty} e^{-y^2} dy.$$

D'altra parte

$$\int_0^{+\infty} e^{-x^2} dx \int_0^{+\infty} e^{-y^2} dy = \iint_D e^{-(x^2+y^2)} dx dy,$$

dove D è il dominio non limitato rappresentato dall'intero primo quadrante. Per calcolare questo integrale doppio improprio consideriamo la successione di domini $D_n = \{(x, y) \in \mathbb{R}^+ \times \mathbb{R}^+ : x^2 + y^2 \leq n^2\}$. Otteniamo

$$\iint_{D_n} e^{-(x^2+y^2)} dx dy = \int_0^{\frac{\pi}{2}} \left(\int_0^n e^{-\rho^2} \rho d\rho \right) d\theta = \frac{\pi}{4} (-e^{-n^2} + 1).$$

da cui passando al limite per $n \rightarrow \infty$ si ha, $I^2 = \frac{\pi}{4}$. Si ottiene così

$$\int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

2) Calcolare, se converge, l'integrale improprio

$$I(\beta) = \int_0^{+\infty} \frac{\sin \beta x}{x} dx.$$

Introduciamo l'integrale ausiliario

$$I(\alpha, \beta) = \int_0^{+\infty} e^{-\alpha x} \frac{\sin \beta x}{x} dx,$$

dove $\alpha > 0$. Come vedremo tra poco, questo integrale converge e si può quindi considerare $I(\alpha, \beta)$ come una funzione di β . La funzione integranda, pensata come funzione di β e x risulta continua in tutto il primo quadrante esclusa l'origine, e ammette derivata parziale rispetto a β continua. Utilizzando la regola di derivazione sotto il segno di integrale, abbiamo

$$\frac{\partial I(\alpha, \beta)}{\partial \beta} = \int_0^{+\infty} \frac{\partial}{\partial \beta} \left(e^{-\alpha x} \frac{\sin \beta x}{x} \right) dx = \int_0^{+\infty} e^{-\alpha x} \cos \beta x dx.$$

Mediante una integrazione per parti si ha

$$\int_0^b e^{-\alpha x} \cos \beta x dx = \frac{1}{\alpha^2 + \beta^2} [\alpha + (\beta \sin b - \alpha \cos \beta b) e^{-\alpha b}].$$

Prendendo il limite per $b \rightarrow +\infty$ si ottiene allora

$$\frac{\partial I(\alpha, \beta)}{\partial \beta} = \frac{\alpha}{\alpha^2 + \beta^2}.$$

D'altra parte $I(\alpha, \beta)|_{\beta=0} = 0$. A questo punto integriamo rispetto a β , ottenendo

$$I(\alpha, \beta) = \int \frac{\alpha}{\alpha^2 + \beta^2} d\beta = \arctan \frac{\beta}{\alpha} + c.$$

Imponendo la condizione $I(\alpha, 0) = 0$, otteniamo $c = 0$. Per la continuità di $I(\alpha, \beta)$ l'integrale $I(\beta)$ si ottiene come limite di $I(\alpha, \beta)$ per $\alpha \rightarrow 0^+$. Si ha

$$I(\beta) = \lim_{\alpha \rightarrow 0^+} \arctan \frac{\beta}{\alpha} = \begin{cases} \frac{\pi}{2} & \text{per } \beta > 0 \\ 0 & \text{per } \beta = 0 \\ -\frac{\pi}{2} & \text{per } \beta < 0. \end{cases}$$

3) Calcolare, se converge, l'integrale improprio

$$I(\alpha, \beta) = \int_0^{+\infty} e^{-\alpha x^2} \cos \beta x \, dx$$

essendo $\alpha > 0$.

Consideriamo l'integrale

$$I_b(\beta) = \int_0^b e^{-\alpha x^2} \cos \beta x \, dx,$$

dove α ha il ruolo di un parametro. Derivando otteniamo

$$\frac{dI_b(\beta)}{d\beta} = - \int_0^b e^{-\alpha x^2} \sin \beta x \, dx.$$

Mediante una integrazione per parti, si ha

$$\frac{dI_b(\beta)}{d\beta} = \frac{1}{2\alpha} e^{-\alpha b^2} \sin \beta b - \frac{\beta}{2\alpha} \int_0^b e^{-\alpha x^2} \cos \beta x \, dx = \frac{1}{2\alpha} e^{-\alpha b^2} \sin \beta b - \frac{\beta}{2\alpha} I_b(\beta).$$

Quest'ultima eguaglianza può essere vista come una equazione differenziale del primo ordine per la funzione $I_b(\beta)$. L'integrale generale di questa equazione è

$$I_b(\beta) = C e^{-\frac{\beta^2}{4\alpha}} + \frac{1}{2\alpha b} e^{-\alpha b^2} \left(\frac{\beta}{2\alpha} \sin \beta b - \cos \beta b \right).$$

Passando al limite per $b \rightarrow +\infty$ si ricava

$$I(\alpha, \beta) = \lim_{b \rightarrow +\infty} I_b(\beta) = C e^{-\frac{\beta^2}{4\alpha}}.$$

Per determinare la costante di integrazione C , osserviamo che, dal risultato dell'esempio 1),

$$I(\alpha, 0) = \int_0^{+\infty} e^{-\alpha x^2} \, dx = \frac{1}{\sqrt{\alpha}} \int_0^{+\infty} e^{-y^2} \, dy = \frac{1}{2} \sqrt{\frac{\pi}{\alpha}}.$$

L'integrale improprio $I(\alpha, \beta)$ risulta quindi convergente e si ha

$$\int_0^{+\infty} e^{-\alpha x^2} \cos \beta x \, dx = \frac{1}{2} \sqrt{\frac{\pi}{\alpha}} e^{-\frac{\beta^2}{4\alpha}}.$$

Capitolo 8

Serie numeriche

8.1 Definizioni e operazioni sulle serie

Data la successione di numeri reali $\{y_n\}_{n \in \mathbb{N}}$ si dice *serie numerica* l'espressione

$$y_1 + y_2 + y_3 + \dots + y_n + \dots = \sum_{n=1}^{\infty} y_n.$$

Al fine di attribuire un significato numerico alla somma infinita $\sum_{n=1}^{\infty} y_n$, consideriamo la successione delle *somme parziali* $\{S_n\}_{n \in \mathbb{N}}$, definite da

$$\begin{aligned} S_1 &= y_1 \\ S_2 &= y_1 + y_2 \\ S_3 &= y_1 + y_2 + y_3 \\ &\vdots \\ S_n &= y_1 + y_2 + \dots + y_n = \sum_{k=1}^n y_k \\ &\vdots \end{aligned}$$

Se tale successione ammette limite finito S allora tale limite si dice somma della serie numerica. La serie si dice allora *convergente* e si scrive

$$S = \lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \sum_{k=1}^n y_k.$$

Se il limite è infinito la serie si dice *divergente*, e se il limite non esiste si dice che la serie è non convergente o *oscillante*. Il comportamento di S_n per $n \rightarrow \infty$ si dice anche *carattere* della serie.

- Esempi

1) La serie

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

è convergente. Infatti si ha

$$\begin{aligned} S_n &= \sum_{k=1}^n \frac{1}{k(k+1)} = \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k+1} \right) = 1 - \frac{1}{2} + \frac{1}{2} - \frac{1}{3} + \frac{1}{3} - \dots + \frac{1}{n} - \frac{1}{n+1} \\ &= 1 - \frac{1}{n+1} = \frac{n}{n+1}, \end{aligned}$$

da cui

$$\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \frac{n}{n+1} = 1.$$

2) Dato $x \in \mathbb{R}$, consideriamo la serie

$$1 + x + x^2 + x^3 + \dots + x^{n-1} + \dots = \sum_{n=1}^{\infty} x^{n-1},$$

detta serie *geometrica*. Osserviamo che

$$S_n = 1 + x + x^2 + \dots + x^{n-1} = \begin{cases} n & \text{se } x = 1 \\ \frac{1-x^n}{1-x} & \text{se } x \neq 1. \end{cases}$$

Quindi, se $x = 1$ la serie diverge. Se $|x| < 1$ risulta $\lim_{n \rightarrow \infty} x^n = 0$. Allora la serie converge e si ha

$$\lim_{n \rightarrow \infty} S_n = \frac{1}{1-x}.$$

Supponiamo ora che $x > 1$. Allora $\lim_{n \rightarrow \infty} x^n = +\infty$, ed essendo $1 - x < 0$, si ottiene

$$\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \frac{1-x^n}{1-x} = +\infty.$$

Infine, se $x \leq -1$, il termine x^n oscilla in quanto risulta positivo per n pari e negativo per n dispari ed inoltre $|x^n| \geq 1$. Ne segue che la serie geometrica risulta oscillante per $x \leq -1$. Riassumendo abbiamo,

$$\sum_{n=1}^{\infty} x^{n-1} \begin{cases} \text{è oscillante per } x \leq -1 \\ \text{è convergente per } -1 < x < 1 \\ \text{è divergente per } x \geq 1. \end{cases}$$

Data la serie $y_1 + y_2 + y_3 + \dots + y_n + \dots$ considero la serie ottenuta da questa eliminando i primi k termini, ovvero

$$y_{k+1} + y_{k+2} + \dots + y_n + \dots = \sum_{n=k+1}^{\infty} y_n.$$

Questa serie si dice *k-esimo resto* della serie data. Vale la seguente proprietà.

Teorema 8.1 *Una serie e tutti i suoi resti hanno lo stesso carattere.*

Dim. Chiamiamo σ_k la somma dei primi k termini della serie. Avremo

$$S_n = \sum_{p=1}^n y_p = \sigma_k + \sum_{p=1}^{n-k} y_{k+p}.$$

Posto $y_{k+p} = \bar{y}_p$, la serie $\sum_{p=1}^{\infty} \bar{y}_p$ è il k -esimo resto della serie data ed $\bar{S}_n = \sum_{p=1}^n \bar{y}_p$ la sua somma parziale. Poiché

$$\lim_{n \rightarrow \infty} S_n = \lim_{n \rightarrow \infty} \left(\sigma_k + \sum_{p=1}^{n-k} \bar{y}_p \right) = \sigma_k + \lim_{n \rightarrow \infty} \bar{S}_n,$$

risulta che se S_n converge, convergerà anche \bar{S}_n e viceversa. Lo stesso si può dire se S_n è divergente o oscillante.

Date le due serie numeriche

$$y_1 + y_2 + \dots + y_n + \dots = \sum_{n=1}^{\infty} y_n$$

$$t_1 + t_2 + \dots + t_n + \dots = \sum_{n=1}^{\infty} t_n$$

e due numeri reali λ, μ , la serie

$$\lambda y_1 + \mu t_1 + \lambda y_2 + \mu t_2 + \dots + \lambda y_n + \mu t_n + \dots = \sum_{n=1}^{\infty} (\lambda y_n + \mu t_n),$$

si dice *combinazione lineare* delle due serie. Si ha il seguente risultato

Teorema 8.2 *La combinazione lineare di due serie convergenti è convergente.*

Dim. Dette S'_n e S''_n le somme parziali delle due serie ed S_n la somma parziale della loro combinazione lineare, si ha

$$S_n = \sum_{k=1}^n (\lambda y_k + \mu t_k) = \lambda \sum_{k=1}^n y_k + \mu \sum_{k=1}^n t_k = \lambda S'_n + \mu S''_n.$$

Ne segue che, detti S' e S'' i limiti delle due somme parziali S'_n e S''_n , esisterà, finito, il limite S della successione S_n , e si avrà

$$S = \lim_{n \rightarrow \infty} S_n = \lambda \lim_{n \rightarrow \infty} S'_n + \mu \lim_{n \rightarrow \infty} S''_n = \lambda S' + \mu S''.$$

Vale la pena osservare che la combinazione lineare di due serie divergenti o oscillanti può essere convergente. Ad esempio, per $|x| < 1$, le due serie

$$1 + 1 + \dots + 1 + \dots$$

$$(x - 1) + (x^2 - 1) + \dots + (x^n - 1) + \dots$$

sono ambedue divergenti ma hanno per somma il primo resto della serie geometrica, che sotto le ipotesi fatte, risulta convergente.

8.2 Criteri di convergenza

Sia $\{y_n\}$ una successione numerica di Cauchy, ovvero tale che $\forall \epsilon > 0 \exists n_\epsilon \in \mathbb{N} : \forall n, m > n_\epsilon \Rightarrow |y_m - y_n| < \epsilon$. Poiché ogni successione numerica di Cauchy in \mathbb{R} è convergente, e viceversa, ogni successione numerica convergente è di Cauchy, possiamo stabilire il seguente criterio di convergenza per una serie numerica.

Teorema 8.3 (Criterio di Cauchy) *Condizione necessaria e sufficiente affinché la serie numerica $\sum_{n=1}^{\infty} y_n$ sia convergente è che, fissato $\epsilon > 0$, esista un $n_\epsilon \in \mathbb{N}$ tale che per $n > n_\epsilon$ si abbia*

$$|y_{n+1} + y_{n+2} + \dots + y_{n+p}| < \epsilon, \quad (p = 0, 1, 2, \dots).$$

Dim. Osserviamo che, posto, $m = n + p$, si può scrivere, per qualunque p ,

$$|y_{n+1} + y_{n+2} + \dots + y_{n+p}| = |S_{n+p} - S_n| = |S_m - S_n|.$$

Allora la condizione del teorema si traduce nella condizione che la successione delle somme parziali della serie sia una successione di Cauchy. Ma questa successione è una successione numerica e la condizione di Cauchy risulta necessaria e sufficiente per la sua convergenza.

Teorema 8.4 *Se la serie numerica $\sum_{n=1}^{\infty} y_n$ converge, allora $\lim_{n \rightarrow \infty} y_n = 0$.*

Dim. Dall'ipotesi di convergenza della serie, applicando il teorema 8.3 per $p = 0$ si ottiene che, fissato $\epsilon > 0$ esiste un n_ϵ tale che per $n > n_\epsilon$ si abbia $|y_n| < \epsilon$. Ma quest'ultima condizione equivale a dire che $\lim_{n \rightarrow \infty} y_n = 0$.

Dal teorema precedente segue, ovviamente che, se $\lim_{n \rightarrow \infty} y_n \neq 0$ oppure se il limite non esiste affatto, la serie $\sum_{n=1}^{\infty} y_n$ sarà divergente o oscillante. Viceversa, il fatto che y_n sia infinitesimo non è sufficiente per affermare che la serie $\sum_{n=1}^{\infty} y_n$ converge. A questo proposito consideriamo la seguente serie,

$$1 + \frac{1}{2} + \frac{1}{3} + \dots + \frac{1}{n} + \dots = \sum_{n=1}^{\infty} \frac{1}{n},$$

detta *serie armonica*. In questo caso si ha $\lim_{n \rightarrow \infty} \frac{1}{n} = 0$. Tuttavia la serie diverge. Per provarlo utilizziamo il criterio di Cauchy. Posto $p = n$, abbiamo

$$y_{n+1} + \dots + y_{2n} = \frac{1}{n+1} + \frac{1}{n+2} + \dots + \frac{1}{2n} > \underbrace{\frac{1}{2n} + \frac{1}{2n} + \dots + \frac{1}{2n}}_{n \text{ volte}} = \frac{1}{2}.$$

Poiché questa disuguaglianza vale per qualunque n , abbiamo $|S_{2n} - S_n| > \frac{1}{2}$. Scelto $\epsilon < \frac{1}{2}$, il criterio di Cauchy non è soddisfatto, quindi la serie non è convergente. Inoltre, la successione S_n risulta crescente e non può essere limitata perché altrimenti sarebbe convergente. Quindi S_n sarà divergente.

Teorema 8.5 (Criterio del confronto) *Siano $\sum_{n=1}^{\infty} y_n$ e $\sum_{n=1}^{\infty} t_n$ due serie a termini non negativi e tali che $y_n \leq t_n$ per ogni n . Allora la convergenza di $\sum_{n=1}^{\infty} t_n$ implica la convergenza di $\sum_{n=1}^{\infty} y_n$ e la divergenza di $\sum_{n=1}^{\infty} y_n$ implica la divergenza di $\sum_{n=1}^{\infty} t_n$.*

Dim. Le somme parziali S_n della serie $\sum_{n=1}^{\infty} t_n$, costituiscono una successione a termini non negativi, monotona non decrescente. Se $\{S_n\}$ è convergente essa sarà limitata. Dalla disuguaglianza $y_n \leq t_n$, risulta allora che $S'_n = \sum_{k=1}^n y_n \leq \sum_{k=1}^n t_n = S_n$, quindi anche S'_n è limitata, ed essendo anch'essa non decrescente, risulterà convergente.

Supponiamo ora che la successione $\{S'_n\}$ sia divergente. Allora essa non sarà limitata superiormente e, per la disuguaglianza $y_n \leq t_n$, anche $\{S_n\}$ non sarà limitata superiormente e quindi divergerà.

• Esempi

1) Studiare la convergenza della serie

$$\sum_{n=1}^{\infty} \frac{1}{2^n + \frac{1}{n}}.$$

Osserviamo che la serie è a termini positivi e risulta $\frac{1}{2^n + \frac{1}{n}} \leq \frac{1}{2^n}, \forall n$. Ma la serie $\sum_{n=1}^{\infty} (\frac{1}{2})^n$ è il primo resto della serie geometrica di argomento $\frac{1}{2}$ ed è quindi convergente a $\frac{1}{1-1/2} - 1 = 1$. Per il criterio del confronto la serie di partenza sarà convergente.

2) Consideriamo la serie

$$1 + \frac{1}{2^x} + \frac{1}{3^x} + \dots + \frac{1}{n^x} + \dots = \sum_{n=1}^{\infty} \frac{1}{n^x},$$

detta *serie armonica generalizzata*. Studiamo il suo carattere al variare di $x \in \mathbb{R}$. Per $x \leq 0$ si ha

$$y_n = \frac{1}{n^x} = n^{-x} \geq 1,$$

quindi $S_n \geq n$ e, utilizzando il criterio del confronto, la serie sarà divergente. Per $0 < x \leq 1$ abbiamo

$$y_n = \frac{1}{n^x} \geq \frac{1}{n},$$

e quindi, sempre per il criterio del confronto la serie diverge. In particolare, per $x = 1$ la serie coincide con la serie armonica. Consideriamo infine il caso $x > 1$. Introdotta la funzione $f(t) = \frac{1}{t^x}$ essa risulta continua e monotona non crescente per $t \geq 1$. Applicando il teorema della media all'integrale

$$\int_n^{n+1} \frac{1}{t^x} dt$$

si ottiene

$$\int_n^{n+1} \frac{1}{t^x} dt = \frac{1}{\xi^x} \geq \frac{1}{(n+1)^x}.$$

D'altra parte, l'integrazione diretta fornisce

$$\int_n^{n+1} \frac{1}{t^x} dt = \frac{1}{x-1} \left[\frac{1}{n^{x-1}} - \frac{1}{(n+1)^{x-1}} \right].$$

Ne segue la disuguaglianza

$$\frac{1}{(n+1)^x} \leq \frac{1}{x-1} \left[\frac{1}{n^{x-1}} - \frac{1}{(n+1)^{x-1}} \right].$$

Consideriamo la serie dei termini a destra di quest'ultima disuguaglianza, a meno del fattore $\frac{1}{x-1}$,

$$1 - \frac{1}{2^{x-1}} + \frac{1}{2^{x-1}} - \frac{1}{3^{x-1}} + \dots + \frac{1}{n^{x-1}} - \frac{1}{(n+1)^{x-1}} + \dots,$$

la cui somma parziale è

$$\sigma_n = 1 - \frac{1}{(n+1)^{x-1}}.$$

Poiché $\lim_{n \rightarrow \infty} \sigma_n = 1$, quest'ultima serie risulta convergente e, per il criterio del confronto risulta convergente anche la serie $\sum_{n=1}^{\infty} \frac{1}{(n+1)^x}$. Ma quest'ultima non è altro che il primo resto della serie armonica generalizzata che risulta così convergente. Riassumendo

$$\sum_{n=1}^{\infty} \frac{1}{n^x} \begin{cases} \text{è divergente per } x \leq 1 \\ \text{è convergente per } x > 1. \end{cases}$$

Teorema 8.6 (Criterio del rapporto o di D'Alembert) Sia $\sum_{n=1}^{\infty} y_n$ una serie a termini positivi. Se accade che

$$\lim_{n \rightarrow \infty} \frac{y_{n+1}}{y_n} = \lambda,$$

essendo $0 \leq \lambda < 1$, allora la serie è convergente. Se invece è $\lambda > 1$ la serie è divergente.

Dim. Supponiamo che $\lim_{n \rightarrow \infty} \frac{y_{n+1}}{y_n} = \lambda < 1$. Considerato un numero q tale che $\lambda < q < 1$ e posto $\epsilon = q - \lambda$, dalla definizione di limite esisterà un $n_\epsilon \in \mathbb{N}$ tale che

$$\frac{y_{n+1}}{y_n} < q, \quad \forall n > n_\epsilon.$$

Ne seguono le disequaglianze

$$y_{n+1} < qy_n, \quad y_{n+2} < qy_{n+1} < q^2y_n, \quad \dots \quad \forall n > n_\epsilon,$$

da cui

$$y_{n+1} + y_{n+2} + \dots + y_{n+l} < y_n \sum_{p=1}^l q^p.$$

Poiché la somma a secondo membro di questa disequaglianza è la somma parziale del primo resto della serie geometrica con argomento minore di uno, che è convergente, la serie

$$y_{n+1} + y_{n+2} + \dots$$

dove n è un qualunque numero intero maggiore di n_ϵ , risulterà convergente. Ma quest'ultima, è l' n -esimo resto della serie di partenza, che dunque risulterà convergente.

Supponiamo ora $\lambda > 1$. In questo caso esisterà un indice ν tale che, per $n > \nu$ si abbia $y_{n+1} > y_n$. Quindi dovrà aversi $\lim_{n \rightarrow \infty} y_n > 0$ e questo implica che la serie non può essere convergente. Infine, essendo a termini positivi, la serie dovrà necessariamente divergere.

• Esempi

1) Studiare il carattere della serie $\sum_{n=1}^{\infty} \frac{n}{2^n}$.

Applicando il criterio del rapporto otteniamo

$$\frac{y_{n+1}}{y_n} = \frac{1}{2} \frac{n+1}{n},$$

il cui limite, per $n \rightarrow \infty$ è $\frac{1}{2}$. La serie è allora convergente.

2) Studiare il carattere della serie $\sum_{n=1}^{\infty} \frac{n^n}{n!}$.

Applico ancora il criterio del rapporto e trovo

$$\frac{y_{n+1}}{y_n} = \left(1 + \frac{1}{n}\right)^n.$$

Poiché $\lim_{n \rightarrow \infty} \left(1 + \frac{1}{n}\right)^n = e$, la serie risulterà divergente.

3) La serie

$$1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^{n-1}}{(n-1)!} + \dots = \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!},$$

si dice *serie esponenziale*. Per $x \geq 0$ la serie risulta essere a termini non negativi. Per $x = 0$ si ha $S_n = 1$ quindi la serie è convergente. Per $x > 0$ si ha

$$\frac{y_{n+1}}{y_n} = \frac{x^n (n-1)!}{n! x^{n-1}} = \frac{x}{n}.$$

Poiché $\lim_{n \rightarrow \infty} \frac{x}{n} = 0$, ne segue che la serie esponenziale è convergente per $x \geq 0$.

4) La serie

$$x + \frac{x^2}{2} + \frac{x^3}{3} + \dots + \frac{x^n}{n} + \dots = \sum_{n=1}^{\infty} \frac{x^n}{n},$$

è detta *serie logaritmica*. Per $x \geq 0$ essa è a termini non negativi. Poiché si ha

$$\frac{y_{n+1}}{y_n} = \frac{nx}{n+1},$$

ne segue $\lim_{n \rightarrow \infty} \frac{y_{n+1}}{y_n} = x$. Quindi essa è convergente per $0 \leq x < 1$ e divergente per $x > 1$. Per $x = 1$ essa coincide con la serie armonica quindi risulta divergente.

Teorema 8.7 (*criterio della radice*). Sia $\sum_{n=1}^{\infty} y_n$ una serie a termini non negativi. Supponiamo che esista, finito, il limite

$$\lim_{n \rightarrow \infty} \sqrt[n]{y_n} = l.$$

Allora, se risulta $0 \leq l < 1$, la serie converge, mentre se risulta $l > 1$ la serie diverge.

Dim. Dall'ipotesi, fissato $\epsilon > 0$, esisterà un indice $\nu \in \mathbb{N}$ tale che

$$l - \epsilon < \sqrt[n]{y_n} < l + \epsilon, \quad \text{per } n > \nu.$$

Se $0 \leq l < 1$, allora, per ϵ sufficientemente piccolo si ha $q = l + \epsilon < 1$ e quindi

$$\sqrt[n]{y_n} < q \implies y_n < q^n.$$

Poiché q^n è il generico termine di una serie geometrica di ragione minore di 1, la serie delle q^n converge, e per il criterio del confronto, convergerà anche la serie di partenza.

Se $l > 1$, sempre per un opportuno ϵ , si avrà $p = l - \epsilon > 1$, da cui

$$\sqrt[n]{y_n} > p, \implies y_n > p^n.$$

Poiché in questo caso la serie di termini p^n sarà divergente, per il teorema del confronto si avrà la divergenza della serie di partenza.

- Esempi

1) Studiare il carattere della serie

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

Si tratta di una serie a termini non negativi. Applicando il criterio della radice si ha

$$\lim_{n \rightarrow \infty} \sqrt[n]{\left(\frac{n-1}{n} \right)^{n^2}} = \lim_{n \rightarrow \infty} \left(\frac{n-1}{n} \right)^n = \lim_{n \rightarrow \infty} \left(1 - \frac{1}{n} \right)^n = \frac{1}{e} < 1.$$

Ne segue la convergenza della serie.

2) Studiare il carattere della serie

$$\sum_{n=1}^{\infty} \frac{5^{n+\frac{2}{n}}}{3^n}.$$

La serie è a termini positivi. Applicando il criterio della radice,

$$\lim_{n \rightarrow \infty} \sqrt[n]{\frac{5^{n+\frac{2}{n}}}{3^n}} = \frac{1}{3} \lim_{n \rightarrow \infty} 5^{\frac{n+2/n}{n}} = \frac{5}{3} > 1.$$

La serie risulta dunque divergente.

OSSERVAZIONE. Come nel caso del criterio del rapporto, se $l = 1$ il criterio della radice non permette di dedurre la convergenza o la divergenza di una serie. Come esempio si consideri la serie armonica generalizzata $\sum_{n=1}^{\infty} \frac{1}{n^x}$. Si ha, per qualunque x ,

$$\lim_{n \rightarrow \infty} \sqrt[n]{\frac{1}{n^x}} = \lim_{n \rightarrow \infty} n^{-x/n} = 1.$$

Come sappiamo, la serie può essere convergente o divergente a seconda del valore di x ma il criterio della radice, in questo caso, non permette di stabilire il carattere della serie.

8.3 Serie alternanti e serie assolutamente convergenti

Si dice *alternante* una serie i cui termini sono alternativamente positivi e negativi. Se, per esempio, il primo termine è positivo, tutti i termini di posto dispari saranno positivi e tutti quelli di posto pari saranno negativi. Ad esempio è alternante la serie

$$1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \dots + \frac{(-1)^{n-1}}{n} + \dots = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}.$$

Si può dimostrare che se i valori assoluti dei termini di una serie alternante formano una successione monotona, allora la serie non è divergente. Più in particolare si ha il seguente risultato

Teorema 8.8 *Data una serie alternante con $y_1 > 0$, supponiamo che la successione $\{|y_n|\}$ sia monotona decrescente e che $\lim_{n \rightarrow \infty} y_n = 0$. Allora la serie converge e la sua somma è un numero positivo $S \leq y_1$.*

Dim. Consideriamo le somme parziali di ordine pari S_{2n} . Si può scrivere

$$S_{2n} = (|y_1| - |y_2|) + (|y_3| - |y_4|) + \dots + (|y_{2n-1}| - |y_{2n}|).$$

Poiché $\{|y_n|\}$ è decrescente, la successione $\{S_{2n}\}$ risulta crescente. D'altra parte si può anche scrivere

$$S_{2n} = |y_1| - (|y_2| - |y_3|) - (|y_4| - |y_5|) - \dots - |y_{2n}|,$$

quindi risulta $0 < S_{2n} < y_1$ cioè $\{S_{2n}\}$ è limitata. Essa risulterà allora convergente. Denotiamo con S il suo limite.

Consideriamo ora le somme parziali di ordine dispari S_{2n+1} . Poiché si ha $S_{2n+1} = S_{2n} + y_{2n+1}$, tenuto conto dell'ipotesi $\lim_{n \rightarrow \infty} y_n = 0$, si ricava

$$\lim_{n \rightarrow \infty} S_{2n+1} = \lim_{n \rightarrow \infty} (S_{2n} + y_{2n+1}) = S + \lim_{n \rightarrow \infty} y_{2n+1} = S.$$

Si conclude allora che le somme parziali pari e le somme parziali dispari convergono allo stesso limite che è la somma della serie.

• **Esempio**

Consideriamo la serie $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$. Poiché si ha $1 > \frac{1}{2} > \frac{1}{3} > \dots$ ed anche $\lim_{n \rightarrow \infty} \frac{(-1)^n}{n} = 0$, per il teorema precedente la serie risulta convergente.

Diremo che la serie $\sum_{n=1}^{\infty} y_n$ è *assolutamente convergente* se risulta convergente la serie $\sum_{n=1}^{\infty} |y_n|$.

Teorema 8.9 *Ogni serie assolutamente convergente è convergente.*

Dim. Osserviamo che, per ogni n si ha $-|y_n| \leq y_n \leq |y_n|$, da cui $0 \leq y_n + |y_n| \leq 2|y_n|$. Per ipotesi abbiamo che $\sum_{n=1}^{\infty} |y_n|$ è convergente e quindi lo sarà anche la serie $\sum_{n=1}^{\infty} 2|y_n|$. Per il criterio del confronto allora sarà convergente anche la serie $\sum_{n=1}^{\infty} (y_n + |y_n|)$. D'altra parte si ha

$$S_n = \sum_{k=1}^n y_k = \sum_{k=1}^n (y_k + |y_k|) - \sum_{k=1}^n |y_k|$$

da cui, passando al limite per $n \rightarrow \infty$, $\{S_n\}$ risulterà convergente perché differenza di due successioni convergenti.

La semplice convergenza non implica l'assoluta convergenza. Per esempio, la serie $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$ è convergente ma la serie dei valori assoluti $\sum_{n=1}^{\infty} \frac{1}{n}$ risulta divergente.

Per le serie assolutamente convergenti vale la disuguaglianza triangolare per somme infinite, ovvero

$$|y_1 + y_2 + \dots + y_n + \dots| \leq |y_1| + |y_2| + \dots + |y_n| + \dots$$

Infatti, data la serie $\sum_{n=1}^{\infty} y_n$ assolutamente convergente, denotate con S_n ed \bar{S}_n , rispettivamente le somme parziali della serie e di quella dei suoi valori assoluti, avremo

$$|S_n| \leq \bar{S}_n \quad \implies \quad -\bar{S}_n \leq S_n \leq \bar{S}_n,$$

in forza della disuguaglianza triangolare per somme finite. Passando al limite per $n \rightarrow \infty$, per la convergenza della serie e denotate con S e \bar{S} rispettivamente la somma della serie e quella dei suoi valori assoluti, otteniamo

$$-\bar{S} \leq S \leq \bar{S}, \quad \implies \quad |S| \leq \bar{S},$$

che è quanto volevamo dimostrare.

Le serie assolutamente convergenti godono inoltre della proprietà di commutatività della somma. Si può dimostrare infatti che se si scambiano di posto i termini di una serie assolutamente convergente la serie risultante continua ad essere assolutamente convergente e la sua somma non cambia. Questa proprietà non vale in generale per una serie non assolutamente convergente.

• Esempio

Abbiamo già visto che la serie $\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n}$ non è assolutamente convergente ma è convergente. Chiamiamo S la sua somma. Cambiamo l'ordine degli addendi della serie in modo tale che ogni termine positivo sia seguito da due termini negativi, ovvero

$$\begin{aligned} & \left(1 - \frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{3} - \frac{1}{6} - \frac{1}{8}\right) + \left(\frac{1}{5} - \frac{1}{10} - \frac{1}{12}\right) + \dots \\ &= \left(\frac{1}{2} - \frac{1}{4}\right) + \left(\frac{1}{6} - \frac{1}{8}\right) + \left(\frac{1}{10} - \frac{1}{12}\right) + \dots \\ &= \frac{1}{2} \left[1 - \frac{1}{2} + \frac{1}{3} - \frac{1}{4} + \frac{1}{5} - \frac{1}{6} + \dots\right] = \frac{1}{2} \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \frac{1}{2} S. \end{aligned}$$

Per effetto del cambiamento di ordine la serie ha cambiato somma.

Capitolo 9

Serie di funzioni

Si dice *serie di funzioni* la serie

$$f_1(x) + f_2(x) + \dots + f_n(x) + \dots = \sum_{n=1}^{\infty} f_n(x),$$

i cui termini sono funzioni definite in un certo intervallo A dei reali. Le serie

$$\begin{aligned} \sum_{n=1}^{\infty} x^{n-1}, & \quad (\text{serie geometrica}), \\ \sum_{n=1}^{\infty} \frac{1}{n^x}, & \quad (\text{serie armonica generalizzata}), \\ \sum_{n=1}^{\infty} \frac{x^{n-1}}{(n-1)!}, & \quad (\text{serie esponenziale}), \\ \sum_{n=1}^{\infty} \frac{x^n}{n}, & \quad (\text{serie logaritmica}), \end{aligned}$$

già considerate nel capitolo precedente, sono esempi di serie di funzioni.

Se una serie di funzioni risulta convergente in tutti i punti di un insieme $D \subseteq A$ diremo che la serie è convergente in D . Diremo poi che $\sum_{n=1}^{\infty} f_n(x)$ è assolutamente convergente in D se è convergente la serie $\sum_{n=1}^{\infty} |f_n(x)|$. La somma $S(x)$ di una serie di funzioni convergente in D è una funzione definita in D .

9.1 Convergenza uniforme

Sia $\sum_{n=1}^{\infty} f_n(x)$ una serie di funzioni, convergente in D . Si dice che la serie è *uniformemente convergente* in $\Omega \subseteq D$ se, $\forall \epsilon > 0, \exists \nu_\epsilon \in \mathbb{N} : \forall n > \nu_\epsilon, \forall x \in \Omega \Rightarrow |S_n(x) - S(x)| < \epsilon$. È importante notare che il numero ν_ϵ dipende qui da ϵ ma *non* da x . La convergenza uniforme implica la convergenza della serie di funzioni punto per punto in Ω . Tuttavia una serie di funzioni può essere convergente in un punto ma non essere uniformemente convergente in quel punto.

- Esempi

1) Verificare che la serie

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{\sqrt{1-x^2+n}},$$

converge uniformemente in $[-1, 1]$.

Osserviamo che la serie data è alternante ed è $f_1(x) = \frac{1}{\sqrt{1-x^2+1}} > 0, \forall x \in [-1, 1]$. Inoltre, la successione dei valori assoluti $\{|f_n(x)|\}$ risulta monotona decrescente per ogni $x \in [-1, 1]$ e si ha $\lim_{n \rightarrow \infty} f_n(x) = 0$. Ne segue che la serie è convergente in $[-1, 1]$. Consideriamo ora il resto n -esimo della serie data,

$$R_n(x) = (-1)^n \left[\frac{1}{\sqrt{1-x^2+n+1}} - \frac{1}{\sqrt{1-x^2+n+2}} + \dots \right].$$

Poiché si ha $|R_n(x)| \leq \frac{1}{\sqrt{1-x^2+n+1}} < \frac{1}{n}$, detta $S(x)$ la somma della serie, risulta

$$|S_n(x) - S(x)| = |R_n(x)| < \frac{1}{n}.$$

Quindi, fissato $\epsilon > 0$ si ha $|S_n(x) - S(x)| < \epsilon$ per $n > \nu_\epsilon = \frac{1}{\epsilon}$, per ogni $x \in [-1, 1]$. Il numero ν_ϵ è indipendente da x quindi la serie è uniformemente convergente in $[-1, 1]$.

2) Consideriamo la serie

$$\sum_{n=1}^{\infty} (x^n - x^{n-1}).$$

La somma parziale n -esima è

$$S_n(x) = (x - 1) + (x^2 - x) + (x^3 - x^2) + \dots + x^n - x^{n-1} = -1 + x^n,$$

quindi

$$\lim_{n \rightarrow \infty} S_n(x) = \begin{cases} -1 & \text{per } 0 \leq x < 1 \\ 0 & \text{per } x = 1 \end{cases}.$$

La serie è dunque convergente in $[0, 1]$. Per esaminare l'uniforme convergenza, osserviamo che, fissato un $\epsilon \in (0, 1)$ la disuguaglianza

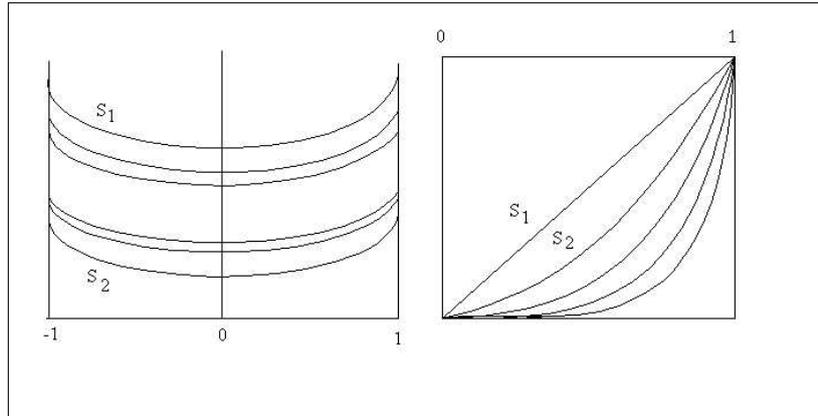
$$|S_n(x) - S(x)| \leq x^n < \epsilon, \quad \forall x \in [0, 1),$$

è soddisfatta per $n > \log_x \epsilon = \nu(\epsilon, x)$. Poiché comunque si scelga ϵ , si ha $\lim_{x \rightarrow 1^-} \log_x \epsilon = +\infty$, non esiste un ν tale che per $n > \nu$ sia soddisfatta la condizione di convergenza per ogni $x \in [0, 1]$. La serie quindi non è uniformemente convergente in $[0, 1]$. Osserviamo infine che, se ci restringiamo a considerare la serie nell'intervallo $[0, 1 - \delta]$, con $0 < \delta < 1$ allora la disuguaglianza $x^n < \epsilon$ risulta soddisfatta per

$$n > \frac{\ln \epsilon}{\ln(1 - \delta)} = \nu_\epsilon.$$

Possiamo dire allora che la serie risulta uniformemente convergente in $[0, 1 - \delta]$. A ulteriore commento dei due esempi precedenti, riportiamo in figura la rappresentazione grafica delle somme parziali delle due serie

Il seguente teorema stabilisce una condizione sufficiente per la convergenza di una serie di funzioni.



Teorema 9.1 (Criterio di Weierstrass). Data la serie di funzioni $\sum_{n=1}^{\infty} f_n(x)$, dove le $f_n(x)$ sono definite in D , supponiamo che per ogni $x \in \Omega \subseteq D$, con Ω chiuso, si abbia $|f_n(x)| \leq a_n$, $n = 1, 2, \dots$, e che la serie numerica $\sum_{n=1}^{\infty} a_n$ sia convergente. Allora la serie $\sum_{n=1}^{\infty} f_n(x)$ risulta assolutamente e uniformemente convergente in Ω .

Dim. Dalla disuguaglianza $|f_n(x)| \leq a_n$, $n = 1, 2, \dots$, mediante il criterio del confronto, deduciamo che la serie $\sum_{n=1}^{\infty} |f_n(x)|$ risulta convergente per ogni $x \in \Omega$ quindi la serie è assolutamente convergente in Ω . Proviamo che essa è anche uniformemente convergente. Posto $\sum_{n=1}^{\infty} f_n(x) = S(x)$ e $\sum_{n=1}^{\infty} a_n = \sigma$, si ha, per ogni $x \in \Omega$,

$$\begin{aligned} |S_n(x) - S(x)| &= |f_{n+1}(x) + f_{n+2}(x) + \dots| \leq |f_{n+1}(x)| + |f_{n+2}(x)| + \dots \\ &\leq a_{n+1} + a_{n+2} + \dots = \sigma - \sigma_n. \end{aligned}$$

Ne segue che, per la convergenza di $\sum_{n=1}^{\infty} a_n$,

$$\forall \epsilon > 0, \exists \nu_\epsilon \in \mathbb{N} : \forall n > \nu_\epsilon \Rightarrow |\sigma_n - \sigma| < \epsilon.$$

Ma questo implica che $|S_n(x) - S(x)| < \epsilon$, $\forall n > \nu_\epsilon$ indipendentemente da x . Concludiamo che la serie di funzioni è uniformemente convergente.

- Esempio

Studiare la convergenza della serie di funzioni

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n^2 + (4 - x^2)^{n/2}}.$$

I termini della serie sono funzioni definite in $[-2, 2]$. Si ha

$$\left| \frac{\sin nx}{n^2 + (4 - x^2)^{n/2}} \right| = \frac{|\sin nx|}{n^2 + \sqrt{(4 - x^2)^n}} \leq \frac{1}{n^2}, \quad \forall n = 1, 2, \dots, \forall x \in [-2, 2].$$

Poiché la serie $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converge, per il criterio di Weierstrass la serie di funzioni converge uniformemente in $[-2, 2]$.

Le serie uniformemente convergenti godono di alcune proprietà notevoli.

Teorema 9.2 *Se si moltiplicano tutti i termini di una serie di funzioni uniformemente convergente in $[a, b]$ per una funzione $g(x)$ limitata in $[a, b]$, allora la serie risultante $\sum_{n=1}^{\infty} g(x)f_n(x)$ è uniformemente convergente in $[a, b]$.*

Dim. Poiché si suppone $g(x)$ limitata in $[a, b]$, esisterà un $C \in \mathbb{R}^+$ tale che $|g(x)| \leq C$, $\forall x \in [a, b]$. Si può scrivere allora

$$\left| \sum_{k=1}^n g(x)f_k(x) - g(x)S(x) \right| = |g(x)||S_n(x) - S(x)| \leq C|S_n(x) - S(x)|, \quad \forall x \in [a, b].$$

Se la serie $\sum_{n=1}^{\infty} f_n(x)$ converge uniformemente, convergerà uniformemente anche la serie $\sum_{n=1}^{\infty} g(x)f_n(x)$ e la sua somma sarà $g(x)S(x)$, definita in $[a, b]$.

Teorema 9.3 *Se la serie $\sum_{n=1}^{\infty} f_n(x)$ è uniformemente convergente in $[a, b]$ e i suoi termini sono funzioni continue in $[a, b]$, allora la somma della serie è una funzione continua in $[a, b]$.*

Dim. Consideriamo un punto $x_0 \in [a, b]$ ed un incremento h tale che $x_0 + h \in [a, b]$. Poiché la serie è uniformemente convergente in $[a, b]$, fissato $\epsilon > 0$ si avrà

$$|S_n(x_0) - S(x_0)| < \frac{\epsilon}{3}, \quad |S_n(x_0 + h) - S(x_0 + h)| < \frac{\epsilon}{3},$$

per ogni $n > \nu_\epsilon \in \mathbb{N}$ indipendentemente dalla scelta di x_0 in $[a, b]$ e di h . Poiché le funzioni $f_n(x)$ sono continue, la somma parziale $S_n(x)$ sarà una funzione continua quindi, preso un qualunque n , esisterà un $\delta_\epsilon(n)$ tale che per $|h| < \delta_\epsilon(n)$ si abbia

$$|S_n(x_0 + h) - S_n(x_0)| \leq \frac{\epsilon}{3}.$$

Se scegliamo allora $n > \nu_\epsilon$ e, in corrispondenza, $|h| < \delta_\epsilon(n)$, risulterà

$$\begin{aligned} |S(x_0 + h) - S(x_0)| &= \\ |S(x_0 + h) - S_n(x_0 + h) + S_n(x_0 + h) - S_n(x_0) + S_n(x_0) - S(x_0)| &\leq \\ |S(x_0 + h) - S_n(x_0 + h)| + |S_n(x_0 + h) - S_n(x_0)| + |S_n(x_0) - S(x_0)| &< \epsilon. \end{aligned}$$

Ma questa non è altro che la condizione di continuità della funzione $S(x)$ in x_0 . Poiché il risultato vale per qualunque $x_0 \in [a, b]$, la somma della serie risulta continua in tutto $[a, b]$.

Osserviamo che se viene a mancare l'ipotesi di convergenza uniforme, la semplice continuità dei termini della serie non è sufficiente a garantire la continuità della somma.

- Esempi

1) Consideriamo la serie precedentemente studiata

$$\sum_{n=1}^{\infty} (x^n - x^{n-1}).$$

Abbiamo già visto che questa serie non è uniformemente convergente in $[0, 1]$. Le funzioni $f_n(x) = x^n - x^{n-1}$ sono continue in $[0, 1]$ mentre la somma risulta discontinua in $x = 1$, in quanto,

$$S(x) = \begin{cases} -1 & \text{per } 0 \leq x < 1 \\ 0 & \text{per } x = 1. \end{cases}$$

2) Abbiamo visto che la serie armonica generalizzata $\sum_{n=1}^{\infty} \frac{1}{n^x}$ è convergente per $x > 1$. Preso un $\alpha > 0$, questa serie risulta uniformemente convergente per $x \geq 1 + \alpha$. Infatti si ha

$$\frac{1}{n^x} \leq \frac{1}{n^{1+\alpha}},$$

e poiché la serie $\sum_{n=1}^{\infty} \frac{1}{n^{1+\alpha}}$ è convergente, per il criterio di Weierstrass la serie armonica generalizzata è uniformemente convergente per $x \geq 1 + \alpha$. Ne segue che la funzione

$$\zeta(x) = \sum_{n=1}^{\infty} \frac{1}{n^x},$$

è continua per $x > 1$. Questa funzione è nota come *zeta di Riemann*.

Si possono dimostrare anche i seguenti teoremi che qui non dimostreremo.

Teorema 9.4 *Se la serie di funzioni $\sum_{n=1}^{\infty} f_n(x)$ converge uniformemente a $S(x)$ in $[a, b]$ e le funzioni $f_n(x)$ sono continue in $[a, b]$, allora la serie è integrabile termine a termine ovvero,*

$$\int_{x_0}^x S(\xi) d\xi = \sum_{n=1}^{\infty} \int_{x_0}^x f_n(\xi) d\xi,$$

per ogni $x_0 \in [a, b]$. Inoltre la serie a secondo membro converge uniformemente in $[a, b]$.

Teorema 9.5 *Sia $\sum_{n=1}^{\infty} f_n(x)$ una serie convergente i cui termini sono funzioni derivabili in $[a, b]$. Supponiamo che la serie delle derivate $\sum_{n=1}^{\infty} f'_n(x)$ converga uniformemente in $[a, b]$. Allora la serie è derivabile termine a termine ovvero,*

$$\frac{d}{dx} S(x) = \sum_{n=1}^{\infty} f'_n(x), \quad \forall x \in [a, b].$$

9.2 Serie di potenze

Sia $x_0 \in \mathbb{R}$ e $c_n \in \mathbb{R}$, $\forall n \in \mathbb{N}$. La serie

$$\sum_{n=0}^{\infty} c_n (x - x_0)^n,$$

si dice *serie di potenze* e i numeri c_n si dicono coefficienti della serie. È evidente che ogni serie di potenze converge per $x = x_0$. In tal caso infatti si ha $\sum_{n=0}^{\infty} c_n (x - x_0)^n = c_0$. Mediante il cambiamento di variabile $x - x_0 = X$ si può sempre scrivere una serie di potenze nella forma $\sum_{n=0}^{\infty} c_n X^n$.

Teorema 9.6 (di Abel). *Se la serie di potenze $\sum_{n=0}^{\infty} c_n x^n$ converge per $x = x_1 \neq 0$, allora essa converge assolutamente per tutti gli x tali che $|x| < |x_1|$. Se essa diverge per $x = x_2$, allora essa non convergerà per tutti gli x tali che $|x| > |x_2|$.*

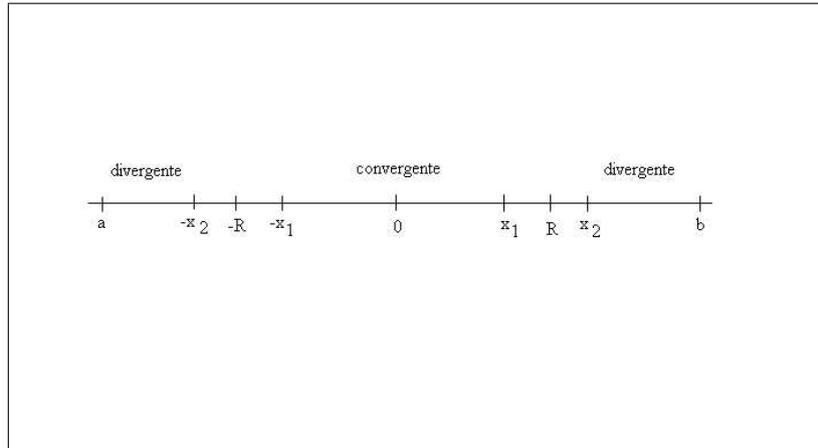
Dim. Se $\sum_{n=0}^{\infty} c_n x^n$ converge in x_1 avremo necessariamente che $\lim_{n \rightarrow \infty} c_n x_1^n = 0$. Ne segue che esiste un maggiorante M dei termini della serie $\sum_{n=0}^{\infty} c_n x_1^n$ cioè $|c_n x_1^n| \leq M, \forall n$. Scelto allora $|x| < |x_1|$ avremo

$$|c_n x^n| = |c_n x_1^n| \left| \frac{x}{x_1} \right|^n \leq M q^n,$$

essendo $q = |x/x_1| < 1$. Ma il termine a secondo membro è il termine generale del primo resto di una serie geometrica di argomento minore di uno, che è convergente. Ne segue, per il criterio del confronto, che la serie data è assolutamente convergente per $|x| < |x_1|$.

Supponiamo ora che la serie di potenze diverga per $x = x_2$ e ragioniamo per assurdo. Supponiamo che per un x tale che $|x| > |x_2|$ la serie converga. Allora, per quanto dimostrato prima, essa deve convergere anche per $x = x_2$. Ma ciò contraddice l'ipotesi, quindi la serie non converge per $|x| > |x_2|$.

Il teorema di Abel ci permette di suddividere l'asse reale in intervalli di convergenza di una serie di potenze ed intervalli di divergenza. Tali intervalli sono simmetrici rispetto all'origine.



Definiamo *raggio di convergenza* di una serie di potenze, l'estremo superiore dell'insieme dei valori x_1 per i quali la serie converge e lo denoteremo con R . Diremo quindi che una serie di potenze ha raggio di convergenza R se essa converge nell'intervallo $(-R, R)$. Tale intervallo può essere chiuso o aperto a seconda che la serie converga o meno per $x = \pm R$. Se la serie converge solo per $x = 0$ allora porremo $R = 0$. Se l'insieme dei valori x_1 per i quali la serie converge non è limitato, porremo $R = \infty$. In generale l'intervallo di convergenza della serie $\sum_{n=0}^{\infty} c_n (x - x_0)^n$ è l'intervallo aperto $(x_0 - R, x_0 + R)$ oppure l'intervallo chiuso $[x_0 - R, x_0 + R]$.

Per determinare il raggio di convergenza di una serie di potenze si può fare uso di uno dei criteri di convergenza studiati per le serie numeriche. Supponiamo per esempio, che esista e sia finito, il seguente limite

$$\lim_{n \rightarrow \infty} \left| \frac{c_{n+1}}{c_n} \right| = L,$$

e applichiamo il criterio del rapporto alla serie dei valori assoluti $\sum_{n=0}^{\infty} |c_n x^n|$. Si ha

$$\lim_{n \rightarrow \infty} \left| \frac{c_{n+1} x^{n+1}}{c_n x^n} \right| = \lim_{n \rightarrow \infty} \left| \frac{c_{n+1}}{c_n} \right| |x| = L|x|,$$

da cui segue che per $|x|L < 1$ la serie converge e per $|x|L > 1$ la serie diverge. dalla definizione di raggio di convergenza si ottiene allora $R = 1/L$. Si ricava così la seguente formula per il calcolo del raggio di convergenza

$$R = \lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right|.$$

• Esempi

1) Determinare l'intervallo di convergenza della serie di potenze $\sum_{n=1}^{\infty} (-1)^{n-1} n x^n$.

Calcoliamo il raggio di convergenza

$$R = \lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right| = \lim_{n \rightarrow \infty} \left| \frac{(-1)^{n-1} n}{(-1)^n (n+1)} \right| = \lim_{n \rightarrow \infty} \frac{n}{n+1} = 1.$$

Ne segue che la serie è assolutamente convergente per $-1 < x < 1$. Studiamo ora la convergenza nei punti estremi di questo intervallo. Per $x = -1$ si ottiene la serie

$$\sum_{n=1}^{\infty} (-1)^{n-1} n (-1)^n = \sum_{n=1}^{\infty} (-1)^{2n-1} n = - \sum_{n=1}^{\infty} n,$$

che è palesemente divergente. Per $x = 1$ si ha

$$\sum_{n=1}^{\infty} (-1)^{n-1} n (1)^n = \sum_{n=1}^{\infty} (-1)^{n-1} n,$$

che è una serie oscillante. Concludiamo che l'intervallo di convergenza della serie data è l'intervallo aperto $(-1, 1)$.

2) Determinare l'intervallo di convergenza della serie di potenze

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n 2^n} (x+3)^n.$$

Calcoliamo R .

$$R = \lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right| = \lim_{n \rightarrow \infty} \frac{(n+1) 2^{n+1}}{n 2^n} = \lim_{n \rightarrow \infty} 2 \left(1 + \frac{1}{n} \right) = 2.$$

Dunque la serie converge assolutamente nell'intervallo $|x+3| < 2$, ovvero per $-5 < x < -1$. Per $x = -5$ si ha la serie

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n 2^n} (-2)^n = \sum_{n=1}^{\infty} \frac{(-1)^{2n-1}}{n} = - \sum_{n=1}^{\infty} \frac{1}{n},$$

che risulta divergente. Per $x = -1$ si ha

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n 2^n} 2^n = \sum_{n=1}^{\infty} \frac{(-1)^n}{n},$$

che sappiamo essere una serie convergente. Ne segue che l'intervallo di convergenza della serie di partenza è $(-5, -1]$.

3) Determinare l'intervallo di convergenza della serie di potenze

$$\sum_{n=1}^{\infty} \frac{(-1)^n (x-2)^n}{n^n}.$$

Si ha

$$\begin{aligned} \lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right| &= \\ \lim_{n \rightarrow \infty} \frac{(n+1)^{n+1}}{n^n} &= \lim_{n \rightarrow \infty} \left(\frac{n+1}{n} \right)^n (n+1) = \lim_{n \rightarrow \infty} \left(1 + \frac{1}{n} \right)^n (n+1) = +\infty. \end{aligned}$$

Ne segue che il raggio di convergenza è infinito e la serie converge in tutto \mathbb{R} .

4) Determinare l'intervallo di convergenza della serie di potenze

$$\sum_{n=0}^{\infty} n! x^n.$$

Poiché si ha

$$\lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right| = \lim_{n \rightarrow \infty} \frac{n!}{(n+1)!} = \lim_{n \rightarrow \infty} \frac{1}{n+1} = 0,$$

il raggio di convergenza è zero. La serie converge allora solo per $x = 0$ e si ha $S(0) = 1$.

Facendo uso dei teoremi sulle serie di funzioni uniformemente convergenti, si può dimostrare che una serie di potenze convergente in $(x_0 - R, x_0 + R)$, converge assolutamente e uniformemente in ogni intervallo chiuso contenuto in $(x_0 - R, x_0 + R)$. Da questo risultato segue immediatamente che la somma di una serie di potenze è una funzione continua in un qualunque intervallo chiuso contenuto in $(x_0 - R, x_0 + R)$. Risulta inoltre che ogni serie di potenze è integrabile termine a termine e si ha

$$\int_0^x \sum_{n=0}^{\infty} c_n \xi^n d\xi = \sum_{n=0}^{\infty} \frac{c_n}{n+1} x^{n+1}.$$

L'intervallo di convergenza \bar{R} di quest'ultima serie è uguale a quello della serie di partenza, R . Infatti si ha

$$\bar{R} = \lim_{n \rightarrow \infty} \left| \frac{c_n}{n+1} \frac{n+2}{c_{n+1}} \right| = \lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right| \frac{n+2}{n+1} = \lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right| = R.$$

Si può inoltre dimostrare che le serie di potenze possono essere derivate termine a termine mantenendo il proprio raggio di convergenza, ovvero

$$\begin{aligned} S'(x) &= \frac{d}{dx} \sum_{n=0}^{\infty} c_n x^n = \sum_{n=1}^{\infty} n c_n x^{n-1}, \\ \bar{R} &= \lim_{n \rightarrow \infty} \left| \frac{n c_n}{(n+1) c_{n+1}} \right| = \lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right| \frac{n}{n+1} = \lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right| = R. \end{aligned}$$

Poiché la serie derivata è una serie di potenze con lo stesso raggio di convergenza essa può essere derivata un numero arbitrario di volte.

9.3 Serie di Taylor

Si dice che una funzione $f(x)$ è sviluppabile nella serie di potenze $\sum_{n=0}^{\infty} c_n(x-x_0)^n$ nell'intervallo $(x_0 - R, x_0 + R)$ se la serie di potenze converge e la sua somma è $f(x)$. Se la funzione è derivabile un numero arbitrario di volte, i coefficienti c_n dello sviluppo sono univocamente determinati da $f(x)$. Infatti, posto

$$f(x) = c_0 + c_1(x-x_0) + c_2(x-x_0)^2 + \dots + c_n(x-x_0)^n + \dots,$$

derivando n volte si ha $f^{(n)}(x_0) = n!c_n$ da cui

$$c_n = \frac{f^{(n)}(x_0)}{n!}.$$

Ne segue che se $f(x)$ è sviluppabile in serie di potenze, il suo sviluppo è

$$f(x) = f(x_0) + f'(x_0)(x-x_0) + \frac{1}{2}f''(x_0)(x-x_0)^2 + \dots + \frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n + \dots$$

Sia $f(x)$ una funzione derivabile un numero arbitrario di volte in $[a, b]$ e sia $x_0 \in (a, b)$. L'espressione

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x-x_0) + \frac{f''(x_0)}{2}(x-x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + \dots \\ &= \sum_{n=0}^{\infty} \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n, \end{aligned}$$

si dice *serie di Taylor* della funzione $f(x)$. Nel caso in cui $x_0 = 0$ questa serie prende anche il nome di *serie di MacLaurin*. Da quanto visto sinora, una funzione sviluppabile in serie di potenze ammette come sviluppo proprio la serie di Taylor. Viceversa, la serie di Taylor di una funzione $f(x)$ può non convergere a $f(x)$ e quindi può non essere lo sviluppo in serie di potenze di $f(x)$.

• Esempio

La funzione

$$f(x) = \begin{cases} e^{-1/x^2} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

è derivabile infinite volte in tutto \mathbb{R} e si ha $f^{(n)}(0) = 0, \forall n$. La sua serie di MacLaurin è dunque

$$0 + 0 \cdot x + 0 \cdot x^2 + \dots + 0 \cdot x^n + \dots = 0.$$

Ne segue che la funzione data non è sviluppabile in serie di Taylor.

Teorema 9.7 *Condizione sufficiente affinché una funzione $f(x)$ sia sviluppabile in serie di potenze (cioè in serie di Taylor), nell'intervallo $(x_0 - R, x_0 + R)$ è che ammetta derivate di qualsiasi ordine in $(x_0 - R, x_0 + R)$ e che in questo stesso intervallo tali derivate siano limitate.*

Dim. Consideriamo la formula di Taylor per la funzione $f(x)$ con resto nella forma di Lagrange

$$f(x) = \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k + \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1}, \quad \xi \in (x_0, x).$$

La dimostrazione del teorema consiste nel provare che

$$f(x) = \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k,$$

ovvero, che

$$\lim_{n \rightarrow \infty} \frac{f^{(n+1)}(\xi)}{(n+1)!} (x - x_0)^{n+1} = 0.$$

Sfruttando la limitatezza delle derivate, abbiamo $|f^{(n+1)}(x)| < M$, $\forall n$, $\forall x \in (x_0 - R, x_0 + R)$ e quindi

$$\left| \frac{(x - x_0)^{n+1}}{(n+1)!} f^{(n+1)}(\xi) \right| \leq M \frac{|x - x_0|^{n+1}}{(n+1)!} \leq M \frac{R^{n+1}}{(n+1)!}.$$

Ma la serie

$$\sum_{n=0}^{\infty} \frac{R^{n+1}}{(n+1)!},$$

converge in quanto è il primo resto di una serie esponenziale, quindi deve essere necessariamente

$$\lim_{n \rightarrow \infty} \frac{R^{n+1}}{(n+1)!} = 0.$$

Ne segue che il resto della serie è un infinitesimo per $n \rightarrow \infty$ e quindi che la serie di Taylor di $f(x)$ converge e ha per somma $f(x)$.

• Esempi

1) Serie esponenziale.

Consideriamo la funzione $f(x) = e^x$ che è definita in \mathbb{R} ed ha derivate di qualunque ordine. Vogliamo studiare la sua sviluppabilità in serie di MacLaurin. Preso $R \in \mathbb{R}^+$, per $x \in [-R, R]$ si ha

$$|f^{(n)}(x)| = e^x \leq e^R,$$

che comporta la limitatezza delle derivate di qualunque ordine. Per il teorema 9.7 la funzione è dunque sviluppabile in serie di potenze in $[-R, R]$. Ma questo risultato vale comunque si scelga R quindi $f(x) = e^x$ è sviluppabile in serie di MacLaurin in tutto \mathbb{R} . Si ricava così la nota serie esponenziale,

$$e^x = 1 + \frac{x}{1!} + \frac{x^2}{2!} + \dots + \frac{x^{n-1}}{(n-1)!} + \dots$$

2) Sviluppo di $\sin x$ e $\cos x$.

Consideriamo la funzione $f(x) = \sin x$ che è definita in tutto \mathbb{R} ed è C^∞ . Comunque preso $R \in \mathbb{R}^+$ si ha

$$|\sin x| \leq 1, \quad |\cos x| \leq 1, \quad \forall x \in [-R, R],$$

quindi, per il teorema 9.7 la funzione è sviluppabile in serie di MacLaurin in $[-R, R]$ e quindi in \mathbb{R} . Si ottiene

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots + \frac{(-1)^n x^{2n+1}}{(2n+1)!} + \dots$$

Un risultato analogo si ricava per la funzione $\cos x$ il cui sviluppo è

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots + \frac{(-1)^n x^{2n}}{(2n)!} + \dots$$

3) Serie binomiale.

Consideriamo la funzione $f(x) = (1+x)^\alpha$ con $\alpha \in \mathbb{R}$, per $x > -1$. Vogliamo esaminare la sviluppabilità di questa funzione in serie di MacLaurin. Osserviamo che, in un intorno dello zero le derivate della funzione non sono limitate. Per esempio, se scegliamo $x = -\frac{1}{2}$, le derivate di f per tale valore sono potenze di 2. Tali potenze crescono al crescere dell'ordine di derivazione in modo esponenziale. La condizione sufficiente del precedente teorema di sviluppabilità non è dunque soddisfatta. Tuttavia si può dimostrare che la funzione data è sviluppabile in serie di MacLaurin in tutto l'intervallo $(-1, 1)$.

Derivando la funzione e calcolando le derivate in $x = 0$, si ha

$$\begin{aligned} f(0) &= 1 \\ f'(0) &= \alpha \\ f''(0) &= \alpha(\alpha - 1) \\ &\vdots \\ f^{(n)}(0) &= \alpha(\alpha - 1)(\alpha - 2)\dots(\alpha - n + 1) \end{aligned}$$

La serie di MacLaurin di f è data quindi da

$$1 + \alpha x + \frac{\alpha(\alpha - 1)}{2!} x^2 + \dots + \frac{\alpha(\alpha - 1)\dots(\alpha - n + 1)}{n!} x^n + \dots$$

Calcoliamo il raggio di convergenza di questa serie.

$$R = \lim_{n \rightarrow \infty} \left| \frac{c_n}{c_{n+1}} \right| = \lim_{n \rightarrow \infty} \left| \frac{n+1}{\alpha - n} \right| = 1.$$

La serie risulta quindi convergente in $(-1, 1)$. Per dimostrare che la serie converge proprio a $f(x)$ osserviamo che si ha

$$(1+x)f'(x) = \alpha f(x), \quad f(0) = 1,$$

In altri termini, la funzione f soddisfa un problema di Cauchy del primo ordine. È facile verificare che la somma $S(x)$ della serie, soddisfa tale problema. Infatti si ha $S(0) = 1$. Inoltre la serie è uniformemente convergente in un qualunque intervallo chiuso contenuto in $(-1, 1)$ ed è quindi derivabile termine a termine. Si ha quindi

$$S'(x) = \alpha + \alpha(\alpha - 1)x + \dots + \frac{\alpha(\alpha - 1)\dots(\alpha - n + 1)}{(n-1)!} x^{n-1} + \dots$$

da cui

$$\begin{aligned} (1+x)S'(x) &= \alpha + \alpha(\alpha-1)x + \dots + \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{(n-1)!}x^{n-1} + \dots \\ &+ x \left[\alpha + \alpha(\alpha-1)x + \dots + \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{(n-1)!}x^{n-1} + \dots \right] \\ &= \alpha S(x) \end{aligned}$$

Poiché il precedente problema di Cauchy ammette una sola soluzione, concludiamo che la somma della serie coincide con la funzione $f(x)$. La funzione è dunque sviluppabile in serie di MacLaurin in tutto l'intervallo $(-1, 1)$. La serie così determinata si chiama *serie binomiale*. Essa si può porre nella forma

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n, \quad x \in (-1, 1),$$

dove i numeri $\binom{\alpha}{n}$, detti *coefficienti binomiali*, sono definiti da

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\dots(\alpha-n+1)}{n!}, \quad \binom{\alpha}{0} = 1,$$

4) Serie logaritmica.

Consideriamo la serie binomiale per $\alpha = -1$. Otteniamo lo sviluppo

$$\frac{1}{1+x} = 1 - x + x^2 - x^3 + \dots + (-1)^n x^n + \dots, \quad x \in (-1, 1).$$

Integrando ambo i membri ricaviamo

$$\int_0^x \frac{d\xi}{1+\xi} = \int_0^x (1 - \xi + \xi^2 - \xi^3 + \dots + (-1)^n \xi^n + \dots) d\xi,$$

da cui, per l'integrabilità termine a termine, otteniamo lo sviluppo

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \dots + \frac{(-1)^n x^{n+1}}{n+1} + \dots$$

che risulta valido per $x \in (-1, 1)$.

Testi per l'esercitazione e l'approfondimento

- [1] L. Recine, M. Romeo, *Esercizi di Analisi Matematica (vol. 2)*, Maggioli (2011).
- [2] M. Bertsch, R. Dal Passo, L. Giacomelli, *Analisi Matematica*, McGraw-Hill (2011).
- [3] E. Giusti, *Analisi Matematica 2*, Terza edizione, Bollati Boringhieri (2003).
- [4] V. Zorich, *Mathematical Analysis II*, Springer (2004).
- [5] A. Jeffrey, *Advanced Engineering Mathematics*, Academic Press (2002).