

Secondo semestre

1^a settimana - dal 25.02.19

Successioni numeriche

Definizione. Una *successione* in un insieme X è un'applicazione (o funzione) $f: \mathbb{N} \rightarrow X$. Se il codominio X è un sottoinsieme di \mathbb{R} , la successione si dice reale (o di numeri reali).

Data una successione $f: \mathbb{N} \rightarrow X$, per motivi di tradizione e di semplicità, l'immagine di un generico $n \in \mathbb{N}$ si denota col simbolo a_n , invece che con $f(n)$. Il valore a_n associato ad n si chiama *termine n -esimo* della successione o elemento di indice n . I numeri naturali sono detti gli *indici* della successione.

Vari modi di indicare una successione $f: \mathbb{N} \rightarrow X$ sono:

- $f: \mathbb{N} \rightarrow X$ (è quello più corretto ma il meno usato);
- $a_1, a_2, \dots, a_n, \dots$ (elencando i termini);
- $\{a_n\}_{n \in \mathbb{N}}$ (specificando che il dominio è \mathbb{N}) o anche $\{a_n\}$ (usato quando risulta chiaro dal contesto che rappresenta una successione).

Da ora in avanti, a meno che non sia diversamente specificato, ci occuperemo di successioni di numeri reali. In questo caso vale la convenzione che abbiamo adottato per le funzioni reali: *per semplicità, a meno che non sia detto esplicitamente, si assume che il codominio coincida con \mathbb{R} .*

Esempi di successioni reali:

- $1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots$
- $-1, 1, -1, \dots, (-1)^n, \dots$
- $2, 4, 6, \dots, 2n, \dots$
- $\frac{1}{2}, \frac{2}{3}, \frac{3}{4}, \dots, \frac{n}{n+1}, \dots$

Alcune successioni sono definite in modo *ricorsivo*, cioè:

- 1) assegnando il primo termine (o i primi k -termini);
- 2) dando una legge per ricavare il termine $(n+1)$ -esimo dai termini precedenti.

Un esempio di successione definita in modo ricorsivo è rappresentato dalla successione di Fibonacci (matematico pisano della prima metà del XIII secolo), dove sono assegnati i primi due termini, $a_1 = 1$ e $a_2 = 1$, ed ogni altro termine è somma dei due precedenti, cioè $a_{n+1} = a_n + a_{n-1}$. Tale successione governa alcuni fenomeni naturali come, ad esempio, il numero dei discendenti dei conigli e il numero degli antenati delle api (nelle varie generazioni).

Un altro esempio di successione ricorsiva si ottiene ponendo $b_n = a_{n+1}/a_n$, dove $\{a_n\}$ è la successione di Fibonacci. Si ha $b_1 = 1$ e $b_{n+1} = 1 + 1/b_n$. Si può provare che tale successione ha come limite (vedere sotto per la definizione di limite) il numero $(1 + \sqrt{5})/2$ che è detto *sezione aurea* e ha importanti applicazioni nell'architettura e nella pittura.

Osservazione. Attenzione a non fare confusione tra i *termini* della successione (che sono sempre infiniti: il primo, il secondo, il terzo, ecc.) e i *valori* assunti dalla successione che possono essere anche un numero finito. Ad esempio, la successione $\{a_n\}_{n \in \mathbb{N}} = \{(-1)^n\}_{n \in \mathbb{N}}$ assume solo i valori 1 e -1 ma, ciascuno, infinite volte (il primo termine è -1 , il secondo è 1, il terzo è -1 , \dots , l' n -esimo è $(-1)^n$, \dots).

Definizione. Si dice che una successione $\{a_n\}$ *tende* (o *converge*) a $l \in \mathbb{R}$, e si scrive $a_n \rightarrow l$ (per $n \rightarrow +\infty$), se per ogni $\epsilon > 0$ esiste un indice N (dipendente da ϵ) tale che per $n > N$ si ha $|a_n - l| < \epsilon$. Il numero l è detto il *limite* di $\{a_n\}$ e si scrive anche

$$\lim_{n \rightarrow +\infty} a_n = l \quad \text{oppure} \quad \lim a_n = l.$$

La seconda notazione, particolarmente sintetica, è giustificata dal fatto che $+\infty$ è l'unico punto di accumulazione per \mathbb{N} e quindi l'unica nozione di limite che ha senso per le successioni è per $n \rightarrow +\infty$. Osserviamo anche che nella definizione di limite affermare che “per ogni $\epsilon > 0$ esiste $N \in \mathbb{N}$ tale che per $n > N$ si abbia $|a_n - l| < \epsilon$ ” è equivalente a verificare che “per ogni $\epsilon > 0$ esiste $N \in \mathbb{R}$ tale che per $n > N$ si abbia $|a_n - l| < \epsilon$ ”. In altre parole è sufficiente che la disuguaglianza $|a_n - l| < \epsilon$ sia soddisfatta per tutti gli indici n maggiori di un certo numero reale N che può anche non essere un intero positivo.

Definizione. Si dice che $\{a_n\}$ *tende* (o *diverge*) a $+\infty$ [$-\infty$] (si scrive $a_n \rightarrow +\infty$ [$a_n \rightarrow -\infty$]) se per ogni $M \in \mathbb{R}$ esiste un indice N (dipendente da M) tale che per $n > N$ si ha $a_n > M$ [$a_n < M$].

Definizione. Una successione $\{a_n\}$ si dice *convergente* se ammette limite finito, *divergente* se il limite è $+\infty$ o $-\infty$ e *non regolare* quando non ammette limite.

Esempio. Verifichiamo che

$$\lim_{n \rightarrow +\infty} \frac{1}{n} = 0.$$

Dato $\epsilon > 0$, dobbiamo determinare $N \in \mathbb{N}$ tale che se $n > N$ si abbia

$$\left| \frac{1}{n} \right| < \epsilon \quad \text{o, equivalentemente,} \quad -\epsilon < \frac{1}{n} < \epsilon.$$

Ovviamente risulta $1/n > -\epsilon$ per ogni $n \in \mathbb{N}$ ed è sufficiente prendere $N = [1/\epsilon]$ perché se $n > N$ si abbia

$$\frac{1}{n} < \epsilon$$

(ricordiamo che $[1/\epsilon]$ denota la parte intera di $1/\epsilon$). Per quanto osservato sopra, se cerchiamo $N \in \mathbb{R}$ e non necessariamente $N \in \mathbb{N}$, sarà sufficiente prendere $N = 1/\epsilon$, senza ricorrere alla parte intera di tale numero.

Esercizio. Verificare, mediante la definizione di limite, che

$$\frac{n}{n+1} \rightarrow 1.$$

Svolgimento. Fissiamo $\epsilon > 0$. Occorre provare che esiste un $N \in \mathbb{N}$ tale che

$$n > N \quad \Rightarrow \quad \left| \frac{n}{n+1} - 1 \right| < \epsilon.$$

La disequazione è equivalente alla coppia di disequazioni

$$-\epsilon < \frac{n}{n+1} - 1 < \epsilon.$$

La disuguaglianza $\frac{n}{n+1} - 1 < \epsilon$ è vera per ogni $n \in \mathbb{N}$ in quanto il primo membro è negativo, mentre $-\epsilon < -\frac{1}{n+1}$ è vera per ogni $n \in \mathbb{N}$ che sia maggiore di $\frac{1}{\epsilon} - 1$.

Esercizio. Verificare, usando la definizione di limite, che

$$\lim_{n \rightarrow +\infty} \frac{3n+1}{n+2} = 3;$$

$$\lim_{n \rightarrow +\infty} \frac{n^3+2n+1}{2n^3-4} = \frac{1}{2},$$

$$\lim_{n \rightarrow +\infty} n^2+n = +\infty,$$

$$\lim_{n \rightarrow +\infty} -n^2+n = -\infty.$$

Esercizio. Provare che la successione $\{(-1)^n\}$ è non regolare.

Teorema. (Unicità del limite). *Il limite di una successione, se esiste (finito o infinito), è unico.*

I concetti di funzione limitata superiormente, limitata inferiormente, limitata sono già stati definiti per arbitrarie funzioni a valori reali. Essi rimangono pertanto validi anche per le successioni reali, essendo queste delle particolari funzioni a valori reali (l'unica distinzione riguarda il dominio, non il codominio). Ad esempio, diremo che la successione $\{a_n\}$ è *limitata* se esiste $M > 0$ tale che $|a_n| \leq M$ per ogni $n \in \mathbb{N}$.

Teorema. *Ogni successione convergente è limitata.*

Osservazione. Esistono successioni limitate ma non convergenti. Ad esempio, la successione $\{a_n\} = \{(-1)^n\}$.

Definizione. Diremo che una successione $\{a_n\}$ soddisfa una certa proprietà \mathcal{P} *definitivamente* se esiste $N \in \mathbb{N}$ tale che i termini a_n soddisfano \mathcal{P} per ogni $n > N$.

In altre parole, se $\{a_n\}$ soddisfa \mathcal{P} definitivamente significa che i termini della successione per i quali \mathcal{P} può non essere soddisfatta sono in numero **finito**.

Esempio. La successione di termine n -esimo

$$a_n = \frac{n-7}{n+3}$$

è definitivamente positiva essendo $a_n > 0$ per ogni $n > 7$.

Esempio. Una successione $\{a_n\}$ converge ad l se per ogni $\epsilon > 0$ si ha $|a_n - l| < \epsilon$ definitivamente.

I risultati che seguono sono stati già enunciati nel contesto delle proprietà dei limiti di funzioni reali di variabile reale. Per completezza li riportiamo interpretandoli nell'ambito delle successioni.

Teorema. (Permanenza del segno per le successioni). *Sia $\{a_n\}$ una successione reale. Se a_n tende ad un limite maggiore di 0 (anche $+\infty$), allora $a_n > 0$ definitivamente (cioè esiste $N \in \mathbb{N}$ tale che $a_n > 0$ per ogni $n > N$).*

Dimostrazione. Supponiamo che il limite l della successione $\{a_n\}$ sia finito e positivo. Fissato $\epsilon = l/2$, dalla definizione di limite si deduce che $l - \epsilon < a_n < l + \epsilon$ definitivamente. Quindi, essendo $l - \epsilon = l - l/2 = l/2 > 0$, si ottiene $0 < a_n$ definitivamente. Analizziamo ora il caso in cui il limite sia $+\infty$. Fissato un qualunque $M > 0$, si ha (sempre per la definizione di limite) $a_n > M$ definitivamente e quindi, a maggior ragione, $a_n > 0$ definitivamente. \square

Teorema. (Operazioni sui limiti per le successioni). Siano $\{a_n\}$ e $\{b_n\}$ due successioni tali che $a_n \rightarrow \lambda$ e $b_n \rightarrow \mu$, dove $\lambda, \mu \in \mathbb{R}^*$. Allora, quando ha senso nei reali estesi (cioè tranne i casi $\infty - \infty$, $0/0$, $0 \cdot \infty$ e ∞/∞), si ha:

- 1) $a_n + b_n \rightarrow \lambda + \mu$;
- 2) $a_n b_n \rightarrow \lambda \mu$;
- 3) $a_n/b_n \rightarrow \lambda/\mu$.

Riportiamo alcuni esempi per mostrare come non sia conveniente dare un senso, nei reali estesi, alle espressioni $\infty - \infty$, $0/0$, $0 \cdot \infty$ e ∞/∞ (dette *forme indeterminate*):

$$\begin{aligned} (\infty - \infty) \quad (n+1) - n &\rightarrow 1 \\ (\infty - \infty) \quad n^2 - n &\rightarrow +\infty \\ (0 \cdot \infty) \quad (1/n)n &\rightarrow 1 \\ (0 \cdot \infty) \quad (1/n)n^2 &\rightarrow +\infty \\ (0/0) \quad (1/n)/(1/n) &\rightarrow 1 \\ (0/0) \quad (1/n^2)/(1/n) &\rightarrow 0 \\ (\infty/\infty) \quad n/n &\rightarrow 1 \\ (\infty/\infty) \quad n^2/n &\rightarrow +\infty \end{aligned}$$

Esercizio. Usando il teorema precedente, calcolare

$$\lim_{n \rightarrow \infty} \frac{n^2 + n + 1}{6n^2 + 3n}; \quad \lim_{n \rightarrow \infty} n - \sqrt{n^2 + 1}.$$

Il risultato che segue è una facile conseguenza del teorema della permanenza del segno.

Teorema (del confronto dei limiti per le successioni). Se $a_n \rightarrow l \in \mathbb{R}$, $b_n \rightarrow m \in \mathbb{R}$ e $a_n \leq b_n$ definitivamente, allora $l \leq m$.

Dimostrazione. Se, per assurdo, l fosse maggiore di m , la successione $\{a_n - b_n\}$ tenderebbe al numero positivo $l - m$. Di conseguenza, per il teorema della permanenza del segno, risulterebbe $a_n - b_n > 0$ definitivamente, in contrasto con l'ipotesi " $a_n \leq b_n, \forall n \in \mathbb{N}$ ". \square

Un caso particolare del Teorema del confronto dei limiti per le successioni è il seguente corollario.

Corollario. Se $\{a_n\}$ è una successione reale a termini non negativi (rispettivamente non positivi) convergente ad l , allora risulta $l \geq 0$ (risp. $l \leq 0$).

Teorema (dei carabinieri per le successioni). Siano $\{a_n\}$, $\{b_n\}$ e $\{c_n\}$ tre successioni tali che

$$a_n \leq b_n \leq c_n.$$

Se $a_n \rightarrow l \in \mathbb{R}$ e $c_n \rightarrow l \in \mathbb{R}$, allora anche $b_n \rightarrow l$.

Dimostrazione. Fissiamo un $\epsilon > 0$. Poiché $a_n \rightarrow l$, esiste un n_1 tale che $l - \epsilon < a_n < l + \epsilon$ per tutti gli $n > n_1$. Per lo stesso ϵ è possibile determinare un n_2 tale che $l - \epsilon < c_n < l + \epsilon$ per $n > n_2$. Quindi, se $n > N := \max\{n_1, n_2\}$, allora $l - \epsilon < a_n \leq b_n \leq c_n < l + \epsilon$, da cui si ottiene $l - \epsilon < b_n < l + \epsilon$ per tutti gli $n > N$. \square

Osservazione. Osserviamo che il teorema dei carabinieri è ancora valido se si suppone che la condizione $a_n \leq b_n \leq c_n$ (ferme restando le altre ipotesi) valga definitivamente (non occorre sia vera per ogni $n \in \mathbb{N}$). L'importante è che i carabinieri $\{a_n\}$ e $\{c_n\}$ prima o poi catturino $\{b_n\}$ e si dirigano entrambi dalla stessa parte.

Esempio. La successione

$$\left\{ \frac{\text{sen } n}{n} \right\}$$

è rapporto di due successioni: $\{\text{sen } n\}$ e $\{n\}$. Il suo limite non si può determinare applicando il Teorema sulle operazioni dei limiti perché la successione $\{\text{sen } n\}$ non ha limite per $n \rightarrow \infty$ (cosa che noi non dimostreremo). D'altra parte si ha

$$-\frac{1}{n} \leq \frac{\text{sen } n}{n} \leq \frac{1}{n},$$

e quindi dal Teorema dei carabinieri si deduce

$$\frac{\text{sen } n}{n} \rightarrow 0.$$

Definizione. Una successione si dice *infinitesima* se è convergente a zero.

E' immediato verificare che

Teorema. La successione $\{a_n\}$ è infinitesima se e solo se la successione $\{|a_n|\}$ è infinitesima.

Osservazione. Nella proposizione precedente non si può sostituire 0 con un valore $l \neq 0$. Infatti, ad esempio, la successione $a_n = (-1)^n$ non ammette limite,

ma $|a_n| = 1$ e quindi converge a 1. In generale, vale però la seguente implicazione:

$$a_n \rightarrow l \Rightarrow |a_n| \rightarrow |l|.$$

Il seguente teorema è utile in molte occasioni:

Teorema. *La successione prodotto di una successione limitata per una infinitesima è infinitesima.*

Dimostrazione. Sia $\{a_n\}$ una successione limitata. Perciò esiste $M > 0$ tale che $|a_n| \leq M$ per ogni $n \in \mathbb{N}$. Sia b_n tale che $\lim_{n \rightarrow +\infty} b_n = 0$. Si ha

$$0 \leq |a_n b_n| \leq M |b_n|,$$

da cui, per il Teorema dei carabinieri e per il teorema precedente, si deduce che $\{a_n b_n\}$ è infinitesima. \square

Esempio. Dal teorema precedente si deduce immediatamente che le successioni

$$\left\{ \frac{(-1)^n}{n} \right\} \quad \text{e} \quad \left\{ \frac{\text{sen } n}{n^2 - n} \right\}$$

sono infinitesime.

Teorema (del carabiniere per le successioni). *Siano $\{a_n\}, \{b_n\}$ due successioni tali che*

$$a_n \leq b_n \quad \text{definitivamente.}$$

Se $a_n \rightarrow +\infty$ ($b_n \rightarrow -\infty$), allora anche $b_n \rightarrow +\infty$ ($a_n \rightarrow -\infty$).

Analogamente a ciò che si è visto per le funzioni, gli estremi inferiore e superiore di $\{a_n\}$ sono gli estremi inferiore e superiore dell'insieme dei valori assunti dalla successione e si denotano, rispettivamente, con

$$\inf_{n \in \mathbb{N}} a_n \quad \text{e} \quad \sup_{n \in \mathbb{N}} a_n,$$

o più semplicemente con $\inf a_n$ e $\sup a_n$. Se una successione $\{a_n\}$ non è limitata superiormente [inferiormente] si pone

$$\sup a_n = +\infty \quad [\inf a_n = -\infty].$$

Osservazione. La seguente caratterizzazione per l'estremo superiore [estremo inferiore] di una successione è analoga a quella data per una funzione reale:

$\sup a_n = l \in \mathbb{R}$ [$\inf a_n = l \in \mathbb{R}$] se e solo se

$$1) a_n \leq l, [a_n \geq l] \forall n \in \mathbb{N};$$

$$2) \forall \epsilon > 0 \text{ esiste un indice } n_\epsilon \text{ tale che } a_{n_\epsilon} > l - \epsilon [a_{n_\epsilon} < l + \epsilon].$$

Esercizio. Provare che

$$\sup_{n \in \mathbb{N}} (-1)^n \left(\frac{n-2}{n+3} \right) = 1 \quad \text{e} \quad \inf_{n \in \mathbb{N}} (-1)^n \left(\frac{n-2}{n+3} \right) = -1.$$

Suggerimento. Per provare la prima delle due uguaglianze usiamo la caratterizzazione richiamata sopra. La 1) è verificata essendo ovviamente

$$(-1)^n \left(\frac{n-2}{n+3} \right) \leq 1$$

per ogni $n \in \mathbb{N}$. Per quanto riguarda la 2) si tratta di verificare che per ogni $\epsilon > 0$ è possibile determinare un indice n_ϵ tale che

$$(-1)^{n_\epsilon} \left(\frac{n_\epsilon-2}{n_\epsilon+3} \right) > 1 - \epsilon.$$

È sufficiente prendere come n_ϵ un n pari e tale che $n > (3(1 - \epsilon) + 2)/\epsilon$.

Definizione. Una successione $\{a_n\}$ si dice *crescente* [*decescente*] se per ogni $m, n \in \mathbb{N}$ con $m < n$ si ha $a_m \leq a_n$ [$a_m \geq a_n$]. Una successione $\{a_n\}$ si dice *strettamente crescente* [*strettamente decrescente*] se per ogni $m, n \in \mathbb{N}$ con $m < n$ si ha $a_m < a_n$ [$a_m > a_n$]. Le successioni (strettamente) crescenti o (strettamente) decrescenti sono dette successioni (*strettamente*) *monotòne*.

Osservazione. È facile verificare che la definizione di successione monotona è equivalente alla seguente:

“ $\{a_n\}$ si dice *crescente* [*decescente*] se, per ogni $n \in \mathbb{N}$ risulta $a_n \leq a_{n+1}$ [$a_n \geq a_{n+1}$].”

Teorema. (Limite per successioni monotone). *Se $\{a_n\}$ è una successione monotona, allora ammette limite. Precisamente si ha $\lim a_n = \sup a_n$ se $\{a_n\}$ è crescente e $\lim a_n = \inf a_n$ se è decrescente. In particolare se $\{a_n\}$, oltre ad essere monotona, è anche limitata, allora è convergente, se invece non è limitata, è divergente.*

Dimostrazione. (Facoltativa) Assumiamo, per fissare le idee, che la successione $\{a_n\}$ sia crescente (il caso “ $\{a_n\}$ decrescente” è analogo).

Supponiamo prima che l'estremo superiore di $\{a_n\}$ sia finito e denotiamolo, per brevità, con la lettera l . Fissiamo un arbitrario $\epsilon > 0$. Poiché (per definizione di

estremo superiore) l è il minimo maggiorante per $\{a_n\}$, il numero $l - \epsilon$ non può essere un maggiorante per $\{a_n\}$. Non è vero quindi che tutti gli a_n verificano la condizione $a_n \leq l - \epsilon$. Ne esiste quindi (almeno) uno, denotiamolo $a_{\bar{n}}$, che non verifica tale condizione. Esiste cioè un indice \bar{n} per il quale risulta $a_{\bar{n}} > l - \epsilon$ (proprietà (2) della caratterizzazione dell'estremo superiore). Dato che abbiamo supposto $\{a_n\}$ crescente, se n è un qualunque indice maggiore di \bar{n} , si ha $a_{\bar{n}} \leq a_n$ e quindi, a maggior ragione, $l - \epsilon < a_n$. D'altra parte l è un maggiorante per gli a_n e, di conseguenza, per ogni n (e non solo per quelli maggiori di \bar{n}) risulta $a_n \leq l$ (proprietà (1) della caratterizzazione dell'estremo superiore). In conclusione, possiamo affermare che per gli $n > \bar{n}$ si ha $l - \epsilon < a_n < l + \epsilon$, e quindi, per la definizione di limite, $a_n \rightarrow l = \sup a_n$.

Supponiamo ora $\sup a_n = +\infty$ e fissiamo un $M > 0$. Poiché (in base al significato della notazione $\sup a_n = +\infty$) la successione non è limitata superiormente, il numero M non può essere un maggiorante per tutti gli a_n . Esiste quindi un indice \bar{n} per il quale risulta $a_{\bar{n}} > M$. Dato che la successione è crescente, quando $n > \bar{n}$ si ha $a_n > M$. Dunque, per la definizione di limite, $a_n \rightarrow +\infty = \sup a_n$.
 \square

Osservazione. La condizione espressa dal teorema precedente è ovviamente solo sufficiente. Ad esempio la successione $\{a_n\} = \left\{\frac{(-1)^n}{n}\right\}$ è convergente (infatti tende a 0), ma non è monotona. Si ha $\sup a_n = \max a_n = 1/2$ e $\inf a_n = \min a_n = -1$.

2^a settimana - dal 04.03.19

Un esempio importante di successione monotona è il seguente:

Esempio. Il numero e .

Consideriamo la successione il cui termine n -esimo è

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

Si può provare che essa è:

- strettamente crescente;
- limitata superiormente (3 è un maggiorante).

Perciò, per il teorema del limite per le successioni monotone, è convergente. Il suo limite si denota con e ed è detto *numero di Nepero*. Tale numero è irrazionale e un suo valore approssimato (a meno di 10^{-9}) è dato da 2,7182818284. Notiamo che questa è una delle possibili definizioni del numero e (un'altra, che vedremo in seguito, è quella basata sulla nozione di serie).

Il logaritmo naturale (o in base e) di un numero x si denota $\ln x$ o $\log x$.

Esempio. Proviamo che

$$\lim_{n \rightarrow +\infty} \left(1 - \frac{1}{n}\right)^n = \frac{1}{e}.$$

Si ha

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

D'altra parte

$$\left(\frac{n}{n-1}\right)^n = \left(\frac{n-1+1}{n-1}\right)^{n-1+1} = \left[\left(1 + \frac{1}{n-1}\right)^{n-1} \left(1 + \frac{1}{n-1}\right)\right] \rightarrow e.$$

Esempio. Proviamo che

$$\lim_{x \rightarrow \pm\infty} \left(1 + \frac{1}{x}\right)^x = e.$$

Per $x > 0$, si ha

$$\left(1 + \frac{1}{[x]+1}\right)^{[x]} \leq \left(1 + \frac{1}{x}\right)^x \leq \left(1 + \frac{1}{[x]}\right)^{[x]+1}.$$

D'altra parte, dal limite della successione $\{(1 + 1/n)^n\}$ considerato sopra, si deduce che il primo e l'ultimo membro della precedente disequazione tendono a e . Di conseguenza, per il Teorema dei carabinieri, si ottiene che il limite per $x \rightarrow +\infty$ è uguale a e . Applicando il Teorema di cambiamento di variabile per i limiti e tenendo conto del fatto che $(1 - 1/n)^n \rightarrow 1/e$, si verifica che anche il limite per $x \rightarrow -\infty$ è uguale a e .

Inoltre, ponendo $x = 1/y$ e applicando nuovamente il Teorema di cambiamento di variabile, si deduce che

$$\lim_{x \rightarrow 0} (1 + x)^{1/x} = e.$$

Infine, tenendo conto della continuità della funzione logaritmo, si ottiene

$$\lim_{x \rightarrow 0} \frac{\log(1 + x)}{x} = 1.$$

Infatti, si ha

$$\lim_{x \rightarrow 0} \frac{\log(1 + x)}{x} = \lim_{x \rightarrow 0} \log(1 + x)^{1/x} = \log e = 1.$$

Il seguente risultato mette in relazione il limite di successione e quello di funzione.

Teorema (“di collegamento”). *Sia $f: X \subseteq \mathbb{R} \rightarrow \mathbb{R}$ e sia $\alpha \in \mathbb{R}^*$ un punto di accumulazione per X . Allora $\lim_{x \rightarrow \alpha} f(x) = \lambda \in \mathbb{R}^*$ se e solo se per ogni successione $\{a_n\}_{n \in \mathbb{N}}$, $a_n \in X$, $a_n \neq \alpha$, $a_n \rightarrow \alpha$ si ha $\lim_{n \rightarrow \infty} f(a_n) = \lambda$.*

Usando il teorema di collegamento e i limiti di funzione provati in precedenza si ottiene immediatamente che:

se $\{a_n\}$ è una qualunque successione infinitesima, allora

$$\lim_{n \rightarrow \infty} \frac{\sin a_n}{a_n} = 1, \quad \lim_{n \rightarrow \infty} \frac{1 - \cos a_n}{a_n^2} = \frac{1}{2}$$

e

$$\lim_{n \rightarrow \infty} \frac{\log(1 + a_n)}{a_n} = 1.$$

In maniera analoga:

se $\{a_n\}$ è una qualunque successione divergente ($a + \infty$ oppure $a - \infty$), allora

$$\lim_{n \rightarrow \infty} \left(1 + \frac{1}{a_n}\right)^{a_n} = e.$$

Da quest'ultimo risultato si deduce facilmente che, fissato $k \in \mathbb{R}$,

$$\lim_{n \rightarrow \infty} \left(\frac{n+k+1}{n+k} \right)^n = e, \quad \lim_{n \rightarrow \infty} \left(\frac{n+k}{n+k+1} \right)^n = \frac{1}{e}.$$

Esempio. Usando il teorema di collegamento si può provare che non esiste il limite per $x \rightarrow +\infty$ di $\sin x$. Basta infatti considerare le successioni di termine n -esimo

$$a_n = \frac{\pi}{2} + 2n\pi, \quad \bar{a}_n = \frac{3}{2}\pi + 2n\pi$$

e osservare che $\lim_{n \rightarrow \infty} \sin a_n = 1$ mentre $\lim_{n \rightarrow \infty} \sin \bar{a}_n = -1$.

Nel calcolo dei limiti di successione sono utili i seguenti teoremi.

Teorema. Sia data una successione $\{a_n\}$ a termini positivi. Allora si ha

$$\lim_{n \rightarrow +\infty} \sqrt[n]{a_n} = \lim_{n \rightarrow +\infty} \frac{a_{n+1}}{a_n},$$

purché esista il limite al secondo membro.

Esempio. Usando il teorema precedente si deduce immediatamente che

$$\sqrt[n]{h} \rightarrow 1, (h > 1); \quad \sqrt[n]{n} \rightarrow 1; \quad \sqrt[n]{n!} \rightarrow +\infty; \quad \sqrt[n]{\frac{n^n}{n!}} \rightarrow e.$$

Ad esempio, si ha

$$\lim_{n \rightarrow +\infty} \sqrt[n]{n} = \lim_{n \rightarrow +\infty} \frac{n+1}{n} = 1.$$

$$\lim_{n \rightarrow +\infty} \sqrt[n]{\frac{n^n}{n!}} = \lim_{n \rightarrow +\infty} \frac{(n+1)^{n+1}}{(n+1)!} \frac{n!}{n^n} = \lim_{n \rightarrow +\infty} \left(\frac{n+1}{n} \right)^n = e.$$

Teorema (criterio del rapporto per le successioni). Sia $\{a_n\}$ una successione a termini positivi. Supponiamo che la successione $\{b_n\} = \{a_{n+1}/a_n\}$ ottenuta da $\{a_n\}$ facendo il rapporto tra un termine e il precedente ammetta limite β (finito o infinito). Allora, se $\beta < 1$ la successione $\{a_n\}$ è infinitesima, se $\beta > 1$ la successione $\{a_n\}$ è infinita.

Esempio. Proviamo che

$$\lim_{n \rightarrow +\infty} r^n = \begin{cases} 0 & \text{se } |r| < 1 \\ 1 & \text{se } r = 1 \\ +\infty & \text{se } r > 1 \\ \cancel{\exists} & \text{se } r \leq -1 \end{cases}$$

Dimostrazione. Consideriamo ad esempio il caso $|r| < 1$. Per il criterio del rapporto, essendo

$$\lim_{n \rightarrow +\infty} \frac{|r^{n+1}|}{|r^n|} = |r| < 1,$$

si ottiene $|r^n| \rightarrow 0$. Di conseguenza, poiché come visto in precedenza una successione è infinitesima se e solo se lo è il suo valore assoluto, allora anche $r^n \rightarrow 0$. Gli altri casi sono lasciati per esercizio. \square

Esempio. Facendo uso del criterio del rapporto si deduce che

$$\frac{n^\alpha}{r^n} \rightarrow 0 \ (\alpha > 0, r > 1), \quad \frac{r^n}{n!} \rightarrow 0, \quad \frac{n!}{n^n} \rightarrow 0.$$

Ad esempio, si ha

$$\lim_{n \rightarrow +\infty} \frac{(n+1)!}{(n+1)^{n+1}} \frac{n^n}{n!} = \lim_{n \rightarrow +\infty} \left(\frac{n}{n+1} \right)^n = \frac{1}{e} < 1,$$

da cui, per il criterio del rapporto, $n!/n^n \rightarrow 0$.

Esempio. Si ha la seguente scala di infiniti:

$$\log n, \quad n^\alpha (\alpha > 0), \quad r^n (r > 1), \quad n!, \quad n^n.$$

La scala va intesa nel senso che il limite tra un infinito e il successivo è 0.

Serie numeriche

Prendiamo ora in esame un esempio particolarmente importante di successione reale. Data una successione reale $\{a_n\}_{n \in \mathbb{N}}$ associamo ad essa la successione $\{s_n\}_{n \in \mathbb{N}}$ di termine n -esimo

$$s_n = a_1 + a_2 + \dots + a_n = \sum_{k=1}^n a_k.$$

Ci interessa studiare il limite

$$\lim_{n \rightarrow +\infty} s_n.$$

Tale limite si indica con il simbolo

$$\sum_{k=1}^{\infty} a_k,$$

detto *serie (numerica)*, e si legge *somma per k che va da 1 a infinito di a_k* . Ovviamente al posto di k si può usare un qualunque altro indice (si pensi al significato di sommatoria).

In altre parole, col simbolo

$$\sum_{n=1}^{\infty} a_n,$$

si intende

$$\lim_{n \rightarrow +\infty} \sum_{k=1}^n a_k.$$

La successione $\{s_n\}$ si dice *successione delle somme parziali* (o *delle ridotte*) della serie, mentre a_n è detto il *termine generale*. Il carattere della serie è, per definizione, il carattere della successione $\{s_n\}$. In altre parole, si dice che la serie *converge* [*diverge*] se $\{s_n\}$ converge [diverge]; se il limite di $\{s_n\}$ non esiste, la serie è *indeterminata* (o *irregolare*). Il limite s (finito o infinito) di $\{s_n\}$, quando esiste, si dice *somma della serie* e si scrive

$$s = \sum_{n=1}^{\infty} a_n.$$

Talvolta, invece di sommare a partire da $n = 1$, si parte da un indice $n_0 \in \mathbb{N}$ (può essere utile anche $n_0 = 0$). Scriveremo allora

$$\sum_{n=n_0}^{\infty} a_n.$$

Si vede facilmente che il carattere di una serie non cambia se si modifica un numero finito di termini.

La serie geometrica (Esempio importante.)

La serie

$$\sum_{n=0}^{\infty} r^n$$

è detta *geometrica*. Osserviamo che in una serie geometrica il rapporto tra un termine e il precedente è costante (ossia, non dipende da n) e vale r . Tale rapporto si chiama *ragione* della serie. *Una serie geometrica converge se e solo se la sua ragione è in valore assoluto minore di uno. Inoltre, la somma è $1/(1-r)$.*

Si ha infatti

$$s_n = 1 + r + r^2 + \dots + r^n.$$

Moltiplicando per r si ottiene

$$r s_n = r + r^2 + \dots + r^{n+1},$$

da cui, sottraendo membro a membro e ricavando s_n , risulta

$$s_n = \frac{1 - r^{n+1}}{1 - r}.$$

Perciò, $\{s_n\}$ è convergente se e solo se $|r| < 1$ (infatti, in tal caso, $r^{n+1} \rightarrow 0$). Di conseguenza, se $|r| < 1$, si ha

$$\sum_{n=0}^{\infty} r^n = \lim_{n \rightarrow +\infty} \frac{1 - r^{n+1}}{1 - r} = \frac{1}{1 - r},$$

o, anche,

$$\sum_{n=1}^{\infty} r^n = r \sum_{n=0}^{\infty} r^n = \frac{r}{1 - r}.$$

Ad esempio, per $r = 1/2$, si ottiene

$$\sum_{n=1}^{\infty} \frac{1}{2^n} = 1.$$

A titolo di esempio, consideriamo il numero decimale periodico $3,1\bar{7} = 3,17777\dots$. Si può scrivere

$$3,1\bar{7} = 3,1 + 0,07 + 0,007 + 0,0007 + \dots = 3,1 + \frac{7}{10^2} + \frac{7}{10^3} + \frac{7}{10^4} + \dots$$

Ovvero

$$3,1\bar{7} = \frac{31}{10} + \sum_{n=2}^{\infty} \frac{7}{10^n}.$$

Si ha

$$\sum_{n=2}^{\infty} \frac{7}{10^n} = \frac{7}{100} \sum_{n=0}^{\infty} \frac{1}{10^n} = \frac{7/100}{1 - 1/10},$$

essendo $\sum_{n=0}^{\infty} 1/10^n$ una serie geometrica di ragione $1/10$. Si ha pertanto

$$3,1\bar{7} = \frac{31}{10} + \frac{7/100}{1 - 1/10} = \frac{286}{90}.$$

In modo analogo si prova che ogni numero decimale periodico è razionale e se ne determina la frazione generatrice.

Esempio. La serie $\sum_{n=1}^{\infty} (-1)^n$ è indeterminata. Infatti si ha $s_n = 0$ se n è pari e $s_n = -1$ se n è dispari. Perciò $\lim_{n \rightarrow +\infty} s_n$ non esiste.

Teorema. *Condizione necessaria affinché una serie sia convergente è che il termine generale tenda a zero.*

Dimostrazione. Sia $\sum_{n=n_0}^{\infty} a_n$ una serie convergente. Ciò significa, per definizione, che la successione $\{s_n\}$ delle somme parziali converge ad un numero (finito) s . Osserviamo che $a_n = s_n - s_{n-1}$ e che (oltre ad $\{s_n\}$) anche $\{s_{n-1}\}$ converge ad s (infatti, se $|s_n - s| < \epsilon$ per $n > N$, allora $|s_{n-1} - s| < \epsilon$ per $n > N + 1$). Si ha allora

$$\lim_{n \rightarrow \infty} a_n = \lim_{n \rightarrow \infty} (s_n - s_{n-1}) = s - s = 0. \quad \square$$

Osservazione. La precedente condizione non è sufficiente. Proveremo infatti, facendo uso del criterio dell'integrale che enunceremo successivamente, che la serie $\sum_{n=1}^{\infty} 1/n$ (detta *armonica*) non è convergente, sebbene il suo termine generale sia infinitesimo.

Esercizio. Sia $\sum_{n=n_0}^{\infty} a_n$ una serie convergente e sia $c \in \mathbb{R}$. Provare che

$$\sum_{n=n_0}^{\infty} ca_n = c \sum_{n=n_0}^{\infty} a_n.$$

Esercizio. Siano $\sum_{n=n_0}^{\infty} a_n$ e $\sum_{n=n_0}^{\infty} b_n$ due serie convergenti. Provare che

$$\sum_{n=n_0}^{\infty} (a_n + b_n) = \sum_{n=n_0}^{\infty} a_n + \sum_{n=n_0}^{\infty} b_n.$$

Osservazione (importante). Se una serie $\sum_{n=n_0}^{\infty} a_n$ è a termini non negativi (ossia $a_n \geq 0$ per ogni n), allora la successione $\{s_n\}$ delle sue somme parziali è crescente. Infatti, essendo $a_n \geq 0$, risulta

$$s_n = s_{n-1} + a_n \geq s_{n-1}.$$

Perciò, per il teorema sul limite delle successioni monotone, il limite di $\{s_n\}$ esiste sempre (finito o infinito) dal momento che coincide con l'estremo superiore di $\{s_n\}$. In tal caso, la somma della serie è ben definita e rappresenta un numero reale esteso (ovviamente positivo). In altre parole, *una serie a termini non negativi o è convergente o è divergente*. Analoghe considerazioni valgono, ovviamente, anche per le serie a termini non positivi.

Esempio. Consideriamo la serie

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha}, \quad \alpha > 0,$$

detta serie *armonica generalizzata*. Si può provare (ad esempio mediante il criterio dell'integrale che enunceremo dopo) che la serie converge se e solo se $\alpha > 1$. Di conseguenza, essendo a termini positivi, essa diverge per $0 < \alpha \leq 1$.

Talvolta non ha interesse o non si riesce a calcolare esplicitamente la somma di una serie; di solito si cerca di stabilirne il carattere. Successivamente, dopo aver provato che una serie converge, la sua somma (se interessa) potrà essere stimata con metodi numerici mediante l'ausilio di un computer.

Diamo ora un criterio utile per stabilire il carattere di una serie **a termini positivi**.

Criterio del confronto per le serie a termini positivi. *Siano $\sum_{n=n_0}^{\infty} a_n$ e $\sum_{n=n_0}^{\infty} b_n$ due serie a termini non negativi. Supponiamo $a_n \leq b_n$ per $n \geq n_0$. Allora, se converge la serie $\sum_{n=n_0}^{\infty} b_n$ (detta maggiorante), converge anche (la minorante) $\sum_{n=n_0}^{\infty} a_n$ e si ha*

$$\sum_{n=n_0}^{\infty} a_n \leq \sum_{n=n_0}^{\infty} b_n ;$$

se diverge la minorante, diverge anche la maggiorante.

Dimostrazione. Poiché le due serie sono a termini non negativi, le successioni delle somme parziali

$$s_n = \sum_{k=n_0}^n a_k \quad \text{e} \quad \sigma_n = \sum_{k=n_0}^n b_k$$

risultano crescenti e, conseguentemente, per il teorema del limite di successioni monotone, per entrambe esiste (finito o infinito) il limite per $n \rightarrow +\infty$. Poiché $a_n \leq b_n$ per $n \geq n_0$, si deduce $s_n \leq \sigma_n$ per ogni $n \geq n_0$. Perciò, se $\sum_{n=n_0}^{\infty} b_n$ è convergente (cioè se la successione $\{\sigma_n\}$ è convergente), dal teorema del confronto dei limiti, si ottiene che la successione $\{s_n\}$ e, quindi, la serie $\sum_{n=n_0}^{\infty} a_n$ sono convergenti. Di conseguenza, se $\sum_{n=n_0}^{\infty} a_n$ diverge, la serie $\sum_{n=n_0}^{\infty} b_n$ non può convergere e quindi necessariamente diverge (ricordarsi che esiste il limite di σ_n). \square

Esercizio. Proviamo che la serie

$$\sum_n \frac{\text{sen}^2 n}{n^3 + n}$$

è convergente. Si ha

$$\frac{\text{sen}^2 n}{n^3 + n} \leq \frac{1}{n^3 + n} \leq \frac{1}{n^3} .$$

La conclusione segue dal criterio del confronto, ricordando che $\sum_n \frac{1}{n^3}$ è convergente.

Esercizio. Proviamo che la serie

$$\sum_n \frac{2 + (-1)^n}{2^n}$$

è convergente. Infatti si ha

$$\frac{2 + (-1)^n}{2^n} \leq \frac{3}{2^n}$$

e $\sum_n (\frac{1}{2})^n$, essendo una serie geometrica di ragione $\frac{1}{2}$, è convergente.

Un utile corollario del criterio del confronto è il seguente:

Criterio del confronto asintotico per le serie a termini positivi. Sia $\sum_{n=n_0}^{\infty} a_n$ una serie a termini non negativi. Supponiamo che per qualche $\alpha > 0$ si abbia

$$\lim_{n \rightarrow +\infty} n^\alpha a_n = \lambda \in \mathbb{R}^*.$$

Allora,

1. se $0 < \lambda < +\infty$ la serie ha lo stesso carattere di $\sum_n \frac{1}{n^\alpha}$ (cioè converge se e solo se $\alpha > 1$);
2. se $\lambda = 0$ e $\alpha > 1$ la serie converge;
3. se $\lambda = +\infty$ e $\alpha \leq 1$ la serie diverge.

Esempio. Consideriamo

$$\sum_n \frac{1}{n+2} \operatorname{arctang} \frac{1}{n}.$$

Poiché

$$\lim_{n \rightarrow +\infty} n^2 \frac{1}{n+2} \operatorname{arctang} \frac{1}{n} = \lim_{n \rightarrow +\infty} n^2 \frac{1}{n+2} \left(\frac{1}{n} + o\left(\frac{1}{n}\right) \right) = 1,$$

per il punto 1) del criterio del confronto asintotico, la serie data ha lo stesso carattere di $\sum_n \frac{1}{n^2}$ e quindi è convergente.

Esempio. Consideriamo

$$\sum_n \frac{\frac{1}{n} - \operatorname{sen} \frac{1}{n}}{\operatorname{tang} \frac{1}{n^2}}.$$

Usando la formula di Taylor si ha

$$\frac{\frac{1}{n} - \left(\frac{1}{n} - \frac{1}{6n^3} + o\left(\frac{1}{n^3}\right)\right)}{\frac{1}{n^2} + o\left(\frac{1}{n^2}\right)} = \frac{\frac{1}{6n^3} + o\left(\frac{1}{n^3}\right)}{\frac{1}{n^2} + o\left(\frac{1}{n^2}\right)}.$$

Perciò, il termine generale della serie data è un infinitesimo di ordine 1. Di conseguenza, per il punto 1) del criterio del confronto asintotico, la serie diverge avendo lo stesso carattere della serie armonica.

Esempio. Consideriamo

$$\sum_n n e^{-n}.$$

Poiché $\lim_{n \rightarrow +\infty} n^{\alpha+1}/e^n = 0$ per ogni $\alpha > 0$ e quindi, in particolare, anche per ogni $\alpha > 1$, per il punto 2) del criterio del confronto asintotico, la serie data converge.

Esempio. Consideriamo

$$\sum_n \frac{1}{\log n}.$$

Poiché $\lim_{n \rightarrow +\infty} n^\alpha / \log n = +\infty$ per ogni $\alpha > 0$ e quindi, in particolare, anche per ogni $\alpha \leq 1$, per il punto 3) del criterio del confronto asintotico, la serie data diverge.

Un utile criterio per stabilire il carattere di una serie numerica a **termini positivi** è il seguente:

Criterio dell'integrale. Sia $n_0 \in \mathbb{N}$ e sia $f: [n_0, +\infty) \rightarrow \mathbb{R}$ una funzione continua positiva e decrescente. Allora

$$\sum_{n=n_0}^{\infty} f(n) \quad e \quad \int_{n_0}^{+\infty} f(x) dx$$

hanno lo stesso carattere. Più precisamente si ha

$$\int_{n_0+1}^{+\infty} f(x) dx \leq \sum_{n=n_0+1}^{\infty} f(n) \leq \int_{n_0}^{+\infty} f(x) dx.$$

Esempio. Dal criterio dell'integrale si deduce immediatamente che la serie armonica

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

è divergente. Più precisamente, sempre usando il criterio dell'integrale, è facile verificare che la serie armonica generalizzata

$$\sum_{n=1}^{\infty} \frac{1}{n^{\alpha}},$$

converge se e solo se $\alpha > 1$. È sufficiente infatti ricordare che l'integrale improprio

$$\int_1^{+\infty} \frac{1}{x^{\alpha}} dx$$

è convergente se e solo se $\alpha > 1$.

Esercizio. Usando il criterio dell'integrale, provare che

$$\sum_{n=2}^{\infty} \frac{1}{n \log n} \quad \text{e} \quad \sum_{n=1}^{\infty} \frac{\log^2 n}{n}$$

sono divergenti.

3^a settimana - dal 11.03.19

Se la serie $\sum_{n=n_0}^{\infty} a_n$ non è di segno costante, in alcuni casi è utile considerare la serie dei valori assoluti, cioè $\sum_{n=n_0}^{\infty} |a_n|$. Vale il seguente risultato

Criterio della convergenza assoluta. *Se converge $\sum_{n=n_0}^{\infty} |a_n|$, converge anche $\sum_{n=n_0}^{\infty} a_n$ e vale la disuguaglianza*

$$\left| \sum_{n=n_0}^{\infty} a_n \right| \leq \sum_{n=n_0}^{\infty} |a_n|.$$

Osservazione. Si osservi che la disuguaglianza nel precedente teorema ha senso in virtù dell'affermazione che la serie $\sum_{n=n_0}^{\infty} a_n$ è convergente. In questo caso, infatti, tale serie rappresenta un numero reale, e quindi ha senso il suo valore assoluto.

Definizione. La serie $\sum_{n=n_0}^{\infty} a_n$ si dice *assolutamente convergente* se è convergente la serie $\sum_{n=n_0}^{\infty} |a_n|$.

Osservazione. Tenendo conto della definizione precedente, il criterio della convergenza assoluta si può enunciare affermando che *se una serie è assolutamente convergente, allora è convergente*. Esistono anche serie convergenti ma non assolutamente convergenti. Ad esempio, la serie $\sum_{n=1}^{\infty} (-1)^n \frac{1}{n}$ è convergente (ciò si può dimostrare facendo uso del criterio di Leibniz che vedremo dopo). D'altra parte, la serie dei valori assoluti $\sum_{n=1}^{\infty} |(-1)^n \frac{1}{n}|$, che è la serie armonica, come sappiamo non converge. Un altro esempio di serie convergente ma non assolutamente convergente è $\sum_{n=1}^{\infty} \frac{\sin n}{n}$.

Il criterio del confronto per le serie a termini positivi è ovviamente un criterio di convergenza assoluta. Infatti, data la serie $\sum_{n=n_0}^{\infty} a_n$ esso si può applicare (come mostrato nell'esempio che segue) alla serie $\sum_{n=n_0}^{\infty} |a_n|$ per studiarne il carattere.

Esempio.

1) La serie

$$\sum_{n=1}^{\infty} \frac{\cos n}{n^3}.$$

è convergente, come si verifica subito applicando il criterio del confronto a

$$\sum_{n=1}^{\infty} \frac{|\cos n|}{n^3}$$

e applicando successivamente alla serie data il criterio della convergenza assoluta.

2) Consideriamo la serie

$$\sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n \operatorname{sen} \frac{\pi}{2} n.$$

Si ha

$$\left(\frac{2}{3}\right)^n \left| \operatorname{sen} \frac{\pi}{2} n \right| \leq \left(\frac{2}{3}\right)^n$$

e quest'ultimo è il termine n -esimo di una serie geometrica di ragione $2/3$ (quindi convergente). Di conseguenza,

$$\sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n \left| \operatorname{sen} \frac{\pi}{2} n \right|$$

è convergente e, quindi, per il criterio della convergenza assoluta,

$$\sum_{n=0}^{\infty} \left(\frac{2}{3}\right)^n \operatorname{sen} \frac{\pi}{2} n$$

è convergente.

Una serie del tipo

$$\sum_{n=n_0}^{\infty} (-1)^n a_n, \quad a_n \geq 0,$$

si dice *a segni alterni*. Per le serie a segni alterni si può dimostrare il seguente criterio di convergenza:

Criterio di Leibniz. Sia $\sum_{n=n_0}^{\infty} (-1)^n a_n$ una serie a segni alterni. Se la successione $\{a_n\}$ è decrescente ed è infinitesima, allora la serie è convergente. Inoltre, denotata con s la sua somma e con s_n la sua somma parziale n -esima, risulta $|s - s_n| < a_{n+1}$, $\forall n \in \mathbb{N}$, ossia l'errore che si commette nella valutazione di s arrestandosi alla somma parziale n -esima è inferiore, in valore assoluto, al valore assoluto del primo termine trascurato.

A titolo di esempio osserviamo che la serie

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n}$$

soddisfa le due ipotesi del criterio di Leibniz e pertanto è convergente. D'altra parte, come già osservato in precedenza, essa non è assolutamente convergente. Inoltre, detta s la sua somma, risulta

$$\left| s - \sum_{k=1}^n \frac{(-1)^k}{k} \right| < \frac{1}{n+1}.$$

Serie di funzioni

Una serie del tipo

$$\sum_{n=n_0}^{\infty} f_n(x),$$

dove le f_n sono funzioni reali di variabile reale, è detta *serie di funzioni* (reali di variabile reale). Si dice che la serie è definita in un insieme $X \subseteq \mathbb{R}$ se il dominio di tutte le funzioni contiene X (ossia, se sono tutte definite per ogni $x \in X$). Ovviamente, ogni volta che si fissa $x \in X$, si ottiene una serie numerica che può essere convergente oppure no, a seconda che la successione di funzioni $\{s_n(x)\}$ delle somme parziali n -esime (dove $s_n(x) = \sum_{k=n_0}^n f_k(x)$) sia convergente o meno. L'insieme dei numeri $x \in X$ per cui la serie converge si chiama *insieme di convergenza (puntuale)*. Ad esempio, l'insieme di convergenza della serie

$$\sum_{n=0}^{\infty} x^n$$

è l'intervallo $(-1, 1)$, visto che si tratta di una serie geometrica di ragione $x \in \mathbb{R}$.

In analogia con le serie numeriche, se $\sum_{n=n_0}^{\infty} |f_n(x)|$ converge in un punto x , diremo che la serie di funzioni $\sum_{n=n_0}^{\infty} f_n(x)$ converge *assolutamente* in x . Dal criterio di convergenza assoluta per le serie numeriche si ottiene subito che la convergenza assoluta di una serie di funzioni implica la sua convergenza (puntuale).

Se una serie di funzioni $\sum_{n=n_0}^{\infty} f_n(x)$ converge puntualmente in un insieme $A \subseteq X$, allora per ogni $x \in A$ la somma della serie è un numero reale che possiamo denotare con $f(x)$. Possiamo scrivere perciò

$$f(x) = \sum_{n=n_0}^{\infty} f_n(x), \quad x \in A.$$

Risulta così definita una funzione reale di variabile reale $f: A \rightarrow \mathbb{R}$. È naturale porsi la domanda se una tale funzione sia continua quando sono continue tutte le f_n . In altre parole: è ancora vero che la somma di funzioni continue è una funzione continua nel caso di infiniti addendi? La risposta è negativa. L'esempio seguente illustra questo fatto.

Esempio. Consideriamo, nell'intervallo $[0, 1]$, la serie di funzioni

$$\sum_{n=1}^{\infty} (x^n - x^{n+1})$$

che si può anche riscrivere nella forma

$$\sum_{n=1}^{\infty} (1-x)x^n.$$

Per $x \in [0, 1)$ la serie è geometrica di ragione x e primo termine $(1-x)x$; pertanto converge e la sua somma è data da

$$\frac{(1-x)x}{1-x} = x.$$

Per $x = 1$ tutte le funzioni $f_n(x) = (1-x)x^n$ sono nulle e, di conseguenza, la serie converge anche in tale punto ed ha somma zero. Si può concludere che la serie converge in tutto l'intervallo chiuso $[0, 1]$ e la sua somma è

$$f(x) = \begin{cases} x & \text{se } x \in [0, 1) \\ 0 & \text{se } x = 1 \end{cases}$$

che è una funzione discontinua, sebbene tutte le f_n siano continue (sono addirittura C^∞).

È utile perciò introdurre un'altro tipo di convergenza, la convergenza totale, che implica la convergenza assoluta e garantisce, tra l'altro, la continuità della funzione somma.

Definizione. Si dice che la serie di funzioni $\sum_{n=n_0}^{\infty} f_n(x)$ converge *totalmente* in un insieme $A \subseteq \mathbb{R}$ se esiste una serie numerica convergente $\sum_{n=n_0}^{\infty} c_n$ tale che $|f_n(x)| \leq c_n, \forall n \geq n_0$ e $\forall x \in A$.

Notiamo che se una serie di funzioni $\sum_{n=n_0}^{\infty} f_n(x)$ converge *totalmente* in un insieme A , allora (come conseguenza dei criteri di convergenza assoluta e del confronto) converge (assolutamente) per ogni $x \in A$. Risulta quindi ben definita la funzione somma

$$f(x) = \sum_{n=n_0}^{\infty} f_n(x).$$

Osserviamo inoltre che, supponendo ogni f_n limitata, la più piccola serie numerica che domina $\sum_{n=n_0}^{\infty} f_n(x)$ è quella il cui termine generale λ_n è dato da

$$\lambda_n = \sup\{|f_n(x)| : x \in A\}.$$

Pertanto, se tale serie numerica converge, allora la serie di funzioni converge *totalmente*. In caso contrario, ossia se $\sum_{n=n_0}^{\infty} \lambda_n = +\infty$, per il criterio del confronto

nessuna serie numerica che domina la serie di funzioni può convergere. Possiamo quindi enunciare il seguente

Teorema. (facoltativo) *Condizione necessaria e sufficiente affinché una serie di funzioni*

$$\sum_{n=n_0}^{\infty} f_n(x)$$

converga totalmente in un insieme A è che sia convergente la serie numerica

$$\sum_{n=n_0}^{\infty} \lambda_n,$$

dove $\lambda_n = \sup\{|f_n(x)| : x \in A\}$.

Esempio. Consideriamo la serie di funzioni

$$\sum_{n=2}^{\infty} n \log \left(1 + \frac{|x|^n}{n(n-1)^2} \right).$$

Osserviamo che il termine generale $f_n(x) = n \log \left(1 + \frac{|x|^n}{n(n-1)^2} \right)$ tende a zero se e solo se $|x| \leq 1$; perciò la serie può convergere in x se e solo se $x \in [-1, 1]$. Fissato $x \in [-1, 1]$, si ha

$$|f_n(x)| = n \log \left(1 + \frac{|x|^n}{n(n-1)^2} \right) \leq n \log \left(1 + \frac{1}{n(n-1)^2} \right) \leq \frac{1}{(n-1)^2},$$

da cui, essendo la serie $\sum_{n=2}^{\infty} \frac{1}{(n-1)^2}$ convergente, dal criterio del confronto si deduce che la serie data è assolutamente convergente in x . D'altra parte, posto $c_n = \frac{1}{(n-1)^2}$, dal fatto che $\sum_{n=2}^{\infty} c_n$ è una serie numerica maggiorante e convergente, si deduce anche che la serie data converge totalmente nell'intervallo $[-1, 1]$.

Consideriamo invece la serie

$$\sum_{n=2}^{\infty} n \log \left(1 + \frac{|x|^n}{n(n-1)} \right).$$

Anche in questo caso il termine generale tende a zero se e solo se $|x| \leq 1$. Però la serie converge puntualmente solo per $x \in (-1, 1)$, in quanto per $x = -1$ e $x = 1$ si ha $\sum_{n=2}^{\infty} n \log \left(1 + \frac{1}{n(n-1)} \right)$ che diverge avendo lo stesso carattere della serie armonica $\sum_{n=1}^{\infty} \frac{1}{n}$. Inoltre, essendo

$$\sup_{x \in (-1, 1)} n \log \left(1 + \frac{|x|^n}{n(n-1)} \right) = n \log \left(1 + \frac{1}{n(n-1)} \right),$$

la serie non converge totalmente in $(-1, 1)$ in quanto, come visto sopra, $\sum_{n=2}^{\infty} n \log \left(1 + \frac{1}{n(n-1)} \right)$ è divergente. Considerando invece un intervallo della forma $[-r, r]$ con $0 < r < 1$, si ha

$$\sup_{x \in [-r, r]} n \log \left(1 + \frac{|x|^n}{n(n-1)} \right) = n \log \left(1 + \frac{r^n}{n(n-1)} \right) = \lambda_n,$$

ed essendo $\sum_n \lambda_n$ convergente si può concludere che la serie data converge totalmente in $[-r, r]$.

Esercizio. Provare che la serie

$$\sum_{n=1}^{\infty} \frac{1}{n^x \log \left(1 + \frac{1}{n} \right)}$$

converge puntualmente se $x > 2$ e converge totalmente in $[r, +\infty)$, con $r > 2$.

Teorema. Sia $\sum_{n=n_0}^{\infty} f_n(x)$ una serie di funzioni continue convergente totalmente in un insieme A e sia

$$f(x) := \sum_{n=n_0}^{\infty} f_n(x)$$

la somma della serie. Allora

- i) (continuità) la funzione f risulta continua in A ;
- ii) (passaggio al limite sotto l'integrale) per ogni intervallo $[a, b] \subseteq A$ si ha $\sum_{n=n_0}^{\infty} \int_a^b f_n(x) dx = \int_a^b \sum_{n=n_0}^{\infty} f_n(x) dx$;
- iii) (derivabilità) se inoltre le funzioni f_n sono di classe C^1 in A e se la serie $\sum_{n=n_0}^{\infty} f'_n(x)$ delle derivate converge totalmente in A , allora f risulta di classe C^1 in A e si ha $f'(x) = \sum_{n=n_0}^{\infty} f'_n(x)$.

Serie di potenze

Un esempio importante di serie di funzioni è costituito dalle serie di potenze. Una serie di funzioni del tipo

$$\sum_{n=0}^{\infty} a_n(x - x_0)^n,$$

dove x_0 e gli a_n sono numeri reali assegnati, si dice una *serie di potenze* in campo reale. Il punto x_0 si chiama *centro della serie*.

Una serie di potenze converge ovviamente in $x = x_0$ (e la sua somma in tal caso vale a_0). Dal teorema che segue si deduce che l'insieme di convergenza di una serie di potenze è un intervallo (eventualmente ridotto al punto $x = x_0$).

Teorema. *Supponiamo che la serie di potenze $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ converga in un punto $\bar{x} \neq x_0$. Allora la serie converge assolutamente in ogni x tale che $|x - x_0| < |\bar{x} - x_0|$. Inoltre converge totalmente in ogni intervallo del tipo $[x_0 - r, x_0 + r]$ con $0 < r < |\bar{x} - x_0|$.*

Dimostrazione. Poiché la serie converge puntualmente in \bar{x} , il termine n -esimo tende a 0 per $n \rightarrow \infty$. Di conseguenza, essendo le successioni convergenti necessariamente limitate, esiste $M > 0$ tale che $|a_n(\bar{x} - x_0)^n| \leq M$. Pertanto, preso x tale che $|x - x_0| < |\bar{x} - x_0|$ risulta

$$|a_n(x - x_0)^n| = |a_n(\bar{x} - x_0)^n| \left(\frac{|x - x_0|}{|\bar{x} - x_0|} \right)^n \leq M \left(\frac{|x - x_0|}{|\bar{x} - x_0|} \right)^n.$$

Per come è stato scelto x , la serie geometrica $\sum_{n=0}^{\infty} \left(\frac{|x - x_0|}{|\bar{x} - x_0|} \right)^n$ ha ragione minore di 1 e quindi converge. Perciò, per il criterio del confronto, $\sum_{n=0}^{\infty} |a_n(x - x_0)|^n$ converge, cioè $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ converge assolutamente in x .

Sia ora r tale che $0 < r < |\bar{x} - x_0|$. Con la stessa dimostrazione fatta sopra si ottiene

$$|a_n(x - x_0)^n| \leq M \left(\frac{r}{|\bar{x} - x_0|} \right)^n.$$

Perciò $\sum_{n=0}^{\infty} a_n(x - x_0)^n$, essendo dominata dalla serie numerica convergente $\sum_{n=0}^{\infty} M \left(\frac{r}{|\bar{x} - x_0|} \right)^n$, risulta totalmente convergente in $[x_0 - r, x_0 + r]$. \square

Dal teorema precedente si ottiene il seguente

Teorema. *Sia data la serie di potenze*

$$\sum_{n=0}^{\infty} a_n(x - x_0)^n.$$

Allora, o la serie converge solo in $x = x_0$, oppure esiste $R > 0$ (eventualmente anche infinito) tale che la serie converge assolutamente se $|x - x_0| < R$ e non converge se $|x - x_0| > R$. Inoltre la serie converge totalmente in ogni intervallo $[x_0 - r, x_0 + r]$, con $0 < r < R$.

Dal risultato precedente si deduce perciò che l'insieme di convergenza di una serie di potenze centrata in x_0 è un intervallo (aperto, chiuso o semiaperto) e che x_0 è equidistante dagli estremi (finiti o infiniti che siano). Osserviamo che non si può dire nulla a priori del comportamento della serie nei punti $x_0 - R$ e $x_0 + R$. Ad esempio: la serie $\sum_n \frac{x^n}{n}$ ha raggio di convergenza 1 e converge (non assolutamente) in $x = -1$ mentre non converge in $x = 1$; la serie $\sum_n \frac{x^n}{n^2}$ ha raggio di convergenza 1 e converge assolutamente sia in $x = 1$ che in $x = -1$; la serie $\sum_n nx^n$ ha raggio di convergenza 1 e non converge negli estremi dell'intervallo $(-1, 1)$ (il termine generale non tende a zero).

Definizione. La semiampiezza R dell'intervallo di convergenza di una serie di potenze si chiama *raggio di convergenza* della serie.

Per calcolare il raggio di convergenza, fissato $x \in \mathbb{R}$, si può ricorrere agli usuali criteri per le serie a termini positivi applicandoli alla serie $\sum_{n=0}^{\infty} a_n |x - x_0|^n$, pensata come una serie numerica dipendente dal parametro x . Altrimenti, si può usare ad esempio il seguente

Teorema. Sia data una serie di potenze $\sum_{n=0}^{\infty} a_n (x - x_0)^n$ e supponiamo che esista $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}$. Allora si ha

$$R = \frac{1}{\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|}}$$

con la convenzione che $R = 0$ se $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = +\infty$ e $R = +\infty$ se $\lim_{n \rightarrow \infty} \sqrt[n]{|a_n|} = 0$

Osservazione. Un altro criterio per il calcolo del raggio di convergenza di una serie di potenze si ottiene ricordando che se esiste $\lim_{n \rightarrow +\infty} \frac{|a_{n+1}|}{|a_n|}$, allora si ha

$$\lim_{n \rightarrow +\infty} \sqrt[n]{|a_n|} = \lim_{n \rightarrow +\infty} \frac{|a_{n+1}|}{|a_n|}.$$

Esempio.

1. La serie

$$\sum_{n=0}^{\infty} \frac{x^n}{n}$$

ha raggio di convergenza $R = 1$.

2. La serie

$$\sum_{n=0}^{\infty} \frac{n^n}{n!} x^n$$

ha raggio di convergenza $R = 1/e$.

3. La serie

$$\sum_{n=0}^{\infty} \frac{x^n}{n!}$$

ha raggio di convergenza $R = +\infty$.

4. La serie

$$\sum_{n=0}^{\infty} n! x^n$$

ha raggio di convergenza $R = 0$.

Serie di Taylor

Sia data una serie di potenze $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ e supponiamo che abbia raggio di convergenza $R > 0$. Allora è ben definita, nell'intervallo $(x_0 - R, x_0 + R)$, la somma della serie, cioè una funzione f tale che

$$f(x) := \sum_{n=0}^{\infty} a_n(x - x_0)^n, \quad x \in (x_0 - R, x_0 + R).$$

Teorema. *La funzione f definita dalla serie di potenze $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ è continua in ogni punto dell'intervallo $(x_0 - R, x_0 + R)$.*

Dimostrazione. Sia \bar{x} un qualunque punto dell'intervallo $(x_0 - R, x_0 + R)$. Esiste un numero r , con $0 < r < R$, tale che l'intervallo $(x_0 - r, x_0 + r)$ contiene \bar{x} . Di conseguenza, poiché la serie è totalmente convergente in $[x_0 - r, x_0 + r]$, la funzione somma f è ivi continua. Questo implica, in particolare, che f è continua anche nel punto \bar{x} . \square

Vediamo altre proprietà delle serie di potenze.

Lemma (di invarianza del dominio di convergenza). *Supponiamo che la serie di potenze*

$$\sum_{n=0}^{\infty} a_n(x - x_0)^n$$

abbia raggio di convergenza $R > 0$. Allora le due serie

$$\sum_{n=1}^{\infty} na_n(x - x_0)^{n-1}$$

e

$$\sum_{n=0}^{\infty} \frac{a_n}{n+1} (x - x_0)^{n+1}$$

hanno raggio di convergenza R .

Il lemma precedente serve per provare che le serie di potenze sono “derivabili termine a termine” e “integrabili termine a termine”; si ha cioè il seguente

Teorema (di derivazione e integrazione delle serie di potenze). *Sia*

$$f(x) := \sum_{n=0}^{\infty} a_n(x - x_0)^n, \quad x \in (x_0 - R, x_0 + R)$$

una funzione definita da una serie di potenze. Allora f è derivabile in $(x_0 - R, x_0 + R)$ e si ha

$$f'(x) = \sum_{n=1}^{\infty} na_n(x - x_0)^{n-1}.$$

Inoltre, per ogni $x \in (x_0 - R, x_0 + R)$ si ha

$$\int_{x_0}^x f(t) dt = \sum_{n=0}^{\infty} \frac{a_n}{n+1} (x - x_0)^{n+1}.$$

Dimostrazione. (facoltativa) Fissato un qualunque punto x appartenente all’intervallo di convergenza $(x_0 - R, x_0 + R)$ della prima serie, esiste un intervallo $(x_0 - r, x_0 + r)$, con $0 < r < R$, contenente x . Per il lemma precedente, R è anche il raggio di convergenza della serie delle derivate, e quindi, in base al teorema della convergenza totale per le serie di potenze, entrambe le serie convergono totalmente in $[x_0 - r, x_0 + r]$. Dal teorema di derivabilità delle serie di funzioni segue che f è derivabile in x e risulta

$$f'(x) = \sum_{n=1}^{\infty} na_n(x - x_0)^{n-1}.$$

Inoltre, dal teorema di passaggio al limite sotto il segno di integrale, si ha

$$\int_{x_0}^x f(t) dt = \int_{x_0}^x \sum_{n=0}^{\infty} a_n(t - x_0)^n dt =$$

$$= \sum_{n=0}^{\infty} \int_{x_0}^x a_n (t - x_0)^n dt = \sum_{n=0}^{\infty} \frac{a_n}{n+1} (x - x_0)^{n+1}.$$

□

Osservazione. Poiché la derivata di una serie di potenze è ancora una serie di potenze, dal teorema precedente segue che le funzioni definite tramite serie di potenze sono di classe C^∞ .

Sia dunque

$$f(x) := a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + \cdots + a_n(x - x_0)^n + \cdots$$

una funzione definita mediante una serie di potenze nell'intervallo $(x_0 - R, x_0 + R)$. Ponendo $x = x_0$, si ottiene $a_0 = f(x_0)$. Derivando e ponendo di nuovo $x = x_0$, si ha $a_1 = f'(x_0)$. Analogamente, mediante derivate successive (vedi l'osservazione precedente), si ottiene

$$a_n = \frac{f^{(n)}(x_0)}{n!}.$$

Pertanto, risulta necessariamente

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

dove $f^{(0)}(x_0)$ denota $f(x_0)$.

In altre parole, la serie di potenze $\sum_{n=0}^{\infty} a_n(x - x_0)^n$ coincide in $(x_0 - R, x_0 + R)$ con la serie di potenze

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

che è detta *serie di Taylor di f di centro x_0* (o di *MacLaurin*, quando $x_0 = 0$).

Il ragionamento precedente prova cioè che se una funzione è definita mediante una serie di potenze, essa risulta di classe C^∞ in x_0 e la sua serie di Taylor ha per somma la funzione stessa.

D'altra parte, data una funzione di classe C^∞ in x_0 , si può considerare la serie di Taylor di f di centro x_0 , cioè

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

Ci chiediamo:

- 1) se esista un intorno di x_0 nel quale questa serie sia convergente;
- 2) supposto che esista un tale intorno, se in esso la somma della serie sia proprio $f(x)$.

Definizione. Una funzione f si dice *svilupabile in serie di Taylor in un intorno di un punto x_0* (o *analitica in x_0*) se esiste un intorno di x_0 in cui vale l'uguaglianza

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

Si dice che f è *analitica* se ogni punto del suo dominio ammette un intorno in cui f è svilupabile in serie di Taylor (ossia, se è analitica in ogni punto del suo dominio).

Osservazione. Esistono funzioni di classe C^∞ ma non analitiche. Una di queste è

$$f(x) = \begin{cases} e^{-1/x^2} & \text{se } x \neq 0 \\ 0 & \text{se } x = 0, \end{cases}$$

le cui derivate successive (come si potrebbe provare usando un corollario del teorema di Lagrange) risultano tutte continue e nulle nel punto $x_0 = 0$. Quindi, se f fosse analitica, in un intorno del punto $x_0 = 0$ dovrebbe valere l'uguaglianza

$$f(x) = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n,$$

e ciò è impossibile perché $f(x) \neq 0$ per $x \neq 0$, mentre la sua serie di Taylor $\sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n$ ha per somma zero (essendo nulli tutti i suoi termini).

Si ha la seguente condizione necessaria e sufficiente affinché una funzione sia svilupabile in serie di Taylor.

Teorema. Una funzione f di classe C^∞ in x_0 è svilupabile in serie di Taylor in un intorno $(x_0 - R, x_0 + R)$ di x_0 se e solo se, per ogni $x \in (x_0 - R, x_0 + R)$, risulta $\lim_{n \rightarrow \infty} R_n(x - x_0) = 0$, dove

$$R_n(x - x_0) = f(x) - \sum_{k=0}^n \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

denota il resto n -esimo della formula di Taylor.

Esempio. La funzione $f(x) = e^x$ è svilupabile in serie di MacLaurin e tale serie ha raggio di convergenza $R = +\infty$. Fissiamo un punto $x \in \mathbb{R}$. Sappiamo che se e^x

è effettivamente sviluppabile in serie di MacLaurin, allora si deve necessariamente avere

$$e^x = \sum_{n=0}^{\infty} \frac{f^{(n)}(0)}{n!} x^n = \sum_{n=0}^{\infty} \frac{x^n}{n!}.$$

Per il teorema precedente, ciò equivale ad affermare che

$$\lim_{n \rightarrow \infty} R_n(x) = \lim_{n \rightarrow \infty} (e^x - P_n(x)) = 0,$$

dove la somma parziale n -esima

$$P_n(x) = \sum_{k=0}^n \frac{x^k}{k!}$$

non è altro che il polinomio di MacLaurin di e^x di ordine n . Scrivendo il resto $R_n(x)$ nella forma di Lagrange, si potrebbe facilmente far vedere che in effetti esso tende a 0 per $n \rightarrow +\infty$. Per l'arbitrarietà del punto x , possiamo concludere che lo sviluppo è valido in tutto \mathbb{R} .

Ponendo $x = 1$ nello sviluppo in serie di MacLaurin di e^x si ottiene il numero e espresso mediante una serie numerica:

$$e = \sum_{n=0}^{\infty} \frac{1}{n!}.$$

Mostriamo ora che la funzione e^x è analitica, cioè che è sviluppabile in serie di Taylor in ogni punto di \mathbb{R} (facoltativo). A tale scopo fissiamo $x_0 \in \mathbb{R}$ e poniamo, per comodità, $x = x_0 + h$. Si ha

$$\begin{aligned} e^x &= e^{x_0+h} = e^{x_0} e^h = e^{x_0} \left(1 + h + \frac{h^2}{2!} + \dots + \frac{h^n}{n!} + \dots \right) = \\ &= e^{x_0} + e^{x_0} h + e^{x_0} \frac{h^2}{2!} + \dots + e^{x_0} \frac{h^n}{n!} + \dots \end{aligned}$$

Quindi

$$e^x = e^{x_0} + e^{x_0}(x - x_0) + e^{x_0} \frac{(x - x_0)^2}{2!} + \dots + e^{x_0} \frac{(x - x_0)^n}{n!} + \dots$$

Esercizio. In maniera analoga a quanto fatto per la funzione esponenziale, si può provare che le funzioni $\cos x$ e $\sin x$ sono sviluppabili in serie di MacLaurin. Determinarne lo sviluppo e l'intorno di validità.

Il metodo usato per sviluppare in serie di MacLaurin le funzioni e^x , $\sin x$ e $\cos x$ (basato su una stima del resto della formula di Taylor) non è adatto per la funzione $f(x) = \log(1+x)$. In questo caso conviene procedere diversamente:

1. si determina prima lo sviluppo della derivata $f'(x)$ di $f(x)$;
2. successivamente, mediante il teorema di integrazione termine a termine delle serie di potenze, si trova una primitiva dello sviluppo di $f'(x)$;
3. infine, tra tutte le primitive di $f'(x)$ espresse in serie di potenze, si sceglie quella che coincide con $f(x)$.

Tale metodo è adatto anche per determinare lo sviluppo di $\arctang x$ e, in generale, di tutte le funzioni di cui è facile sviluppare la derivata. A tale proposito ricordiamo che due primitive di una stessa funzione (definita in un intervallo) differiscono per una costante e, di conseguenza, se coincidono in un punto, coincidono in tutto l'intervallo di definizione.

Cominciamo col determinare, col metodo appena esposto, lo sviluppo di MacLaurin di $\log(1+x)$. La derivata $(1+x)^{-1}$ di $\log(1+x)$ rappresenta, per $x \in (-1, 1)$, la somma di una serie geometrica di ragione $-x$ e primo termine 1. Quindi, per $x \in (-1, 1)$, si ha

$$(1+x)^{-1} = 1 - x + x^2 - x^3 + \dots + (-1)^n x^n + \dots$$

Dal teorema di derivazione delle serie di potenze si deduce che

$$g(x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots + (-1)^n \frac{x^{n+1}}{n+1} + \dots$$

è una primitiva di $(1+x)^{-1}$; ma, è bene precisare, soltanto per x appartenente al comune dominio di convergenza $(-1, 1)$ delle due serie. Dunque, $\log(1+x)$ e $g(x)$ hanno la stessa derivata per $x \in (-1, 1)$. Poiché coincidono per $x = 0$, si può concludere che

$$\log(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \dots + (-1)^n \frac{x^{n+1}}{n+1} + \dots \quad \forall x \in (-1, 1).$$

Esercizio. Provare che la funzione $\arctang x$ è sviluppabile in serie di MacLaurin, determinarne lo sviluppo e l'intervallo di validità.

Suggerimento. Sviluppare prima la derivata di $\arctang x$.

Esercizio. Provare che la funzione e^{-x^2} è sviluppabile in serie di MacLaurin, determinarne lo sviluppo e l'intervallo di validità.

Suggerimento. Si ricorda che l'uguaglianza

$$e^x = 1 + x + \frac{x^2}{2!} + \dots + \frac{x^n}{n!} + \dots$$

è valida per ogni numero reale x , e quindi, in particolare, è valida per ogni numero reale $-x^2$.

Esercizio. Provare che la funzione degli errori,

$$\operatorname{erf} x = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt,$$

è sviluppabile in serie di MacLaurin, determinarne lo sviluppo e l'intervallo di validità.

Suggerimento. Sviluppare prima la derivata di $\operatorname{erf} x$.

Esercizio. Provare che la funzione

$$f(x) = \begin{cases} \frac{\operatorname{sen} x}{x} & \text{se } x \neq 0 \\ 1 & \text{se } x = 0 \end{cases}$$

è sviluppabile in serie di MacLaurin, determinarne lo sviluppo e l'intervallo di validità.

Suggerimento. Sviluppare prima $\operatorname{sen} x$.

Dato $\alpha \in \mathbb{R}$, consideriamo la funzione $f(x) = (1+x)^\alpha$ che è senz'altro definita e C^∞ nell'intervallo $(-1, +\infty)$. Si ha

$$f^{(n)}(0) = \alpha(\alpha-1) \cdots (\alpha-n+1),$$

per cui la serie di MacLaurin di f è

$$\sum_{n=0}^{\infty} \binom{\alpha}{n} x^n,$$

ove ricordiamo che $\binom{\alpha}{n}$ è il *coefficiente binomiale*

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1) \cdots (\alpha-n+1)}{n!}, \quad \binom{\alpha}{0} = 1.$$

Si può provare che per $|x| < 1$ e qualunque sia α si ha

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n.$$

Questa serie di potenze è detta *serie binomiale*. Se α è un numero naturale la serie è una somma finita. Infatti si riduce alla somma dei primi $\alpha+1$ termini essendo i coefficienti binomiali tutti nulli per $n > \alpha$. In questo caso prende il nome di *binomio di Newton*.

Esercizio. Scrivere i primi quattro termini della serie di MacLaurin di $f(x) = \sqrt{1+x}$ ($\alpha = 1/2$) e di $f(x) = \frac{1}{\sqrt{1+x}}$ ($\alpha = -(1/2)$).

4^a settimana - dal 18.03.19

Seconda parte
(Argomenti di Analisi Matematica 2)

Funzioni di più variabili

Lo spazio \mathbb{R}^2 è l'insieme delle coppie ordinate di numeri reali; ossia delle coppie di numeri (x, y) , con $x, y \in \mathbb{R}$. Una coppia (x, y) si dice “ordinata” perché dei due numeri x e y è importante conoscere quale sia il primo e quale il secondo (ovvero, l'ordine in cui si susseguono). In altre parole, una coppia (x, y) , quando $x \neq y$, si considera diversa da (y, x) . Analogamente, lo spazio \mathbb{R}^3 è l'insieme delle terne ordinate di numeri reali. Più in generale, dato $k \in \mathbb{N}$, con \mathbb{R}^k si denota l'insieme delle k -uple (si legge “cappauple”) di numeri reali.

Gli elementi di \mathbb{R}^2 [di \mathbb{R}^3 , di \mathbb{R}^k] si dicono punti (o vettori). Dato un punto $P = (x, y) \in \mathbb{R}^2$, il numero x si dice la prima *coordinata* (o *componente*) di P e il numero y la seconda. Analogamente, dato $P = (x_1, x_2, \dots, x_k) \in \mathbb{R}^k$, i numeri x_1, x_2, \dots, x_k sono le coordinate di P (la prima, la seconda, ... la k -esima). Il punto di \mathbb{R}^k con componenti tutte nulle si chiama *origine* di \mathbb{R}^k (o *vettore nullo*, o *vettore banale*) e si indica con 0 (come lo 0 dei reali).

Dati due punti $P = (x, y)$ e $Q = (\bar{x}, \bar{y})$ di \mathbb{R}^2 , la loro somma si definisce “componente per componente”:

$$P + Q = (x + \bar{x}, y + \bar{y}).$$

Dato un punto $P = (x, y) \in \mathbb{R}^2$ e dato un numero $\lambda \in \mathbb{R}$ (detto scalare, per distinguerlo dal vettore P) si definisce il prodotto λP moltiplicando per λ ogni componente di P . Ossia, $\lambda P = (\lambda x, \lambda y)$. Ovviamente anche la differenza tra due punti di \mathbb{R}^2 , che è definita come l'operazione inversa della somma, si fa componente per componente. Analoghe definizioni si danno in \mathbb{R}^3 e in \mathbb{R}^k (i dettagli sono lasciati allo studente).

Se $P = (x, y)$ è un punto di \mathbb{R}^2 , la sua *norma* (o *modulo*) è il numero

$$\|P\| = \sqrt{x^2 + y^2},$$

che rappresenta la distanza di P dall'origine di \mathbb{R}^2 . Più in generale, dato $P = (x_1, x_2, \dots, x_k) \in \mathbb{R}^k$, la sua norma è il numero

$$\|P\| = \sqrt{x_1^2 + x_2^2 + \dots + x_k^2} = \sqrt{\sum_{i=1}^k x_i^2}.$$

Si potrebbe provare che la norma gode di proprietà simili a quelle del valore assoluto.

La *distanza* tra due punti P e Q di \mathbb{R}^k è, per definizione, il numero

$$d(P, Q) = \|P - Q\|.$$

In particolare, se $P = (x, y)$ e $Q = (\bar{x}, \bar{y})$ sono due punti di \mathbb{R}^2 , si ha

$$d(P, Q) = \|P - Q\| = \sqrt{(x - \bar{x})^2 + (y - \bar{y})^2}.$$

Nello spazio \mathbb{R}^k si definisce il prodotto tra due arbitrari vettori $u = (u_1, u_2, \dots, u_k)$ e $v = (v_1, v_2, \dots, v_k)$, detto *prodotto scalare* e denotato col simbolo $u \cdot v$ (si legge “ u scalare v ”), ponendo

$$u \cdot v = u_1 v_1 + u_2 v_2 + \dots + u_k v_k.$$

Il significato geometrico è il seguente: il prodotto scalare tra due vettori u e v è dato dal prodotto dei moduli (dei due vettori) per il coseno dell’angolo θ compreso (tra i due vettori). Si ha pertanto

$$|u \cdot v| = \|u\| \|v\| |\cos \theta| \leq \|u\| \|v\|,$$

da cui si deduce la nota *disuguaglianza di Schwarz*:

$$|u \cdot v| \leq \|u\| \|v\|.$$

Si osservi che la norma di un vettore $v \in \mathbb{R}^k$ può essere definita anche attraverso il prodotto scalare. Si ha infatti $\|v\| = \sqrt{v \cdot v}$.

Dato un punto $P_0 \in \mathbb{R}^k$ e dato $r > 0$, l’*intorno (sferico)* di centro P_0 e raggio r è l’insieme

$$B_r(P_0) = \{P \in \mathbb{R}^k : \|P - P_0\| < r\}$$

dei punti P di \mathbb{R}^k che distano da P_0 meno di r . In \mathbb{R}^2 , l’intorno sferico di un punto P_0 si dice anche *intorno circolare*, ed è costituito da un cerchio di centro P_0 privato della circonferenza (la frontiera del cerchio). In \mathbb{R}^3 , l’intorno $B_r(P_0)$ è una palla di centro P_0 privata della superficie sferica, e in \mathbb{R} è l’intervallo aperto (di ampiezza $2r$) $(P_0 - r, P_0 + r)$.

L’insieme ottenuto da $B_r(P_0)$ togliendo il punto P_0 , è detto *intorno forato* (di centro P_0 e raggio r).

Analogamente a quanto si è visto per lo spazio \mathbb{R} , dato $A \subseteq \mathbb{R}^k$ e dato $P_0 \in \mathbb{R}^k$, si dice che P_0 è un punto di *accumulazione* per A se **ogni** suo intorno forato

contiene punti di A (o, in formule, se per ogni $r > 0$ si ha $A \cap B_r(P_0) \setminus \{P_0\} \neq \emptyset$). In maniera equivalente, P_0 è un punto di accumulazione per A se ogni suo intorno contiene infiniti punti di A .

Un punto P_0 di A si dice *isolato* se non è di accumulazione per A (cioè se **esiste** un intorno forato di P_0 la cui intersezione con A è l'insieme vuoto o, in simboli, se esiste $r > 0$ tale che $A \cap B_r(P_0) \setminus \{P_0\} = \emptyset$). In maniera equivalente, un punto P_0 di A è isolato se esiste un intorno di P_0 che non contiene punti di A diversi da P_0 stesso (in simboli, se esiste $r > 0$ tale che $A \cap B_r(P_0) = \{P_0\}$).

Esempio. Il punto $(0, 0)$ è di accumulazione per l'insieme

$$A = \{(x, y) \in \mathbb{R}^2 : xy > 0\}.$$

Infatti, per ogni $r > 0$, i punti del tipo (x, x) con $0 < |x| < r/\sqrt{2}$ appartengono a $B_r((0, 0)) \cap A$.

Esempio. L'insieme

$$A = \{(x, y) \in \mathbb{R}^2 : (x, y) = \left(\frac{1}{n}, \frac{1}{n}\right), n \in \mathbb{N}\}$$

è costituito tutto da punti isolati. Infatti, sia $P_0 = \left(\frac{1}{n_0}, \frac{1}{n_0}\right)$ un qualunque punto di A . Denotato con P_1 il punto $\left(\frac{1}{n_0+1}, \frac{1}{n_0+1}\right)$, è facile verificare che si ha $B_{r_0}(P_0) \cap A = \{P_0\}$ pur di scegliere $r_0 < \|P_1 - P_0\|$.

D'altra parte, il punto $(0, 0)$ non appartiene ad A ma è di accumulazione per A . Infatti fissato $r > 0$, il punto $P = \left(\frac{1}{n}, \frac{1}{n}\right) \in B_r((0, 0))$ non appena $n > \sqrt{2}/r$.

Diremo che P_0 è *interno* ad A se **esiste** un intorno di P_0 contenuto in A , cioè se esiste $r > 0$ tale che $B_r(P_0) \subseteq A$.

Osservazione. Un punto interno ad un insieme deve necessariamente appartenere a tale insieme, dato che l'intorno di un punto contiene il punto stesso. Non è vero però il contrario. Ad esempio, se

$$A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$$

il punto $(1, 0)$, pur appartenendo ad A , non è interno perché ogni suo intorno contiene degli elementi che non stanno in A (osserviamo che negare che “esiste un intorno interamente contenuto in A ” equivale ad affermare che “ogni intorno non è interamente contenuto in A ”, e quindi contiene punti del complementare).

Un sottoinsieme A di \mathbb{R}^k si dice *aperto* se ogni suo punto è interno; si dice *chiuso* se il suo complementare è aperto. Si potrebbe provare che un insieme è chiuso se e solo se contiene tutti i suoi punti di accumulazione.

Osservazione. Gli intervalli aperti, (a, b) , $(a, +\infty)$, $(-\infty, b)$ di \mathbb{R} sono insiemi aperti in \mathbb{R} (ma non in \mathbb{R}^k , $k \geq 2$), mentre gli intervalli chiusi $[a, b]$, $[a, +\infty)$, $(-\infty, b]$ sono insiemi chiusi.

Si dice che $P_0 \in \mathbb{R}^k$ è di *frontiera* per A se ogni suo intorno contiene sia punti di A sia punti del complementare di A . L'insieme dei punti di frontiera di A è detto la *frontiera* di A e si denota con ∂A . Discende dalla definizione che un insieme e il suo complementare hanno la stessa frontiera.

Dato $A \subseteq \mathbb{R}^k$ la *chiusura* di A (denotata \bar{A}) è l'insieme $\bar{A} = A \cup \partial A$. L'insieme \bar{A} è un chiuso e, anzi, è il più piccolo chiuso contenente A . Si può provare che un insieme è chiuso se e solo se coincide con la sua chiusura. Inoltre, si può far vedere che la chiusura di un insieme A si ottiene aggiungendo ad A tutti i suoi punti di accumulazione. Si ha anche $\partial A = \bar{A} \cap \overline{A^c}$, dove A^c denota il complementare di A in \mathbb{R}^k .

Ovviamente, esistono anche insiemi che non sono né aperti né chiusi; basti pensare ad un insieme che contiene soltanto alcuni punti della sua frontiera, ma non tutti. Ad esempio, ogni intervallo di \mathbb{R} della forma $(a, b]$. Altri esempi sono i tre che seguono.

Esempio. L'insieme

$$A = \{(x, y) \in \mathbb{R}^2 : 1 \leq x^2 + y^2 < 4\}$$

non è né aperto (i punti (x, y) tali che $x^2 + y^2 = 1$ appartengono ad A ma non sono interni) né chiuso (infatti i punti (x, y) tali che $x^2 + y^2 = 4$ non appartengono ad A ma sono di accumulazione per A). Si ha

$$\partial A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \cup \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 4\}$$

e

$$\bar{A} = \{(x, y) \in \mathbb{R}^2 : 1 \leq x^2 + y^2 \leq 4\}.$$

Esempio. L'insieme

$$A = \{(x, y) \in \mathbb{R}^2 : 0 < x^2 + y^2 \leq 1\}$$

non è né aperto (i punti (x, y) tali che $x^2 + y^2 = 1$ appartengono ad A ma non sono interni) né chiuso (infatti $(0, 0) \notin A$ ma è di accumulazione per A). Si ha

$$\partial A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\} \cup \{(0, 0)\}$$

e

$$\bar{A} = A \cup \{(0, 0)\} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}.$$

Esempio. Abbiamo provato in precedenza che l'insieme

$$A = \{(x, y) \in \mathbb{R}^2 : (x, y) = \left(\frac{1}{n}, \frac{1}{n}\right), n \in \mathbb{N}\}$$

è costituito tutto da punti isolati e che $(0, 0)$ è di accumulazione per A . Si ha

$$\partial A = \bar{A} = A \cup \{(0, 0)\}.$$

Esempio. L'insieme

$$A = \{(x, y) \in \mathbb{R}^2 : xy > 0\}$$

è un insieme aperto. Si ha

$$\bar{A} = \{(x, y) \in \mathbb{R}^2 : xy \geq 0\}, \quad \partial A = \{(x, y) \in \mathbb{R}^2 : x = 0 \vee y = 0\}.$$

Un sottoinsieme A di \mathbb{R}^k si dice *limitato* se esiste $r > 0$ tale che $A \subseteq \overline{B_r(0)} = \{P \in \mathbb{R}^k : \|P\| \leq r\}$, cioè se esiste $r > 0$ tale che $\|P\| \leq r$ per ogni $P \in A$.

Esempio. L'insieme

$$A = \{(x, y) \in \mathbb{R}^2 : (x - 2)^2 + (y - 1)^2 < 1\}$$

è un insieme limitato. Esso infatti è l'intorno di centro $(2, 1)$ e raggio 1 ed è contenuto ad esempio in $\overline{B_4(0)} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 < 16\}$. Osserviamo inoltre che A è aperto e si ha

$$\partial A = \{(x, y) \in \mathbb{R}^2 : (x - 2)^2 + (y - 1)^2 = 1\}$$

e

$$\bar{A} = \{(x, y) \in \mathbb{R}^2 : (x - 2)^2 + (y - 1)^2 \leq 1\}.$$

Esempio. L'insieme

$$A = \{(x, y) \in \mathbb{R}^2 : y \geq x^2\}$$

è un insieme non limitato. Infatti, preso un qualunque $r > 0$, il punto $(0, 2r)$ appartiene ad A ma non a $\overline{B_r(0)}$. Pertanto, non esiste $r > 0$ tale che $A \subseteq \overline{B_r(0)}$. Osserviamo inoltre che A è chiuso e

$$\partial A = \{(x, y) \in \mathbb{R}^2 : y = x^2\}.$$

Ovviamente in questo caso risulta $A = \overline{A}$.

Una *successione* $\{P_n\}_{n \in \mathbb{N}}$ in \mathbb{R}^k è un'applicazione che associa ad ogni $n \in \mathbb{N}$ il punto $P_n \in \mathbb{R}^k$.

Si dice che una successione $\{P_n\}$ di punti di \mathbb{R}^k *tende* (o *converge*) ad un punto $P_0 \in \mathbb{R}^k$, e si scrive $P_n \rightarrow P_0$ (per $n \rightarrow +\infty$), se e solo se la distanza tra P_n e P_0 tende a zero, cioè se e solo se la successione reale $\{\|P_n - P_0\|\}$ è infinitesima. In formule avremo che $P_n \rightarrow P_0$ (per $n \rightarrow +\infty$) se per ogni $\epsilon > 0$ esiste un indice N (dipendente da ϵ) tale che per $n > N$ si ha $\|P_n - P_0\| < \epsilon$.

Se una successione $\{P_n\}$ converge ad un punto P_0 , si dice anche che P_0 è il *limite* di $\{P_n\}$ e si scrive

$$\lim_{n \rightarrow \infty} P_n = P_0 \quad \text{oppure} \quad \lim P_n = P_0.$$

La seconda notazione, particolarmente sintetica, è giustificata dal fatto che nei reali estesi $+\infty$ è l'unico punto di accumulazione per \mathbb{N} e quindi l'unica nozione di limite che ha senso per le successioni è per $n \rightarrow +\infty$. In generale, se in un insieme X è definita una distanza, si dice che una successione $\{P_n\}$ di punti di X converge ad un punto $P_0 \in X$ se la distanza tra P_n e P_0 tende a zero. In tal modo il concetto di successione convergente in un insieme dotato di distanza (si chiama *spazio metrico*) è ricondotto alla classica nozione di successione convergente (a zero) di numeri reali (la distanza tra due punti è infatti un numero reale). Ad esempio, nell'insieme dei numeri complessi \mathbb{C} è definita una distanza (la distanza tra due punti è il modulo della differenza) e quindi una successione $\{z_n\}$ di numeri complessi converge a z_0 se la successione di numeri reali $\{\|z_n - z_0\|\}$ tende a zero.

Osservazione. Una successione $\{P_n\} = \{(x_n, y_n)\}$ di punti di \mathbb{R}^2 converge a $P_0 = (x_0, y_0)$ se e solo se $x_n \rightarrow x_0$ e $y_n \rightarrow y_0$. Più in generale, una successione $\{P_n\}$ in \mathbb{R}^k converge a P_0 se e solo se ogni componente di $\{P_n\}$ converge alla corrispondente componente di P_0 . In \mathbb{C} una successione $\{z_n\} = \{\alpha_n + i\beta_n\}$ tende a $z_0 = \alpha_0 + i\beta_0$ se e solo se $\alpha_n \rightarrow \alpha_0$ e $\beta_n \rightarrow \beta_0$.

Esempio. La successione $\{P_n\} = \{(\frac{n+1}{n}, \frac{2}{n})\}$ di punti di \mathbb{R}^2 converge a $P_0 = (1, 0)$. Infatti $x_n = \frac{n+1}{n} \rightarrow 1$ e $y_n = \frac{2}{n} \rightarrow 0$.

Si dice che una successione $\{P_n\}$ di punti di \mathbb{R}^k *tende* a ∞ , e si scrive $P_n \rightarrow \infty$ (per $n \rightarrow +\infty$), se e solo se la norma di P_n tende a $+\infty$, cioè se e solo se la successione reale $\{\|P_n\|\}$ è divergente a $+\infty$. In formule avremo che $P_n \rightarrow \infty$ (per $n \rightarrow +\infty$) se per ogni $M > 0$ esiste un $N \in \mathbb{N}$ (dipendente da M) tale che per $n > N$ si ha $\|P_n\| > M$.

Esempio. La successione $\{P_n\} = \{(n^2, \frac{n+1}{n})\}$ di punti di \mathbb{R}^2 tende a ∞ . Infatti $\|P_n\| = \sqrt{n^4 + (\frac{n+1}{n})^2} \rightarrow +\infty$.

Funzioni reali di più variabili

Una funzione $f: A \rightarrow \mathbb{R}$ si dice reale di due [di tre, di k] variabili reali se il suo dominio A è un sottoinsieme di \mathbb{R}^2 [di \mathbb{R}^3 , di \mathbb{R}^k]. Se f è una funzione di due [tre, k] variabili, il valore che assume in un punto (x, y) [(x, y, z) , (x_1, x_2, \dots, x_k)] si denota con $f(x, y)$ [$f(x, y, z)$, $f(x_1, x_2, \dots, x_k)$] o, più semplicemente, con $f(P)$, dove P sta per (x, y) [per (x, y, z) , per (x_1, x_2, \dots, x_k)]. Il grafico di una funzione di due variabili è il sottoinsieme di \mathbb{R}^3 dato da $\{(x, y, z) \in A \times \mathbb{R} : z = f(x, y)\}$. In maniera analoga il grafico di una funzione di tre [k] variabili è un sottoinsieme di \mathbb{R}^4 [\mathbb{R}^{k+1}].

In alcuni casi le funzioni reali si possono comporre. Ad esempio si può comporre la funzione reale di due variabili reali $f(x, y) = 9 - x^2 - y^2$ con la funzione reale di variabile reale $g(z) = \sqrt{z}$, ottenendo così la funzione

$$(g \circ f)(x, y) = \sqrt{9 - x^2 - y^2}.$$

Il dominio, in questo caso, è l'insieme dei punti (x, y) per cui ha senso scrivere

$$\sqrt{9 - x^2 - y^2},$$

cioè il cerchio di raggio $r = 3$ con il centro situato nell'origine di \mathbb{R}^2 . In generale il dominio della composizione di una funzione reale di due variabili reali $f(x, y)$ con una funzione reale di variabile reale $g(x)$ è l'insieme dei punti $(x, y) \in \text{dom } f$ tali che $f(x, y) \in \text{dom } g$.

Si osservi che se si compone una successione in \mathbb{R}^2 (cioè una funzione da \mathbb{N} in \mathbb{R}^2) con una funzione $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, si ottiene una funzione da \mathbb{N} in \mathbb{R} , ossia una successione di numeri reali. Ad esempio la composizione della successione $\{(x_n, y_n)\} = \{(1/n^2, -2/n)\}$ con la funzione $f(x, y) = 3x + y^2$ dà la successione reale $\{f(x_n, y_n)\} = \{7/n^2\}$. In generale una successione $\{P_n\}$ in un insieme (arbitrario) A si può comporre con una funzione $f: A \rightarrow \mathbb{R}$, ottenendo così una successione di numeri reali.

Esempio. Il dominio della funzione di due variabili

$$f(x, y) = \log(\sqrt{2x^2 + y^2} - 3 - 1)$$

è l'insieme

$$\{(x, y) \in \mathbb{R}^2 : \frac{x^2}{2} + \frac{y^2}{4} > 1\},$$

cioè è la parte di piano esterna all'ellisse di equazione $x^2/2 + y^2/4 = 1$. Perciò tale dominio è un insieme aperto e non limitato. La frontiera è l'insieme $\{(x, y) : x^2/2 + y^2/4 = 1\}$, mentre la chiusura è l'insieme $\{(x, y) : x^2/2 + y^2/4 \geq 1\}$.

Esempio. Il dominio della funzione di due variabili

$$f(x, y) = \arcsen \frac{x - y}{x + y}$$

è l'insieme $\{(x, y) \in \mathbb{R}^2 : xy \geq 0\} \setminus \{(0, 0)\}$ che non né aperto né chiuso. La sua chiusura è l'insieme $\{(x, y) \in \mathbb{R}^2 : xy \geq 0\}$ e la frontiera è costituita dai due assi $\{(x, 0) : x \in \mathbb{R}\} \cup \{0, y\} : y \in \mathbb{R}\}$.

Esercizio. Determinare e disegnare il dominio delle seguenti funzioni di due variabili:

$$f(x, y) = \arcsen(x^2 + y), \quad f(x, y) = \frac{1}{\sqrt{\sen x \cos y}}.$$

Stabilire se tale dominio è un insieme aperto, chiuso, limitato. Determinarne la frontiera e la chiusura.

Esercizio. Determinare e disegnare il dominio delle funzioni di tre variabili:

$$f(x, y, z) = \log(x^2 + 4y^2 + z^2), \quad f(x, y, z) = \sqrt{1 - (x - 3)^2 - y^2 - z^2}.$$

Talvolta, invece di rappresentare il grafico (tridimensionale) di una funzione di due variabili si preferisce tracciare nel piano gli insiemi

$$S_c = \{(x, y) \in A : f(x, y) = c\},$$

dove c è una costante reale. Tali insiemi sono detti *insiemi (o linee) di livello* di f .

Esempio. Determiniamo le linee di livello della funzione di due variabili: $f(x, y) = x^2 + y^2 - 4x$. Fissato $c \in \mathbb{R}$, si ha $x^2 + y^2 - 4x = c$ o, equivalentemente, $(x - 2)^2 + y^2 = c + 4$. Perciò, se $c > -4$ l'insieme di livello S_c è la circonferenza di centro $(2, 0)$ e raggio $\sqrt{c + 4}$, se $c = -4$ si ha $S_c = \{(2, 0)\}$, mentre se $c < -4$ si ha $S_c = \emptyset$.

Esercizio. Determinare e disegnare le linee di livello delle funzioni di due variabili:

$$f(x, y) = |x| + |y|, \quad f(x, y) = \log(x + y).$$

Sia $f: A \rightarrow \mathbb{R}$ una funzione reale di k variabili reali e sia P_0 un punto di accumulazione per il dominio A di f . Si dice che $f(P)$ tende ad un numero reale l per P che tende ad P_0 se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che da

$$0 < \|P - P_0\| < \delta$$

e $P \in A$ segue $|f(P) - l| < \varepsilon$. In tal caso, si scrive

$$\lim_{P \rightarrow P_0} f(P) = l.$$

Se ad esempio f è di due variabili, posto $P_0 = (x_0, y_0)$ e $P = (x, y)$ la definizione precedente è equivalente alla seguente. Si dice che $f(x, y)$ tende ad un numero reale l per (x, y) che tende ad (x_0, y_0) se per ogni $\varepsilon > 0$ esiste $\delta > 0$ (dipendente da ε) tale che da

$$0 < \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta$$

e $(x, y) \in A$ segue $|f(x, y) - l| < \varepsilon$. In tal caso si scrive

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = l.$$

Esempio. Calcoliamo

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\text{sen}(4xy)}{y}.$$

Osserviamo intanto che, ponendo $x = 0$, cioè restringendo la funzione $f(x, y) = \text{sen}(4xy)/y$ all'asse y , si ottiene $f(0, y) = 0$ e quindi il limite di tale restrizione è 0. Pertanto il limite della funzione data, se esiste, vale 0. Fissato $\varepsilon > 0$, si ha

$$\left| \frac{\text{sen}(4xy)}{y} \right| \leq \frac{4|xy|}{|y|} \leq 4|x| < 4\sqrt{x^2 + y^2} < 4\delta,$$

da cui

$$\left| \frac{\text{sen}(4xy)}{y} \right| < \varepsilon,$$

pur di prendere $\delta \leq \varepsilon/4$.

5^a settimana - dal 25.03.19

Si dice che $f(x, y)$ tende a $+\infty$ [a $-\infty$] per (x, y) che tende ad (x_0, y_0) , e si scrive $f(x, y) \rightarrow +\infty$ [$f(x, y) \rightarrow -\infty$] per $(x, y) \rightarrow (x_0, y_0)$, se per ogni $M > 0$ esiste $\delta > 0$ (dipendente da M) tale che da

$$0 < \sqrt{(x - x_0)^2 + (y - y_0)^2} < \delta$$

e $(x, y) \in A$ segue $f(x, y) > M$ [$f(x, y) < -M$].

Analoghe definizioni valgono più in generale, se $f: A \rightarrow \mathbb{R}$ è una funzione reale di k variabili reali.

Esempio. Proviamo che

$$\lim_{(x,y) \rightarrow (0,0)} \frac{1}{(x^2 + y^2)^\alpha} = +\infty \quad (\alpha > 0).$$

Fissato $M > 0$, si ha

$$\frac{1}{(x^2 + y^2)^\alpha} = \frac{1}{(\sqrt{x^2 + y^2})^{2\alpha}} \geq \frac{1}{\delta^{2\alpha}} > M$$

pur di prendere $\delta < 1/M^{1/2\alpha}$.

Esercizio. Provare che

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\text{sen } x^2}{\sqrt{x^2 + y^2}} = 0, \quad \lim_{(x,y) \rightarrow (0,0)} \frac{\text{sen } y}{x^2 y} = +\infty.$$

Osservazione. Supponiamo che **esista** il limite per $P \rightarrow P_0$ di $f(P)$. Allora, fissata una qualunque *direzione* v , cioè un vettore $v \in \mathbb{R}^k$ tale che $\|v\| = 1$, è facile verificare che esiste ed è uguale al precedente anche

$$\lim_{t \rightarrow 0} f(P_0 + tv).$$

Nel caso ad esempio di due variabili, se esiste il limite per $(x, y) \rightarrow (x_0, y_0)$ di $f(x, y)$, allora, fissato un qualunque vettore $v = (h, k) \in \mathbb{R}^2$ tale che $h^2 + k^2 = 1$, è facile verificare che esiste ed è uguale al precedente anche

$$\lim_{t \rightarrow 0} f(x_0 + th, y_0 + tk).$$

In altre parole, **se esiste** il limite per $P \rightarrow P_0$ di $f(P)$, allora il *limite direzionale*

$$\lim_{t \rightarrow 0} f(P_0 + tv)$$

esiste per ogni direzione v ed è indipendente da essa. Di conseguenza, se il limite direzionale (in P_0) dipende dalla direzione (o non esiste per qualche direzione), allora $f(P)$ non ammette limite per $P \rightarrow P_0$.

Esempio. Consideriamo il

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy}{x^2 + y^2}.$$

Fissiamo un vettore $(h, k) \in \mathbb{R}^2$ tale che $h^2 + k^2 = 1$ ed eseguiamo le sostituzioni $x = th$, $y = tk$. Si ha

$$\lim_{t \rightarrow 0} \frac{t^2 hk}{t^2 h^2 + t^2 k^2} = \frac{hk}{h^2 + k^2} = hk,$$

ricordando che $h^2 + k^2 = 1$. Ad esempio, nelle direzioni degli assi coordinati il limite viene zero, mentre nella direzione $(h, k) = (1/\sqrt{2}, 1/\sqrt{2})$, cioè nella direzione della retta $y = x$ il limite viene $1/2$. Il limite direzionale dipende dunque dalla direzione (h, k) . Si può pertanto concludere che la funzione $xy/(x^2 + y^2)$ non ammette limite per $(x, y) \rightarrow (0, 0)$.

Osservazione (importante). L'esistenza del limite direzionale non garantisce tuttavia l'esistenza del limite. Consideriamo ad esempio

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x}{y} \sqrt{x^2 + y^2}.$$

Il limite direzionale

$$\lim_{t \rightarrow 0} \frac{th}{tk} \sqrt{t^2 h^2 + t^2 k^2} = \lim_{t \rightarrow 0} \frac{h}{k} |t| \sqrt{h^2 + k^2} = \lim_{t \rightarrow 0} \frac{h}{k} |t|$$

esiste ed è uguale a 0 lungo ogni direzione $v = (h, k)$, $k \neq 0$. Perciò, per l'osservazione precedente, se esistesse il limite della funzione per $(x, y) \rightarrow (0, 0)$, esso sarebbe necessariamente uguale a 0. D'altra parte, considerando il limite della restrizione della funzione data alla curva $y = x^2$, $x > 0$, si ottiene

$$\lim_{x \rightarrow 0^+} \frac{x}{x^2} \sqrt{x^2 + x^4} = 1.$$

Perciò il limite considerato non esiste.

Per le funzioni reali di più variabili si hanno teoremi analoghi a quelli già incontrati nel caso di una variabile. Con le opportune modifiche, valgono infatti i seguenti risultati (si invita lo studente a formularne gli enunciati):

1. teorema sulle operazioni tra limiti;
2. teorema di unicità del limite;
3. teorema della permanenza segno;
4. teorema dei carabinieri (e del carabiniere).

Diremo che una funzione $f: A \rightarrow \mathbb{R}$, definita su un sottoinsieme A di \mathbb{R}^k è *limitata* [limitata superiormente, limitata inferiormente] se lo è la sua immagine. Non è difficile verificare che f è limitata se e solo se esiste una costante $M > 0$ tale che $|f(P)| \leq M$ per ogni $P \in A$.

Corollario (del teorema dei carabinieri). *Siano f e g due funzioni reali definite in un sottoinsieme A di \mathbb{R}^k . Supponiamo che f sia limitata e che $g(P) \rightarrow 0$ per $P \rightarrow P_0$. Allora $f(P)g(P) \rightarrow 0$ per $P \rightarrow P_0$.*

Dimostrazione. Poichè f è limitata, esiste $M > 0$ tale che $|f(P)| \leq M$ per ogni $P \in A$. Si ha

$$0 \leq |f(P)g(P)| \leq M |g(P)|,$$

da cui, per il teorema dei carabinieri, $|f(P)g(P)| \rightarrow 0$ per $P \rightarrow P_0$. Di conseguenza, anche $f(P)g(P) \rightarrow 0$ per $P \rightarrow P_0$. \square

Esempio. Come applicazione del precedente corollario, calcoliamo il seguente limite:

$$\lim_{(x,y) \rightarrow (0,0)} \frac{xy^2}{x^2 + y^2}.$$

Osserviamo che la funzione $f(x, y) = xy^2/(x^2 + y^2)$ è il prodotto di tre funzioni:

$$\frac{x}{\sqrt{x^2 + y^2}}, \quad \frac{y}{\sqrt{x^2 + y^2}} \quad \text{e} \quad y.$$

Le prime due non ammettono limite per $(x, y) \rightarrow (0, 0)$, come si vede facilmente controllando il limite direzionale (che, in questo caso, dipende dalla direzione). La terza funzione, invece, tende a zero. Non è dunque applicabile il teorema del prodotto dei limiti. Tuttavia, si ha

$$\frac{|x|}{\sqrt{x^2 + y^2}} \leq 1, \quad \frac{|y|}{\sqrt{x^2 + y^2}} \leq 1.$$

Pertanto, per il corollario precedente, si può concludere che $xy^2/(x^2 + y^2) \rightarrow 0$ per $(x, y) \rightarrow (0, 0)$.

La seguente osservazione illustra un metodo a volte utile per calcolare i limiti di funzioni di due variabili.

Osservazione (facoltativa). Dato un punto $(x_0, y_0) \in \mathbb{R}^2$, ogni punto (x, y) del piano si può rappresentare nel seguente modo:

$$\begin{cases} x = x_0 + \rho \cos \theta \\ y = y_0 + \rho \sin \theta \end{cases}$$

dove $\rho \geq 0$ e $0 \leq \theta < 2\pi$ sono dette *coordinate polari*.

Usando la definizione di limite, si può verificare che

$$\lim_{(x,y) \rightarrow (x_0,y_0)} f(x, y) = l \in \mathbb{R}$$

se e solo se

$$\lim_{\rho \rightarrow 0} \sup_{\theta \in [0, 2\pi)} |f(x_0 + \rho \cos \theta, y_0 + \rho \sin \theta) - l| = 0.$$

Una condizione sufficiente a garantire che $\lim_{\rho \rightarrow 0} \sup_{\theta \in [0, 2\pi)} |f(x_0 + \rho \cos \theta, y_0 + \rho \sin \theta) - l| = 0$ si ottiene se è possibile trovare una funzione $g(\rho)$ infinitesima per $\rho \rightarrow 0$ (e che non dipende da θ) tale che

$$|f(x_0 + \rho \cos \theta, y_0 + \rho \sin \theta) - l| \leq g(\rho), \quad \text{per ogni } \theta \in [0, 2\pi).$$

Ad esempio, proviamo che

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\sin(x^2 + 3y^2)}{\sqrt{x^2 + y^2}} = 0.$$

Si ha

$$\frac{|\sin(\rho^2 \cos^2 \theta + 3\rho^2 \sin^2 \theta)|}{\rho} \leq |\rho \cos^2 \theta + 3\rho \sin^2 \theta| \leq 4\rho,$$

dove ovviamente $|\cos^2 \theta + 3 \sin^2 \theta| \leq 4$ e $g(\rho) = 4\rho$ tende a 0 per $\rho \rightarrow 0$. \square

La continuità di una funzione di più variabili si definisce analogamente al caso di una sola variabile:

Definizione (di continuità in un punto). Data $f: A \rightarrow \mathbb{R}$, con $A \subseteq \mathbb{R}^k$, e dato un punto $P_0 \in A$ diremo che f è *continua in* P_0 se P_0 è un punto isolato di A oppure, nel caso che P_0 sia un **punto di accumulazione** di A , se $f(P) \rightarrow f(P_0)$ per $P \rightarrow P_0$. In ogni caso, f è continua in P_0 se per ogni $\epsilon > 0$ esiste un $\delta > 0$ tale che da $P \in A$ e $\|P - P_0\| < \delta$ segue $|f(P) - f(P_0)| < \epsilon$.

In modo equivalente, usando il Teorema di collegamento tra limiti di funzioni e limiti di successioni, si ha:

Definizione (di continuità in un punto con le successioni). Data $f: A \rightarrow \mathbb{R}$, con $A \subseteq \mathbb{R}^k$, e dato un punto $P_0 \in A$ di accumulazione per A diremo che f è

continua in P_0 se per ogni successione $\{P_n\}_{n \in \mathbb{N}}$ in A tale che $P_n \rightarrow P_0$ risulta $f(P_n) \rightarrow f(P_0)$.

Si osservi che, in base alla precedente definizione, una funzione reale $f(x, y)$ definita in un insieme $A \subseteq \mathbb{R}^2$ è continua in un punto $P_0 = (x_0, y_0) \in A$ se e solo se è vera la seguente proposizione:

$$(x_n \rightarrow x_0) \wedge (y_n \rightarrow y_0) \Rightarrow f(x_n, y_n) \rightarrow f(x_0, y_0).$$

Esempio. La funzione $f(x, y) = \sin xy$ è continua in ogni $(x_0, y_0) \in \mathbb{R}^2$. Infatti, prese due successioni $x_n \rightarrow x_0$ e $y_n \rightarrow y_0$, dobbiamo provare che $\sin x_n y_n \rightarrow \sin x_0 y_0$. Questo segue immediatamente dal fatto che, posto $z_n = x_n y_n$ e $z_0 = x_0 y_0$, si ha $\sin z_n \rightarrow \sin z_0$ come provato nella parte riguardante le funzioni di una variabile.

Esempio. La funzione

$$f(x, y) = \begin{cases} \frac{\sin(x^2+y^2)}{x^2+y^2} & \text{se } (x, y) \neq (0, 0) \\ 1 & \text{se } (x, y) = (0, 0) \end{cases}$$

è continua in $(0, 0)$. Infatti, prese due successioni $x_n \rightarrow 0$ e $y_n \rightarrow 0$, dobbiamo provare che

$$\frac{\sin(x_n^2 + y_n^2)}{x_n^2 + y_n^2} \rightarrow 1.$$

Questo segue immediatamente dal fatto che, posto $z_n = x_n^2 + y_n^2$, si ha $\frac{\sin z_n}{z_n} \rightarrow 1$ come provato nella parte riguardante le funzioni di una variabile.

Definizione (di funzione continua). Sia $f: A \rightarrow \mathbb{R}$ una funzione reale di k variabili reali e sia B un sottoinsieme del dominio A di f . Si dice che f è *continua in B* se è continua in ogni punto di B . Si dice semplicemente che f è *continua* (senza ulteriori precisazioni), se è continua in ogni punto del suo dominio A .

Esercizio. Provare che se $f: \mathbb{R}^k \rightarrow \mathbb{R}$ è una funzione costante, allora è continua.

Osserviamo che la funzione di una sola variabile x può essere pensata anche come una funzione di due variabili (costante rispetto alla seconda variabile y). Precisamente, la funzione x , pensata definita in \mathbb{R}^2 , è quella legge che ad ogni punto $(x, y) \in \mathbb{R}^2$ associa la sua ascissa (cioè la prima coordinata). Analogamente la funzione y è quell'applicazione che ad ogni $(x, y) \in \mathbb{R}^2$ fa corrispondere la sua ordinata, cioè la seconda coordinata di (x, y) . Le funzioni x e y si chiamano, rispettivamente, *prima* e *seconda funzione coordinata* (cartesiana).

Esercizio. Provare che le funzioni coordinate x e y sono continue.

Teorema (di continuità delle funzioni combinate). *Se una funzione reale di una o più variabili reali è ottenuta combinando funzioni continue mediante operazioni di somma, prodotto, quoziente e composizione, allora è continua.*

Si osservi che dal precedente teorema si deduce che i monomi di due variabili, cioè le funzioni del tipo $ax^n y^m$ (dove a è una costante ed n ed m sono interi non negativi) sono funzioni continue. Di conseguenza anche i polinomi di due variabili (essendo somma di monomi) e le funzioni razionali (cioè rapporto di polinomi) sono funzioni continue.

Teorema di Weierstrass in \mathbb{R}^k . *Sia $f: A \rightarrow \mathbb{R}$ una funzione continua in un sottoinsieme chiuso e limitato $A \subseteq \mathbb{R}^k$. Allora f ammette minimo e massimo assoluti in A .*

Derivate parziali

Sia $f(x, y)$ una funzione reale di due variabili reali. Se si fissa una delle due variabili, ad esempio se si fissa $y = y_0$, si ottiene la funzione di una sola variabile $x \mapsto f(x, y_0)$, detta *funzione parziale* della x (per $y = y_0$).

In modo analogo si può fare per funzioni di tre variabili [di k variabili].

Sia $f: A \rightarrow \mathbb{R}$ una funzione reale definita su un aperto A di \mathbb{R}^2 e sia $P_0 = (x_0, y_0)$ un punto di A . La *derivata (parziale)* nel punto P_0 di f rispetto alla x (o meglio, rispetto alla prima variabile) è (se esiste) la derivata in x_0 della funzione parziale (reale di variabile reale) $x \mapsto f(x, y_0)$, cioè è il limite (se esiste finito) per $x \rightarrow x_0$ del rapporto incrementale (nella variabile x)

$$\frac{f(x, y_0) - f(x_0, y_0)}{x - x_0}.$$

In modo analogo si definisce la derivata (parziale) rispetto alla seconda variabile. La derivata parziale di f rispetto ad x in $P_0 = (x_0, y_0)$ si denota con uno dei seguenti simboli:

$$\frac{\partial f}{\partial x}(x_0, y_0), \quad \frac{\partial f}{\partial x}(P_0), \quad f_x(x_0, y_0), \quad f_x(P_0).$$

Un'analogia notazione vale per la derivata rispetto ad y . Diremo poi che f è *derivabile (parzialmente)* rispetto ad x se è derivabile (rispetto ad x) in ogni punto del dominio. In questo caso risulta ben definita la funzione, detta derivata rispetto ad x , che ad ogni punto $(x, y) \in A$ associa il numero

$$f_x(x, y).$$

La definizione della funzione derivata rispetto ad y è analoga.

Si dice che f è *derivabile* se è derivabile sia rispetto ad x sia rispetto ad y .

Più in generale, se $f(x_1, x_2, \dots, x_k)$ è una funzione di k variabili, la derivata (parziale) rispetto alla variabile x_i , $i = 1, \dots, k$, in $P_0 = (x_i^0, \dots, x_k^0)$ si denota

$$\frac{\partial f}{\partial x_i}(x_i^0, \dots, x_k^0), \quad \frac{\partial f}{\partial x_i}(P_0), \quad f_{x_i}(x_i^0, \dots, x_k^0), \quad f_{x_i}(P_0).$$

Osservazione. A differenza di quanto accade per le funzioni di una sola variabile, una funzione di due (o più) variabili può essere derivabile senza che risulti continua come mostra l'esempio che segue. Tuttavia, se una funzione è derivabile e le sue derivate sono continue, allora si può provare che è continua (come conseguenza del Teorema del Differenziale Totale che enunceremo in seguito).

Esempio (di funzione derivabile ma non continua). Consideriamo la funzione

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

Dalla definizione di derivata parziale segue subito che è derivabile in $(0, 0)$ e che le due derivate parziali (in tale punto) sono nulle. D'altra parte è immediato verificare che f non è continua nell'origine: infatti, come visto in precedenza, il limite direzionale dipende dalla direzione e, di conseguenza, non esiste il limite per $(x, y) \rightarrow (0, 0)$ di $f(x, y)$. Per quanto osservato sopra, notiamo che f , non essendo continua, non può neppure avere le derivate continue (si invita lo studente a verificare direttamente la discontinuità in $(0, 0)$ delle derivate parziali di f).

Sia $f: A \rightarrow \mathbb{R}$ un'applicazione definita su un aperto A di \mathbb{R}^k e sia $P_0 \in A$. Fissata una qualunque direzione v , cioè un vettore $v \in \mathbb{R}^k$ tale che $\|v\| = 1$, la *derivata (direzionale) di f in P_0 lungo il vettore v* è (se esiste) la derivata della funzione di una variabile $t \in (-\delta, \delta) \mapsto f(P_0 + tv)$ in $t_0 = 0$. In altre parole, f è derivabile in P_0 lungo il vettore v se esiste finito

$$\lim_{t \rightarrow 0} \frac{f(P_0 + tv) - f(P_0)}{t}.$$

Ad esempio, se f è di due variabili, $P_0 = (x_0, y_0)$ e $v = (h, k)$ con $h^2 + k^2 = 1$, diremo che f è derivabile in (x_0, y_0) lungo il vettore v se esiste finito

$$\lim_{t \rightarrow 0} \frac{f(x_0 + th, y_0 + tk) - f(x_0, y_0)}{t}.$$

La derivata direzionale di f in P_0 si denota con uno dei seguenti simboli:

$$\frac{\partial f}{\partial v}(P_0), \quad f_v(P_0).$$

Esempio. Calcoliamo la derivata di

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

in $(0, 0)$ lungo una generica direzione $v = (h, k)$ con $h^2 + k^2 = 1$. Si ha

$$\lim_{t \rightarrow 0} \frac{f(th, tk) - f(0, 0)}{t} = \lim_{t \rightarrow 0} \frac{t^2 h^2 tk}{t(t^2 h^2 + t^2 k^2)} = h^2 k.$$

Si ottiene perciò,

$$\frac{\partial f}{\partial v}((0, 0)) = h^2 k.$$

Ad esempio, lungo la direzione della bisettrice $y = x$ (scegliendo il verso crescente delle x) si ha $v = (1/\sqrt{2}, 1/\sqrt{2})$ e quindi la derivata precedente vale $1/(2\sqrt{2})$.

Osservazione. E' facile verificare che la funzione dell'esempio precedente è continua in $(0, 0)$ (la verifica è lasciata per esercizio). Non è vero però, in generale, che l'esistenza delle derivate lungo tutte le direzioni implichi la continuità di una funzione. Ad esempio, la funzione

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

è derivabile in $(0, 0)$ lungo tutte le direzioni $v = (h, k)$, $k \neq 0$, essendo

$$\lim_{t \rightarrow 0} \frac{th}{t t^2 k^2} (t^2 h^2 + t^2 k^2) = \frac{h}{k^2},$$

ma non è continua in $(0, 0)$ come si può subito verificare osservando che ad esempio $\lim_{y \rightarrow 0} f(0, y) = 0$, mentre $\lim_{x \rightarrow 0^+} f(x, x^2) = +\infty$ (notiamo che in questo esempio ogni limite direzionale, cioè $\lim_{t \rightarrow 0} f(th, tk)$, vale 0).

Osservazione. Le derivate parziali di f in P_0 sono le derivate direzionali di f in P_0 lungo le direzioni degli assi. Ad esempio, se f è di due variabili, la derivata parziale $f_x(P_0)$ si ottiene per $v = (1, 0)$, mentre $f_y(P_0)$ si ottiene per $v = (0, 1)$.

Introduciamo ora il concetto di differenziabilità. Consideriamo dapprima il caso di una funzione di due variabili.

Definizione. Sia $f: A \rightarrow \mathbb{R}$ una funzione reale definita su un aperto A di \mathbb{R}^2 e sia $P_0 = (x_0, y_0)$ un punto di A . Supponiamo che esistano $f_x(P_0)$ e $f_y(P_0)$. Diremo che f è differenziabile in P_0 , se

$$\lim_{(x, y) \rightarrow (x_0, y_0)} \frac{f(x, y) - f(x_0, y_0) - f_x(P_0)(x - x_0) - f_y(P_0)(y - y_0)}{\sqrt{(x - x_0)^2 + (y - y_0)^2}} = 0$$

Perciò, f è differenziabile in P_0 , se e solo se risulta

$$f(x, y) = f(x_0, y_0) + f_x(P_0)(x - x_0) + f_y(P_0)(y - y_0) + o\left(\sqrt{(x - x_0)^2 + (y - y_0)^2}\right).$$

Il polinomio omogeneo di primo grado

$$df(P_0)(x - x_0, y - y_0) = f_x(P_0)(x - x_0) + f_y(P_0)(y - y_0)$$

si chiama il *differenziale* della funzione f calcolato in P_0 e applicato al vettore incremento $P - P_0 = (x - x_0, y - y_0)$.

Osserviamo che, se una funzione è differenziabile in P_0 , il suo differenziale calcolato in P_0 e applicato al vettore incremento $P - P_0 = (x - x_0, y - y_0)$ approssima l'incremento $f(x, y) - f(x_0, y_0)$ di f in modo tale che la differenza tende a zero più velocemente rispetto alla norma dell'incremento stesso.

Geometricamente l'equazione

$$z = f(x_0, y_0) + f_x(P_0)(x - x_0) + f_y(P_0)(y - y_0)$$

rappresenta un piano passante per il punto $(x_0, y_0, f(x_0, y_0))$ che è chiamato *piano tangente* al grafico di f in $(x_0, y_0, f(x_0, y_0))$.

Se una funzione f ammette derivate parziali in un punto (x_0, y_0) , in tale punto si definisce il vettore *gradiente* di f

$$\text{grad}f(x_0, y_0) = (f_x(x_0, y_0), f_y(x_0, y_0)).$$

Più in generale, se $f(x_1, x_2, \dots, x_k)$ è una funzione di k variabili che ammette derivate parziali in un punto P_0 di un aperto A di \mathbb{R}^k , il gradiente di f in P_0 è il vettore

$$\text{grad}f(P_0) = (f_{x_1}(P_0), f_{x_2}(P_0), \dots, f_{x_k}(P_0)).$$

Il gradiente di f in P_0 si indica anche $\nabla f(P_0)$. Avendo introdotto la nozione di gradiente, possiamo dire, in maniera equivalente alla definizione data, che f è *differenziabile* in P_0 se

$$\lim_{P \rightarrow P_0} \frac{f(P) - f(P_0) - \text{grad}f(P_0) \cdot (P - P_0)}{\|P - P_0\|} = 0,$$

o anche se

$$\lim_{H \rightarrow 0} \frac{f(P_0 + H) - f(P_0) - \text{grad}f(P_0) \cdot H}{\|H\|} = 0,$$

dove $H = P - P_0$.

In questo caso il differenziale della funzione f calcolato in P_0 e applicato al vettore incremento H è il polinomio omogeneo di primo grado

$$df(P_0)H = \text{grad}f(P_0) \cdot H.$$

Per funzioni reali di una variabile, abbiamo visto che la derivabilità in un punto x_0 implica la differenziabilità della funzione in x_0 (anzi, si può provare che è equivalente ad essa). In più variabili, non è più vero che l'esistenza delle derivate parziali (o, anche, delle derivate direzionali) implichi la differenziabilità, come mostra l'esempio che segue.

Esempio. Consideriamo la funzione $f(x, y) = \sqrt{|xy|}$. La funzione f è ovviamente continua in $(0, 0)$, essendo prodotto e composizione di funzioni continue. Poiché la restrizione di f sia all'asse x che all'asse y è la funzione nulla, le derivate parziali $f_x(0, 0)$ e $f_y(0, 0)$ esistono e valgono zero. D'altra parte,

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\sqrt{|xy|}}{\sqrt{x^2 + y^2}}$$

non esiste come si verifica immediatamente considerando i limiti direzionali. Perciò f non è differenziabile in $(0, 0)$.

Osserviamo esplicitamente che è il concetto di differenziabilità (e non quello di derivabilità) che permette di estendere alle funzioni di più variabili la nozione di derivabilità introdotta nella prima parte del corso per funzioni di una variabile, in modo tale che vengano ancora garantite importanti proprietà come, ad esempio, la continuità. Infatti, come già mostrato in precedenza con un esempio, per funzioni di più variabili, la derivabilità non implica la continuità. Si deduce, invece, immediatamente dalla nozione di differenziabilità il seguente:

Teorema. Sia $f: A \rightarrow \mathbb{R}$ una funzione reale definita su un aperto A di \mathbb{R}^k differenziabile in $P_0 \in A$. Allora f è continua in P_0 .

Dimostrazione. Si tratta di provare che $f(P) \rightarrow f(P_0)$ per $P \rightarrow P_0$ o, equivalentemente, che $\lim_{P \rightarrow P_0} (f(P) - f(P_0)) = 0$. Dalla differenziabilità di f in P_0 si ha, per P in un conveniente intorno di P_0 ,

$$f(P) - f(P_0) = \text{grad}f(P_0) \cdot (P - P_0) + o(\|P - P_0\|).$$

Tenendo conto della disuguaglianza triangolare e della disuguaglianza di Schwarz, si ottiene

$$|f(P) - f(P_0)| \leq |\text{grad}f(P_0) \cdot (P - P_0)| + |o(\|P - P_0\|)| \leq$$

$$\|\operatorname{grad}f(P_0)\| \|P - P_0\| + |o(\|P - P_0\|)|.$$

Poiché sia $\|P - P_0\|$ che $|o(\|P - P_0\|)|$ sono infinitesimi per $P \rightarrow P_0$, dal Teorema dei carabinieri si deduce $\lim_{P \rightarrow P_0} |(f(P) - f(P_0))| = 0$. Di conseguenza, anche $\lim_{P \rightarrow P_0} (f(P) - f(P_0)) = 0$. \square

6^a settimana - dal 1.04.19

La differenziabilità di una funzione fornisce anche un'utile formula per il calcolo delle derivate direzionali:

Teorema. Sia $f: A \rightarrow \mathbb{R}$ una funzione reale definita su un aperto A di \mathbb{R}^k differenziabile in $P_0 \in A$. Allora f ammette derivate direzionali in P_0 lungo ogni direzione v e si ha

$$\frac{\partial f}{\partial v}(P_0) = \text{grad}f(P_0) \cdot v$$

Dimostrazione. Poiché f è differenziabile in P_0 che è interno ad A , esiste $\delta > 0$ tale che, data una direzione v , si ha per ogni $|t| < \delta$,

$$f(P_0 + tv) - f(P_0) = \text{grad}f(P_0) \cdot (tv) + o(\|tv\|),$$

da cui

$$\frac{f(P_0 + tv) - f(P_0)}{t} = \frac{\text{grad}f(P_0) \cdot (tv) + o(\|tv\|)}{t} = \frac{t \text{grad}f(P_0) \cdot v}{t} + \frac{o(\|tv\|)}{t}.$$

Passando al limite per $t \rightarrow 0$ e osservando che

$$\frac{o(\|tv\|)}{t} = \frac{o(\|tv\|)}{\|tv\|} \frac{\|tv\|}{t} = \frac{o(\|tv\|)}{\|tv\|} \frac{|t|}{t} \rightarrow 0,$$

si ottiene

$$\lim_{t \rightarrow 0} \frac{f(P_0 + tv) - f(P_0)}{t} = \text{grad}f(P_0) \cdot v.$$

Perciò f ammette derivata direzionale in P_0 lungo la direzione v e risulta

$$\frac{\partial f}{\partial v}(P_0) = \text{grad}f(P_0) \cdot v.$$

□

Osservazione. Il gradiente di una funzione, nei punti in cui è non nullo, indica la direzione di massima e minima crescita della funzione stessa. Infatti, la massima variazione di f si ha in quelle direzioni nelle quali $\text{grad}f(P_0)$ e v sono paralleli. In particolare la direzione di massima crescita si ha quando $v = \text{grad}f(P_0)/\|\text{grad}f(P_0)\|$ (cioè $\text{grad}f(P_0)$ e v concordi) mentre quella di minima crescita si ha quando $v = -\text{grad}f(P_0)/\|\text{grad}f(P_0)\|$ (cioè $\text{grad}f(P_0)$ e v discordi).

Osservazione. L'uguaglianza $\frac{\partial f}{\partial v}(P_0) = \text{grad}f(P_0) \cdot v$ non è più valida se f non è differenziabile in P_0 . Consideriamo ad esempio la funzione

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

Poiché la restrizione di f sia all'asse x che all'asse y è la funzione nulla, si ha ovviamente $f_x(0,0) = f_y(0,0) = 0$ e, quindi, $\text{grad}f(\mathbf{0}) \cdot v = 0$ qualunque sia la direzione v (con $\mathbf{0}$ denotiamo il vettore nullo di \mathbb{R}^2 , cioè $\mathbf{0} = (0,0)$). D'altra parte, per ogni direzione $v = (h, k)$ con $h \neq 0$ e $k \neq 0$, $h^2 + k^2 = 1$, si ha

$$\frac{\partial f}{\partial v}(0,0) = \lim_{t \rightarrow 0} \frac{|th|tk}{t\sqrt{t^2(h^2+k^2)}} = |h|k \neq 0.$$

Ciò prova che in questo caso l'uguaglianza $\frac{\partial f}{\partial v}(P_0) = \text{grad}f(P_0) \cdot v$ non vale. Di conseguenza, f non può essere differenziabile in $(0,0)$. (Si invita lo studente a verificare ciò anche applicando direttamente la definizione di differenziabilità, cioè calcolando il limite, per $(x,y) \rightarrow (0,0)$, di $|x|y/(x^2+y^2)$).

Per funzioni di più variabili, mentre, come già osservato, l'**esistenza** delle derivate parziali non garantisce la differenziabilità, si può invece provare che la **continuità** delle derivate parziali implica la differenziabilità. Si ha infatti il seguente

Teorema (del Differenziale Totale). *Sia $f: A \rightarrow \mathbb{R}$ una funzione reale definita su un aperto A di \mathbb{R}^k con le derivate parziali continue in $P_0 \in A$. Allora f è differenziabile (e, quindi, continua) in P_0 .*

La condizione espressa dal teorema del Differenziale Totale non è però necessaria alla differenziabilità. Consideriamo infatti il seguente

Esempio. La funzione

$$f(x,y) = \begin{cases} x^2y^2 \operatorname{sen} \frac{1}{x^2y^2} & \text{se } xy \neq 0 \\ 0 & \text{se } xy = 0 \end{cases}$$

è differenziabile in $(0,0)$. Infatti $f_x(0,0) = f_y(0,0) = 0$ e

$$\lim_{(x,y) \rightarrow (0,0)} \frac{f(x,y)}{\sqrt{x^2+y^2}} = 0.$$

D'altra parte, considerando ad esempio la derivata rispetto ad x , si ha

$$f_x(x,y) = 2xy^2 \operatorname{sen} \frac{1}{x^2y^2} - \frac{2}{x} \cos \frac{1}{x^2y^2}$$

che non ammette limite finito per $(x,y) \rightarrow (0,0)$ come si può verificare considerando, ad esempio, il limite della restrizione di f_x alla retta $y = x$. Perciò $\lim_{(x,y) \rightarrow (0,0)} f_x(x,y) \neq f_x(0,0)$, cioè f_x non è continua in $(0,0)$.

Definizione. Sia $f: A \rightarrow \mathbb{R}$ una funzione reale definita in un sottoinsieme A di \mathbb{R}^k . Un punto $P_0 \in A$ si dice *di minimo* [*di massimo*] *relativo* (o locale) per f in

A se esiste un intorno $B_\delta(P_0)$ di centro P_0 e raggio $\delta > 0$ tale che $f(P) \geq f(P_0)$ [$f(P) \leq f(P_0)$] per ogni $P \in B_\delta(P_0) \cap A$. Un punto di minimo o di massimo relativo per f (in A) si dice *estremante* per f (in A).

Teorema di Fermat (per funzioni di due variabili). *Sia $f: A \rightarrow \mathbb{R}$ una funzione reale di due variabili reali e sia $P_0 = (x_0, y_0) \in A$. Supponiamo che siano soddisfatte le seguenti tre ipotesi:*

1. P_0 è interno ad A ;
2. f è derivabile in P_0 ;
3. P_0 è un punto estremante per f in A .

Allora

$$f_x(P_0) = 0, \quad f_y(P_0) = 0$$

o, equivalentemente,

$$\text{grad}f(P_0) = (f_x(P_0), f_y(P_0)) = (0, 0).$$

Dimostrazione. Consideriamo la funzione parziale $\varphi: x \mapsto f(x, y_0)$. Dall'ipotesi 1), essendo il punto (x_0, y_0) interno ad A , essa risulta definita almeno in un intervallo aperto di \mathbb{R} contenente x_0 . Dalle ipotesi 2) e 3) segue immediatamente che φ è derivabile in x_0 e che x_0 è estremante per φ . Di conseguenza, il Teorema di Fermat per funzioni di una variabile ci assicura che $\varphi'(x_0) = 0$. Pertanto, ricordandosi la definizione di derivata parziale, si ha

$$f_x(x_0, y_0) = \varphi'(x_0) = 0.$$

In modo analogo si prova che in (x_0, y_0) si annulla anche la derivata parziale di f rispetto ad y . \square

Osserviamo che, in base al teorema precedente, i punti estremanti di una funzione $f(x, y)$ vanno cercati tra le seguenti tre categorie:

1. punti di frontiera;
2. punti in cui la funzione non è derivabile;
3. punti in cui si annullano entrambe le derivate parziali.

Nessuna di queste tre condizioni ci assicura che un punto sia estremante. Tanto per fare un esempio, per la funzione $f(x, y) = x^2 - y^2$ l'origine non è né di massimo né di minimo (per provarlo basta considerare la restrizione di f agli assi), ma $f_x(0, 0) = f_y(0, 0) = 0$. Tuttavia, se un punto è estremante, almeno una delle tre condizioni precedenti deve essere soddisfatta.

Definizione. Un punto $P_0 \in A \subseteq \mathbb{R}^2$ tale che $f_x(P_0) = f_y(P_0) = 0$ è detto punto *critico* per f . Un punto critico che non è estremante è detto punto di *sella*.

Nella ricerca dei punti estremanti di una funzione di due variabili definita in un insieme $A \subseteq \mathbb{R}^2$ si procede per esclusione: come primo provvedimento, tenendo conto del Teorema di Fermat, si escludono tutti i punti interni ad A in cui la funzione risulta derivabile ed una delle due derivate parziali è diversa da zero. Rimangono da analizzare i punti di frontiera, i punti interni in cui la funzione non è derivabile e i punti interni in cui si annullano entrambe le derivate. Per escludere altri punti di frontiera si tiene conto del fatto che *se $P \in \partial A$ non è estremante per la restrizione di f alla frontiera di A , allora non lo è neppure per f in A* . Di solito, lo studio della restrizione di f a ∂A consente di escludere la grande maggioranza dei punti di frontiera. Con tale procedimento, quasi sempre rimangono pochi punti “sospetti” che possono essere analizzati a parte con metodi vari; ad esempio, **se si cercano gli estremi assoluti** di una funzione continua f in un insieme chiuso e limitato A (la cui esistenza è garantita dal teorema di Weierstrass) si può procedere direttamente mediante il confronto dei valori assunti.

Esempio. Determinare gli estremi assoluti di $f(x, y) = x^2 + 4y^2$ in

$$A = \{(x, y) \in \mathbb{R}^2 : (x - 3)^2 + y^2 \leq 1\}.$$

Osserviamo innanzitutto che, essendo f continua e A chiuso e limitato, l'esistenza degli estremi assoluti di f in A è assicurata dal teorema di Weierstrass. È immediato verificare che l'unico punto critico di f è $(0, 0)$ che non appartiene ad A . Il massimo e il minimo sono quindi assunti su ∂A . La restrizione di f alla semicirconferenza superiore è la funzione di una variabile $g(x) = f(x, \sqrt{1 - (x - 3)^2}) = x^2 + 4(1 - (x - 3)^2)$ con $2 \leq x \leq 4$. Come è facile verificare, la funzione g nell'intervallo $[2, 4]$ ha minimo in $x = 2$ e massimo in $x = 4$. Analogamente, la restrizione di f alla semicirconferenza inferiore è $h(x) = f(x, -\sqrt{1 - (x - 3)^2}) = x^2 + 4(1 - (x - 3)^2)$ con $2 \leq x \leq 4$ e in questo caso coincide con $g(x)$, essendo la funzione f pari (rispetto a y). Di conseguenza, $\min_A f(x, y) = f(2, 0) = 4$ e $\max_A f(x, y) = f(4, 0) = 16$.

Lo stesso risultato si può ottenere anche in modo geometrico prendendo in esame gli insiemi di livello di f . Per $c > 0$, gli insiemi di livello

$$S_c = \{(x, y) \in \mathbb{R}^2 : x^2 + 4y^2 = c\} = \{(x, y) \in \mathbb{R}^2 : \frac{x^2}{c} + \frac{y^2}{c/4} = 1\}$$

sono ellissi di semiassi \sqrt{c} e $\sqrt{c}/2$. Perciò il minimo di f in A si ottiene quando $\sqrt{c} = 2$ (e, quindi, è $c = 4$), mentre il massimo si ha quando $\sqrt{c} = 4$ (cioè $c = 16$).

Esercizio. Determinare gli estremi assoluti della funzione dell'esempio precedente nell'insieme

$$A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}.$$

Sugg. In questo caso il punto critico $P_0 = (0, 0)$ è interno ad A e risulta di minimo. Il massimo invece è assunto sulla frontiera di A .

Esempio. Determinare gli estremi assoluti di $f(x, y) = x^2 - y$ nel triangolo di vertici $A = (0, 0)$, $B = (1, 0)$, $C = (1, 1)$.

Poiché $f_y(x, y) = -1$ per ogni (x, y) , la f non ha punti critici. Gli estremi assoluti sono assunti perciò sulla frontiera del triangolo. La restrizione di f al lato AB è $g(x) = f(x, 0) = x^2$ con $0 \leq x \leq 1$ e quindi ha minimo in $x = 0$ e massimo in $x = 1$. La restrizione di f al lato AC è $h(x) = f(x, x) = x^2 - x$ con $0 \leq x \leq 1$ e quindi ha massimo in $x = 0$ e $x = 1$ e minimo in $x = 1/2$. Infine, la restrizione di f al lato BC è $k(y) = f(1, y) = 1 - y$ con $0 \leq y \leq 1$ e quindi ha massimo in $y = 0$ e minimo in $y = 1$. Confrontando i valori assunti si ottiene $f(A) = 0$, $f(B) = 1$, $f(C) = 0$, $f(1/2, 1/2) = -1/4$. Concludendo, il minimo di f nel triangolo considerato è $-1/4$ ed è assunto in $(1/2, 1/2)$, mentre il massimo è 1 ed è assunto in $(1, 0)$.

Esercizio. Determinare gli estremi assoluti di $f(x, y) = 2x^2 + y^2$ in

$$A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \geq 4, |x| \leq 4, |y| \leq 4\}.$$

Verificare il risultato trovato prendendo in esame gli insiemi di livello di f .

Esercizio. Provare che esistono gli estremi assoluti della funzione $f(x, y) = \sqrt{1 - x^2 - y^2} - x$ e determinarli.

Sugg. Il dominio di f è l'insieme chiuso e limitato $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$ e f risulta continua in tale insieme. Perciò, per il teorema di Weierstrass, essa ammette massimo e minimo assoluti in A . Per determinarli, osserviamo che l'unico punto critico è $P_0 = (-1/\sqrt{2}, 0)$, mentre, considerando la restrizione di f a $\partial A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 = 1\}$ si ottengono i punti $P_1 = (-1, 0)$ e $P_2 = (1, 0)$. La conclusione segue da un confronto tra i valori assunti da f in P_0 , P_1 e P_2 .

Derivate parziali di ordine superiore

Se la derivata di f rispetto ad x è di nuovo derivabile rispetto ad x , allora f si dice derivabile due volte rispetto ad x (o che ammette *derivata seconda* rispetto ad x due volte). La derivata rispetto ad x della derivata rispetto ad x , calcolata in un punto (x, y) , cioè $\frac{\partial}{\partial x} \frac{\partial f}{\partial x}(x, y)$, si indica con uno dei seguenti simboli:

$$\frac{\partial^2 f}{\partial x^2}(x, y), \quad f_{xx}(x, y).$$

In maniera analoga, la derivata rispetto ad y della derivata rispetto ad x , calcolata in un punto (x, y) , cioè $\frac{\partial}{\partial y} \frac{\partial f}{\partial x}(x, y)$, si indica con uno dei simboli:

$$\frac{\partial^2 f}{\partial y \partial x}(x, y), \quad f_{xy}(x, y).$$

In modo simile si definiscono le altre derivate seconde:

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

Esercizio. Definire le derivate parziali di ordine superiore al secondo per una funzione di due variabili.

Definizione (di funzione di classe C^n). Una funzione $f(x, y)$ si dice di classe C^0 (o che è C^0 , o che appartiene a C^0) se è continua; si dice di classe C^1 se è derivabile e le sue derivate parziali sono C^0 . In generale, si dice che f è di classe C^n se è derivabile e le sue derivate parziali sono di classe C^{n-1} . Infine, si dice che f è di classe C^∞ se è C^n per ogni $n \in \mathbb{N}$.

La nozione di funzione di classe C^n può essere banalmente estesa alle funzioni di tre o più variabili.

Sappiamo che se una funzione di una variabile è derivabile, allora è anche continua. Come mostrato con un esempio, un risultato analogo è falso per le funzioni di due (o più) variabili. Tuttavia, come conseguenza del Teorema del Differenziale Totale, se una funzione di due (o più) variabili è derivabile e le sue derivate parziali sono continue, allora si può affermare che essa è continua. In altre parole, dal Teorema del Differenziale Totale si può dedurre il seguente risultato:

Teorema. Se f è una funzione di classe C^1 , allora è anche di classe C^0 .

Come conseguenza si hanno le seguenti proprietà:

Teorema. Se f è una funzione di classe C^n , allora è anche di classe C^{n-1} .

Teorema(di regolarità delle funzioni combinate). *La somma, il prodotto, il quoziente e la composizione di funzioni C^n , quando (e dove) ha senso, è ancora una funzione C^n .*

Osservazione. Si osservi che in \mathbb{R}^2 le funzioni costanti e le funzioni coordinate x e y sono di classe C^∞ , quindi, in base al teorema di regolarità delle funzioni combinate, i polinomi (e le funzioni razionali) di due variabili sono funzioni C^∞ .

Esercizio. Dedurre dal teorema di regolarità delle funzioni combinate che le seguenti funzioni di due variabili sono di classe C^∞ :

$$\frac{y^2 \cos(x-y)}{1-xy}, \quad (2^y - x^2y)^3, \quad \frac{\log(xy)}{x+y}, \quad \arctang(y/x), \quad \frac{1}{x^2+y^2}.$$

Esercizio. Dedurre dal teorema di regolarità delle funzioni combinate che la funzione

$$f(x, y) = xy\sqrt{|y|}$$

è di classe C^∞ nell'insieme (aperto) costituito da \mathbb{R}^2 meno l'asse delle x , cioè nell'insieme $A = \{(x, y) \in \mathbb{R}^2 : y \neq 0\}$.

Esercizio (facoltativo.) Mostrare che la funzione

$$f(x, y) = xy\sqrt{|y|}$$

è di classe C^1 su tutto il suo dominio \mathbb{R}^2 (cioè anche nei punti dell'asse x) ma non è di classe C^2 .

Suggerimento. Sia $(x_0, 0)$ un punto fissato sull'asse x . È facile verificare che $f_x(x_0, 0) = f_y(x_0, 0) = 0$. Per provare che f è di classe C^1 in $(x_0, 0)$ si tratta di verificare che $\lim_{(x,y) \rightarrow (x_0,0)} f_x(x, y) = f_x(x_0, 0) = 0$ e che $\lim_{(x,y) \rightarrow (x_0,0)} f_y(x, y) = f_y(x_0, 0) = 0$. D'altra parte, la derivata parziale $f_y(x, y)$ non è derivabile rispetto a y in $(x_0, 0)$ e quindi f non può essere di classe C^2 in tale punto.

Teorema di Schwarz. *Se $f : A \rightarrow \mathbb{R}$ è una funzione di classe C^2 su un aperto A di \mathbb{R}^2 , allora*

$$\frac{\partial^2 f}{\partial y \partial x}(x, y) = \frac{\partial^2 f}{\partial x \partial y}(x, y), \quad \forall (x, y) \in A.$$

Osservazione. Se una funzione è sufficientemente regolare, il Teorema di Schwarz ci permette di scambiare l'ordine di due qualunque delle sue derivate. Si considerino, ad esempio, le seguenti due derivate terze:

$$f_{xyx}(x, y), \quad f_{xxy}(x, y).$$

Si può affermare che sono uguali? Mostriamo che se f è di classe C^3 , la risposta è affermativa. Infatti, in tale ipotesi, la funzione

$$g(x, y) := f_x(x, y)$$

è di classe C^2 . Di conseguenza, le due derivate terze di $f(x, y)$ considerate sopra risultano uguali, essendo le derivate seconde miste della funzione $g(x, y)$. Non c'è quindi nessuna ambiguità nell'adottare per tali due derivate la seguente notazione:

$$\frac{\partial^3 f}{\partial y \partial x^2}(x, y),$$

che rappresenta la funzione che si ottiene derivando la f due volte rispetto alla x e una volta rispetto alla y , indipendentemente dall'ordine in cui si eseguono le tre derivate.

Abbiamo visto che, nel caso di funzioni di una variabile aventi derivata seconda continua, il segno della derivata seconda in un punto critico fornisce informazioni sulla natura di tale punto. In analogia, per una funzione di due variabili, vogliamo ora dare una condizione sulle derivate seconde sufficiente a stabilire la natura di un punto critico.

Sia $f: A \rightarrow \mathbb{R}$ una funzione di classe C^2 su un aperto A di \mathbb{R}^2 e sia $P_0 \in A$. Consideriamo la matrice

$$Hf(P_0) = \begin{pmatrix} f_{xx}(P_0) & f_{xy}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) \end{pmatrix}$$

che è detta la *matrice hessiana* di f in P_0 . Notiamo che, in conseguenza del Teorema di Schwarz, si ha $f_{xy}(P_0) = f_{yx}(P_0)$ e, pertanto, la matrice è simmetrica. Si ha il seguente

Teorema. *Sia $f: A \rightarrow \mathbb{R}$ una funzione definita su un aperto A di \mathbb{R}^2 e di classe C^2 in A . Supponiamo che in un punto $P_0 \in A$ si annullino le due derivate parziali di f . Se il determinante della matrice hessiana in P_0 , detto hessiano della f in P_0 , cioè il numero*

$$\det Hf(P_0) = f_{xx}(P_0)f_{yy}(P_0) - (f_{xy}(P_0))^2,$$

è positivo, allora P_0 è un punto estremante. Se è negativo, P_0 non è estremante. In particolare, nel caso che $\det Hf(P_0)$ sia positivo, P_0 è di minimo relativo o di massimo relativo a seconda che la derivata $f_{xx}(P_0)$ sia positiva o negativa.

Osservazione. Se $\det Hf(P_0) = 0$ non si ha nessuna informazione sulla natura del punto critico P_0 , come si vede immediatamente considerando i seguenti esempi:

1. $f(x, y) = x^4 + y^4$ ($(0, 0)$ è di minimo);
2. $f(x, y) = -(x^4 + y^4)$ ($(0, 0)$ è di massimo);
3. $f(x, y) = x^2 + y^3$ ($(0, 0)$ è di sella).

Esempio. Stabiliamo la natura dei punti critici di $f(x, y) = 2x^4 - y^2 - x^8$ in \mathbb{R}^2 .

Si ha $\text{grad}f(x, y) = (8x^3 - 8x^7, -2y) = (0, 0)$ nei punti $P_0 = (0, 0)$, $P_1 = (1, 0)$, $P_2 = (-1, 0)$. Calcolando l'hessiano in un generico punto $P = (x, y)$ si ha $\det Hf((x, y)) = -16x^2(3 - 7x^4)$. Perciò, valutandolo nei punti critici si ottiene $\det Hf(P_0) = 0$ e $\det Hf(P_1) = \det Hf(P_2) = 64$. Poiché $f_{xx}(P_1) = f_{xx}(P_2) = -32$, i punti P_1 e P_2 sono di massimo. Il teorema precedente non dà informazioni sul punto $P_0 = (0, 0)$ ma, considerando le restrizioni di f agli assi, cioè le due funzioni $x \mapsto f(x, 0)$ e $y \mapsto f(0, y)$, si verifica facilmente che esso è un punto di sella.

Esercizio. Stabilire la natura dei punti critici di $f(x, y) = \cos x + y^2$ in \mathbb{R}^2 .

Funzioni vettoriali di più variabili

Passiamo ora a considerare il caso di funzioni vettoriali di variabile vettoriale. Tali funzioni sono anche dette *campi vettoriali*. Sia $f: A \rightarrow \mathbb{R}^m$, $f(P) = (f_1(P), f_2(P), \dots, f_m(P))$ un'applicazione definita in un sottoinsieme A di \mathbb{R}^k .

Il limite di una funzione vettoriale di variabile vettoriale si definisce in modo analogo a come è stato fatto per una funzione reale di più variabili reali. Sia P_0 un punto di accumulazione per il dominio A di f . Si dice che $f(P)$ tende ad un vettore $l \in \mathbb{R}^m$ per P che tende a P_0 , e si scrive $f(P) \rightarrow l$ per $P \rightarrow P_0$, se per ogni $\epsilon > 0$ esiste $\delta > 0$ tale che da $0 < \|P - P_0\|_k < \delta$ e $P \in A$ segue $\|f(P) - l\|_m < \epsilon$ ($\|\cdot\|_k$ e $\|\cdot\|_m$ denotano la norma in \mathbb{R}^k e in \mathbb{R}^m rispettivamente). In questo caso si può anche scrivere

$$\lim_{P \rightarrow P_0} f(P) = l.$$

Si può provare che

$$\lim_{P \rightarrow P_0} f(P) = l = (l_1, l_2, \dots, l_m) \iff \lim_{P \rightarrow P_0} f_i(P) = l_i \quad \forall i = 1, 2, \dots, m.$$

Diremo che f è *continua* in un punto $P_0 \in A$ se sono continue (in P_0) le sue m funzioni componenti: f_1, f_2, \dots, f_m . Per quanto visto sopra sul limite di

una funzione vettoriale di variabile vettoriale, nel caso che P_0 sia un punto di accumulazione per A , ciò equivale ad affermare che

$$\lim_{P \rightarrow P_0} f(P) = f(P_0).$$

(o, equivalentemente, che $\lim_{P \rightarrow P_0} f_i(P) = f_i(P_0) \quad \forall i = 1, 2, \dots, m$.)

Analogamente, diremo che f è *derivabile* in un punto $P_0 \in A$ se sono derivabili (in P_0) le sue m funzioni componenti: f_1, f_2, \dots, f_m .

Sia $f: A \rightarrow \mathbb{R}^m$ un'applicazione derivabile definita su un aperto A di \mathbb{R}^k . Denotiamo con f_1, f_2, \dots, f_m le m funzioni reali che compongono la f (ricordiamo che sono funzioni reali di k variabili reali). Fissato un punto $P_0 \in A$, la matrice

$$Jf(P_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_k} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_k} \\ \vdots & & & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \dots & \frac{\partial f_m}{\partial x_k} \end{pmatrix}_{P_0}$$

si chiama *matrice jacobiana* dell'applicazione f in P_0 (il simbolo P_0 in basso a destra significa che tutti gli elementi della matrice si considerano calcolati nel punto P_0).

Altre notazioni usate per indicare la matrice jacobiana sono $J_f(P_0)$ oppure $f'(P_0)$.

Osserviamo che le m righe della matrice jacobiana sono rispettivamente i gradienti delle m componenti f_1, f_2, \dots, f_m della funzione f .

Quando $k = m$, la matrice $Jf(P_0)$ è quadrata. In questo caso ha senso il suo determinante, $\det Jf(P_0)$, chiamato *jacobiano* dell'applicazione f in P_0 .

Descriviamo ora alcuni esempi di funzioni vettoriali di variabile vettoriale che saranno particolarmente utili nel calcolo di integrali doppi e tripli.

Coordinate polari

$$f(\rho, \theta) = (\rho \cos \theta, \rho \sin \theta), \quad \rho \geq 0, \quad 0 \leq \theta \leq 2\pi.$$

La corrispondente matrice jacobiana è

$$Jf(\rho, \theta) = \begin{pmatrix} \cos \theta & -\rho \operatorname{sen} \theta \\ \operatorname{sen} \theta & \rho \cos \theta \end{pmatrix}$$

e

$$\det Jf(\rho, \theta) = \rho.$$

Coordinate sferiche

$$f(\rho, \varphi, \theta) = (\rho \operatorname{sen} \varphi \cos \theta, \rho \operatorname{sen} \varphi \operatorname{sen} \theta, \rho \cos \varphi),$$

$$\rho \geq 0, 0 \leq \varphi \leq \pi, 0 \leq \theta \leq 2\pi$$

(la coordinata φ è detta *colatitudine*, la coordinata θ è detta *longitudine*).

La corrispondente matrice jacobiana è

$$Jf(\rho, \varphi, \theta) = \begin{pmatrix} \cos \theta \operatorname{sen} \varphi & \rho \cos \theta \cos \varphi & -\rho \operatorname{sen} \theta \operatorname{sen} \varphi \\ \operatorname{sen} \theta \operatorname{sen} \varphi & \rho \operatorname{sen} \theta \cos \varphi & \rho \cos \theta \operatorname{sen} \varphi \\ \cos \varphi & -\rho \operatorname{sen} \varphi & 0 \end{pmatrix}$$

e

$$\det Jf(\rho, \varphi, \theta) = \rho^2 \operatorname{sen} \varphi.$$

Coordinate cilindriche

$$f(\rho, \theta, z) = (\rho \cos \theta, \rho \operatorname{sen} \theta, z), \quad \rho \geq 0, 0 \leq \theta \leq 2\pi, z \in \mathbb{R}.$$

La corrispondente matrice jacobiana è

$$Jf(\rho, \theta, z) = \begin{pmatrix} \cos \theta & -\rho \operatorname{sen} \theta & 0 \\ \operatorname{sen} \theta & \rho \cos \theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

e

$$\det Jf(\rho, \theta, z) = \rho.$$

7^a settimana - dal 08.04.19

Divergenza e rotore

Sia $f: A \subseteq \mathbb{R}^k \rightarrow \mathbb{R}^k$,

$$f(x_1, \dots, x_k) = (f_1(x_1, \dots, x_k), f_2(x_1, \dots, x_k), \dots, f_k(x_1, \dots, x_k))$$

un campo vettoriale di classe C^1 su un aperto A di \mathbb{R}^k . Si definisce *divergenza* di f e si indica con $\operatorname{div} f$ la funzione a **valori reali**

$$\operatorname{div} f(x_1, \dots, x_k) := \sum_{i=1}^k \frac{\partial f_i}{\partial x_i}(x_1, \dots, x_k)$$

Esempio. Sia $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ il campo vettoriale

$$f(x, y, z) = (xy, \operatorname{sen}(xyz), yz^2).$$

Si ha

$$\operatorname{div} f(x, y, z) = \frac{\partial f_1}{\partial x}(x, y, z) + \frac{\partial f_2}{\partial y}(x, y, z) + \frac{\partial f_3}{\partial z}(x, y, z) = y + xz \cos(xyz) + 2yz.$$

Data una funzione $u: A \rightarrow \mathbb{R}$ di classe C^2 su un aperto $A \subseteq \mathbb{R}^k$, la funzione reale $\operatorname{div} \operatorname{grad} u: A \rightarrow \mathbb{R}$ è detta *laplaciano* di u e si indica con Δu . Si ha, posto $P = (x_1, \dots, x_k)$,

$$\Delta u(P) = \operatorname{div} \operatorname{grad} u(P) = \operatorname{div} \left(\frac{\partial u}{\partial x_1}(P), \dots, \frac{\partial u}{\partial x_k}(P) \right) = \sum_{i=1}^k \frac{\partial^2 u}{\partial x_i^2}(P).$$

Esempio. Sia $u(x, y) = x^2 - y^2$. Si ha

$$\Delta u(x, y) = \frac{\partial^2 u}{\partial x^2}(x, y) + \frac{\partial^2 u}{\partial y^2}(x, y) = 2 - 2 = 0.$$

Sia $f: A \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}^3$,

$$f(x, y, z) = (f_1(x, y, z), f_2(x, y, z), f_3(x, y, z)),$$

un campo vettoriale di classe C^1 su un aperto A di \mathbb{R}^3 . Si definisce *rotore* di f e si indica con $\operatorname{rot} f$ la funzione a **valori vettoriali**

$$\operatorname{rot} f(x, y, z) := \left(\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}, \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}, \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y} \right),$$

dove le derivate parziali sono calcolate in $(x, y, z) \in A$. Se $\text{rot} f = 0$, il campo vettoriale f è detto *irrotazionale*.

Un metodo utile per calcolare le componenti di $\text{rot} f$ è quello di sviluppare rispetto alla prima riga il seguente determinante “formale”

$$\det \begin{pmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_1 & f_2 & f_3 \end{pmatrix}$$

dove e_1, e_2, e_3 rappresentano i versori degli assi x, y, z rispettivamente.

Esempio. Sia $f: \mathbb{R}^3 \rightarrow \mathbb{R}^3$ il campo vettoriale

$$f(x, y, z) = (yz^2, xyz, x^2y).$$

Si ha

$$\text{rot} f(x, y, z) = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ yz^2 & xyz & x^2y \end{pmatrix} = (x^2 - xy, 2yz - 2xy, yz - z^2).$$

Osservazione. E' facile verificare che se $f: A \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}^3$ è di classe C^2 , si ha

$$\text{div}(\text{rot} f) = 0.$$

Considerando, ad esempio, il campo vettoriale dell'esempio precedente, si ottiene

$$\text{div}(\text{rot} f) = \text{div}(x^2 - xy, 2yz - 2xy, yz - z^2) = (2x - y) + (2z - 2x) + (y - 2z) = 0.$$

Osservazione. Introduciamo in \mathbb{R}^3 il seguente operatore formale di derivazione, detto operatore *nabla* e denotato con il simbolo ∇ ,

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z} \right)$$

Allora,

- data una funzione reale $f: A \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}$ si ha

$$\nabla f(x, y, z) = \left(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y}, \frac{\partial f}{\partial z} \right) = \text{grad} f(x, y, z);$$

- dato un campo vettoriale $f: A \subseteq \mathbb{R}^3 \rightarrow \mathbb{R}^3$, $f = (f_1, f_2, f_3)$, si ha

$$\nabla \cdot f(x, y, z) = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z} = \operatorname{div} f(x, y, z)$$

$$\nabla \times f(x, y, z) = \det \begin{pmatrix} e_1 & e_2 & e_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_1 & f_2 & f_3 \end{pmatrix} = \operatorname{rot} f(x, y, z)$$

(dove, dati due vettori u e v , $u \cdot v$ denota il prodotto scalare di u con v mentre $u \times v$ ne denota il prodotto vettoriale).

Integrali doppi

Una *partizione* di un rettangolo $R = [a, b] \times [c, d] \subseteq \mathbb{R}^2$ è una coppia $P = (P_1, P_2)$ di partizioni degli intervalli $[a, b]$ e $[c, d]$, rispettivamente.

Date due partizioni $P_1 = \{x_0, x_1, \dots, x_{n_1}\}$ di $[a, b]$ e $P_2 = \{y_0, y_1, \dots, y_{n_2}\}$ di $[c, d]$, il rettangolo R viene suddiviso in $n_1 n_2$ sottorettangoli

$$R_{ij} = [x_{i-1}, x_i] \times [y_{j-1}, y_j], \quad i = 1, \dots, n_1, \quad j = 1, \dots, n_2,$$

di area $(\Delta x)_i (\Delta y)_j = (x_i - x_{i-1})(y_j - y_{j-1})$. In ogni sottorettangolo R_{ij} scegliamo un punto (ξ_{ij}, η_{ij}) . L'insieme S dei punti (ξ_{ij}, η_{ij}) si dice una *scelta di punti* nella partizione $P = (P_1, P_2)$ di R . La coppia $\alpha = (P, S)$, costituita dalla partizione P di R e dalla scelta S , si dice una *partizione puntata* di R . Il parametro di finezza di $\alpha = (P, S)$, denotato con $|\alpha|$, è la massima ampiezza dei lati di tutti i possibili rettangoli individuati dalla partizione P .

Sia ora assegnata una funzione $f: [a, b] \times [c, d] \rightarrow \mathbb{R}$. Ad ogni partizione puntata α di $R = [a, b] \times [c, d]$ possiamo associare il numero

$$\Sigma(\alpha) = \sum_{\substack{i=1, \dots, n_1 \\ j=1, \dots, n_2}} f(\xi_{ij}, \eta_{ij}) (\Delta x)_i (\Delta y)_j.$$

Si ha così una funzione reale $\Sigma: \mathcal{P} \rightarrow \mathbb{R}$ definita nell'insieme \mathcal{P} delle partizioni puntate del rettangolo R .

Per meglio comprendere il significato del numero $\Sigma(\alpha)$ individuato dalla partizione puntata α , è bene osservare che se la funzione f è positiva, ogni termine $f(\xi_{ij}, \eta_{ij}) (\Delta x)_i (\Delta y)_j$ della sommatoria rappresenta il volume di un parallelepipedo di altezza $f(\xi_{ij}, \eta_{ij})$ che ha per base il rettangolo R_{ij} . Quindi, se tutti

i rettangoli R_{ij} sono abbastanza piccoli, c'è da aspettarsi che il numero $\Sigma(\alpha)$ rappresenti una buona approssimazione del volume del solido

$$\left\{ (x, y, z) \in \mathbb{R}^3 : (x, y) \in R, 0 \leq z \leq f(x, y) \right\}$$

costituito dai punti che stanno sopra il rettangolo R e sotto il grafico $z = f(x, y)$ della funzione f .

Intuitivamente l'integrale doppio della funzione f nel rettangolo R è, quando esiste, il valore limite che si ottiene facendo tendere a zero i lati dei sottorettangoli individuati dalle possibili partizioni puntate di R . Diamo la definizione precisa.

Definizione (di integrale doppio in un rettangolo). Sia $f(x, y)$ una funzione reale definita (almeno) in un rettangolo R con i lati paralleli agli assi cartesiani. Si dice che un numero $l \in \mathbb{R}$ è l'*integrale doppio di f in R* se, fissato un arbitrario "errore" $\epsilon > 0$, esiste un $\delta > 0$ tale che, comunque si assegni una partizione puntata α con parametro di finezza $|\alpha|$ minore di δ , la distanza $|\Sigma(\alpha) - l|$ tra la somma $\Sigma(\alpha)$ e il numero l è minore di ϵ . Se ciò accade, si scrive

$$\lim_{|\alpha| \rightarrow 0} \Sigma(\alpha) = l$$

e la funzione f si dice *integrabile* in R (secondo Cauchy-Riemann).

Il numero l si chiama "integrale (doppio)" di f in R e si denota con uno dei seguenti simboli:

$$\iint_R f(x, y) \, dx dy, \quad \iint_R f \, dx dy, \quad \int_R f(x, y) \, dx dy.$$

Si osservi che il numero

$$l = \iint_R f(x, y) \, dx dy$$

non dipende dai simboli usati per indicare le variabili. Ad esempio al posto di x e y si possono usare le lettere u e v (il limite di $\Sigma(\alpha)$ per $|\alpha| \rightarrow 0$ non cambia).

Dalla precedente definizione segue facilmente che l'integrale doppio, quando esiste, è unico (unicità del limite). Inoltre, dalle note proprietà del limite si deduce che se f e g sono due funzioni integrabili in un rettangolo R ed λ e μ sono due numeri, allora anche la funzione $\lambda f + \mu g$ è integrabile e si ha

$$\iint_R (\lambda f + \mu g) \, dx dy = \lambda \iint_R f \, dx dy + \mu \iint_R g \, dx dy,$$

cioè l'integrale gode della **proprietà di linearità**. Più precisamente: l'integrale è un *funzionale lineare* sullo spazio vettoriale delle funzioni integrabili (nel rettangolo R).

Sempre dalla definizione di integrale si deduce che se f è integrabile in un rettangolo R e $f(x, y) \geq 0, \forall (x, y) \in R$, allora

$$\iint_R f \, dx \, dy \geq 0,$$

e da ciò segue (tenendo conto della linearità) la seguente proprietà dell'integrale doppio.

Proprietà di monotonia. *Siano f, g due funzioni integrabili in un rettangolo R . Se $f(x, y) \leq g(x, y), \forall (x, y) \in R$, allora*

$$\iint_R f \, dx \, dy \leq \iint_R g \, dx \, dy.$$

Un sottoinsieme di \mathbb{R}^2 si dice *trascurabile (in \mathbb{R}^2)* oppure di *misura (bidimensionale) nulla* se per ogni $\epsilon > 0$ esso può essere ricoperto con una famiglia (al più) numerabile di rettangoli di area totale minore di ϵ (nel senso che la somma, o la serie, delle aree dei rettangoli è minore di ϵ). Si potrebbe dimostrare che il grafico ($y = g(x)$ o $x = g(y)$) di una funzione continua in un intervallo è un insieme trascurabile di \mathbb{R}^2 . Inoltre l'unione di un numero finito (o, addirittura, di un'infinità numerabile) di insiemi trascurabili è ancora un insieme trascurabile. In particolare gli insiemi costituiti da un numero finito (o da un'infinità numerabile) di punti sono trascurabili.

Analogamente a quanto si è visto per gli integrali semplici, si ha il seguente

Teorema (di integrabilità). *Una funzione $f: [a, b] \times [c, d] \rightarrow \mathbb{R}$ è integrabile nel rettangolo $[a, b] \times [c, d]$ se e solo se è limitata e l'insieme dei suoi punti di discontinuità è trascurabile.*

Una prima conseguenza del teorema di integrabilità è che la somma, il prodotto e la composizione di funzioni integrabili è ancora integrabile (il quoziente potrebbe essere una funzione non limitata, e quindi non integrabile). Facciamo notare, inoltre, che se una funzione è continua in un rettangolo chiuso R , allora è anche integrabile (in tale rettangolo), essendo limitata (per il Teorema di Weierstrass) ed avendo un insieme vuoto (quindi trascurabile) di punti di discontinuità. Più in generale, se una funzione ha un numero finito (o un'infinità numerabile) di punti di discontinuità, allora, purché sia limitata, è integrabile (la limitatezza, questa volta, non è assicurata).

Teorema di equivalenza. *Siano f e g due funzioni integrabili in un rettangolo R . Se esse differiscono soltanto in un insieme trascurabile di punti di R , allora*

$$\iint_R f(x, y) \, dx \, dy = \iint_R g(x, y) \, dx \, dy.$$

Osservazione. Per integrare una funzione $f(x, y)$ in un rettangolo R non occorre che questa sia necessariamente definita in tutti i punti del rettangolo. Ad esempio, se è definita in tutto R tranne un numero finito di punti, può essere estesa assegnandole dei valori arbitrari in tali punti (per esempio il valore zero). In base al teorema di equivalenza, due differenti estensioni hanno lo stesso integrale.

In pratica tutte le funzioni che uno studente di ingegneria può incontrare nello svolgere gli esercizi hanno un insieme trascurabile di punti di discontinuità. Il motivo è dovuto al fatto che ogni “ragionevole funzione” è deducibile (con un numero finito di operazioni di somma, prodotto, quoziente, composizione, restrizione ad un intervallo e inversione) dalle cosiddette funzioni *fondamentali* (vedi appunti sugli integrali di funzioni di una variabile), ciascuna delle quali ha un insieme trascurabile di punti di discontinuità (spesso è addirittura continua). Stabilire quindi se una funzione è integrabile (in un rettangolo) o se non lo è si riduce a controllare se (in tale rettangolo) è limitata oppure no.

Esempio. La funzione

$$f(x, y) = \frac{1}{x^2 + y^2}$$

è integrabile in un rettangolo (chiuso) R se e solo se R non contiene il punto $(0, 0)$. Infatti, se R non contiene l’origine, allora, essendo continua in tutti punti del suo dominio $\mathbb{R}^2 \setminus \{(0, 0)\}$, è continua anche in R ed è quindi integrabile in un tale R . Se invece R contiene l’origine, allora la funzione non può essere limitata in tale rettangolo, dato che $f(x, y) \rightarrow +\infty$ per $(x, y) \rightarrow (0, 0)$. Si fa notare che in questo caso la non integrabilità non dipende dal fatto che non è definita in $(0, 0)$: può essere estesa assegnandole un valore qualunque nell’origine, ma ogni estensione non potrà renderla limitata (casomai la renderà discontinua in un punto, ma un punto costituisce un insieme trascurabile).

Una risposta al problema del calcolo di un integrale doppio è data dal seguente risultato che riconduce il calcolo di un integrale doppio a due successive integrazioni semplici:

Teorema di Fubini (per gli integrali doppi). *Sia f una funzione reale delle variabili (x, y) definita in un rettangolo $R = [a, b] \times [c, d]$. Allora, “quando ha senso”, risulta*

$$\iint_R f(x, y) \, dx dy = \int_a^b \left(\int_c^d f(x, y) \, dy \right) dx$$
$$\left[\iint_R f(x, y) \, dx dy = \int_c^d \left(\int_a^b f(x, y) \, dx \right) dy \right].$$

In sostanza, il Teorema di Fubini afferma che per calcolare l'integrale doppio di $f(x, y)$ in $[a, b] \times [c, d]$ è possibile integrare prima in $[c, d]$ la funzione $f(x, y)$ rispetto alla variabile y , ottenendo così una funzione

$$g(x) = \int_c^d f(x, y) dy,$$

ed integrare poi $x \mapsto g(x)$ nell'intervallo $[a, b]$.

Con l'affermazione “quando ha senso” vogliamo significare quando siano soddisfatte ipotesi che garantiscano che gli integrali in questione siano ben definiti. Tanto per fare un esempio, se f è continua nel rettangolo $[a, b] \times [c, d]$, gli integrali hanno senso (ma ci sono ovviamente situazioni molto più generali).

Esempio. Sia $R = [0, 1] \times [2, 3]$. Calcoliamo

$$\iint_R x^2 y \, dx dy.$$

Applicando il teorema di Fubini si ha

$$\iint_A x^2 y \, dx dy = \int_0^1 \left(\int_2^3 x^2 y \, dy \right) dx = \int_0^1 \frac{x^2}{2} (9 - 4) dx = \frac{5}{2} \int_0^1 x^2 dx = \frac{5}{6}.$$

Vogliamo ora estendere la nozione di integrale doppio ad insiemi che non siano necessariamente rettangoli. Dato un insieme A di \mathbb{R}^2 e data $f: A \rightarrow \mathbb{R}$, la funzione $\hat{f}: \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da

$$\hat{f}(x, y) = \begin{cases} f(x, y) & \text{se } (x, y) \in A \\ 0 & \text{se } (x, y) \notin A \end{cases}$$

si chiama *estensione standard* di f (relativa ad A).

Sia $(x, y) \mapsto f(x, y)$ una funzione di due variabili definita in un sottoinsieme **limitato** A di \mathbb{R}^2 . Consideriamo un (arbitrario) rettangolo R contenente A . Diremo che f è *integrabile in A* se è integrabile in R la sua estensione standard \hat{f} . In tal caso l'integrale di f in A si definisce nel modo seguente:

$$\iint_A f(x, y) \, dx dy := \iint_R \hat{f}(x, y) \, dx dy.$$

Dal fatto che \hat{f} è nulla fuori da A si potrebbe dedurre che il secondo integrale non dipende dal rettangolo R contenente A . Pertanto, la definizione è ben posta.

Definizione. Un sottoinsieme limitato A di \mathbb{R}^2 si dice *misurabile* (secondo Peano-Jordan) quando è integrabile in A la funzione $f(x, y) \equiv 1$. In tal caso la *misura* (bidimensionale) di A , detta anche *area*, è il numero

$$\mu(A) = \iint_A dx dy.$$

Osservazione. Si può provare che un sottoinsieme limitato A è misurabile se e solo se la sua frontiera è trascurabile. Ad esempio, è misurabile ogni insieme limitato la cui frontiera è unione finita di grafici ($y = \varphi(x)$ o $x = \psi(y)$) di funzioni continue.

Elenchiamo alcune proprietà dell'integrale doppio che discendono facilmente dalla definizione.

Proprietà di linearità. Siano $f, g: A \rightarrow \mathbb{R}$ due funzioni integrabili in un insieme limitato $A \subseteq \mathbb{R}^2$ e siano $\lambda, \mu \in \mathbb{R}$. Allora risulta

$$\iint_A (\lambda f(x, y) + \mu g(x, y)) dx dy = \lambda \iint_A f(x, y) dx dy + \mu \iint_A g(x, y) dx dy.$$

Proprietà di monotonia. Siano $f, g: A \rightarrow \mathbb{R}$ due funzioni integrabili in un insieme limitato $A \subseteq \mathbb{R}^2$ e supponiamo $f(x, y) \leq g(x, y)$ per ogni $(x, y) \in A$. Allora risulta

$$\iint_A f(x, y) dx dy \leq \iint_A g(x, y) dx dy.$$

Proprietà di additività (rispetto all'insieme di integrazione). Supponiamo che f sia una funzione integrabile sia in un insieme A che in un insieme B , con $A \cap B = \emptyset$ (o, più in generale, $A \cap B$ trascurabile). Allora f è integrabile in $A \cup B$ e si ha

$$\iint_{A \cup B} f(x, y) dx dy = \iint_A f(x, y) dx dy + \iint_B f(x, y) dx dy.$$

Sia $A \subseteq \mathbb{R}^2$ un insieme del tipo

$$A = \{(x, y) : a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x)\},$$

dove $\varphi_1, \varphi_2: [a, b] \rightarrow \mathbb{R}$ sono due funzioni continue. Si dice che A è *semplice* (o *normale*) rispetto all'asse delle y . Infatti ogni retta parallela a tale asse lo interseca in un intervallo (di estremi $\varphi_1(x)$ e $\varphi_2(x)$, per $x \in [a, b]$).

Supponiamo che f sia una funzione integrabile in A . Dato un rettangolo R contenente A , per definizione, l'integrale di f in A è

$$\iint_R \hat{f}(x, y) \, dx dy.$$

Per quanto osservato in precedenza, tale integrale non dipende dalla scelta del rettangolo R che contiene A . Per semplicità prendiamo allora $R = [a, b] \times [c, d]$ con $c \leq \min\{\varphi_1(x) : x \in [a, b]\}$ e $d \geq \max\{\varphi_2(x) : x \in [a, b]\}$. Dal Teorema di Fubini si ha

$$\iint_R \hat{f}(x, y) \, dx dy = \int_a^b \int_c^d \hat{f}(x, y) \, dy \, dx.$$

D'altra parte, per l'additività degli integrali di funzioni di una variabile rispetto agli intervalli,

$$\int_c^d \hat{f}(x, y) \, dy = \int_c^{\varphi_1(x)} \hat{f}(x, y) \, dy + \int_{\varphi_1(x)}^{\varphi_2(x)} \hat{f}(x, y) \, dy + \int_{\varphi_2(x)}^d \hat{f}(x, y) \, dy,$$

e, tenendo conto che \hat{f} è nulla fuori da A , si ottiene

$$\int_c^d \hat{f}(x, y) \, dy = \int_{\varphi_1(x)}^{\varphi_2(x)} \hat{f}(x, y) \, dy.$$

Poiché in A le due funzioni f ed \hat{f} coincidono, si ha

$$\int_c^d \hat{f}(x, y) \, dy = \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) \, dy.$$

Si ottiene così la seguente importante **formula di riduzione**, valida per gli insiemi che sono semplici rispetto all'asse delle y :

$$\iint_A f(x, y) \, dx dy = \int_a^b \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) \, dy \, dx.$$

Analogamente, se $A \subseteq \mathbb{R}^2$ è un insieme del tipo

$$A = \{(x, y) : c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y)\},$$

dove $\psi_1, \psi_2: [c, d] \rightarrow \mathbb{R}$ sono due funzioni continue, ed f è integrabile in A , si ha l'altra **formula di riduzione**, valida quando A è semplice rispetto all'asse delle x :

$$\iint_A f(x, y) \, dx dy = \int_c^d \int_{\psi_1(y)}^{\psi_2(y)} f(x, y) \, dx \, dy.$$

Esempio. Usando le formule di riduzione, calcolare

$$\iint_A x^2 y \, dx dy \quad \text{dove } A = \{(x, y) \in \mathbb{R}^2 : x^2 \leq y \leq 2x\}.$$

Applichiamo la formula di riduzione con $\varphi_1(x) = x^2$ e $\varphi_2(x) = 2x$, osservando inoltre che $\varphi_1(x) \leq \varphi_2(x)$ per $0 \leq x \leq 2$. Si ha

$$\iint_A x^2 y \, dx dy = \int_0^2 \left(\int_{x^2}^{2x} x^2 y \, dy \right) dx = \int_0^2 \frac{x^2}{2} (4x^2 - x^4) dx.$$

L'ultimo termine è l'integrale definito di una funzione di una variabile e il suo calcolo è lasciato per esercizio.

Esempio. Usando le formule di riduzione, calcolare

$$\iint_A (x - y) \, dx dy \quad \text{dove } A = \{(x, y) \in \mathbb{R}^2 : x^2 \leq y, x^2 + y^2 \leq 1\}.$$

Si ha

$$\begin{aligned} \iint_A (x - y) \, dx dy &= \int_a^b \left(\int_{x^2}^{\sqrt{1-x^2}} (x - y) \, dy \right) dx \\ &= \int_a^b \left(x\sqrt{1-x^2} - x^3 - (1/2)(1 - x^2 - x^4) \right) dx, \end{aligned}$$

dove $a = -\sqrt{(-1 + \sqrt{5})/2}$ e $b = \sqrt{(-1 + \sqrt{5})/2}$.

Esercizio. Usando le formule di riduzione, calcolare

$$\iint_A xy \, dx dy \quad \text{nei due casi } A = \{(x, y) \in \mathbb{R}^2 : y \geq x^2, y \leq \sqrt{x}\}$$

$$\text{e } A = \{(x, y) \in \mathbb{R}^2 : x + y \geq 1, y \geq x^2, y \leq \sqrt{x}\}.$$

Un altro importante risultato spesso utilizzato nel calcolo degli integrali doppi è il seguente

Teorema (di cambiamento di variabile per gli integrali doppi). *Sia $\varphi(u, v) = (\varphi_1(u, v), \varphi_2(u, v))$ un'applicazione continua da un insieme chiuso e limitato $D \subseteq \mathbb{R}^2$ in \mathbb{R}^2 . Supponiamo che le frontiere di D e di $\varphi(D)$ siano trascurabili e che φ sia C^1 e iniettiva nell'interno di D . Allora, data una funzione di due variabili f continua su $\varphi(D)$, risulta*

$$\iint_{\varphi(D)} f(x, y) \, dx dy = \iint_D f(\varphi_1(u, v), \varphi_2(u, v)) |\det J\varphi(u, v)| \, dudv.$$

Esempio. Facendo uso delle coordinate polari $x = \rho \cos \theta$, $y = \rho \sin \theta$, e del teorema di cambiamento di variabile per gli integrali doppi, calcoliamo

$$\iint_A \frac{y e^{\sqrt{x^2+y^2}}}{x^2+y^2} dx dy$$

ove $A = \{(x, y) \in \mathbb{R}^2 : 0 < r^2 \leq x^2 + y^2 \leq R^2, x \geq 0, y \geq 0\}$.

Sia $\varphi: \mathbb{R}^2 \rightarrow \mathbb{R}^2$ l'applicazione definita da

$$\varphi(\rho, \theta) = (\rho \cos \theta, \rho \sin \theta).$$

I punti di A si ottengono tramite φ , facendo variare la coppia di numeri (ρ, θ) nel rettangolo chiuso e limitato $D = [r, R] \times [0, \pi/2]$ del piano $\rho\theta$. In altre parole, l'immagine $\varphi(D)$ del rettangolo D tramite φ coincide col dominio d'integrazione A . Ricordando che lo jacobiano di φ (in un generico punto (ρ, θ)) coincide con ρ , dalla formula di cambiamento di variabile per gli integrali doppi si ha

$$\iint_D \frac{\rho \sin \theta e^\rho}{\rho^2} \rho d\rho d\theta = \int_0^{\pi/2} \int_r^R \sin \theta e^\rho d\rho d\theta = e^R - e^r.$$

Si osservi che le ipotesi del teorema di cambiamento di variabili sono soddisfatte. Infatti D è chiuso e limitato, ∂D e $\partial A = \partial\varphi(D)$ sono insiemi trascurabili, φ è continua in D , è C^1 nell'interno di D (è addirittura C^∞) ed è iniettiva nell'interno di D (in questo caso lo è anche su ∂D). Notiamo che se consideriamo invece la corona circolare $A = \{(x, y) \in \mathbb{R}^2 : 0 < r^2 \leq x^2 + y^2 \leq R^2\}$, allora si ottiene $D = [r, R] \times [0, 2\pi]$ e φ è iniettiva solo nell'interno di D e non lo è su ∂D in quanto per $\theta = 0$ e $\theta = 2\pi$ si ottengono gli stessi punti di A .

8^a settimana - dal 29.04.19

Esempio. (facoltativo) Calcoliamo

$$\iint_A \sqrt{x^2 + y^2} \, dx dy$$

ove $A = \{(x, y) \in \mathbb{R}^2 : (x - 1)^2 + y^2 \leq 1, x^2 + (y - 1)^2 \leq 1\}$. Usando coordinate polari, l'insieme A è immagine dell'insieme

$$D = \{(\rho, \theta) \in \mathbb{R}^2 : 0 \leq \rho \leq 2 \operatorname{sen} \theta, 0 \leq \theta \leq \pi/4\}$$

$$\cup \{(\rho, \theta) \in \mathbb{R}^2 : 0 \leq \rho \leq 2 \cos \theta, \pi/4 \leq \theta \leq \pi/2\}.$$

Si ha

$$\begin{aligned} \iint_A \sqrt{x^2 + y^2} \, dx dy &= \iint_D \rho \, d\rho d\theta = \\ &= \int_0^{\pi/4} \int_0^{2 \operatorname{sen} \theta} \rho^2 \, d\rho d\theta + \int_{\pi/4}^{\pi/2} \int_0^{2 \cos \theta} \rho^2 \, d\rho d\theta = \frac{8}{3} \left(\int_0^{\pi/4} \operatorname{sen}^3 \theta \, d\theta + \int_{\pi/4}^{\pi/2} \cos^3 \theta \, d\theta \right) \end{aligned}$$

Esempio. Tenendo conto della proprietà di additività dell'integrale, calcoliamo

$$\iint_A (x^2 + y^2) \, dx dy$$

ove

$$A = \{(x, y) \in \mathbb{R}^2 : (x - \frac{1}{2})^2 + y^2 \geq \frac{1}{4}, x^2 + y^2 \leq 1\}.$$

Si ha

$$\iint_A (x^2 + y^2) \, dx dy = \iint_{A_1} (x^2 + y^2) \, dx dy - \iint_{A_2} (x^2 + y^2) \, dx dy$$

essendo

$$A_1 = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1\}$$

e

$$A_2 = \{(x, y) \in \mathbb{R}^2 : (x - \frac{1}{2})^2 + y^2 \leq \frac{1}{4}\}.$$

Si ha

$$\iint_{A_1} (x^2 + y^2) \, dx dy = 2\pi \int_0^1 \rho^3 \, d\rho = \frac{\pi}{2}$$

e, considerando il cambiamento di variabile $x = \frac{1}{2} + \rho \cos \theta$, $y = \rho \sin \theta$ con $0 \leq \theta \leq 2\pi$ e $0 \leq \rho \leq 1/2$, si ottiene

$$\iint_{A_2} (x^2 + y^2) dx dy = \int_0^{2\pi} \int_0^{\frac{1}{2}} \left(\left(\frac{1}{2} + \rho \cos \theta \right)^2 + \rho^2 \sin^2 \theta \right) \rho d\rho d\theta = \frac{3}{32} \pi.$$

Pertanto

$$\iint_A (x^2 + y^2) dx dy = \frac{\pi}{2} - \frac{3\pi}{32} = \frac{13}{32} \pi.$$

Esempio. Calcoliamo

$$\iint_A \sqrt{1 - \frac{x^2}{a^2} - \frac{y^2}{b^2}} dx dy$$

ove

$$A = \{(x, y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1\}.$$

Introducendo le coordinate ellittiche

$$x = a\rho \cos \theta, \quad y = b\rho \sin \theta, \quad 0 \leq \rho \leq 1, \quad 0 \leq \theta \leq 2\pi,$$

e osservando che il determinante della matrice jacobiana è $ab\rho$, si ottiene

$$2\pi \int_0^1 \sqrt{1 - \rho^2} ab\rho d\rho = \frac{2}{3} \pi a b.$$

Dato un insieme di misura non nulla $A \subseteq \mathbb{R}^2$, il suo *centro di massa (geometrico)* è il punto (x_c, y_c) che ha per ascissa la media delle ascisse e per ordinata le media delle ordinate. Si ha pertanto

$$x_c = \frac{1}{\mu(A)} \iint_A x dx dy, \quad y_c = \frac{1}{\mu(A)} \iint_A y dx dy,$$

dove ricordiamo che con $\mu(A)$ abbiamo indicato l'area di A .

Esempio. Determiniamo il centro di massa (geometrico) del semicerchio

$$A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2, y \geq 0\}$$

Per ragioni di simmetria risulta $x_c = 0$. Occorre quindi calcolare soltanto l'ordinata y_c . L'area $\mu(A)$ del semicerchio è $\pi r^2/2$ e quindi, usando le formule di riduzione, si ottiene

$$y_c = \frac{2}{\pi r^2} \iint_A y dx dy = \frac{2}{\pi r^2} \int_{-r}^r dx \int_0^{\sqrt{r^2 - x^2}} y dy = \frac{1}{\pi r^2} \int_{-r}^r (r^2 - x^2) dx = \frac{4r}{3\pi}.$$

Si osservi che $4r/3\pi$ è un numero tra 0 ed r ; anzi, è addirittura minore di $r/2$.

Più in generale, determiniamo il centro di massa del settore circolare di raggio r e di angolo al centro 2α

$$S = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2, x > 0, |y/x| \leq \tan \alpha\}.$$

L'area di S è αr^2 e, per motivi di simmetria, risulta ovviamente $y_c = 0$. Usando il teorema di cambiamento di variabile e le coordinate polari, si ottiene

$$x_c = \frac{1}{\alpha r^2} \int_{-\alpha}^{\alpha} \int_0^r \rho \cos \theta \rho d\rho d\theta = \frac{2}{3} r \frac{\sin \alpha}{\alpha}.$$

Osserviamo che facendo tendere α a 0 si ottiene $x_c \rightarrow \frac{2}{3}r$ che ovviamente non è il centro di massa di una sbarra omogenea.

Esercizio. Determinare il centro di massa (geometrico) del cerchio forato

$$A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 16, (x - 1)^2 + y^2 \geq 1\}.$$

Suggerimento. Usare la proprietà di additività dell'integrale rispetto all'insieme di integrazione.

La massa m di una piastra $A \subseteq \mathbb{R}^2$ (non necessariamente omogenea) di densità (superficiale) $\delta(x, y)$ è data da

$$m = \iint_A \delta(x, y) dx dy.$$

Talvolta il prodotto della densità (superficiale) $\delta(x, y)$ per l'elemento di area $d\sigma = dx dy$ si denota col simbolo dm , detto *elemento di massa*. Perciò, si usa anche scrivere

$$m = \iint_A dm.$$

Esempio. Calcoliamo la massa di un disco di raggio r sapendo che la densità superficiale è proporzionale alla distanza dal centro del disco. Si può supporre che il disco considerato sia l'insieme ove $A = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2\}$. La densità superficiale risulta allora $\delta(x, y) = k\sqrt{x^2 + y^2}$ dove k è una costante positiva. Si ha perciò

$$m = \iint_A k\sqrt{x^2 + y^2} dx dy.$$

Data la simmetria circolare della funzione integranda $\sqrt{x^2 + y^2}$ e del dominio di integrazione A , è conveniente individuare i punti di A mediante le coordinate

polari ed esprimere la funzione integranda in tali coordinate. Dalla formula di cambiamento di variabile per gli integrali doppi si ha

$$m = 2\pi \int_0^r k\rho \rho d\rho = \frac{2}{3}\pi k r^3$$

Se un sottoinsieme (limitato) $A \subseteq \mathbb{R}^2$ rappresenta una piastra (non necessariamente omogenea) di densità superficiale $\delta(x, y)$, le coordinate del centro di massa sono date da

$$x_c = \frac{1}{m} \iint_A x \delta(x, y) dx dy, \quad y_c = \frac{1}{m} \iint_A y \delta(x, y) dx dy,$$

dove m è la massa della piastra.

Sia $A \subseteq \mathbb{R}^2$ una piastra (non necessariamente omogenea) di massa m . Fissato un punto $P_0 = (x_0, y_0) \in \mathbb{R}^2$, il *momento d'inerzia* di A rispetto a P_0 (o, equivalentemente, rispetto ad una retta passante per P_0 e perpendicolare al piano) è il numero

$$I = \iint_A d(P, P_0)^2 \delta(x, y) dx dy,$$

dove $d(P, P_0) = \sqrt{(x - x_0)^2 + (y - y_0)^2}$ denota la funzione distanza di un generico punto $P = (x, y) \in A$ da $P_0 = (x_0, y_0)$ e $\delta(x, y)$ rappresenta la densità (superficiale). Ricordando che il prodotto della densità (superficiale) $\delta(x, y)$ per l'elemento di area $d\sigma = dx dy$ si indica anche con l'elemento di massa dm , si può anche scrivere

$$I = \iint_A d(P, P_0)^2 dm.$$

Se la piastra è omogenea, la densità δ è costante ed è data da $\delta = m/\mu(A)$.

Come esempio, calcoliamo il momento d'inerzia di un disco omogeneo di massa m e raggio r rispetto al centro. Denotiamo con A il disco e poniamolo, per semplicità, nel piano xy col centro nell'origine degli assi. Poiché il disco è omogeneo, la sua densità superficiale è $\delta = m/\pi r^2$. Occorre calcolare

$$I = \iint_A (x^2 + y^2) \delta dx dy.$$

Data la simmetria circolare della funzione integranda $x^2 + y^2$ e del dominio di integrazione A , è conveniente individuare i punti di A mediante le coordinate polari ed esprimere la funzione integranda in tali coordinate. Dalla formula di cambiamento di variabile per gli integrali doppi si ha

$$I = \iint_A (x^2 + y^2) \delta dx dy = 2\pi\delta \int_0^r \rho^3 d\rho = \frac{1}{2}\pi r^4 \delta = \frac{1}{2} m r^2.$$

Teorema di Pappo-Guldino (per i solidi di rotazione). *Sia A un insieme misurabile e limitato contenuto in un semipiano delimitato da una retta α . Il volume del solido che si ottiene ruotando A di un angolo 2π intorno alla retta α è dato dal prodotto dell'area di A per la lunghezza della circonferenza percorsa dal centro di massa di A .*

Come applicazione del Teorema di Pappo calcoliamo il volume di una sfera di raggio r . Tale sfera si può ottenere (ad esempio) facendo fare un giro completo intorno all'asse x al semicerchio A di raggio r di cui abbiamo appena determinato il centro di massa. Poiché la distanza y_c del centro di massa di A dall'asse delle x è $4r/3\pi$, tale punto, nella rotazione, descrive una circonferenza di lunghezza $(2\pi)(4r/3\pi) = 8r/3$. Dato che l'area di A è $\pi r^2/2$, il volume della sfera risulta

$$\frac{\pi r^2}{2} \frac{8r}{3} = \frac{4}{3}\pi r^3.$$

Esempio (importante). Nel capitolo degli integrali impropri, abbiamo affermato che

$$\int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

Mediante l'uso degli integrali doppi è possibile provarlo. A questo scopo, sia $r > 0$ e consideriamo il rettangolo $R = [0, r] \times [0, r]$. Denotiamo inoltre

$$A_r = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2, x \geq 0, y \geq 0\}$$

e

$$A_{\sqrt{2}r} = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 2r^2, x \geq 0, y \geq 0\}.$$

Si ha ovviamente

$$A_r \subset R \subset A_{\sqrt{2}r}$$

da cui

$$\iint_{A_r} e^{-(x^2+y^2)} dx dy \leq \iint_R e^{-(x^2+y^2)} dx dy \leq \iint_{A_{\sqrt{2}r}} e^{-(x^2+y^2)} dx dy.$$

Passando in coordinate polari, risulta

$$\iint_{A_r} e^{-(x^2+y^2)} dx dy = \int_0^{\frac{\pi}{2}} \int_0^r e^{-\rho^2} \rho d\rho d\theta = \frac{\pi}{4}(-e^{-r^2} + 1)$$

e, in maniera analoga,

$$\iint_{A_{\sqrt{2}r}} e^{-(x^2+y^2)} dx dy = \frac{\pi}{4}(-e^{-2r^2} + 1).$$

D'altra parte, per il teorema di Fubini, si ha

$$\iint_R e^{-(x^2+y^2)} dx dy = \left(\int_0^r e^{-x^2} dx \right) \left(\int_0^r e^{-y^2} dy \right) = \left(\int_0^r e^{-x^2} dx \right)^2$$

e, quindi,

$$\iint_{A_r} e^{-(x^2+y^2)} dx dy \leq \left(\int_0^r e^{-x^2} dx \right)^2 \leq \iint_{A_{\sqrt{2}r}} e^{-(x^2+y^2)} dx dy,$$

da cui, per il calcolo precedente,

$$\sqrt{\frac{\pi}{4}(-e^{-r^2} + 1)} \leq \int_0^r e^{-x^2} dx \leq \sqrt{\frac{\pi}{4}(-e^{-2r^2} + 1)}.$$

Passando al limite per $r \rightarrow +\infty$ e ricordando il teorema dei carabinieri, si ottiene

$$\lim_{r \rightarrow +\infty} \int_0^r e^{-x^2} dx = \int_0^{+\infty} e^{-x^2} dx = \frac{\sqrt{\pi}}{2}.$$

Integrali tripli

Una *partizione* di un parallelepipedo

$$Q = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3] \subseteq \mathbb{R}^3$$

è una terna $P = (P_1, P_2, P_3)$ di partizioni degli intervalli $[a_1, b_1]$, $[a_2, b_2]$ e $[a_3, b_3]$, rispettivamente.

Date tre partizioni, $P_1 = \{x_0, x_1, \dots, x_{n_1}\}$ di $[a_1, b_1]$, $P_2 = \{y_0, y_1, \dots, y_{n_2}\}$ di $[a_2, b_2]$ e $P_3 = \{z_0, z_1, \dots, z_{n_3}\}$ di $[a_3, b_3]$, il parallelepipedo Q viene suddiviso in $n_1 n_2 n_3$ sottoparallelepipedo

$$Q_{ijk} = [x_{i-1}, x_i] \times [y_{j-1}, y_j] \times [z_{k-1}, z_k]$$

di volume $(\Delta x)_i (\Delta y)_j (\Delta z)_k = (x_i - x_{i-1})(y_j - y_{j-1})(z_k - z_{k-1})$. In ogni parallelepipedo Q_{ijk} scegliamo un punto $(\xi_{ijk}, \eta_{ijk}, \zeta_{ijk})$. L'insieme S dei punti $(\xi_{ijk}, \eta_{ijk}, \zeta_{ijk})$ si dice una *scelta di punti* nella partizione $P = (P_1, P_2, P_3)$ di Q . Ogni parallelepipedo Q_{ijk} della partizione con il punto $(\xi_{ijk}, \eta_{ijk}, \zeta_{ijk})$ scelto si dice un *parallelepipedo puntato*. La coppia $\alpha = (P, S)$, costituita dalla partizione $P = (P_1, P_2, P_3)$ di Q e dalla scelta S , si dice una *partizione puntata* di Q . Il parametro di finezza di $\alpha = (P, S)$, denotato con $|\alpha|$, è la massima ampiezza dei lati di tutti i possibili parallelepipedo individuati dalla partizione P .

Sia ora $f: Q \rightarrow \mathbb{R}$ una funzione definita in Q . Ad ogni partizione puntata $\alpha = (P, S)$ di Q possiamo associare il numero

$$\Sigma(\alpha) = \sum_{(i,j,k) \in \mathcal{K}} f(\xi_{ijk}, \eta_{ijk}, \zeta_{ijk})(\Delta x)_i (\Delta y)_j (\Delta z)_k,$$

dove la terna di indici (i, j, k) varia nell'insieme

$$\mathcal{K} = \{(i, j, k) \in \mathbb{N}^3 : 1 \leq i \leq n_1, 1 \leq j \leq n_2, 1 \leq k \leq n_3\}.$$

Intuitivamente l'integrale triplo in Q della funzione f è, quando esiste, il valore limite che si ottiene facendo tendere a zero i lati dei sottoparallelepipedi individuati dalle possibili partizioni puntate di Q . Diremo infatti che il numero l è l'integrale triplo di f in Q se, fissato un "errore" $\epsilon > 0$, esiste un $\delta > 0$ tale che, comunque si assegni una partizione puntata α con parametro di finezza $|\alpha|$ minore di δ , la somma $\Sigma(\alpha)$ sopra definita dista da l meno di ϵ . Se ciò accade, si scrive

$$\lim_{|\alpha| \rightarrow 0} \Sigma(\alpha) = l$$

e la funzione f si dice *integrabile* (in Q) (secondo Cauchy-Riemann). Il numero l si chiama "integrale (triplo)" di f in Q e si denota con uno dei seguenti simboli:

$$\iiint_Q f(x, y, z) dx dy dz, \quad \iiint_Q f dx dy dz, \quad \int_Q f(x, y, z) dx dy dz.$$

È ovvio che il numero

$$l = \iiint_Q f(x, y, z) dx dy dz$$

non dipende dai simboli usati per indicare le variabili. Ad esempio al posto di x, y e z si possono usare le lettere u, v e w (il limite di $\Sigma(\alpha)$ per $|\alpha| \rightarrow 0$ non cambia).

Un sottoinsieme di \mathbb{R}^3 si dice *trascurabile* (in \mathbb{R}^3) oppure di *misura (tridimensionale) nulla* se per ogni $\epsilon > 0$ può essere ricoperto con famiglia (al più) numerabile di parallelepipedi di volume totale minore o uguale ad ϵ . Si potrebbe dimostrare che il grafico di una funzione continua (reale di due variabili reali) è un insieme trascurabile di \mathbb{R}^3 . Inoltre l'unione di un numero finito (o, addirittura, di un'infinità numerabile) di insiemi trascurabili è ancora un insieme trascurabile. In particolare gli insiemi costituiti da un numero finito (o da un'infinità numerabile) di punti sono trascurabili.

Analogamente a quanto si è visto per gli integrali doppi, **una funzione $f(x, y, z)$ è integrabile in un parallelepipedo Q se e solo se è limitata e l'insieme dei suoi punti di discontinuità è trascurabile.**

Dalla definizione segue facilmente che l'integrale, quando esiste, è unico (unicità del limite). Inoltre, dalle note proprietà del limite si deduce che se f e g sono due funzioni integrabili in un parallelepipedo Q ed α e β sono due numeri, allora anche la funzione $\alpha f + \beta g$ è integrabile e si ha

$$\iiint_Q (\alpha f + \beta g) dx dy dz = \alpha \iiint_Q f dx dy dz + \beta \iiint_Q g dx dy dz,$$

cioè l'integrale gode della **proprietà di linearità**. Più precisamente: l'integrale è un *funzionale lineare* sullo spazio vettoriale delle funzioni integrabili (nel parallelepipedo Q).

Sempre dalla definizione di integrale si deduce che se f è integrabile in Q e $f(x, y, z) \geq 0, \forall (x, y, z) \in Q$, allora

$$\iiint_Q f dx dy dz \geq 0,$$

e da ciò segue (tenendo conto della linearità) la seguente proprietà dell'integrale triplo:

Proprietà di monotonìa. *Siano f, g due funzioni integrabili in un parallelepipedo Q . Se $f(x, y, z) \leq g(x, y, z), \forall (x, y, z) \in Q$, allora*

$$\iiint_Q f dx dy dz \leq \iiint_Q g dx dy dz.$$

Il seguente risultato riconduce il calcolo di un integrale triplo a due successive integrazioni: una semplice seguita da una doppia, o una doppia seguita da una semplice.

Teorema di Fubini (per gli integrali tripli). *Sia f una funzione reale delle variabili (x, y, z) definita in un parallelepipedo $Q = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$. Allora, "quando ha senso", risulta*

$$\iiint_Q f(x, y, z) dx dy dz = \iint_R \left(\int_{a_3}^{b_3} f(x, y, z) dz \right) dx dy,$$

$$\iiint_Q f(x, y, z) dx dy dz = \int_{a_3}^{b_3} \left(\iint_R f(x, y, z) dx dy \right) dz,$$

dove R denota il rettangolo $[a_1, b_1] \times [a_2, b_2]$ nel piano xy .

La prima formula del Teorema di Fubini afferma che per calcolare l'integrale triplo di $f(x, y, z)$ in Q è possibile integrare prima in $[a_3, b_3]$ la funzione $f(x, y, z)$

rispetto alla variabile z , ottenendo così una funzione

$$g(x, y) = \int_{a_3}^{b_3} f(x, y, z) dz,$$

ed integrare poi $g(x, y)$ nel rettangolo $[a_1, b_1] \times [a_2, b_2]$.

La seconda formula afferma che si ottiene lo stesso risultato facendo prima l'integrale doppio in $R = [a_1, b_1] \times [a_2, b_2]$ della funzione $f(x, y, z)$ rispetto alle variabili x ed y , ottenendo così una funzione

$$h(z) = \iint_R f(x, y, z) dx dy,$$

ed integrando poi $h(z)$ nell'intervallo $[a_3, b_3]$.

Ovviamente, nel Teorema di Fubini in \mathbb{R}^3 i ruoli delle variabili x , y e z possono essere permutati.

Per l'affermazione “quando ha senso” valgono osservazioni analoghe a quelle fatte in precedenza per gli integrali doppi.

Dato un insieme A di \mathbb{R}^3 e data $f: A \rightarrow \mathbb{R}$, la funzione $\hat{f}: \mathbb{R}^3 \rightarrow \mathbb{R}$ definita da

$$\hat{f}(x, y, z) = \begin{cases} f(x, y, z) & \text{se } (x, y, z) \in A \\ 0 & \text{se } (x, y, z) \notin A \end{cases}$$

si chiama *estensione standard* di f (relativa ad A).

Sia f una funzione di tre variabili definita in un sottoinsieme limitato A di \mathbb{R}^3 . Consideriamo un (arbitrario) parallelepipedo Q contenente A . Diremo che f è *integrabile in A* se è integrabile in Q la sua estensione standard \hat{f} . In tal caso l'integrale di f in A si definisce nel modo seguente:

$$\iiint_A f(x, y, z) dx dy dz := \iiint_Q \hat{f}(x, y, z) dx dy dz.$$

Dal fatto che \hat{f} è nulla fuori da A si può dedurre che il secondo integrale non dipende dal parallelepipedo Q contenente A . Pertanto, la definizione è ben posta.

Definizione. Un sottoinsieme limitato A di \mathbb{R}^3 si dice *misurabile (secondo Peano-Jordan)* quando è integrabile in A la funzione $f(x, y, z) \equiv 1$. In tal caso, la *misura (tridimensionale)* di A , detta anche *volume*, è il numero

$$\mu_3(A) = \iiint_A dx dy dz.$$

Teorema (di additività rispetto all'insieme di integrazione). *Supponiamo che una funzione $f(x, y, z)$ sia integrabile sia in un insieme A che in un insieme B , con $A \cap B = \emptyset$ (o, più in generale, trascurabile). Allora f è integrabile in $A \cup B$ e*

$$\iiint_{A \cup B} f(x, y, z) \, dx \, dy \, dz = \iiint_A f(x, y, z) \, dx \, dy \, dz + \iiint_B f(x, y, z) \, dx \, dy \, dz.$$

Diamo ora un'applicazione del Teorema di Fubini al calcolo di integrali tripli in insiemi che hanno una forma particolare.

Sia $A \subseteq \mathbb{R}^3$ un insieme del tipo

$$A = \{(x, y, z) : (x, y) \in C, \varphi_1(x, y) \leq z \leq \varphi_2(x, y)\},$$

dove $\varphi_1, \varphi_2: C \rightarrow \mathbb{R}$ sono due funzioni continue definite in un sottoinsieme chiuso e limitato C di \mathbb{R}^2 . Si dice che A è *semplice (o normale) rispetto all'asse delle z* , perché ogni retta parallela a tale asse lo interseca in un intervallo (di estremi $\varphi_1(x, y)$ e $\varphi_2(x, y)$, per $(x, y) \in C$).

Supponiamo che f sia una funzione integrabile in A . Dal Teorema di Fubini si ottiene la seguente importante **formula di riduzione**, detta anche **formula degli spaghetti** (paralleli all'asse z), valida per gli insiemi che sono semplici rispetto all'asse delle z :

$$\iiint_A f(x, y, z) \, dx \, dy \, dz = \iint_C \left(\int_{\varphi_1(x, y)}^{\varphi_2(x, y)} f(x, y, z) \, dz \right) \, dx \, dy,$$

dove C è la proiezione ortogonale di A sul piano xy e $\varphi_1, \varphi_2: C \rightarrow \mathbb{R}$ sono due funzioni i cui grafici delimitano A .

Ovviamente si hanno altre due formule degli spaghetti: una con spaghetti paralleli all'asse x e l'altra con spaghetti paralleli all'asse y . I dettagli sono lasciati allo studente.

Esempio. Calcoliamo

$$\iiint_A z \, dx \, dy \, dz$$

ove

$$A = \{(x, y, z) \in \mathbb{R}^3 : x^2 + 4y^2 \leq 1, 0 \leq z \leq x^2 + 4y^2\}.$$

Usando la formula degli spaghetti si ottiene

$$\iiint_A z \, dx \, dy \, dz = \iint_C \left(\int_0^{x^2+4y^2} z \, dz \right) \, dx \, dy = \iint_C \frac{(x^2 + 4y^2)^2}{2} \, dx \, dy,$$

ove $C = \{(x, y) \in \mathbb{R}^2 : x^2 + 4y^2 \leq 1\}$. L'ultimo integrale si calcola facilmente facendo uso di coordinate ellittiche, cioè ponendo $x = \rho \cos \theta$, $y = \frac{1}{2}\rho \sin \theta$. Si ha cioè

$$\iint_C \frac{(x^2 + 4y^2)^2}{2} dx dy = 2\pi \int_0^1 \frac{\rho^4}{2} \frac{1}{2} \rho d\rho = \frac{\pi}{12}.$$

Esempio. Calcoliamo il volume V di una sfera di raggio r con la formula degli spaghetti. Si ha

$$\frac{V}{2} = \iint_C \left(\int_0^{\sqrt{r^2 - x^2 - y^2}} dz \right) dx dy = \iint_C \sqrt{r^2 - x^2 - y^2} dx dy,$$

ove $C = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2\}$. Passando alle coordinate polari si ottiene

$$\iint_C \sqrt{r^2 - x^2 - y^2} dx dy = 2\pi \int_0^r \sqrt{r^2 - \rho^2} \rho d\rho = \frac{2}{3}\pi r^3.$$

Dal Teorema di Fubini si deduce anche un'altra formula di riduzione per il calcolo di un integrale triplo in un insieme limitato A : la **formula delle fette**. Anche in questo caso, in realtà, si avranno tre formule, a seconda che le fette siano perpendicolari all'asse z , all'asse x o all'asse y .

Illustriamo ora la formula delle fette (perpendicolari all'asse z). Sia $(x, y, z) \mapsto f(x, y, z)$ una funzione integrabile in un insieme limitato $A \subseteq \mathbb{R}^3$. Fissato $z \in \mathbb{R}$, denotiamo con

$$A_z = \{(x, y) \in \mathbb{R}^2 : (x, y, z) \in A\}$$

la “fetta” (eventualmente vuota) che si ottiene “tagliando” A con il piano perpendicolare all'asse z e passante per il punto $(0, 0, z)$. Sia $[a, b]$ un intervallo contenente la proiezione ortogonale di A sull'asse z . Allora vale la seguente **formula delle fette**:

$$\iiint_A f(x, y, z) dx dy dz = \int_a^b \left(\iint_{A_z} f(x, y, z) dx dy \right) dz.$$

Osservazione. Dalla formula delle fette si deduce che il volume $\mu_3(A)$ di un solido A la cui proiezione ortogonale sull'asse z risulti contenuta in un intervallo $[a, b]$ si ottiene integrando tra a e b l'area $\mu_2(A_z)$ della generica fetta A_z . Se l'intervallo $[a, b]$ è troppo grande, alcune fette A_z sono vuote, e quindi, per tali fette, risulta $\mu_2(A_z) = 0$ (pertanto, tanto vale scegliere $a = \inf\{z : (x, y, z) \in A\}$ e $b = \sup\{z : (x, y, z) \in A\}$). Risulta perciò

$$\mu_3(A) = \iiint_A dx dy dz = \int_a^b \left(\iint_{A_z} dx dy \right) dz = \int_a^b \mu_2(A_z) dz.$$

Esempio. Calcoliamo il volume V di una sfera di raggio r con la formula delle fette. Si ha

$$\frac{V}{2} = \int_0^r \mu_2(A_z) dz = \int_0^r \pi(r^2 - z^2) dz = \frac{2}{3}\pi r^3,$$

dove A_z è la fetta alla quota z , ossia $A_z = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2 - z^2\}$.

Esempio. Calcoliamo il volume del “cono col gelato”

$$A = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 - z \leq 0, z^2 \geq x^2 + y^2\}.$$

Usando la formula delle fette si ha

$$V = \int_0^{\frac{1}{2}} \mu_2(A_z) dz + \int_{\frac{1}{2}}^1 \mu_2(\hat{A}_z) dz = \int_0^{\frac{1}{2}} \pi z^2 dz + \int_{\frac{1}{2}}^1 \pi(z - z^2) dz,$$

ove

$$A_z = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq z^2\}$$

e

$$\hat{A}_z = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq z - z^2\}.$$

Con la formula degli spaghetti si avrebbe

$$\begin{aligned} V &= \iint_C \left(\int_{\sqrt{x^2+y^2}}^{\frac{1}{2} + \sqrt{\frac{1}{4} - x^2 - y^2}} dz \right) dx dy \\ &= \iint_C \left(\frac{1}{2} + \sqrt{\frac{1}{4} - x^2 - y^2} - \sqrt{x^2 + y^2} \right) dx dy, \end{aligned}$$

ove $C = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1/4\}$. In questo caso, la conclusione si ottiene usando coordinate polari.

Esercizio. Calcolare il volume di un cono circolare retto di altezza h e raggio di base r .

Suggerimento: Sia

$$A = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \leq \frac{r^2}{h^2}(z - h)^2, 0 \leq z \leq h\}.$$

Usando la formula di integrazione per fette e denotato con A_z l'insieme

$$A_z = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq \frac{r^2}{h^2}(z - h)^2\},$$

si ottiene

$$V = \int_0^h \mu_2(A_z) dz = \int_0^h \pi \frac{r^2}{h^2} (z-h)^2 dz = \frac{\pi}{3} r^2 h.$$

Allo stesso risultato si perviene ovviamente anche usando la formula degli spaghetti. Si invita lo studente a verificarlo.

Esempio. Sia $f: [a, b] \rightarrow \mathbb{R}$ una funzione continua e non negativa. Vogliamo calcolare il volume V del solido ottenuto dalla rotazione intorno all'asse z dell'insieme $\{(y, z) \in \mathbb{R}^2 : a \leq z \leq b, y = f(z)\}$. Usando la formula di riduzione per fette, si ha

$$V = \int_a^b \mu_2(A_z) dz = \pi \int_a^b f^2(z) dz,$$

dove $A_z = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq f^2(z)\}$.

Ad esempio, il volume del paraboloide ottenuto facendo ruotare il grafico di $f(z) = \sqrt{z}$ in $[0, 1]$ è

$$V = \pi \int_0^1 z dz = \pi [z^2/2]_0^1 = \frac{\pi}{2}.$$

In maniera analoga, facendo ruotare $y = z$ intorno all'asse z si ottiene il volume del cono circolare retto di altezza 1

$$V = \pi \int_0^1 z^2 dz = \pi [z^3/3]_0^1 = \frac{\pi}{3}.$$

Teorema (di cambiamento di variabile per gli integrali tripli). Sia $\varphi(u, v, w) = (\varphi_1(u, v, w), \varphi_2(u, v, w), \varphi_3(u, v, w))$ un'applicazione continua da un insieme chiuso e limitato $D \subseteq \mathbb{R}^3$ in \mathbb{R}^3 . Supponiamo che le frontiere di D e di $\varphi(D)$ siano trascurabili e che φ sia C^1 e iniettiva nell'interno di D . Allora, data una funzione di tre variabili f continua su $\varphi(D)$, risulta

$$\begin{aligned} & \iiint_{\varphi(D)} f(x, y, z) dx dy dz = \\ & \iiint_D f(\varphi_1(u, v, w), \varphi_2(u, v, w), \varphi_3(u, v, w)) |\det J\varphi(u, v, w)| du dv dw. \end{aligned}$$

Esempio. Come esempio calcoliamo il volume V di una sfera di raggio r usando coordinate sferiche. Poniamo $A = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2\}$ e osserviamo che i punti di A sono immagine tramite l'applicazione φ che definisce le coordinate

sferiche, dei punti del parallelepipedo chiuso e limitato $D = \{(\rho, \theta, \varphi) \in \mathbb{R}^3 : 0 \leq \rho \leq r, 0 \leq \theta \leq 2\pi, 0 \leq \varphi \leq \pi\}$. L'applicazione φ è continua in D e di classe C^1 e iniettiva nell'interno di D (non è iniettiva su ∂D). Applicando il teorema di cambiamento di variabile si ottiene

$$\begin{aligned} V &= \iiint_A dx dy dz = \iiint_D \rho^2 |\sin \varphi| d\rho d\theta d\varphi = \\ &= \int_0^\pi \int_0^{2\pi} \int_0^r \rho^2 \sin \varphi d\rho d\theta d\varphi = \\ &= \frac{2}{3} \pi r^3 \int_0^\pi \sin \varphi d\varphi = \frac{2}{3} \pi r^3 [-\cos \varphi]_0^\pi = \frac{4}{3} \pi r^3. \end{aligned}$$

Le definizioni di massa, di centro di massa fisico di un solido (non necessariamente omogeneo) sono analoghe a quelle date per una piastra piana.

Esempio. Determiniamo la massa e il centro di massa di una semisfera di raggio r sapendo che la densità è proporzionale alla distanza dal centro. In un sistema di assi cartesiani, il solido considerato può essere descritto nel modo seguente:

$$A = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 \leq r^2, z \geq 0\}.$$

e la densità di volume è data da $\delta(x, y, z) = k\sqrt{x^2 + y^2 + z^2}$ dove k è una costante positiva. Si ha perciò (introducendo coordinate sferiche e facendo uso del teorema di cambiamento di variabile)

$$\begin{aligned} m &= \iiint_A \delta(x, y, z) dx dy dz = k \iiint_A \sqrt{x^2 + y^2 + z^2} dx dy dz \\ &= 2\pi k \int_0^r \int_0^{\pi/2} \rho^3 \sin \varphi d\varphi d\rho = \frac{1}{2} k \pi r^4. \end{aligned}$$

Per quanto riguarda il centro di massa, per motivi di simmetria risulta, ovviamente, $x_c = y_c = 0$. Basta dunque calcolare

$$z_c = \frac{1}{m} \iiint_A z \delta(x, y, z) dx dy dz.$$

In coordinate sferiche si ottiene

$$z_c = \frac{1}{m} 2\pi k \int_0^r \int_0^{\pi/2} \rho^4 \cos \varphi \sin \varphi d\varphi d\rho = \frac{2}{5} r.$$

9^a settimana - dal 06.05.19

Curve e integrali curvilinei

Una *curva parametrica* in \mathbb{R}^k è una funzione continua $\gamma(t)$ da un intervallo $I \subseteq \mathbb{R}$ in \mathbb{R}^k . Si avrà perciò $\gamma(t) = (\gamma_1(t), \dots, \gamma_k(t))$, dove le funzioni $\gamma_i: I \rightarrow \mathbb{R}$, dette le *componenti* di γ , sono ovviamente funzioni continue per ogni $i = 1, \dots, k$. Se I è un intervallo chiuso e limitato $[a, b]$, i punti $\gamma(a)$ e $\gamma(b)$ si dicono, rispettivamente, *primo* e *secondo estremo* della curva. Se $\gamma(a) = \gamma(b)$ (ossia, se gli estremi coincidono) la curva si dice *chiusa*. La variabile t di $\gamma(t)$ si chiama il *parametro* della curva. Ovviamente, al posto di t si può usare una qualunque altra lettera (le più usate sono τ, s, θ, φ , mentre, di solito, si preferisce non denotare con x il parametro della curva).

Le equazioni

$$\begin{cases} x_1 = \gamma_1(t), \\ \vdots \\ x_k = \gamma_k(t) \end{cases}$$

si chiamano *equazioni parametriche* della curva $\gamma(t)$. Se $k = 2$ la curva si dice *piana* e le componenti talvolta vengono denotate $x(t)$ e $y(t)$. Una curva parametrica γ in \mathbb{R}^2 (o in \mathbb{R}^3) rappresenta dal punto di vista cinematico il moto di un punto nel piano (o nello spazio) e il parametro t rappresenta il tempo (si chiama anche *curva oraria* del moto).

L'immagine di una curva parametrica, cioè l'insieme $\gamma(I)$ è detto il *sostegno* della curva. Ad esempio, la curva $\gamma(t) = (r \cos t, r \sin t)$, $t \in [0, \pi]$, ha come sostegno la semicirconferenza di centro l'origine e raggio r , con primo estremo $\gamma(0) = (r, 0)$ e secondo estremo $\gamma(\pi) = (-r, 0)$.

È importante non confondere la curva parametrica con il suo sostegno. Tornando all'esempio della cinematica, la curva oraria γ rappresenta non solo la traiettoria (che è il sostegno di γ) percorsa dal punto $\gamma(t)$, ma l'intera legge di percorrenza: uno stesso sottoinsieme di \mathbb{R}^2 (o di \mathbb{R}^3) può essere sostegno di differenti curve parametriche; può essere infatti percorso in più modi e con velocità diverse. Il seguente esempio illustra quanto detto:

Esempio. Le tre curve

$$\gamma(t) = (\cos t, \sin t), \quad t \in [0, 2\pi],$$

$$\tilde{\gamma}(t) = (\cos t, \sin t), \quad t \in [0, 4\pi]$$

e

$$\hat{\gamma}(t) = (\cos 2t, \sin 2t), \quad t \in [0, \pi]$$

hanno tutte come sostegno la circonferenza di centro l'origine e raggio 1. Tale circonferenza però è percorsa, rispettivamente, una volta, due volte e una volta ma con velocità doppia.

Una curva parametrica $\gamma: [a, b] \rightarrow \mathbb{R}^k$ si dice *semplice* se comunque si prendano due punti $t_1, t_2 \in [a, b]$ di cui almeno uno appartenente ad (a, b) si ha $\gamma(t_1) \neq \gamma(t_2)$. Osserviamo che se $\gamma(a) \neq \gamma(b)$ la definizione equivale all'iniettività di γ in $[a, b]$. Nella precedente definizione rientrano però anche le curve chiuse. In tal caso, essendo $\gamma(a) = \gamma(b)$, la funzione γ non potrà ovviamente essere iniettiva nell'intervallo chiuso $[a, b]$.

Osservazione. Il sostegno di una curva parametrica piana, semplice e chiusa è detto *curva di Jordan*. Un esempio di curva di Jordan è dato da una circonferenza, o anche dalla frontiera di un triangolo o, più in generale di un poligono. Un famoso risultato di topologia, intuitivo nell'enunciato ma la cui dimostrazione non è semplice, afferma che *ogni curva di Jordan divide il piano in due insiemi aperti di cui uno è limitato (insieme dei punti racchiusi dalla curva) e l'altro è non limitato*.

Esempio.

1. La curva $\gamma: [0, 2\pi] \rightarrow \mathbb{R}^2$, $\gamma(t) = (a \cos t, b \sin t)$, $a > 0, b > 0$, è semplice e chiusa e il suo sostegno è l'ellisse $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$. In particolare, se $a = b$, il sostegno è la circonferenza di centro l'origine e raggio a .
2. La curva $\gamma: \mathbb{R} \rightarrow \mathbb{R}^2$, $\gamma(t) = (a \cosh t, b \sinh t)$, $a > 0, b > 0$, è semplice e il suo sostegno è il ramo per $x > a$ dell'iperbole $\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1$.
3. Sia $g: I \rightarrow \mathbb{R}$ una funzione continua. La curva parametrica $\gamma: I \rightarrow \mathbb{R}^2$ definita da $\gamma(t) = (t, g(t))$ è semplice e piana e il suo sostegno coincide con il grafico di g .
4. Le equazioni $x(t) = x_0 + pt$, $y(t) = y_0 + qt$, $z(t) = z_0 + rt$, $t \in \mathbb{R}$, sono le equazioni parametriche di una retta nello spazio passante per il punto (x_0, y_0, z_0) .
5. La curva $\gamma: \mathbb{R} \rightarrow \mathbb{R}^3$, $\gamma(t) = (r \cos t, r \sin t, kt)$, $r, k \in \mathbb{R}$, è semplice, non è chiusa e il suo sostegno è un'elica che si appoggia al cilindro $\{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 = r^2\}$. L'innalzamento, dopo un periodo 2π , della quota z di un punto che percorre l'elica è detto *passo dell'elica* e vale $2\pi k$.

Una curva semplice γ induce una *orientazione* del suo sostegno, cioè un verso di percorrenza di esso al crescere del parametro t . Diremo che il punto

$\gamma(t_1)$ precede $\gamma(t_2)$ sul sostegno di γ , se $t_1 < t_2$. Ad esempio, la circonferenza che è sostegno della parametrizzazione $\gamma(t) = (\cos t, \sin t)$, $t \in [0, 2\pi]$, è percorsa in senso antiorario. Se consideriamo invece la parametrizzazione $\tilde{\gamma}(t) = (\cos(-t), \sin(-t))$, $t \in [-2\pi, 0]$, lo stesso sostegno è percorso in senso orario.

Data una curva parametrica γ , la sua *derivata* in un punto t_0 è, quando esiste, il limite per $t \rightarrow t_0$ del rapporto incrementale

$$\frac{\gamma(t) - \gamma(t_0)}{t - t_0}.$$

Ovviamente, tale rapporto ha senso perché è il prodotto del vettore $\gamma(t) - \gamma(t_0)$ per lo scalare $1/(t - t_0)$. È facile verificare che γ è derivabile in t_0 se e solo se sono derivabili in t_0 le sue funzioni componenti. In tal caso la derivata di γ può essere eseguita componente per componente. La derivata di γ in t_0 si denota col simbolo $\gamma'(t_0)$. Diremo inoltre che γ è di classe C^n se sono C^n tutte le sue funzioni componenti.

Quando γ è la curva oraria del moto, il vettore $\gamma'(t)$ rappresenta la velocità istantanea, mentre la sua norma $\|\gamma'(t)\|$ è la velocità scalare.

Una curva $\gamma: [a, b] \rightarrow \mathbb{R}^k$ si dice C^1 a tratti se è continua in $[a, b]$ (questo è già implicito nella definizione di curva) e se è possibile dividere $[a, b]$ in un numero finito di intervalli chiusi $J_i = [t_{i-1}, t_i]$, $a = t_0 < t_1 < \dots < t_n = b$, in ognuno dei quali γ è C^1 .

Esempio. La curva piana

$$\gamma(t) = (t, |t|), \quad t \in [-1, 1]$$

è C^1 a tratti ma non C^1 .

Sia $\gamma: [a, b] \rightarrow A$ una curva parametrica con sostegno in un aperto A di \mathbb{R}^k e sia $f: A \rightarrow \mathbb{R}$ una funzione continua definita su A . Vogliamo introdurre la nozione di *integrale curvilineo (non orientato)* di f lungo la curva γ .

Fissiamo una partizione puntata α di $[a, b]$ costituita da una partizione $\{t_0, t_1, \dots, t_n\}$ di $[a, b]$ e da una scelta $\{\theta_1, \theta_2, \dots, \theta_n\}$ di punti tali che $\theta_i \in I_i = [t_{i-1}, t_i]$. Alla partizione puntata α possiamo associare il numero

$$\Sigma(\alpha) = \sum_{i=1}^n f(\gamma(\theta_i)) \|\gamma(t_i) - \gamma(t_{i-1})\|,$$

dove $\|\gamma(t_i) - \gamma(t_{i-1})\|$ denota la lunghezza del segmento che ha per estremi i due punti $\gamma(t_{i-1})$ e $\gamma(t_i)$ sul sostegno della curva γ .

In questo modo risulta definita, nell'insieme \mathcal{P} delle partizioni puntate di $[a, b]$, una funzione reale Σ che ad ogni $\alpha \in \mathcal{P}$ associa la somma $\Sigma(\alpha)$.

L'integrale curvilineo (non orientato) lungo γ di f in ds è, quando esiste, il valore l a cui tende $\Sigma(\alpha)$ quando il parametro di finezza $|\alpha|$ della partizione α tende a zero. Più precisamente scriveremo che

$$\lim_{|\alpha| \rightarrow 0} \Sigma(\alpha) = l \in \mathbb{R}$$

se fissato $\epsilon > 0$ esiste $\delta > 0$ tale che da $|\alpha| < \delta$ segue $|\Sigma(\alpha) - l| < \epsilon$.

Denoteremo tale limite col simbolo

$$\int_{\gamma} f(P) ds \quad \text{o, anche,} \quad \int_{\gamma} f ds.$$

Il termine ds viene detto *elemento di lunghezza d'arco*.

Dalle proprietà dei limiti discende facilmente che se λ_1 e λ_2 sono due costanti e $f_1, f_2: A \rightarrow \mathbb{R}$ due funzioni continue, allora

$$\int_{\gamma} (\lambda_1 f_1(P) + \lambda_2 f_2(P)) ds = \lambda_1 \int_{\gamma} f_1(P) ds + \lambda_2 \int_{\gamma} f_2(P) ds.$$

Dalla proprietà del confronto dei limiti discende la monotonia:

se $f_1(P) \leq f_2(P)$ per ogni P sul sostegno di γ , allora

$$\int_{\gamma} f_1(P) ds \leq \int_{\gamma} f_2(P) ds.$$

Un'altra proprietà importante è l'additività.

Siano $\gamma_1: [a, b] \rightarrow A$ e $\gamma_2: [b, c] \rightarrow A$ due curve parametriche tali che $\gamma_1(b) = \gamma_2(b)$. La curva $\gamma: [a, c] \rightarrow A$ definita da

$$\gamma(t) = \begin{cases} \gamma_1(t) & \text{se } t \in [a, b] \\ \gamma_2(t) & \text{se } t \in [b, c] \end{cases}$$

è detta *unione* di γ_1 e γ_2 e denotata $\gamma_1 \cup \gamma_2$. Notiamo che se γ_1 e γ_2 sono regolari, in generale $\gamma_1 \cup \gamma_2$ è solo regolare a tratti. Data $f: A \rightarrow \mathbb{R}$ continua, si ha

$$\int_{\gamma_1 \cup \gamma_2} f(P) ds = \int_{\gamma_1} f(P) ds + \int_{\gamma_2} f(P) ds.$$

Con riferimento alla definizione di integrale curvilineo non orientato, si osservi che quando $f(P) \equiv 1$, la somma $\Sigma(\alpha)$ associata ad una partizione puntata α non è altro che la lunghezza della poligonale inscritta al sostegno di γ con vertici $\gamma(t_0), \gamma(t_1), \dots, \gamma(t_n)$. Pertanto, se esiste il limite per $|\alpha| \rightarrow 0$ delle lunghezze delle poligonali associate alle partizioni α di $[a, b]$, tale limite può essere assunto come definizione della *lunghezza* di γ . Ovviamente, in questo caso, essendo $f(P) \equiv 1$, è superfluo che le partizioni di $[a, b]$ siano puntate.

Definizione. Una curva parametrica γ si dice *rettificabile* se esiste l'integrale lungo γ della funzione $f(P) \equiv 1$ in ds . In tal caso, si definisce *lunghezza* di γ il numero

$$L(\gamma) = \int_{\gamma} ds.$$

Non tutte le curve parametriche (continue) sono rettificabili. Per avere un'idea di ciò, si fa presente che nel 1890 l'illustre matematico torinese Giuseppe Peano (1858-1932) fornì uno esempio di curva parametrica (continua) il cui sostegno è un intero quadrato (pieno!): una tale curva non è certo rettificabile. Nel 1935, tuttavia, A. B. Brown mostrò che un tale fenomeno non può verificarsi se la curva è di classe C^1 , perché il suo sostegno risulta privo di punti interni (sono appunto le curve di classe C^1 , le curve che più aderiscono alla nostra concezione intuitiva di curva). Come vedremo nel teorema che segue, le curve di classe C^1 (o di classe C^1 a tratti) sono rettificabili.

Il seguente importante risultato riconduce il calcolo di un integrale curvilineo (non orientato) a quello di un integrale definito.

Teorema (di riduzione agli integrali semplici per gli integrali curvilinei non orientati). *Se $f: A \subseteq \mathbb{R}^k \rightarrow \mathbb{R}$ è continua e $\gamma: [a, b] \rightarrow A$ è C^1 a tratti, allora*

$$\int_{\gamma} f(P) ds = \int_a^b f(\gamma(t)) \|\gamma'(t)\| dt.$$

In particolare, γ è rettificabile e si ha

$$L(\gamma) = \int_a^b \|\gamma'(t)\| dt.$$

Esempio. Calcoliamo la lunghezza della curva $\gamma_r(t) = (r \cos t, r \sin t)$, $t \in [0, 2\pi]$, il cui sostegno è ovviamente la circonferenza di centro l'origine e raggio r . Si ha

$$L(\gamma_r) = \int_0^{2\pi} \|\gamma_r'(t)\| dt = \int_0^{2\pi} \sqrt{r^2 \sin^2 t + r^2 \cos^2 t} dt = 2\pi r.$$

Calcoliamo invece la lunghezza della curva $\tilde{\gamma}_r(t) = (r \cos t, r \sin t)$, $t \in [0, 4\pi]$, il cui sostegno è ancora la circonferenza di centro l'origine e raggio r , ma tale sostegno è percorso due volte. Si ha

$$L(\tilde{\gamma}_r) = \int_0^{4\pi} \|\tilde{\gamma}'_r(t)\| dt = 4\pi r.$$

In generale perciò, calcolare la lunghezza di una curva non significa calcolare la lunghezza del suo sostegno. Infatti, ad esempio, quando γ è la curva oraria del moto, il numero $L(\gamma)$ rappresenta proprio la lunghezza della traiettoria percorsa (ricordare che $\|\gamma'(t)\|$ è la velocità scalare).

Esempio. Consideriamo l'arco di parabola $g(t) = t^2$, $t \in [0, 1]$, e sia $\gamma(t) = (t, t^2)$, $t \in [0, 1]$, la curva il cui sostegno coincide con il grafico di g . Calcoliamo $\int_\gamma x ds$. Si ha

$$\int_\gamma x ds = \int_0^1 t \sqrt{1 + 4t^2} dt = \frac{1}{12} \left[(1 + 4t^2)^{3/2} \right]_0^1 = \frac{1}{12} (5\sqrt{5} - 1).$$

Esempio. Calcoliamo la lunghezza dell'arco di elica cilindrica $\gamma(t) = (r \cos t, r \sin t, kt)$, $t \in [0, 2\pi]$. Si ha

$$L(\gamma) = \int_0^{2\pi} \|\gamma'(t)\| dt = \int_0^{2\pi} \sqrt{r^2 \sin^2 t + r^2 \cos^2 t + k^2} dt = 2\pi \sqrt{r^2 + k^2}.$$

Se invece vogliamo calcolare l'integrale curvilineo della funzione $f(x, y, z) = z$ lungo l'elica γ si avrà

$$\begin{aligned} \int_\gamma z ds &= \int_0^{2\pi} kt \sqrt{r^2 \sin^2 t + r^2 \cos^2 t + k^2} dt = \\ &= k \sqrt{r^2 + k^2} \int_0^{2\pi} t dt = 2\pi^2 k \sqrt{r^2 + k^2}. \end{aligned}$$

Una curva γ si dice *regolare* se è C^1 in I e se $\|\gamma'(t)\| \neq 0$ per ogni valore t del parametro interno ad I . In altre parole, γ è regolare se è C^1 e le derivate delle sue funzioni componenti non si annullano mai simultaneamente nei punti interni di I . In tali punti risulta ben definito il versore

$$T(t) := \frac{\gamma'(t)}{\|\gamma'(t)\|},$$

detto *versore tangente*.

Osservazione. Quando γ rappresenta la curva oraria del moto di un punto materiale nel piano o nello spazio e il moto è C^1 , dire che è regolare significa affermare che il punto $\gamma(t)$ non si ferma mai.

Esempio. Sia $g: [a, b] \rightarrow \mathbb{R}$ una funzione di classe C^1 . La curva $\gamma(t) = (t, g(t))$, $t \in [a, b]$, il cui sostegno è il grafico di g , è ovviamente una curva regolare essendo $\gamma'(t) = (1, g'(t)) \neq (0, 0)$ per ogni $t \in [a, b]$. Inoltre la lunghezza di γ è

$$L(\gamma) = \int_a^b \sqrt{1 + g'^2(t)} dt.$$

Una curva $\gamma: [a, b] \rightarrow \mathbb{R}^k$ si dice *regolare a tratti* (o *generalmente regolare*) se è continua in $[a, b]$ (questo è già implicito nella definizione di curva) e se è possibile dividere $[a, b]$ in un numero finito di intervalli chiusi $J_i = [t_{i-1}, t_i]$, $a = t_0 < t_1 < \dots < t_n = b$, in ognuno dei quali γ è regolare.

Esempio. La curva piana

$$\gamma(t) = (t^2, t^3), t \in [-1, 1]$$

è regolare a tratti ma non regolare.

Definizione. Un'applicazione $\sigma: [\alpha, \beta] \rightarrow [a, b]$ tra due intervalli chiusi e limitati si dice un *cambiamento di parametro* se è di classe C^1 , suriettiva, con derivata sempre diversa da zero. Un cambiamento di parametro $t = \sigma(\tau)$ è detto *concorde* se $\sigma'(\tau) > 0$ e *discorde* in caso contrario.

Osserviamo che se $t = \sigma(\tau)$ è un cambiamento di parametro, allora σ è invertibile e anche $\tau = \sigma^{-1}(t)$ è un cambiamento di parametro.

Definizione. Due curve parametriche γ e δ si dicono *equivalenti* se una si ottiene dall'altra (mediante la composizione) con un cambiamento di parametro. In altre parole, $\gamma: [a, b] \rightarrow \mathbb{R}^k$ e $\delta: [\alpha, \beta] \rightarrow \mathbb{R}^k$ sono equivalenti se esiste un cambiamento di parametro $\sigma: [\alpha, \beta] \rightarrow [a, b]$ tale che $\delta(\tau) = \gamma(\sigma(\tau))$ per ogni $\tau \in [\alpha, \beta]$. Due curve parametriche equivalenti si dicono *concordi* o *discordi* a seconda che sia concorde o discorde il cambiamento di parametro che fa passare dall'una all'altra.

Osservazione. Se $t \in [a, b] \mapsto \gamma(t)$ è una curva parametrica, la curva $\tau \in [-b, -a] \mapsto \gamma(-\tau)$ è equivalente a γ e discorde con essa. In questo caso si ha $\sigma(\tau) = -\tau$. Per motivi evidenti, di solito la curva $\gamma \circ \sigma$ viene denotata con $-\gamma$.

Osservazione. Ovviamente curve equivalenti hanno lo stesso sostegno. Viceversa, non è vero in generale che curve con lo stesso sostegno siano equivalenti. Ad esempio le due curve $\gamma(t) = (\cos t, \sin t)$, $t \in [0, 2\pi]$ e $\delta(t) = (\cos t, \sin t)$, $t \in [0, 4\pi]$ hanno lo stesso sostegno (come già osservato in precedenza) ma non sono equivalenti in quanto la prima è semplice mentre la seconda no.

Rafforzando le ipotesi, si può però dimostrare il seguente risultato:

Teorema. *Se due curve semplici e regolari hanno lo stesso sostegno, allora sono equivalenti.*

Teorema (di indipendenza per gli integrali curvilinei non orientati). *Sia $f: A \subseteq \mathbb{R}^k \rightarrow \mathbb{R}$ un'applicazione continua e siano γ e δ due curve parametriche regolari (o, regolari a tratti), con sostegno in A ed equivalenti. Allora*

$$\int_{\gamma} f(P) ds = \int_{\delta} f(P) ds,$$

indipendentemente dal fatto che γ e δ siano concordi o discordi. In particolare, le due curve γ e δ hanno la stessa lunghezza.

Dimostrazione. Sia $\sigma: [\alpha, \beta] \rightarrow [a, b]$ un cambiamento di parametro tale che $\delta(\tau) = \gamma(\sigma(\tau))$ per ogni $\tau \in [\alpha, \beta]$. Si ha

$$\int_{\delta} f(P) ds = \int_{\alpha}^{\beta} f(\delta(\tau)) \|\delta'(\tau)\| d\tau = \int_{\alpha}^{\beta} f(\gamma(\sigma(\tau))) \|\gamma'(\sigma(\tau))\| |\sigma'(\tau)| d\tau.$$

Se γ e δ sono concordi, si ha $\sigma'(\tau) > 0$ e quindi $|\sigma'(\tau)| = \sigma'(\tau)$ e $\sigma(\alpha) = a, \sigma(\beta) = b$. Di conseguenza, operando nell'ultimo integrale la sostituzione $t = \sigma(\tau)$, si ottiene

$$\int_{\alpha}^{\beta} f(\gamma(\sigma(\tau))) \|\gamma'(\sigma(\tau))\| \sigma'(\tau) d\tau = \int_a^b f(\gamma(t)) \|\gamma'(t)\| dt = \int_{\gamma} f(P) ds.$$

Se invece γ e δ sono discordi, si ha $\sigma'(\tau) < 0$ e quindi $|\sigma'(\tau)| = -\sigma'(\tau)$, $\sigma(\alpha) = b, \sigma(\beta) = a$, da cui

$$\int_{\alpha}^{\beta} f(\gamma(\sigma(\tau))) \|\gamma'(\sigma(\tau))\| (-\sigma'(\tau)) d\tau = - \int_b^a f(\gamma(t)) \|\gamma'(t)\| dt = \int_{\gamma} f(P) ds.$$

Perciò, si ha in ogni caso $\int_{\gamma} f(P) ds = \int_{\delta} f(P) ds$. \square

I due teoremi precedenti giustificano l'uso di espressioni come “lunghezza di una circonferenza”, “lunghezza di un filo”, “integrale curvilineo sulla frontiera di un quadrato”, ecc. Si intende cioè riferirsi a sottoinsiemi di \mathbb{R}^2 o di \mathbb{R}^3 che siano sostegno di curve semplici e regolari (o regolari a tratti).

Esempio. Calcoliamo $\int_{\gamma} y ds$ dove γ è una curva (semplice e regolare) il cui sostegno è $S = \{(x, y) : x^2 + y^2 - x = 0, x \geq 0, y \geq 0\}$ di estremi $P_1 = (0, 0)$ e $P_2 = (1, 0)$. Delle equazioni parametriche per S sono

$$\gamma(t) = (x(t), y(t)) = \left(\frac{1}{2} + \frac{1}{2} \cos t, \frac{1}{2} \sin t\right), \quad 0 \leq t \leq \pi.$$

Osserviamo che queste equazioni parametriche inducono sul sostegno S il verso di percorrenza da P_2 a P_1 . Per il teorema precedente il risultato è indipendente dal verso di percorrenza del sostegno di γ . Si ha

$$\int_{\gamma} y \, ds = \int_0^{\pi} \frac{1}{2} \sin t \sqrt{\frac{1}{4} \sin^2 t + \frac{1}{4} \cos^2 t} \, dt = \frac{1}{2}.$$

Esempio. Calcoliamo $\int_{\gamma} x^2 \, ds$ dove γ è una curva semplice il cui sostegno è il triangolo di vertici $A = (-1, 1)$, $B = (1, 1)$, $C = (0, 0)$ percorso in senso antiorario. Per l'additività dell'integrale si ha

$$\int_{\gamma} x^2 \, ds = \sum_{i=1}^3 \int_{\gamma_i} x^2 \, ds,$$

dove γ_1 ha come sostegno il segmento \overline{AB} , γ_2 ha come sostegno il segmento \overline{AC} e γ_3 ha come sostegno il segmento \overline{BC} . Una parametrizzazione di \overline{AB} è, ad esempio, $\gamma_1(t) = (t, 1)$ con $-1 \leq t \leq 1$ e quindi, essendo $\|\gamma_1'(t)\| = 1$ per ogni t ,

$$\int_{\gamma_1} x^2 \, ds = \int_{-1}^1 t^2 \, dt = \frac{2}{3}.$$

In maniera analoga, definiamo $\gamma_2(t) = (t, -t)$ con $-1 \leq t \leq 0$ e $\gamma_3(t) = (t, t)$ con $0 \leq t \leq 1$. Si ha $\|\gamma_2'(t)\| = \|\gamma_3'(t)\| = \sqrt{2}$ per ogni t , da cui

$$\int_{\gamma_2} x^2 \, ds = \int_{-1}^0 t^2 \sqrt{2} \, dt = \frac{\sqrt{2}}{3}$$

e

$$\int_{\gamma_3} x^2 \, ds = \int_0^1 t^2 \sqrt{2} \, dt = \frac{\sqrt{2}}{3}.$$

Di conseguenza,

$$\int_{\gamma} x^2 \, ds = \frac{2 + 2\sqrt{2}}{3}.$$

Come già osservato, il risultato non dipende dall'orientazione di γ_i , $i = 1, 2, 3$.

Esempio: L'ascissa curvilinea. (Facoltativo) Data una curva regolare $\gamma: [a, b] \rightarrow \mathbb{R}^k$, definiamo

$$s = s(t) = \int_a^t \|\gamma'(\theta)\| \, d\theta$$

Per il teorema fondamentale del calcolo integrale, la funzione $s(t)$ è di classe C^1 in $[a, b]$ e $s'(t) = \|\gamma'(t)\| \neq 0$. Perciò essa è un cambiamento di parametro tra gli intervalli $[a, b]$ e $s([a, b]) = [0, L(\gamma)]$. Di conseguenza, risulta ben definito il cambiamento di parametro inverso $t = t(s)$ e la curva $\delta(s) = \gamma(t(s))$ è equivalente a γ . Si ha, ricordando la regola di derivazione della funzione inversa,

$$\delta'(s) = \gamma'(t(s)) t'(s) = \gamma'(t(s)) \frac{1}{s'(t(s))} = \gamma'(t(s)) \frac{1}{\|\gamma'(t(s))\|},$$

da cui

$$\|\delta'(s)\| = \|\gamma'(t(s))\| \frac{1}{\|\gamma'(t(s))\|} = 1.$$

Il valore s misura la lunghezza dell'arco di curva tra $\gamma(a)$ e $\gamma(t)$ ed è detto *ascissa curvilinea*. In particolare, data $f: A \subseteq \mathbb{R}^k \rightarrow \mathbb{R}$, usando come parametro l'ascissa curvilinea, si ha

$$\int_{\gamma} f(P) ds = \int_0^{L(\gamma)} f(\delta(s)) ds.$$

Un integrale curvilineo non orientato può avere vari significati fisici o geometrici. Ad esempio:

- la massa di un filo (quando $f(P) ds$ è l'elemento di massa dm , ossia quando la funzione (di due o di tre variabili) f rappresenta la densità lineare di massa);
- la carica elettrica su un filo (quando f è una densità di carica);
- il momento d'inerzia di un filo rispetto ad una retta (se f rappresenta il quadrato della distanza dalla retta per la densità di massa o, equivalentemente, quando $f(P) ds$ è il prodotto del quadrato della distanza dalla retta per l'elemento di massa dm);

Data una curva parametrica γ in \mathbb{R}^2 , il *centro di massa geometrico* del sostegno di γ è il punto (x_c, y_c) dato da

$$x_c = \frac{1}{L(\gamma)} \int_{\gamma} x ds \quad \text{e} \quad y_c = \frac{1}{L(\gamma)} \int_{\gamma} y ds.$$

Il *centro di massa fisico* di un filo (non necessariamente omogeneo) che sia il sostegno di una curva γ in \mathbb{R}^2 è quel punto le cui coordinate (x_c, y_c) si ottengono facendo la media ponderata delle funzioni coordinate omologhe mediante la

densità lineare di massa $\delta(x, y)$. Ad esempio, la prima coordinata è data da

$$x_c = \frac{1}{m} \int_{\gamma} x \delta(x, y) ds = \frac{1}{m} \int_{\gamma} x dm,$$

dove $dm = \delta(x, y) dx dy$ si chiama *elemento di massa* e

$$m = \int_{\gamma} \delta(x, y) ds = \int_{\gamma} dm$$

è la massa del filo.

Esempio. Determinare la massa di un filo che ha la forma di una semicirconferenza di raggio 1 sapendo che la densità di massa è $\delta(P) = d^2(P, M)$ ove M è il punto di mezzo del filo e $d(P, M)$ denota la distanza del generico punto P del filo da M . Nel piano xy si può pensare che il filo sia il sostegno di $\gamma: [0, \pi] \rightarrow \mathbb{R}^2$ data da $\gamma(t) = (x(t), y(t)) = (\cos t, \sin t)$. Perciò $\delta(x, y) = x^2 + (y - 1)^2$ e $M = \gamma(\pi/2) = (0, 1)$.

Si ha

$$\begin{aligned} m &= \int_{\gamma} dm = \int_{\gamma} \delta(x, y) ds = \int_{\gamma} (x^2 + (y - 1)^2) ds = \\ &= \int_0^{\pi} (\cos^2 t + (\sin t - 1)^2) \sqrt{(-\sin t)^2 + \cos^2 t} dt = 2 \int_0^{\pi} (1 - \sin t) dt = 2(\pi - 2). \end{aligned}$$

Esercizio. Determinare il centro di massa fisico del filo dell'esempio precedente.

10^a settimana - dal 13.5.19

I numeri complessi (\mathbb{C})

Introduzione ai numeri complessi. L'unità immaginaria. Somma e prodotto di numeri complessi. Dato un numero complesso $z = a + ib$, i numeri reali a e b si dicono rispettivamente *parte reale* e *parte immaginaria* di z ; il numero reale $\sqrt{a^2 + b^2}$ è detto *modulo* di z e si indica $|z|$, mentre il *coniugato* di $z = a + ib$ è il numero complesso $a - ib$ e si denota \bar{z} .

Esistenza delle operazioni inverse della somma e del prodotto.

Osservazione. *Il prodotto di un numero complesso per il suo coniugato è uguale al quadrato del modulo.* Infatti, se $z = a + ib$ si ha

$$z\bar{z} = (a + ib)(a - ib) = a^2 + b^2 = |z|^2.$$

Esercizio. Provare che coniugato di un prodotto [di una somma] è il prodotto [la somma] dei coniugati.

Osservazione. Metodo pratico per eseguire il quoziente di due numeri complessi: si moltiplica numeratore e denominatore per il coniugato del denominatore. Ad esempio

$$\frac{a + ib}{c + id} = \frac{(a + ib)(c - id)}{(c + id)(c - id)} = \frac{(a + ib)(c - id)}{c^2 + d^2} = \frac{ac + bd}{c^2 + d^2} + i \frac{bc - ad}{c^2 + d^2}$$

In particolare, il reciproco di un numero complesso $z \neq 0$ è dato da

$$\frac{1}{z} = \frac{\bar{z}}{|z|^2}.$$

Infatti, se $z = a + ib$, si ha

$$\frac{1}{a + ib} = \frac{a - ib}{(a + ib)(a - ib)} = \frac{a - ib}{a^2 + b^2}.$$

I numeri complessi sono in corrispondenza biunivoca con i punti del piano cartesiano. È naturale perciò rappresentarli in un piano, detto *piano complesso* (o anche piano di Argand-Gauss).

Forma trigonometrica di un numero complesso. Argomenti di un numero complesso.

Teorema. Siano $\rho(\cos \theta + i \operatorname{sen} \theta)$ e $r(\cos \varphi + i \operatorname{sen} \varphi)$ due numeri complessi in forma trigonometrica. Se

$$\rho(\cos \theta + i \operatorname{sen} \theta) = r(\cos \varphi + i \operatorname{sen} \varphi),$$

allora $\rho = r$ ed esiste $k \in \mathbb{Z}$ tale che $\theta - \varphi = 2k\pi$.

Tra tutti gli argomenti di un numero complesso, il più piccolo in valore assoluto (con la preferenza di π rispetto a $-\pi$) viene detto *argomento principale*. In altre parole, l'argomento principale di un numero complesso è l'unico, tra gli infiniti argomenti, che appartiene all'intervallo $(-\pi, \pi]$.

Teorema. Il prodotto di due numeri complessi in forma trigonometrica è un numero complesso che ha per modulo il prodotto dei moduli e per argomento la somma degli argomenti.

Esercizio. Dedurre dal teorema precedente che il modulo del rapporto di due numeri complessi è il rapporto dei moduli e l'argomento è la differenza degli argomenti.

Suggerimento. Ricordarsi del significato di quoziente (come operazione inversa del prodotto).

Sia $\cos \theta + i \operatorname{sen} \theta$ un numero complesso di modulo 1 in forma trigonometrica. La potenza n -esima di tale numero è data dalla seguente relazione, detta *formula di De Moivre*,

$$(\cos \theta + i \operatorname{sen} \theta)^n = \cos n\theta + i \operatorname{sen} n\theta.$$

Di conseguenza, la potenza n -esima del numero $\rho(\cos \theta + i \operatorname{sen} \theta)$ è data da

$$(\rho(\cos \theta + i \operatorname{sen} \theta))^n = \rho^n(\cos n\theta + i \operatorname{sen} n\theta).$$

Definizione. Dato un numero complesso w e dato un numero naturale $n > 1$, le soluzioni dell'equazione $z^n = w$ si chiamano *radici n -esime* di w . Nel caso particolare $w = 1$, le soluzioni dell'equazione $z^n = 1$ sono dette *radici n -esime dell'unità*.

Per il calcolo delle radici n -esime di un numero complesso si ha il seguente

Teorema. Sia data l'equazione $z^n = w$. Allora

i) se $w = 0$, si ha $z = 0$;

ii) se $w \neq 0$, l'equazione ha esattamente le seguenti n soluzioni complesse distinte

$$z_k = \sqrt[n]{|w|} \left(\cos \frac{\varphi + 2k\pi}{n} + i \operatorname{sen} \frac{\varphi + 2k\pi}{n} \right) \quad k = 0, 1, \dots, n-1,$$

dove φ è l'argomento principale di w .

Dimostrazione. Posto $w = r(\cos \varphi + i \operatorname{sen} \varphi)$, si tratta di trovare i numeri $z = \rho(\cos \theta + i \operatorname{sen} \theta)$ tali che

$$(\rho(\cos \theta + i \operatorname{sen} \theta))^n = \rho^n (\cos n\theta + i \operatorname{sen} n\theta) = r(\cos \varphi + i \operatorname{sen} \varphi).$$

Per il teorema sull'uguaglianza di due numeri complessi in forma trigonometrica, si ha che i due numeri precedenti sono uguali se e solo se hanno lo stesso modulo e se i loro argomenti differiscono per un multiplo di 2π . Si ottiene perciò $\rho = \sqrt[n]{r}$ e $n\theta = \varphi + 2k\pi$, $k \in \mathbb{Z}$. La conclusione segue immediatamente osservando che si ottengono n valori di θ distinti e appartenenti all'intervallo $[0, 2\pi)$ se e solo se $k = 0, 1, \dots, n-1$. \square

È facile verificare che, nel piano complesso, le radici n -esime ($n > 2$) di un numero complesso w sono i vertici di un poligono regolare di n lati iscritto in una circonferenza centrata nell'origine e di raggio $\sqrt[n]{|w|}$.

Esempio. Risolviamo l'equazione $z^3 = 1$ o, in maniera equivalente, determiniamo le radici cubiche dell'unità. Scrivendo 1 in forma trigonometrica si ha $\cos 0 + i \operatorname{sen} 0$, da cui, per il teorema precedente, si ottengono le tre soluzioni

$$z_k = \cos \frac{0 + 2k\pi}{3} + i \operatorname{sen} \frac{0 + 2k\pi}{3}, \quad k = 0, 1, 2,$$

e, cioè,

$$\begin{aligned} z_0 &= \cos 0 + i \operatorname{sen} 0 = 1, \\ z_1 &= \cos \frac{0 + 2\pi}{3} + i \operatorname{sen} \frac{0 + 2\pi}{3} = -\frac{1}{2} + i \frac{\sqrt{3}}{2}, \\ z_2 &= \cos \frac{0 + 4\pi}{3} + i \operatorname{sen} \frac{0 + 4\pi}{3} = -\frac{1}{2} - i \frac{\sqrt{3}}{2}. \end{aligned}$$

Esercizio. Risolvere le seguenti equazioni:

$$z^4 = i; \quad z^3 = 1 + i.$$

Monomi in campo complesso. Polinomi in campo complesso.

Come in campo reale, una *radice* di un polinomio complesso $P(z)$ è (per definizione) una soluzione dell'equazione $P(z) = 0$ (cioè un numero complesso z_1 che verifica la condizione $P(z_1) = 0$). Si dice che una radice z_1 di $P(z)$ è *semplice* (o di *molteplicità* 1) se si può scrivere $P(z) = (z - z_1)Q(z)$, con $Q(z_1) \neq 0$; ossia se z_1 è radice di $P(z)$ ma non lo è per il quoziente $Q(z)$ della divisione di $P(z)$ per $z - z_1$. Si dice che z_1 è una radice *doppia* (o di molteplicità 2) se $P(z) = (z - z_1)^2Q(z)$ e $Q(z_1) \neq 0$. In modo analogo si definisce il concetto di radice tripla e, più in generale, di molteplicità k .

Teorema Fondamentale dell'Algebra. *In campo complesso ogni polinomio non costante (cioè, di grado non nullo) ammette almeno una radice.*

Dal Teorema Fondamentale dell'Algebra si deduce immediatamente il seguente risultato.

Corollario. *Un polinomio di grado n ammette (in campo complesso) esattamente n radici, contate con la loro molteplicità.*

Osservazione. Le radici n -esime di un numero complesso w non sono altro che le radici del polinomio $P(z) = z^n - w$ (e, come abbiamo visto, sono esattamente n e tutte semplici).

Notazione esponenziale dei numeri complessi: si pone

$$e^{i\theta} = \cos \theta + i \operatorname{sen} \theta, \quad \theta \in \mathbb{R}.$$

Si ha: $e^{i\theta_1}e^{i\theta_2} = e^{i(\theta_1+\theta_2)}$. Inoltre $e^{2k\pi i} = 1$, $k \in \mathbb{Z}$, e $|e^{i\theta}| = 1$.

Più in generale, per ogni numero complesso $z = a + ib$ poniamo

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

Pertanto, e^{a+ib} è un numero complesso di modulo e^a e argomento b . La funzione, detta *esponenziale complessa*, che a $z = a + ib$ associa $e^z = e^a(\cos b + i \operatorname{sen} b)$ è un esempio di funzione complessa di variabile complessa, ossia di una funzione con dominio e codominio in \mathbb{C} . Dalla definizione precedente si ottiene immediatamente

$$e^{a-ib} = e^a(\cos b - i \operatorname{sen} b),$$

da cui sommando e, rispettivamente, sottraendo membro a membro si ricavano le cosiddette *formule di Eulero*

$$\frac{e^{a+ib} + e^{a-ib}}{2} = e^a \cos b$$

$$\frac{e^{a+ib} - e^{a-ib}}{2i} = e^a \operatorname{sen} b.$$

Equazioni differenziali ordinarie

Sia $f: A \rightarrow \mathbb{R}$ una funzione continua su un aperto A di \mathbb{R}^2 . Un'uguaglianza del tipo

$$y' = f(x, y)$$

si chiama *equazione differenziale (ordinaria del prim'ordine) in forma normale*. Si fa presente che per capire cosa sia un'equazione (anche non differenziale), a parte il modo di chiamarla o di scriverla, è indispensabile aver ben definito il concetto di soluzione. In altre parole, è necessario avere un criterio chiaro per decidere quando, in un assegnato insieme in cui si cercano le soluzioni, un elemento di tale insieme è o non è una soluzione. Per quanto riguarda la precedente equazione, l'insieme in cui si cercano le soluzioni è l'insieme delle funzioni reali definite in un intervallo non banale e derivabili con continuità. Più precisamente,

Definizione. Una funzione reale di una variabile $y: I \rightarrow \mathbb{R}$ definita in un intervallo non banale I e di classe C^1 in I è una *soluzione* dell'equazione differenziale $y' = f(x, y)$ se, per ogni $x \in I$, si ha $(x, y(x)) \in A$ e

$$y'(x) = f(x, y(x)).$$

L'insieme di tutte le soluzioni di un'equazione differenziale si dice *soluzione generale* o *integrale generale*.

Esempio. 1) Il più banale esempio di equazione differenziale è

$$y' = f(x),$$

dove f è una funzione continua definita in un intervallo $I \subseteq \mathbb{R}$. In questo caso la soluzione generale è data dall'insieme delle primitive di f .

2) Un altro esempio semplice di equazione differenziale si ha quando ci poniamo il problema di cercare una funzione definita in un intervallo che ivi coincida con la sua derivata, cioè

$$y' = y.$$

In questo caso si ha $f(x, y) = y$ e una funzione che risolve tale equazione è ovviamente $y(x) = e^x$. Si verifica immediatamente che anche $y(x) = ce^x$, $c \in \mathbb{R}$, è soluzione. Come nell'esempio precedente, anche in questo caso troviamo infinite soluzioni, dipendenti da una costante $c \in \mathbb{R}$. Come vedremo meglio in seguito questo fatto non è casuale ma ha un preciso riscontro teorico.

11^a settimana - dal 20.5.19

Osserviamo che, data una soluzione $y: I \rightarrow \mathbb{R}$, la sua restrizione ad un sottointervallo non banale di I è ancora una soluzione. Tra tutte le soluzioni, quelle che non sono restrizione di altre soluzioni si dicono *massimali* (o *non prolungabili*). Si potrebbe dimostrare che ogni soluzione non massimale è la restrizione di una massimale. In altre parole, ogni soluzione non massimale si può prolungare fino ad ottenere una soluzione massimale.

Non sempre la variabile di un'equazione differenziale viene indicata con x , e non sempre la funzione incognita si denota con y . Ad esempio,

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

è un'equazione differenziale dove t è la variabile e $x(t)$ la funzione incognita. Spesso, quando t denota la variabile “tempo”, si scrive \dot{x} invece che x' .

Sia $y' = f(x, y)$ un'equazione differenziale del prim'ordine in forma normale e sia $A \subseteq \mathbb{R}^2$ l'aperto su cui è definita la funzione f . Dato un punto $(x_0, y_0) \in A$, ci si pone il problema di trovare, tra tutte le soluzioni $y(x)$ dell'equazione, quella che verifica (o quelle che verificano) la condizione $y(x_0) = y_0$. In altre parole, tra tutte le soluzioni, si cercano quelle il cui grafico contiene il punto (x_0, y_0) assegnato. Tale *problema* viene detto *di Cauchy* e si scrive

$$\begin{cases} y' = f(x, y) \\ y(x_0) = y_0. \end{cases}$$

La condizione $y(x_0) = y_0$ si chiama *condizione di Cauchy* o anche *condizione iniziale*. Il punto x_0 si dice *punto* (o *istante*) *iniziale* (della soluzione cercata) e y_0 è il *valore iniziale*.

Il seguente risultato dà una risposta al problema posto e asserisce che la continuità della funzione f assicura l'esistenza di almeno una soluzione del problema di Cauchy per l'equazione $y' = f(x, y)$.

Teorema (di esistenza di Peano). *Sia $f: A \rightarrow \mathbb{R}$ una funzione continua su un aperto $A \subseteq \mathbb{R}^2$. Allora, per ogni $(x_0, y_0) \in A$, l'equazione $y' = f(x, y)$ ammette almeno una soluzione che verifica la condizione $y(x_0) = y_0$.*

Esempio (di non unicità della soluzione di un problema di Cauchy). Si osservi che le funzioni $y_1(x) \equiv 0$ e $y_2(x) = x^3$ sono due soluzioni (massimali) dell'equazione differenziale

$$y' = 3\sqrt[3]{y^2}$$

e verificano entrambe la condizione di Cauchy $y(0) = 0$.

Il risultato che segue fornisce una condizione sufficiente affinché il problema di Cauchy ammetta un'unica soluzione massimale. Ovviamente se si considerano le soluzioni non massimali non si può avere unicità perché la restrizione di una soluzione ad un sottointervallo non banale del dominio è ancora una soluzione (ed è diversa dalla precedente).

Teorema (di esistenza e unicità per le equazioni del prim'ordine). *Consideriamo l'equazione differenziale*

$$y' = f(x, y),$$

dove f è una funzione continua in un aperto $A \subseteq \mathbb{R}^2$. Se f è derivabile rispetto ad y e la derivata parziale f_y è continua, allora, per ogni $(x_0, y_0) \in A$, l'equazione ammette un'unica soluzione massimale che verifica la condizione $y(x_0) = y_0$.

Dal teorema di esistenza e unicità si deduce un'importante conseguenza:

Corollario. *Consideriamo l'equazione differenziale*

$$y' = f(x, y),$$

dove f è una funzione continua in un aperto $A \subseteq \mathbb{R}^2$. Supponiamo che f sia derivabile rispetto ad y con derivata continua. Allora i grafici di due differenti soluzioni massimali non possono intersecarsi.

(Facoltativo) Una proprietà significativa delle soluzioni di un'equazione differenziale è espressa dal teorema che segue. Per enunciarlo occorre introdurre la seguente nozione: una soluzione $y: I \rightarrow \mathbb{R}$ che non sia restrizione di un'altra soluzione definita in un intervallo più ampio a destra (sinistra) si dice *massimale a destra (a sinistra)*, o *non prolungabile a destra (a sinistra)*. Ovviamente, una soluzione è massimale se e solo se è massimale sia a destra sia a sinistra.

Teorema (di Kamke). *Sia $f: A \rightarrow \mathbb{R}$ una funzione continua su un aperto $A \subseteq \mathbb{R}^2$. Il grafico di una soluzione massimale a destra (a sinistra) dell'equazione differenziale*

$$y' = f(x, y)$$

non può essere contenuto in un sottoinsieme limitato e chiuso di A .

In un certo senso il Teorema di Kamke afferma che il grafico di ogni soluzione massimale (che, ricordiamo, è un sottoinsieme del dominio A della funzione f) è una curva che “prosegue” (sia verso destra sia verso sinistra) finché le è consentito proseguire. In parole povere prosegue verso destra (ma anche verso sinistra) fino a raggiungere la frontiera di A (e allora si deve arrestare), oppure se ne va all'infinito

(sempre rimanendo dentro A). Ciò che non può accadere è che il grafico di una soluzione non prolungabile si arresti in un punto interno ad A : il Teorema di Peano gli consentirebbe di proseguire, contraddicendo la non prolungabilità.

Equazione a variabili separabili.

Un'equazione differenziale del tipo

$$y' = a(x)h(y),$$

dove a e h sono funzioni di una variabile definite in aperti di \mathbb{R} , si dice a *variabili separabili*. Cerchiamo di spiegarne il motivo, illustrandone il metodo di risoluzione. Supponiamo a continua e h di classe C^1 . Con tali ipotesi la funzione $f(x, y) := a(x)h(y)$ soddisfa le condizioni del teorema di esistenza e unicità.

Supponiamo inoltre che a si annulli soltanto in punti isolati. Se $y_0 \in \mathbb{R}$ è un punto tale che $h(y_0) = 0$, allora la funzione costante $y(x) \equiv y_0$ è chiaramente una soluzione dell'equazione differenziale. Viceversa, (avendo supposto che a si possa annullare soltanto in punti isolati), ogni soluzione costante $y(x) \equiv y_0$ è tale $h(y_0) = 0$. Le soluzioni costanti sono quindi in corrispondenza biunivoca con gli zeri di h .

Occupiamoci quindi di determinare le soluzioni non costanti. Se $x \mapsto y(x)$ è una tale soluzione, per il teorema di esistenza e unicità si deve avere $h(y(x)) \neq 0$ per ogni x nell'intervallo I in cui è definita y (altrimenti il grafico di y intersecherebbe il grafico di una soluzione costante). Dividendo l'uguaglianza

$$y'(x) = a(x)h(y(x))$$

per $h(y(x))$ si ha allora

$$\frac{y'(x)}{h(y(x))} = a(x).$$

Integrando entrambi i membri dell'uguaglianza si ottiene

$$\int \frac{y'(x)}{h(y(x))} dx = \int a(x) dx.$$

Dunque, denotando con $H(y)$ una primitiva di $1/h(y)$ (in un intervallo in cui h non si annulla) e con $A(x)$ una primitiva di $a(x)$, si ottiene

$$H(y(x)) = A(x) + c,$$

dove c è un'arbitraria costante. (Ovviamente, per verificare che $H(y(x))$ è una primitiva di $y'(x)/h(y(x))$, occorre tener conto del teorema di derivazione di una funzione composta).

Ricavando la y (osserviamo esplicitamente che H è iniettiva perché la stiamo considerando in un intervallo in cui la sua derivata $H'(y) = 1/h(y)$ ha segno costante) si ha la formula

$$y(x) = H^{-1}(A(x) + c)$$

che dà **le soluzioni non costanti** dell'equazione a variabili separabili considerata. Si lascia per esercizio la verifica che ogni funzione del tipo

$$y(x) = H^{-1}(A(x) + c),$$

purché la si consideri definita in un intervallo, è effettivamente una soluzione dell'equazione differenziale

$$y' = a(x)h(y).$$

Si avverte che nell'eseguire la verifica, la presenza di H^{-1} rende indispensabile l'uso del teorema di derivazione di una funzione inversa.

Torniamo per un momento all'uguaglianza

$$\frac{y'(x)}{h(y(x))} = a(x),$$

Essa esprime il fatto che la funzione $y(x)$ verifica l'equazione differenziale

$$\frac{1}{h(y)} \frac{dy}{dx} = a(x),$$

che, per abuso di notazioni (e per tradizione), viene talvolta scritta nella forma

$$\frac{1}{h(y)} dy = a(x)dx,$$

dove la variabile dipendente y è separata dalla variabile indipendente x , nel senso che una sta soltanto nel primo membro dell'equazione e l'altra nel secondo (ed ecco perché l'equazione iniziale si dice “a variabili separabili”). Il metodo tradizionale (ma poco ortodosso) per risolvere l'ultima equazione (quella con le variabili separate) consiste nell'integrare entrambi i membri.

Si ha quindi

$$\int \frac{dy}{h(y)} = \int a(x)dx$$

da cui, con le notazioni introdotte sopra, si ottiene

$$H(y) = A(x) + c,$$

o, anche,

$$y = H^{-1}(A(x) + c).$$

Osservazione. Il metodo per risolvere le equazioni a variabili separabili esposto sopra si può applicare anche quando h è solo continua (non sono cioè soddisfatte le ipotesi del teorema di esistenza e unicità), purché ci si limiti alla ricerca delle soluzioni $y(x)$ tali che $h(y(x)) \neq 0$. Si fa presente che se la funzione reale $y \mapsto h(y)$ non ha la derivata continua, possono esistere soluzioni non costanti il cui grafico incontra il grafico di una soluzione costante, come, ad esempio, accade per la soluzione $y(x) = x^3$ dell'equazione a variabili separabili $y' = 3\sqrt[3]{y^2}$.

Esempio. Consideriamo l'equazione differenziale

$$y' = y^2.$$

Essa è a variabili separabili con $a(x) = 1$ e $h(y) = y^2$. Ovviamente l'equazione possiede la soluzione nulla che è l'unica soluzione costante. Cerchiamo ora le soluzioni non costanti. Sia y una soluzione non costante. Come già osservato, essa non potrà annullarsi in nessun punto e quindi sarà o sempre positiva o sempre negativa. Dividendo per $y^2(x)$ entrambi i membri dell'uguaglianza $y'(x) = y^2(x)$ e integrando si ottiene

$$\int \frac{y'(x)}{y^2(x)} dx = \int 1 dx$$

da cui

$$-\frac{1}{y(x)} = x + c, \quad c \in \mathbb{R}.$$

Di conseguenza, le soluzioni non costanti dell'equazione sono date dalla formula

$$y(x) = -\frac{1}{x + c}$$

e l'intervallo massimale di definizione è $(-\infty, -c)$ se $y(x) > 0$ e $(-c, +\infty)$ se $y(x) < 0$. Ad esempio, la soluzione (massimale) dell'equazione con dato iniziale $y(0) = 1$ si ottiene per $c = -1$ ed è quindi data dalla restrizione della funzione

$$y(x) = \frac{1}{1 - x}$$

all'intervallo $(-\infty, 1)$. Si osservi che anche la soluzione con dato iniziale $y(2) = -1$ si ottiene per $c = -1$, ma non coincide con la precedente soluzione perché in questo caso l'intervallo di definizione è la semiretta $(1, +\infty)$. Se invece consideriamo il dato iniziale $y(3) = 0$, otteniamo la soluzione costante $y(x) = 0$ per ogni x , il cui intervallo di definizione è \mathbb{R} .

Osservazione. Nel caso (molto frequente) in cui la funzione f di un'equazione differenziale in forma normale $y' = f(x, y)$ sia definita in una striscia $A = (a, b) \times \mathbb{R}$ (con a e b reali estesi), il dominio di una qualunque soluzione è necessariamente contenuto nella base (a, b) della striscia. Talvolta però può coincidere con l'intervallo (a, b) stesso. In tal caso si dice che la soluzione è *persistente* (o *globale*). Ovviamente ogni soluzione persistente è necessariamente massimale (ossia non è prolungabile). L'equazione differenziale $y' = y^2$ considerata in precedenza mostra che possono esistere soluzioni massimali non persistenti (il Teorema di Kamke implica che tali soluzioni non possono essere limitate).

Esempio. Consideriamo il problema di Cauchy

$$\begin{cases} y' = \frac{1+y^2}{1+x^2} \\ y(0) = 1. \end{cases}$$

L'equazione differenziale è a variabili separabili con $h(y) = 1 + y^2$ che ovviamente è derivabile infinite volte e quindi il problema di Cauchy ammette una e una sola soluzione massimale. Essendo $h(y) > 0$ per ogni y , possiamo dividere per $h(y) > 0$ e integrare. Si ha

$$\int \frac{y'(x)}{1+y^2(x)} dx = \int \frac{1}{1+x^2} dx$$

da cui

$$\arctang y(x) = \arctang x + c, \quad c \in \mathbb{R}.$$

Considerando la condizione iniziale $y(0) = 1$, si ricava $c = \arctang 1 = \pi/4$, per cui la soluzione massimale del problema di Cauchy è

$$\arctang y(x) = \arctang x + \frac{\pi}{4}$$

definita se $|\arctang x + \pi/4| < \pi/2$, cioè nell'intervallo $(-\infty, 1)$. In questo caso, facendo uso di formule note di trigonometria, si riesce anche a ricavare l'espressione esplicita della soluzione. Si ottiene

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

Esercizio. Trovare le soluzioni dell'equazione differenziale

$$y' = 1 - y^2.$$

Esempio. Consideriamo l'equazione

$$y' = (2x - y)^2.$$

Essa soddisfa le ipotesi del teorema di esistenza e unicità essendo la funzione $f(x, y) = (2x - y)^2$ addirittura derivabile infinite volte. Sia $y(x)$ una soluzione dell'equazione. Poniamo

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

Poiché $z'(x) = 2 - y'(x)$, si ricava che $z(x)$ è soluzione dell'equazione

$$z' = 2 - z^2.$$

Quest'ultima equazione è a variabili separabili e la sua risoluzione è lasciata per esercizio.

Equazioni lineari del prim'ordine.

Una classe particolarmente importante di equazioni differenziali del prim'ordine in forma normale sono le equazioni lineari.

Un'equazione differenziale del prim'ordine si dice *lineare* se è della forma

$$y' = a(x)y + b(x),$$

dove a e b sono due funzioni continue definite in un intervallo I . In particolare, quando il termine noto b è identicamente nullo, l'equazione si dice *lineare omogenea*, e in questo caso la funzione identicamente nulla è soluzione dell'equazione differenziale (si chiama soluzione *banale* o *nulla*).

Cominciamo con lo studiare l'equazione lineare omogenea. Il teorema seguente descrive l'integrale generale di tale equazione.

Teorema. *Sia a una funzione continua in un intervallo $I \subseteq \mathbb{R}$. Le soluzioni (massimali) dell'equazione (lineare omogenea del prim'ordine)*

$$y' = a(x)y$$

sono le funzioni del tipo

$$y(x) = ce^{A(x)},$$

dove $A(x)$ è una primitiva di $a(x)$ in I e c un'arbitraria costante.

Dimostrazione. L'equazione è a variabili separabili. Si osserva che l'unica soluzione costante è quella banale (cioè la funzione identicamente nulla). Per trovare le soluzioni non costanti separiamo le variabili. Sia $y(x)$ una soluzione non costante.

Per quanto osservato in precedenza, si ha $y(x) \neq 0$ per ogni x e, dividendo, si ottiene

$$\frac{y'(x)}{y(x)} = a(x),$$

da cui,

$$\int \frac{y'(x)}{y(x)} dx = \int a(x) dx.$$

Quindi

$$\log |y(x)| = A(x) + k,$$

dove $A(x)$ è una primitiva di $a(x)$ e k è un'arbitraria costante. Pertanto

$$|y(x)| = e^{A(x)+k} = e^k e^{A(x)},$$

cioè

$$y(x) = \pm e^k e^{A(x)}$$

o, equivalentemente, tenendo conto che e^k è un'arbitraria costante positiva, una qualunque soluzione non costante è data da

$$y(x) = ce^{A(x)}, \quad \text{con } c \neq 0.$$

Poiché ponendo $c = 0$ nella precedente equazione si ottiene la soluzione banale (che avevamo considerato a parte), si può affermare che la soluzione generale dell'equazione differenziale $y' = a(x)y$ è data da

$$y(x) = ce^{A(x)},$$

con c costante arbitraria, anche nulla. \square

Consideriamo ora l'equazione lineare non omogenea.

Teorema. *Supponiamo che \bar{y} sia una soluzione dell'equazione differenziale lineare*

$$y' = a(x)y + b(x),$$

dove a e b sono funzioni continue in un intervallo $I \subseteq \mathbb{R}$. Allora la soluzione generale dell'equazione non omogenea si ottiene aggiungendo ad \bar{y} la soluzione generale dell'equazione omogenea associata $y' = a(x)y$, cioè ogni soluzione dell'equazione non omogenea è del tipo

$$y(x) = \bar{y}(x) + ce^{A(x)}, \quad c \in \mathbb{R}.$$

Dimostrazione. Proviamo prima che se u è una soluzione dell'equazione omogenea, allora la funzione $\bar{y} + u$ è soluzione della non omogenea. Dalle uguaglianze

$$\bar{y}'(x) = a(x)\bar{y}(x) + b(x) \quad \text{e} \quad u'(x) = a(x)u(x)$$

segue

$$\begin{aligned} (\bar{y}(x) + u(x))' &= \bar{y}'(x) + u'(x) = (a(x)\bar{y}(x) + b(x)) + a(x)u(x) = \\ &= a(x)(\bar{y}(x) + u(x)) + b(x). \end{aligned}$$

Ciò prova che $\bar{y} + u$ è soluzione della non omogenea.

Viceversa, rimane da provare che se \tilde{y} è una qualunque soluzione dell'equazione non omogenea, allora esiste $\bar{c} \in \mathbb{R}$ tale che $\tilde{y}(x) = \bar{y}(x) + \bar{c}e^{A(x)}$. In altre parole, rimane da provare che, posta $\bar{u}(x) := \tilde{y}(x) - \bar{y}(x)$, la funzione \bar{u} è soluzione dell'equazione omogenea. Si ha

$$\begin{aligned} \bar{u}(x)' &= (\tilde{y}(x) - \bar{y}(x))' = \tilde{y}'(x) - \bar{y}'(x) = \\ &= a(x)\tilde{y}(x) + b(x) - (a(x)\bar{y}(x) + b(x)) = a(x)(\tilde{y}(x) - \bar{y}(x)) = a(x)\bar{u}(x). \end{aligned}$$

□

Osservazione. Il teorema precedente, nel caso particolare in cui la funzione a sia nulla, si riduce ad un risultato ben noto: *data una primitiva \bar{y} di b , ogni altra primitiva si ottiene aggiungendo ad \bar{y} un'arbitraria costante (ossia, la soluzione generale dell'equazione differenziale omogenea $y' = 0$).*

Osservazione. Notiamo che nel caso di un'equazione differenziale lineare la funzione $(x, y) \mapsto a(x)y + b(x)$ è definita nella striscia $I \times \mathbb{R}$, essendo $I \subseteq \mathbb{R}$ l'intervallo di definizione delle funzioni a e b . Si può dimostrare che per tale equazione le soluzioni massimali sono definite in I , cioè, ricordando la definizione data in precedenza, che esse sono persistenti.

Abbiamo visto che per trovare la soluzione generale di un'equazione non omogenea occorre prima trovarne almeno una (comunemente detta *soluzione particolare*). Un metodo per trovare una soluzione particolare è il cosiddetto *metodo di variazione della costante* (per equazioni di ordine superiore al primo si chiama *metodo di variazione delle costanti*).

Consideriamo l'equazione

$$y' = a(x)y + b(x).$$

Sappiamo che la soluzione generale dell'omogenea associata è data da

$$u(x) = ce^{A(x)},$$

dove A è una primitiva di a e c una costante arbitraria. Il metodo consiste nel cercare una soluzione particolare dell'equazione non omogenea, pensando “variabile” la costante c . In altre parole, si cerca una soluzione del tipo

$$\bar{y}(x) = c(x)e^{A(x)}.$$

Derivando si ottiene

$$\bar{y}'(x) = c'(x)e^{A(x)} + c(x)a(x)e^{A(x)}.$$

Sostituendo l'espressione trovata nell'equazione differenziale, si ha

$$c'(x)e^{A(x)} + a(x)c(x)e^{A(x)} = a(x)c(x)e^{A(x)} + b(x),$$

da cui si deduce che \bar{y} è soluzione se (e solo se)

$$c'(x) = e^{-A(x)}b(x),$$

ossia se (e solo se) $c(x)$ è una primitiva di $e^{-A(x)}b(x)$.

Di conseguenza, la soluzione generale dell'equazione non omogenea è data da

$$y(x) = ce^{A(x)} + e^{A(x)} \int e^{-A(x)}b(x) dx,$$

dove c è un'arbitraria costante, ed è definita nell'intervallo I .

Esempio. Consideriamo l'equazione differenziale

$$y' + 2xy = x.$$

Supponiamo di voler trovare, tra tutte le soluzioni, quella che verifica la condizione di Cauchy $y(0) = 0$. Poiché $A(x) = -x^2$, si ha

$$y(x) = ce^{-x^2} + c(x)e^{-x^2},$$

con

$$c'(x) = e^{x^2} x.$$

Integrando, si ottiene perciò

$$c(x) = \frac{1}{2}e^{x^2}$$

e, quindi,

$$y(x) = ce^{-x^2} + \frac{1}{2}.$$

Dobbiamo ancora determinare la costante c in modo che sia verificata la condizione $y(0) = 0$. Abbiamo

$$y(0) = c + \frac{1}{2} = 0,$$

da cui si ricava $c = -1/2$. La soluzione cercata è dunque

$$y(x) = \frac{1}{2} (1 - e^{-x^2}),$$

come si può facilmente verificare (si invita lo studente a farlo). Osserviamo che l'equazione precedente è a variabili separabili e quindi può anche essere risolta con il metodo visto per tali tipi di equazioni.

Esempio. Consideriamo l'equazione

$$y' = xy + e^{x^2/2}.$$

Le soluzioni dell'equazione omogenea $y' = xy$ sono della forma $y(x) = ce^{x^2/2}$, $c \in \mathbb{R}$. Per trovare una soluzione dell'equazione non omogenea usiamo il metodo della variazione della costante, cioè cerchiamola nella forma $\bar{y}(x) = c(x)e^{x^2/2}$. Si ha $c'(x) = 1$, per cui si può prendere come soluzione particolare $\bar{y}(x) = xe^{x^2/2}$. Pertanto, l'integrale generale dell'equazione data è

$$y(x) = (x + c)e^{x^2/2}, c \in \mathbb{R}.$$

Esempio. Consideriamo il problema di Cauchy

$$\begin{cases} y' = xy + x^3 \\ y(0) = 1. \end{cases}$$

Applichiamo direttamente la formula risolutiva delle equazioni lineari. Una primitiva di $a(x) = x$ è $A(x) = x^2/2$; l'integrale generale dell'equazione $y' = xy + x^3$ è perciò

$$y(x) = ce^{x^2/2} + \left(e^{x^2/2} \int x^3 e^{-x^2/2} dx \right), c \in \mathbb{R}.$$

Usando due volte la formula di integrazione per parti e tenendo conto della condizione $y(0) = 1$, si ottiene

$$y(x) = -(x^2 + 2) + 3e^{x^2/2}.$$

Esercizio. Data l'equazione $Ly' + Ry = E$, con $L, R, E > 0$, trovare la soluzione tale che $y(0) = I_0$ e tracciarne il grafico se $I_0 > E/R$. Mostrare che ogni soluzione

tende a E/R per $x \rightarrow +\infty$. Osserviamo che questa equazione può rappresentare ad esempio un modello di circuito elettrico con resistenza R , induttanza L e forza elettromotrice E ; in tal caso la soluzione $y(x)$ è l'intensità di corrente.

Equazioni di ordine superiore al primo.

Passiamo ora a considerare equazioni differenziali di ordine superiore al primo. Un'espressione del tipo

$$y'' = f(x, y, y'),$$

dove f è una funzione continua definita su un aperto A di \mathbb{R}^3 , si dice un'equazione differenziale del second'ordine (in forma normale). Come precedentemente affermato, se di un'equazione non è ben definito il concetto di soluzione, non è ben definita l'equazione stessa; e per introdurre in modo corretto la nozione di equazione occorrono due ingredienti: 1) un insieme in cui si cercano le soluzioni; 2) un criterio chiaro per stabilire quando un elemento di tale insieme abbia il diritto di chiamarsi soluzione.

Per quanto riguarda l'equazione del second'ordine considerata sopra, le soluzioni si cercano nell'insieme delle funzioni definite in un intervallo e aventi derivata seconda continua in tale intervallo. Una funzione y di tale insieme si dirà una soluzione se per ogni x appartenente all'intervallo I in cui è definita risulta

$$(x, y(x), y'(x)) \in A \quad \text{e} \quad y''(x) = f(x, y(x), y'(x)).$$

Dal punto di vista fisico, un'equazione del secondo ordine può rappresentare la legge di moto di un punto materiale di massa unitaria, vincolato a muoversi in una retta e sottoposto ad una forza f dipendente dal tempo (in questo caso la variabile "tempo" si denota con t invece che con x), dalla posizione e dalla velocità (denotate rispettivamente con x e con \dot{x}). Ovviamente, non è l'unica interpretazione fisica: un'equazione del second'ordine ne può avere molte altre o, più precisamente, molti fenomeni fisici (e non solo di dinamica) sono governati da equazioni differenziali del second'ordine (e non solo del second'ordine).

Più in generale, un'equazione differenziale di ordine n (in forma normale) è un'espressione del tipo

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}),$$

dove $f: A \rightarrow \mathbb{R}$ è una funzione continua da un aperto A di \mathbb{R}^{n+1} in \mathbb{R} . Una soluzione è una funzione y definita in un intervallo I , con derivata n -esima continua in I e tale che, $\forall x \in I$,

$$(x, y(x), y'(x), \dots, y^{(n-1)}(x)) \in A$$

e

$$y^{(n)}(x) = f(x, y(x), y'(x), \dots, y^{(n-1)}(x)).$$

Come per le equazioni del prim'ordine, anche per quelle di ordine n il *problema di Cauchy* consiste nella ricerca delle soluzioni che “passano” per un punto assegnato dell'aperto A in cui è definita la f . In altre parole, dato un punto $p = (x_0, y_0, y_1, \dots, y_{n-1}) \in A$, tra tutte le soluzioni y dell'equazione

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}),$$

si cerca quella che verifica (o quelle che verificano) le n condizioni

$$y(x_0) = y_0, y'(x_0) = y_1, \dots, y^{(n-1)}(x_0) = y_{n-1}.$$

Anche per le equazioni di ordine n si può definire il concetto di soluzione massimale e vale ancora un teorema di esistenza e unicità. Ci limitiamo a dire che se la funzione di $n + 1$ variabili $(x, y, y', \dots, y^{(n-1)}) \mapsto f(x, y, y', \dots, y^{(n-1)})$ è continua e se è derivabile rispetto alle n variabili $y, y', \dots, y^{(n-1)}$ con derivate continue, allora il problema di Cauchy ammette una ed una sola soluzione massimale. Tornando all'esempio fisico dell'equazione di moto di un punto vincolato ad una retta, ciò significa che se ad un certo istante t_0 si assegna la posizione x_0 e la velocità \dot{x}_0 , il moto è determinato.

Esaminiamo ora delle **particolari equazioni del secondo ordine**.

Caso 1: Equazioni della forma

$$y'' = f(x, y')$$

(cioè f non dipende esplicitamente da y) con f continua nelle due variabili e derivabile con continuità rispetto alla seconda variabile. Si procede trasformando l'equazione del second'ordine in una del prim'ordine come è illustrato negli esempi che seguono

Esempio. Consideriamo l'equazione

$$y'' = \frac{y'}{x}.$$

Essa è un'equazione lineare (vedi dopo la trattazione generale di tali equazioni) e le soluzioni massimali sono definite o per $x > 0$ o per $x < 0$. Consideriamo ad esempio il caso $x > 0$. Sia $y(x)$ una soluzione dell'equazione in $(0, +\infty)$. Poiché nel secondo membro non compare esplicitamente la variabile y , la funzione $z(x) = y'(x)$ soddisfa l'equazione del prim'ordine

$$z' = \frac{z}{x}.$$

Perciò

$$z(x) = c_1 x$$

per qualche costante $c_1 \in \mathbb{R}$, da cui

$$y(x) = c_1 \frac{x^2}{2} + c_2$$

con $c_2 \in \mathbb{R}$.

Esempio. Consideriamo il problema di Cauchy

$$\begin{cases} y'' = 3(x y')^2 \\ y(1) = 1 \\ y'(1) = -1. \end{cases}$$

Questo problema ha una e una sola soluzione massimale y definita in un opportuno intorno del punto $x_0 = 1$. Ponendo $y'(x) = z(x)$ si ottiene che z è la soluzione del problema di Cauchy

$$\begin{cases} z' = 3(x z)^2 \\ z(1) = -1. \end{cases}$$

Risolvendo l'equazione si ottiene $z(x) = -\frac{1}{x^3 + c_1}$ per qualche costante reale c_1 da cui, essendo $z(1) = -1$, si ricava $c_1 = 0$ e quindi

$$z(x) = y'(x) = -\frac{1}{x^3}.$$

Di conseguenza,

$$y(x) = \frac{1}{2x^2} + c_2$$

ed essendo $y(1) = 1$ si ottiene infine

$$y(x) = \frac{1}{2} \left(\frac{1}{x^2} + 1 \right)$$

che, nell'intervallo $(0, +\infty)$ è la soluzione massimale cercata. Osserviamo che la stessa equazione differenziale con condizioni iniziali $y(x_0) = y_0$, $y'(x_0) = 0$ ha come soluzione (che si vede per immediata sostituzione) la funzione costante $y(x) \equiv y_0$.

In maniera analoga si possono trattare le equazioni di ordine n del tipo $y^{(n)} = f(x, y^{(n-1)})$. Ad esempio, lasciamo per esercizio la risoluzione dell'equazione del terz'ordine $y^{(3)} = (xy'')^2$.

Caso 2: Equazioni della forma

$$y'' = f(y)$$

con f di classe C^1 . Questa equazione, nell'interpretazione fisica come equazione fondamentale della dinamica, rappresenta il caso interessante in cui la forza dipende solo dalla posizione. Illustriamo con un esempio un metodo per determinare una soluzione dell'equazione.

Esempio. Consideriamo il problema di Cauchy

$$\begin{cases} y'' = -2 \operatorname{sen} y \cos^3 y \\ y(1) = 0 \\ y'(1) = 1. \end{cases}$$

Essendo soddisfatte le ipotesi del teorema di esistenza e unicità, il problema ha una e una sola soluzione massimale. Se $y(x)$ è questa soluzione, poiché $y'(1) \neq 0$ si ha $y'(x) \neq 0$ in un intorno di $x_0 = 1$. Moltiplicando entrambi i membri dell'equazione per $y'(x)$ e integrando si ottiene

$$\int y''(x)y'(x) dx = -2 \int \operatorname{sen} y(x) \cos^3 y(x)y'(x) dx$$

da cui

$$\frac{y'(x)^2}{2} = \frac{\cos^4 y(x)}{2} + c_1.$$

Dalle condizioni iniziali si ricava $c_1 = 0$ e quindi, essendo $y'(1) > 0$, si ha che la soluzione $y(x)$ soddisfa l'equazione del prim'ordine (a variabili separabili)

$$y' = \cos^2 y$$

Di conseguenza y è tale che

$$\operatorname{tang} y(x) = x + c_2$$

per un'opportuna costante c_2 . Ponendo $y(1) = 0$, si ottiene $0 = 1 + c_2$ e quindi

$$y(x) = \operatorname{arctang}(x - 1)$$

12^a settimana - dal 27.5.19

Equazioni lineari del second'ordine.

Un'equazione differenziale del second'ordine si dice *lineare* se è del tipo

$$y'' + a_1(x)y' + a_0(x)y = b(x),$$

dove a_0 , a_1 e b sono funzioni continue in un intervallo I . Le funzioni a_0 e a_1 si dicono i *coefficienti* dell'equazione e b rappresenta il *termine noto*. Quando $b(x) \equiv 0$, l'equazione si dice *omogenea*.

Più in generale un'equazione differenziale lineare di ordine n sarà del tipo

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + a_{n-2}(x)y^{(n-2)} + \dots + a_1(x)y' + a_0(x)y = b(x),$$

con a_0, \dots, a_{n-1} e b funzioni continue in un intervallo I .

Per le equazioni lineari vale il teorema di esistenza e unicità e si potrebbe dimostrare che ogni soluzione massimale è *globale* (o *persistente*), il che ricordiamo significa che è definita in tutto l'intervallo I in cui sono definite le funzioni a_0, \dots, a_{n-1} e b .

L'integrale generale di un'equazione lineare del second'ordine a coefficienti continui è descritto dal seguente teorema. Un analogo risultato è valido per le equazioni di ordine n .

Teorema. *Tutte le soluzioni dell'equazione differenziale lineare non omogenea*

$$y'' + a_1(x)y' + a_0(x)y = b(x)$$

si ottengono sommando ad una soluzione dell'equazione non omogenea tutte le possibili soluzioni dell'equazione omogenea associata. In altre parole, se \bar{y} è una soluzione dell'equazione non omogenea, ogni altra soluzione è del tipo $y = \bar{y} + u$, dove u è una soluzione dell'equazione omogenea associata.

Dimostrazione. La dimostrazione è analoga a quella fatta per le equazioni lineari del prim'ordine. La inseriamo per completezza. Mostriamo prima che, fissata una soluzione (detta particolare) \bar{y} dell'equazione non omogenea, ogni funzione del tipo $\bar{y} + u$, dove u risolve l'equazione omogenea $y'' + a_1(x)y' + a_0(x)y = 0$, è ancora una soluzione dell'equazione non omogenea. Per la linearità degli operatori di derivazione, si ha infatti

$$\begin{aligned} (\bar{y}(x) + u(x))'' + a_1(x)(\bar{y}(x) + u(x))' + a_0(x)(\bar{y}(x) + u(x)) = \\ \bar{y}''(x) + a_1(x)\bar{y}'(x) + a_0(x)\bar{y}(x) + u''(x) + a_1(x)u'(x) + a_0(x)u(x) = \end{aligned}$$

$$b(x) + 0 = b(x).$$

Rimane da provare che se \tilde{y} è una qualunque soluzione dell'equazione non omogenea, allora la differenza $\tilde{u} := \tilde{y} - \bar{y}$ è una soluzione dell'omogenea. Risulta infatti

$$(\tilde{y}(x) - \bar{y}(x))'' + a_1(x)(\tilde{y}(x) - \bar{y}(x))' + a_0(x)(\tilde{y}(x) - \bar{y}(x)) = b(x) - b(x) = 0. \quad \square$$

Denotiamo con $C^\infty(\mathbb{R})$ l'insieme costituito dalle funzioni di classe C^∞ da \mathbb{R} in sé. Osserviamo che due funzioni di $C^\infty(\mathbb{R})$ si possono sommare ottenendo ancora una funzione di $C^\infty(\mathbb{R})$. Inoltre, se si moltiplica una funzione di $C^\infty(\mathbb{R})$ per uno scalare reale (ossia per una costante appartenente ad \mathbb{R}) si ottiene ancora una funzione dello stesso insieme. Si ha così quello che viene chiamato uno *spazio vettoriale* sui reali (i reali si dicono gli scalari e gli elementi dello spazio i vettori). In tale spazio c'è un vettore che è neutro rispetto alla somma (cioè, sommato ad un qualunque vettore dà il vettore stesso): è la funzione identicamente nulla (chiamata *zero dello spazio*).

Definizione. Si dice che due funzioni $y_1, y_2 \in C^\infty(\mathbb{R})$ sono *linearmente indipendenti* se dall'uguaglianza

$$c_1 y_1(x) + c_2 y_2(x) = 0, \quad \forall x \in \mathbb{R}$$

segue $c_1 = c_2 = 0$, ossia se l'unica combinazione lineare che dà la funzione (identicamente) nulla è quella con i coefficienti tutti nulli.

Esempio. Proviamo che se λ_1 e λ_2 sono due numeri reali distinti, allora le funzioni $e^{\lambda_1 x}$ e $e^{\lambda_2 x}$ sono linearmente indipendenti. Supponiamo infatti che la funzione

$$y(x) := c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}$$

sia identicamente nulla. Di conseguenza, lo è anche la sua derivata. In particolare si ha $y(0) = 0$ e $y'(0) = 0$, da cui si ottiene il sistema (di due equazioni in due incognite)

$$\begin{cases} c_1 + c_2 = 0 \\ \lambda_1 c_1 + \lambda_2 c_2 = 0 \end{cases}$$

che, come si verifica subito, ha come unica soluzione $c_1 = c_2 = 0$ essendo $\lambda_1 \neq \lambda_2$.

Esempio. Mostriamo che le funzioni $\cos \omega x$ e $\sin \omega x$ (dove $\omega \in \mathbb{R}$) sono linearmente indipendenti. Supponiamo infatti che la funzione

$$y(x) := c_1 \cos \omega x + c_2 \sin \omega x$$

sia (identicamente) nulla. Poiché $y(x)$ è zero per ogni x , deve essere anche per $x = 0$. Ponendo $x = 0$ si ottiene $c_1 = 0$. Per provare che anche il coefficiente c_2 è nullo, basta porre $x = \pi/(2\omega)$.

Esercizio. Dato $\lambda \in \mathbb{R}$, provare che le funzioni $e^{\lambda x}$ e $xe^{\lambda x}$ sono linearmente indipendenti.

Esercizio. Dati due numeri reali α e β , provare che se $\beta \neq 0$, allora le due funzioni $e^{\alpha x} \cos \beta x$ e $e^{\alpha x} \sin \beta x$ sono linearmente indipendenti.

Studiamo ora in dettaglio il caso di un'equazione lineare del second'ordine a **coefficienti costanti** con termine noto che supporremo per semplicità definito in tutto \mathbb{R} e di classe C^∞ , ossia consideriamo un'equazione della forma

$$y'' + a_1 y' + a_0 y = b(x),$$

con $a_0, a_1 \in \mathbb{R}$ e $b : \mathbb{R} \rightarrow \mathbb{R}$ di classe C^∞ .

In base al teorema precedente, il problema di risolvere l'equazione lineare presa in esame si scinde in due sottoproblemi:

- 1) risolvere l'equazione omogenea associata;
- 2) trovare almeno una soluzione dell'equazione non omogenea.

Cominciamo perciò col **risolvere l'equazione omogenea associata**.

Il risultato che segue è una facile conseguenza dei teoremi di esistenza e unicità. Per brevità ne omettiamo la dimostrazione.

Teorema (della dimensione). *L'insieme delle soluzioni (massimali) di un'equazione differenziale omogenea di ordine 2 a coefficienti costanti (o, più in generale, a coefficienti di classe C^∞) è un sottospazio vettoriale 2-dimensionale di $C^\infty(\mathbb{R})$.*

Il teorema che segue descrive l'integrale generale dell'equazione omogenea e non è altro che una riformulazione del Teorema della dimensione.

Teorema. *Le soluzioni massimali dell'equazione differenziale lineare omogenea*

$$y'' + a_1 y' + a_0 y = 0$$

sono della forma

$$y(x) = c_1 y_1(x) + c_2 y_2(x)$$

dove c_1 e c_2 sono coefficienti in \mathbb{R} , e y_1 e y_2 sono due soluzioni linearmente indipendenti.

Osservazione. È facile verificare che l'applicazione $L: C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ definita da

$$Ly = y'' + a_1y' + a_0y$$

è lineare e che il suo nucleo, $\ker L$, non è altro che lo spazio vettoriale delle soluzioni dell'equazione differenziale omogenea. Il teorema della dimensione afferma perciò che $\ker L$ è uno spazio di dimensione 2.

Torniamo al problema di risolvere l'equazione omogenea associata ad una equazione lineare del second'ordine a coefficienti costanti. Come osservato in precedenza, il problema di trovare **tutte** le soluzioni dell'equazione omogenea è ricondotto (per il teorema della dimensione) a quello di trovarne **due** linearmente indipendenti.

Un ruolo fondamentale nella ricerca di due soluzioni linearmente indipendenti è giocato dal *polinomio caratteristico*

$$P(\lambda) = \lambda^2 + a_1\lambda + a_0$$

associato all'equazione omogenea.

Il seguente risultato fornisce una regola pratica per trovare due soluzioni linearmente indipendenti di un'equazione differenziale omogenea del second'ordine a coefficienti costanti. La sua dimostrazione risulta chiara se si introduce il concetto di soluzione complessa di un'equazione differenziale lineare, cosa che faremo subito dopo i quattro esempi seguenti.

Teorema (risoluzione delle equazioni differenziali del second'ordine lineari, omogenee, a coefficienti costanti). *Sia $y'' + a_1y' + a_0y = 0$ un'equazione omogenea del second'ordine a coefficienti (reali) costanti. Se il polinomio caratteristico $P(\lambda) = \lambda^2 + a_1\lambda + a_0$ ha due radici reali e distinte λ_1 e λ_2 , allora e^{λ_1x} e e^{λ_2x} sono due soluzioni linearmente indipendenti dell'equazione differenziale. Se $P(\lambda)$ ha una radice doppia $\hat{\lambda}$, allora due soluzioni linearmente indipendenti sono $e^{\hat{\lambda}x}$ e $xe^{\hat{\lambda}x}$. Se $P(\lambda)$ ha una radice complessa $\alpha + i\beta$ (con $\beta \neq 0$) allora ammette anche la radice coniugata $\alpha - i\beta$, e le funzioni reali $e^{\alpha x} \cos \beta x$ e $e^{\alpha x} \sin \beta x$ sono due soluzioni linearmente indipendenti.*

Esempio. Consideriamo l'equazione differenziale

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

Il polinomio caratteristico è $P(\lambda) = \lambda^2 - 1$ che ha le due radici reali e distinte $\lambda_1 = -1$ e $\lambda_2 = 1$. Pertanto ogni soluzione dell'equazione è del tipo

$$y(x) = c_1e^{-x} + c_2e^x,$$

con c_1 e c_2 costanti arbitrarie.

Esempio. Consideriamo l'equazione differenziale

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

Il polinomio caratteristico è $P(\lambda) = (\lambda - 1)^2$ che ha la sola radice (di molteplicità due) $\hat{\lambda} = 1$. Pertanto ogni soluzione dell'equazione è del tipo

$$y(x) = c_1 e^x + c_2 x e^x,$$

con c_1 e c_2 costanti arbitrarie.

Esempio (equazione del moto armonico). Consideriamo l'equazione differenziale

$$y'' + \omega^2 y = 0, \quad \omega > 0.$$

Il polinomio caratteristico è $P(\lambda) = \lambda^2 + \omega^2$ che ha le due radici complesse e coniugate $-i\omega$ e $+i\omega$. Pertanto, ogni soluzione è del tipo

$$y(x) = c_1 \cos \omega x + c_2 \sin \omega x,$$

con c_1 e c_2 costanti arbitrarie.

Da elementari considerazioni di trigonometria si deduce facilmente che ogni soluzione può essere scritta anche nella forma

$$y(x) = A \sin(\omega x + \varphi),$$

dove le costanti reali $A \geq 0$ e φ (dette, rispettivamente, *ampiezza* e *fase* della oscillazione) sono arbitrarie e ω (detta *pulsazione*) è assegnata.

Esempio (equazione delle oscillazioni smorzate). Consideriamo l'equazione differenziale

$$y'' + 2\epsilon y' + \omega^2 y = 0,$$

dove $0 < \epsilon < \omega$. Il polinomio caratteristico è $p(\lambda) = \lambda^2 + 2\epsilon\lambda + \omega^2$, le cui radici sono $-\epsilon \pm i\sqrt{\omega^2 - \epsilon^2}$. Dunque la soluzione generale è data da

$$y(x) = c_1 e^{-\epsilon x} \cos\left(\sqrt{\omega^2 - \epsilon^2} x\right) + c_2 e^{-\epsilon x} \sin\left(\sqrt{\omega^2 - \epsilon^2} x\right)$$

o, equivalentemente come osservato nell'esempio precedente, da

$$y(x) = A e^{-\epsilon x} \sin\left(\sqrt{\omega^2 - \epsilon^2} x + \varphi\right),$$

dove A e φ sono costanti arbitrarie.

=====

Cenno alle soluzioni complesse delle equazioni differenziali lineari omogenee. (Facoltativo)

Per lo studio delle equazioni differenziali lineari a coefficienti costanti è opportuno introdurre uno spazio più ampio di $C^\infty(\mathbb{R})$, cioè lo spazio vettoriale $C^\infty(\mathbb{R}, \mathbb{C})$ costituito dalle funzioni di classe C^∞ da \mathbb{R} in \mathbb{C} . Una funzione z di tale spazio si scrive nella forma $z(x) = \alpha(x) + i\beta(x)$ dove α e β , dette rispettivamente *parte reale* e *parte immaginaria* della funzione $z(x)$, appartengono a $C^\infty(\mathbb{R})$. La *derivata* di z è la funzione $z'(x) = \alpha'(x) + i\beta'(x)$, e quindi appartiene ancora allo spazio $C^\infty(\mathbb{R}, \mathbb{C})$. Ovviamente, ogni funzione di $C^\infty(\mathbb{R})$ può essere pensata anche in $C^\infty(\mathbb{R}, \mathbb{C})$ (con parte immaginaria nulla).

Si fa notare che le funzioni di $C^\infty(\mathbb{R}, \mathbb{C})$, non solo si possono moltiplicare per dei numeri reali, ma anche per dei numeri complessi, ottenendo ancora delle funzioni di classe C^∞ da \mathbb{R} in \mathbb{C} . Per questo motivo si usa dire che $C^\infty(\mathbb{R}, \mathbb{C})$ è uno *spazio vettoriale sui complessi* (o uno *spazio complesso*) e gli elementi di \mathbb{C} rappresentano gli *scalari* dello spazio. In maniera analoga a quanto fatto in $C^\infty(\mathbb{R})$, si può introdurre la nozione di funzioni linearmente indipendenti per gli elementi di $C^\infty(\mathbb{R}, \mathbb{C})$.

Una volta introdotto lo spazio $C^\infty(\mathbb{R}, \mathbb{C})$, diremo che una funzione $z \in C^\infty(\mathbb{R}, \mathbb{C})$ è una *soluzione complessa* dell'equazione a coefficienti costanti se

$$z''(x) + a_1z'(x) + a_0z(x) = 0$$

per ogni $x \in \mathbb{R}$.

Esercizio. Siano $z(x) = \alpha(x) + i\beta(x)$ e $\bar{z}(x) = \alpha(x) - i\beta(x)$ due funzioni complesse e coniugate di $C^\infty(\mathbb{R}, \mathbb{C})$. Allora si ha

$$\frac{z(x) + \bar{z}(x)}{2} = \alpha(x) \quad \text{e} \quad \frac{z(x) - \bar{z}(x)}{2i} = \beta(x).$$

Dedurre da ciò che se z e \bar{z} sono soluzioni di un'equazione differenziale omogenea, allora lo sono anche α e β (notiamo esplicitamente che $\alpha(x)$ e $\beta(x)$ sono funzioni reali).

Esercizio. Provare che se λ_1 e λ_2 sono due numeri complessi distinti, allora le funzioni $e^{\lambda_1 x}$ e $e^{\lambda_2 x}$ sono linearmente indipendenti (nello spazio $C^\infty(\mathbb{R}, \mathbb{C})$).

Esercizio. Dato $\lambda \in \mathbb{C}$, provare che le funzioni $e^{\lambda x}$ e $xe^{\lambda x}$ sono linearmente indipendenti (in $C^\infty(\mathbb{R}, \mathbb{C})$).

Cerchiamo di illustrare con un esempio la definizione di soluzione complessa.

Esempio. Consideriamo l'equazione differenziale

$$y'' + y = 0.$$

Proviamo che essa ammette soluzioni complesse. Facciamo vedere che ammette soluzioni del tipo $z(x) = e^{\mu x}$, dove μ è un numero complesso. Si ha $z'(x) = \mu e^{\mu x}$ e $z''(x) = \mu^2 e^{\mu x}$. Quindi $z(x)$ è soluzione se e solo se

$$(\mu^2 + 1)e^{\mu x} = 0, \quad \forall x \in \mathbb{R},$$

ossia (essendo $e^{\mu x} \neq 0$) se e solo se $\mu^2 + 1 = 0$, da cui si ricava $\mu = \pm i$. Pertanto e^{ix} e e^{-ix} sono soluzioni dell'equazione differenziale considerata, e sono le uniche del tipo $e^{\mu x}$. Si osservi che non solo

$$e^{ix} = \cos x + i \sin x \quad \text{e} \quad e^{-ix} = \cos x - i \sin x$$

sono soluzioni dell'equazione in esame, ma lo è anche $z(x) = c_1 e^{ix} + c_2 e^{-ix}$ qualunque siano $c_1, c_2 \in \mathbb{C}$. Ciò dipende dal fatto che l'equazione $y'' + y = 0$ è lineare omogenea, e quindi (come è facile verificare) la somma di due soluzioni è ancora una soluzione e se si moltiplica una soluzione per una costante si ottiene ancora una soluzione. L'insieme delle soluzioni dell'equazione $y'' + y = 0$ è dunque uno spazio vettoriale. È facile verificare che questo fatto è vero per una qualunque equazione differenziale lineare omogenea di ordine n .

Torniamo dall'esempio al caso generale. Una semplice verifica mostra che la funzione complessa $z(x) = e^{\mu x}$ è soluzione dell'equazione differenziale $y'' + a_1 y' + a_0 y = 0$ se e solo se μ è radice del polinomio, detto **polinomio caratteristico**,

$$P(\lambda) = \lambda^2 + a_1 \lambda + a_0$$

(ossia, se e solo se $\mu^2 + a_1 \mu + a_0 = 0$).

Prendiamo allora in esame le radici (in \mathbb{C}) del polinomio caratteristico. Osserviamo esplicitamente che, essendo il polinomio che stiamo considerando a coefficienti reali (infatti abbiamo supposto $a_0, a_1 \in \mathbb{R}$), se esso ha una radice complessa $\alpha + i\beta$, allora necessariamente ha anche la coniugata $\alpha - i\beta$. Il polinomio è di secondo grado; perciò i casi che si possono presentare sono i seguenti:

- $P(\lambda)$ ha due radici reali e distinte λ_1, λ_2 . Allora le due funzioni $y_1(x) = e^{\lambda_1 x}$ e $y_2(x) = e^{\lambda_2 x}$ sono soluzioni dell'equazione differenziale e, per quanto osservato in precedenza, sono linearmente indipendenti. In questo caso ogni soluzione (complessa) y dell'equazione differenziale è della forma

$$y(x) = c_1 e^{\lambda_1 x} + c_2 e^{\lambda_2 x}$$

con $c_1, c_2 \in \mathbb{C}$. In particolare, se $c_1, c_2 \in \mathbb{R}$ si ottengono tutte le soluzioni reali.

- $P(\lambda)$ ha un'unica radice reale $\hat{\lambda}$ (ovviamente) con molteplicità due. Allora le due funzioni $y_1(x) = e^{\hat{\lambda}x}$ e $y_2(x) = xe^{\hat{\lambda}x}$ sono due soluzioni linearmente indipendenti dell'equazione. In questo caso ogni soluzione complessa [reale] y dell'equazione differenziale è della forma

$$y(x) = c_1 e^{\hat{\lambda}x} + c_2 x e^{\hat{\lambda}x},$$

con $c_1, c_2 \in \mathbb{C}$ [$c_1, c_2 \in \mathbb{R}$].

- $P(\lambda)$ ha due radici complesse e coniugate $\alpha + i\beta$ e $\alpha - i\beta$. Allora le funzioni $e^{(\alpha+i\beta)x}$ e $e^{(\alpha-i\beta)x}$ sono due soluzioni complesse dell'equazione differenziale e (vedi esercizio) sono linearmente indipendenti. In questo caso ogni soluzione complessa y dell'equazione differenziale è della forma

$$y(x) = c_1 e^{(\alpha+i\beta)x} + c_2 e^{(\alpha-i\beta)x}$$

con $c_1, c_2 \in \mathbb{C}$. D'altra parte, usando le formule di Eulero introdotte sopra e tenendo conto che l'insieme delle soluzioni di un'equazione lineare è uno spazio vettoriale, si ottiene che l'equazione ammette anche come soluzioni le due funzioni reali $e^{\alpha x} \cos \beta x$ e $e^{\alpha x} \sin \beta x$. Si può provare che anch'esse sono linearmente indipendenti. In particolare, perciò, tutte le soluzioni reali dell'equazione sono della forma

$$y(x) = e^{\alpha x} (c_1 \cos \beta x + c_2 \sin \beta x),$$

con $c_1, c_2 \in \mathbb{R}$.

Fine cenni sulle soluzioni complesse.

=====

Una volta trovato l'integrale generale dell'equazione differenziale omogenea, possiamo passare al secondo problema che avevamo da risolvere cioè quello di determinare almeno **una soluzione dell'equazione non omogenea**, la cosiddetta soluzione particolare.

Il metodo generale per determinare una soluzione particolare di un'equazione lineare (anche non a coefficienti costanti) è il metodo di variazione delle costanti che illustreremo nel seguito. Quando però l'equazione differenziale è a coefficienti costanti e quando il termine noto è di un certo tipo, si può usare un metodo più rapido per trovare una soluzione della non omogenea. Il metodo consiste nel cercare una soluzione in una classe di funzioni dello stesso tipo del termine noto. C'è un motivo teorico che giustifica questo procedimento, ma lo spiegarlo esula dai nostri scopi. In questi appunti perciò il metodo che descriveremo costituisce semplicemente una "regola pratica".

Supponiamo pertanto che il termine noto sia della forma

$$b(x) = e^{\alpha x}(p(x) \cos \beta x + q(x) \sin \beta x),$$

dove α e β sono due numeri reali (eventualmente nulli), mentre $p(x)$ e $q(x)$ sono polinomi (eventualmente di grado zero, cioè costanti). Si potrebbe dimostrare che, in questo caso, otteniamo una soluzione particolare ancora dello stesso tipo. Per decidere in che forma cercarla è fondamentale il ruolo giocato dal numero complesso $\alpha + i\beta$.

Impariamo, innanzi tutto, a riconoscere quando il termine noto $b(x)$ si presenta in una di tali forme e a determinarne il corrispondente numero $\lambda = \alpha + i\beta$. Ecco alcuni esempi:

$b(x)$	$\alpha + i\beta$	$p(x)$	$q(x)$
$\sin x$	i	0	1
$x^2 e^{-2x}$	-2	x^2	0
-3	0	-3	0
$-e^x \cos 2x$	$1 + 2i$	-1	0
$x e^{-2x} \sin \pi x$	$-2 + \pi i$	0	x
$x^3 - x$	0	$x^3 - x$	0
$2e^x$	1	2	0
$2 \sin \omega x - x \cos \omega x$	$i\omega$	$-x$	2

Regola pratica (per determinare una soluzione particolare). Supponiamo che il termine noto $b(x)$ sia del tipo

$$e^{\alpha x}(p(x) \cos \beta x + q(x) \sin \beta x),$$

dove $p(x)$ e $q(x)$ sono due polinomi di grado k_1 e k_2 rispettivamente (non occorre che abbiano lo stesso grado) e $\alpha, \beta \in \mathbb{R}$.

- Se $\alpha + i\beta$ non è radice del polinomio caratteristico, allora si cerca una soluzione particolare della forma

$$e^{\alpha x}(r(x) \cos \beta x + s(x) \sin \beta x),$$

dove $r(x)$ e $s(x)$ sono polinomi di grado $k = \max\{k_1, k_2\}$, i cui coefficienti sono da determinare.

- Se $\alpha + i\beta$ è radice semplice del polinomio caratteristico, allora si cerca una soluzione moltiplicando per x la forma relativa al caso precedente.

- Se $\alpha + i\beta$ è radice doppia del polinomio caratteristico, allora si cerca una soluzione moltiplicando per x^2 la forma relativa al primo caso (cioè quella in cui $\alpha + i\beta$ non è radice).

In tutti i casi si tratta perciò di determinare i coefficienti dei polinomi r e s . Come già detto sopra, si potrebbe dimostrare che ciò è sempre possibile.

Illustriamo ora con qualche esempio l'applicazione di questo metodo.

Esempio. Determiniamo una soluzione particolare dell'equazione

$$y'' - y = 2e^{3x}.$$

Il polinomio caratteristico ha le due radici reali $\lambda_1 = -1$ e $\lambda_2 = 1$. Poiché $\alpha = 3$ non è una di tali radici, cerchiamo una soluzione dell'equazione non omogenea nella forma (si ha $\beta = 0$, $p(x) \equiv 2$)

$$\bar{y}(x) = ae^{3x}, a \in \mathbb{R}.$$

Derivando e sostituendo nell'equazione si ottiene

$$9ae^{3x} - ae^{3x} = 2e^{3x},$$

che porta all'equazione algebrica (lineare) $8a = 2$. In conclusione, si ricava

$$\bar{y}(x) = \frac{1}{4}e^{3x}.$$

Se consideriamo invece l'equazione

$$y'' - y = e^x,$$

ci troviamo nel caso in cui $\alpha = 1$ è radice semplice del polinomio caratteristico. Cerchiamo perciò

$$\bar{y}(x) = axe^x, a \in \mathbb{R}.$$

Con calcolo analogo al precedente, si ottiene

$$2ae^x + axe^x - axe^x = e^x,$$

da cui $2ae^x = e^x$ e quindi $a = 1/2$.

Consideriamo infine l'equazione

$$y'' - y = (x + 3)e^x.$$

Questa volta il polinomio p non è costante ma è il polinomio di primo grado $p(x) = x + 3$. Si pone perciò

$$\bar{y}(x) = x(ax + b)e^x, \quad a, b \in \mathbb{R}.$$

Derivando e sostituendo si ha

$$(2ax + 2a + b)e^x + (ax^2 + 2ax + bx + b)e^x - x(ax + b)e^x = (x + 3)e^x,$$

da cui si ottiene il sistema lineare nelle due incognite a e b

$$\begin{cases} 4a = 1 \\ 2a + 2b = 3 \end{cases}$$

che ha per soluzione $a = 1/4$ e $b = 5/4$.

Esempio. Determiniamo una soluzione particolare dell'equazione

$$y'' - 2y' + y = e^x.$$

In questo caso, $\alpha = 1$ è radice di molteplicità 2 del polinomio caratteristico. Ponendo

$$\bar{y}(x) = ax^2e^x, \quad a \in \mathbb{R},$$

e sostituendo, si ricava

$$\bar{y}(x) = (1/2)x^2e^x.$$

Esempio. Determiniamo una soluzione particolare dell'equazione

$$y'' + y = 2e^{3x} \cos x.$$

Poiché $3 + i$ non è radice del polinomio caratteristico, cerchiamo una soluzione della non omogenea nella forma

$$\bar{y}(x) = e^{3x}(a \cos x + b \sin x), \quad a, b \in \mathbb{R}.$$

Procedendo come in precedenza, si ottiene il seguente sistema algebrico nell'incognite a e b

$$\begin{cases} 9a + 6b = 2 \\ -6a + 9b = 0 \end{cases}$$

che ha per soluzione $a = 2/13$ e $b = 4/39$. Si ha perciò

$$\bar{y}(x) = e^{3x} \left(\frac{2}{13} \cos x + \frac{4}{39} \sin x \right).$$

Esempio (Oscillatore armonico con termine forzante, senza o con risonanza). Consideriamo l'equazione non omogenea

$$y'' + \omega^2 y = \text{sen } \beta x,$$

dove ω e β sono costanti positive. Come abbiamo visto l'integrale generale dell'equazione omogenea associata è

$$c_1 \cos \omega x + c_2 \text{sen } \omega x.$$

Supponiamo $\beta \neq \omega$. In questo caso, poiché βi non è radice del polinomio caratteristico, una soluzione particolare dell'equazione non omogenea si può cercare nella forma

$$\bar{y}(x) = a \cos \beta x + b \text{sen } \beta x.$$

Svolgendo i calcoli si trova $a = 0$ e $b = 1/(\omega^2 - \beta^2)$ e quindi l'integrale generale dell'equazione considerata è

$$y(x) = c_1 \cos \omega x + c_2 \text{sen } \omega x + \frac{\text{sen } \beta x}{\omega^2 - \beta^2}.$$

Quando $\beta = \omega$ si dice che c'è **risonanza**. L'equazione non omogenea diventa perciò

$$y'' + \omega^2 y = \text{sen } \omega x$$

ed essendo ωi radice di molteplicità 1 del polinomio caratteristico una soluzione particolare è della forma

$$\bar{y}(x) = x(a \cos \omega x + b \text{sen } \omega x).$$

Derivando due volte e sostituendo si ottiene $a = -1/(2\omega)$ e $b = 0$. Di conseguenza, l'integrale generale dell'equazione nel caso di risonanza risulta

$$y(x) = c_1 \cos \omega x + c_2 \text{sen } \omega x - x \frac{\cos \omega x}{2\omega}.$$

Osserviamo che in presenza di risonanza le soluzioni risultano sempre **non limitate** a causa del fattore x (vibrazione forzata).

(Facoltativo) Possiamo determinare una soluzione particolare dell'equazione $y'' + \omega^2 y = \text{sen } \omega x$ anche per altra via.

Fra tutte le soluzioni dell'equazione non in risonanza, cioè tra tutte le soluzioni della forma $c_1 \cos \omega x + c_2 \text{sen } \omega x + \frac{\text{sen } \beta x}{\omega^2 - \beta^2}$, scegliamo la soluzione y_β che soddisfa

le condizioni iniziali $y_\beta(0) = y'_\beta(0) = 0$. Calcolando le costanti c_1 e c_2 otteniamo $c_1 = 0$ e $c_2 = -\frac{\beta}{\omega(\omega^2 - \beta^2)}$, da cui

$$y_\beta(x) = \frac{\beta \operatorname{sen} \omega x - \omega \operatorname{sen} \beta x}{\omega(\beta^2 - \omega^2)}.$$

Questa soluzione è definita per ogni valore $\beta \neq \omega$; possiamo però cercare di calcolare (se esiste) il limite per β tendente a ω di y_β . Si ha

$$\lim_{\beta \rightarrow \omega} y_\beta(x) = \lim_{\beta \rightarrow \omega} \frac{\beta \operatorname{sen} \omega x - \omega \operatorname{sen}(\beta x)}{\omega(\beta^2 - \omega^2)} = \frac{1}{2\omega^2} \lim_{\beta \rightarrow \omega} \frac{\beta \operatorname{sen} \omega x - \omega \operatorname{sen}(\beta x)}{\beta - \omega}$$

Applicando la regola di de l'Hôpital si trova, derivando rispetto a β ,

$$\lim_{\beta \rightarrow \omega} y_\beta(x) = \frac{1}{2\omega^2} \lim_{\beta \rightarrow \omega} \frac{\operatorname{sen} \omega x - x\omega \cos \beta x}{1} = \frac{\operatorname{sen} \omega x - x\omega \cos \omega x}{2\omega^2} = y_\omega(x)$$

Con un calcolo diretto si può verificare che la funzione

$$y_\omega = \frac{\operatorname{sen} \omega x}{2\omega^2} - x \frac{\cos \omega x}{2\omega}$$

soddisfa l'equazione $y'' + \omega^2 y = \operatorname{sen} \omega x$ (e le condizioni iniziali $y_\omega(0) = y'_\omega(0) = 0$). Inoltre, essendo la funzione $\frac{\operatorname{sen} \omega x}{2\omega^2}$ soluzione dell'equazione omogenea, si riottiene la soluzione $-x \frac{\cos \omega x}{2\omega}$ della non omogenea, già trovata sopra in altro modo.

Esempio. Tra tutte le soluzioni dell'equazione differenziale

$$y'' = xe^{-x},$$

determinare quella che verifica le condizioni $y(0) = 0$ e $y'(0) = 0$. Si tratta perciò di risolvere il seguente problema di Cauchy

$$\begin{cases} y'' = xe^{-x} \\ y(0) = 0 \\ y'(0) = 0. \end{cases}$$

L'equazione omogenea associata è $y'' = 0$ e il suo polinomio caratteristico, $P(\lambda) = \lambda^2$, ha due radici coincidenti: $\lambda_1 = 0$ e $\lambda_2 = 0$. Quindi, la soluzione generale dell'equazione omogenea è $u(x) = c_1 + c_2 x$. Occorre trovare una soluzione particolare dell'equazione non omogenea. Osserviamo che $\alpha = -1$ non è radice del polinomio caratteristico. Si cerca quindi una soluzione del tipo $\bar{y}(x) = (ax + b)e^{-x}$. Si ha

$$\bar{y}''(x) = (b - 2a)e^{-x} + axe^{-x}.$$

Quindi $\bar{y}(x)$ è soluzione se (e solo se) è verificata la condizione

$$(b - 2a)e^{-x} + (a - 1)xe^{-x} = 0, \quad \forall x \in \mathbb{R},$$

ossia se (e solo se) $b - 2a = 0$ e $a - 1 = 0$ (per quanto riguarda il “solo se”, anche se non ci interessa, ricordarsi che le funzioni e^{-x} e xe^{-x} sono linearmente indipendenti). La funzione $\bar{y}(x) = (x + 2)e^{-x}$ è dunque una soluzione dell’equazione non omogenea. Pertanto, la soluzione generale dell’equazione in esame è

$$y(x) = c_1 + c_2x + (x + 2)e^{-x}.$$

Determiniamo ora c_1 e c_2 in modo che siano soddisfatte le condizioni assegnate. Si ha $y'(x) = c_2 + e^{-x} - (x + 2)e^{-x}$, e quindi $y(0) = c_1 + 2$ e $y'(0) = c_2 - 1$. Ponendo $y(0) = 0$ e $y'(0) = 0$ si ricava $c_1 = -2$ e $c_2 = 1$. Dunque, la soluzione che verifica il problema di Cauchy assegnato è

$$y(x) = -2 + x + (x + 2)e^{-x}.$$

Esercizio. Provare che una soluzione particolare dell’equazione differenziale

$$y'' + a_1y' + a_0y = b_1(x) + b_2(x)$$

si può ottenere sommando una soluzione di $y'' + a_1y' + a_0y = b_1(x)$ con una soluzione di $y'' + a_1y' + a_0y = b_2(x)$.

Esempio. Troviamo tutte le soluzioni dell’equazione

$$y'' + 9y = \sin 3x + e^x.$$

Il polinomio caratteristico ha le due radici complesse (e coniugate) $3i$ e $-3i$. La soluzione dell’equazione omogenea è perciò $c_1 \cos 3x + c_2 \sin 3x$. Per l’esercizio precedente per determinare una soluzione dell’equazione non omogenea possiamo cercare separatamente una soluzione particolare \bar{y}_1 di $y'' + 9y = \sin 3x$ e una \bar{y}_2 di $y'' + 9y = e^x$ e poi sommarle. Cerchiamo \bar{y}_1 nella forma (si ha $\alpha = 0, \beta = 3, p(x) \equiv 0, q(x) \equiv 1$)

$$\bar{y}_1(x) = x(a \cos 3x + b \sin 3x), \quad a, b \in \mathbb{R}.$$

Si trova $a = -(1/6)$ e $b = 0$, da cui $\bar{y}_1(x) = -(1/6)x \cos 3x$. Cerchiamo ora \bar{y}_2 nella forma (si ha $\alpha = 1, \beta = 0, p(x) \equiv 1$)

$$\bar{y}_2(x) = ae^x, \quad a \in \mathbb{R}.$$

Si trova $a = (1/10)$ e quindi $\bar{y}_2(x) = (1/10)e^x$. In definitiva, l'integrale generale dell'equazione data è

$$y(x) = c_1 \cos 3x + c_2 \sin 3x - \frac{1}{6} x \cos 3x + \frac{1}{10} e^x, c_1, c_2 \in \mathbb{R}.$$

Nel caso in cui il termine noto b dell'equazione differenziale non sia del tipo considerato sopra, per trovare una soluzione particolare si usa il metodo di **variazione delle costanti**. Come già osservato per le equazioni del prim'ordine, il metodo consiste nel cercare una soluzione dell'equazione non omogenea, pensando "variabili" le costanti c_1 e c_2 che compaiono nell'integrale generale dell'omogenea. In altre parole, si cerca una soluzione dell'equazione non omogenea nella forma

$$\bar{y}(x) = c_1(x)y_1(x) + c_2(x)y_2(x)$$

dove y_1 e y_2 sono due soluzioni linearmente indipendenti dell'equazione omogenea.

Derivando, si ottiene

$$\bar{y}'(x) = c_1'(x)y_1(x) + c_1(x)y_1'(x) + c_2'(x)y_2(x) + c_2(x)y_2'(x).$$

Imponendo la condizione $c_1'(x)y_1(x) + c_2'(x)y_2(x) = 0$ e derivando una seconda volta si ha

$$\bar{y}''(x) = c_1'(x)y_1'(x) + c_1(x)y_1''(x) + c_2'(x)y_2'(x) + c_2(x)y_2''(x).$$

Sostituendo le espressioni trovate nell'equazione si ottiene

$$c_1'(x)y_1'(x) + c_1(x)y_1''(x) + c_2'(x)y_2'(x) + c_2(x)y_2''(x) + a_1(c_1(x)y_1'(x) + c_2(x)y_2'(x)) + a_0(c_1(x)y_1(x) + c_2(x)y_2(x)),$$

da cui, tenendo conto che y_1 e y_2 sono soluzioni dell'omogenea, si ricava

$$c_1'(x)y_1'(x) + c_2'(x)y_2'(x) = b(x).$$

Perciò \bar{y} è soluzione dell'equazione differenziale se risolve il sistema

$$\begin{cases} c_1'(x)y_1(x) + c_2'(x)y_2(x) = 0 \\ c_1'(x)y_1'(x) + c_2'(x)y_2'(x) = b(x). \end{cases}$$

Questo sistema ha sempre una e una sola soluzione $c_1'(x)$, $c_2'(x)$ (dipende dal fatto che y_1 e y_2 sono linearmente indipendenti) data da

$$c_1'(x) = \frac{\det \begin{pmatrix} 0 & y_2(x) \\ b(x) & y_2'(x) \end{pmatrix}}{\det \begin{pmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{pmatrix}}, \quad c_2'(x) = \frac{\det \begin{pmatrix} y_1(x) & 0 \\ y_1'(x) & b(x) \end{pmatrix}}{\det \begin{pmatrix} y_1(x) & y_2(x) \\ y_1'(x) & y_2'(x) \end{pmatrix}}$$

A questo punto, prendendo delle primitive di c'_1 e di c'_2 , si ricava l'espressione della soluzione \bar{y} .

Nei due esempi che seguono si applica il metodo di variazione delle costanti.

Esempio. Consideriamo l'equazione

$$y'' + y = \frac{1}{\cos x}.$$

Ovviamente le soluzioni massimali sono definite negli intervalli tra due zeri successivi del coseno. Per semplicità, restringiamoci all'intervallo $(-\pi/2, \pi/2)$. L'integrale generale dell'equazione omogenea associata è

$$c_1 \cos x + c_2 \sin x, \quad c_1, c_2 \in \mathbb{R}.$$

Cerchiamo una soluzione particolare nella forma

$$\bar{y}(x) = c_1(x) \cos x + c_2(x) \sin x.$$

Si ottiene il sistema nelle incognite $c'_1(x)$ e $c'_2(x)$

$$\begin{cases} c'_1(x) \cos x + c'_2(x) \sin x = 0 \\ -c'_1(x) \sin x + c'_2(x) \cos x = \frac{1}{\cos x}. \end{cases}$$

Moltiplicando la prima equazione per $\sin x$, la seconda per $\cos x$ e sommando si ottiene $c'_2(x) = 1$ da cui, sostituendo nella prima equazione,

$$c'_1(x) = -\frac{\sin x}{\cos x}.$$

Integrando e ricordando che $x \in (-\pi/2, \pi/2)$, si ricava

$$c_1(x) = \log \cos x, \quad c_2(x) = x.$$

Pertanto, una soluzione particolare è

$$\bar{y}(x) = (\log \cos x) \cos x + x \sin x.$$

Esempio. Consideriamo l'equazione

$$y'' - y = \frac{1}{x}.$$

Ovviamente le soluzioni sono definite o per $x > 0$ o per $x < 0$. L'integrale generale dell'equazione omogenea associata è

$$c_1 e^x + c_2 e^{-x}, \quad c_1, c_2 \in \mathbb{R}.$$

Cerchiamo una soluzione particolare nella forma

$$\bar{y}(x) = c_1(x) e^x + c_2(x) e^{-x}.$$

Si ottiene il sistema nelle incognite $c'_1(x)$ e $c'_2(x)$

$$\begin{cases} c'_1(x) e^x + c'_2(x) e^{-x} = 0 \\ c'_1(x) e^x - c'_2(x) e^{-x} = \frac{1}{x}. \end{cases}$$

Risolvendo il sistema, si ricava

$$c'_1(x) = \frac{e^{-x}}{2x}, \quad c'_2(x) = -\frac{e^x}{2x},$$

per cui una soluzione particolare è

$$\bar{y}(x) = e^x \int \frac{e^{-x}}{2x} dx - e^{-x} \int \frac{e^x}{2x} dx.$$

Osserviamo che le funzioni $c'_1(x)$ e $c'_2(x)$ trovate in questo esempio non hanno primitiva elementare (ricordarsi le osservazioni fatte nelle lezioni su integrali indefiniti e primitive). Un modo per esprimere gli integrali precedenti è quello, ad esempio, di “calcolarli per serie” facendo uso dello sviluppo in serie di potenze della funzione esponenziale.

Cenno ai problemi ai limiti.

Per concludere, osserviamo che, accanto ai problemi di Cauchy, esiste una vasta parte della teoria delle equazioni differenziali dedicata allo studio dei cosiddetti *problemi ai limiti*. Una trattazione generale di tali problemi esula dagli scopi di questo corso. Solo per darne un'idea, consideriamo l'equazione del second'ordine lineare a coefficienti costanti omogenea

$$y'' + y = 0.$$

Vogliamo vedere se esiste (almeno) una soluzione y dell'equazione che soddisfi le condizioni $y(0) = y(\pi) = 0$. Un problema di questo tipo è detto *problema dei due punti*. Come abbiamo già provato, l'integrale generale dell'equazione precedente è dato da

$$y(x) = c_1 \cos x + c_2 \sin x, \quad c_1, c_2 \in \mathbb{R}.$$

Imponendo la condizione $y(0) = 0$ otteniamo $c_1 = 0$, mentre ponendo $y(\pi) = 0$ non si ha nessuna limitazione sulla costante c_2 . Pertanto il problema ha **infinite** soluzioni della forma $y(x) = c \sin x$, $c \in \mathbb{R}$.

Se consideriamo invece la condizione $y(0) = 0$, $y(\pi) = \alpha$, $\alpha \neq 0$, si verifica subito che il problema non ha **nessuna** soluzione, mentre con $y(0) = 0$, $y(\pi/2) = \alpha$ si ottiene **una e una sola** soluzione, cioè la soluzione $y(x) = \alpha \sin x$.