

Seconda parte: Analisi Matematica II

71 - Mercoledì 03/03/10 (per i dettagli consultare il testo di riferimento)

Introduzione ai numeri complessi. L'unità immaginaria i . Parte reale, parte immaginaria e modulo di un numero complesso. Coniugato \bar{z} di un numero complesso z .

Notazione. L'insieme dei numeri complessi si denota col simbolo \mathbb{C} .

Osservazione. *Il prodotto di un numero complesso per il suo coniugato è uguale al quadrato del modulo.*

Esercizio. Provare che il coniugato di un prodotto [di una somma] coincide col prodotto [la somma] dei coniugati.

Argomenti di un numero complesso. Forma trigonometrica di un numero complesso.

Concetto di quoziente di due numeri complessi (come operazione inversa del prodotto).

Metodo pratico per eseguire il quoziente di due numeri complessi: si moltiplica numeratore e denominatore per il coniugato del denominatore.

Esercizio. Porre in forma trigonometrica i seguenti numeri complessi:

$$2, \quad 1 + i, \quad -1, \quad i, \quad -1 + i\sqrt{3}, \quad -3(\cos(\pi/3) + i\sin(\pi/3)).$$

72 - Mercoledì 03/03/10 (per i dettagli consultare il testo di riferimento)

Teorema. *Siano $\rho(\cos \theta + i \sin \theta)$ e $r(\cos \varphi + i \sin \varphi)$ due numeri complessi in forma trigonometrica. Se*

$$\rho(\cos \theta + i \sin \theta) = r(\cos \varphi + i \sin \varphi),$$

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 (del prodotto di due numeri complessi). *Il prodotto di due numeri complessi ha per modulo il prodotto dei moduli e per argomento la somma degli argomenti.*

Esercizio (sul quoziente di due numeri complessi). 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).

Esercizio. Utilizzare il teorema del prodotto di due numeri complessi, nonché l'esercizio precedente, per determinare (in modo rapido) il modulo del seguente numero complesso:

$$\frac{i(1-2i)(1+i)}{(1+3i)(2+i)}.$$

Esercizio. Determinare (in modo rapido) l'argomento dei seguenti numeri complessi:

$$-3(\cos(\pi/3) + i \operatorname{sen}(\pi/3)), \quad \frac{(1-i)(1+i)}{(2-2i)i}.$$

Esercizio. Porre in forma trigonometrica i seguenti numeri complessi:

$$-i(\cos(\pi/6) + i \operatorname{sen}(\pi/6)), \quad -3(\cos(\pi/3) + i \operatorname{sen}(\pi/3)).$$

Dal teorema del prodotto di due numeri complessi si deduce facilmente la seguente *Formula di De Moivre per la potenza n -esima di un numero complesso*:

$$[\rho(\cos \theta + i \operatorname{sen} \theta)]^n = \rho^n(\cos n\theta + i \operatorname{sen} n\theta).$$

Definizione (di radice n -esima di un numero complesso). Dato un numero complesso w e dato un numero naturale n , le soluzioni dell'equazione $z^n = w$ si chiamano *radici n -esime* di w .

Dalla Formula di De Moivre per il calcolo della potenza n -esima segue facilmente la seguente formula (anch'essa dovuta a De Moivre) per il calcolo delle radici n -esime di un numero complesso $z = \rho(\cos \theta + i \operatorname{sen} \theta)$:

$$\sqrt[n]{\rho(\cos \theta + i \operatorname{sen} \theta)} = \sqrt[n]{\rho} \left(\cos(\theta/n + 2k\pi/n) + i \operatorname{sen}(\theta/n + 2k\pi/n) \right), \quad k \in \mathbb{Z}.$$

Osservazione (sulla disposizione geometrica delle radici n -esime di un numero complesso). Nel piano complesso, le radici n -esime di un numero $z \in \mathbb{C}$ sono i vertici di un poligono regolare di n lati inscritto ad una circonferenza centrata nell'origine il cui raggio è la radice n -esima aritmetica del modulo $|z|$ di z . La differenza angolare tra due radici consecutive (cioè ottenute in base alla formula precedente con due valori consecutivi di k) è di un n -esimo dell'angolo giro (in altre parole, due radici consecutive sono sfasate di $2\pi/n$). Quindi, assegnando n valori consecutivi all'intero k , si ottengono tutte le possibili radici n -esime del numero complesso z .

Esercizio. Risolvere le seguenti equazioni in campo complesso:

$$(z+1)^3 + i = 0, \quad (z+i)^4 + 1 = 0.$$

73 - Venerdì 05/03/09 (per i dettagli consultare il testo di riferimento)

ATTENZIONE! In generale $\arctang(b/a)$ non coincide con l'argomento di $a + ib$ (a meno che la parte reale non sia positiva). Il motivo è dovuto al fatto che la funzione arcotangente assume soltanto valori appartenenti all'intervallo $(-\pi/2, \pi/2)$; di conseguenza,

se nessuno degli argomenti di un numero complesso $z = a + ib$ cade in tale intervallo (come accade, ad esempio, per $z = -1 + i$), il numero arctang (b/a) non può essere un argomento di z .

Esercizio^{fac}. Scrivere un programma che abbia come *input* i due numeri a e b di un numero complesso $z = a + ib$ e come *output* il modulo ρ e l'argomento (principale) θ di detto numero. Controllare, eventualmente con un altro programma, che effettivamente si abbia $a = \rho \cos \theta$ e $b = \rho \sin \theta$.

Definizione (di distanza in \mathbb{C}). Dati due numeri complessi z_1 e z_2 , il numero non negativo $|z_2 - z_1|$ si chiama *distanza* tra z_1 e z_2 .

Definizione (di intorno di un punto in \mathbb{C}). Dato un punto $z_0 \in \mathbb{C}$ ed assegnato un numero $r > 0$, l'*intorno* di z_0 di raggio r è l'insieme

$$I(z_0, r) = \left\{ z \in \mathbb{C} : |z - z_0| < r \right\}$$

costituito dai punti z che distano da z_0 meno di r . Pertanto, nel piano complesso, $I(z_0, r)$ coincide con un cerchio aperto (cioè senza la circonferenza) di centro z_0 e raggio r .

74 - Venerdì 05/03/10 (per i dettagli consultare il testo di riferimento)

Un *monomio* in campo complesso è una funzione (da \mathbb{C} in sé) del tipo az^n , dove $a \in \mathbb{C}$ è una costante e n è un intero non negativo (detto *grado del monomio*). Un *polinomio* in campo complesso non è altro che una somma finita di monomi (il suo *grado* è, per definizione, il massimo grado dei monomi di cui è composto). Una funzione ottenuta facendo il rapporto di due polinomi si chiama *funzione razionale*.

Richiami sul concetto di quoziente e di resto di una divisione tra due polinomi.

Come in campo reale, una *radice* (o uno *zero*) 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 $P(z_1) = 0$).

Osservazione. Un numero complesso z_1 è una radice di un polinomio $P(z)$ se e solo se $P(z)$ è divisibile (con resto zero) per il polinomio di primo grado $z - z_1$; cioè se e solo se esiste un polinomio $Q(z)$ tale che $P(z) = (z - z_1)Q(z)$.

Definizione (di molteplicità di una radice). Si dice che una radice z_1 di un polinomio $P(z)$ 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 del 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 esiste un polinomio $Q(z)$ tale che $P(z) = (z - z_1)^2 Q(z)$ e $Q(z_1) \neq 0$. In generale, un numero complesso z_1 è una *radice di molteplicità* $k \in \mathbb{N}$ di un polinomio $P(z)$ se esiste un polinomio $Q(z)$ tale che $Q(z_1) \neq 0$ e $P(z) = (z - z_1)^k Q(z)$.

Teorema. *Un polinomio a coefficienti reali, se ammette una radice complessa, ammette anche la sua coniugata.*

Esercizio. Provare il suddetto teorema.

Suggerimento. Usare i seguenti fatti: 1) il coniugato di un numero reale coincide con se stesso; 2) il coniugato della somma è uguale alla somma dei coniugati; 3) il coniugato di un prodotto coincide col prodotto dei coniugati.

Teorema Fondamentale dell’Algebra^{sd}. *In campo complesso ogni polinomio non costante (cioè, di grado non nullo) ammette almeno una radice.*

Esercizio. Dedurre, dal Teorema Fondamentale dell’Algebra, che 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 sono le radici del polinomio $P(z) = z^n - w$ (e sono tutte semplici).

Esercizio. Dedurre, dal teorema di esistenza degli zeri, che un polinomio a coefficienti reali di grado dispari ammette almeno una radice reale.

Esercizio. Provare che la formula risolutiva per le equazioni di secondo grado è ancora valida in campo complesso.

Esercizio. Determinare le radici dei seguenti polinomi e stabilirne la molteplicità:

$$z^2 + z + 1, \quad z^3 + 1, \quad z^6 + 4z^4, \quad z^4 - 1.$$

Esercizio. Determinare le radici del seguente polinomio:

$$P(z) = z^3 + z^2 - z - 1.$$

Suggerimento. Si osservi che $P(1) = 0$, e quindi $P(z)$ è divisibile per $z - 1$.

ATTENZIONE! *Nei numeri complessi non è possibile introdurre una relazione d’ordine totale compatibile con le operazioni algebriche di somma e prodotto.* In parole povere (più intuitive ma meno precise), qualunque nozione di “minore o uguale” si introducesse non godrebbe della seguente proprietà fondamentale (vera per i reali):

$$(0 \leq z_1) \wedge (0 \leq z_2) \implies 0 \leq z_1 z_2.$$

Quindi, espressioni del tipo

$$1 - i < 2 + 3i, \quad 0 \leq i, \quad 1 + 2i > 0, \quad \text{ecc.}$$

sono prive di significato, e non dovranno mai comparire nei compiti d’esame.

75 - Mercoledì 10/03/10 (per i dettagli consultare il testo di riferimento)

Riportiamo una definizione di funzione esponenziale in campo complesso. È probabile che in seguito gli studenti abbiano occasione di incontrarne un’altra equivalente, ma più naturale, basata sulla teoria delle serie di potenze in campo complesso.

Definizione (di funzione esponenziale in campo complesso). Dato un numero complesso $z = x + iy$, si pone $e^z = e^x(\cos y + i \operatorname{sen} y)$. Ossia, dato $z \in \mathbb{C}$, e^z è quel numero complesso che ha per modulo il numero e elevato alla parte reale di z e per argomento la parte immaginaria di z .

Osservazione (forma esponenziale dei numeri complessi). In base alla suddetta definizione, ogni numero complesso può essere rappresentato nel seguente modo: $z = \rho e^{i\theta}$, dove ρ è il modulo e θ un (qualunque) argomento.

Esercizio. Porre in forma esponenziale i seguenti numeri complessi:

$$2, \quad 1 + i, \quad -1, \quad i, \quad -1 + i\sqrt{3}, \quad -3(\cos(\pi/3) + i \operatorname{sen}(\pi/3)), \quad -2e^{i\frac{\pi}{2}}.$$

Teorema (proprietà fondamentale della funzione esponenziale in campo complesso). Per ogni $z_1, z_2 \in \mathbb{C}$, si ha $e^{z_1+z_2} = e^{z_1}e^{z_2}$. In particolare, dato $z = x + iy$, risulta $e^{x+iy} = e^x e^{iy}$.

Esercizio. Dedurre la suddetta proprietà fondamentale della funzione esponenziale in campo complesso dal fatto che il prodotto di due numeri complessi ha per modulo il prodotto dei moduli e per argomento la somma degli argomenti.

76 - Mercoledì 10/03/10

Ricordiamo che, in generale, una funzione (o applicazione) $f: A \rightarrow B$ tra due insiemi è una “legge” che ad ogni elemento $p \in A$ del primo insieme (chiamato *dominio*) associa un unico elemento $f(p) \in B$ del secondo insieme (detto *codominio*). Ricordiamo inoltre che il codominio di f non va confuso con l’immagine (denotata $\operatorname{Im} f$), che è l’insieme dei valori $f(p)$ assunti da f al variare di p nel dominio; cioè, per definizione,

$$\operatorname{Im} f = \{q \in B : \text{esiste (almeno) un } p \in A \text{ per il quale risulta } f(p) = q\}.$$

Quando l’immagine coincide col codominio, allora f si dice suriettiva.

Il *grafico* di una funzione $f: A \rightarrow B$ è il sottoinsieme del prodotto cartesiano $A \times B$ costituito dalle coppie (p, q) che verificano la relazione $q = f(p)$, detta *equazione del grafico* di f .

L’espressione “legge che ad ogni elemento del dominio associa un unico elemento del codominio” è intuitiva ma non del tutto chiara. La definizione che segue introduce il concetto di funzione in modo rigoroso, la riportiamo soltanto per le esigenze degli studenti desiderosi di una più profonda comprensione dei concetti astratti.

Definizione formale di funzione. Una *funzione* (o *applicazione*) è una terna (ordinata) $f = (A, B, G)$, dove A è un insieme, detto *dominio*, B è un insieme, chiamato *codominio*, e G , il cosiddetto *grafico* di f , è un sottoinsieme del prodotto cartesiano $A \times B$ che gode della seguente proprietà: per ogni $p \in A$ esiste un unico $q \in B$ per il quale si ha $(p, q) \in G$; tale elemento, univocamente associato a p , si chiama *immagine di p* e si denota $f(p)$.

Se il dominio di una funzione f è un sottoinsieme di \mathbb{R} , allora (qualunque sia il suo codominio) si dice che f è una *funzione di una variabile reale*; se è un sottoinsieme di \mathbb{R}^2

(di \mathbb{R}^3 , di \mathbb{R}^k), allora si dice che f è una *funzione di due (di tre, di k) variabili reali*. Quando il codominio di f è \mathbb{R} (o un suo sottoinsieme), allora la funzione f (qualunque sia il suo dominio) è detta *reale* (o *a valori reali*); quando il codominio è \mathbb{C} (o un suo sottoinsieme), allora si dice che f è una *funzione complessa*. Se il codominio di una funzione f è uno spazio vettoriale, allora f è detta una *funzione vettoriale*.

Ad esempio, la legge che ad ogni cilindro circolare retto di altezza h e raggio di base r fa corrispondere il volume v del cilindro è una funzione reale di due variabili reali definita dall'equazione $v = \pi hr^2$ (l'equazione del grafico). In questo caso h ed r si dicono le variabili indipendenti, mentre v è la variabile dipendente (da h e da r).

Ovviamente le funzioni reali (o complesse) si possono sommare tra loro, si possono moltiplicare, se ne può fare il quoziente. Osserviamo però che il dominio $\mathcal{D}(f_1 + f_2)$ della somma di due funzioni reali (o complesse) f_1 e f_2 è dato dall'intersezione $\mathcal{D}(f_1) \cap \mathcal{D}(f_2)$ dei domini; ossia è l'insieme dei punti p per cui hanno senso entrambi i numeri $f_1(p)$ e $f_2(p)$. Analogamente, per il prodotto, risulta $\mathcal{D}(f_1 f_2) = \mathcal{D}(f_1) \cap \mathcal{D}(f_2)$. Per quanto riguarda il quoziente f_1/f_2 , il dominio si ottiene togliendo dall'intersezione dei due domini l'insieme $f_2^{-1}(0)$ in cui si annulla il denominatore. Ad esempio il dominio della funzione $f(x, y) = (x + y)/(x - y)$ è \mathbb{R}^2 meno la retta $y = x$, mentre il dominio di $f(x, y, z) = (x + z)/(x - y)$ è \mathbb{R}^3 meno il piano $y = x$.

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 $g(x) = \sqrt{x}$, 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 $(x, y) \in \mathcal{D}(f)$ tali che $f(x, y) \in \mathcal{D}(g)$.

Esercizio. Determinare l'immagine della funzione

$$f(x) = \frac{1}{x + 2}$$

usando esclusivamente la definizione di immagine.

Esercizio. Determinare l'immagine della funzione

$$f(x) = \frac{1}{|x| + 1}$$

usando soltanto la definizione di immagine.

Esercizio. Determinare l'immagine della funzione $f: \mathbb{C} \rightarrow \mathbb{R}$, definita da

$$f(z) = 2 + |z^4 - 3z^3 + 2i|,$$

usando la definizione di immagine e il Teorema Fondamentale dell'Algebra.

Esercizio. Provare, usando esclusivamente la definizione di immagine, che il numero 6 appartiene all'immagine della funzione (reale di due variabili reali) definita da

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

Esercizio^{fac.} Provare, usando solo la definizione di immagine, che la seguente funzione reale è suriettiva:

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

La definizione di continuità per le funzioni reali di due (o più) variabili reali è analoga a quella già vista per il caso di una sola variabile. Prima di introdurre tale concetto, è utile ricordare la nozione di successione in un insieme arbitrario ed è necessario dare la definizione di successione convergente in \mathbb{R}^2 (o, più in generale, in \mathbb{R}^k).

Definizione (di successione). Sia X un insieme arbitrario. Una *successione in X* è un'applicazione $a: \mathbb{N} \rightarrow X$.

Se il codominio X di una successione $a: \mathbb{N} \rightarrow X$ è uguale a \mathbb{R} , la successione si dice reale; se $X = \mathbb{C}$, si dice complessa; se $X = \mathbb{N}$ si dice di numeri naturali, se $X = \mathbb{R}^2$, si dice di punti di \mathbb{R}^2 , ecc.

Data una successione $a: \mathbb{N} \rightarrow X$, 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 $a(n)$. Il valore a_n associato ad n si dice *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 $a: \mathbb{N} \rightarrow X$:

- $a: \mathbb{N} \rightarrow X$ (corretto, ma poco usato);
- $a_1, a_2, \dots, a_n, \dots$ (elencando i termini);
- $\{a_n\}$ (il più comune).

Osservazione. Attenzione a non fare confusione tra i *termini* di una successione (che sono sempre infiniti) e i suoi *valori* assunti, che possono essere anche un numero finito. Ad esempio, la successione reale $\{a_n\} = \{(-1)^n\}$ assume solo i valori 1 e -1 ma, ciascuno, infinite volte.

Ricordiamo, dal corso di Geometria, che dati due punti $p_1 = (x_1, y_1)$ e $p_2 = (x_2, y_2)$ di \mathbb{R}^2 , la loro distanza $d(p_1, p_2)$ è la norma della loro differenza. Pertanto:

$$d(p_1, p_2) = \|p_1 - p_2\| = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}.$$

In maniera analoga si definisce la distanza di due punti in \mathbb{R}^3 o in \mathbb{R}^k .

77 - Mercoledì 10/03/10

Definizione (di successione convergente). 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 per ogni $\varepsilon > 0$ esiste un $\bar{n} \in \mathbb{N}$ tale che per $n > \bar{n}$ si ha $\|p_n - p_0\| < \varepsilon$.

Se una successione $\{p_n\}$ in \mathbb{R}^k 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 $n \rightarrow +\infty$. Per lo stesso motivo si può scrivere anche

$$\lim_{n \rightarrow \infty} p_n = p_0,$$

dato che ∞ equivale a $\pm\infty$.

Si osservi che, in base alla suddetta definizione, una successione $\{p_n\}$ di punti di \mathbb{R}^k converge a p_0 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 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). Poiché anche in \mathbb{C} è definita una distanza (la distanza tra due punti è il modulo della differenza) diremo che una successione $\{z_n\}$ di numeri complessi converge a z_0 se la successione di numeri reali $\{|z_n - z_0|\}$ tende a zero.

Teorema. 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 $\beta_n \rightarrow \beta_0$.

Dimostrazione. Per esercizio^{fac}.

Esercizio. Provare che se una successione in \mathbb{R}^k è definitivamente costante (ossia, a partire da un certo indice \bar{n} tutti i suoi termini sono uguali), allora è convergente.

Teorema (di unicità del limite). Il limite di una successione in \mathbb{R}^k (o in \mathbb{C}), se esiste, è unico.

Dimostrazione. Per esercizio^{fac}.

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 $\{f(x_n, y_n)\} = \{7/n^2\}$. 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.

Definizione (sequenziale di continuità in un punto). Sia $f: A \rightarrow \mathbb{R}$ una funzione reale di due (o più) variabili reali e sia p_0 un punto del suo dominio A . Si dice che f è *continua in* p_0 se per ogni successione $\{p_n\}$ in A tale che $p_n \rightarrow p_0$ risulta $f(p_n) \rightarrow f(p_0)$.

Si osservi che, in base alla suddetta definizione, una funzione $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ è continua in un punto $p_0 = (x_0, y_0)$ se e solo se è vera la seguente proposizione:

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

Definizione (topologica di continuità in un punto). Sia $f: A \rightarrow \mathbb{R}$ una funzione reale di due (o più) variabili reali e sia p_0 un punto del suo dominio A . Si dice che f è *continua in* p_0 se per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che

$$(p \in A) \wedge (\|p - p_0\| < \delta) \implies |f(p) - f(p_0)| < \varepsilon.$$

Teorema (di collegamento per la continuità). *Le due definizioni di continuità, quella sequenziale e quella topologica, sono equivalenti.*

Dimostrazione. Per esercizio^{fac}.

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 .

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 $p \in \mathbb{R}^2$ associa la sua ascissa (cioè la prima coordinata). Analogamente la funzione y è quell'applicazione che ad ogni $p \in \mathbb{R}^2$ fa corrispondere la seconda coordinata di p . Le funzioni x e y si chiamano, rispettivamente, *prima e seconda funzione coordinata* (cartesiana).

Esercizio. Provare che se $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ è una funzione costante, allora è continua.

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

Teorema^{sd} (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.

78 - Venerdì 12/03/10

Definizione (di intorno di un punto). Dato un punto p_0 di \mathbb{R}^k e dato $r > 0$, l'insieme dei punti che distano da p_0 meno di r si chiama *intorno di* p_0 *di raggio* r e si denota $I(p_0, r)$.

In simboli:

$$I(p_0, r) = \{p \in \mathbb{R}^k : \|p - p_0\| < r\}.$$

Si osservi che $I(p_0, r)$ in \mathbb{R}^2 è un cerchio (di raggio r e centro p_0), in \mathbb{R}^3 è una palla e in \mathbb{R} è un intervallo (di ampiezza $2r$).

Teorema (della permanenza del segno per funzioni continue di più variabili). *Sia f una funzione reale definita in un sottoinsieme A di \mathbb{R}^k . Supponiamo che f sia continua in un punto $p_0 \in A$ e che $f(p_0)$ sia diverso da 0. Allora esiste un intorno U di p_0 tale che per tutti i punti p di tale intorno (e appartenenti al dominio di f) il numero $f(p)$ ha lo stesso segno di $f(p_0)$, cioè $f(p)f(p_0) > 0$ per ogni $p \in U \cap A$.*

Dimostrazione. Per esercizio.

Definizione (di alcune nozioni topologiche basilari). Dato un sottoinsieme A di \mathbb{R}^2 (o di \mathbb{R}^k), un punto p_0 si dice *interno* ad A se **esiste** un intorno $I(p_0, r)$ interamente contenuto in A , si dice di *frontiera* di A se **ogni** suo intorno contiene sia punti di A sia punti del complementare A^c . L'insieme dei punti di frontiera di A si dice *frontiera* di A e si denota ∂A . Un insieme A di \mathbb{R}^2 (o di \mathbb{R}^k) si dice *aperto* se ogni suo punto è interno (o, equivalentemente, se ogni suo punto di frontiera appartiene ad A^c). Se tutti i punti di frontiera di A appartengono ad A , allora A si dice *chiuso* (o, equivalentemente, A è chiuso se il suo complementare è aperto).

Definizione (di insieme limitato). Un sottoinsieme A di \mathbb{R}^2 (o di \mathbb{R}^k) si dice *limitato* se il numero reale esteso $\text{diam}(A) = \sup\{\|p - q\| : p, q \in A\}$, detto *diametro* di A , è finito.

Esercizio. Calcolare il diametro di un rettangolo $[a, b] \times [c, d] \subseteq \mathbb{R}^2$.

Esercizio. Siano A e B due sottoinsiemi di \mathbb{R}^k . Provare che se $A \subseteq B$, allora

$$\text{diam}(A) \leq \text{diam}(B).$$

Dedurre da ciò che se un insieme è limitato, allora lo è anche ogni suo sottoinsieme.

Esercizio^{fac}. Provare che il diametro di un intorno $I(p_0, r)$ di un punto $p_0 \in \mathbb{R}^k$ è $2r$.

Convenzione. Da ora in poi, a meno che non venga esplicitamente affermato il contrario, supporremo che le funzioni di due o più variabili siano definite su insiemi aperti o, più in generale, che siano la restrizione (ad insiemi non necessariamente aperti) di funzioni definite su insiemi aperti.

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 $y = y_0$). Si osservi che esistono infinite funzioni parziali della prima variabile: una per ogni valore assegnato della y . Analogamente, se si fissa un valore della prima variabile, si ottiene una delle tante funzioni parziali della seconda variabile. È opportuno osservare che le funzioni parziali (delle funzioni di due variabili), anche se sono infinite, si dividono in due sole classi: quelle della prima variabile e quelle della seconda.

Il concetto di funzione parziale si estende in modo naturale al caso di tre o più variabili. I dettagli sono lasciati allo studente.

Definizione (di derivata parziale in un punto). Sia $f: A \rightarrow \mathbb{R}$ una funzione reale di due variabili reali e sia (x_0, y_0) un punto di A . La *derivata (parziale)* in (x_0, y_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) $x \mapsto f(x, y_0)$. Analogamente, la *derivata (parziale)* in (x_0, y_0) di f rispetto alla y è la derivata in y_0 della funzione parziale $y \mapsto f(x_0, y)$.

Data una funzione f di due variabili, la sua derivata rispetto alla x in un punto $p_0 = (x_0, y_0)$ si denota con uno dei seguenti simboli:

$$\frac{\partial f}{\partial x}(x_0, y_0), \quad \frac{\partial f}{\partial x}(p_0), \quad \partial_1 f(x_0, y_0), \quad \partial_1 f(p_0), \quad D_1 f(x_0, y_0), \quad D_1 f(p_0).$$

Notazioni analoghe valgono per la derivata rispetto alla y .

Se si è capito il concetto di derivata parziale si dovrebbe saper svolgere l'esercizio che segue eseguendo pochi semplicissimi calcoli.

Esercizio. Determinare le derivate parziali nel punto $(0, 0)$ della funzione

$$f(x, y) = \begin{cases} \sin(x + 2y) - \frac{xy \sin(x - 2y)}{\sqrt{x^2 + 3y^2}} & \text{se } (x, y) \neq (0, 0), \\ 0 & \text{se } (x, y) = (0, 0). \end{cases}$$

Si dice che una funzione f di due variabili è *derivabile (parzialmente)* rispetto alla x (senza ulteriori precisazioni) se è derivabile rispetto alla x in ogni punto del dominio. In questo caso risulta ben definita la funzione $\partial f / \partial x$ (denotata anche $\partial_1 f$ o $D_1 f$), detta (*funzione*) *derivata* rispetto alla x (o rispetto alla prima variabile) che ad ogni punto (x, y) del dominio di f associa il numero

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

La definizione della funzione derivata rispetto alla seconda variabile (denotata $\partial f / \partial y$ o $\partial_2 f$ o $D_2 f$) è analoga.

Esercizio. Definire la nozione di derivata parziale in un punto (e di funzione derivata parziale) per le funzioni di tre (o, più in generale, di k) variabili.

Si dice che una funzione di due variabili f è *derivabile* se è derivabile sia rispetto alla x sia rispetto alla y . Se la derivata di f rispetto alla x è di nuovo derivabile rispetto alla x , allora f si dice derivabile due volte rispetto alla x (o che ammette *derivata seconda* rispetto alla x). La derivata rispetto alla x della derivata rispetto alla x , calcolata in un punto (x, y) , si indica con uno dei seguenti simboli:

$$\frac{\partial}{\partial x} \frac{\partial f}{\partial x}(x, y), \quad \frac{\partial^2 f}{\partial x^2}(x, y), \quad \partial_1 \partial_1 f(x, y), \quad D_1 D_1 f(x, y).$$

In modo simile si definiscono le altre derivate seconde:

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

79 - Venerdì 12/03/10

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. I dettagli sono lasciati allo studente.

Esercizio. Definire le derivate parziali di ordine superiore al secondo (per le funzioni di due variabili).

Sappiamo che se una funzione di una variabile è derivabile, allora è anche continua. Un risultato analogo è falso per le funzioni di due (o più) variabili. Tuttavia, se una funzione di due (o più) variabili è derivabile e le sue derivate parziali sono continue, allora si può affermare che è continua. Dal Teorema di Lagrange si potrebbe infatti dedurre il seguente risultato.

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

Esercizio^{fac}. Provare che se f è una funzione di classe C^n , allora è anche di classe C^{n-1} .

Suggerimento. Denotare con P_n la proposizione “ $f \in C^n \implies f \in C^{n-1}$ ”. Dopo aver osservato che (per il precedente teorema) P_1 è vera, procedere per induzione.

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 .

Dimostrazione^{fac}.

(Somma) Sia P_n la proposizione “ $f, g \in C^n \implies (f + g) \in C^n$ ”. Poiché la somma di funzioni continue è una funzione continua, P_0 è vera. Assumiamo vera P_{n-1} e dimostriamo che allora è vera anche P_n . Supponiamo quindi che f e g siano C^n (ovvero che tutte le derivate parziali $D_i f$ e $D_i g$ siano C^{n-1}) e mostriamo che $(f + g) \in C^n$ (ossia che $D_i(f + g) \in C^{n-1}$). Poiché $D_i(f + g) = D_i f + D_i g$, dall’ipotesi induttiva si ha $D_i(f + g) \in C^{n-1}$, che per definizione significa $(f + g) \in C^n$.

(Prodotto) Analogamente alla dimostrazione precedente, denotiamo con P_n la proposizione “ $f, g \in C^n \implies fg \in C^n$ ”. Poiché il prodotto di funzioni continue è una funzione continua, P_0 è vera. Assumiamo (per ipotesi induttiva) vera P_{n-1} e supponiamo che f e g siano C^n . Vogliamo provare che il prodotto fg è di classe C^n , ovvero che le derivate parziali $D_i(fg) = (D_i f)g + f(D_i g)$, sono di classe C^{n-1} . Questo segue facilmente dall’esercizio precedente, dall’ipotesi induttiva, e da quanto già provato per la somma.

(Quoziente) La dimostrazione è simile alle due precedenti ed è lasciata per esercizio ^{fac} allo studente.

(Composizione) La dimostrazione, che riportiamo per semplicità soltanto per le funzioni di una sola variabile, è basata sulla formula della derivata di una funzione composta: $(g \circ f)'(x) = g'(f(x))f'(x)$. Questa ci dice che se f e g sono di classe C^n , la derivata di $g \circ f$ è prodotto e composizione di funzioni di classe C^{n-1} (si ha infatti $(g \circ f)' = (g' \circ f)f'$). Pertanto, posto $P_n = “f, g \in C^n \implies g \circ f \in C^n”$, il risultato segue facilmente per induzione, dopo aver osservato che P_0 è vera. \square

Si osservi che in \mathbb{R}^2 le funzioni costanti e le funzioni coordinate x e y sono di classe C^∞ (verificarlo per esercizio), 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. Mostrare che la funzione

$$f(x, y) = |x|y$$

è derivabile in $(0, 0)$.

Esercizio. Mostrare che la funzione

$$f(x, y) = |x|y$$

è derivabile rispetto ad y in tutti i punti di \mathbb{R}^2 .

Esercizio. Mostrare che l'insieme dei punti in cui la funzione

$$f(x, y) = |x|y$$

non è derivabile è

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

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 .

Esercizio. Mostrare che la funzione

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

è di classe C^1 (in tutto il suo dominio \mathbb{R}^2).

80 - Mercoledì 17/03/10

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

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

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

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

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

$$g(x, y) := \frac{\partial f}{\partial x}(x, y)$$

è di classe C^2 . Di conseguenza, le suddette derivate terze risultano uguali, visto che non sono altro che le derivate seconde miste della funzione $g(x, y)$. Non c'è quindi nessuna ambiguità nell'adottare per le suddette 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.

In maniera analoga si prova che se una funzione di k variabili è di classe C^n , in una sua derivata n -esima si può invertire l'ordine di derivazione di due arbitrarie derivate (si osservi che l'inversione di due derivate non adiacenti si riconduce ad un numero finito di inversioni di derivate adiacenti).

Esercizio. Dedurre dal Teorema di Schwarz che se $f: A \rightarrow \mathbb{R}$ è una funzione di classe C^2 su un aperto A di \mathbb{R}^k , allora, qualunque siano $i, j \in \{1, 2, \dots, k\}$, risulta

$$\frac{\partial^2 f}{\partial x_i \partial x_j} = \frac{\partial^2 f}{\partial x_j \partial x_i}.$$

81 - Mercoledì 17/03/10

Sia $g: W \rightarrow \mathbb{R}$ una funzione continua su un aperto W di \mathbb{R}^3 . Un'uguaglianza del tipo

$$g(x, y, y') = 0$$

o del tipo

$$g(x, y(x), y'(x)) = 0$$

si chiama *equazione differenziale ordinaria del prim'ordine*.

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 detto insieme è o non è una soluzione. Per quanto riguarda la suddetta equazione, l'insieme in cui si cercano le soluzioni è dato dalle funzioni reali di classe C^1 definite in un intervallo non banale (l'intervallo può variare da funzione a funzione). Una funzione $y(x)$, di classe C^1 in un intervallo J , è una *soluzione* se per ogni $x \in J$ si ha $(x, y(x), y'(x)) \in W$ e

$$g(x, y(x), y'(x)) = 0.$$

In particolare, data una soluzione $y: J \rightarrow \mathbb{R}$, la sua restrizione ad un sottointervallo non banale di J è ancora una soluzione. Tra tutte le soluzioni, quelle che non sono restrizione propria 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. L'insieme di tutte le soluzioni di un'equazione differenziale si dice *soluzione generale* o *integrale generale*.

Talvolta la funzione $g(x, y, y')$ è del tipo $y' - f(x, y)$, con f definita (e continua) in un aperto U di \mathbb{R}^2 . Si osservi che in questo caso l'aperto W in cui è definita la funzione $g(x, y, y') := y' - f(x, y)$ è $U \times \mathbb{R}$, e l'equazione può essere scritta nella forma

$$y' = f(x, y),$$

detta *forma normale*.

Da ora in avanti, a meno che non sia altrimenti specificato, ci occuperemo di equazioni differenziali in forma normale.

82 - Mercoledì 17/03/10

Il più banale esempio di equazione differenziale è

$$y' = f(x),$$

dove f è una funzione continua definita in un intervallo $J \subseteq \mathbb{R}$. In questo caso la soluzione generale è data dall'insieme delle primitive di f .

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.

Esercizio. Stabilire quali, tra le seguenti funzioni, sono soluzioni dell'equazione differenziale $x + y - y' = 0$:

$$x - e^x, \quad x^2, \quad -x - 1, \quad \cos x, \quad 3e^x - x - 1, \quad 1 - x.$$

Esercizio. Stabilire se tra le soluzioni dell'equazione differenziale

$$y' = xy + 1 - x^2$$

c'è un polinomio di primo grado.

Esercizio. Stabilire se l'equazione differenziale

$$y' + 2y = 0$$

ammette una soluzione del tipo $e^{\lambda x}$, con λ costante reale.

Esercizio. Determinare (se esiste) una funzione $g(x)$ in modo che la funzione $e^{g(x)}$ risulti soluzione dell'equazione differenziale $y' = 2xy$.

Esercizio. Stabilire se l'equazione differenziale

$$y' = x^2(1 - y^2)$$

ammette soluzioni costanti.

83 - Venerdì 19/03/10

Esercizio. Sia $a(x)$ una funzione continua definita in un intervallo e sia $A(x)$ una sua primitiva. Provare che ogni funzione del tipo $y(x) = ce^{A(x)}$, dove $c \in \mathbb{R}$ è un'arbitraria costante, è una soluzione dell'equazione differenziale $y' = a(x)y$ che, come vedremo, si dice *lineare, omogenea, del prim'ordine* (in seguito proveremo che, viceversa, ogni soluzione della suddetta equazione è della forma $y(x) = ce^{A(x)}$, cioè tale espressione dà tutte le soluzioni dell'equazione).

Esercizio. Siano $a(x)$ e $b(x)$ due funzioni continue definite in un intervallo (lo stesso per entrambe) e sia $A(x)$ una primitiva di $a(x)$. Provare che se $\hat{y}(x)$ è una soluzione dell'equazione differenziale $y' = a(x)y + b(x)$, detta *lineare, non omogenea, del prim'ordine*, allora anche le funzioni del tipo $y(x) = ce^{A(x)} + \hat{y}(x)$, dove c è un'arbitraria costante, sono soluzioni di tale equazione (in seguito proveremo che tutte le soluzioni della suddetta equazione si esprimono in tal modo).

Sia $y' = f(x, y)$ un'equazione differenziale del prim'ordine in forma normale e sia $U \subseteq \mathbb{R}^2$ l'aperto in cui è definita la funzione f . Dato un punto $(x_0, y_0) \in U$, ci si pone il problema di trovare, tra tutte le soluzioni $y(x)$ della suddetta 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* (si pronuncia "coscì"); così come *di Cauchy* si chiama la *condizione* $y(x_0) = y_0$. Il punto x_0 si dice *punto* (o *istante*) *iniziale* (della soluzione cercata) e y_0 è il *valore iniziale*.

Il risultato che segue 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^{sd} (di esistenza di Peano). *Sia $f: U \rightarrow \mathbb{R}$ una funzione continua su un aperto $U \subseteq \mathbb{R}^2$. Allora, per ogni $(x_0, y_0) \in U$, l'equazione $y' = f(x, y)$ ammette almeno una soluzione (massimale) 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 $y(0) = 0$.

84 - Venerdì 19/03/10

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^{sd} (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 $U \subseteq \mathbb{R}^2$. Se f è derivabile rispetto ad y e $\partial f / \partial y$ è continua, allora, per ogni $(x_0, y_0) \in U$, la suddetta 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 $U \subseteq \mathbb{R}^2$. Supponiamo che f sia derivabile rispetto ad y con derivata continua. Allora i grafici di due differenti soluzioni massimali non possono intersecarsi.

Una proprietà significativa delle soluzioni di un'equazione differenziale è espressa dal teorema che segue. Per enunciarlo occorre introdurre la seguente nozione: una soluzione $y: J \rightarrow \mathbb{R}$ 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^{sd} (di Kamke). *Sia $f: U \rightarrow \mathbb{R}$ una funzione continua su un aperto $U \subseteq \mathbb{R}^2$. 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 U .

In un certo senso il Teorema di Kamke afferma che il grafico di ogni soluzione massimale (che, ricordiamo, è un sottoinsieme del dominio U 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 U (e allora si deve arrestare), oppure se ne va all’infinito (sempre rimanendo dentro U). Ciò che non può accadere è che il grafico di una soluzione non prolungabile si arresti in un punto interno ad U : il Teorema di Peano gli consentirebbe di proseguire, contraddicendo la non prolungabilità.

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 $U = (a, b) \times \mathbb{R}$ (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’esempio che segue mostra che possono esistere soluzioni massimali non persistenti (il Teorema di Kamke implica che tali soluzioni non possono essere limitate).

Esempio. Consideriamo il seguente problema di Cauchy:

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

In questo caso la funzione $f(x, y)$ coincide con y^2 , e quindi la striscia $(a, b) \times \mathbb{R}$ del piano xy in cui f è definita è $\mathbb{R} \times \mathbb{R}$ (cioè $a = -\infty$ e $b = +\infty$). Il metodo per risolvere il suddetto problema sarà visto in seguito. Per ora si osservi che la funzione

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

soddisfa sia l’equazione differenziale sia la condizione di Cauchy (si invitano gli studenti a verificarlo per esercizio). Non possiamo però affermare che detta funzione sia una soluzione del problema assegnato, perché il suo dominio non è un intervallo (come invece è richiesto dalla definizione di soluzione di un’equazione differenziale). Tuttavia la restrizione $y(x)$ di tale funzione all’intervallo $(-\infty, 1)$ è certamente una soluzione del problema, ed è anche massimale dato che non è prolungabile né a sinistra (perché l’estremo inferiore del suo dominio è $-\infty$) né a destra (perché $\lim_{x \rightarrow 1^-} y(x) = +\infty$).

Esercizio^{fac}. Provare che se $f(x, y)$ è una funzione di classe C^∞ (in un aperto U di \mathbb{R}^2), allora sono C^∞ anche tutte le soluzioni dell’equazione differenziale $y' = f(x, y)$.

Suggerimento. Usare la definizione di soluzione, la definizione di funzione C^n , il teorema di regolarità delle funzioni combinate, il principio di induzione.

85 - Mercoledì 24/03/10

Un'equazione differenziale del tipo

$$y' = a(x)h(y),$$

dove $a(x)$ e $h(y)$ sono funzioni di una sola variabile, si dice a *variabili separabili*. Supponiamo $a(x)$ continua e $h(y)$ 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 (perché?). È inoltre conveniente supporre che $a(x)$ si possa annullare soltanto in punti isolati.

Si fa notare che se $c \in \mathbb{R}$ è un punto in cui $h(y)$ si annulla, allora la funzione costante $y(x) \equiv c$ è una soluzione dell'equazione differenziale. Viceversa (avendo supposto che $a(x)$ si possa annullare soltanto in punti isolati) ogni soluzione costante $y(x) \equiv c$ è tale che $h(c) = 0$. Le soluzioni costanti sono quindi in corrispondenza biunivoca con gli zeri di $h(y)$.

Cerchiamo di determinare le soluzioni non costanti. Se $y(x)$ è una tale soluzione, per il teorema di esistenza e unicità si deve avere $h(y(x)) \neq 0$ per ogni x nell'intervallo J in cui è definita $y(x)$, altrimenti il grafico di $y(x)$ intersecherebbe il grafico di una soluzione costante. Dividendo l'uguaglianza

$$y'(x) = a(x)h(y(x))$$

per $h(y(x))$ si ha allora

$$\frac{y'(x)}{h(y(x))} = a(x).$$

Quindi 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

$$\int \frac{dy}{h(y)} = \int a(x) dx.$$

Dunque, denotando con $H(y)$ una primitiva di $1/h(y)$ e con $A(x)$ una primitiva di $a(x)$, si ottiene la seguente equazione (del grafico delle soluzioni non costanti):

$$H(y) = A(x) + c,$$

dove c è un'arbitraria costante. Ricavando la y si ha la formula

$$y = H^{-1}(A(x) + c)$$

che **dà soltanto le soluzioni non costanti** dell'equazione in esame (le soluzioni costanti le avevamo già trovate, ed è bene ricordarselo quando c'è da risolvere un problema di Cauchy). Si fa notare che $H(y)$ è strettamente monotona, quindi invertibile, perché la stiamo considerando in un intervallo in cui la sua derivata $H'(y) = 1/h(y)$ ha segno costante (altrimenti $h(y)$ si annullerebbe in qualche punto di tale intervallo). Si lascia per esercizio^{fac} 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

$$y' = a(x)h(y).$$

Si avverte che nell'eseguire la verifica, la presenza di H^{-1} rende indispensabile l'uso del teorema di derivazione di una funzione inversa.

Un metodo più convincente per ottenere la suddetta formula risolutiva (per le soluzioni non costanti) è quello di integrare direttamente entrambi i membri dell'uguaglianza

$$\frac{y'(x)}{h(y(x))} = a(x),$$

che avevamo precedentemente ottenuto. Con le notazioni introdotte, si ha

$$H(y(x)) = A(x) + c,$$

come segue facilmente dal fatto che $H(y(x))$ e $A(x)$ hanno la stessa derivata (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 $y(x)$ si ha infine

$$y(x) = H^{-1}(A(x) + c).$$

86 - Mercoledì 24/03/10

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 ammette la soluzione nulla, che è l'unica soluzione costante. Sia quindi $y(x)$ una soluzione non costante. Come già osservato, essa non potrà mai annullarsi, 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

$$\int \frac{y'(x)}{y^2(x)} dx = \int dx,$$

ossia

$$\int \frac{dy(x)}{y^2(x)} = \int dx,$$

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 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 $(1, +\infty)$.

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 è di classe C^1 e quindi il problema di Cauchy ammette una e una sola soluzione massimale. Seguendo il metodo precedentemente illustrato 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 $y(x)$ del problema di Cauchy verifica la condizione

$$\operatorname{arctang} y(x) = \operatorname{arctang} x + \frac{\pi}{4},$$

definita se $|\operatorname{arctang} x + \pi/4| < \pi/2$, cioè nell'intervallo $(-\infty, 1)$. In questo caso, facendo uso delle formule di addizione delle funzioni seno e coseno, si riesce anche a ricavare l'espressione esplicita della soluzione. Si ottiene

$$y(x) = \operatorname{tang} \left(\operatorname{arctang} x + \frac{\pi}{4} \right) = \frac{x+1}{1-x}.$$

Esercizio. Trovare le soluzioni dell'equazione differenziale

$$y' = 1 - y^2.$$

87 - Mercoledì 24/03/10

Un'equazione differenziale del prim'ordine si dice *lineare* se è della forma

$$y' = a(x)y + b(x),$$

dove $a(x)$ e $b(x)$ sono due funzioni continue definite in un intervallo J . In particolare, quando il *termine noto* $b(x)$ 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*).

Teorema. Sia $a(x)$ una funzione continua in un intervallo $J \subseteq \mathbb{R}$. 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 $a(x)$ 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 con il metodo pratico (anche se poco ortodosso). Si ottiene

$$\frac{1}{y} dy = a(x) dx,$$

da cui

$$\int \frac{1}{y} dy = \int a(x) dx.$$

Quindi

$$\log |y| = A(x) + k,$$

dove $A(x)$ è una primitiva di $a(x)$ e k è un'arbitraria costante. Pertanto

$$|y| = e^{A(x)+k} = e^k e^{A(x)},$$

cioè

$$y = \pm e^k e^{A(x)}$$

o, equivalentemente, tenendo conto che e^k è un'arbitraria costante positiva, il grafico di una qualunque soluzione non costante verifica l'equazione

$$y = ce^{A(x)}, \quad \text{con } c \neq 0.$$

Poiché ponendo $c = 0$ nella suddetta 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

$$y = ce^{A(x)},$$

con c costante arbitraria, anche nulla. □

Teorema. *Supponiamo che $\hat{y}(x)$ sia una soluzione dell'equazione differenziale lineare*

$$y' = a(x)y + b(x),$$

dove $a(x)$ e $b(x)$ sono funzioni continue in un intervallo $J \subseteq \mathbb{R}$. Allora ogni altra soluzione si ottiene aggiungendo ad $\hat{y}(x)$ la soluzione generale dell'equazione omogenea associata $y' = a(x)y$.

Dimostrazione. Proviamo prima che se $u(x)$ è una soluzione dell'equazione omogenea, allora $y(x) := \hat{y}(x) + u(x)$ è soluzione della non omogenea. Dalle uguaglianze

$$\hat{y}'(x) = a(x)\hat{y}(x) + b(x) \quad \text{e} \quad u'(x) = a(x)u(x)$$

segue

$$\begin{aligned} \hat{y}'(x) + u'(x) &= (a(x)\hat{y}(x) + b(x)) + a(x)u(x) = \\ &= a(x)(\hat{y}(x) + u(x)) + b(x). \end{aligned}$$

Quindi

$$y'(x) = a(x)y(x) + b(x),$$

e ciò prova che $y(x) := \hat{y}(x) + u(x)$ è soluzione della non omogenea.

Rimane da verificare che se $\tilde{y}(x)$ è una soluzione dell'equazione non omogenea, allora esiste una soluzione $u(x)$ dell'omogenea tale che $\tilde{y}(x) = \hat{y}(x) + u(x)$. In altre parole, rimane da provare che la funzione $u(x) := \tilde{y}(x) - \hat{y}(x)$ è soluzione dell'equazione omogenea, e questo è un esercizio che lasciamo allo studente. □

Osserviamo che il suddetto teorema, nel caso particolare in cui la funzione $a(x)$ sia nulla, si riduce ad un risultato ben noto: *data una primitiva $\hat{y}(x)$ di $b(x)$, ogni altra primitiva si ottiene aggiungendo ad $\hat{y}(x)$ un'arbitraria costante (ossia, la soluzione generale di $y' = 0$).*

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, che incontreremo più avanti, 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 $A(x)$ è una primitiva di $a(x)$ e c una costante arbitraria. Il metodo consiste nel cercare una soluzione particolare dell'equazione non omogenea, pensando variabile la costante c (è una contraddizione in termini, ma l'espressione “variazione delle costanti” fa ormai parte del folklore matematico). In altre parole, si cerca una soluzione del tipo

$$\hat{y}(x) = c(x)e^{A(x)}.$$

Derivando si ottiene

$$\hat{y}'(x) = c'(x)e^{A(x)} + c(x)a(x)e^{A(x)}.$$

Quindi $\hat{y}(x)$ è soluzione se (e anche solo se, ma non ci interessa)

$$c'(x)e^{A(x)} + a(x)c(x)e^{A(x)} = a(x)c(x)e^{A(x)} + b(x),$$

ossia se

$$c'(x) = e^{-A(x)}b(x),$$

ovvero se $c(x)$ è una primitiva di $e^{-A(x)}b(x)$. Di conseguenza, la soluzione generale dell'equazione non omogenea è data dalla seguente

Formula risolutiva per le equazioni differenziali lineari del prim'ordine (che si consiglia di imparare a memoria):

$$y(x) = ce^{A(x)} + e^{A(x)} \int e^{-A(x)}b(x) dx, \quad c \in \mathbb{R}.$$

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

$$\begin{aligned} y(x) &= ce^{-x^2} + e^{-x^2} \int e^{x^2} x dx = ce^{-x^2} + \frac{1}{2} e^{-x^2} \int e^{x^2} dx^2 = \\ &= ce^{-x^2} + \frac{1}{2} e^{-x^2} e^{x^2} = ce^{-x^2} + \frac{1}{2}. \end{aligned}$$

Dobbiamo ancora determinare la costante c in modo che sia verificata la condizione iniziale $y(0) = 0$. Abbiamo

$$y(0) = c + \frac{1}{2} = 0,$$

da cui si ricava $c = -1/2$. La soluzione cercata è dunque

$$y(x) = \frac{1}{2} \left(1 - e^{-x^2} \right),$$

come si può facilmente verificare (si invita lo studente a farlo).

Il metodo di variazione della costante per determinare una soluzione particolare di un'equazione lineare non omogenea può essere, talvolta, troppo elaborato. Esistono dei metodi, detti “rapidi”, che funzionano bene quando l'equazione omogenea associata ha il coefficiente costante (ossia, è del tipo $y' = ay$, con a costante) e quando il termine noto si presenta in una forma molto particolare. Il metodo (che possiamo chiamare “del pescatore”) consiste nel “tirare ad indovinare”, ovvero nel cercare una soluzione dell'equazione in una classe di funzioni dove si suppone debba essercene almeno una. Se poi non si trova, pazienza, si può sempre procedere col metodo di variazione della costante. Analizziamo alcuni casi in cui il metodo funziona.

Consideriamo l'equazione

$$y' = ay + b(x).$$

Supponiamo che $b(x)$ sia un polinomio di grado n . Se $a \neq 0$, si cerca una soluzione particolare tra i polinomi di grado n . Se, invece, $a = 0$, una soluzione particolare è una primitiva di $b(x)$, che in tal caso è un polinomio di grado $n + 1$.

Supponiamo ora che $b(x)$ sia del tipo $\alpha \cos \omega x + \beta \sin \omega x$, con $\omega > 0$. In questo caso si cerca una soluzione particolare dello stesso tipo, dove, ovviamente, al posto delle costanti α e β si scrivono dei coefficienti da determinare.

Un terzo caso importante si ha quando $b(x) = \alpha e^{\gamma x}$, con $\gamma \in \mathbb{R}$. Se $a \neq \gamma$, si cerca una soluzione particolare del tipo $ce^{\gamma x}$, con c costante da determinare. Se $a = \gamma$ (ossia se $e^{\gamma x}$ è soluzione dell'equazione omogenea associata), si determina una soluzione particolare moltiplicando per x la funzione $ce^{\gamma x}$ e si trova c in modo da ottenere una soluzione della non omogenea.

88 - Venerdì 26/03/10

Un'espressione del tipo

$$y'' = f(x, y, y'),$$

dove f è una funzione continua definita su un aperto U 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, detto *universo*, in cui si cercano le soluzioni; 2) un criterio chiaro per stabilire quando un elemento dell'universo ha il diritto di chiamarsi soluzione.

Per quanto riguarda la suddetta equazione differenziale, le soluzioni si cercano nell'insieme delle funzioni di classe C^2 , ciascuna delle quali è definita in un intervallo (dipendente dalla funzione stessa). Una funzione $y(x)$ di tale insieme si dirà una soluzione se per ogni x appartenente all'intervallo J in cui è definita risulta

$$(x, y(x), y'(x)) \in U \quad \text{e} \quad y''(x) = f(x, y(x), y'(x)).$$

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 (che in questo caso si denota con t invece che con x), dalla posizione e dalla velocità (spesso denotate rispettivamente con x e con \dot{x}). Ovviamente questa non è l'unica possibile interpretazione della suddetta equazione: molti fenomeni fisici (non solo di dinamica) sono governati da equazioni differenziali 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 $f: U \rightarrow \mathbb{R}$ è una funzione continua da un aperto U di \mathbb{R}^{n+1} in \mathbb{R} . Una soluzione è una funzione $y(x)$ di classe C^n in un intervallo J tale che

$$y^{(n)}(x) = f(x, y(x), y'(x), \dots, y^{(n-1)}(x)), \quad \forall x \in J.$$

Ovviamente, affinché abbia senso la suddetta uguaglianza, si sottintende che

$$(x, y(x), y'(x), \dots, y^{(n-1)}(x)) \in U, \quad \forall x \in J.$$

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 U in cui è definita la f . Più precisamente, dato un punto $p_0 = (x_0, y_0, y_1, \dots, y_{n-1}) \in U$, tra tutte le soluzioni $y(x)$ dell'equazione

$$y^{(n)} = f(x, y, y', \dots, y^{(n-1)}),$$

si cercano quelle che verificano le condizioni

$$y(x_0) = y_0, \quad y'(x_0) = y_1, \quad \dots, \quad y^{(n-1)}(x_0) = y_{n-1}.$$

Il numero x_0 si dice *istante* (o *punto*) *iniziale* (del problema di Cauchy) e i numeri y_0, y_1, \dots, y_{n-1} sono i *valori iniziali*.

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 $f(x, y, y', \dots, y^{(n-1)})$ è continua e derivabile rispetto alle variabili $y, y', \dots, y^{(n-1)}$ con derivate continue (ad esempio, se è C^1), allora il problema di Cauchy ammette una ed una sola soluzione massimale. Per l'equazione di moto di un punto vincolato ad una retta, significa che se ad

un certo istante t_0 si conoscono sia la posizione sia la velocità, il moto è univocamente determinato.

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*).

L'applicazione $D: C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ che ad ogni funzione $y(x)$ dello spazio $C^\infty(\mathbb{R})$ associa la funzione derivata $Dy := y'$ gode delle seguenti due proprietà:

(*additività*) $D(y_1 + y_2) = Dy_1 + Dy_2$, per ogni $y_1, y_2 \in C^\infty(\mathbb{R})$;

(*omogeneità*) $D(cy) = cDy$, per ogni $c \in \mathbb{R}$ e per ogni $y \in C^\infty(\mathbb{R})$.

Poiché D gode di tali proprietà, si dice che è un'applicazione *lineare* (o un *operatore lineare*) dallo spazio $C^\infty(\mathbb{R})$ in sé. Un altro esempio di operatore lineare da $C^\infty(\mathbb{R})$ in sé è l'*operatore identico* (o *identità*) $I: C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$; ossia quell'applicazione che ad ogni funzione associa la funzione stessa. Due operatori lineari si possono sommare, ottenendo ancora un operatore lineare, così come un operatore lineare si può moltiplicare per una costante (e il risultato è ancora un operatore lineare). Nel caso di applicazioni da uno spazio in sé, come quello che stiamo considerando, ha senso anche la composizione. Ad esempio, la composizione di D con D , denotata con D^2 , è l'operatore che ad ogni funzione di $C^\infty(\mathbb{R})$ associa la sua derivata seconda (ossia, $D^2y = y''$). Più in generale, D^n rappresenta l'operatore lineare che ad ogni funzione $y \in C^\infty(\mathbb{R})$ associa la sua derivata n -esima $D^n y = y^{(n)}$. Allo scopo di semplificare le notazioni si fa la convenzione che D^0 rappresenti l'identità I dello spazio $C^\infty(\mathbb{R})$. In un certo senso è come se $D^0 y$ significasse derivare zero volte la funzione y (cioè lasciarla invariata).

In generale, dato un intervallo J e dato un intero n non negativo, con $C^n(J)$ si denota lo spazio vettoriale delle funzioni (reali) di classe C^n definite in J . Analogamente $C^\infty(J)$ rappresenta lo spazio delle funzioni C^∞ in J . Si osservi che $C^\infty(J)$ è un sottospazio di $C^n(J)$ qualunque sia n e che $C^n \subseteq C^{n-1}$ per ogni $n \in \mathbb{N}$ (in base alla definizione di funzione C^∞ si può addirittura affermare che $C^\infty(J)$ è l'intersezione di tutti gli spazi $C^n(J)$). Si osservi che l'operatore D è definito anche in $C^1(J)$, e non solo in $C^\infty(J)$, e manda $C^1(J)$ in $C^0(J)$. In generale D manda $C^n(J)$ in $C^{n-1}(J)$ qualunque sia $n \in \mathbb{N}$ (ed è per questo motivo che manda $C^\infty(J)$ in sé).

Si fa notare che l'equazione lineare del prim'ordine

$$y' = a(x)y + b(x),$$

dove $a(x)$ e $b(x)$ sono funzioni continue in un intervallo J , può essere scritta nella forma

$$Ly = b,$$

dove $y \in C^1(J)$ rappresenta la funzione incognita, $b \in C^0(J)$ è il termine noto ed $L: C^1(J) \rightarrow C^0(J)$ è quell'operatore lineare che ad ogni funzione $y(x)$ di classe C^1 (in J) associa la funzione continua $y'(x) - a(x)y(x)$.

89 - Venerdì 26/03/10

Un'equazione differenziale di ordine n si dice *lineare* se è del tipo

$$y^{(n)} + a_{n-1}(x)y^{(n-1)} + a_{n-2}(x)y^{(n-2)} + \dots + a_1(x)y' + a_0(x)y = h(x),$$

dove $a_0(x), \dots, a_{n-1}(x)$ e $h(x)$ sono funzioni continue in un intervallo J (di solito $J = \mathbb{R}$). Le funzioni $a_0(x), \dots, a_{n-1}(x)$ si dicono i *coefficienti* dell'equazione e $h(x)$ rappresenta il *termine noto*. Quando $h(x) \equiv 0$ (cioè $h(x) = 0$ per ogni $x \in J$), l'equazione si dice *omogenea*.

Si osservi che per le equazioni lineari vale il teorema di esistenza e unicità, dato che le derivate rispetto alle variabili $y, y', \dots, y^{(n-1)}$ sono i coefficienti $a_0(x), a_1(x), \dots, a_{n-1}(x)$, che abbiamo supposto continui.

Teorema^{sd} (di persistenza). *Le soluzioni massimali di un'equazione differenziale lineare sono tutte persistenti, ossia sono definite in tutto l'intervallo J in cui sono definiti i coefficienti e il termine noto.*

Da ora in avanti ci occuperemo esclusivamente di equazioni differenziali lineari a *coefficienti costanti* con termine noto definito in tutto \mathbb{R} e di classe C^∞ ; ossia di equazioni del tipo

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = h(x),$$

dove a_0, \dots, a_{n-1} sono numeri reali e $h \in C^\infty(\mathbb{R})$. In questo caso, per il precedente teorema, le soluzioni massimali sono definite in tutto \mathbb{R} . Si potrebbe provare^{fac}, usando il principio di induzione, che tali soluzioni sono tutte di classe C^∞ .

Notiamo che un'equazione lineare a coefficienti costanti

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = h(x)$$

si può scrivere in modo sintetico nella forma

$$Ly = h,$$

dove $L: C^\infty(\mathbb{R}) \rightarrow C^\infty(\mathbb{R})$ è l'operatore lineare, detto *operatore differenziale* (di ordine n), che ad ogni $y(x)$ di classe C^∞ in \mathbb{R} associa la funzione

$$y^{(n)}(x) + a_{n-1}y^{(n-1)}(x) + \dots + a_1y'(x) + a_0y(x),$$

anch'essa di classe C^∞ .

Si osservi che il suddetto operatore lineare può essere rappresentato nel modo seguente:

$$L = D^n + a_{n-1}D^{n-1} + a_{n-2}D^{n-2} + \dots + a_1D + a_0I,$$

mettendo così in evidenza come L si possa esprimere mediante somma e composizione di operatori lineari più elementari.

L'operatore L si può ottenere a partire dal polinomio

$$p(\lambda) = \lambda^n + a_{n-1}\lambda^{n-1} + a_{n-2}\lambda^{n-2} + \dots + a_1\lambda + a_0,$$

detto *polinomio caratteristico* dell'equazione differenziale considerata, semplicemente sostituendo D al posto della variabile λ (notiamo che il termine a_0 si può scrivere $a_0\lambda^0$, e sostituendo D al posto di λ si ha $a_0D^0 = a_0I$).

In generale, se $p(\lambda)$ è un polinomio qualunque (anche non associato ad un'equazione differenziale), l'operatore che si ottiene sostituendo D al posto di λ si denota con $p(D)$.

Esercizio. Scrivere il polinomio caratteristico dell'equazione differenziale $y'' + \omega^2 y = 0$ e calcolarne le radici.

Osservazione. Ogni equazione differenziale lineare a coefficienti costanti con termine noto $h \in C^\infty(\mathbb{R})$ si può scrivere nella forma $p(D)y = h$, dove $p(\lambda)$ è il polinomio caratteristico dell'equazione.

Ad esempio, il polinomio caratteristico dell'equazione differenziale

$$y''' - 2y'' - y' + 3y = x - \cos x$$

è

$$p(\lambda) = \lambda^3 - 2\lambda^2 - \lambda + 3$$

e l'operatore differenziale associato è

$$p(D) = D^3 - 2D^2 - D + 3I.$$

L'equazione può quindi essere scritta nella forma $p(D)y = h$, dove $h(x) = x - \cos x$.

Teorema. *Le soluzioni di un'equazione differenziale lineare non omogenea*

$$P(D)y = h$$

si ottengono sommando ad una soluzione dell'equazione non omogenea tutte le possibili soluzioni dell'equazione omogenea associata. In altre parole, se \hat{y} è una soluzione dell'equazione non omogenea, ogni altra soluzione è del tipo $y = \hat{y} + u$, dove $P(D)u = 0$.

Dimostrazione. Mostriamo prima che, fissata una soluzione (detta particolare) \hat{y} dell'equazione non omogenea, ogni funzione del tipo $y = \hat{y} + u$, dove u soddisfa la condizione $P(D)u = 0$, è ancora una soluzione dell'equazione non omogenea. Per la linearità di $P(D)$ si ha infatti $P(D)(\hat{y} + u) = P(D)\hat{y} + P(D)u = P(D)\hat{y} = h$.

Rimane da provare che se y è una qualunque soluzione dell'equazione non omogenea, allora la differenza $u := y - \hat{y}$ è una soluzione dell'omogenea. Risulta infatti

$$P(D)u = P(D)(y - \hat{y}) = P(D)y - P(D)\hat{y} = h - h = 0,$$

e l'affermazione è dimostrata. □

In base al precedente risultato, il problema di determinare tutte le soluzioni di un'equazione differenziale lineare non omogenea si scinde in due sottoproblemi:

- 1) risolvere l'equazione omogenea associata;
- 2) trovare almeno una soluzione dell'equazione non omogenea.

Occupiamoci prima dei metodi per risolvere le equazioni omogenee a coefficienti costanti. È necessario prima richiamare alcune nozioni di algebra lineare.

Definizione. Una *combinazione lineare* di n funzioni $y_1(x), y_2(x), \dots, y_n(x)$ dello spazio $C^\infty(\mathbb{R})$ è una funzione del tipo

$$c_1 y_1(x) + c_2 y_2(x) + \dots + c_n y_n(x),$$

dove c_1, c_2, \dots, c_n sono n costanti (detti coefficienti della combinazione lineare).

Ad esempio, un polinomio di grado minore o uguale ad n non è altro che una combinazione lineare delle funzioni $1, x, x^2, \dots, x^n$ (il polinomio è di grado n se il coefficiente di x^n è diverso da zero).

Definizione. Si dice che n funzioni di $C^\infty(\mathbb{R})$, $y_1(x), y_2(x), \dots, y_n(x)$, sono *linearmente indipendenti* se dall'uguaglianza

$$c_1 y_1(x) + c_2 y_2(x) + \dots + c_n y_n(x) = 0, \quad \forall x \in \mathbb{R}$$

segue $c_1 = c_2 = \dots = c_n = 0$; ossia, se l'unica combinazione lineare che dà la funzione (identicamente) nulla è quella con i coefficienti tutti nulli.

Esempio. Mostriamo che le funzioni $\cos \omega x$ e $\sin \omega x$ (dove $\omega > 0$) sono linearmente indipendenti. Supponiamo infatti che la funzione

$$y(x) := a \cos \omega x + b \sin \omega x$$

sia (identicamente) nulla. Poiché $y(x)$ è zero per ogni x , deve esserlo anche per $x = 0$. Ponendo $x = 0$ si ottiene $a = 0$. Per provare che anche il coefficiente b è nullo, basta porre $x = \pi/(2\omega)$.

Esercizio. Provare che se α_1 e α_2 sono due numeri reali differenti, allora le funzioni $e^{\alpha_1 x}$ e $e^{\alpha_2 x}$ sono linearmente indipendenti.

Suggerimento. Se la funzione $y(x) := a e^{\alpha_1 x} + b e^{\alpha_2 x}$ è identicamente nulla, allora lo è anche la sua derivata. In particolare si ha $y(0) = 0$ e $y'(0) = 0$. Applicare il Teorema di Cramer al sistema (di due equazioni in due incognite) così ottenuto per dedurre che $a = b = 0$.

Esercizio. Provare che le funzioni $1, x$ e x^2 sono linearmente indipendenti.

Esercizio. Dato $\alpha \in \mathbb{R}$, provare che le funzioni $e^{\alpha x}$ e $x e^{\alpha x}$ 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.

Il risultato che segue è una facile conseguenza dei teoremi di esistenza e unicità e di persistenza. Per brevità ne omettiamo la dimostrazione.

Teorema^{sd} (della dimensione del nucleo di un operatore differenziale). *L'insieme delle soluzioni (massimali) di un'equazione differenziale omogenea di ordine n a coefficienti costanti (o, più in generale, a coefficienti di classe C^∞) è un sottospazio vettoriale n -dimensionale di $C^\infty(\mathbb{R})$.*

Osservazione (riguardante il metodo per risolvere le equazioni lineari omogenee). In base al suddetto teorema possiamo affermare che se di un'equazione differenziale lineare omogenea di ordine n a coefficienti costanti se ne determinano n soluzioni linearmente indipendenti, tutte le altre si ottengono combinando linearmente le n soluzioni trovate.

Ad esempio, è immediato verificare che le funzioni $\cos \omega x$ e $\sin \omega x$ (con $\omega > 0$) sono soluzioni dell'equazione differenziale $y'' + \omega^2 y = 0$ (detta *equazione del moto armonico*). Poiché, come abbiamo visto, tali funzioni sono linearmente indipendenti, ogni soluzione della suddetta equazione è del tipo $y(x) = a \cos \omega x + b \sin \omega x$, con a e b costanti arbitrarie.

Esercizio. Provare che ogni funzione del tipo $y(x) = a \sin \omega x + b \cos \omega x$ (con $\omega > 0$) può essere rappresentata nella forma $y(x) = A \sin(\omega x + \varphi)$, con $A \geq 0$ (la costante ω si dice *pulsazione*, A si chiama *ampiezza* e φ è la *fase* dell'*oscillazione* $y(x)$). Si osservi che $y(x)$ è una funzione periodica e se ne determini il periodo T in funzione della pulsazione ω .

Esercizio. Sia $\omega > 0$. Mostrare che l'applicazione da \mathbb{C} in $C^\infty(\mathbb{R})$ che ad ogni numero complesso $a + ib$ associa la funzione $a \sin \omega x + b \cos \omega x$ è lineare ed iniettiva. Pertanto, è un isomorfismo tra \mathbb{C} e il sottospazio bidimensionale di $C^\infty(\mathbb{R})$ costituito dalle soluzioni dell'equazione differenziale (del moto armonico) $y'' + \omega^2 y = 0$.

Esercizio. Consideriamo la funzione periodica $y(x) = a \sin \omega x + b \cos \omega x$ (con $\omega > 0$). Mostrare che l'ampiezza e la fase di tale funzione coincidono, rispettivamente, col modulo e l'argomento del numero complesso $a + ib$.

Esercizio. Provare che e^x e e^{-x} sono due soluzioni linearmente indipendenti dell'equazione differenziale $y'' = y$. Trovare la soluzione generale di detta equazione e determinare quella che verifica il seguente problema di Cauchy: $y(0) = 0$, $y'(0) = 1$ (risposta: $\sinh x$).

Esercizio. Verificare che $\sinh x$ e $\cosh x$ sono soluzioni dell'equazione differenziale $y'' = y$ e provare che sono linearmente indipendenti. Determinare la soluzione generale di $y'' = y$ combinando linearmente le suddette funzioni e risolvere il seguente problema di Cauchy: $y(0) = 1$, $y'(0) = 1$ (è la funzione e^x).

Il polinomio caratteristico di un'equazione differenziale (lineare) a coefficienti costanti di ordine n gioca un ruolo fondamentale nella risoluzione dell'equazione omogenea associata. Come vedremo, la conoscenza delle radici di detto polinomio ci permetterà di determinare facilmente n soluzioni linearmente indipendenti (dell'omogenea associata). Dato che l'insieme delle soluzioni di un'equazione differenziale omogenea è uno spazio vettoriale di dimensione uguale all'ordine dell'equazione, la soluzione generale (dell'omogenea associata) si otterrà combinando linearmente le n soluzioni trovate.

90 - Mercoledì 31/03/10

Il seguente risultato, la cui dimostrazione risulterà chiara nella prossima lezione (dopo aver introdotto il concetto di soluzione complessa di un'equazione differenziale lineare), fornisce una regola pratica per trovare due soluzioni linearmente indipendenti di un'equazione differenziale omogenea del second'ordine a coefficienti costanti.

Teorema fondamentale per la risoluzione delle edo del second'ordine (lineari, omogenee, a coefficienti costanti). *Sia $y'' + by' + cy = 0$ un'equazione omogenea del second'ordine a coefficienti (reali) costanti. Se il polinomio caratteristico $p(\lambda) = \lambda^2 + b\lambda + c$ ha due radici reali e distinte α_1 e α_2 , allora $e^{\alpha_1 x}$ e $e^{\alpha_2 x}$ sono due soluzioni linearmente indipendenti dell'equazione differenziale. Se $p(\lambda)$ ha una radice doppia α , allora due soluzioni linearmente indipendenti sono $e^{\alpha x}$ e $x e^{\alpha x}$. 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 delle oscillazioni smorzate:

$$y'' + 2\varepsilon y' + \omega^2 y = 0,$$

dove $0 < \varepsilon < \omega$. Il polinomio caratteristico è $p(\lambda) = \lambda^2 + 2\varepsilon\lambda + \omega^2$, le cui radici sono $-\varepsilon \pm i\sqrt{\omega^2 - \varepsilon^2}$. Dunque la soluzione generale è data da

$$y(x) = c_1 e^{-\varepsilon x} \operatorname{sen} \left(\sqrt{\omega^2 - \varepsilon^2} x \right) + c_2 e^{-\varepsilon x} \operatorname{cos} \left(\sqrt{\omega^2 - \varepsilon^2} x \right)$$

o, equivalentemente, da

$$y(x) = A e^{-\varepsilon x} \operatorname{sen} \left(\sqrt{\omega^2 - \varepsilon^2} x + \varphi \right),$$

dove A e φ sono costanti arbitrarie.

Il metodo illustrato nel precedente teorema riguardante le equazioni del second'ordine si estende facilmente al caso generale. Data un'equazione differenziale

$$y^{(n)} + a_{n-1} y^{(n-1)} + a_{n-2} y^{(n-2)} + \dots + a_1 y' + a_0 y = 0,$$

ad ogni radice reale α del polinomio caratteristico

$$p(\lambda) = \lambda^n + a_{n-1} \lambda^{n-1} + a_{n-2} \lambda^{n-2} + \dots + a_1 \lambda + a_0,$$

si fa corrispondere:

- la soluzione $e^{\alpha x}$ se α è semplice;
- le due soluzioni $e^{\alpha x}$ e $x e^{\alpha x}$ se α è doppia;
- ...
- le k soluzioni $e^{\alpha x}, x e^{\alpha x}, \dots, x^{k-1} e^{\alpha x}$ se α è di molteplicità k .

Ad ogni coppia di radici complesse $\alpha \pm i\beta$ di $p(\lambda)$ si fa corrispondere:

- le due soluzioni $e^{\alpha x} \sin \beta x$ e $e^{\alpha x} \cos \beta x$ se $\alpha + i\beta$ è semplice (in tal caso lo è anche $\alpha - i\beta$);
- le quattro soluzioni

$$e^{\alpha x} \sin \beta x, \quad e^{\alpha x} \cos \beta x, \quad x e^{\alpha x} \sin \beta x, \quad x e^{\alpha x} \cos \beta x$$

se $\alpha + i\beta$ è doppia (in tal caso lo è anche $\alpha - i\beta$);

...

- le $2k$ soluzioni

$$e^{\alpha x} \sin \beta x, \quad e^{\alpha x} \cos \beta x, \quad \dots, \quad x^{k-1} e^{\alpha x} \sin \beta x, \quad x^{k-1} e^{\alpha x} \cos \beta x$$

se $\alpha + i\beta$ è di molteplicità k (in tal caso lo è anche $\alpha - i\beta$).

Poiché il polinomio caratteristico $p(\lambda)$ ammette esattamente n radici in campo complesso (contate con la loro molteplicità), le soluzioni dell'equazione differenziale che si determinano col suddetto metodo sono esattamente n . Si potrebbe provare (ma non lo facciamo) che le soluzioni così determinate sono linearmente indipendenti. Pertanto, dato che l'insieme delle soluzioni di

$$y^{(n)} + a_{n-1}y^{(n-1)} + a_{n-2}y^{(n-2)} + \dots + a_1y' + a_0y = 0$$

è uno spazio vettoriale n -dimensionale, l'integrale generale della suddetta equazione si ottiene combinando linearmente le n soluzioni trovate.

91 - Mercoledì 31/03/10

Esempio. Consideriamo l'equazione differenziale $y^{(n)} = 0$. Il polinomio caratteristico è $p(\lambda) = \lambda^n$, e quindi $\alpha = 0$ è una radice di molteplicità n (n radici coincidenti). In base al procedimento illustrato si può affermare che le funzioni

$$e^{0x}, \quad x e^{0x}, \quad x^2 e^{0x}, \quad \dots, \quad x^{n-1} e^{0x}$$

sono n soluzioni linearmente indipendenti dell'equazione differenziale considerata. Dunque, l'integrale generale è

$$y(x) = a_0 + a_1x + a_2x^2 + \dots + a_{n-1}x^{n-1}.$$

Di conseguenza, le funzioni con derivata n -esima nulla sono (soltanto) i polinomi di grado minore o uguale ad $n - 1$. Si osservi che ciò estende una nota conseguenza del Teorema di Lagrange: le funzioni con derivata nulla sono (soltanto) i polinomi di grado zero, cioè le costanti.

Esempio. Determiniamo l'integrale generale dell'equazione differenziale

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

Il polinomio caratteristico è $p(\lambda) = \lambda(\lambda^2 + 2\lambda + 2)$, le cui radici sono $\lambda_1 = 0$, $\lambda_2 = -1 + i$ e $\lambda_3 = -1 - i$. Quindi, per quanto visto, la soluzione generale è data da

$$y(x) = c_1 + c_2 e^{-x} \sin x + c_3 e^{-x} \cos x;$$

che può essere anche scritta nella forma

$$y(x) = c + Ae^{-x} \operatorname{sen}(x + \varphi),$$

con c , A e φ costanti arbitrarie.

Esempio. Risolviamo la seguente equazione differenziale:

$$y^{(4)} + 8y^{(2)} + 16y = 0.$$

Occorre trovare le quattro radici del polinomio caratteristico $p(\lambda) = \lambda^4 + 8\lambda^2 + 16$. Consideriamo quindi l'equazione algebrica $\lambda^4 + 8\lambda^2 + 16 = 0$. Ponendo $\lambda^2 = \mu$ si ha $\mu^2 + 8\mu + 16 = (\mu + 4)^2 = 0$, da cui si ricavano due soluzioni coincidenti: $\mu_1 = -4$ e $\mu_2 = -4$. Avendo posto $\lambda^2 = \mu$, si ottiene $\lambda^2 = -4$, dove il valore -4 va considerato due volte. Le quattro radici del polinomio caratteristico sono quindi $\lambda_1 = 2i$, $\lambda_2 = -2i$, $\lambda_3 = 2i$ e $\lambda_4 = -2i$. Poiché $2i$ è radice doppia (così come lo è la sua coniugata $-2i$) la soluzione generale dell'equazione differenziale è

$$y(x) = c_1 \operatorname{sen} 2x + c_2 \cos 2x + c_3 x \operatorname{sen} 2x + c_4 x \cos 2x$$

o, equivalentemente,

$$y(x) = A_1 \operatorname{sen}(2x + \varphi_1) + A_2 x \operatorname{sen}(2x + \varphi_2).$$

Esempio. Tra tutte le soluzioni dell'equazione differenziale

$$y^{(4)} + y = 0$$

determiniamo quelle che tendono a zero per $x \rightarrow +\infty$. Le soluzioni dell'equazione algebrica $\lambda^4 + 1 = 0$ sono le radici quarte del numero -1 (che è un numero complesso di modulo 1 e argomento π). Risolviamo quindi l'equazione

$$[r(\cos \varphi + i \operatorname{sen} \varphi)]^4 = \cos \pi + i \operatorname{sen} \pi$$

(o, equivalentemente, l'equazione $(re^{i\varphi})^4 = e^{i\pi}$). Si ha

$$r^4 \cos(4\varphi) + ir^4 \operatorname{sen}(4\varphi) = \cos \pi + i \operatorname{sen} \pi,$$

da cui si deduce $r = 1$ e $4\varphi = \pi + 2k\pi$, con $k \in \mathbb{Z}$. Per ottenere le quattro radici basta dare a k quattro valori consecutivi. Per $k = 0$ si ottiene $\lambda_0 = \cos(\pi/4) + i \operatorname{sen}(\pi/4) = \sqrt{2}/2 + i\sqrt{2}/2$, e per $k = 1$ si ha $\lambda_1 = \cos(3\pi/4) + i \operatorname{sen}(3\pi/4) = -\sqrt{2}/2 + i\sqrt{2}/2$. Le altre due radici si ottengono ponendo $k = -1$ e $k = 2$ oppure, più semplicemente, considerando le coniugate delle due radici trovate (dipende dal fatto che il polinomio caratteristico è reale). Pertanto, la soluzione generale è

$$y(x) = e^{-\frac{\sqrt{2}}{2}x} \left(c_1 \operatorname{sen} \frac{\sqrt{2}}{2}x + c_2 \cos \frac{\sqrt{2}}{2}x \right) + e^{\frac{\sqrt{2}}{2}x} \left(c_3 \operatorname{sen} \frac{\sqrt{2}}{2}x + c_4 \cos \frac{\sqrt{2}}{2}x \right)$$

o, equivalentemente,

$$y(x) = Ae^{-\frac{\sqrt{2}}{2}x} \operatorname{sen} \left(\frac{\sqrt{2}}{2}x + \varphi \right) + Be^{\frac{\sqrt{2}}{2}x} \operatorname{sen} \left(\frac{\sqrt{2}}{2}x + \psi \right).$$

Quindi, le soluzioni che tendono a zero per $x \rightarrow +\infty$ sono quelle con $B = 0$. Cioè

$$y(x) = Ae^{-\frac{\sqrt{2}}{2}x} \operatorname{sen} \left(\frac{\sqrt{2}}{2}x + \varphi \right),$$

dove A e φ sono due arbitrarie costanti.

Esempio. Consideriamo l'equazione

$$y''' - 3y'' + 4y = 0.$$

Il polinomio caratteristico è $p(\lambda) = \lambda^3 - 3\lambda^2 + 4$. Si vede subito che $\lambda_1 = -1$ è una radice, e quindi il polinomio è divisibile per $\lambda + 1$. Il quoziente della divisione è $\lambda^2 - 4\lambda + 4$, perciò $p(\lambda) = (\lambda^2 - 4\lambda + 4)(\lambda + 1)$. Le altre due radici sono $\lambda_2 = 2$ e $\lambda_3 = 2$ (ossia, 2 è una radice doppia). Dunque, la soluzione generale è

$$y(x) = c_1e^{-x} + c_2e^{2x} + c_3xe^{2x}.$$

Supponiamo ora di voler trovare, tra tutte le soluzioni, quella che verifica il seguente problema di Cauchy: $y(1) = 0$, $y'(1) = 0$, $y''(1) = 0$. La risposta è semplice, non c'è bisogno di fare calcoli: la soluzione si vede a occhio (ed è unica, perché l'equazione è lineare e sono perciò verificate le ipotesi del teorema di esistenza e unicità).

92 - Mercoledì 31/03/10

Ricordiamo che, dato un arbitrario numero complesso $z = x + iy$, si definisce

$$e^z = e^x(\cos y + i \operatorname{sen} y).$$

Ossia, e^{x+iy} è un numero complesso di modulo e^x e argomento y . Pertanto, in base alla suddetta definizione, ogni numero complesso z può essere rappresentato nella forma $\rho e^{i\theta}$ (detta *forma esponenziale dei numeri complessi*), dove ρ è il modulo (di z) e θ è un qualunque argomento. Ricordiamo inoltre che la funzione e^z gode della seguente *proprietà fondamentale*: $e^{z_1+z_2} = e^{z_1}e^{z_2}$, $\forall z_1, z_2 \in \mathbb{C}$.

La funzione e^z , denotata anche $\exp z$, è un esempio di funzione complessa di variabile complessa; ossia, di una funzione con dominio e codominio in \mathbb{C} . Un altro esempio è dato dalla potenza z^n o, più in generale, da un polinomio di variabile complessa a coefficienti reali o complessi. Anche le funzioni razionali complesse, ossia le funzioni ottenute mediante il quoziente di polinomi complessi, sono esempi di funzioni complesse di variabile complessa.

Per provare il teorema fondamentale per la risoluzione delle edo del second'ordine (o, più in generale, per provare la validità del metodo illustrato nella lezione precedente) occorre

introdurre la nozione di *soluzione complessa* di una edo lineare a coefficienti costanti (il motivo risulterà chiaro in seguito).

Osserviamo che una funzione $z(x)$ da \mathbb{R} in \mathbb{C} si scrive nella forma $z(x) = \alpha(x) + i\beta(x)$, dove $\alpha(x)$ e $\beta(x)$, dette rispettivamente *parte reale* e *parte immaginaria* della funzione $z(x)$, sono funzioni reali di variabile reale. La derivata di $z(x)$ è, per definizione, la funzione $z'(x) = \alpha'(x) + i\beta'(x)$. Si dice che $z(x)$ è di classe C^∞ (o che appartiene allo spazio $C^\infty(\mathbb{R}, \mathbb{C})$) se sono di classe C^∞ entrambe le funzioni $\alpha(x)$ e $\beta(x)$. Ovviamente, ogni funzione dello spazio $C^\infty(\mathbb{R})$ può essere pensata anche in $C^\infty(\mathbb{R}, \mathbb{C})$, con parte immaginaria nulla (quindi $C^\infty(\mathbb{R})$ è un sottospazio vettoriale di $C^\infty(\mathbb{R}, \mathbb{C})$).

Esercizio^{fac.} Fissato un numero complesso $\lambda = \alpha + i\beta$, si consideri la funzione da \mathbb{R} in \mathbb{C} definita da $e^{\lambda x}$. Provare che tale funzione è derivabile e risulta $De^{\lambda x} = \lambda e^{\lambda x}$.

Definizione (di soluzione complessa di una edo lineare a coefficienti costanti). Sia

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_1y' + a_0y = h(x)$$

un'equazione differenziale (lineare a coefficienti costanti) con termine noto $h \in C^\infty(\mathbb{R})$. Una funzione $z \in C^\infty(\mathbb{R}, \mathbb{C})$ è una *soluzione complessa* di tale equazione se

$$z^{(n)}(x) + a_{n-1}z^{(n-1)}(x) + \dots + a_1z'(x) + a_0z(x) = h(x),$$

per ogni $x \in \mathbb{R}$.

Consideriamo, ad esempio, l'equazione differenziale

$$y'' + y = 0.$$

Proviamo a vedere se ammette soluzioni del tipo $z(x) = e^{\lambda x}$, dove $\lambda = \alpha + i\beta$ è un numero complesso. Si ha $z'(x) = \lambda e^{\lambda x}$ e $z''(x) = \lambda^2 e^{\lambda x}$. Quindi $z(x)$ è soluzione se e solo se

$$(\lambda^2 + 1)e^{\lambda x} = 0, \quad \forall x \in \mathbb{R},$$

ossia (essendo $e^{\lambda x} \neq 0$) se e solo se $\lambda^2 + 1 = 0$, da cui si ricava $\lambda = \pm i$. Pertanto

$$e^{ix} = \cos x + i \sin x \quad \text{e} \quad e^{-ix} = \cos x - i \sin x$$

sono soluzioni dell'equazione differenziale considerata, e sono le uniche del tipo $e^{\lambda x}$.

Esercizio. Dato $\lambda \in \mathbb{C}$, mostrare che la funzione complessa $z(x) = e^{\lambda x}$ è soluzione dell'equazione differenziale $y'' + by' + cy = 0$ se e solo se $\lambda^2 + b\lambda + c = 0$ (cioè se e solo se λ è radice del polinomio caratteristico dell'equazione).

Esercizio^{fac.} Provare che la combinazione lineare a coefficienti complessi di due (o più) soluzioni complesse di una edo lineare omogenea è ancora una soluzione.

In base al suddetto esercizio si può affermare che l'insieme delle soluzioni complesse di una edo lineare omogenea è uno spazio vettoriale sui complessi (sono gli scalari dello spazio).

Si potrebbe dimostrare (ma non lo facciamo) che, come nel caso reale, la dimensione di tale spazio è uguale all'ordine dell'equazione.

Esercizio. Siano $z(x) = \alpha(x) + i\beta(x)$ e $\bar{z}(x) = \alpha(x) - i\beta(x)$ due funzioni complesse e coniugate di $C^\infty(\mathbb{R}, \mathbb{C})$. Si osservi che

$$\frac{z(x) + \bar{z}(x)}{2} = \alpha(x) \quad \text{e} \quad \frac{z(x) - \bar{z}(x)}{2i} = \beta(x).$$

Dedurre da ciò che se $z(x) = \alpha(x) + i\beta(x)$ è soluzione di un'equazione differenziale omogenea, allora lo sono anche la parte reale $\alpha(x)$ e la parte immaginaria $\beta(x)$.

Esercizio. Dedurre dagli esercizi precedenti che se $\alpha + i\beta$, con $\beta \neq 0$, è una radice del polinomio $\lambda^2 + b\lambda + c$, allora le due funzioni $e^{\alpha x} \cos \beta x$ e $e^{\alpha x} \sin \beta x$ sono soluzioni dell'equazione differenziale $y'' + by' + cy = 0$.

Esercizio ^{fac.} Fissato $\alpha \in \mathbb{R}$, provare che le due funzioni $e^{\alpha x}$ e $xe^{\alpha x}$ sono soluzioni dell'equazione differenziale $y'' + by' + cy = 0$ se e solo se α è radice doppia del polinomio $\lambda^2 + b\lambda + c$ (in tal caso risulta $\Delta = b^2 - 4c = 0$ e $\alpha = -b/2$).

93 - Venerdì 09/04/10

Abbiamo già visto come sia possibile, conoscendo le radici del polinomio caratteristico, determinare l'integrale generale di un'equazione differenziale del tipo

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y = h(x),$$

quando $h(x) \equiv 0$, cioè nel caso omogeneo. Ricordiamo che se il termine noto $h(x)$ non è identicamente nullo, l'integrale generale si ottiene sommando ad una soluzione della non omogenea (comunemente detta *soluzione particolare*) tutte le possibili soluzioni dell'omogenea associata. Tuttavia, a meno che il termine noto non sia di un tipo molto speciale, non è possibile in generale determinare in modo esplicito una soluzione dell'equazione non omogenea. È necessario allora ricorrere a metodi numerici, tramite i quali si riesce a tabulare, con una buona approssimazione, la soluzione (particolare) dell'equazione (non omogenea) che soddisfa una determinata condizione di Cauchy (si può, ad esempio, fissare i valori iniziali $y(0) = y'(0) = \dots = y^{(n-1)}(0) = 0$).

Ci sono alcuni casi (piuttosto speciali, ma importanti per le applicazioni) in cui è possibile determinare con carta e penna una soluzione particolare dell'equazione non omogenea. Ad esempio, quando il termine noto si presenta nella seguente forma:

$$h(x) = r(x)e^{\alpha x} \sin \beta x + s(x)e^{\alpha x} \cos \beta x,$$

dove α e β sono due numeri reali (eventualmente nulli), mentre $r(x)$ e $s(x)$ sono polinomi (eventualmente di grado zero, cioè costanti, uno dei quali può essere nullo). In questo caso si cerca una soluzione in una forma molto particolare (ecco l'origine dell'espressione "soluzione particolare"!). 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 $h(x)$ si presenta in una delle suddette forme e a determinarne il corrispondente numero $\lambda = \alpha + i\beta$. Ecco alcuni esempi:

$h(x)$	$\alpha + i\beta$	$r(x)$	$s(x)$
$\text{sen } x$	i	1	0
$x^2 e^{-2x}$	-2	0	x^2
-3	0	0	-3
$-e^x \cos 2x$	$1 + 2i$	0	-1
$x e^{-2x} \text{sen } \pi x$	$-2 + \pi i$	x	0
$x^3 - x$	0	0	$x^3 - x$
$2e^x$	1	0	2
$2 \text{sen } \omega x - x \cos \omega x$	$i\omega$	2	$-x$

Regola pratica (per determinare una soluzione particolare). Consideriamo l’equazione differenziale

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y = h(x)$$

e supponiamo che il termine noto $h(x)$ sia del tipo

$$e^{\alpha x} (r(x) \text{sen } \beta x + s(x) \cos \beta x),$$

dove $r(x)$ e $s(x)$ sono due polinomi (non occorre che abbiano lo stesso grado) e $\alpha, \beta \in \mathbb{R}$. Denotiamo con k il massimo tra il grado di $r(x)$ e il grado di $s(x)$ (cioè k è il grado del polinomio $r(x) + s(x)$).

Se $\alpha + i\beta$ non è radice del polinomio caratteristico, allora si cerca una soluzione particolare della forma

$$e^{\alpha x} (a(x) \text{sen } \beta x + b(x) \cos \beta x),$$

dove $a(x)$ e $b(x)$ sono polinomi di grado k , 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 la forma relativa al caso precedente (cioè moltiplicando per x^2 quella in cui $\alpha + i\beta$ non è radice).

Se $\alpha + i\beta$ è radice tripla... uffa!

Esempio. Tra tutte le soluzioni dell’equazione differenziale

$$y'' = x e^{-x},$$

determiniamo quella che verifica il problema di Cauchy $y(0) = 0$ e $y'(0) = 0$.

L’equazione omogenea associata è $y'' = 0$ e il suo polinomio caratteristico ha due radici coincidenti: $\lambda_1 = 0$ e $\lambda_2 = 0$. Quindi, la soluzione generale dell’equazione omogenea è $u(x) = c_0 + c_1x$. Occorre trovare una soluzione particolare dell’equazione non omogenea (per poi sommarla alla soluzione generale dell’omogenea). Osserviamo che $\lambda = -1$ non è radice del polinomio caratteristico. Quindi, in base alla regola pratica appena illustrata, cerchiamo una soluzione del tipo

$$\hat{y}(x) = (a + bx)e^{-x}.$$

Derivando due volte si ha

$$\hat{y}''(x) = (a - 2b)e^{-x} + bxe^{-x}.$$

Quindi $\hat{y}(x)$ è soluzione se (e solo se) è verificata la condizione

$$(a - 2b)e^{-x} + (b - 1)xe^{-x} = 0, \quad \forall x \in \mathbb{R},$$

ossia se (e solo se) $a - 2b = 0$ e $b - 1 = 0$ (si osservi che l'affermazione “ $\hat{y}(x)$ è soluzione solo se $a - 2b = 0$ e $b - 1 = 0$ ”, anche se non è importante per il nostro scopo, è conseguenza del fatto che le funzioni e^{-x} e xe^{-x} sono linearmente indipendenti). Dunque, una soluzione (particolare) dell'equazione non omogenea è data da $\hat{y}(x) = (2 + x)e^{-x}$ e la soluzione generale è

$$y(x) = c_0 + c_1x + (2 + x)e^{-x}.$$

Determiniamo ora c_0 e c_1 in modo che siano soddisfatte le condizioni assegnate. Poiché $y(0) = c_0 + 2$ e $y'(0) = c_1 - 1$, ponendo $y(0) = 0$ e $y'(0) = 0$ si ricava $c_0 = -2$ e $c_1 = 1$. Pertanto, la soluzione che verifica il problema di Cauchy assegnato è

$$y(x) = -2 + x + (2 + x)e^{-x}.$$

94 - Venerdì 09/04/10

Esempio. Tra tutte le soluzioni dell'equazione differenziale

$$y'' - 2y' + 2y = 2x,$$

determiniamo quella che verifica il problema di Cauchy $y(0) = 2$, $y'(0) = 2$.

Troviamo prima tutte le soluzioni dell'equazione differenziale e imponiamo poi le condizioni iniziali assegnate. Il polinomio caratteristico ha due radici complesse coniugate: $\lambda_1 = 1 + i$ e $\lambda_2 = 1 - i$. Quindi, in campo complesso, la soluzione generale dell'equazione omogenea associata è data da

$$u(x) = c_1e^{(1+i)x} + c_2e^{(1-i)x}.$$

A questo punto, volendo, si potrebbe determinare anche la soluzione generale in campo reale, ma per risolvere il nostro problema non è necessario: il teorema di esistenza e unicità (valido anche per le soluzioni complesse) ci assicura che il problema di Cauchy ha una e una sola soluzione, e questa deve essere reale (visto che sia l'equazione sia le condizioni iniziali sono reali). Procediamo quindi in campo complesso. Occorre trovare una soluzione particolare dell'equazione non omogenea e sommarla alla soluzione generale dell'omogenea. Poiché $\lambda = 0$ non è radice del polinomio caratteristico, in base alla regola pratica si cerca una soluzione del tipo $\hat{y}(x) = ax + b$. Con semplici calcoli si vede che $\hat{y}(x)$ è soluzione se (e solo se) $a = 1$ e $b = 1$. La soluzione generale (in campo complesso) dell'equazione non omogenea è dunque

$$y(x) = c_1e^{(1+i)x} + c_2e^{(1-i)x} + x + 1,$$

dove c_1 e c_2 sono arbitrarie costanti complesse. Occorre determinare c_1 e c_2 in modo che si abbia $y(0) = 2$, $y'(0) = 2$. Poiché

$$y'(x) = (1+i)c_1e^{(1+i)x} + (1-i)c_2e^{(1-i)x} + 1,$$

le costanti c_1 e c_2 devono verificare il sistema

$$\begin{cases} c_1 + c_2 = 1 \\ (1+i)c_1 + (1-i)c_2 = 1, \end{cases}$$

da cui si ricava l'unica soluzione $(c_1, c_2) = (1/2, 1/2)$. Possiamo concludere che la soluzione dell'equazione differenziale che verifica le condizioni iniziali assegnate è

$$y(x) = \frac{e^{(1+i)x} + e^{(1-i)x}}{2} + x + 1 = e^x \cos x + x + 1$$

che, come ci aspettavamo, è una funzione reale. Si invita a verificare che la funzione $y(x) = e^x \cos x + x + 1$ è effettivamente una soluzione dell'equazione differenziale considerata e che soddisfa le condizioni iniziali $y(0) = 2$, $y'(0) = 2$.

Esercizio. Sia

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y = h_1(x) + h_2(x) \quad (18.1)$$

un'equazione differenziale lineare con termine noto $h(x) = h_1(x) + h_2(x)$. Provare che se $\hat{y}_1(x)$ è una soluzione (particolare) di

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y = h_1(x)$$

e $\hat{y}_2(x)$ è una soluzione (particolare) di

$$y^{(n)} + a_{n-1}y^{(n-1)} + \dots + a_0y = h_2(x),$$

allora $\hat{y}(x) = \hat{y}_1(x) + \hat{y}_2(x)$ è una soluzione (particolare) della (18.1).

Esempio. Determiniamo l'integrale generale della seguente equazione differenziale:

$$y'' - 4y = e^x - xe^{-2x} + 1.$$

Il polinomio caratteristico è $p(\lambda) = \lambda^2 - 4$ e le sue radici sono $\lambda_1 = -2$ e $\lambda_2 = 2$. Quindi, la soluzione generale dell'equazione omogenea associata è

$$u(x) = c_1e^{-2x} + c_2e^{2x}.$$

Per determinare una soluzione particolare osserviamo che il termine noto è somma di tre termini, tutti del tipo $q(x)e^{\alpha x}$, dove $q(x)$ è un polinomio. Poiché 1 non è radice del polinomio caratteristico $p(\lambda)$, il primo termine ci induce a cercare una soluzione dell'equazione $p(D)y = e^x$ del tipo $\hat{y}_1(x) = ae^x$. Riguardo al secondo termine, osserviamo che -2 è una radice semplice del polinomio caratteristico, e quindi si cerca una soluzione di $p(D)y = -xe^{-2x}$ nella forma $\hat{y}_2(x) = x(bx + c)e^{-2x}$. Infine, per quanto riguarda il terzo

termine, si cerca una soluzione di $p(D)y = 1$ del tipo $\hat{y}_3(x) = d$, cioè costante (infatti, 0 non è radice di $p(\lambda)$). Cerchiamo quindi una soluzione dell'equazione differenziale non omogenea nella forma

$$\hat{y}(x) = ae^x + x(bx + c)e^{-2x} + d.$$

Sostituendo $\hat{y}(x)$ nell'equazione differenziale, si ricava $a = -1/3$, $b = 1/8$, $c = 1/16$ e $d = -1/4$. Pertanto, l'integrale generale dell'equazione non omogenea è

$$y(x) = c_1e^{-2x} + c_2e^{2x} - \frac{e^x}{3} + \left(\frac{x^2}{8} + \frac{x}{16}\right)e^{-2x} - \frac{1}{4},$$

dove c_1 e c_2 sono due arbitrarie costanti.

Esercizio. Determinare l'integrale generale (in campo reale) delle seguenti equazioni differenziali:

$$y'' + y = \sin x, \quad y'' = x^3 + e^x \cos 2x, \quad y''' + y'' = x, \quad y''' + y = e^{-x}.$$

Esercizio^{fac}. Determinare per quali valori del parametro $\varepsilon > 0$ le soluzioni dell'equazione differenziale

$$y'' + 2\varepsilon y' + \omega^2 y = 0,$$

sono (tutte) non oscillanti (cioè non cambiano segno infinite volte).

Esempio (di calcolo di una formula di MacLaurin di una soluzione di un problema di Cauchy). Determiniamo la formula di MacLaurin del secondo ordine della soluzione del seguente problema di Cauchy:

$$y' = 1 + \sin y, \quad y(0) = 0.$$

Il teorema di esistenza e unicità ci assicura che il suddetto problema ammette un'unica soluzione (massimale) $y(x)$. La formula di MacLaurin del second'ordine di $y(x)$ è data da

$$y(x) = y(0) + y'(0)x + \frac{y''(0)}{2}x^2 + \varepsilon(x)x^2.$$

Pertanto, per determinarla è sufficiente calcolare $y(0)$, $y'(0)$ e $y''(0)$. Il valore $y(0)$ è già noto ed è assegnato dalla condizione di Cauchy (i.e. $y(0) = 0$). Per ricavare gli altri due ricordiamo che, in base alla definizione di soluzione, risulta

$$y'(x) = 1 + \sin y(x),$$

per ogni x nell'intervallo J in cui $y(x)$ è definita. Pertanto, per $x = 0$ si ottiene

$$y'(0) = 1 + \sin y(0) = 1 + \sin 0 = 1.$$

Infine, dall'uguaglianza $y'(x) = 1 + \sin y(x)$, tenendo conto della regola della derivata di una funzione composta, si ricava

$$y''(x) = y'(x) \cos y(x),$$

che è ancora un'uguaglianza, la quale, per $x = 0$ ci dà

$$y''(0) = y'(0) \cos y(0) = 1(\cos 0) = 1.$$

Dunque

$$y(x) = x + \frac{x^2}{2} + \varepsilon(x)x^2.$$

Esempio (di calcolo di una formula di MacLaurin di una soluzione di un problema di Cauchy). Supponiamo di voler determinare la formula di MacLaurin del terzo ordine della soluzione $y(x)$ del seguente problema di Cauchy:

$$\begin{cases} y' = 1 - 2x + x^2 \cos y \\ y(0) = 0. \end{cases}$$

L'unicità della soluzione è assicurata dal fatto che la funzione $f(x, y) = 1 - 2x + x^2 \cos y$ è di classe C^1 (è addirittura di classe C^∞). Osserviamo, innanzi tutto, che è sufficiente ricavare la formula di MacLaurin del second'ordine della derivata $y'(x)$. Infatti, dai coefficienti del polinomio di MacLaurin di $y'(x)$ si deducono facilmente le derivate $g'(0)$ e $g''(0)$ della funzione $g(x) := y'(x)$, e queste (tenendo conto che $y(0) = 0$ e $y'(0) = g(0)$) ci permettono di scrivere la formula (di MacLaurin) del terz'ordine di $y(x)$, dato che la derivata k -esima di $g(x)$ non è altro che la derivata $(k+1)$ -esima di $y(x)$. Si osservi ora che per ricavare la formula del second'ordine di $g(x) = 1 - 2x + x^2 \cos(y(x))$ è sufficiente conoscere la formula di ordine zero della funzione composta $\cos(y(x))$. Infatti il resto $\varepsilon(x)$ di tale formula, moltiplicato per x^2 , dà $\varepsilon(x)x^2$. Dato che $\cos(y(x))$ è una funzione continua, si ha $\cos(y(x)) = \cos(y(0)) + \varepsilon(x) = 1 + \varepsilon(x)$. Quindi $g(x) = 1 - 2x + x^2 + \varepsilon(x)x^2$, da cui (in base all'unicità della formula di MacLaurin) si deduce $g(0) = 1$, $g'(0)/1! = -2$ e $g''(0)/2! = 1$. Pertanto, $y'(0) = 1$, $y''(0) = -2$ e $y^{(3)}(0) = 2$. Dunque, tenendo conto che $y(0) = 0$, si ottiene

$$y(x) = \sum_{k=0}^3 \frac{y^{(k)}(0)}{k!} x^k + \varepsilon(x)x^3 = x - x^2 + \frac{x^3}{3} + \varepsilon(x)x^3.$$

95 - Mercoledì 14/04/10

Definizione (di punto di accumulazione). Dato $A \subseteq \mathbb{R}^k$ e dato $p_0 \in \mathbb{R}^k$ (non necessariamente appartenente ad A), si dice che p_0 è un punto di *accumulazione* per A se esiste una successione $\{p_n\}$ in $A \setminus \{p_0\}$ (cioè di punti di A diversi da p_0) che converge a p_0 .

Si potrebbe provare^{fac} che un punto $p_0 \in \mathbb{R}^k$ è di accumulazione per un insieme $A \subseteq \mathbb{R}^k$ se e solo se ogni intorno di p_0 contiene infiniti punti di A (o, equivalentemente, contiene almeno un punto di A diverso da p_0).

Esercizio. Provare che se un punto p_0 è interno ad un insieme $A \subseteq \mathbb{R}^k$, allora è anche di accumulazione (mostrare con un esempio che non è vero il contrario).

Definizione (di limite con le successioni). Sia $f: A \rightarrow \mathbb{R}$ una funzione reale di k variabili reali. Dato un punto p_0 di accumulazione per il dominio A di f e dato un numero reale esteso γ (i.e. $\gamma = \ell \in \mathbb{R}$, $\gamma = -\infty$ oppure $\gamma = +\infty$), si dice che γ è il *limite di $f(p)$ per p che tende a p_0* , e si scrive

$$\lim_{p \rightarrow p_0} f(p) = \gamma,$$

se per ogni successione $\{p_n\}$ in $A \setminus \{p_0\}$ che converge a p_0 , risulta

$$\lim_{n \rightarrow +\infty} f(p_n) = \gamma.$$

In tal caso si dice anche che $f(p)$ *tende a γ per $p \rightarrow p_0$* , e si scrive

$$f(p) \rightarrow \gamma \quad \text{per} \quad p \rightarrow p_0.$$

C'è un modo diverso per definire il concetto di limite, che evita di ricorrere alla nozione di successione. Occorre però distinguere tre casi (come prima, p_0 è un punto di accumulazione per il dominio A di f):

- 1) $f(p) \rightarrow \ell \in \mathbb{R}$ per $p \rightarrow p_0$;
- 2) $f(p) \rightarrow +\infty$ per $p \rightarrow p_0$;
- 3) $f(p) \rightarrow -\infty$ per $p \rightarrow p_0$.

Definizione (di limite con gli intorni). Siano f , A e p_0 come nella precedente definizione.

- 1) Si dice che $f(p) \rightarrow \ell \in \mathbb{R}$ per $p \rightarrow p_0$ se per ogni $\varepsilon > 0$ esiste un $\delta > 0$ tale che se $p \in A$, $p \neq p_0$ e $\|p - p_0\| < \delta$, allora $|f(p) - \ell| < \varepsilon$.
- 2) Si dice che $f(p) \rightarrow +\infty$ per $p \rightarrow p_0$ se per ogni $c > 0$ esiste un $\delta > 0$ tale che se $p \in A$, $p \neq p_0$ e $\|p - p_0\| < \delta$, allora $f(p) > c$.
- 3) Si dice che $f(p) \rightarrow -\infty$ per $p \rightarrow p_0$ se per ogni $c > 0$ esiste un $\delta > 0$ tale che se $p \in A$, $p \neq p_0$ e $\|p - p_0\| < \delta$, allora $f(p) < -c$.

Teorema^{sd} (di collegamento per i limiti). *La definizione di limite con le successioni e la definizione di limite con gli intorni sono equivalenti.*

Definizione (di punto isolato). Sia A un sottoinsieme di \mathbb{R}^k . Un punto $p_0 \in A$ si dice un *punto isolato* di A se non è di accumulazione (ossia, se esiste un intorno di p_0 il cui unico punto di A è p_0 stesso).

Osservazione. Sia $f: A \rightarrow \mathbb{R}$ una funzione definita in un sottoinsieme $A \subseteq \mathbb{R}^k$ e sia p_0 un punto del dominio A di f . In base alla definizione di continuità (data prima di introdurre il concetto di limite), se p_0 è di accumulazione per A , allora f è continua in p_0 se e solo se $f(p) \rightarrow f(p_0)$ per $p \rightarrow p_0$. Se invece p_0 è un punto isolato di A , allora (sempre in base alla definizione) f risulta continua in p_0 (provarlo per esercizio^{fac}), ma in tal caso non ha senso parlare di limite per $p \rightarrow p_0$.

Osservazione. Dato che tutti i punti di un aperto sono di accumulazione, se si considerano soltanto le funzioni definite in insiemi aperti, la nozione di continuità può essere data mediante il concetto di limite.

Esercizio. Provare, mediante la definizione di limite, che

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

Ricordiamo che con la notazione $\overline{\mathbb{R}}$ si intende l'insieme dei numeri reali estesi, cioè l'insieme dei numeri reali \mathbb{R} più due simboli: $-\infty$ e $+\infty$. Per convenzione, dato un qualunque $x \in \mathbb{R}$, risulta $-\infty < x < +\infty$. Si definiscono inoltre le seguenti operazioni (la lettera a rappresenta un arbitrario numero reale e ∞ sta per $-\infty$ o $+\infty$): $-\infty + a = -\infty$, $+\infty + a = +\infty$, $(-\infty) + (-\infty) = -\infty$, $(+\infty) + (+\infty) = +\infty$, $a(\pm\infty) = \pm\infty$, se $a > 0$ e $a(\pm\infty) = \mp\infty$ se $a < 0$, $a/\infty = 0$, $a/0 = \infty$ se $a \neq 0$, $\infty/0 = \infty$. Ogni eventuale definizione delle espressioni $(+\infty) + (-\infty)$, $0/0$, $0 \cdot \infty$ e ∞/∞ , chiamate *forme indeterminate*, porterebbe a delle incoerenze, e quindi non conviene introdurla.

Teorema^{sd} (fondamentale dei limiti per funzioni di due o più variabili). *Siano f_1 ed f_2 due funzioni reali di k variabili reali. Supponiamo che, nei reali estesi, si abbia*

$$\lim_{p \rightarrow p_0} f_1(p) = \gamma_1 \quad e \quad \lim_{p \rightarrow p_0} f_2(p) = \gamma_2.$$

Allora, quando ha senso, per $p \rightarrow p_0$ risulta:

- 1) $f_1(p) + f_2(p) \rightarrow \gamma_1 + \gamma_2$;
- 2) $f_1(p)f_2(p) \rightarrow \gamma_1\gamma_2$;
- 3) $f_1(p)/f_2(p) \rightarrow \gamma_1/\gamma_2$.

Esempio. Calcoliamo il

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\operatorname{sen} xy - xy}{2x^3y^3 + x^4y^3}.$$

Per la formula di MacLaurin del terz'ordine di $\operatorname{sen} x$ risulta

$$\operatorname{sen} x = x - \frac{x^3}{6} + \varepsilon(x)x^3, \quad \forall x \in \mathbb{R}.$$

Poiché la suddetta uguaglianza è verificata qualunque sia il numero x , sostituendo x col numero xy si ottiene

$$\operatorname{sen} xy = xy - \frac{x^3y^3}{6} + \varepsilon(xy)x^3y^3, \quad \forall (x,y) \in \mathbb{R}^2,$$

dove $\varepsilon(xy)$ è la funzione (da \mathbb{R}^2 in \mathbb{R}) ottenuta mediante la composizione della funzione (da \mathbb{R}^2 in \mathbb{R}) xy con la funzione (da \mathbb{R} in \mathbb{R}) $\varepsilon(x)$. Si ha quindi

$$\begin{aligned} \lim_{(x,y) \rightarrow (0,0)} \frac{\operatorname{sen} xy - xy}{2x^3y^3 + x^4y^3} &= \lim_{(x,y) \rightarrow (0,0)} \frac{xy - x^3y^3/6 + \varepsilon(xy)x^3y^3 - xy}{2x^3y^3 + x^4y^3} \\ &= \lim_{(x,y) \rightarrow (0,0)} \frac{-x^3y^3/6 + \varepsilon(xy)x^3y^3}{2x^3y^3 + x^4y^3} = \lim_{(x,y) \rightarrow (0,0)} \frac{-1/6 + \varepsilon(xy)}{2 + x} = \frac{-1/6 + \varepsilon(0)}{2 + 0} = -\frac{1}{12}. \end{aligned}$$

Ricordiamo che i vettori di norma uno di \mathbb{R}^k si chiamano *direzioni* o *versori* (di \mathbb{R}^k). In particolare, in \mathbb{R}^2 una direzione è una coppia di numeri (h, k) tale che $h^2 + k^2 = 1$, e in tal caso esiste un angolo $\theta \in \mathbb{R}$ per il quale risulta $h = \cos \theta$ e $k = \sin \theta$.

Definizione (di limite direzionale). Sia $f: A \rightarrow \mathbb{R}$ una funzione definita in un sottoinsieme $A \subseteq \mathbb{R}^k$ e sia p_0 un punto del dominio A di f . Fissata una direzione $v \in \mathbb{R}^k$, il

$$\lim_{t \rightarrow 0} f(p_0 + tv),$$

ammesso che esista, si dice *limite direzionale di f in p_0 nella direzione v* .

Si osservi che il limite direzionale riguarda una ordinaria funzione reale di variabile reale.

Osservazione. Se il limite per $(x, y) \rightarrow (x_0, y_0)$ di $f(x, y)$ è uguale a $\lambda \in \overline{\mathbb{R}}$, allora il *limite direzionale*

$$\lim_{t \rightarrow 0} f(x_0 + t \cos \theta, y_0 + t \sin \theta)$$

esiste ed è uguale a λ qualunque sia la *direzione* $v = (\cos \theta, \sin \theta) \in \mathbb{R}^2$. Di conseguenza, se per una funzione $f(x, y)$ il limite direzionale (in (x_0, y_0)) dipende dalla direzione (o non esiste per qualche direzione), allora $f(x, y)$ non ammette limite per $(x, y) \rightarrow (x_0, y_0)$.

Dimostrazione. Per esercizio^{fac}.

Esempio. Consideriamo il

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

Fissata un'arbitraria direzione $v = (\cos \theta, \sin \theta) \in \mathbb{R}^2$, eseguiamo le sostituzioni $x = t \cos \theta$ e $y = t \sin \theta$ nella funzione

$$f(x, y) = \frac{xy}{x^2 + y^2}$$

e calcoliamo il limite per $t \rightarrow 0$. Si ha

$$\lim_{t \rightarrow 0} \frac{t^2 \cos \theta \sin \theta}{t^2 \cos^2 \theta + t^2 \sin^2 \theta} = \cos \theta \sin \theta.$$

Il limite direzionale dipende dunque dalla direzione $(\cos \theta, \sin \theta)$. Si può pertanto concludere che la funzione $xy/(x^2 + y^2)$ non ammette limite per $(x, y) \rightarrow (0, 0)$.

Vedremo in seguito, con un esempio, che il limite direzionale può esistere ed essere uguale in ogni direzione, ma il limite può non esistere.

96 - Mercoledì 14/04/10

Per le funzioni 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):

- teorema di unicità del limite;
- teorema della permanenza del segno;

- teorema del confronto dei limiti;
- teorema dei carabinieri.

Come per il caso di una variabile, una funzione di più variabili $f: A \rightarrow \mathbb{R}$ si dice *limitata* se esiste una costante $c > 0$ tale che $|f(p)| \leq c$ per ogni $p \in A$ (in questo caso si dice che f è *dominata* da c).

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 $p \rightarrow p_0$.*

Come applicazione del precedente corollario, determiniamo il seguente limite:

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

dove con ρ , come è consuetudine, abbiamo denotato la funzione $\|(x, y)\| = \sqrt{x^2 + y^2}$, cioè la distanza di (x, y) dall'origine. Osserviamo che la funzione $f(x, y) = xy^2/\rho^2$ è il prodotto di tre funzioni: x/ρ , y/ρ e 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, essendo limitate le prime due funzioni (verificare che sono dominate dalla costante $c = 1$), si può concludere che $xy^2/\rho^2 \rightarrow 0$ per $(x, y) \rightarrow (0, 0)$.

Esercizio. Sia $f: A \rightarrow \mathbb{R}$ una funzione reale di k variabili reali e sia p_0 un punto di accumulazione per A . Provare che se esiste ed è finito il limite di $f(p)$ per $p \rightarrow p_0$, allora esiste un intorno di p_0 in cui f è limitata (cioè esistono due costanti positive c e δ tali che $|f(p)| \leq c$ per ogni $p \in I(p_0, \delta) \cap A$).

Esercizio. Mostrare che la funzione x/ρ (dove, ricordiamo, $\rho = \sqrt{x^2 + y^2}$), pur essendo limitata, non ammette limite per $(x, y) \rightarrow (0, 0)$. Non ci si lasci ingannare dall'assonanza (peculiare della lingua italiana) tra l'aggettivo "limitato" e il sostantivo "limite", due concetti completamente differenti che in inglese e in francese corrispondono a vocaboli foneticamente distanti tra loro.

Esercizio. Sia $f: A \rightarrow \mathbb{R}$ una funzione reale di k variabili reali e sia p_0 un punto di accumulazione per A . Provare che se $|f(p)| \rightarrow +\infty$ per $p \rightarrow p_0$, allora f non è una funzione limitata (in A).

Esercizio. Provare che la funzione

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

non ammette un intorno di $(0, 0)$ in cui è limitata.

Suggerimento. Fissare un arbitrario intorno di $(0, 0)$ e mostrare che in tale intorno esiste almeno un punto p_0 di accumulazione per il dominio di f per il quale è verificata l'ipotesi dell'esercizio precedente (esistono addirittura infiniti punti con tale proprietà).

Esercizio. Provare che non esiste il

$$\lim_{(x,y) \rightarrow (0,0)} \frac{x^3 y^2}{x - y^2}.$$

Suggerimento. Se il limite esistesse, allora dovrebbe essere finito perché la funzione è nulla lungo gli assi cartesiani (più precisamente, il limite non potrebbe essere altro che zero). Quindi (in base ad un precedente esercizio) la funzione dovrebbe essere limitata in un intorno di $(0, 0)$, ma ciò contraddice l'esercizio precedente.

Esercizio. Provare che il limite direzionale in $(0, 0)$ della funzione

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

esiste ed è lo stesso in ogni direzione (anche se, in base all'esercizio precedente, la funzione $f(x, y)$ non ammette limite per $(x, y) \rightarrow (0, 0)$).

Esempio. Calcoliamo il

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

Consideriamo una qualunque successione $\{p_n\} = \{(x_n, y_n)\}$ nel dominio della funzione

$$f(x, y) = \frac{\text{sen}(x^2 + y)}{x^2 + y}$$

tale che $(x_n, y_n) \rightarrow (0, 0)$. Ovviamente è anche verificata la condizione $(x_n, y_n) \neq (0, 0)$, perché $(0, 0)$ non sta nel dominio di f (pur essendo di accumulazione). Poiché

$$\lim_{x \rightarrow 0} \frac{\text{sen } x}{x} = 1,$$

per la condizione necessaria del teorema di collegamento (quello di Analisi Matematica I), si ha

$$\lim_{n \rightarrow +\infty} \frac{\text{sen } \alpha_n}{\alpha_n} = 1,$$

dove $\{\alpha_n\}$ è la successione reale definita da $\alpha_n = x_n^2 + y_n$ (che è sempre diversa da zero perché, per ipotesi, (x_n, y_n) sta nel dominio di f). Quindi, data l'arbitrarietà della successione $\{p_n\} = \{(x_n, y_n)\}$, in base alla definizione di limite (con le successioni), possiamo concludere che

$$\lim_{(x,y) \rightarrow (0,0)} \frac{\text{sen}(x^2 + y)}{x^2 + y} = 1.$$

Esercizio. Provare che non esiste il limite per $(x, y) \rightarrow (0, 0)$ della funzione

$$f(x, y) = \text{sen} \left(\frac{1}{x^2 + y^2} \right).$$

97 - Mercoledì 14/04/10

Ricordiamo che una funzione si dice *di variabile reale* se il suo dominio è un sottoinsieme di \mathbb{R} (indipendentemente dal suo codominio), si dice *di due variabili reali* se il suo dominio è un sottoinsieme di \mathbb{R}^2 . Più in generale, si dice che una funzione è *di k variabili reali* se il suo dominio è un sottoinsieme di \mathbb{R}^k (di solito si richiede che il dominio abbia almeno un punto interno).

Se il codominio di una funzione è \mathbb{R} allora, indipendentemente dal suo dominio, si dice che è una *funzione reale*, se il codominio è \mathbb{C} si dice che è una *funzione complessa*. In generale una funzione si dice *vettoriale* o *a valori vettoriali* se il suo codominio è uno spazio vettoriale (indipendentemente dal suo dominio e dalla sua immagine). Ad esempio, la funzione che ad ogni $t \in \mathbb{R}$ associa la coppia di numeri $(1 - t, t^2)$ è una funzione vettoriale (a valori in \mathbb{R}^2). Un altro esempio è dato da $f: \mathbb{R}^2 \rightarrow \mathbb{R}^3$ così definita: $(x, y) \mapsto (x - y, \sin x, x + y^2)$.

In generale una funzione $f: A \rightarrow \mathbb{R}^s$ definita in un sottoinsieme $A \subseteq \mathbb{R}^k$ si dice *vettoriale di variabile vettoriale* (di solito si richiede che il dominio abbia almeno un punto interno). Si osservi che una funzione a valori in \mathbb{R}^s è costituita da s funzioni reali, f_1, f_2, \dots, f_s , dette *funzioni componenti*. Ad esempio, la funzione $(x, y) \mapsto (x - y, \sin x, x - y^2)$ è composta da tre funzioni reali di due variabili reali: $f_1(x, y) = x - y$, $f_2(x, y) = \sin x$, $f_3(x, y) = x - y^2$.

Il limite di una funzione vettoriale di variabile vettoriale si definisce in modo analogo a come si è fatto per le funzioni reali di più variabili reali.

Definizione (di limite per una funzione vettoriale di variabile vettoriale). Sia $f: A \rightarrow \mathbb{R}^s$ un'applicazione definita in un sottoinsieme A di \mathbb{R}^k e sia p_0 un punto di accumulazione per il dominio A di f . Si dice che $f(p)$ *tende ad un elemento* $q_0 \in \mathbb{R}^s$ *per* p *che tende a* p_0 , e si scrive $f(p) \rightarrow q_0$ per $p \rightarrow p_0$, se **per ogni** successione $\{p_n\}$ in $A \setminus \{p_0\}$ che converge a p_0 , risulta

$$\lim_{n \rightarrow +\infty} f(p_n) = q_0.$$

In tal caso si dice anche che q_0 è *il limite di* $f(p)$ *per* p *che tende a* p_0 , e si scrive

$$\lim_{p \rightarrow p_0} f(p) = q_0.$$

Si potrebbe dimostrare che $f(p) \rightarrow q_0$ per $p \rightarrow p_0$ se e solo se la prima componente $f_1(p)$ di $f(p)$ tende alla prima componente di q_0 , la seconda componente di $f(p)$ tende alla seconda componente di q_0 , e così via per tutte le altre componenti.

Sia $f: A \rightarrow \mathbb{R}^s$ un'applicazione definita in un sottoinsieme A di \mathbb{R}^k . Diremo che f è continua in un punto $p_0 \in A$ se sono continue (in p_0) le sue s funzioni componenti. Per quanto visto prima 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

$$\lim_{p \rightarrow p_0} f(p) = f(p_0).$$

Ovviamente, dire che $f: A \rightarrow \mathbb{R}^s$ è continua, senza aggiungere altro, significa che è continua in ogni punto del suo dominio A .

Si dirà che un'applicazione $f: A \rightarrow \mathbb{R}^s$ definita in un sottoinsieme A di \mathbb{R}^k è derivabile (di classe C^n , di classe C^∞) se sono derivabili (di classe C^n , di classe C^∞) tutte le sue funzioni componenti. Insomma f è “quello che uno vuole” se sono “quello che uno vuole” tutte le sue componenti.

98 - Venerdì 16/04/10

Dato un aperto U di \mathbb{R}^3 (o di \mathbb{R}^2 o di \mathbb{R}^k), un *campo vettoriale* (definito su U) è una “legge” \mathbf{v} che ad ogni punto $p \in U$ assegna un vettore $\mathbf{v}(p) \in \mathbb{R}^3$. Quindi, se $\{\mathbf{i}, \mathbf{j}, \mathbf{k}\}$ denota la base canonica di \mathbb{R}^3 , ogni campo vettoriale $\mathbf{v}: U \rightarrow \mathbb{R}^3$ si può rappresentare nel seguente modo:

$$\mathbf{v} = A(x, y, z)\mathbf{i} + B(x, y, z)\mathbf{j} + C(x, y, z)\mathbf{k},$$

dove A , B e C sono tre funzioni reali (che supporremo continue) definite in U (dette *componenti del campo*).

Ovviamente, un campo vettoriale nel piano avrà due sole componenti e si rappresenterà nel seguente modo:

$$\mathbf{v} = A(x, y)\mathbf{i} + B(x, y)\mathbf{j}.$$

Un campo vettoriale si dice di classe C^n (risp. C^∞) se sono di classe C^n (risp. C^∞) le sue funzioni componenti.

Una funzione reale, continua su un aperto U di \mathbb{R}^3 (o di \mathbb{R}^2 o di \mathbb{R}^k), si dice anche un *campo scalare* (definito su U).

Il gradiente (di cui ora diamo la definizione) è un importante operatore (lineare) che trasforma campi scalari in campi vettoriali. Come vedremo, il gradiente di una funzione reale indica la direzione di massima crescita della funzione e la sua norma (o modulo) misura il tasso di crescita della funzione nella direzione del gradiente.

Definizione (di gradiente di una funzione reale). Data una funzione $f: U \rightarrow \mathbb{R}$ di classe C^1 su un aperto U di \mathbb{R}^3 (o di \mathbb{R}^2 o di \mathbb{R}^k), il campo vettoriale che ha per componenti le derivate parziali di f si chiama *gradiente* di f (denotato $\text{grad } f$ o ∇f).

Ad esempio, in \mathbb{R}^3 si ha

$$\nabla f = \frac{\partial f}{\partial x}\mathbf{i} + \frac{\partial f}{\partial y}\mathbf{j} + \frac{\partial f}{\partial z}\mathbf{k}.$$

Il simbolo ∇ si chiama “nabla” e rappresenta un operatore lineare dallo spazio vettoriale delle funzioni di classe C^1 su un aperto U di \mathbb{R}^3 (o di \mathbb{R}^2 o di \mathbb{R}^k) a valori nello spazio vettoriale dei campi vettoriali definiti su U (se il concetto non è chiaro, non importa). L'operatore nabla può essere rappresentato nel seguente modo (come somma di tre operatori lineari):

$$\nabla = \frac{\partial}{\partial x}\mathbf{i} + \frac{\partial}{\partial y}\mathbf{j} + \frac{\partial}{\partial z}\mathbf{k}.$$

Definizione (di potenziale di un campo vettoriale e nozione di campo conservativo). Un campo vettoriale in \mathbb{R}^3 , $\mathbf{v} = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$, si dice *conservativo* se ammette un *potenziale*; cioè una funzione f tale che $\nabla f = \mathbf{v}$ (vale a dire $A = \partial f/\partial x$, $B = \partial f/\partial y$ e $C = \partial f/\partial z$). Analogamente, in \mathbb{R}^2 un campo vettoriale $\mathbf{v} = A\mathbf{i} + B\mathbf{j}$ è *conservativo* se esiste una funzione derivabile f (detta *potenziale*) tale che $\partial f/\partial x = A$ e $\partial f/\partial y = B$.

Ovviamente, se f è un potenziale di un campo \mathbf{v} , allora lo è anche $f + c$, qualunque sia la costante $c \in \mathbb{R}$.

Si osservi che una funzione reale di variabile reale può essere pensata come un campo vettoriale in \mathbb{R} . In tal caso il concetto di potenziale coincide con quello di primitiva.

Definizione (di campo irrotazionale). Un campo vettoriale $\mathbf{v} = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$ (di classe C^1 su un aperto di \mathbb{R}^3) si chiama *irrotazionale* se (supposto di classe C^1) sono verificate le seguenti tre condizioni:

$$\frac{\partial C}{\partial y} = \frac{\partial B}{\partial z}, \quad \frac{\partial A}{\partial z} = \frac{\partial C}{\partial x}, \quad \frac{\partial B}{\partial x} = \frac{\partial A}{\partial y}.$$

Analogamente, in due variabili, un campo vettoriale $\mathbf{v} = A\mathbf{i} + B\mathbf{j}$ è *irrotazionale* se

$$\frac{\partial B}{\partial x} = \frac{\partial A}{\partial y}.$$

Teorema. *Condizione necessaria affinché un campo vettoriale di classe C^1 sia conservativo è che sia irrotazionale.*

Dimostrazione. Sia $\mathbf{v} = A\mathbf{i} + B\mathbf{j}$ un campo vettoriale (in \mathbb{R}^2) di classe C^1 (ricordiamo che in tal caso $A(x, y)$ e $B(x, y)$ sono funzioni di classe C^1 definite su aperto di \mathbb{R}^2). Supponiamo che $f(x, y)$ sia un potenziale di \mathbf{v} . Allora $f(x, y)$ è necessariamente di classe C^2 , dato che le sue derivate parziali (rispetto alla prima e alla seconda variabile) coincidono (rispettivamente) con $A(x, y)$ e con $B(x, y)$, che abbiamo supposto di classe C^1 . Si può quindi applicare il Teorema di Schwarz alla funzione f ottenendo così la seguente uguaglianza:

$$\frac{\partial}{\partial y} \frac{\partial f}{\partial x}(x, y) = \frac{\partial}{\partial x} \frac{\partial f}{\partial y}(x, y).$$

Il risultato segue subito tenendo conto che $\partial f/\partial x = A$ e $\partial f/\partial y = B$. La dimostrazione nel caso in cui \mathbf{v} sia un campo in \mathbb{R}^3 è lasciata per esercizio^{fac} allo studente. \square

Vedremo in seguito (dopo aver introdotto gli integrali curvilinei) che la condizione che un campo vettoriale sia irrotazionale, senza opportune ipotesi sul dominio del campo, non assicura che sia conservativo.

Esempio. Il campo vettoriale

$$\mathbf{v} = (y^2 + \cos y)\mathbf{i} + (2xy + 1)\mathbf{j}$$

non ammette un potenziale. Per quale motivo?

99 - Venerdì 16/04/10

Esempio. Consideriamo il campo vettoriale

$$\mathbf{v} = (y^2 + \cos x)\mathbf{i} + (2xy + 1)\mathbf{j}.$$

Risulta $\partial(y^2 + \cos x)/\partial y = 2y$ e $\partial(2xy + 1)/\partial x = 2y$. Quindi \mathbf{v} è irrotazionale; ossia è soddisfatta la condizione necessaria affinché \mathbf{v} sia conservativo. Proviamo a vedere se effettivamente \mathbf{v} ammette un potenziale $f: \mathbb{R}^2 \rightarrow \mathbb{R}$. Se f esiste, si deve avere $\nabla f = \mathbf{v}$, ossia

$$\frac{\partial f}{\partial x} = y^2 + \cos x, \quad \frac{\partial f}{\partial y} = 2xy + 1.$$

Dalla prima uguaglianza si deduce che $f(x, y)$ è uguale a $xy^2 + \sin x$ più una costante rispetto alla variabile x (cioè più una funzione della sola y). Quindi

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

dove $\varphi(y)$ è un'arbitraria funzione della variabile y che occorre determinare. Derivando rispetto alla y l'espressione della f che abbiamo appena determinato, si ha $\partial f/\partial y = 2xy + \varphi'(y)$. Pertanto, tenendo conto che (se f è un potenziale) deve essere $\partial f/\partial y = 2xy + 1$, si ottiene $2xy + \varphi'(y) = 2xy + 1$. Quindi $\varphi'(y) = 1$ e, di conseguenza, $\varphi(y) = y + c$ (dove c è un'arbitraria costante). Si può concludere che se f è un potenziale di \mathbf{v} , allora necessariamente

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

Un semplice controllo mostra che effettivamente il gradiente di $xy^2 + \sin x + y + c$ è proprio $(y^2 + \cos x)\mathbf{i} + (2xy + 1)\mathbf{j}$.

Esempio. Studiamo il seguente campo vettoriale:

$$\mathbf{w} = 2yz\mathbf{i} + xz\mathbf{j} + (xy - 2z)\mathbf{k}.$$

Verifichiamo, innanzi tutto, se si tratta di un campo irrotazionale. Si ha $\partial(2yz)/\partial y = 2z$ e $\partial(xz)/\partial x = z$. Poiché le due funzioni $2z$ e z non coincidono, è inutile proseguire con il calcolo delle altre derivate: il campo vettoriale non è irrotazionale e quindi non ammette un potenziale (non è conservativo).

Esempio. Studiamo il seguente campo vettoriale:

$$\mathbf{w} = yz\mathbf{i} + xz\mathbf{j} + (xy - 2z)\mathbf{k}.$$

Verifichiamo, innanzi tutto, se si tratta di un campo irrotazionale. Si ha

$$\frac{\partial(yz)}{\partial y} = \frac{\partial(xz)}{\partial x} = z, \quad \frac{\partial(xz)}{\partial z} = \frac{\partial(xy - 2z)}{\partial y} = x, \quad \frac{\partial(xy - 2z)}{\partial x} = \frac{\partial yz}{\partial z} = y.$$

Il campo è quindi irrotazionale; ossia è verificata la condizione necessaria affinché il campo sia conservativo. Il campo, quindi, potrebbe ammettere un potenziale (vedremo in seguito che quando le tre componenti del campo sono definite in tutto lo spazio \mathbb{R}^3 , allora la

condizione necessaria per l'esistenza di un potenziale è anche sufficiente). Supponiamo quindi che esista $f: \mathbb{R}^3 \rightarrow \mathbb{R}$ tale che $\nabla f = \mathbf{w}$. Poiché la derivata di f rispetto alla x coincide con yz , la funzione $f(x, y, z)$ risulta uguale a xyz più una costante rispetto alla x (cioè più una funzione delle sole variabili y e z). Quindi $f(x, y, z) = xyz + \varphi(y, z)$. Occorre determinare la funzione $\varphi(y, z)$. Derivando rispetto alla y l'espressione della f che abbiamo appena determinato, si ottiene la funzione

$$xz + \frac{\partial \varphi}{\partial y}(y, z),$$

che deve coincidere con la seconda componente del campo \mathbf{w} . Pertanto

$$\frac{\partial \varphi}{\partial y}(y, z) = 0,$$

e ciò implica che $\varphi(y, z)$ è una funzione che dipende soltanto dalla z . Denotiamola con $\psi(z)$. Si ha quindi $f(x, y, z) = xyz + \psi(z)$. Poiché la derivata di f rispetto alla z deve coincidere con la terza componente del campo, si ottiene $\psi'(z) = -2z$. Quindi $\psi(z) = -z^2 + c$. Concludendo si ha $f(x, y, z) = xyz - z^2 + c$. Un semplice controllo ci assicura che effettivamente la funzione $xyz - z^2 + c$ è un potenziale di \mathbf{w} .

La divergenza è un operatore lineare che trasforma campi vettoriali in campi scalari, ed è una nozione di fondamentale importanza per le applicazioni alla fluidodinamica e alla teoria del potenziale (gravitazionale, elettrico, magnetico). Intuitivamente, il valore in un punto della divergenza di un campo vettoriale misura la tendenza del campo a divergere dal punto (quando il valore è positivo) o a convergere verso il punto (quando è negativo).

Definizione (di divergenza di un campo vettoriale e nozione di campo solenoidale). Sia

$$\mathbf{w} = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$$

un campo vettoriale di classe C^1 su un aperto di \mathbb{R}^3 . La *divergenza* di \mathbf{w} è la funzione (reale di tre variabili reali) definita da

$$\operatorname{div} \mathbf{w} = \frac{\partial A}{\partial x} + \frac{\partial B}{\partial y} + \frac{\partial C}{\partial z}.$$

Analogamente, se $\mathbf{w} = A\mathbf{i} + B\mathbf{j}$ è un campo vettoriale (di classe C^1) in \mathbb{R}^2 , la sua *divergenza* è il campo scalare

$$\operatorname{div} \mathbf{w} = \frac{\partial A}{\partial x} + \frac{\partial B}{\partial y}.$$

Un campo vettoriale si dice *solenoidale* se la sua divergenza è nulla.

Il rotore è un operatore (lineare) che trasforma campi vettoriali in \mathbb{R}^3 in campi vettoriali in \mathbb{R}^3 e che misura la tendenza di un campo vettoriale ad ammettere vortici. È un operatore di basilare importanza per le applicazioni alla teoria dell'elettromagnetismo.

Definizione (di rotore di un campo vettoriale in \mathbb{R}^3). Sia $\mathbf{w} = A\mathbf{i} + B\mathbf{j} + C\mathbf{k}$ un campo vettoriale di classe C^1 su un aperto di \mathbb{R}^3 . Si chiama *rotore* di \mathbf{w} il campo vettoriale $\text{rot } \mathbf{w}$ di componenti

$$\frac{\partial C}{\partial y} - \frac{\partial B}{\partial z}, \quad \frac{\partial A}{\partial z} - \frac{\partial C}{\partial x}, \quad \frac{\partial B}{\partial x} - \frac{\partial A}{\partial y}.$$

Come regola mnemonica, si usa scrivere

$$\text{rot } \mathbf{w} = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ A & B & C \end{pmatrix}$$

Si osservi che, in base alla nozione di campo irrotazionale precedentemente introdotta (è a pagina 183), un campo vettoriale in \mathbb{R}^3 è irrotazionale se e solo se il suo rotore è nullo.

Si osservi inoltre che in base alla condizione necessaria affinché un campo vettoriale sia conservativo (è a pagina 183), data una funzione f di classe C^2 su un aperto di \mathbb{R}^3 , risulta $\text{rot grad } f = 0$.

Un altro operatore di fondamentale importanza per le molteplici applicazioni alla Fisica e all'Ingegneria (per lo studio della propagazione ondosa, del flusso del calore, della deformazione dei materiali, del potenziale elettrico, ecc.) è l'operatore di Laplace (detto *laplaciano*), denotato con la lettera greca Δ (delta maiuscola). Esso trasforma campi scalari in campi scalari e si ottiene componendo il gradiente con la divergenza. Ad esempio, se f è una funzione di tre variabili (di classe C^2), si ha

$$\Delta f = \text{div grad } f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}.$$

Analoghe espressioni si ottengono in \mathbb{R}^2 e, più in generale, in \mathbb{R}^k .

100 - Mercoledì 21/04/10

Una funzione continua $\gamma: J \rightarrow \mathbb{R}^2$ definita in un intervallo (non banale) $J \subseteq \mathbb{R}$ si dice una *curva (parametrica)* in \mathbb{R}^2 . Una curva $\gamma(t)$ in \mathbb{R}^2 è quindi composta da due funzioni reali di variabile reale (dette funzioni componenti) che possiamo denotare, per esempio, $x(t)$ e $y(t)$, oppure $x_1(t)$ e $x_2(t)$, oppure $\gamma_1(t)$ e $\gamma_2(t)$, ecc. La variabile $t \in J$ si dice il *parametro* della curva, e anche questo può essere denotato con una qualunque lettera (dipende dal significato fisico o geometrico che si vuol dare alla curva). Ad esempio, la variabile t può essere interpretata come il tempo, e in tal caso $\gamma(t)$ rappresenta la posizione di un punto in \mathbb{R}^2 all'istante $t \in J$.

Analogamente, una curva in \mathbb{R}^3 è una funzione continua $\gamma: J \rightarrow \mathbb{R}^3$ definita in un intervallo (non banale) $J \subseteq \mathbb{R}$, ed è composta da tre funzioni (componenti) denotate $\gamma_1(t)$, $\gamma_2(t)$ e $\gamma_3(t)$, oppure $x(t)$, $y(t)$ e $z(t)$, oppure $x_1(t)$, $x_2(t)$ e $x_3(t)$, ecc.

Esercizio. Definire il concetto di curva in \mathbb{R}^s .

In base a quanto detto prima per le funzioni da \mathbb{R}^k in \mathbb{R}^s , una curva in \mathbb{R}^s è continua se è continua ogni sua componente, è derivabile se è derivabile ogni sua componente, è C^n se è C^n ogni sua componente, è C^∞ se è C^∞ ogni sua componente.

Definizione. Data una curva $\gamma: J \rightarrow \mathbb{R}^s$ di componenti $\gamma_1(t), \gamma_2(t), \dots, \gamma_s(t)$, la sua derivata $\gamma'(t_0)$ in un punto $t_0 \in J$ è (se esiste) il vettore di componenti $\gamma'_1(t_0), \gamma'_2(t_0), \dots, \gamma'_s(t_0)$. Si dice anche che $\gamma'(t_0)$ è il vettore tangente alla curva (parametrica) $\gamma(t)$ in $\gamma(t_0)$, o nel punto che corrisponde a t_0 . Anche ogni altro vettore proporzionale a $\gamma'(t_0)$ si dice tangente alla curva in $\gamma(t_0)$, ma con l'espressione "il vettore tangente" (con l'articolo determinativo) si intende proprio $\gamma'(t_0)$.

Si osservi che se $p(t)$ rappresenta la posizione all'istante t di un punto in \mathbb{R}^3 , allora $p'(t)$ è il vettore velocità e $p''(t)$ il vettore accelerazione.

Esercizio. Data una curva $p(t)$ in \mathbb{R}^3 (o in \mathbb{R}^k), provare che la funzione reale $\|p'(t)\|^2 = p'(t) \cdot p'(t)$ è costante se e solo se il vettore $p''(t)$ è ortogonale a $p'(t)$.

Il precedente esercizio ha una chiara interpretazione fisica: se la forza che agisce su un punto materiale è ortogonale al moto, allora l'energia cinetica del punto non varia (e viceversa).

Esercizio. Sia $p: \mathbb{R} \rightarrow \mathbb{R}^k$ la curva parametrica $p(t) = p_0 + tv$, dove p_0 è un punto (di \mathbb{R}^k) e v è un vettore (di \mathbb{R}^k). Provare che $p'(t) = v$ per ogni $t \in \mathbb{R}$ (se il vettore v è non nullo, allora $p(t)$ è una curva un po' speciale: è una retta parametrica, il cui significato cinematico, per $k = 2$ o $k = 3$, è quello di moto rettilineo uniforme con velocità v).

Esercizio. Data una curva $\gamma: J \rightarrow \mathbb{R}^k$, mostrare che la sua derivata $\gamma'(t_0)$ in un punto $t_0 \in J$ coincide col limite per $t \rightarrow t_0$ del rapporto incrementale in t_0 . In simboli:

$$\gamma'(t_0) = \lim_{t \rightarrow t_0} \frac{\gamma(t) - \gamma(t_0)}{t - t_0} = \lim_{h \rightarrow 0} \frac{\gamma(t_0 + h) - \gamma(t_0)}{h}.$$

Suggerimento. Usare il fatto che il limite di una funzione vettoriale si può eseguire componente per componente.

Teorema della derivata di una funzione composta^{sd} (caso speciale: regola della derivata a catena). Sia $f(x, y)$ una funzione reale di classe C^1 (in un aperto di \mathbb{R}^2) e siano $x(t)$ e $y(t)$ due funzioni (reali di variabile reale) derivabili (in un intervallo). Allora la funzione composta $\varphi(t) = f(x(t), y(t))$ è derivabile e risulta

$$\varphi'(t) = \frac{\partial f}{\partial x}(x(t), y(t)) x'(t) + \frac{\partial f}{\partial y}(x(t), y(t)) y'(t).$$

Più in generale, con notazioni vettoriali, se $f(p)$ è una funzione reale di k variabili reali di classe C^1 e $p(t) = (x_1(t), x_2(t), \dots, x_k(t))$ è una curva derivabile a valori in \mathbb{R}^k , allora la funzione (reale di variabile reale) $\varphi(t) = f(p(t))$ è derivabile e risulta

$$\varphi'(t) = \frac{\partial f}{\partial x_1}(p(t)) x'_1(t) + \frac{\partial f}{\partial x_2}(p(t)) x'_2(t) + \dots + \frac{\partial f}{\partial x_k}(p(t)) x'_k(t).$$

Si osservi che, mediante il prodotto scalare, la suddetta espressione della derivata può essere scritta nella forma

$$\varphi'(t) = \nabla f(p(t)) \cdot p'(t),$$

dove (ricordiamo) $\nabla f(p(t))$ denota il gradiente di f nel punto $p(t)$.

Un altro modo per esprimere la derivata di una funzione composta è il seguente (si può omettere di scrivere la variabile t , dato che l'uguaglianza vale per ogni t):

$$\frac{d\varphi}{dt} = \sum_{i=1}^k \frac{\partial f}{\partial x_i} \frac{dx_i}{dt}.$$

Talvolta (per abuso di notazione) si scrive anche

$$\frac{df}{dt} = \sum_{i=1}^k \frac{\partial f}{\partial x_i} \frac{dx_i}{dt},$$

oppure, quando $y = f(x_1, x_2, \dots, x_k)$ denota l'equazione del grafico di f , si usa anche una delle seguenti espressioni:

$$\frac{dy}{dt} = \sum_{i=1}^k \frac{\partial f}{\partial x_i} \frac{dx_i}{dt}, \quad \frac{dy}{dt} = \sum_{i=1}^k \frac{\partial y}{\partial x_i} \frac{dx_i}{dt}.$$

Il precedente teorema permette di calcolare la derivata di funzioni composte più complesse di $f(x(t), y(t))$. Calcoliamo, ad esempio, la derivata parziale rispetto alla prima variabile della funzione composta

$$\varphi(u, v) = f(x(u, v), y(u, v)),$$

dove $f(x, y)$, $x(u, v)$ e $y(u, v)$ sono di classe C^1 . Ricordando che la derivata parziale di $\varphi(u, v)$ rispetto alla variabile u non è altro che la derivata della funzione parziale $u \mapsto \varphi(u, v)$, che è di una sola variabile, dal teorema precedente segue immediatamente

$$\frac{\partial \varphi}{\partial u}(u, v) = \frac{\partial f}{\partial x}(x(u, v), y(u, v)) \frac{\partial x}{\partial u}(u, v) + \frac{\partial f}{\partial y}(x(u, v), y(u, v)) \frac{\partial y}{\partial u}(u, v).$$

Poiché tale uguaglianza vale per ogni (u, v) , più sinteticamente possiamo scrivere

$$\frac{\partial \varphi}{\partial u} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u},$$

oppure

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x} \frac{\partial x}{\partial u} + \frac{\partial f}{\partial y} \frac{\partial y}{\partial u}.$$

101 - Mercoledì 21/04/10

Sia $f: A \rightarrow \mathbb{R}^s$ un'applicazione di classe C^1 definita su un aperto A di \mathbb{R}^k . Denotiamo con f_1, f_2, \dots, f_s le s funzioni reali che compongono la f (ricordiamo che sono funzioni reali

di k variabili reali). Fissato un punto $p \in A$, la matrice

$$f'(p) = \left(\frac{\partial f_i}{\partial x_j} \right)_p = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \dots & \frac{\partial f_1}{\partial x_k} \\ \frac{\partial f_2}{\partial x_1} & \frac{\partial f_2}{\partial x_2} & \dots & \frac{\partial f_2}{\partial x_k} \\ \vdots & & & \vdots \\ \frac{\partial f_s}{\partial x_1} & \frac{\partial f_s}{\partial x_2} & \dots & \frac{\partial f_s}{\partial x_k} \end{pmatrix}_p$$

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

Un'altra notazione per la matrice jacobiana di f in p è la seguente:

$$\frac{\partial(f_1, f_2, \dots, f_s)}{\partial(x_1, x_2, \dots, x_k)}(p).$$

Quest'ultima notazione è particolarmente utile per rappresentare una sottomatrice della matrice jacobiana. Ad esempio,

$$\frac{\partial(f_1, f_2)}{\partial(x_1, x_2)}$$

è la derivata di f_1 e f_2 rispetto alle variabili x_1 e x_2 (cioè la sottomatrice di f' costituita dalle prime due righe e le prime due colonne), mentre

$$\frac{\partial(f_1, f_2, \dots, f_s)}{\partial x_j}$$

rappresenta la j -esima colonna della matrice f' (è il vettore tangente alla j -esima funzione parziale vettoriale, cioè il vettore tangente alla curva che si ottiene facendo variare soltanto la j -esima coordinata della funzione vettoriale f). Un altro modo per rappresentare la colonna j -esima della matrice jacobiana di f è il seguente:

$$\frac{\partial f}{\partial x_j},$$

dove f sta per (f_1, f_2, \dots, f_s) .

Si osservi che se f è una funzione reale di una sola variabile reale (cioè $k = s = 1$), la matrice jacobiana di f in un punto p è composta da un solo elemento, ossia è un numero e coincide con la classica derivata di f in p . Pertanto, la matrice jacobiana rappresenta una generalizzazione della nozione di derivata, ed è per questo motivo che, anche nel caso vettoriale, possiamo continuare a chiamarla *derivata*.

Quando $k = s$, la matrice $f'(p)$ è quadrata, e in questo caso ha senso il suo determinante, $\det f'(p)$, chiamato *jacobiano* di f in p .

Teorema della derivata di una funzione composta^{sd} (caso generale). *La derivata (cioè la matrice jacobiana) in un generico punto p della composizione $\varphi(p) = g(f(p))$ di*

due applicazioni (vettoriali di variabile vettoriale) f e g (di classe C^1) coincide col prodotto righe per colonne delle derivate $g'(f(p))$ e $f'(p)$. Ovvero

$$\varphi'(p) = g'(f(p))f'(p).$$

102 - Mercoledì 21/04/10

Definizione (di derivata direzionale). Sia $f: A \rightarrow \mathbb{R}$ una funzione reale definita in un aperto di \mathbb{R}^k . Fissato un punto $p \in A$ e una direzione $v \in \mathbb{R}^k$, il

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

si chiama *derivata direzionale* di f in p nella direzione v .

L'esercizio che segue (non facoltativo) mostra che il concetto di derivata direzionale estende la nozione di derivata parziale.

Esercizio. Sia f una funzione reale definita in un aperto A di \mathbb{R}^k e sia p un punto di A . Provare che se e_i denota l' i -esimo elemento della base canonica di \mathbb{R}^k , allora la derivata di f in p nella direzione e_i coincide con la i -esima derivata parziale di f in p .

La derivata direzionale di f in p nella direzione v si denota con uno dei seguenti simboli:

$$\frac{\partial f}{\partial v}(p), \quad df(p, v), \quad df_p(v).$$

Di solito alle ultime due notazioni, a differenza della prima, si attribuisce una validità più ampia: denotano la *derivata lungo un vettore*, la cui definizione è uguale a quella di derivata direzionale, senza però richiedere che v sia una direzione (può essere un qualunque vettore, anche nullo).

Definizione (di derivata lungo un vettore). Sia $f: A \rightarrow \mathbb{R}$ una funzione reale definita in un aperto di \mathbb{R}^k . Fissato un punto $p \in A$ e un qualunque vettore $v \in \mathbb{R}^k$, il

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

si chiama *derivata lungo il vettore v* di f in p e si denota con uno dei seguenti simboli:

$$df(p, v), \quad df_p(v).$$

L'osservazione che segue, la cui dimostrazione è lasciata per esercizio (non facoltativo) agli studenti, è di fondamentale utilità per il calcolo della derivata lungo un vettore.

Osservazione. La derivata di f in p lungo un vettore v è la derivata per $t = 0$ della funzione composta (reale di variabile reale) $\varphi(t) = f(p + tv)$. Pertanto se, in particolare, $\|v\| = 1$, la derivata direzionale di f in p nella direzione v si può interpretare come la derivata in p della restrizione della f alla retta passante per i punti p e $p + v$. Il suo valore misura il tasso di crescita (in p) della funzione f nella direzione v .

Tenendo conto della suddetta osservazione, se f è C^1 , in base al teorema della derivata di una funzione composta, si ha

$$df(p, v) = \varphi'(0) = \nabla f(p) \cdot v.$$

In particolare, se v è una generica direzione, si ottiene

$$\frac{\partial f}{\partial v}(p) = \nabla f(p) \cdot v.$$

Da ciò si deduce il significato di gradiente (in un punto p) di una funzione f : è il vettore che indica la direzione di massima crescita (in p) della funzione f , e il suo modulo (o norma) è la derivata in p lungo tale direzione. Si osservi infatti che il massimo (al variare della direzione v) di $\nabla f(p) \cdot v$ si ottiene quando la direzione v è allineata con $\nabla f(p)$, cioè quando l'angolo θ tra $\nabla f(p)$ e v è zero, ovvero quando $\cos \theta = 1$. A tale proposito ricordiamo che

$$|\nabla f(p) \cdot v| = \|\nabla f(p)\| \|v\| \cos \theta,$$

ed essendo $\|v\| = 1$, risulta

$$|\nabla f(p) \cdot v| = \|\nabla f(p)\| \cos \theta.$$

103 - Venerdì 23/04/10

Definizione (di insieme di livello). Sia $f: A \rightarrow \mathbb{R}$ una funzione di due variabili e sia $c \in \mathbb{R}$. Il sottoinsieme

$$f^{-1}(c) = \{(x, y) \in A : f(x, y) = c\}$$

di \mathbb{R}^2 costituito dai punti (x, y) in cui f assume valore c si dice *curva di livello c* (di f) o, meglio ancora, *insieme di livello c* .

Non sempre una curva di livello rispetta ciò che la nostra intuizione ci suggerisce essere una curva. Si pensi ad esempio ad una funzione costante: l'unica "curva" di livello non vuota coincide con tutto il dominio (col cavolo che assomiglia a una curva!). Si potrebbe dimostrare (ma non lo facciamo) che una condizione affinché un insieme di livello sia effettivamente ciò che noi riteniamo essere una curva è che il gradiente sia non nullo in tutti i punti dell'insieme.

Definizione (di curva di livello regolare). Sia $f: A \rightarrow \mathbb{R}$ una funzione di due variabili di classe C^1 in un aperto A di \mathbb{R}^2 . Una curva di livello $f^{-1}(c) = \{(x, y) \in A : f(x, y) = c\}$ si dice *regolare* se in ogni suo punto p il gradiente $\nabla f(p)$ è non nullo (cioè se almeno una delle due derivate parziali della f è diversa da zero).

Ricordiamo che il grafico di una funzione $f: X \rightarrow Y$ tra due arbitrari insiemi è l'insieme delle coppie ordinate (x, y) del prodotto cartesiano $X \times Y$ che soddisfano la relazione $y = f(x)$, detta equazione del grafico. In base a tale definizione, data una funzione $f: \mathbb{R}^2 \rightarrow \mathbb{R}$, una coppia $((x, y), z) \in \mathbb{R}^2 \times \mathbb{R}$ sta nel grafico di f se e solo se verifica

l'equazione $z = f(x, y)$. Dato che il primo elemento (x, y) della coppia $((x, y), z)$ è a sua volta una coppia (di numeri reali) e dato che $\mathbb{R}^2 \times \mathbb{R}$ si può identificare con \mathbb{R}^3 , è conveniente pensare il grafico di f costituito dalle terne (x, y, z) di \mathbb{R}^3 che verificano la condizione $z = f(x, y)$. Se la funzione f è continua, il suo grafico è una superficie in \mathbb{R}^3 .

Data una funzione reale di variabile reale f , consideriamo un punto (x_0, y_0) del suo grafico (ossia, supponiamo che x_0 stia nel dominio di f e che y_0 sia uguale a $f(x_0)$). Ricordiamo che se f è derivabile in x_0 , la *retta tangente* al grafico della f in (x_0, y_0) è la retta passante per (x_0, y_0) con coefficiente angolare $f'(x_0)$. Ossia, è la retta di equazione

$$y - y_0 = f'(x_0)(x - x_0).$$

In un certo senso, la retta tangente al grafico di f in (x_0, y_0) è, tra tutte le rette del piano \mathbb{R}^2 , quella che meglio approssima il grafico di f in un intorno del punto (x_0, y_0) .

Analogamente, se $f(x, y)$ è una funzione reale di due variabili reali (che per semplicità supponiamo sia di classe C^1 su aperto A di \mathbb{R}^2), il piano tangente al grafico di f in un punto (x_0, y_0, z_0) è, tra tutti i piani dello spazio \mathbb{R}^3 , quello che meglio approssima il grafico di f in un intorno di (x_0, y_0, z_0) . Vedremo più avanti (con il concetto di differenziale) il preciso significato dell'espressione "piano che meglio approssima". Per ora definiamo il piano tangente come quel piano di \mathbb{R}^3 che passa per il punto (x_0, y_0, z_0) e contiene le rette tangenti alle due funzioni parziali $x \mapsto f(x, y_0)$ e $y \mapsto f(x_0, y)$. Quindi, l'equazione di detto piano è data da

$$z - z_0 = \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0).$$

104 - Venerdì 23/04/10

Definizione. Sia $f: A \rightarrow \mathbb{R}$ una funzione reale definita in un sottoinsieme A di \mathbb{R}^k e sia p_0 un punto di A . Si dice che $p_0 \in A$ è un *punto di minimo* [di massimo] *assoluto* per f in A se $f(p) \geq f(p_0)$ [$f(p) \leq f(p_0)$] per ogni $p \in A$; cioè se $f(p_0)$ coincide col (valore) minimo [massimo] di f in A . Si dice che $p_0 \in A$ è un *punto di minimo* [di massimo] *relativo* (o *locale*) per f in A se esiste un intorno $I(p_0, r)$ di p_0 tale che $f(p) \geq f(p_0)$ [$f(p) \leq f(p_0)$] per ogni $p \in I(p_0, r) \cap A$ (se, inoltre, per $p \neq p_0$ la disuguaglianza vale in senso stretto, allora il punto di minimo [di massimo] si dice che è *stretto* o *forte*). Un punto di minimo o di massimo relativo per f (in A) si dice *estremante* per f (in A). I valori assunti nei punti estremanti si dicono gli *estremi* (relativi o assoluti) di f in A .

Ovviamente se un punto è di minimo [massimo] per una funzione f , allora è di massimo [minimo] per $-f$.

Esempio. Consideriamo la funzione $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da

$$f(x, y) = \frac{2}{3 + \sqrt{1 + (x + 1)^2 + |y|}}.$$

Determiniamone gli estremi superiore e inferiore, e stabiliamo se questi sono, rispettivamente, massimo e minimo.

Iniziamo con la ricerca degli eventuali punti di massimo assoluto. Poiché f è un rapporto di due quantità positive, e il numeratore è costante, la funzione assumerà il valore massimo in corrispondenza dei punti di minimo assoluto del denominatore (ammesso che esistano). Il denominatore è minimo quando è minimo l'argomento della radice, cioè quando sono minimi entrambi i termini $(x+1)^2$ e $|y|$. Dunque il **punto** $(-1, 0)$ è di massimo assoluto per f , e il **valore** $f(-1, 0) = 1/2$ è il massimo assoluto per f (ricordarsi della distinzione tra punti e valori: i primi stanno nel dominio e i secondi nel codominio).

Passiamo ora alla ricerca degli eventuali punti di minimo assoluto. Questi si trovano in corrispondenza dei punti di massimo assoluto della funzione $g(x, y) = (x+1)^2 + |y|$, ma tali punti non esistono, dato che l'estremo superiore di g è $+\infty$ (verificarlo per esercizio). Quindi $\inf f = 0$ (verificarlo per esercizio).

Tirando le somme possiamo affermare che gli estremi (inferiore e superiore) di f sono 0 e $1/2$. Il primo però non è un valore assunto mentre il secondo sì. Cioè 0 è solo estremo inferiore ma non minimo, mentre $1/2$ è addirittura il massimo di f (e quindi anche estremo superiore). Vedremo in seguito come da ciò si possa dedurre che la funzione (essendo continua e definita in tutto \mathbb{R}^2) ha per immagine tutto l'intervallo $(0, 1/2]$.

Esercizio. Siano $A \subseteq \mathbb{R}^k$, $f: A \rightarrow \mathbb{R}$, $B \subseteq A$ e $p_0 \in B$. Provare che se p_0 è un punto di minimo [di massimo] relativo per f in A , allora lo è anche per f in B . Mostrare con un esempio che non è vero il contrario.

Esercizio. Siano V ed R due costanti positive. Determinare (se esiste) il massimo valore assunto dalla seguente funzione di tre variabili:

$$I(\omega, L, C) = \frac{V}{\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2}}.$$

105 - Mercoledì 28/04/10

Teorema di Fermat in \mathbb{R}^k . Sia $f: A \rightarrow \mathbb{R}$ una funzione reale di k variabili reali e sia $p \in A$. Supponiamo che siano soddisfatte le seguenti tre ipotesi:

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

Allora $\nabla f(p) = 0$.

Dimostrazione. Per provare che la derivata di f in p rispetto alla i -esima variabile è nulla, basta considerare la funzione reale di variabile reale $\varphi(t) = f(p + te_i)$, dove e_i è l' i -esimo vettore della base canonica di \mathbb{R}^k , e applicare il Teorema di Fermat per le funzioni di una sola variabile nel punto $t = 0$ (ricordarsi che $\varphi'(0)$ è la derivata direzionale lungo e_i nel punto p e coincide con la derivata parziale di f in p rispetto alla variabile i -esima). \square

I punti (interni al dominio) in cui si annulla il gradiente di una funzione reale f vengono detti *critici* o *stazionari* per f . I punti in cui f non è derivabile si dicono *singolari*.

Osserviamo che, in base al Teorema di Fermat, se un punto p è estremante per una funzione $f: A \rightarrow \mathbb{R}$, allora verifica almeno una delle seguenti tre condizioni:

- 1) è di frontiera per A ;
- 2) è singolare per f ;
- 3) è critico per f .

La ricerca dei punti estremanti va quindi ristretta ai punti che verificano almeno una delle suddette condizioni, nessuna delle quali ci assicura che un punto sia estremante.

Tre esempi (di punti estremanti che verificano soltanto una delle suddette condizioni):

- 1) $(0, 0)$ è estremante per $x + y$ in $[0, +\infty) \times [0, +\infty)$;
- 2) $(0, 0)$ è estremante per $|x| + |y|$ in \mathbb{R}^2 ;
- 3) $(0, 0)$ è estremante per $x^2 + y^2$ in \mathbb{R}^2 .

Nella ricerca dei punti estremanti di una funzione definita in un insieme $A \subseteq \mathbb{R}^k$ si procede per esclusione: come primo provvedimento, tenendo conto del Teorema di Fermat, si scagionano subito “dall'accusa di essere estremanti” tutti i punti interni ad A in cui la funzione risulta derivabile ed una delle 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 tutte le derivate. Per “discolpare” 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 di successive esclusioni, rimangono spesso soltanto pochi punti “indiziati” che possono essere analizzati a parte con metodi vari. Ad esempio, se si cercano gli estremi assoluti di f in A , mediante il confronto dei valori assunti da f in tali punti.

Per comprendere meglio il procedimento per esclusione nella ricerca dei punti estremanti, riformuliamo il Teorema di Fermat in una versione equivalente: la versione “garantista”.

Teorema di Fermat in \mathbb{R}^k (riformulato). *Sia $f: A \rightarrow \mathbb{R}$ una funzione reale di k variabili reali e sia $p \in A$. Supponiamo che siano soddisfatte le seguenti tre ipotesi:*

- 1) p è interno ad A ;
- 2) f è derivabile in p ;
- 3) una delle derivate parziali di f in p è diversa da zero.

Allora p non è estremante per f in A .

Teorema di Weierstrass^{sd}. *Sia $f: A \rightarrow \mathbb{R}$ una funzione continua in un sottoinsieme compatto (cioè limitato e chiuso) $A \subseteq \mathbb{R}^k$. Allora f ammette minimo e massimo assoluti. Esistono quindi almeno due punti, p_1 e p_2 in A , per i quali si ha*

$$f(p_1) \leq f(p) \leq f(p_2), \quad \forall p \in A.$$

Attenzione! Data $f: A \rightarrow \mathbb{R}$, la condizione “ $f(p_1) \leq f(p) \leq f(p_2)$, $\forall p \in A$ ” non implica

che f assume tutti i valori compresi tra $f(p_1)$ e $f(p_2)$. È un errore molto frequente nei compiti scritti.

106 - Mercoledì 28/04/10

Un *arco di curva parametrica* in \mathbb{R}^2 (o in \mathbb{R}^k) è una funzione continua $\gamma(t)$ definita in un intervallo compatto $[a, b]$ a valori in \mathbb{R}^2 (o in \mathbb{R}^k). I punti $\gamma(a)$ e $\gamma(b)$ si dicono, rispettivamente, *primo* e *secondo estremo* della curva. La curva $\gamma(t)$ si dice *chiusa* se i suoi estremi coincidono (ossia, se $\gamma(a) = \gamma(b)$). La variabile t di $\gamma(t)$ si chiama il *parametro* della curva. Ovviamente, il parametro può essere denotato con una qualunque lettera (le più comuni sono $t, \tau, s, \theta, \varphi$). Talvolta l'immagine di un arco di curva parametrica γ viene anche detta *sostegno* di γ .

Ricordiamo, dal corso di Geometria, che se p_0 e p_1 sono due punti di \mathbb{R}^k , l'insieme

$$\overline{p_0 p_1} = \{tp_1 + (1-t)p_0 : t \in [0, 1]\} = \{p_0 + t(p_1 - p_0) : t \in [0, 1]\}$$

si dice *segmento* che congiunge p_0 con p_1 (se si fa variare t in \mathbb{R} si ottiene addirittura la retta che passa per p_0 e p_1). Si osservi che la funzione $\gamma: [0, 1] \rightarrow \mathbb{R}^k$ definita da $\gamma(t) = p_0 + t(p_1 - p_0)$ è un arco di curva parametrica il cui sostegno coincide col segmento $\overline{p_0 p_1}$ (si dice che $\gamma(t)$ è una parametrizzazione di $\overline{p_0 p_1}$).

Definizione. Un sottoinsieme A di \mathbb{R}^k si dice *connesso (per archi)* se dati due qualunque punti $p_0, p_1 \in A$ esiste un arco di curva (parametrica) a valori in A congiungente p_0 con p_1 (cioè, con estremi p_0 e p_1).

Osserviamo che se un insieme $A \subseteq \mathbb{R}^k$ è *convesso*, cioè ha la proprietà che se contiene due punti contiene anche il segmento che li congiunge, allora è anche connesso. Ovviamente esistono connessi che non sono convessi (si pensi, ad esempio, ad una banana in \mathbb{R}^3 , ad una corona circolare in \mathbb{R}^2 , oppure all'insieme che si ottiene rimuovendo una retta da \mathbb{R}^3).

Si potrebbe dimostrare (come conseguenza del teorema di esistenza degli zeri) che in \mathbb{R} i connessi e i convessi sono gli stessi insiemi, cioè gli intervalli.

Teorema dei valori intermedi. Sia $f: A \rightarrow \mathbb{R}$ una funzione continua in un insieme connesso $A \subseteq \mathbb{R}^k$. Allora l'immagine di f è un intervallo.

Dimostrazione^{fac}. Occorre provare che se y_1 e y_2 sono due punti dell'immagine $\text{Im } f$ di f e $c \in \mathbb{R}$ verifica la condizione $y_1 < c < y_2$, allora anche c appartiene a $\text{Im } f$. Poiché $y_1, y_2 \in \text{Im } f$, esistono (almeno) due punti $p_1, p_2 \in A$ tali che $f(p_1) = y_1$ e $f(p_2) = y_2$. Essendo A un insieme connesso, esiste un arco di curva parametrica $\gamma: [a, b] \rightarrow A$ tale che $\gamma(a) = p_1$ e $\gamma(b) = p_2$. Consideriamo la funzione $g: [a, b] \rightarrow \mathbb{R}$ definita da $g(t) = f(\gamma(t))$. Osserviamo che g è continua (perché composizione di funzioni continue) ed assume i valori y_1 (per $t = a$) e y_2 (per $t = b$). Pertanto, in base al teorema dei valori intermedi per le funzioni di una variabile, g assume anche il valore c in (almeno) un punto $t_0 \in [a, b]$ (in realtà t_0 sta addirittura in (a, b) , perché?). Ne consegue che anche f assume il valore c (nel punto $p_0 = \gamma(t_0) \in A$). □

Esempio (con metafora giudiziaria). Determiniamo l'insieme dei valori assunti dalla funzione $f(x, y) = x^2 + y^2 + x$ nel triangolo chiuso T di vertici $(0, 0)$, $(1, 0)$ e $(0, 1)$. Prima di iniziare ad eseguire calcoli è utile vedere dove si va a parare: conviene controllare se è possibile applicare il Teorema di Weierstrass o il Teorema dei valori intermedi (o entrambi). La funzione è continua in \mathbb{R}^2 e quindi anche nel triangolo chiuso T (che è anche un insieme limitato), pertanto, per il Teorema di Weierstrass, ammette minimo assoluto e massimo assoluto in T . Il triangolo T è anche connesso (è addirittura convesso) e quindi, per il Teorema dei valori intermedi, la sua immagine (tramite f) è un intervallo. Possiamo concludere che l'insieme dei valori assunti da f in T (cioè l'immagine di T tramite f) è un intervallo compatto $[m, M]$ i cui estremi coincidono con i valori minimo e massimo (assoluti) di f in T . Determiniamo tali valori (cioè gli estremi assoluti della f in T). A tale scopo è necessario selezionare i possibili punti estremanti di f e individuare tra questi i colpevoli di uno dei seguenti reati: essere un punto di minimo assoluto o essere un punto di massimo assoluto (l'esistenza dei reati è accertata dal Teorema di Weierstrass). Procediamo per esclusione. Verifichiamo prima l'alibi dei punti interni a T (ossia dei punti di T che non sono di frontiera). Dato che il gradiente di f è ovunque definito (non ci sono punti singolari) e non si annulla mai in T (l'unico punto critico di f non sta in T), il Teorema di Fermat ci permette di scagionare dall'infamante accusa tutti i punti interni a T . Tuttavia, dato che nel triangolo chiuso ci devono essere almeno due colpevoli, la ricerca deve proseguire tra i punti di frontiera. Analizziamo prima il lato L_1 che congiunge $(0, 0)$ con $(1, 0)$. Se in tale lato c'è un punto estremante (per f in T), questo deve essere estremante anche per la restrizione di f a L_1 (anche se non è vero il contrario). La restrizione di f a L_1 è rappresentata dalla funzione (parziale) $\varphi_1(x) = f(x, 0) = x^2 + x$, con $x \in [0, 1]$. Poiché φ_1 è strettamente crescente in $[0, 1]$, non ci sono punti estremanti interni al lato L_1 . Quindi, gli unici indiziati, per ora, sono gli estremi del lato L_1 (cioè i vertici $(0, 0)$ e $(1, 0)$ del triangolo). Proseguiamo l'analisi lungo il lato L_2 di vertici $(0, 1)$ e $(1, 0)$. Detto lato è contenuto nella retta di equazione $x + y = 1$. Quindi, la restrizione di f ad L_2 si può rappresentare con la funzione $\varphi_2(x) = f(x, 1 - x) = x^2 + (1 - x)^2 + x$, con $x \in [0, 1]$. Tale funzione decresce per $x \in [0, 1/4]$ e cresce per $x \in [1/4, 1]$, come si vede dal segno della sua derivata. Pertanto, ai due indiziati di prima bisogna aggiungere il vertice $(0, 1)$ e il punto $(1/4, 3/4)$. L'altro vertice lo avevamo già considerato e rimane tra gli indiziati (viene rinviato a giudizio). Occorre infine analizzare il comportamento di f nel lato L_3 di vertici $(0, 0)$ e $(0, 1)$. Studiamo quindi la funzione (parziale) $\varphi_3(y) = y^2$, per $y \in [0, 1]$. Si tratta di una funzione strettamente crescente (per $y \in [0, 1]$) e, di conseguenza, agli indiziati di prima non se ne aggiungono altri, dato che i vertici $(0, 0)$ e $(0, 1)$ erano già stati considerati (si può solo affermare che quest'ultima analisi non li scagiona dall'accusa di essere punti estremanti assoluti). Per concludere possiamo affermare che tra i quattro indiziati (cioè i tre vertici del triangolo più il punto $(1/4, 3/4)$) ci sono senz'altro (almeno) due colpevoli (ce lo dice il Teorema di Weierstrass): almeno uno deve essere un punto di minimo assoluto e un altro deve essere di massimo assoluto. Calcolando i valori assunti da f nei quattro punti rinviati a giudizio si può concludere che $(0, 0)$ è un punto di minimo assoluto (ed è unico) mentre $(1, 0)$ è un punto di massimo assoluto (e anch'esso è unico). L'insieme dei valori assunti da f in T è dunque l'intervallo compatto $[0, 2]$, come risulta dai valori calcolati nei due punti giudicati colpevoli.

107 - Mercoledì 28/04/10

Definizione (di limite per $\|p\| \rightarrow +\infty$). Sia $f: \mathbb{R}^k \rightarrow \mathbb{R}$ una funzione reale di k variabili reali. Dato un numero reale esteso γ (i.e. $\gamma = \ell \in \mathbb{R}$, $\gamma = -\infty$ oppure $\gamma = +\infty$), si dice che γ è il limite di $f(p)$ per $\|p\| \rightarrow +\infty$, e si scrive

$$\lim_{\|p\| \rightarrow +\infty} f(p) = \gamma,$$

se per ogni successione $\{p_n\}$ in \mathbb{R}^k tale che $\|p_n\| \rightarrow +\infty$, risulta

$$\lim_{n \rightarrow +\infty} f(p_n) = \gamma.$$

C'è un modo diverso per definire il concetto di limite per $\|p\| \rightarrow +\infty$, che evita di ricorrere alla nozione di successione. Occorre però distinguere tre casi:

- 1) $f(p) \rightarrow \ell \in \mathbb{R}$;
- 2) $f(p) \rightarrow +\infty$;
- 3) $f(p) \rightarrow -\infty$.

Definizione (di limite per $\|p\| \rightarrow +\infty$ senza le successioni). Sia $f: \mathbb{R}^k \rightarrow \mathbb{R}$ una funzione reale di k variabili reali.

- 1) Si dice che $f(p) \rightarrow \ell \in \mathbb{R}$ per $\|p\| \rightarrow +\infty$ se per ogni $\varepsilon > 0$ esiste un $h > 0$ tale che se $\|p\| > h$, allora $|f(p) - \ell| < \varepsilon$.
- 2) Si dice che $f(p) \rightarrow +\infty$ per $\|p\| \rightarrow +\infty$ se per ogni $c > 0$ esiste un $h > 0$ tale che se $\|p\| > h$, allora $f(p) > c$.
- 3) Si dice che $f(p) \rightarrow -\infty$ per $\|p\| \rightarrow +\infty$ se per ogni $c > 0$ esiste un $h > 0$ tale che se $\|p\| > h$, allora $f(p) < -c$.

Teorema^{sd} (di collegamento per il limite di $f(p)$ per $\|p\| \rightarrow +\infty$). *Le due precedenti definizioni di limite per $\|p\| \rightarrow +\infty$ sono equivalenti.*

Esercizio. Provare che la funzione

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

non ammette limite per $\sqrt{x^2 + y^2} \rightarrow +\infty$.

Esercizio. Provare che la funzione

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

è limitata (determinando una costante $c > 0$ tale che $|f(x, y)| \leq c$ per ogni $(x, y) \in \mathbb{R}^2$).

Esercizio. Dedurre dall'esercizio precedente che la funzione

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

tende a zero per $\sqrt{x^2 + y^2} \rightarrow +\infty$.

Suggerimento. Si osservi che è il prodotto di una funzione limitata per una infinitesima.

Esercizio. Provare che $x^2 + y^2 + x \rightarrow +\infty$ per $\sqrt{x^2 + y^2} \rightarrow +\infty$.

Suggerimento. Si osservi che per $(x, y) \neq (0, 0)$ risulta

$$x^2 + y^2 + x = (x^2 + y^2) \left(1 + \frac{x}{x^2 + y^2} \right).$$

Esercizio ^{fac.}. Sia $f: \mathbb{R}^k \rightarrow \mathbb{R}$ continua e tale che $f(p) \rightarrow +\infty$ per $\|p\| \rightarrow +\infty$. Provare che f ammette minimo assoluto in \mathbb{R}^k .

Suggerimento. Applicare il Teorema di Weierstrass ad una opportuna restrizione di f ad un sottoinsieme limitato e chiuso di \mathbb{R}^k .

Esercizio. Determinare l'insieme dei valori assunti dalla funzione $f(x, y) = x^2 + y^2 + x$.

Esercizio. Determinare l'insieme dei valori assunti dalla funzione $f(x, y) = x^2 + y^2 + x$ nel cerchio chiuso $\{(x, y) : x^2 + y^2 \leq 1\}$.

Esercizio. Determinare l'insieme dei valori assunti dalla funzione $f(x, y) = x^2 + y^2 + x$ nel cerchio aperto $\{(x, y) : x^2 + y^2 < 1\}$.

Esercizio. Determinare, in funzione di $r > 0$, l'insieme dei valori assunti dalla funzione $f(x, y) = x^2 + y^2 + x$ nel cerchio chiuso avente centro nell'origine e raggio r .

108 - Venerdì 30/04/10

Ricordiamo che un polinomio di due variabili (o, più in generale, di k variabili) si dice *omogeneo* se tutti i suoi monomi sono dello stesso grado. Ad esempio $x^3 - xy^2$ è un polinomio omogeneo (di due variabili) di terzo grado, mentre $y - x^2$ è un polinomio non omogeneo di secondo grado.

Un polinomio omogeneo di primo grado di k variabili si dice una *forma lineare* in \mathbb{R}^k (si chiama così perché è un'applicazione lineare da \mathbb{R}^k in \mathbb{R}).

Definizione (di forma quadratica). Un polinomio omogeneo di secondo grado di k variabili si dice una *forma quadratica* in \mathbb{R}^k .

Osservazione. Ogni forma quadratica è nulla nell'origine.

Una forma quadratica si dice *definita positiva* (*definita negativa*) se è maggiore di zero (minore di zero) tranne il caso in cui tutte le variabili sono nulle. Una forma quadratica si dice *indefinita* se assume valori sia positivi sia negativi. Infine, una forma quadratica si dice *semidefinita* positiva (negativa) se è sempre maggiore (minore) o uguale a zero. Ovviamente, una forma quadratica definita positiva (o definita negativa) è anche semidefinita.

Esercizio. Sia g una forma quadratica in \mathbb{R}^k . Mostrare che se g è definita positiva (negativa), allora l'origine $0 \in \mathbb{R}^k$ è un punto di minimo (massimo) per g . Provare inoltre che se g è indefinita allora l'origine non è né di massimo né di minimo.

Esercizio. Sia g una forma quadratica in \mathbb{R}^k . Mostrare che g è una funzione omogenea di grado 2. Ossia, per ogni $v \in \mathbb{R}^k$ e per ogni $\lambda \in \mathbb{R}$ risulta $g(\lambda v) = \lambda^2 g(v)$.

Esempi di forme quadratiche in \mathbb{R}^2 :

$x^2 + 2y^2$ è definita positiva (e quindi anche semidefinita positiva);

$x^2 - y^2$ è indefinita;

$x^2 - 2xy + y^2 = (x - y)^2$ è semidefinita positiva (ma non definita positiva).

Esempi di forme quadratiche in \mathbb{R}^3 :

$x^2 + 2y^2$ è semidefinita positiva (si osservi che si annulla nei vettori del tipo $(0, 0, z)$);

$x^2 - y^2$ è indefinita;

$x^2 + 2y^2 + z^2$ è definita positiva.

Vediamo ora come sia possibile classificare una generica forma quadratica in \mathbb{R}^2 , cioè una funzione reale di due variabili reali del tipo

$$f(x, y) = ax^2 + 2bxy + cy^2.$$

Nel caso in cui il coefficiente a sia zero, lo studio è banale ed è lasciato per esercizio (non facoltativo) allo studente. Supponiamo quindi $a \neq 0$. Osserviamo che se $y = 0$, la forma si annulla solo se $x = 0$. Non è pertanto restrittivo supporre $y \neq 0$. In tal caso si ha

$$ax^2 + 2bxy + cy^2 = y^2 \left(a \left(\frac{x}{y} \right)^2 + 2b \left(\frac{x}{y} \right) + c \right).$$

Il segno di $ax^2 + 2bxy + cy^2$ dipende quindi dal segno del trinomio $at^2 + 2bt + c$, ove si è posto $x/y = t$. Si può quindi concludere che se il discriminante

$$\Delta = 4(b^2 - ac)$$

è negativo, allora il trinomio non si annulla mai e, di conseguenza, la forma quadratica è definita (positiva o negativa, a seconda che il coefficiente a sia maggiore o minore di zero). Se Δ è positivo, allora il trinomio cambia segno, e la forma quadratica risulta indefinita. Se, infine, $\Delta = 0$, allora la forma quadratica è semidefinita (ma non è definita, perché si annulla nei punti della retta $x/y = t_0$, dove t_0 è la radice doppia del trinomio $at^2 + 2bt + c$).

Si osservi che la forma quadratica

$$ax^2 + 2bxy + cy^2$$

può essere rappresentata nel seguente modo: $Av \cdot v$, dove A è la matrice simmetrica

$$\begin{pmatrix} a & b \\ b & c \end{pmatrix},$$

detta *matrice associata alla forma quadratica*, v è il vettore colonna di componenti x e y , Av è il prodotto righe per colonne di A per v , $Av \cdot v$ è il prodotto scalare di Av per v .

Osservazione. Sia $ax^2 + 2bxy + cy^2$ una forma quadratica in \mathbb{R}^2 e sia

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}$$

la matrice associata. Denotiamo con Δ il discriminante del polinomio $at^2 + 2bt + c$. Si osservi che $\Delta = -4\det A$; pertanto, per quanto appena visto, se $\det A > 0$ allora la forma quadratica è definita (positiva o negativa, in accordo col segno di a), se $\det A < 0$ è indefinita, e se $\det A = 0$ è “irrimediabilmente” semidefinita (ovvero, è semidefinita ma non definita).

109 - Venerdì 30/04/10

Prima di proseguire il nostro studio, conviene osservare quanto segue. Come abbiamo già visto, per funzioni di più variabili reali a valori reali continua ad avere senso il problema della ricerca dei massimi e minimi (non ha invece più senso se consideriamo funzioni a valori in \mathbb{R}^s , con $s > 1$). In Analisi 1 lo studio dei massimi e minimi riceve un grosso aiuto dallo studio delle derivate prima e seconda della funzione in esame. In più variabili abbiamo visto il Teorema di Fermat che coinvolge le derivate parziali. C'è un analogo della derivata seconda per funzioni di più variabili? Se la risposta è sì, come si definisce e come si usa tale concetto? Ebbene, vedremo tra breve che i ragionamenti appena fatti sulle forme quadratiche si collegano direttamente a quella che sarà la derivata seconda della funzione. Prima di questo, introduciamo il concetto di differenziale.

Sia $f: U \rightarrow \mathbb{R}$ una funzione C^1 su un aperto U di \mathbb{R}^2 . Consideriamo due punti $p = (x, y)$ e $p + v = (x + h, y + k)$ del dominio U di f . L'incremento che subisce f passando dal punto p al punto $p + v$ è il numero $\Delta f = f(p + v) - f(p)$. Tale incremento (che dovremmo a rigore denotare con $\Delta f(p, v)$ o con $\Delta f_p(v)$, perché dipende sia dal punto iniziale p sia dall'incremento v della variabile vettoriale) ha il difetto di non essere facile da valutare. Tuttavia, se Δf ci interessa soltanto per valori piccoli di $\|v\|$ (come, ad esempio, per il calcolo degli errori), al posto dell'incremento vero è preferibile utilizzare un incremento “virtuale” che, pur non essendo vero (a meno che il grafico di f non sia un piano), ha il duplice pregio di essere facile da calcolare (perché risulta una funzione lineare del vettore v) e di approssimare bene l'incremento vero (se f è, come abbiamo supposto, di classe C^1). Tale incremento virtuale, detto *differenziale di f in p (relativo all'incremento v)*, è così definito:

$$df_p(v) = \lim_{t \rightarrow 0} \frac{f(p + tv) - f(p)}{t} = \varphi'(0),$$

dove $\varphi(t)$ è la funzione reale di variabile reale $t \mapsto f(p + tv)$ ottenuta componendo la curva parametrica (rettilinea) $t \mapsto p + tv$ con la funzione f . Di conseguenza, in base alla regola di derivazione di una funzione composta, si ha

$$df_p(v) = \nabla f(p) \cdot v = \frac{\partial f}{\partial x}(p)h + \frac{\partial f}{\partial y}(p)k.$$

Si potrebbe dimostrare (ma non lo facciamo) che, avendo supposto f di classe C^1 , l'incremento virtuale $df_p(v)$ approssima bene l'incremento vero $\Delta f_p(v)$ per piccoli valori di $\|v\|$, nel senso che il rapporto

$$\frac{\Delta f_p(v) - df_p(v)}{\|v\|}$$

tende a zero per $\|v\|$ che tende a zero. Pertanto, per valori piccoli di h e k , lo scarto tra l'incremento vero e quello virtuale è piccolo rispetto alla norma $\sqrt{h^2 + k^2}$ dell'incremento $v = (h, k)$ della variabile (vettoriale). In un certo senso, fissato il punto p , il differenziale di f in p è, tra tutte le applicazioni lineari da \mathbb{R}^2 in \mathbb{R} , quella che meglio approssima la funzione incremento $v \mapsto \Delta f_p(v)$. Quindi, se la funzione f è lineare (o, più in generale, affine), l'incremento vero e quello virtuale coincidono.

Come per le funzioni di una variabile, il simbolo $\varepsilon(v)$ denota una qualunque funzione di k variabili, continua e nulla nell'origine $0 \in \mathbb{R}^k$. Ovviamente, la variabile (vettoriale) di detta funzione potrà essere indicata con una qualunque lettera (o k -pla di lettere, quando si preferisce far comparire le componenti della variabile vettoriale). Ad esempio, se $\varepsilon(v)$ è di due variabili, si potrà scrivere $\varepsilon(p)$ se si desidera denotare con p il generico punto di \mathbb{R}^2 , oppure si potrà scrivere $\varepsilon(x, y)$ se si preferisce far comparire le coordinate di p , oppure andrà bene anche scrivere $\varepsilon(v_1, v_2)$ al posto di $\varepsilon(v)$, mettendo così in evidenza le componenti del vettore v . Ovviamente se si moltiplica una funzione $\varepsilon(v)$ per una funzione continua si ottiene ancora una funzione $\varepsilon(v)$, così come la somma (e anche la differenza) di due funzioni $\varepsilon(v)$ è una funzione $\varepsilon(v)$.

Spesso, nei calcoli, si ottengono funzioni $\varepsilon(x_1, x_2, \dots, x_k)$ componendo funzioni di k variabili con una $\varepsilon(x)$ di una sola variabile: quando si sostituisce al posto della x una funzione di k variabili del tipo $\varepsilon(x_1, x_2, \dots, x_k)$.

Esercizio^{fac.} Provare che se una funzione $f(v)$ di k variabili è nulla nell'origine di \mathbb{R}^k ed è un o-piccolo di $\|v\|^n$ (cioè $f(v)/\|v\|^n \rightarrow 0$ per $v \rightarrow 0$), allora $f(v) = \varepsilon(v)\|v\|^n$.

Esercizio^{fac.} Sia $f: U \rightarrow \mathbb{R}$ una funzione C^1 in un aperto U di \mathbb{R}^2 (o, più in generale, di \mathbb{R}^k) e sia $p \in U$. Dal fatto che

$$\lim_{\|v\| \rightarrow 0} \frac{\Delta f_p(v) - df_p(v)}{\|v\|} = 0,$$

dedurre la seguente uguaglianza (detta *formula di Taylor del prim'ordine di f in p*):

$$\Delta f_p(v) = df_p(v) + \|v\|\varepsilon(v).$$

Osserviamo ora che gli incrementi veri h e k che subiscono le funzioni coordinate x e y passando dal punto $p = (x_0, y_0)$ al punto $p + v = (x_0 + h, y_0 + k)$ coincidono con i loro rispettivi incrementi virtuali $dx_p(v)$ e $dy_p(v)$. In altre parole, per le funzioni x e y si ha, rispettivamente, $\Delta x_p(v) = h = dx_p(v)$ e $\Delta y_p(v) = k = dy_p(v)$. Pertanto risulta

$$df_p(v) = \frac{\partial f}{\partial x}(p)dx_p(v) + \frac{\partial f}{\partial y}(p)dy_p(v),$$

dove dx_p e dy_p sono i differenziali (nel punto p) di due particolari funzioni da \mathbb{R}^2 ad \mathbb{R} (che risultano lineari): le cosiddette funzioni coordinate (cartesiane) x ed y . Ovvero, le due “leggi” che ad ogni punto di \mathbb{R}^2 assegnano, rispettivamente, la prima e la seconda coordinata di detto punto.

La suddetta uguaglianza, visto che vale per ogni p e per ogni v , si potrà scrivere anche

$$df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

o, più semplicemente, con notazioni vettoriali, nella forma

$$df = \nabla f \cdot dp.$$

Tenendo conto che $z = f(x, y)$ è l'equazione del grafico della funzione f , talvolta si usa scrivere

$$dz = \frac{\partial z}{\partial x} dx + \frac{\partial z}{\partial y} dy$$

oppure

$$dz = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$$

oppure

$$dz = \nabla f \cdot dp.$$

Esercizio. Estendere la nozione di differenziale al caso delle funzioni di k variabili.

Esercizio^{fac.} Siano $f, g : U \rightarrow \mathbb{R}$ due funzioni C^1 in un aperto U di \mathbb{R}^2 e sia c una costante. Provare le seguenti proprietà del differenziale:

- 1) $d(cf) = c df$;
- 2) $d(fg) = g df + f dg$;
- 3) $d(f + g) = df + dg$;
- 4) $d(f/g) = (g df - f dg)/g^2$.

110 - Mercoledì 05/05/10

Definizione (di vettore applicato). Dato un sottoinsieme A di \mathbb{R}^k , una coppia (p, v) , dove $p \in A$ e $v \in \mathbb{R}^k$ si dice un *vettore applicato* in (un punto di) A . Il punto p si chiama *punto di applicazione* e v è il *vettore libero*.

Vedremo in seguito che una qualunque funzione (a valori reali) che ha per dominio un insieme di vettori applicati si chiama *espressione differenziale* (reale). Il differenziale df di una funzione f di classe C^1 in un aperto U di \mathbb{R}^k , essendo una legge che ad ogni vettore applicato (p, v) in U associa un numero (i.e. $df(p, v) = \nabla f(p) \cdot v$), è un esempio di espressione differenziale. Ricordiamo che tale legge, quando il punto di applicazione p è fissato, è lineare rispetto al vettore libero v (in questo caso si denota $df_p : \mathbb{R}^k \rightarrow \mathbb{R}$ e si chiama *differenziale di f in p*).

Osservazione. Sia $f : U \rightarrow \mathbb{R}$ una funzione reale di classe C^1 in un aperto $U \subseteq \mathbb{R}^k$ e sia (p, v) un vettore applicato in un punto $p \in U$. Col linguaggio delle matrici, il differenziale della funzione f calcolato in (p, v) si può rappresentare nel modo seguente: $df(p, v) = f'(p)v$, dove (in questo caso) la matrice jacobiana $f'(p)$ è un vettore riga e v è un vettore colonna.

In sostanza, il Teorema di Fermat afferma che se un punto p è estremante (cioè di massimo o di minimo relativo) per una funzione reale f (di classe C^1 in un aperto U contenente p), allora il differenziale di f in p è necessariamente nullo (ossia p è un punto critico). Abbiamo visto con degli esempi che non è vero il contrario: il differenziale in un punto p può essere nullo, senza che p sia estremante. Per le funzioni di una variabile sappiamo che se un punto x_0 è critico per una funzione $g(x)$ e $g''(x_0) > 0$ [$g''(x_0) < 0$], allora il punto è di minimo [di massimo] stretto. Vedremo in seguito che un risultato analogo vale anche per le funzioni di più variabili (di classe C^2): se un punto p è critico per una funzione f , cioè per ogni direzione v la derivata prima $\varphi'(0)$ della funzione $\varphi(t) = f(p + tv)$ è nulla, allora una condizione sufficiente affinché p sia di minimo [di massimo] stretto è che si abbia $\varphi''(0) > 0$ [$\varphi''(0) < 0$] **per ogni** direzione v . È utile, in questa fase del discorso, introdurre la nozione di differenziale secondo (che costituisce un ulteriore esempio di espressione differenziale).

Definizione (di differenziale secondo). Sia f una funzione di classe C^2 in un aperto U di \mathbb{R}^k e sia $p \in U$. Il *differenziale secondo* di f in p calcolato lungo un vettore (libero) v è la derivata seconda $\varphi''(0)$ della funzione composta $\varphi(t) = f(p + tv)$. Detto differenziale si denota $d^2f(p, v)$, oppure $d^2f_p(v)$ quando si pensa fissato il punto di applicazione p .

Si potrebbe dimostrare (ma non lo facciamo) che, avendo supposto f di classe C^2 , si ha

$$\Delta f_p(v) = df_p(v) + \frac{1}{2} d^2f_p(v) + \|v\|^2 \varepsilon(v),$$

per ogni v ammissibile (cioè tale che $p + v \in U$); dove, ricordiamo, $\varepsilon(v)$ è una funzione di k variabili, continua e nulla nell'origine $0 \in \mathbb{R}^k$. La suddetta formula è un'uguaglianza (non un'approssimazione) dell'incremento $\Delta f_p(v)$ e si dice *formula di Taylor di f in p del second'ordine* (col resto di Peano). Se il punto di applicazione p è 0 (cioè l'origine di \mathbb{R}^k) la formula di Taylor si chiama anche di MacLaurin.

Sia $f: U \rightarrow \mathbb{R}$ una funzione C^2 in un aperto U di \mathbb{R}^2 . Consideriamo un punto $p = (x, y) \in U$ e un vettore $v = (h, k) \in \mathbb{R}^2$. Abbiamo già mostrato che, come conseguenza del teorema di derivazione di una funzione composta, risulta

$$df_p(v) = \frac{\partial f}{\partial x}(p)h + \frac{\partial f}{\partial y}(p)k.$$

Pertanto (fissato p) df_p è un polinomio omogeneo di primo grado nelle due variabili h e k (cioè una forma lineare in \mathbb{R}^2).

Mostriamo ora che il differenziale secondo di f in p è un polinomio omogeneo di secondo grado (cioè una forma quadratica), e calcoliamone l'espressione. Dato un vettore applicato (p, v) , con $p = (x, y) \in U$ e $v = (h, k) \in \mathbb{R}^2$, occorre calcolare $\varphi''(0)$, dove

$$\varphi(t) = f(p(t)) = f(x(t), y(t)), \quad p(t) = (x(t), y(t)) = (x + th, y + tk).$$

Per la regola della catena, risulta

$$\varphi'(t) = \frac{\partial f}{\partial x}(x(t), y(t)) x'(t) + \frac{\partial f}{\partial y}(x(t), y(t)) y'(t) = \frac{\partial f}{\partial x}(p(t)) h + \frac{\partial f}{\partial y}(p(t)) k.$$

Derivando ancora $\varphi'(t)$ si ha

$$\varphi''(t) = \left(\frac{\partial^2 f}{\partial x^2}(p(t)) h + \frac{\partial^2 f}{\partial y \partial x}(p(t)) k \right) h + \left(\frac{\partial^2 f}{\partial x \partial y}(p(t)) h + \frac{\partial^2 f}{\partial y^2}(p(t)) k \right) k.$$

Quindi, tenendo conto del Teorema di Schwarz e ponendo $t = 0$, si ottiene

$$d^2 f_p(v) = \varphi''(0) = \frac{\partial^2 f}{\partial x^2}(p) h^2 + 2 \frac{\partial^2 f}{\partial x \partial y}(p) h k + \frac{\partial^2 f}{\partial y^2}(p) k^2.$$

Possiamo concludere che, fissato p , $d^2 f_p$ è una forma quadratica nelle due componenti h e k del vettore libero v (il quale, ricordiamo, rappresenta l'incremento della variabile vettoriale).

Ovviamente, quando f è una funzione di k variabili, con lo stesso metodo si prova che $d^2 f_p$ è ancora un polinomio omogeneo di secondo grado (cioè una forma quadratica), ma di k variabili (costituite dalle componenti v_1, v_2, \dots, v_k del vettore libero v).

Ricordiamo che una funzione reale definita in un intervallo si dice *convessa* [*concava*] se la corda (cioè il segmento) che congiunge due punti qualunque del suo grafico sta sopra [sotto] il grafico. Ovviamente, la restrizione di una funzione convessa $f: J \rightarrow \mathbb{R}$ ad un sottointervallo I di J è ancora una funzione convessa.

Si fa notare che una funzione f è concava se e solo se è convessa $-f$.

Si osservi che una funzione può essere contemporaneamente convessa e concava, ma solo in un caso: quando il suo grafico è rettilineo (ovvero, quando la sua derivata è costante).

Teorema^{sd}. *Sia $f: J \rightarrow \mathbb{R}$ derivabile in un intervallo J . Allora f è convessa [concava] se e solo se la retta tangente ad un punto qualunque del suo grafico sta sotto [sopra] il grafico.*

Corollario. *Sia $f: J \rightarrow \mathbb{R}$ convessa [concava] e derivabile in un intervallo J e sia $x_0 \in J$ un punto critico per f (ossia, un punto in cui si annulla la derivata di f). Allora x_0 è un punto di minimo [massimo] assoluto per f (in J).*

Dimostrazione (caso in cui f è convessa). La retta tangente al grafico di f nel punto $(x_0, f(x_0))$ è orizzontale (ha equazione $y = f(x_0)$). Dal fatto che il grafico di f sta sopra tale retta si deduce che $f(x) \geq f(x_0)$ per ogni $x \in J$. □

Il risultato che segue fornisce un metodo pratico per stabilire quando una funzione è convessa.

Teorema^{sd}. *Sia $f: J \rightarrow \mathbb{R}$ di classe C^2 in un intervallo J . Allora f è convessa [concava] se e solo se $f''(x) \geq 0$ [$f''(x) \leq 0$] per ogni x in J .*

Il concetto di funzione convessa di due o più variabili si introduce riconducendolo in modo naturale a quello di una sola variabile.

Definizione (di funzione convessa in \mathbb{R}^k). Sia $f: Q \rightarrow \mathbb{R}$ una funzione continua definita in un sottoinsieme convesso di \mathbb{R}^k . Si dice che f è *convessa* [*concava*] se è convessa [concava]

la sua “restrizione” ad ogni “intervallo” contenuto in Q . Più precisamente, se dato un arbitrario punto $p \in Q$ e un arbitrario vettore $v \in \mathbb{R}^k$, risulta convessa la funzione di una variabile $\varphi(t) = f(p + tv)$.

Come per il caso di una sola variabile, una funzione di due o più variabili f è (per definizione) *concava* se è convessa $-f$.

Per le funzioni di due variabili vale un risultato analogo a quello che per le funzioni di una sola variabile permette di caratterizzare la convessità mediante le rette tangenti al grafico. In questo caso il grafico non è una curva nel piano xy ma una superficie nello spazio xyz e il ruolo importante è giocato dai piani tangenti al grafico.

Teorema^{sd}. *Sia $f: Q \rightarrow \mathbb{R}$ di classe C^1 in un convesso Q di \mathbb{R}^2 . Allora f è convessa [concava] se e solo se il piano tangente ad un punto qualunque del suo grafico sta sotto [sopra] il grafico.*

Il suddetto teorema potrebbe essere esteso alle funzioni di tre o più variabili, ma al posto del piano tangente si dovrebbe usare il concetto di varietà lineare tangente. È meglio lasciar perdere; ma nel caso (auspicabile e poco probabile) che uno studente sia curioso di sapere che razza di aggeggio sia la varietà lineare tangente al grafico di una funzione $f: \mathbb{R}^k \rightarrow \mathbb{R}$ in un punto $(x_0, f(x_0))$, eccolo accontentato: è il sottoinsieme di $\mathbb{R}^k \times \mathbb{R}$ costituito dalle coppie (x, y) tali che $y - y_0 = f'(x_0)(x - x_0)$, dove $y_0 = f(x_0)$ e $f'(x_0)$ è la matrice jacobiana di f in x_0 (ricordiamo che $f'(x_0)$ è una matrice riga e $x - x_0$ è un vettore colonna).

Teorema. *Sia $f: Q \rightarrow \mathbb{R}$ una funzione convessa [concava] di classe C^1 su un sottoinsieme aperto e convesso di \mathbb{R}^k . Se $p_0 \in Q$ è un punto critico per f (cioè se $\nabla f(p_0) = 0$), allora è di minimo [massimo] assoluto (per f in Q).*

Dimostrazione^{fac} (caso in cui f è convessa). Sia p un arbitrario punto di Q . Occorre mostrare che $f(p_0) \leq f(p)$. Consideriamo la funzione convessa (di una sola variabile) $\varphi(t) = f(p_0 + t(p - p_0))$, il cui dominio è un intervallo J contenente $[0, 1]$. Poiché la derivata $\varphi'(0) = \nabla f(p_0) \cdot (p - p_0)$ di φ nel punto $t = 0$ è nulla, per il precedente corollario si ha $\varphi(0) \leq \varphi(t)$ per ogni $t \in J$, ossia il punto $t = 0$ è di minimo assoluto per la φ in J . Quindi, in particolare, $f(p_0) = \varphi(0) \leq \varphi(1) = f(p)$. □

Il suddetto teorema, nel caso delle funzioni di due variabili, ha una chiara interpretazione geometrica: se $f(x, y)$ è convessa e (x_0, y_0) è un suo punto critico, allora il piano tangente al grafico di f nel punto $(x_0, y_0, f(x_0, y_0))$ è orizzontale e quindi, dato che tale piano sta sotto il grafico, risulta $f(x, y) \geq f(x_0, y_0)$ per ogni (x, y) nel dominio di f .

Esercizio^{fac}. Sia f una funzione di classe C^2 in un aperto $U \subseteq \mathbb{R}^k$. Dati $p \in U$ e $v \in \mathbb{R}^k$, consideriamo la funzione di una variabile $\varphi(t) = f(p + tv)$. Provare che, fissato t_0 nel dominio di φ , si ha

$$\varphi'(t_0) = df(p + t_0v, v) \quad \text{e} \quad \varphi''(t_0) = d^2f(p + t_0v, v).$$

Suggerimento. Fissato t_0 nel dominio di φ , si consideri la funzione

$$\psi(t) = f(p + t_0v + tv) = \varphi(t_0 + t).$$

Il risultato che segue (la cui dimostrazione è lasciata per esercizio^{fac}) può essere considerato come una naturale estensione (e una facile conseguenza), al caso delle funzioni di più variabili, del fatto che una funzione di una variabile di classe C^2 è convessa se e solo se la sua derivata seconda è maggiore o uguale a zero. Per semplicità, a meno che non sia esplicitamente affermato il contrario, da ora in poi supporremo che le funzioni convesse siano definite in insiemi aperti (oltre che convessi). Pertanto, se p è un punto del dominio $Q \subseteq \mathbb{R}^k$ di una funzione convessa f e v è un arbitrario vettore di \mathbb{R}^k , il dominio della funzione composta $\varphi(t) = f(p + tv)$ è un intervallo aperto (costituito dall'insieme dei $t \in \mathbb{R}$ per cui $p + tv \in Q$).

Teorema. *Sia $f: Q \rightarrow \mathbb{R}$ di classe C^2 su un convesso $Q \subseteq \mathbb{R}^k$. Condizione necessaria e sufficiente affinché f sia una funzione convessa è che per ogni $p \in Q$ la forma quadratica d^2f_p sia semidefinita positiva.*

Dimostrazione^{fac}. Ricordiamo che f è convessa (in Q) se e solo se per ogni $v \in \mathbb{R}^k$ risulta convessa la funzione di una variabile $\varphi(t) = f(p + tv)$. La tesi segue subito dal caso unidimensionale tenendo conto che $\varphi''(t) = d^2f(p + tv, v)$. □

Sia $f: U \rightarrow \mathbb{R}$ una funzione di classe C^2 su un aperto U di \mathbb{R}^k . Dato che la forma quadratica d^2f_p dipende soltanto dalle derivate seconde di f in p , si potrebbe provare che (essendo f di classe C^2) se d^2f_p è definita positiva [definita negativa] in un punto, allora rimane definita positiva [definita negativa] in tutto un intorno del punto (provarlo per esercizio^{fac} per le funzioni di due variabili usando il teorema della permanenza del segno per le funzioni continue).

111 - Mercoledì 05/05/10

Teorema^{sd} (di rappresentazione delle forme quadratiche). *Ogni forma quadratica $g(v)$ in \mathbb{R}^k si può rappresentare nel modo seguente: $g(v) = Av \cdot v$, dove A è una matrice simmetrica $k \times k$ (detta “matrice associata alla forma quadratica”) e v (la variabile indipendente) è un generico vettore (colonna) di \mathbb{R}^k .*

Esercizio. Si consideri (in \mathbb{R}^3) la forma quadratica $ax^2 + by^2 + cz^2 + 2dxy + 2exz + 2fyz$ e se ne determini la matrice associata.

Suggerimento. Determinare, innanzi tutto, gli elementi della diagonale principale.

Con riferimento al precedente teorema, essendo la matrice A simmetrica, i suoi autovalori sono tutti reali (quindi ci sarà un minimo autovalore e un massimo autovalore). Proveremo in seguito (col metodo dei moltiplicatori di Lagrange) che i valori minimo e massimo assunti da $Av \cdot v$ al variare di v in tutte le possibili direzioni (cioè i vettori di norma 1) coincidono, rispettivamente, con gli autovalori minimo e massimo di A e le corrispondenti autodirezioni (cioè gli autovettori di norma 1) sono tra loro ortogonali, purché il massimo sia diverso dal minimo. Per inciso, si fa notare che, per il Teorema di Weierstrass, la forma quadratica $Av \cdot v$ deve necessariamente ammettere minimo e massimo (assoluti) sull'insieme dei vettori di norma uno, dato che detto insieme (denotato S^{k-1} e chiamato *sfera unitaria di \mathbb{R}^k*) è compatto (cioè limitato e chiuso).

Dal fatto che il minimo e il massimo autovalore di una matrice simmetrica A coincidono con gli estremi assoluti di $Av \cdot v$ in S^{k-1} si deduce facilmente il seguente importantissimo risultato:

Teorema (di classificazione delle forme quadratiche). *Sia $g(v) = Av \cdot v$ una forma quadratica in \mathbb{R}^k . Allora:*

- $Av \cdot v$ è definita positiva [negativa] se e solo se gli autovalori della matrice A sono tutti positivi [negativi];
- $Av \cdot v$ è semidefinita positiva [negativa] se e solo se tutti gli autovalori di A sono maggiori [minori] o uguali a zero;
- $Av \cdot v$ è indefinita se e solo se la matrice A ammette sia autovalori negativi sia autovalori positivi.

Esercizio^{fac}. Dimostrare il suddetto teorema.

Suggerimento. Si tenga presente che il minimo e il massimo autovalore di A coincidono, rispettivamente, con i valori minimo e massimo assunti della funzione $g(v) = Av \cdot v$ in S^{k-1} . Si osservi poi che $g(v)$ è una funzione omogenea di grado 2 (cioè $g(\lambda v) = \lambda^2 g(v)$, qualunque siano $\lambda \in \mathbb{R}$ e $v \in \mathbb{R}^k$). Quindi, ad esempio, se $g(v)$ è positiva per tutti i vettori di S^{k-1} , allora è positiva per tutti i vettori non nulli.

Teorema (di classificazione dei punti critici per le funzioni di k variabili). *Sia f una funzione reale di classe C^2 su un aperto U di \mathbb{R}^k e sia $p_0 \in U$ un punto critico per f (i.e. $\nabla f(p_0) = 0$). Se la forma quadratica $d^2f_{p_0}$ è definita positiva [definita negativa], allora p_0 è un punto di minimo [massimo] relativo per f . Se $d^2f_{p_0}$ è indefinita, allora p_0 non è un punto estremante per f .*

Dimostrazione^{fac}. Supponiamo che la forma quadratica $d^2f_{p_0}$ sia definita positiva [negativa]. Allora (essendo f di classe C^2) esiste un intorno $I(p_0, \delta)$ in cui la forma quadratica d^2f_p risulta definita positiva [negativa]. Pertanto, per il teorema precedente, f è convessa [concava] in $I(p_0, \delta)$ e, di conseguenza, il punto critico p_0 è di minimo [massimo] assoluto in tale intorno (cioè è di minimo [massimo] relativo per f). Supponiamo ora che la forma quadratica $d^2f_{p_0}$ sia indefinita. Allora, dal significato di differenziale secondo si deduce che esistono delle direzioni v per le quali la funzione $\varphi(t) = f(p_0 + tv)$ ha un minimo stretto per $t = 0$ (quelle per le quali $\varphi''(0) = d^2f_{p_0}(v) > 0$) e delle direzioni v per le quali $\varphi(t) = f(p_0 + tv)$ ha un massimo stretto per $t = 0$. Di conseguenza, in ogni intorno di p_0 ci sono punti p per cui $f(p) > f(p_0)$ e punti p per cui $f(p) < f(p_0)$. Il punto p_0 , quindi, non è né di massimo né di minimo (si dice che è un *punto sella* o un *passo di montagna* o un *punto iperbolico*). □

Il risultato che segue è un importantissimo caso particolare del precedente.

Corollario (di classificazione dei punti critici per le funzioni di due variabili). *Sia f una funzione (reale) di classe C^2 in un aperto U di \mathbb{R}^2 . Supponiamo che in un punto $p \in U$ si annullino le due derivate parziali di f . Consideriamo la forma quadratica*

$$d^2f_p(h, k) = \frac{\partial^2 f}{\partial x^2}(p)h^2 + 2\frac{\partial^2 f}{\partial x \partial y}(p)hk + \frac{\partial^2 f}{\partial y^2}(p)k^2.$$

Se $d^2f_p(h, k)$ è definita positiva, allora p è un punto di minimo relativo (per f); se è definita negativa, allora p è un punto di massimo relativo; se è indefinita, allora p non è né di massimo né di minimo.

Dal suddetto corollario e dalle considerazioni riguardanti lo studio del segno di una forma quadratica di due variabili, si deduce facilmente la seguente regola per lo studio dei punti critici di una funzione di due variabili.

Regola pratica (per la classificazione dei punti critici in \mathbb{R}^2). Sia $f: U \rightarrow \mathbb{R}$ una funzione di classe C^2 in un aperto U di \mathbb{R}^2 . Supponiamo che in un punto $p \in U$ si annullino le due derivate parziali di f . Allora, se il numero

$$Hf(p) = \frac{\partial^2 f}{\partial x^2}(p) \frac{\partial^2 f}{\partial y^2}(p) - \left(\frac{\partial^2 f}{\partial x \partial y}(p) \right)^2,$$

detto hessiano di f in p , è positivo, p è un punto estremante; se $Hf(p)$ è negativo, p non è estremante. In particolare, nel caso che $Hf(p)$ sia positivo, p è di minimo o di massimo a seconda che la derivata $\partial^2 f / \partial x^2(p)$ sia positiva o negativa.

Osservazione. Il numero $Hf(p)$ è il determinante della matrice

$$f''(p) = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}_p$$

detta *matrice hessiana* (o *derivata seconda*) di f in p . Si osservi che, in base al Teorema di Schwarz, tale matrice è simmetrica.

Si potrebbe dimostrare che la forma quadratica $d^2f_p(v)$ è definita positiva quando gli autovalori della matrice hessiana $f''(p)$ sono entrambi positivi, è definita negativa quando sono entrambi negativi, ed è indefinita se uno è negativo e l'altro è positivo.

Si osservi che $d^2f_p(v)$ coincide col seguente prodotto di matrici (righe per colonne):

$$v^t f''(p) v,$$

dove v^t denota il vettore (riga) trasposto di v . Un altro modo per rappresentare $d^2f_p(v)$ è

$$f''(p) v \cdot v,$$

cioè mediante il prodotto scalare dei due vettori (colonna) $f''(p)v$ e v (il primo dei quali si ottiene mediante il prodotto righe per colonne di $f''(p)$ con v).

Le suddette rappresentazioni di $d^2f_p(v)$ si prestano ad essere interpretate anche in \mathbb{R}^k (ma lasciamo perdere). Diciamo solo che l'elemento di riga i e colonna j della matrice hessiana in un punto p di una funzione f di k variabili è il numero

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

e che tale matrice è simmetrica grazie al Teorema di Schwarz.

ATTENZIONE! Il metodo che abbiamo illustrato (quello di considerare il determinante della matrice hessiana per stabilire se un punto è estremante) è valido solo per le funzioni di due variabili. Per convincersene si considerino i seguenti due esempi in \mathbb{R}^3 :

$$f_1(x, y, z) = x^2 + y^2 + z^2, \quad f_2(x, y, z) = x^2 - y^2 - z^2.$$

Entrambe le funzioni hanno un punto critico in $(0, 0, 0)$. Le loro matrici hessiane hanno determinante positivo (in detto punto), ma f_1 ha un punto estremante nell'origine mentre f_2 no (verificarlo, per esercizio, in base alla definizione di punto estremante).

Esercizio. Determinare i punti estremanti (cioè di massimo e di minimo relativo) della seguente funzione: $f(x, y) = x^3 + y^3 + x^2 + y^2 - x$.

Esercizio. Determinare i punti estremanti della seguente funzione:

$$f(x, y, z) = x^3 + y^3 + z^3 + 3x^2 + 3y^2 + 3z^2 - 1.$$

Esercizio. Svolgere un po' di esercizi sui massimi e minimi tratti dai libri consigliati (vedere la prima pagina del registro).

Data una funzione f di classe C^n in un aperto U di \mathbb{R}^k , il differenziale n -esimo $d^n f(p, v)$ di f valutato in un vettore applicato $(p, v) \in U \times \mathbb{R}^k$ è la derivata n -esima nel punto $t = 0$ della funzione composta $\varphi(t) = f(p + tv)$. È conveniente definire $d^n f$ anche per $n = 0$: significa derivare zero volte la funzione $\varphi(t)$ e valutarla nel punto $t = 0$, cioè $d^0 f(p, 0) = \varphi(0) = f(p)$. Si osservi che per $n = 1$ si ottiene $d^1 f(p, v) = \varphi'(0) = df(p, v)$. Come per i casi $n = 1$ e $n = 2$, $d^n f(p, v)$ si denota $d^n f_p(v)$ quando si pensa p fissato e si vuole mettere in evidenza la dipendenza da v (cioè la variabile indipendente è il vettore libero v invece del vettore applicato (p, v)).

Si potrebbe provare (ma non lo facciamo) che, fissato $p \in U$, $d^n f_p(v)$ è un polinomio omogeneo di grado n nelle variabili v_1, v_2, \dots, v_k , che costituiscono le componenti del vettore libero v .

A questo punto (data $f: U \rightarrow \mathbb{R}$ come sopra) è facile che lo studente faccia confusione tra le notazioni $d^n f$, $d^n f_p$, $d^n f(p, v)$ e $d^n f_p(v)$. Ricapitoliamo:

- $d^n f$ è un'applicazione dall'insieme $U \times \mathbb{R}^k$ dei vettori applicati di U in \mathbb{R} ;
- $d^n f_p$ è un'applicazione dall'insieme \mathbb{R}^k dei vettori liberi in \mathbb{R} (è lineare se $n = 1$);
- $d^n f(p, v)$ è il valore che assume $d^n f$ nel vettore applicato (p, v) ;
- $d^n f_p(v)$ è il valore che assume $d^n f_p$ nel vettore libero v (quindi $d^n f_p(v) = d^n f(p, v)$).

Si potrebbe dimostrare (ma non lo facciamo) che, avendo supposto f di classe C^n , è vera la seguente uguaglianza (detta *formula di Taylor di ordine n di f in p*):

$$f(p + v) = \sum_{i=0}^n \frac{1}{i!} d^i f_p(v) + \varepsilon(v) \|v\|^n.$$

Quindi, la formula di Taylor di ordine n di f in p è un'uguaglianza del tipo

$$f(p + v) = p_n(v) + \varepsilon(v) \|v\|^n,$$

dove $p_n(v)$ è un polinomio di grado minore o uguale ad n nella variabile vettoriale v (cioè nelle variabili reali v_1, v_2, \dots, v_k che compongono v). Si potrebbe provare (ma non lo facciamo) che vale un teorema di unicità della formula di Taylor. Ossia se vale l'uguaglianza

$$f(p+v) = p_n(v) + \varepsilon(v)\|v\|^n,$$

dove $p_n(v)$ è un polinomio di grado minore o uguale ad n in v , allora $p_n(v)$ è identico a

$$\sum_{i=0}^n \frac{1}{i!} d^i f_p(v),$$

nel senso che i monomi analoghi dei due polinomi hanno gli stessi coefficienti.

112 - Mercoledì 05/05/10

Consideriamo l'equazione $x^2 + y^2 - 1 = 0$. Le sue soluzioni sono l'insieme di livello zero

$$g^{-1}(0) = \{(x, y) : g(x, y) = 0\}$$

della funzione di due variabili $g(x, y) = x^2 + y^2 - 1$. Tale insieme, che rappresenta una circonferenza (in \mathbb{R}^2) con centro nell'origine e raggio 1, non può essere considerato il grafico di una funzione reale di una variabile reale $y = y(x)$. Il motivo è che per $x \in (-1, 1)$ esistono due valori di y con la proprietà che $(x, y) \in g^{-1}(0)$, e non uno solo, come è richiesto dalla definizione di funzione. Per la stessa ragione $g^{-1}(0)$ non è neppure il grafico di una funzione del tipo $x = x(y)$. Osserviamo però che $g^{-1}(0)$ ammette dei sottoinsiemi che possono essere considerati dei grafici di funzioni continue. Ad esempio, l'insieme definito dall'equazione $y = \sqrt{1-x^2}$ è un sottoinsieme di $g^{-1}(0)$ e rappresenta il grafico di una funzione continua per $x \in [-1, 1]$ (è addirittura C^∞ se la si restringe all'intervallo aperto $(-1, 1)$). Si dice allora che $y = \sqrt{1-x^2}$ è una *funzione implicita* definita dall'equazione $x^2 + y^2 - 1 = 0$ (più precisamente si dovrebbe dire che $y = \sqrt{1-x^2}$ è l'equazione del grafico di una funzione ma, visto che ogni funzione è univocamente determinata dal suo grafico, non c'è pericolo di confusione). Osserviamo che anche la funzione $y = -\sqrt{1-x^2}$ è implicitamente definita da $x^2 + y^2 - 1 = 0$. Altre due si ottengono ricavando x in funzione di y : $x = \sqrt{1-y^2}$ e $x = -\sqrt{1-y^2}$. Quindi, in un certo senso, si può dire che l'equazione $x^2 + y^2 - 1 = 0$ definisce implicitamente quattro funzioni continue: due che si ottengono ricavando la y in funzione della x e le altre due ricavando la x in funzione della y . In realtà le funzioni implicite definite da $x^2 + y^2 - 1 = 0$ sono infinite, perché si dovrebbero considerare anche tutte le restrizioni delle suddette quattro funzioni ai sottointervalli del loro dominio. Tuttavia le quattro funzioni trovate sono le uniche *massimali* (continue); nel senso che non sono restrizioni (proprie) di altre funzioni implicite. Non sempre la situazione è così semplice come nel suddetto esempio. Non è detto infatti che, data un'equazione del tipo $g(x, y) = 0$, sia facile ricavare esplicitamente una variabile in funzione dell'altra. Spesso, una volta accertata l'esistenza di una funzione implicita, per ricavarla si dovrà ricorrere a metodi numerici, o accontentarsi di trovarne la formula di Taylor di un dato

ordine. Riguardo all'esistenza della funzione implicita, riportiamo, senza dimostrazione, il seguente importante risultato attribuito al matematico pisano Ulisse Dini (1845-1918).

Teorema^{sd} (di Dini, o della funzione implicita in \mathbb{R}^2). *Sia $g:U \rightarrow \mathbb{R}$ una funzione di classe C^n ($n \geq 1$) su un aperto U di \mathbb{R}^2 . Dato $(x_0, y_0) \in g^{-1}(0)$, supponiamo che la derivata parziale $\partial g/\partial y(x_0, y_0)$ di g in (x_0, y_0) sia diversa da zero. Allora in un intorno di (x_0, y_0) l'insieme di livello $g^{-1}(0)$ è il grafico di una funzione $y(x)$ di classe C^n . Più precisamente, esistono $\varepsilon > 0$ e $\delta > 0$ tali che per ogni $x \in (x_0 - \delta, x_0 + \delta)$ esiste un unico $y = y(x) \in (y_0 - \varepsilon, y_0 + \varepsilon)$ che verifica la condizione $g(x, y(x)) = 0$. Inoltre, la funzione $y(x)$, detta funzione implicita definita dall'equazione $g(x, y) = 0$, ha la stessa regolarità di g (ovvero, è di classe C^n).*

Con riferimento al Teorema di Dini, osserviamo che la funzione composta $g(x, y(x))$ risulta identicamente nulla nell'intervallo $J = (x_0 - \delta, x_0 + \delta)$. Pertanto, derivandola rispetto alla x , si ottiene l'identità

$$\frac{\partial g}{\partial x}(x, y(x)) + \frac{\partial g}{\partial y}(x, y(x)) y'(x) = 0, \quad \forall x \in J.$$

In particolare, tenendo conto dell'ipotesi $\partial g/\partial y(x_0, y_0) \neq 0$, si ha

$$y'(x_0) = -\frac{\frac{\partial g}{\partial x}(x_0, y_0)}{\frac{\partial g}{\partial y}(x_0, y_0)}.$$

Derivando ancora la suddetta espressione si ottiene l'identità

$$\frac{\partial^2 g}{\partial x^2} + 2\frac{\partial^2 g}{\partial x \partial y} y' + \frac{\partial^2 g}{\partial y^2} y'^2 + \frac{\partial g}{\partial y} y'' = 0,$$

che permette di ricavare y'' (visto che $\partial g/\partial y \neq 0$).

La conoscenza dei tre numeri $y(x_0) = y_0$, $y'(x_0)$ e $y''(x_0)$ ci consente di scrivere la formula di Taylor della funzione implicita:

$$y(x_0 + h) = y_0 + y'(x_0)h + \frac{y''(x_0)}{2}h^2 + \varepsilon(h)h^2.$$

Osservazione. È ovvio che nel teorema della funzione implicita i ruoli di x e y possono essere scambiati. Pertanto, se si suppone che in un punto $(x_0, y_0) \in g^{-1}(0)$ si abbia $\partial g/\partial x(x_0, y_0) \neq 0$, allora $g^{-1}(0)$, in un intorno di (x_0, y_0) , è il grafico di una funzione $x(y)$ classe C^n (definita in un intorno di y_0).

113 - Venerdì 07/05/10

Consideriamo l'equazione $g(x, y, z) = 0$ definita dalla funzione

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

Questa definisce una superficie sferica con centro nell'origine di \mathbb{R}^3 e raggio uno. L'insieme di livello $g^{-1}(0)$ non è il grafico di una funzione reale di due variabili reali $z = z(x, y)$. Tuttavia, se dall'equazione

$$x^2 + y^2 + z^2 - 1 = 0$$

si ricava la z , si ottengono due funzioni reali (e continue) di due variabili:

$$z = \sqrt{1 - x^2 - y^2} \quad \text{e} \quad z = -\sqrt{1 - x^2 - y^2},$$

ciascuna delle quali è definita nel cerchio chiuso $x^2 + y^2 \leq 1$. Entrambe le funzioni hanno il grafico che soddisfa l'equazione $x^2 + y^2 + z^2 - 1 = 0$, ossia contenuto nell'insieme di livello $g^{-1}(0)$. Altre due funzioni si ottengono ricavando la y , e altre due ricavando la x . L'insieme $g^{-1}(0)$ contiene quindi il grafico di sei funzioni (continue) reali di due variabili reali (che rappresentano le funzioni implicite massimali definite da $x^2 + y^2 + z^2 - 1 = 0$).

In generale, data un'equazione del tipo $g(x, y, z) = 0$, non è sempre facile ricavare esplicitamente una variabile in funzione delle altre due. Una volta accertata l'esistenza di una funzione implicita, per calcolarla si dovrà ricorrere a metodi numerici, o accontentarsi di determinarne la formula di Taylor di un certo ordine. Un risultato che ci assicura l'esistenza di tale funzione è costituito dalla seguente versione in \mathbb{R}^3 del Teorema di Dini:

Teorema^{sd} (della funzione implicita in \mathbb{R}^3). *Sia $g: U \rightarrow \mathbb{R}$ una funzione di classe C^n ($n \geq 1$) su un aperto U di \mathbb{R}^3 e sia $(x_0, y_0, z_0) \in g^{-1}(0)$. Se $\partial g / \partial z(x_0, y_0, z_0) \neq 0$, allora, in un conveniente intorno di (x_0, y_0, z_0) , $g^{-1}(0)$ è il grafico di una funzione (implicita) $z = z(x, y)$ di classe C^n in un intorno del punto $(x_0, y_0) \in \mathbb{R}^2$.*

È ovvio che nel suddetto teorema il ruolo della variabile z potrà essere assunto da una qualunque delle altre variabili. Se, ad esempio, si suppone che in un punto (x_0, y_0, z_0) di $g^{-1}(0)$ si abbia $\partial g / \partial x(x_0, y_0, z_0) \neq 0$, allora, in un intorno di tale punto, $g^{-1}(0)$ è il grafico di una funzione $x = x(y, z)$ di classe C^n , ed è quindi verificata la condizione $g(x(y, z), y, z) = 0$ per ogni (y, z) in un intorno di (y_0, z_0) . Possiamo concludere che se in un punto p di $g^{-1}(0)$ il gradiente di g non è nullo, allora (localmente) almeno una delle tre variabili può essere pensata funzione delle altre due.

Definizione (di superficie di livello regolare). Sia $g: U \rightarrow \mathbb{R}$ una funzione di classe C^1 su un aperto U di \mathbb{R}^3 . Dato (un livello) $c \in \mathbb{R}$, l'insieme (di livello)

$$g^{-1}(c) = \{(x, y, z) \in U : g(x, y, z) = c\}$$

si dice una *superficie di livello regolare* se in ogni punto $p \in g^{-1}(c)$ risulta $\nabla g(p) \neq 0$; ossia, se l'insieme di livello c non contiene punti critici di g .

Esercizio. Mostrare che l'insieme $S^2 = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + z^2 = 1\}$ è una superficie di livello regolare.

Consideriamo il seguente sistema di due equazioni in tre incognite:

$$\begin{cases} x^2 + y^2 + z^2 - 1 & = & 0 \\ x - y & & = & 0 \end{cases}$$

L'idea è che tre incognite sono troppe per due sole equazioni. Se ad una delle incognite si assegna un valore, forse c'è speranza di risolvere il sistema rispetto alle altre due. Quando ciò è possibile, due delle incognite risultano funzione della terza. In questo particolare esempio è facile ricavare y e z in funzione di x . Risulta infatti:

$$y = x, \quad z = \pm\sqrt{1 - 2x^2}.$$

Si ottengono così due applicazioni continue, a valori in \mathbb{R}^2 , definite per x nell'intervallo chiuso $[-\sqrt{2}/2, \sqrt{2}/2]$:

$$x \mapsto (x, \sqrt{1 - 2x^2}); \quad x \mapsto (x, -\sqrt{1 - 2x^2}).$$

Osserviamo che, per come sono state ricavate, entrambe le applicazioni hanno il grafico contenuto nell'insieme delle soluzioni del suddetto sistema (si dice pertanto che sono implicitamente definite da tale sistema).

114 - Venerdì 07/05/10

In generale, data un'applicazione $g: U \rightarrow \mathbb{R}^s$, di classe C^1 su un aperto di \mathbb{R}^k , consideriamo l'equazione (vettoriale) $g(p) = 0$, detta anche "sistema di s equazioni in k incognite". Se $k > s$, ci sono più incognite che equazioni. Assegnando dei valori a $k - s$ incognite, può darsi che le rimanenti s incognite possano essere ricavate in funzione dei $k - s$ valori fissati. Da qui nasce l'idea di funzione implicita associata all'equazione (vettoriale) $g(p) = 0$.

Per precisare meglio il concetto di funzione implicita è conveniente rappresentare i vettori k -dimensionali di \mathbb{R}^k come coppie di vettori $(x, y) \in \mathbb{R}^{k-s} \times \mathbb{R}^s$. In tal modo la precedente equazione diventa $g(x, y) = 0$, e una funzione implicita (della y in funzione della x) altro non è che un'applicazione continua φ , a valori in \mathbb{R}^s , definita in un sottoinsieme connesso V di \mathbb{R}^{k-s} , il cui grafico

$$G = \left\{ (x, y) \in \mathbb{R}^{k-s} \times \mathbb{R}^s : y = \varphi(x) \right\}$$

è costituito da punti $p = (x, y)$ che verificano l'equazione assegnata. Ossia, G è un sottoinsieme dell'insieme di livello $g^{-1}(0)$. In altre parole, si ha $g(x, \varphi(x)) = 0, \forall x \in V$.

Diamo ora, senza dimostrazione, la versione più generale del teorema della funzione implicita (di Ulisse Dini).

Teorema^{sd} (della funzione implicita in \mathbb{R}^k). *Sia $g: U \rightarrow \mathbb{R}^s$ un'applicazione C^n su un aperto U di $\mathbb{R}^{k-s} \times \mathbb{R}^s$ e sia $p_0 = (x_0, y_0)$ un punto di $g^{-1}(0)$. Se il determinante della matrice*

$$\frac{\partial g}{\partial y}(p_0) = \frac{\partial(g_1, g_2, \dots, g_s)}{\partial(y_1, y_2, \dots, y_s)}(p_0),$$

costituita dalle ultime s colonne della matrice jacobiana di g in p_0 , è diverso da zero, allora $g^{-1}(0)$, in un intorno di p_0 , è il grafico di una funzione $y(x)$ di classe C^n , definita in un intorno di $x_0 \in \mathbb{R}^{k-s}$, a valori in \mathbb{R}^s .

Sia $g: U \rightarrow \mathbb{R}$ una funzione C^1 in un aperto di \mathbb{R}^2 . Abbiamo visto che se $y = y(x)$ è una funzione implicita (di classe C^1) per l'equazione $g(x, y) = 0$, allora

$$\frac{\partial g}{\partial x}(x, y(x)) + \frac{\partial g}{\partial y}(x, y(x)) y'(x) = 0.$$

Tale formula, se è soddisfatta la condizione $\partial g / \partial y(x, y(x)) \neq 0$, permette di ricavare $y'(x)$ in funzione delle derivate parziali di g . L'uguaglianza può essere interpretata anche quando g è un'applicazione da un aperto U di $\mathbb{R}^{k-s} \times \mathbb{R}^s$ in \mathbb{R}^s . In questo caso i simboli $\partial g / \partial x$ e $\partial g / \partial y$ sono sottomatrici della matrice jacobiana di g e y' è la matrice jacobiana della funzione implicita. Se è soddisfatta la condizione $\det(\partial g / \partial y) \neq 0$, allora la matrice $\partial g / \partial y$ è invertibile, e moltiplicando a sinistra entrambi i membri della suddetta formula per l'inversa di $\partial g / \partial y$ si ottiene

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

È ovvio che nel suddetto teorema il ruolo delle ultime s variabili (rappresentate dal vettore y) potrà essere assunto da qualunque scelta di s variabili tra le k disponibili; non occorre che siano le ultime s . L'importante, per poter applicare il teorema della funzione implicita (in un punto $p_0 \in g^{-1}(0)$), è che la sottomatrice quadrata della matrice jacobiana che si ottiene considerando le derivate parziali delle funzioni g_1, g_2, \dots, g_s rispetto alle variabili scelte (cioè quelle che si vogliono ricavare) sia invertibile (per $p = p_0$). Pertanto, se si suppone che in un punto $p_0 \in g^{-1}(0)$ il rango della matrice jacobiana $g'(p_0)$ sia s (ossia, uguale alla dimensione dello spazio di arrivo), allora (in un intorno di p_0) s variabili possono essere pensate funzioni delle rimanenti $k - s$.

115 - Mercoledì 12/05/10

Abbiamo già incontrato i concetti di “curva di livello regolare” (in \mathbb{R}^2) e di “superficie di livello regolare” (in \mathbb{R}^3). In generale, per applicazioni da \mathbb{R}^k in \mathbb{R}^s non ci sono sufficienti nomi per distinguere tutti i casi possibili. Occorre perciò una nuova definizione.

Definizione (di punto critico, di punto regolare e di varietà di livello regolare). Sia $g: U \rightarrow \mathbb{R}^s$ un'applicazione C^1 su un aperto U di \mathbb{R}^k . Un punto $p \in U$ si dice *critico* per g se la matrice jacobiana $g'(p)$ ha rango minore di s (cioè minore della dimensione dello spazio bersaglio), in caso contrario si dice *regolare*. Dato (un livello vettoriale) $c \in \mathbb{R}^s$, l'insieme (di livello)

$$g^{-1}(c) = \{p \in U : g(p) = c\}$$

si dice una *varietà di livello regolare* ($k - s$)-dimensionale se non contiene punti critici (o, equivalentemente, se tutti i punti di $g^{-1}(c)$ sono regolari). In particolare, una varietà di livello regolare si dice una curva (di livello regolare) se ha dimensione 1 e si dice una superficie (di livello regolare) se ha dimensione 2.

Esercizio. Mostrare che l'insieme definito dal sistema

$$\begin{cases} x^2 + y^2 & = & 1 \\ x + y + z & = & 0 \end{cases}$$

è una curva di livello regolare in \mathbb{R}^3 (è l'intersezione di un cilindro con un piano).

Suggerimento. Determinare i punti critici dell'applicazione $g: \mathbb{R}^3 \rightarrow \mathbb{R}^2$ definita da

$$(x, y, z) \mapsto (x^2 + y^2, x + y + z)$$

e verificare che nessuno di essi verifica il sistema.

Definizione (di vettore tangente e di spazio tangente). Data una varietà di livello (regolare) $M \subseteq \mathbb{R}^k$ e dato un punto $p \in M$, un vettore $v \in \mathbb{R}^k$ si dice *tangente ad M in p* se esiste una curva parametrica $\gamma: (-\varepsilon, \varepsilon) \rightarrow M$, di classe C^1 , passante per p quando $t = 0$ e tale che $\gamma'(0) = v$. L'insieme di tutti i vettori tangenti ad M in p si chiama *spazio tangente ad M in p* e si denota $T_p(M)$.

Osserviamo che, se $M \subseteq \mathbb{R}^k$ è una varietà di livello e $p \in M$, il vettore nullo di \mathbb{R}^k appartiene a $T_p(M)$. Per provarlo basta considerare la curva costante $\gamma: (-1, 1) \rightarrow M$ definita da $\gamma(t) = p$ per ogni $t \in (-1, 1)$. Più in generale, vale il seguente risultato:

Teorema^{sd}. *Lo spazio tangente in un punto p ad una varietà di livello regolare $M \subseteq \mathbb{R}^k$ è un sottospazio di \mathbb{R}^k della stessa dimensione di M .*

116 - Mercoledì 12/05/10

Ricordiamo che due vettori u e v di \mathbb{R}^k si dicono *ortogonali tra loro* (si scrive $u \perp v$) se il loro prodotto scalare $u \cdot v$ è nullo. Se un vettore $v \in \mathbb{R}^k$ è ortogonale a tutti i vettori di \mathbb{R}^k , allora, in particolare, è ortogonale a se stesso, e quindi $v \cdot v = \|v\|^2 = 0$, ossia $v = 0$. Si dice che $v \in \mathbb{R}^k$ è ortogonale (o normale) ad un sottospazio E di \mathbb{R}^k (si scrive $v \perp E$) se è ortogonale ad ogni vettore di E . L'insieme dei vettori ortogonali ad E è un sottospazio di \mathbb{R}^k e si denota E^\perp . Ricordiamo infine che, dato un qualunque sottospazio E di \mathbb{R}^k , risulta $\mathbb{R}^k = E \oplus E^\perp$; ossia, ogni vettore di \mathbb{R}^k è univocamente decomponibile nella somma di due vettori: uno di E e uno di E^\perp . In questo caso si ha $\dim E + \dim E^\perp = k$.

Definizione (di vettore normale e di spazio normale). Data una varietà di livello (regolare) $M \subseteq \mathbb{R}^k$ e dato un punto $p \in M$, un vettore $v \in \mathbb{R}^k$ si dice *normale (o ortogonale) ad M in p* se è ortogonale a $T_p(M)$. L'insieme di tutti i vettori normali ad M in p si chiama *spazio normale (o ortogonale) ad M in p* e si denota $T_p(M)^\perp$.

Un metodo pratico per determinare se un vettore è tangente in un punto ad una varietà di livello regolare è fornito dal seguente risultato:

Teorema (di caratterizzazione dei vettori tangenti). *Sia $g: U \rightarrow \mathbb{R}^s$ una funzione C^1 su un aperto $U \subseteq \mathbb{R}^k$ e sia $M = g^{-1}(0)$ una varietà di livello regolare. Allora, dato $p \in M$, un vettore $v \in \mathbb{R}^k$ è tangente ad M in p se e solo se $g'(p)v = 0$ (cioè, se e solo se $v \perp \nabla g_i(p)$ per ogni $i = 1, 2, \dots, s$, dove g_1, g_2, \dots, g_s sono le funzioni componenti di g).*

Dimostrazione. Sia $v \in \mathbb{R}^k$ un vettore tangente ad M in p . Esiste una curva parametrica $\gamma: (-\varepsilon, \varepsilon) \rightarrow M$, di classe C^1 , tale che $\gamma(0) = p$ e $\gamma'(0) = v$. Poiché g è nulla nei punti di M , la funzione composta $\varphi(t) = g(\gamma(t))$ è identicamente nulla in $(-\varepsilon, \varepsilon)$. Quindi, per il teorema di derivazione di una funzione composta, si ha $\varphi'(t) = g'(\gamma(t))\gamma'(t) = 0$ per

ogni $t \in (-\varepsilon, \varepsilon)$. Dunque, per $t = 0$, si ottiene $g'(\gamma(0))\gamma'(0) = g'(p)v = 0$, e questo prova che $v \in T_p(M)$ **solo se** $g'(p)v = 0$. Cioè lo spazio tangente ad M in p è contenuto nel nucleo dell'applicazione lineare $dg_p: \mathbb{R}^k \rightarrow \mathbb{R}^s$ definita dalla matrice jacobiana $g'(p)$, detta *differenziale di g in p* (ossia $T_p(M) \subseteq \text{Ker } dg_p$).

Viceversa^{fac}, proviamo che **se** v verifica la condizione $g'(p)v = 0$, allora v è tangente ad M in p . Ossia, proviamo che $\text{Ker } dg_p \subseteq T_p(M)$. Sappiamo, per il teorema precedente, che $T_p(M)$ è uno spazio $(k - s)$ -dimensionale contenuto in \mathbb{R}^k . Anche il nucleo di dg_p è uno spazio $(k - s)$ -dimensionale, dato che, per ipotesi, dg_p è un'applicazione suriettiva (cioè, di rango s). Dunque i due spazi, $T_p(M)$ e $\text{Ker } dg_p$, devono necessariamente coincidere, visto che hanno la stessa dimensione e (per la parte necessaria già provata) il primo è contenuto nel secondo. □

Illustriamo l'utilità del precedente teorema con degli esempi.

Consideriamo, in \mathbb{R}^3 , un'equazione del tipo $g(x, y, z) = 0$, dove g è C^1 (in un aperto di \mathbb{R}^3) e soddisfa la condizione $\nabla g(x, y, z) \neq 0$ per ogni $(x, y, z) \in M = g^{-1}(0)$. Quindi M è una superficie di livello regolare in \mathbb{R}^3 . In base al precedente teorema, lo spazio tangente ad M in un punto $p = (x_0, y_0, z_0) \in M$ coincide con lo spazio ortogonale al vettore (non nullo) $\nabla g(p)$ o, equivalentemente, col nucleo del differenziale dg_p . In altre parole, lo spazio tangente ad M in p è costituito dai vettori $v = (v_1, v_2, v_3)$ che verificano l'equazione

$$\frac{\partial g}{\partial x}(p)v_1 + \frac{\partial g}{\partial y}(p)v_2 + \frac{\partial g}{\partial z}(p)v_3 = 0,$$

che rappresenta un piano passante per l'origine.

Non dobbiamo confondere lo spazio tangente col piano tangente. Lo spazio tangente infatti, essendo un sottospazio, passa sempre per l'origine, mentre il piano tangente in p ad $M = g^{-1}(0)$ è, per definizione, il piano passante per p parallelo allo spazio tangente. Ossia, è il piano costituito dai punti (x, y, z) con la proprietà che i vettori $v = (x - x_0, y - y_0, z - z_0)$ sono tangenti ad M in $p = (x_0, y_0, z_0)$. Di conseguenza la sua equazione è data da

$$\frac{\partial g}{\partial x}(p)(x - x_0) + \frac{\partial g}{\partial y}(p)(y - y_0) + \frac{\partial g}{\partial z}(p)(z - z_0) = 0.$$

Consideriamo ora un altro esempio, sempre in \mathbb{R}^3 . Siano $g_1(x, y, z) = 0$ e $g_2(x, y, z) = 0$ due equazioni, dove $g_1(x, y, z)$ e $g_2(x, y, z)$ sono di classe C^1 su un aperto U di \mathbb{R}^3 . Se si suppone che la matrice jacobiana dell'applicazione $g = (g_1, g_2)$ abbia rango due in ogni punto (x, y, z) che verifica le due suddette equazioni, allora il sistema

$$\begin{cases} g_1(x, y, z) = 0 \\ g_2(x, y, z) = 0 \end{cases}$$

definisce una varietà di livello regolare di dimensione uno $M \subseteq U$ (cioè una curva di livello regolare). In questo caso lo spazio tangente ad M in un punto $p \in M$ è costituito dai vettori ortogonali ai due gradienti $\nabla g_1(p)$ e $\nabla g_2(p)$. Cioè un vettore $v = (v_1, v_2, v_3)$ appartiene a $T_p(M)$ se e solo se verifica le due equazioni

$$\frac{\partial g_1}{\partial x}(p)v_1 + \frac{\partial g_1}{\partial y}(p)v_2 + \frac{\partial g_1}{\partial z}(p)v_3 = 0,$$

$$\frac{\partial g_2}{\partial x}(p)v_1 + \frac{\partial g_2}{\partial y}(p)v_2 + \frac{\partial g_2}{\partial z}(p)v_3 = 0,$$

che definiscono una retta passante per l'origine (data dall'intersezione di due piani). Da non confondere con la retta tangente ad M in p che, per definizione, è la retta passante per p , parallela allo spazio tangente. Si fa notare che $\nabla g_1(p)$ e $\nabla g_2(p)$ sono le righe della matrice jacobiana $g'(p)$ che, per ipotesi, ha rango due. Pertanto i due vettori $\nabla g_1(p)$ e $\nabla g_2(p)$ sono linearmente indipendenti e generano quindi lo spazio bidimensionale $T_p(M)^\perp$.

A titolo di esempio, mostriamo come l'insieme delle posizioni (detto *spazio delle configurazioni*) di un'asta rigida in \mathbb{R}^2 (di lunghezza assegnata l) possa essere pensato come una varietà di livello tridimensionale in \mathbb{R}^4 . La posizione dell'asta in \mathbb{R}^2 è univocamente individuata dai suoi estremi p_1 e p_2 . Quindi, ad ogni posizione dell'asta corrisponde (ossia, si può associare) un punto $(p_1, p_2) \in \mathbb{R}^2 \times \mathbb{R}^2$, e tale corrispondenza è iniettiva. Possiamo dunque identificare lo spazio delle configurazioni dell'asta con il sottoinsieme M di \mathbb{R}^4 costituito dall'immagine dell'applicazione appena definita. I punti di $M \subseteq \mathbb{R}^4$ sono le coppie (p_1, p_2) di punti di \mathbb{R}^2 che verificano la condizione $\|p_2 - p_1\|^2 = l^2$. Pertanto

$$M = \left\{ (x_1, y_1, x_2, y_2) \in \mathbb{R}^4 : (x_2 - x_1)^2 + (y_2 - y_1)^2 = l^2 \right\}.$$

In altre parole $M = g^{-1}(0)$, dove $g: \mathbb{R}^4 \rightarrow \mathbb{R}$ è l'applicazione

$$g(x_1, y_1, x_2, y_2) = (x_2 - x_1)^2 + (y_2 - y_1)^2 - l^2.$$

Si lascia allo studente il compito di verificare che nessuno dei punti critici di g verifica la condizione $g(x_1, y_1, x_2, y_2) = 0$. Pertanto lo spazio M delle configurazioni dell'asta è una varietà di livello regolare in \mathbb{R}^4 . In questo esempio fisico i vettori dello spazio tangente ad M in un punto $p = (p_1, p_2) \in M$ si chiamano anche *spostamenti virtuali* di p . Euristicamente parlando, si chiamano virtuali perché “partono per la tangente”, e se gli si stesse dietro “si uscirebbe dal vincolo” M (anche se di poco) e non si compirebbe quindi uno “spostamento reale” (ma quasi).

Con lo stesso metodo non è difficile verificare che le configurazioni di un'asta rigida in \mathbb{R}^3 costituiscono una varietà di dimensione 5 in \mathbb{R}^6 . In modo analogo, ma con più calcoli, si potrebbe provare che lo spazio delle configurazioni di un corpo rigido è una varietà di dimensione 6 in \mathbb{R}^9 (per provarlo occorre tener conto che la posizione di un corpo rigido è individuata dai vertici di un triangolo solidale con il corpo).

Concludiamo dicendo (ma senza provarlo) che le configurazioni di un pendolo piano si identificano con una circonferenza in \mathbb{R}^2 , di un pendolo sferico con una superficie sferica in \mathbb{R}^3 , di un pendolo doppio con un toro in \mathbb{R}^4 (un toro è una superficie a forma di ciambella).

117 - Mercoledì 12/05/10

Ricordiamo che, data una funzione reale f definita in un aperto U di \mathbb{R}^k e dato un sottoinsieme X di U , un punto $p \in X$ si dice *estremante per f in X* se è di massimo o di minimo relativo per la restrizione di f ad X . Sappiamo (per il Teorema di Fermat in \mathbb{R}^k)

che se $X = U$ ed f è C^1 , allora una condizione necessaria affinché un punto $p \in X$ sia estremante è che sia un punto critico per f (sia nullo, cioè, il gradiente di f in p).

Il problema della ricerca dei punti estremanti di una funzione reale f in un aperto U di \mathbb{R}^k si dice un *problema di massimi e minimi liberi* per distinguerlo da quello che consiste nella ricerca dei punti estremanti della restrizione di f ad un assegnato sottoinsieme (non aperto) X di U (come, ad esempio, un quadrato in \mathbb{R}^2 , una circonferenza in \mathbb{R}^2 , una superficie in \mathbb{R}^3 ecc.). Se, in particolare, l'insieme in cui si cercano i punti estremanti è una varietà di livello regolare $M = g^{-1}(0) \subseteq U$, allora il problema si dice di *massimi e minimi condizionati* o *vincolati*, ed M si dice il *vincolo* e l'equazione $g(x) = 0$ si dice la *condizione* a cui devono sottostare i punti cercati.

Teorema di Fermat (per gli estremi condizionati). *Sia f una funzione reale di classe C^1 su un aperto U di \mathbb{R}^k e sia M una varietà di livello regolare contenuta in U . Condizione necessaria affinché $p \in M$ sia un punto estremante per f in M è che il gradiente di f in p sia ortogonale a $T_p(M)$.*

Dimostrazione. Sia $p \in M$ un punto estremante per la restrizione di f ad M . Per fissare le idee supponiamo sia di massimo relativo. In base alla definizione di spazio tangente, è sufficiente provare che se γ è una curva C^1 a valori in M che passa per p quando $t = 0$, allora $\nabla f(p) \cdot \gamma'(0) = 0$. Sia γ una tale curva. Consideriamo la funzione composta $\varphi(t) = f(\gamma(t))$. Poiché l'immagine di γ è contenuta in M e $\gamma(0) = p$, il punto $t = 0$ è di massimo relativo per φ e, di conseguenza, la derivata $\varphi'(t) = \nabla f(\gamma(t)) \cdot \gamma'(t)$ si annulla per $t = 0$. Pertanto $\nabla f(p) \cdot \gamma'(0) = 0$. □

Corollario. *Sia $g: U \rightarrow \mathbb{R}$ una funzione di classe C^1 su un aperto U di \mathbb{R}^k . Se g è costante su una varietà $M \subseteq U$, allora $\nabla g(p)$ è ortogonale a $T_p(M)$ per ogni $p \in M$.*

Dimostrazione. La tesi segue dal fatto che ogni punto di M è estremante per g (è addirittura sia di massimo sia di minimo). □

Per illustrare il Teorema di Fermat (per gli estremi condizionati) consideriamo, ad esempio, due funzioni reali f e g di classe C^1 su un aperto U di \mathbb{R}^3 . Supponiamo che l'insieme $\Sigma = g^{-1}(0)$ sia una superficie di livello regolare e che $p = (x, y, z)$ sia un punto estremante per f in Σ . Dal Teorema di Fermat (e dal precedente corollario) segue $\nabla f(p), \nabla g(p) \in T_p(\Sigma)^\perp$. D'altra parte $\mathbb{R}^3 = T_p(\Sigma) \oplus T_p(\Sigma)^\perp$, e quindi $T_p(\Sigma)^\perp$ è uno spazio unidimensionale, visto che è bidimensionale $T_p(\Sigma)$. Poiché, per ipotesi, il vettore $\nabla g(p)$ è non nullo, esso costituisce una base di $T_p(\Sigma)^\perp$. Esiste quindi un (unico) $\lambda \in \mathbb{R}$ tale che $\nabla f(p) = \lambda \nabla g(p)$. Ossia, il punto $(p, \lambda) = (x, y, z, \lambda)$ di $\mathbb{R}^3 \times \mathbb{R}$ verifica il sistema

$$\begin{cases} \frac{\partial f}{\partial x}(x, y, z) = \lambda \frac{\partial g}{\partial x}(x, y, z) \\ \frac{\partial f}{\partial y}(x, y, z) = \lambda \frac{\partial g}{\partial y}(x, y, z) \\ \frac{\partial f}{\partial z}(x, y, z) = \lambda \frac{\partial g}{\partial z}(x, y, z) \\ g(x, y, z) = 0. \end{cases}$$

In altre parole, il punto (x, y, z, λ) è una soluzione del sistema che si ottiene uguagliando

a zero le quattro derivate parziali della funzione

$$F(x, y, z, \lambda) = f(x, y, z) - \lambda g(x, y, z).$$

118 - Venerdì 14/05/10

Un metodo pratico per la ricerca dei massimi e minimi condizionati è dovuto al matematico italo-francese Lagrange (1736-1813). Tale metodo si basa sul seguente risultato che, in un certo senso, può essere considerato una riformulazione del Teorema di Fermat per gli estremi condizionati.

Teorema (metodo dei moltiplicatori di Lagrange per la ricerca dei massimi e minimi vincolati). *Sia $g: U \rightarrow \mathbb{R}^s$ un'applicazione di classe C^1 su un aperto U di \mathbb{R}^k . Supponiamo che g non abbia punti critici in $M = g^{-1}(0)$. Se $f: U \rightarrow \mathbb{R}$ è una funzione reale di classe C^1 su U e $p \in M$ è un punto estremante per f in M , allora esiste $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_s) \in \mathbb{R}^s$ con la proprietà che nel punto $(p, \lambda) \in U \times \mathbb{R}^s$ si annullano tutte le $k + s$ derivate parziali della funzione $F: U \times \mathbb{R}^s \rightarrow \mathbb{R}$ definita da $F(x, \lambda) = f(x) - \lambda \cdot g(x)$, dove x denota $(x_1, x_2, \dots, x_k) \in U$ e $\lambda \cdot g(x)$ è il prodotto scalare $\lambda_1 g_1(x) + \lambda_2 g_2(x) + \dots + \lambda_s g_s(x)$.*

Dimostrazione^{fac}. Osserviamo che le prime k equazioni che si ottengono uguagliando a zero le derivate parziali della funzione

$$F(x, \lambda) = f(x) - \lambda \cdot g(x),$$

ossia le derivate rispetto ad x_1, x_2, \dots, x_k , equivalgono ad affermare che il gradiente di f in p , $\nabla f(p)$, è combinazione lineare dei gradienti in p delle funzioni g_1, g_2, \dots, g_s componenti la g , mentre le ultime s ci dicono che p sta nel vincolo $g^{-1}(0)$. Poiché la condizione $p \in g^{-1}(0)$ è vera per ipotesi, è sufficiente provare che $\nabla f(p)$ è combinazione lineare di $\nabla g_1(p), \nabla g_2(p), \dots, \nabla g_s(p)$. Dato che p è un punto estremante, in base al Teorema di Fermat (per gli estremi vincolati), $\nabla f(p)$ appartiene allo spazio $T_p(M)^\perp$ ortogonale a $T_p(M)$, che ha ovviamente dimensione s , visto che $\dim T_p(M) = \dim M = k - s$. Essendo le funzioni g_i costanti su M , anche gli s vettori $\nabla g_i(p)$ appartengono a $T_p(M)^\perp$. Il risultato segue quindi se si mostra che tali vettori sono linearmente indipendenti (in questo caso, essendo esattamente s , generano tutto lo spazio $T_p(M)^\perp$). Per provare ciò basta osservare che $\nabla g_1(p), \nabla g_2(p), \dots, \nabla g_s(p)$ sono le s righe della matrice jacobiana, valutata in p , dell'applicazione g . Tale matrice, infatti, avendo per ipotesi rango s , ha le righe linearmente indipendenti. □

Per illustrare il metodo di Lagrange, supponiamo, ad esempio, $k = 3$ ed $s = 2$. Ossia supponiamo che si debbano trovare i punti estremanti di una funzione $f(x, y, z)$ con le condizioni $g_1(x, y, z) = 0$ e $g_2(x, y, z) = 0$, che definiscono il vincolo. Assumiamo che le funzioni f, g_1 e g_2 siano di classe C^1 (su un aperto di \mathbb{R}^3) e che la matrice jacobiana di $g = (g_1, g_2)$ abbia rango 2 in ogni punto (x, y, z) che verifica le suddette condizioni. La funzione di Lagrange in questo caso ha $3 + 2$ variabili ed è così definita:

$$F(x, y, z, \lambda_1, \lambda_2) = f(x, y, z) - \lambda_1 g_1(x, y, z) - \lambda_2 g_2(x, y, z).$$

Le due variabili λ_1 e λ_2 si chiamano *moltiplicatori di Lagrange*. Per trovare gli eventuali punti estremanti per la restrizione di f al vincolo, basta uguagliare a zero le cinque derivate parziali della F e risolvere il sistema (di cinque equazioni in cinque incognite) così ottenuto. Se $(p, \mu) = (x_0, y_0, z_0, \mu_1, \mu_2)$ è una soluzione, le prime tre equazioni ci dicono che $\nabla f(p) = \mu_1 \nabla g_1(p) + \mu_2 \nabla g_2(p)$ (e quindi $\nabla f(p) \in T_p(M)^\perp$), e le rimanenti due esprimono le equazioni del vincolo (e quindi $p \in g^{-1}(0)$).

Osserviamo che, col metodo di Lagrange, per risolvere un problema di massimi e minimi condizionati si opera come se si dovessero cercare dei massimi e minimi liberi. Pagando tuttavia un prezzo: da una funzione di k variabili si passa ad una di $k + s$ variabili.

È bene ricordarsi che, come per il Teorema di Fermat, il metodo di Lagrange dà soltanto una condizione necessaria affinché un punto sia estremante. In altre parole, quello che possiamo affermare è che se in un punto p del vincolo M il gradiente di f non è combinazione lineare dei gradienti delle funzioni g_i che definiscono il vincolo, allora p non può essere estremante. Fisicamente questo significa che se si interpreta f come il potenziale di una forza (conservativa), allora il vettore $\nabla f(p)$ rappresenta la forza che agisce sul punto (materiale) p vincolato alla varietà $g^{-1}(0)$, e se la sua componente parallela al vincolo non è nulla (ossia se $\nabla f(p)$ non è ortogonale a $T_p(M)$), allora p non può stare in equilibrio.

119 - Venerdì 14/05/10

Esercizio. Determinare, mediante il metodo dei moltiplicatori di Lagrange, il punto della retta $x + y = 0$ più vicino a $(0, 1)$.

Esercizio. Determinare, mediante il metodo dei moltiplicatori di Lagrange, il punto del piano $x + y - z = 0$ più vicino a $(0, 0, 2)$.

Esercizio. Mostrare che, tra tutti i rettangoli di area assegnata, il quadrato è quello di perimetro minimo.

Esercizio. Mostrare che, tra tutti i rettangoli di perimetro assegnato, il quadrato è quello di area massima.

Esercizio. Mostrare che, tra tutti i parallelepipedi di area assegnata, il cubo è quello di volume massimo.

Esercizio. Determinare, mediante il metodo dei moltiplicatori di Lagrange, il punto della retta

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

più vicino all'origine di \mathbb{R}^3 .

Esercizio. Tra tutti i rettangoli inscritti in un'ellisse, trovare quello di area massima (i lati del rettangolo si suppongono paralleli agli assi dell'ellisse).

Esercizio. Tra tutti i parallelepipedi inscritti in un ellissoide, determinare quello di volume massimo (gli spigoli si assumono paralleli agli assi dell'ellissoide).

Esercizio. Tra tutte le scatole senza coperchio di volume assegnato, trovare, col metodo dei moltiplicatori di Lagrange, quella di area minima.

Esercizio. Determinare l'insieme dei valori assunti dalla funzione $f(x, y, z) = 1 + 2z$ nella curva di livello

$$\begin{cases} x^2 + y^2 & = & 1 \\ x + y + z & = & 0 \end{cases}$$

120 - Mercoledì 19/05/10

Ricordiamo che un *vettore applicato* in \mathbb{R}^2 è una coppia $(p, v) \in \mathbb{R}^2 \times \mathbb{R}^2$. Il punto p si dice *punto di applicazione* di (p, v) e v è il *vettore libero*. Più in generale, un vettore applicato in \mathbb{R}^k è una coppia (p, v) di $\mathbb{R}^k \times \mathbb{R}^k$.

Un'espressione differenziale (reale di grado uno) in \mathbb{R}^2 è una funzione continua $\omega : X \rightarrow \mathbb{R}$ definita in un sottoinsieme (generalmente aperto) X di $\mathbb{R}^2 \times \mathbb{R}^2$. In altre parole, ω è una "legge" che ad ogni vettore applicato (p, v) di X associa un numero reale $\omega(p, v)$. In modo analogo si definisce il concetto di espressione differenziale in \mathbb{R}^3 (o, più in generale, in \mathbb{R}^k).

Chiaramente, le espressioni differenziali, come tutte le funzioni reali, si possono sommare, moltiplicare, dividere tra loro e, nell'ordine giusto, anche comporre con funzioni reali di variabile reale. Le convenzioni che si fanno sul dominio della somma, del quoziente, ecc., sono analoghe a quelle già viste per funzioni reali di variabile reale. Quindi, il dominio della somma o del prodotto di due espressioni differenziali è dato dall'intersezione dei domini delle due espressioni; il dominio di un quoziente è l'intersezione dei due domini meno i vettori (applicati) in cui si annulla il denominatore; il dominio di una composizione $f \circ \omega$ di un'espressione differenziale $\omega : X \rightarrow \mathbb{R}$ con una funzione reale di variabile reale $f : A \rightarrow \mathbb{R}$ è il sottoinsieme $\omega^{-1}(A) = \{(p, v) \in X : \omega(p, v) \in A\}$ di X .

Data un'espressione differenziale ω in \mathbb{R}^2 (o in \mathbb{R}^k), se si fissa un punto p di \mathbb{R}^2 (o di \mathbb{R}^k) si ottiene una funzione, denotata con ω_p , che dipende soltanto dal vettore libero v . In altre parole, fissato p , si ha $\omega_p(v) = \omega(p, v)$; cioè con ω_p si denota l'applicazione parziale che si ottiene fissando p e facendo variare soltanto v .

Esempi di espressioni differenziali:

- 1) una funzione continua $f : A \rightarrow \mathbb{R}$ definita su un sottoinsieme A di \mathbb{R}^2 (o di \mathbb{R}^k) può essere pensata come una particolare espressione differenziale dipendente soltanto dal punto di applicazione (si pone $f(p, v) := f(p)$);
- 2) l'incremento Δf di un'applicazione continua $f : A \rightarrow \mathbb{R}$ è l'espressione differenziale $\Delta f(p, v) := f(p + v) - f(p)$;
- 3) il differenziale df di una $f : U \rightarrow \mathbb{R}$ di classe C^1 su un aperto U di \mathbb{R}^2 (o di \mathbb{R}^k) è l'espressione differenziale che ad ogni $(p, v) \in U \times \mathbb{R}^2$ (o $(p, v) \in U \times \mathbb{R}^k$) associa il differenziale $df_p(v)$ di f in p relativo a v (ricordarsi che $df_p(v) = \nabla f(p) \cdot v$);
- 4) l'*elemento di lunghezza* ds è l'espressione differenziale che ad ogni vettore applicato (p, v) associa la lunghezza di v (ossia $ds(p, v) = \|v\|$);

5) il differenziale secondo d^2f di una $f:U \rightarrow \mathbb{R}$ di classe C^2 su un aperto U di \mathbb{R}^2 è l'espressione differenziale che ad ogni $(p, v) \in U \times \mathbb{R}^2$ associa il numero $d^2f(p, v)$ ottenuto calcolando per $t = 0$ la derivata seconda della funzione composta $\varphi(t) := f(p + tv)$.

Esercizio. Provare che il differenziale dx della funzione x , pensata da \mathbb{R}^2 (o da \mathbb{R}^3) in \mathbb{R} , è l'espressione differenziale che ad ogni vettore applicato (p, v) associa la prima componente del vettore libero v . In altre parole: se $p = (x_0, y_0)$ e $v = (h, k)$, si ha $dx(p, v) = h$.

Esercizio. Siano (in \mathbb{R}^3) $p = (1, -1, 2)$ e $v = (2, 1, -1)$. Calcolare i seguenti numeri: $dx(p, v)$, $dy(p, v)$, $dz(p, v)$, $dx^2(p, v)$, $ds^2(p, v)$, $(dx^2 + dy^2 + dz^2)(p, v)$.

Esercizio. Sia $f(x, y, z) := x^2 - 2yz$, e siano $p = (1, -1, 2)$ e $v = (2, 1, -1)$. Calcolare i seguenti numeri: $f(p, v)$, $df(p, v)$, $d^2f(p, v)$, $\Delta f(p, v)$, $(f + df + \frac{1}{2}d^2f)(p, v)$.

Esercizio. Data $f:\mathbb{R} \rightarrow \mathbb{R}$ di classe C^1 , mostrare che le espressioni differenziali df/dx e d^2f/dx^2 dipendono soltanto dal punto di applicazione (quindi possono essere considerate delle ordinarie funzioni reali di variabile reale).

Esercizio. Data $f:\mathbb{R}^3 \rightarrow \mathbb{R}$ di classe C^1 , provare la seguente uguaglianza tra espressioni differenziali:

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz.$$

Proviamo che (nel senso delle espressioni differenziali in \mathbb{R}^2) vale l'uguaglianza

$$ds = \sqrt{dx^2 + dy^2}.$$

Ossia, mostriamo che per ogni $(p, v) \in \mathbb{R}^2 \times \mathbb{R}^2$ risulta

$$ds(p, v) = \left(\sqrt{dx^2 + dy^2} \right) (p, v).$$

A tale scopo fissiamo un punto $p = (x_0, y_0)$ e un vettore libero $v = (h, k)$. Per definizione di ds si ha $ds(p, v) = \|v\|$. Mostriamo ora che se si applica l'espressione differenziale $\sqrt{dx^2 + dy^2}$ a (p, v) si ottiene ancora $\|v\|$. Ricordiamo infatti che il differenziale dx della funzione x (pensata in \mathbb{R}^2) associa ad ogni vettore applicato (p, v) la prima componente di v . Quindi, se $v = (h, k)$, si ha $dx(p, v) = h$. Analogamente $dy(p, v) = k$. Pertanto, in base alla nozione di somma, prodotto e composizione di espressioni differenziali, risulta

$$\left(\sqrt{dx^2 + dy^2} \right) (p, v) = \sqrt{(dx(p, v))^2 + (dy(p, v))^2} = \sqrt{h^2 + k^2} = \|v\|.$$

In modo analogo si prova che in \mathbb{R}^3 risulta $ds^2 = dx^2 + dy^2 + dz^2$ o, equivalentemente,

$$ds = \sqrt{dx^2 + dy^2 + dz^2}.$$

Supponiamo ora di voler esprimere l'elemento d'arco ds di \mathbb{R}^2 tramite le coordinate polari. Per far ciò basta calcolare le espressioni dx e dy in coordinate polari e sostituirle nell'uguaglianza $ds^2 = dx^2 + dy^2$. Ricordiamo che le coordinate cartesiane (x, y) sono legate

alle coordinate polari (ρ, θ) dalle seguenti relazioni: $x = \rho \cos \theta$ e $y = \rho \sin \theta$. Quindi, differenziando, si ottiene

$$dx = \cos \theta d\rho - \rho \sin \theta d\theta, \quad dy = \sin \theta d\rho + \rho \cos \theta d\theta.$$

Pertanto,

$$ds^2 = (\cos \theta d\rho - \rho \sin \theta d\theta)^2 + (\sin \theta d\rho + \rho \cos \theta d\theta)^2 = d\rho^2 + \rho^2 d\theta^2$$

o, equivalentemente, $ds = \sqrt{d\rho^2 + \rho^2 d\theta^2}$, che è l'espressione cercata.

Definizione (di forma differenziale). Un'espressione differenziale $\omega: U \times \mathbb{R}^2 \rightarrow \mathbb{R}$, dove U è un aperto di \mathbb{R}^2 , si dice una *forma differenziale* (di grado uno) su U se è lineare rispetto al vettore libero. Ossia se, fissato un qualunque $p \in U$, l'applicazione parziale $\omega_p: \mathbb{R}^2 \rightarrow \mathbb{R}$ definita da $\omega_p(v) = \omega(p, v)$ è lineare.

Il compito di definire la nozione di forma differenziale in \mathbb{R}^3 (o in \mathbb{R}^k) è lasciato per esercizio (non facoltativo) agli studenti.

Si osservi che, data una funzione f di classe C^1 su un aperto U di \mathbb{R}^2 , il suo differenziale df è una forma differenziale. Ricordiamo infatti che, fissato $p \in U$, df_p opera sui vettori liberi $v \in \mathbb{R}^2$ nel seguente modo:

$$df_p(v) = \nabla f(p) \cdot v.$$

Pertanto, per le note proprietà del prodotto scalare, df_p risulta una funzione lineare di v . In pratica il differenziale di f è un'espressione della forma

$$A(x, y)dx + B(x, y)dy,$$

dove le funzioni $A(x, y)$ e $B(x, y)$ sono le derivate parziali di f rispetto alla x e alla y , dx è il differenziale della prima funzione coordinata e dy della seconda. Si potrebbe dimostrare ^{fac} che, non solo i differenziali delle funzioni, ma addirittura ogni forma differenziale ω su un aperto U di \mathbb{R}^2 si può rappresentare nel seguente modo:

$$\omega = A(x, y)dx + B(x, y)dy,$$

dove $A(x, y)$ e $B(x, y)$ sono due funzioni definite in U . In altre parole, fissato $p \in U$, l'applicazione lineare $\omega_p: \mathbb{R}^2 \rightarrow \mathbb{R}$ è combinazione lineare di due sole applicazioni lineari: dx e dy (che costituiscono una base dello spazio delle applicazioni lineari da \mathbb{R}^2 in \mathbb{R}). I coefficienti della combinazione lineare dipendono dal punto $p = (x, y)$ fissato, e sono quindi due funzioni di (x, y) : $A(x, y)$ e $B(x, y)$.

Diremo che una forma differenziale

$$\omega = A(x, y)dx + B(x, y)dy$$

è di classe C^n (o C^∞) se sono di classe C^n (o C^∞) entrambe le funzioni $A(x, y)$ e $B(x, y)$.

Abbiamo già osservato che un'ordinaria funzione $f: A \rightarrow \mathbb{R}$ definita in un sottoinsieme A di \mathbb{R}^k può essere pensata come un'espressione differenziale che dipende soltanto dal punto di applicazione. Il dominio di tale espressione è costituito da tutti i vettori applicati $(p, v) \in A \times \mathbb{R}^k$ con punto di applicazione in A . Viceversa, un'espressione differenziale ω il cui dominio sia il prodotto cartesiano $A \times \mathbb{R}^k \subseteq \mathbb{R}^k \times \mathbb{R}^k$, nel caso dipenda soltanto dal punto di applicazione (cioè fissato un qualunque $p \in A$, $\omega(p, v)$ sia costante rispetto al vettore libero v) può essere vista come un'ordinaria funzione reale definita in A .

Non è difficile provare^{fac} che se $f: \mathbb{R} \rightarrow \mathbb{R}$ è una funzione di classe C^1 , allora il rapporto tra il suo differenziale df e il differenziale dx della funzione coordinata x dipende soltanto dal punto di applicazione. Pertanto, nello spirito della suddetta identificazione, tale rapporto rappresenta un'ordinaria funzione (reale di variabile reale). Precisamente, vale l'uguaglianza

$$\frac{df}{dx} = f'.$$

Esercizio^{fac}. Provare che se $f: \mathbb{R} \rightarrow \mathbb{R}$ è di classe C^n , allora risulta

$$\frac{d^n f}{dx^n} = f^{(n)}.$$

121 - Mercoledì 19/05/10

Ci poniamo la seguente domanda: data una forma differenziale

$$\omega = A(x, y)dx + B(x, y)dy,$$

si può affermare che questa è il differenziale di una funzione $f(x, y)$? Il risultato che segue mostra che la risposta è in generale negativa.

Teorema. *Sia $\omega = A(x, y)dx + B(x, y)dy$ una forma differenziale di classe C^1 su un aperto U di \mathbb{R}^2 . Se esiste una funzione $f: U \rightarrow \mathbb{R}$ tale che $df = \omega$ (ossia, tale che $\partial f/\partial x = A$ e $\partial f/\partial y = B$), allora $\partial A/\partial y = \partial B/\partial x$.*

Dimostrazione. Sia f una funzione tale che $\partial f/\partial x = A$ e $\partial f/\partial y = B$. Poiché A e B sono di classe C^1 , la funzione f risulta di classe C^2 . Di conseguenza, tenendo conto del Teorema di Schwarz, si ha

$$\frac{\partial A}{\partial y} = \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial^2 f}{\partial x \partial y} = \frac{\partial B}{\partial x},$$

e la tesi è dimostrata. □

Esempio. Consideriamo la forma differenziale $\omega = xdx - xydy$. Poiché $\partial x/\partial y = 0$ e $\partial(-xy)/\partial x = -y$, non esiste una funzione $f: \mathbb{R}^2 \rightarrow \mathbb{R}$ tale che $df = \omega$.

Definizione. Sia $\omega = A(x, y)dx + B(x, y)dy$ una forma differenziale definita su un aperto U di \mathbb{R}^2 . Si dice che ω è una *forma esatta* (in U) se esiste una funzione $f: U \rightarrow \mathbb{R}$, detta

primitiva di ω , tale che $df = \omega$. Si dice che ω è una *forma chiusa* (in U) se è di classe C^1 e $\partial A/\partial y = \partial B/\partial x$.

In base alla suddetta definizione, il precedente teorema può essere riformulato nel modo seguente: “*condizione necessaria affinché una forma differenziale di classe C^1 sia esatta è che sia chiusa*”.

Vedremo in seguito, dopo aver introdotto gli integrali curvilinei, che la condizione che una forma differenziale sia chiusa non ci assicura che sia anche esatta (a meno che non siano verificate delle opportune ipotesi sul suo dominio).

Esempio. Consideriamo la forma differenziale

$$\omega = (y^2 + \cos x)dx + (2xy + 1)dy.$$

Risulta $\partial(y^2 + \cos x)/\partial y = 2y$ e $\partial(2xy + 1)/\partial x = 2y$. Quindi ω è chiusa; ossia è soddisfatta la condizione necessaria affinché ω sia una forma esatta. Proviamo a vedere se effettivamente ω ammette una primitiva $f: \mathbb{R}^2 \rightarrow \mathbb{R}$. Se una tale f esiste, si deve avere $df = \omega$, ossia

$$\frac{\partial f}{\partial x} = y^2 + \cos x, \quad \frac{\partial f}{\partial y} = 2xy + 1.$$

Dalla prima uguaglianza si deduce che $f(x, y)$ è uguale a $xy^2 + \sin x$ più una costante rispetto alla variabile x (cioè più una funzione della sola y). Quindi $f(x, y) = xy^2 + \sin x + \varphi(y)$. Occorre determinare la funzione $\varphi(y)$. Derivando rispetto alla y l'espressione della f che abbiamo appena determinato, si ha $\partial f/\partial y = 2xy + \varphi'(y)$. Pertanto, tenendo conto che (se f è una primitiva) deve essere $\partial f/\partial y = 2xy + 1$, si ottiene $2xy + \varphi'(y) = 2xy + 1$. Quindi $\varphi'(y) = 1$ e, di conseguenza, $\varphi(y) = y + c$ (dove c è un'arbitraria costante). Si può concludere che se f è una primitiva, allora necessariamente

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

Un semplice controllo mostra che effettivamente $xy^2 + \sin x + y + c$ è una primitiva di ω .

Ricordiamo che se $f(x, y, z)$ è una funzione di classe C^1 su un aperto U di \mathbb{R}^3 , il suo differenziale è l'espressione

$$df = \frac{\partial f}{\partial x}dx + \frac{\partial f}{\partial y}dy + \frac{\partial f}{\partial z}dz$$

o, con notazioni vettoriali, l'espressione

$$df = \nabla f \cdot dp,$$

dove dp è il vettore incremento (di componenti dx , dy e dz).

In generale, un'espressione del tipo

$$\omega = A(x, y, z)dx + B(x, y, z)dy + C(x, y, z)dz,$$

dove A , B e C sono funzioni continue su un aperto U di \mathbb{R}^3 , rappresenta una *forma differenziale* in \mathbb{R}^3 . Come per le forme nel piano, ω è di classe C^n (o C^∞) se tali sono le sue tre *funzioni componenti*: A , B e C . Diremo che ω è una forma *esatta* se esiste una funzione f , detta *primitiva* di ω , tale che $df = \omega$. Diremo che ω è *chiusa* se sono verificate le seguenti tre condizioni:

$$\frac{\partial A}{\partial y} = \frac{\partial B}{\partial x}, \quad \frac{\partial B}{\partial z} = \frac{\partial C}{\partial y}, \quad \frac{\partial C}{\partial x} = \frac{\partial A}{\partial z}.$$

Come per le forme in \mathbb{R}^2 , dal Teorema di Schwarz discende che (anche in \mathbb{R}^3) *ogni forma esatta di classe C^1 è chiusa*. In generale, tuttavia, non è vero il contrario (lo vedremo con un esempio, dopo aver introdotto gli integrali curvilinei).

Esiste un perfetto parallelismo tra le forme differenziali e i campi vettoriali. Ci limitiamo ad un confronto in \mathbb{R}^3 . Ad ogni forma differenziale $\omega = A dx + B dy + C dz$, definita su un aperto U di \mathbb{R}^3 , associamo il campo vettoriale $\mathbf{w}: U \rightarrow \mathbb{R}^3$ con le stesse componenti di ω ; ossia, dato $p \in U$, $\mathbf{w}(p)$ è il vettore $A(p)\mathbf{i} + B(p)\mathbf{j} + C(p)\mathbf{k}$. In altre parole (questa volta indipendenti dal sistema di coordinate), fissato un punto $p \in U$, $\mathbf{w}(p)$ è quel vettore (è facile provare^{fac} che è unico) che gode della seguente proprietà: $\omega_p(v) = \mathbf{w}(p) \cdot v$ per ogni $v \in \mathbb{R}^3$. È ovvio che si ha una corrispondenza biunivoca tra forme differenziali e campi vettoriali. In tale corrispondenza i seguenti concetti sono accoppiati:

$$\begin{array}{ll} f \text{ è una primitiva di } \omega & \iff f \text{ è un potenziale di } \mathbf{w} \\ \omega \text{ è esatta} & \iff \mathbf{w} \text{ è conservativo} \\ \omega \text{ è chiusa} & \iff \mathbf{w} \text{ è irrotazionale} \end{array}$$

122 - Mercoledì 19/05/10

Una condizione che assicura che una forma chiusa sia anche esatta è che il suo dominio sia *semplicemente connesso*. La definizione formale di insieme semplicemente connesso richiede concetti topologici che vanno al di là degli scopi del corso. Ci limitiamo pertanto a darne un'idea intuitiva, corredandola poi con alcuni esempi esplicativi.

Definizione (euristica di insieme semplicemente connesso). Un sottoinsieme connesso U di \mathbb{R}^2 (o, più in generale, di \mathbb{R}^k) si dice *semplicemente connesso* se ogni curva chiusa contenuta in U può essere deformata con continuità riducendola ad un punto, senza che nella deformazione si tocchino punti del complementare di U (si pensi ad un elastico che si contrae, rimanendo sempre dentro U , fino a diventare un punto).

Ricordiamo che un sottoinsieme Q di \mathbb{R}^2 (o di \mathbb{R}^k) è *convesso* se presi due qualunque punti di Q , il segmento che li congiunge è contenuto in Q . Ad esempio, i cerchi, i triangoli e i rettangoli sono convessi di \mathbb{R}^2 , le sfere (piene) e i parallelepipedi sono convessi di \mathbb{R}^3 . Ovviamente, l'intero spazio \mathbb{R}^2 è convesso, così come è convesso un semipiano. Si potrebbe dimostrare (se si fosse data una definizione rigorosa di semplicemente connesso) che gli insiemi convessi sono anche semplicemente connessi.

Esempi di insiemi non semplicemente connessi si ottengono togliendo dal piano un punto, o un numero finito di punti o, addirittura, un arbitrario insieme limitato. Se, invece, dallo spazio \mathbb{R}^3 si toglie un punto (o un numero finito di punti), ciò che resta è ancora un insieme semplicemente connesso (si pensi ad un elastico che si contrae senza mai toccare i punti rimossi). Se da \mathbb{R}^3 si toglie una retta, o una circonferenza (o un numero finito di rette e circonferenze) ciò che rimane non è semplicemente connesso (si pensi ad un elastico che circonda una retta o che è concatenato con una circonferenza).

Teorema di Poincaré^{sd}. *Se una forma differenziale è chiusa ed è definita in un insieme semplicemente connesso, allora è esatta.*

Esercizio. Enunciare l’analogo del precedente teorema per i campi vettoriali.

Esercizio. Studiare la seguente forma differenziale:

$$\omega = yzdx + xzdy + (xy - 2z)dz.$$

Esercizio. Mostrare che il campo irrotazionale

$$\mathbf{w} = \frac{x}{x^2 + y^2} \mathbf{i} + \frac{y}{x^2 + y^2} \mathbf{j}$$

è conservativo, sebbene non sia definito in un insieme semplicemente connesso.

Suggerimento. Procedere come nell’esempio precedente, cercando f tale che $\nabla f = \mathbf{w}$.

Vogliamo definire il concetto di integrale curvilineo di un’espressione differenziale lungo una curva parametrica di classe C^1 . Allo scopo è conveniente introdurre alcune notazioni alternative per rappresentare l’integrale non orientato, cioè quello direttamente legato alla prima definizione di integrale (con le sommatorie).

Sia f una funzione integrabile in un intervallo non banale $[a, b] \subseteq \mathbb{R}$. Dato che $a < b$, in base alla definizione di integrale orientato, il numero

$$\int_{[a,b]} f(x) dx$$

può essere anche denotato col simbolo

$$\int_a^b f(x) dx.$$

Le seguenti sono altre due possibili notazioni:

$$\int_a^b f(x) |dx|, \quad \int_{[a,b]} f(x) |dx|,$$

da usare esclusivamente quando $a < b$. L’espressione differenziale $|dx|$ rappresenta l’elemento di lunghezza in \mathbb{R} . In una variabile infatti si ha

$$ds = \sqrt{dx^2} = |dx|.$$

Definizione (di integrale curvilineo di un’espressione differenziale). Sia $\gamma: [a, b] \rightarrow \mathbb{R}^k$ una curva parametrica di classe C^1 e sia ω un’espressione differenziale su un aperto U di \mathbb{R}^k contenente l’immagine di γ . Si chiama *integrale curvilineo* di ω lungo γ (o su γ) il numero

$$\int_{\gamma} \omega = \int_{[a,b]} \omega(\gamma(t), \gamma'(t)) dt.$$

Si fa notare che il suddetto integrale ha senso perché, essendo ω continua (per definizione di espressione differenziale) e γ di classe C^1 , la funzione reale di variabile reale $\omega(\gamma(t), \gamma'(t))$ è continua, e quindi integrabile nell’intervallo compatto $[a, b]$.

Con la suddetta definizione di integrale curvilineo le regole di calcolo risultano particolarmente naturali e facili da ricordare. Consideriamo alcuni esempi.

Definizione (di lunghezza di una curva parametrica). Data una curva $\gamma: [a, b] \rightarrow \mathbb{R}^k$ di classe C^1 , la sua *lunghezza* è il numero

$$L(\gamma) = \int_{\gamma} ds,$$

dove ds denota l’elemento di lunghezza in \mathbb{R}^k (ricordiamo che $ds(p, v) := \|v\|$).

Dal punto di vista fisico, pensando a γ come ad una traiettoria, $L(\gamma)$ rappresenta la “strada totale percorsa”, anche se alcuni tratti di strada possono essere ripetuti più volte.

Esempio. Dati due punti $p_0, p_1 \in \mathbb{R}^3$, calcoliamo la lunghezza della curva parametrica

$$\gamma(t) = p_0 + \frac{t}{T}(p_1 - p_0), \quad t \in [0, T].$$

Detta curva si può interpretare come il moto di un punto che parte da p_0 all’istante $t = 0$, si dirige verso p_1 con velocità (vettoriale) costante $\gamma'(t) = (p_1 - p_0)/T$ e raggiunge p_1 all’istante T . Dalla definizione si ottiene immediatamente

$$L(\gamma) = \int_{\gamma} ds = \int_0^T ds(\gamma(t), \gamma'(t)) dt = \int_0^T \|\gamma'(t)\| dt = \int_0^T (\|p_1 - p_0\|/T) dt = \|p_1 - p_0\|.$$

Esempio. Consideriamo la curva $\gamma(t)$ in \mathbb{R}^2 di equazioni parametriche $x = r \cos t$ e $y = r \sin t$, con $t \in [0, 2\pi]$. Questa rappresenta una circonferenza (parametrica) di raggio r , col centro nell’origine, percorsa una sola volta con velocità (scalare) costante. La sua lunghezza è data da

$$L(\gamma) = \int_{\gamma} ds = \int_0^{2\pi} \|\gamma'(t)\| dt = \int_0^{2\pi} \sqrt{x'(t)^2 + y'(t)^2} dt = \int_0^{2\pi} r dt = 2\pi r.$$

In modo equivalente possiamo calcolare la lunghezza di γ mediante l’uguaglianza $ds^2 = dx^2 + dy^2$. Risulta

$$L(\gamma) = \int_{\gamma} ds = \int_{\gamma} \sqrt{dx^2 + dy^2}.$$

Differenziando le equazioni parametriche di γ si ottiene $dx = -r \sin t dt$ e $dy = r \cos t dt$. Sostituendo le espressioni di dx e dy nella precedente uguaglianza si ha infine

$$L(\gamma) = \int_{\gamma} \sqrt{dx^2 + dy^2} = \int_0^{2\pi} \sqrt{r^2 dt^2} = \int_0^{2\pi} r |dt| = 2\pi r.$$

Supponiamo ora che la suddetta circonferenza sia data in coordinate polari (ρ, θ) . Ossia, consideriamo le equazioni parametriche $\rho = r$, $\theta = t$, con $t \in [0, 2\pi]$. In questo caso $d\rho = 0$ e $d\theta = dt$. Quindi, dall'uguaglianza $ds^2 = d\rho^2 + \rho^2 d\theta^2$ (che si consiglia di verificare per esercizio) si ottiene

$$L(\gamma) = \int_{\gamma} \sqrt{d\rho^2 + \rho^2 d\theta^2} = \int_0^{2\pi} \sqrt{r^2 dt^2} = \int_0^{2\pi} r |dt| = 2\pi r.$$

Esercizio. Provare che se una curva γ è costante, allora la sua lunghezza è nulla.

Integrale curvilineo di una forma differenziale. Sia $\omega = A(x, y)dx + B(x, y)dy$ una forma differenziale in \mathbb{R}^2 e sia $\gamma(t) = (x(t), y(t))$, $t \in [a, b]$, una curva parametrica di classe C^1 . In base alla definizione generale di integrale curvilineo di un'espressione differenziale si ha

$$\int_{\gamma} \omega = \int_{\gamma} A(x, y) dx + B(x, y) dy = \int_a^b \left(A(x(t), y(t)) x'(t) + B(x(t), y(t)) y'(t) \right) dt.$$

In pratica, per calcolare l'integrale curvilineo di una forma differenziale, basta integrare, nell'intervallo in cui varia il parametro, la funzione che si ottiene sostituendo $x(t)$ al posto di x e $y(t)$ al posto di y ; ricordandosi però che $dx(t) = x'(t)dt$ e $dy(t) = y'(t)dt$. Il calcolo dell'integrale curvilineo di una forma differenziale in \mathbb{R}^3 o, più in generale, in \mathbb{R}^k è analogo.

Per esempio, calcoliamo

$$\int_{\gamma} x dx + (y + 2) dy,$$

dove γ è la curva di equazioni parametriche $x = \cos \theta$, $y = \sin \theta$, $\theta \in [0, 2\pi]$. Differenziando le equazioni parametriche si ottiene $dx = -\sin \theta d\theta$ e $dy = \cos \theta d\theta$. Quindi, sostituendo si ottiene

$$\int_0^{2\pi} \left(-\cos \theta \sin \theta + (\sin \theta + 2) \cos \theta \right) d\theta = 2 \int_0^{2\pi} \cos \theta d\theta = 0.$$

Esercizio. Provare che se una curva $\gamma: [a, b] \rightarrow \mathbb{R}^2$ è costante, allora l'integrale lungo γ di una qualunque forma differenziale è nullo.

L'integrale curvilineo di un campo vettoriale $\mathbf{f}(p) = A(p)\mathbf{i} + B(p)\mathbf{j}$ lungo una curva γ si scrive nel seguente modo:

$$\int_{\gamma} \mathbf{f}(p) \cdot d\mathbf{p},$$

dove \mathbf{dp} sta per $dx \mathbf{i} + dy \mathbf{j}$. Si osservi che esso coincide con l'integrale lungo γ della forma differenziale associata $\omega = A(p) dx + B(p) dy$, dove p denota il generico punto (x, y) . Si ha infatti

$$\int_{\gamma} \mathbf{f}(p) \cdot \mathbf{dp} = \int_{\gamma} A(p) dx + B(p) dy = \int_{\gamma} A(x, y) dx + B(x, y) dy.$$

Quando il campo vettoriale \mathbf{f} rappresenta una forza, il significato di detto integrale è di capitale importanza per la Fisica: rappresenta il lavoro compiuto da \mathbf{f} lungo la curva γ .

123 - Venerdì 21/05/10

Teorema (formula fondamentale per gli integrali curvilinei). *Sia ω una forma differenziale esatta su un aperto U di \mathbb{R}^k e sia f una sua primitiva. Se $\gamma: [a, b] \rightarrow U$ è una curva C^1 a valori in U , allora risulta*

$$\int_{\gamma} \omega = f(p_2) - f(p_1),$$

dove p_1 e p_2 sono, rispettivamente, il primo ed il secondo estremo di γ . In particolare, l'integrale curvilineo di una forma differenziale esatta non dipende dal cammino, ma soltanto dagli estremi della curva.

Dimostrazione. Si ha

$$\int_{\gamma} \omega = \int_{\gamma} df := \int_a^b df(\gamma(t), \gamma'(t)) dt = \int_a^b f'(\gamma(t)) \gamma'(t) dt.$$

Il risultato segue immediatamente dalla formula fondamentale del calcolo integrale, tenendo conto che la funzione (reale di variabile reale) $\varphi(t) := f(\gamma(t))$ è una primitiva della funzione integranda $f'(\gamma(t)) \gamma'(t)$. □

Corollario. *Se ω è una forma differenziale esatta, allora l'integrale curvilineo di ω lungo una qualunque curva (parametrica) chiusa è nullo. Analogamente, se \mathbf{w} è un campo vettoriale conservativo, allora l'integrale curvilineo di \mathbf{w} lungo ogni curva chiusa è nullo.*

In sostanza, il suddetto corollario afferma che “una condizione necessaria affinché una forma differenziale sia esatta è che l'integrale lungo ogni curva chiusa sia zero”. Tale condizione è anche sufficiente. Vale infatti il seguente

Teorema^{sd}. *Condizione necessaria e sufficiente affinché una forma differenziale sia esatta è che l'integrale curvilineo della forma lungo una qualunque curva (parametrica) chiusa sia zero. Analogamente, un campo vettoriale \mathbf{w} è conservativo se e solo se l'integrale curvilineo di \mathbf{w} lungo ogni curva chiusa è nullo.*

Su alcuni testi di Fisica un campo di forze viene detto conservativo se il suo integrale curvilineo lungo ogni curva chiusa è nullo. Il suddetto teorema mostra che tale definizione è equivalente a quella precedentemente data (cioè un campo è conservativo se è il gradiente di una funzione).

Notazione. L'integrale curvilineo di un'espressione differenziale ω esteso ad una curva chiusa γ viene spesso denotato con

$$\oint_{\gamma} \omega.$$

Ricordiamo che, come conseguenza del Teorema di Schwarz, ogni forma differenziale esatta (di classe C^1) è anche chiusa. Inoltre, se una forma chiusa è definita in un aperto semplicemente connesso, allora è anche esatta. Mostriamo con un esempio che negli aperti non semplicemente connessi possono esistere forme chiuse che non sono esatte. Consideriamo la forma differenziale

$$\omega = A(x, y) dx + B(x, y) dy = -\frac{y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy.$$

Si ha

$$\frac{\partial A}{\partial y}(x, y) = \frac{y^2 - x^2}{(x^2 + y^2)^2} \quad \text{e} \quad \frac{\partial B}{\partial x}(x, y) = \frac{y^2 - x^2}{(x^2 + y^2)^2}.$$

Pertanto ω è una forma chiusa. Osserviamo che il dominio di ω è l'aperto $U = \mathbb{R}^2 \setminus \{(0, 0)\}$, che non è semplicemente connesso. Ciò non implica che la forma debba essere non esatta ("il dominio semplicemente connesso" è soltanto una condizione sufficiente affinché una forma chiusa sia esatta). Tuttavia, se fosse esatta, il suo integrale lungo una qualunque curva chiusa (con sostegno in U) dovrebbe essere zero. Proviamo ad integrarla lungo la circonferenza $\gamma_r(t)$ definita dalle seguenti equazioni parametriche:

$$x = r \cos t, \quad y = r \sin t, \quad 0 \leq t \leq 2\pi.$$

Si ha

$$\begin{aligned} & \oint_{\gamma_r} \left(-\frac{y}{x^2 + y^2} dx + \frac{x}{x^2 + y^2} dy \right) \\ &= \int_0^{2\pi} \left(-\frac{r \sin t}{r^2} d(r \cos t) + \frac{r \cos t}{r^2} d(r \sin t) \right) \\ &= \int_0^{2\pi} (\sin^2 t + \cos^2 t) dt = \int_0^{2\pi} dt = 2\pi. \end{aligned}$$

La forma differenziale non è pertanto esatta.

Per inciso osserviamo che il suddetto integrale non dipende dal raggio r della circonferenza $\gamma_r(t)$ considerata. Questo fatto ha una spiegazione teorica di cui diamo soltanto un'idea intuitiva. La dimostrazione rigorosa compete ad una moderna disciplina matematica, la Topologia, ed è basata sul seguente risultato (da cui si può dedurre il Teorema di Poincaré per le forme differenziali sugli aperti semplicemente connessi, che abbiamo già visto):

Lemma di Poincaré^{sd}. *Se una forma differenziale chiusa è definita in un insieme convesso, allora è esatta.*

Vediamo una spiegazione intuitiva di come dal Lemma di Poincaré si possa dedurre che l'integrale di una forma chiusa non muta se la curva viene "deformata con continuità".

Supponiamo che ω sia una forma differenziale chiusa su un aperto U di \mathbb{R}^2 . Consideriamo l'integrale curvilineo di ω esteso ad una curva chiusa γ (con sostegno in U). Supponiamo di deformare (in una zona) la curva γ , trasformandola in una nuova curva γ_1 , in modo che la deformazione avvenga dentro un cerchio aperto C interamente contenuto in U . Se si considera la differenza dei due integrali curvilinei (quello esteso a γ meno quello esteso a γ_1), il risultato è come se si facesse un integrale curvilineo esteso ad una curva chiusa interamente contenuta nel cerchio C . Si osservi infatti che le due curve coincidono fuori da C , e quindi, facendo la differenza dei due integrali, il contributo dei tratti di curva che stanno fuori da C è nullo. D'altra parte, il cerchio C è un insieme convesso; dunque, essendo ω una forma chiusa, la sua restrizione a C è una forma esatta. Ciò prova che la differenza dei due integrali curvilinei, essendo equivalente ad un integrale lungo una curva chiusa contenuta in C , è zero. Possiamo concludere che se ω è una forma chiusa e γ è una curva chiusa, l'integrale curvilineo di ω esteso a γ non muta se si deforma γ in un piccolo tratto. È un fatto intuitivo, e dimostrabile rigorosamente, che se due curve chiuse, entrambe con sostegno in un aperto U di \mathbb{R}^2 , differiscono di poco (non solo in un piccolo tratto), allora è possibile deformare una nell'altra con un numero finito di piccole deformazioni in modo che ciascuna di queste avvenga dentro un cerchio contenuto in U . Da ciò si deduce che se due curve differiscono di poco, l'integrale di una forma chiusa esteso a una curva o all'altra è lo stesso.

Immaginiamo ora di avere (in un aperto U di \mathbb{R}^2) una famiglia di curve chiuse $\gamma_\lambda(t)$ che dipendono con continuità da un parametro λ che varia in un intervallo (si pensi, ad esempio, alla famiglia di circonferenze $\gamma_r(t)$ considerate prima: in tal caso il parametro che distingue una curva da un'altra è il raggio r). Per la dipendenza continua da λ , nel passare da una curva ad un'altra della famiglia, si può dare al parametro una sequenza di valori in modo da ottenere delle curve intermedie con la proprietà che due qualunque curve consecutive siano sufficientemente vicine tra loro. Per quanto visto prima, se ω è una forma chiusa in U , passando da una curva alla curva successiva della sequenza, l'integrale curvilineo non cambia. Ciò prova, almeno intuitivamente, che l'integrale curvilineo di ω lungo una qualunque curva della famiglia non dipende dalla curva considerata.

Definizione (intuitiva di curve omotope). Sia U un aperto di \mathbb{R}^2 (o di \mathbb{R}^k) e siano γ_1 e γ_2 due curve chiuse con sostegno in U . Se γ_1 e γ_2 sono deformabili con continuità l'una nell'altra in modo che nella deformazione tutte le curve intermedie abbiano sostegno contenuto in U , allora si dice che γ_1 e γ_2 sono *omotope in U* .

Le suddette “chiacchiere” si concretizzano nel seguente risultato:

Teorema^{sd} (di invarianza per omotopia). Se ω è una forma chiusa in un aperto U di \mathbb{R}^k e se γ_1 e γ_2 sono due curve chiuse omotope in U , allora

$$\oint_{\gamma_1} \omega = \oint_{\gamma_2} \omega.$$

In particolare, se una curva chiusa γ è omotopa (in U) ad una curva costante, risulta

$$\oint_{\gamma} \omega = 0.$$

Quindi, se U è semplicemente connesso l'integrale lungo ogni curva chiusa è nullo e, di conseguenza, ω è una forma esatta.

124 - Venerdì 21/05/10

L'integrale curvilineo di un'espressione differenziale del tipo $f(p)ds$ si dice anche *integrale non orientato*. Il motivo intuitivo è dovuto al fatto che tale integrale non dipende dal verso di percorrenza della curva di integrazione (una spiegazione precisa di tale affermazione richiederebbe il concetto di parametrizzazioni equivalenti che, per il momento, omettiamo).

Per gli integrali non orientati vale la seguente **proprietà di monotonia** (immediata conseguenza dell'analoga proprietà dell'integrale di Cauchy-Riemann): Sia $\gamma: [a, b] \rightarrow \mathbb{R}^k$ una curva parametrica di classe C^1 e siano $f(p)$ e $g(p)$ due funzioni continue in un aperto U di \mathbb{R}^k contenente il sostegno $\text{Im } \gamma$ di γ . Se $f(p) \leq g(p)$ per ogni p nel sostegno di γ , allora

$$\int_{\gamma} f(p) ds \leq \int_{\gamma} g(p) ds.$$

Teorema (della media per gli integrali curvilinei). Sia $\gamma: [a, b] \rightarrow \mathbb{R}^k$ un arco di curva (parametrica di classe C^1) e sia $f(p)$ una funzione continua in un aperto U di \mathbb{R}^k contenente il sostegno di γ . Allora la media di f lungo γ , ossia il rapporto

$$\frac{1}{L(\gamma)} \int_{\gamma} f(p) ds$$

tra l'integrale curvilineo lungo γ di $f(p)$ e la lunghezza di γ , è un numero compreso tra l'estremo inferiore e l'estremo superiore di f . Quindi, essendo f continua, esiste un punto $c \in \text{Im } \gamma$ per il quale si ha

$$\int_{\gamma} f(p) ds = f(c)L(\gamma).$$

Dimostrazione. Denotiamo, rispettivamente, con m e M l'estremo inferiore e l'estremo superiore della funzione $f(p)$ per $p \in \text{Im } \gamma$. Si ha

$$m \leq f(p) \leq M, \quad \forall p \in \text{Im } \gamma.$$

Quindi, per la proprietà di monotonia, risulta

$$\int_{\gamma} m ds \leq \int_{\gamma} f(p) ds \leq \int_{\gamma} M ds.$$

Dividendo i tre membri della suddetta disuguaglianza per la lunghezza

$$L(\gamma) = \int_{\gamma} ds$$

della curva si ottiene la tesi. □

Definizione (di curva parametrica semplice). Una curva parametrica $\gamma: [a, b] \rightarrow \mathbb{R}^k$ si dice *semplice* se esistono al più due punti con la stessa immagine, e quando ciò accade tali

punti sono soltanto gli estremi a e b dell'intervallo di definizione (in tal caso, ricordiamo, la curva si dice *chiusa*).

Definizione (di curva parametrica regolare). Una curva parametrica $\gamma: [a, b] \rightarrow \mathbb{R}^k$ si dice *regolare* se è di classe C^1 e la sua derivata $\gamma'(t)$ non si annulla mai (significa che le derivate delle sue funzioni componenti non si annullano mai simultaneamente).

Definizione (di arco di curva regolare). Un sottoinsieme C di \mathbb{R}^k si dice un *arco* (di curva regolare) se è il sostegno (cioè l'immagine) di una curva parametrica semplice e regolare. Una qualunque curva parametrica semplice e regolare il cui sostegno sia un arco di curva regolare C si dice una *parametrizzazione* di C .

Si potrebbe dimostrare, ma non lo facciamo, che se un arco di curva regolare C ammette una parametrizzazione chiusa $\gamma: [a, b] \rightarrow \mathbb{R}^k$ (cioè tale che $\gamma(a) = \gamma(b)$), allora ogni altra parametrizzazione di C è chiusa. In tal caso si dice che C è un *arco di curva chiusa*.

Teorema^{sd} (di indipendenza dalla parametrizzazione per gli integrali curvilinei non orientati). Se γ_1 e γ_2 sono due parametrizzazioni di uno stesso arco di curva regolare C , allora, data una qualunque funzione continua $f(p)$ definita su C , risulta

$$\int_{\gamma_1} f(p) ds = \int_{\gamma_2} f(p) ds.$$

Il suddetto risultato giustifica la seguente

Definizione. Sia C un arco di curva regolare e sia $f(p)$ una funzione continua definita su C . Si definisce

$$\int_C f(p) ds = \int_{\gamma} f(p) ds,$$

dove γ è una qualunque parametrizzazione di C . In particolare, se $f(p) \equiv 1$, si ottiene un numero che dipende soltanto da C , denotato $L(C)$ e detto *lunghezza di C* .

Definizione (di centro di massa di una curva). Dato un arco di curva regolare C in \mathbb{R}^2 , il suo *centro di massa geometrico* è quel punto (x_0, y_0) che ha per ascissa la media della funzione ascissa e per ordinata la media della funzione ordinata. In simboli:

$$x_0 = \frac{1}{L(C)} \int_C x ds, \quad y_0 = \frac{1}{L(C)} \int_C y ds.$$

Si osservi che, come conseguenza del teorema della media per gli integrali curvilinei, se la curva è contenuta in un rettangolo, allora anche il suo centro di massa deve stare in detto rettangolo.

Determiniamo, per esempio, il centro di massa (geometrico) della semicirconferenza C definita dalla parametrizzazione $x = r \cos t$ e $y = r \sin t$, $0 \leq t \leq \pi$. Per ragioni di simmetria risulta $x_0 = 0$. Occorre quindi calcolare soltanto l'ordinata y_0 del centro di massa. La lunghezza $L(\gamma)$ della curva è πr e quindi

$$y_0 = \frac{1}{\pi r} \int_{\gamma} y ds = \frac{1}{\pi r} \int_{\gamma} y \sqrt{dx^2 + dy^2} = \frac{r^2}{\pi r} \int_0^{\pi} \sin t |dt| = \frac{2r}{\pi}.$$

Si osservi che $2r/\pi$ è un numero tra 0 ed r (in accordo col teorema della media).

Teorema di Pappo^{sd} (per le superfici di rotazione). *Sia $C \subseteq \mathbb{R}^2$ un arco di curva regolare interamente contenuto in un semipiano delimitato da una retta α . L'area della superficie che si ottiene ruotando C di un angolo 2π intorno alla retta α è uguale al prodotto della lunghezza di C per la lunghezza della circonferenza percorsa dal centro di massa di C .*

Come applicazione del Teorema di Pappo calcoliamo l'area di una superficie sferica di raggio r . Tale superficie si può ottenere (ad esempio) facendo fare un giro completo intorno all'asse x alla semicirconferenza C di raggio r di cui abbiamo appena determinato il centro di massa. Poiché la distanza y_0 del centro di massa di C dall'asse delle x è $2r/\pi$, tale punto, nella rotazione, descrive una circonferenza di lunghezza $(2\pi)(2r/\pi) = 4r$. Dato che la semicirconferenza C ha lunghezza πr , l'area della sfera risulta $(\pi r)(4r) = 4\pi r^2$.

Il *momento d'inerzia* rispetto ad un punto $p_0 \in \mathbb{R}^2$ di un filo omogeneo C di peso m , sostegno di una curva semplice e regolare in \mathbb{R}^2 , è il numero

$$I = \int_C d(p, p_0)^2 \delta ds,$$

dove $d(p, p_0)$ è la funzione distanza di un generico punto p dal punto di riferimento p_0 e $\delta = m/L(C)$ è la densità lineare del filo.

Se il filo non è omogeneo, il suddetto integrale dà ancora il momento d'inerzia del filo, ma in tal caso la densità è una funzione $\delta(p)$ del generico punto $p \in C$. Talvolta l'espressione differenziale $\delta(p) ds$ si denota col simbolo dm , detto *elemento di massa*.

Analogamente, il *momento d'inerzia* rispetto ad una retta $\alpha \subseteq \mathbb{R}^2$ di un filo C (non necessariamente omogeneo) è il numero

$$I = \int_C d(p, \alpha)^2 dm,$$

dove $d(p, \alpha)$ è la funzione distanza di un generico punto p dalla retta di riferimento α .

Calcoliamo, ad esempio, il momento d'inerzia di una circonferenza omogenea di raggio r e massa m rispetto al suo centro. Tutti i punti della circonferenza hanno distanza r dal centro, e quindi il contributo di un elemento di massa dm è $dI = r^2 dm$. Se la nostra intuizione non sbaglia, sommando i vari contributi, il momento d'inerzia dovrebbe essere $I = mr^2$. Vediamo se è vero. Poniamo la circonferenza col centro nell'origine $0 \in \mathbb{R}^2$ e parametrizziamola nel modo seguente: $x = r \cos \theta$, $y = r \sin \theta$, $0 \leq \theta \leq 2\pi$. Si ha

$$I = \int_{\gamma} d(p, 0)^2 dm = \delta \int_{\gamma} (x^2 + y^2) ds = \frac{m}{2\pi r} \int_0^{2\pi} (r^2) r d\theta = mr^2.$$

Esercizio. Calcolare il momento d'inerzia rispetto al centro di massa di un'asta omogenea di lunghezza l e massa m .

Esercizio. Dato un filo omogeneo, semicircolare, di raggio r e massa m , determinarne il momento d'inerzia rispetto al diametro congiungente i due estremi.

Definizione (di arco orientato). Un arco (di curva regolare) C si dice *orientato* se si è scelto uno dei due sensi di percorrenza. Se l'arco non è chiuso (ossia, è il sostegno di una curva semplice e regolare con estremi distinti), ciò equivale ad aver deciso quale dei due estremi è il primo e quale il secondo.

Si osservi che ogni parametrizzazione di un arco orientato percorre l'arco in modo *concorde* o *discorde* con l'orientazione scelta, e che due parametrizzazioni di uno stesso arco (non necessariamente orientato) possono essere *tra loro concordi* o *discordi*.

Le suddette nozioni (di arco orientato, di parametrizzazione concorde con l'orientazione e di parametrizzazioni tra loro concordi) dovrebbero risultare sufficientemente chiare e intuitive ad ogni studente di media preparazione. Per renderle precise (in modo che si possano adoperare nelle dimostrazioni) occorrerebbe il concetto di *cambiamento di parametro*, che preferiamo non introdurre per non appesantire troppo il carico didattico. Gli studenti desiderosi di comprendere più a fondo i concetti introdotti non hanno che da chiedere ulteriori spiegazioni al docente.

Concludiamo le disquisizioni sugli integrali curvilinei con il seguente importante risultato, di cui omettiamo la dimostrazione:

Teorema^{sd} (di indipendenza dalla parametrizzazione per gli integrali curvilinei orientati). Sia C un arco di curva regolare (non necessariamente orientato) e siano γ_1 e γ_2 due parametrizzazioni di C . Allora, data una forma differenziale ω su C , risulta

$$\int_{\gamma_1} \omega = \pm \int_{\gamma_2} \omega,$$

a seconda che le due parametrizzazioni siano tra loro concordi o discordi. Analogamente, dato un campo vettoriale \mathbf{w} su C , si ha

$$\int_{\gamma_1} \mathbf{w} \cdot d\mathbf{p} = \pm \int_{\gamma_2} \mathbf{w} \cdot d\mathbf{p},$$

a seconda che γ_1 e γ_2 siano tra loro concordi o discordi.

Il precedente teorema permette di definire il concetto di integrale di una forma differenziale (o di un campo vettoriale) su un arco orientato C senza bisogno che questo sia *a priori* parametrizzato (la definizione è lasciata per esercizio agli studenti). In pratica, per calcolare l'integrale su C basta scegliere una qualunque parametrizzazione γ di C ; se questa è concorde con l'orientazione, allora l'integrale sull'arco coincide con l'integrale sulla curva parametrica γ , altrimenti basterà cambiare di segno all'integrale su γ .

Esercizio. Studiare le seguenti forme differenziali:

$$\frac{x dx + y dy + z dz}{(x^2 + y^2 + z^2)^{3/2}}; \quad \frac{x^2 - y^2}{(x^2 + y^2)^{1/2}} dx - \frac{2xy}{(x^2 + y^2)^{1/2}} dy;$$

$$x dx + y^2 dy + z dz; \quad (x + y) dx - y^2 dy; \quad \frac{y + x^2 + y^2}{x^2 + y^2} dx - \frac{x}{x^2 + y^2} dy.$$

Esercizio. Calcolare i seguenti integrali curvilinei di forme differenziali (con orientazione antioraria):

$$\oint_{\gamma} (x^2 + y) dx + xy dy, \quad \gamma = \{(x, y) : x^2 + y^2 - 4x + 3 = 0\};$$

$$\oint_{\gamma} \frac{dx + dy}{|x| + |y|}, \quad \gamma = \{(x, y) : |x| + |y| = 1\}.$$

Esercizio. Calcolare

$$\int_{\gamma} \frac{x}{y} dx + \frac{x+2}{y^2} dy,$$

dove γ è il segmento di primo estremo $(1, 1)$ e secondo estremo $(2, 3)$.

Esercizio. Verificare che la forma differenziale

$$\omega = \left(2xy + \frac{x}{x^2 + y^2}\right) dx + \left(x^2 + \frac{y}{x^2 + y^2}\right) dy$$

è esatta e determinarne una primitiva.

Esercizio. Calcolare

$$\int_{\gamma} (1 + ye^{xy}) dx + (xe^{xy} + \cos y) dy$$

lungo una qualunque curva di estremi $(-\pi, -\pi/2)$ e $(\pi/2, \pi)$.

Esercizio. Calcolare il lavoro compiuto dal campo di forze

$$\mathbf{f} = (x^2 - y^2) \mathbf{i} + 2xy \mathbf{j}$$

su una particella che percorre (una sola volta) in senso antiorario il quadrato di lato a e di vertici $(0, 0)$, $(a, 0)$, (a, a) e $(0, a)$.

Esercizio. Determinare il centro di massa (geometrico) della curva γ definita dalle seguenti equazioni parametriche: $x = r(t - \sin t)$, $y = r(1 - \cos t)$, $t \in [0, 2\pi]$.

125 - Mercoledì 26/05/10

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, $p_1 = \{x_0, x_1, \dots, x_n\}$ di $[a, b]$ e $p_2 = \{y_0, y_1, \dots, y_m\}$ di $[c, d]$, il rettangolo R viene suddiviso in nm sottorettangoli

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

di area $\mu(R_{ij}) = (\Delta x)_i (\Delta y)_j$, dove $(\Delta x)_i = x_i - x_{i-1}$ e $(\Delta y)_j = y_j - y_{j-1}$ denotano, rispettivamente, l' i -esimo incremento della x e lo j -esimo (si pronuncia "iotesimo") incremento della y definiti dalla partizione $p = (p_1, p_2)$. In ogni sottorettangolo R_{ij} scegliamo

un punto $q_{ij} = (x_{ij}, y_{ij})$. L'insieme s dei punti q_{ij} si dice una *scelta di punti* nella partizione $p = (p_1, p_2)$ di R . Ogni rettangolo R_{ij} della partizione col punto q_{ij} scelto si dice un *rettangolo puntato*. La coppia $\alpha = (p, s)$, costituita dalla partizione $p = (p_1, p_2)$ 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 $\alpha = (p, s)$ di $R = [a, b] \times [c, d]$ possiamo associare il numero

$$S_f(\alpha) = \sum_{\substack{i=1, \dots, n \\ j=1, \dots, m}} f(q_{ij}) \mu(R_{ij}) = \sum_{i=1}^n \sum_{j=1}^m f(x_{ij}, y_{ij}) (\Delta x)_i (\Delta y)_j,$$

dove, ricordiamo, $(\Delta x)_i (\Delta y)_j = \mu(R_{ij})$ denota l'area del generico sottorettangolo R_{ij} individuato dalla partizione (la lettera μ , cioè la m greca, significa *misura*, e in \mathbb{R}^2 si chiama area o misura bidimensionale) e $(x_{ij}, y_{ij}) = q_{ij}$ è il punto scelto in R_{ij} . Si ha così una funzione reale $S_f: \mathcal{P} \rightarrow \mathbb{R}$ definita nell'insieme \mathcal{P} delle partizioni puntate del rettangolo R .

Per meglio comprendere il significato del numero $S_f(\alpha)$ individuato dalla partizione puntata α , è bene osservare che se la funzione f è positiva, ogni termine $f(q_{ij}) \mu(R_{ij})$ della sommatoria rappresenta il volume di un parallelepipedo di altezza $f(q_{ij})$ 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 $S_f(\alpha)$ 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 (secondo Cauchy-Riemann) in R della funzione $f(x, y)$ 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 $I \in \mathbb{R}$ è l'*integrale doppio di f in R* se, fissato un arbitrario "errore" $\varepsilon > 0$, esiste un $\delta > 0$ tale che, comunque si assegni una partizione puntata α con parametro di finezza $|\alpha|$ minore di δ , la distanza $|S_f(\alpha) - I|$ tra la somma $S_f(\alpha)$ e il numero I è minore di ε . Se ciò accade, si scrive

$$\lim_{|\alpha| \rightarrow 0} S_f(\alpha) = I$$

e la funzione f si dice *integrabile* in R (secondo Cauchy-Riemann).

L'integrale doppio di una funzione $f(x, y)$ in un rettangolo R si denota con uno dei seguenti simboli:

$$\int_R f, \quad \int_R f \, d\mu, \quad \int_R f(p) \, d\mu, \quad \int_R f(x, y) \, dx dy,$$

$$\iint_R f, \quad \iint_R f \, d\mu, \quad \iint_R f(p) \, d\mu, \quad \iint_R f(x, y) \, dx dy.$$

Quest'ultima notazione è dovuta a Leibniz. Ecco come si giustifica:

$$\sum_{i=1}^n \sum_{j=1}^m f(x_{ij}, y_{ij})(\Delta x)_i(\Delta y)_j \longrightarrow \iint_R f(x, y) \, dx dy.$$

Si osservi che il numero

$$I = \iint_R f(x, y) \, dx dy$$

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 $S_f(\alpha)$ per $|\alpha| \rightarrow 0$ non cambia.

Esercizio. Dedurre, dalla definizione di integrale doppio, che, dato $k \in \mathbb{R}$, la funzione costante $f(x, y) \equiv k$ è integrabile in ogni rettangolo R e risulta

$$\iint_R k \, dx dy = k\mu(R).$$

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 α e β sono due numeri, allora anche la funzione $\alpha f + \beta g$ è integrabile e si ha

$$\int_R (\alpha f + \beta g) \, d\mu = \alpha \int_R f \, d\mu + \beta \int_R g \, d\mu,$$

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

$$\int_R f \, d\mu \geq 0,$$

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

$$\int_R f \, d\mu \leq \int_R g \, d\mu.$$

126 - Mercoledì 26/05/10

Ricordiamo che un insieme si dice *numerabile* se ha la stessa cardinalità dei numeri naturali (cioè se può essere messo in corrispondenza biunivoca con \mathbb{N}). È noto che l'insieme dei razionali è numerabile, ma non lo è l'insieme dei reali.

Definizione. Un sottoinsieme C di \mathbb{R}^2 si dice *trascurabile* (in \mathbb{R}^2), o di *misura (bidimensionale) nulla* secondo Lebesgue (si legge “lebeg”), se per ogni $\varepsilon > 0$ esiste una famiglia contabile (cioè finita o numerabile) di rettangoli che copre C (ossia, la cui unione contiene C) ed ha 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 ($y = \varphi(x)$ o $x = \psi(y)$) di una funzione continua (definita 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.

Teorema di integrabilità^{sd}. *Una funzione $f(x, y)$ è integrabile in un rettangolo R se e solo se (in detto rettangolo) 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^{sd}. *Siano $f(x, y)$ e $g(x, y)$ due funzioni integrabili in un rettangolo R . Se dette funzioni 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 detti 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” si ottiene combinando tra loro le note funzioni elementari con operazioni di somma, prodotto, quoziente, composizione, restrizione ad un intervallo e inversione, ed ogni funzione elementare, se non è continua, ha al più un insieme trascurabile di punti di discontinuità. Quindi, nella pratica, il compito di verificare se una funzione è integrabile (in un rettangolo) si riduce a controllare se (in detto rettangolo) è limitata (cioè, se esiste una costante che la maggia in valore assoluto).

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 detto rettangolo). 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 che importa: un punto è trascurabile).

Esercizio^{fac.}. Determinare il dominio della funzione

$$F(\lambda) = \iint_{R(\lambda)} \frac{1}{x^2 + y^2} dx dy,$$

dove $R(\lambda)$ è il quadrato $[\lambda - 1, \lambda + 1] \times [\lambda - 1, \lambda + 1]$.

Suggerimento. Trovare l'insieme dei numeri $\lambda \in \mathbb{R}$ per i quali il suddetto integrale ha senso (cioè rappresenta un numero). Per esempio, $F(0)$ è un numero reale ben definito? Cosa si può dire riguardo a $F(2)$? Ha senso?

Il risultato che segue riconduce il calcolo di un integrale doppio (in un rettangolo) a due successive integrazioni semplici.

Teorema di Fubini^{sd} (per gli integrali doppi). *Sia $f(x, y)$ una funzione reale definita in un rettangolo $R = [a, b] \times [c, d]$. Allora, quando ha senso, risulta*

$$\iint_R f(x, y) dx dy = \int_c^d \left(\int_a^b f(x, y) dx \right) dy = \int_a^b \left(\int_c^d f(x, y) dy \right) dx.$$

In sostanza, la prima uguaglianza del Teorema di Fubini afferma che (quando è possibile) per calcolare l'integrale doppio di $f(x, y)$ in $[a, b] \times [c, d]$ si integra prima in $[a, b]$ la funzione $f(x, y)$ rispetto alla variabile x , ottenendo così una funzione

$$g(y) := \int_a^b f(x, y) dx,$$

e poi si integra $g(y)$ nell'intervallo $[c, d]$. Ovviamente occorre che tali operazioni abbiano senso; cioè che per ogni $y \in [c, d]$ la funzione parziale $x \mapsto f(x, y)$ sia integrabile (in $[a, b]$) e che la funzione $g: [c, d] \rightarrow \mathbb{R}$ che si ottiene dopo aver eseguito la prima integrazione sia a sua volta integrabile. Si invitano gli studenti ad interpretare la seconda uguaglianza.

In modo equivalente, quando ha senso, si può prima integrare rispetto alla variabile y , ottenendo una funzione della sola x , e integrare poi rispetto alla x .

Per convenzione un'espressione del tipo $\varphi(x) dx$ si può scrivere anche $dx \varphi(x)$. Tenendo conto di ciò, la tesi del Teorema di Fubini si può esprimere nel modo seguente:

$$\iint_R f(x, y) dx dy = \int_c^d dy \int_a^b f(x, y) dx = \int_a^b dx \int_c^d f(x, y) dy.$$

Ecco un modo (esclusivamente) euristico per giustificare il Teorema di Fubini:

$$\sum_{i=1}^n \sum_{j=1}^m f(x_{ij}, y_{ij})(\Delta x)_i (\Delta y)_j = \sum_{i=1}^n (\Delta x)_i \sum_{j=1}^m f(x_{ij}, y_{ij})(\Delta y)_j \longrightarrow \int_a^b dx \int_c^d f(x, y) dy.$$

Esercizio. Calcolare

$$\iint_R xy dx dy,$$

dove $R = [0, 1] \times [-1, 2]$.

127 - Mercoledì 26/05/10

Definizione (di estensione standard). Dato un insieme A di \mathbb{R}^2 e data $f(x, y)$ definita (almeno) in A , la funzione

$$\hat{f}_A(x, y) = \begin{cases} f(x, y) & \text{se } (x, y) \in A \\ 0 & \text{se } (x, y) \notin A \end{cases}$$

si chiama *estensione standard di f (relativa ad A)*.

Spesso risulterà evidente dal contesto rispetto a quale insieme A si sta considerando l'estensione standard di una funzione f . In tal caso scriveremo \hat{f} al posto di \hat{f}_A .

Definizione (di integrale doppio in un arbitrario insieme limitato). Sia $f(x, y)$ una funzione di due variabili definita (almeno) 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:

$$\iint_A f(x, y) dx dy := \iint_R \hat{f}(x, y) dx dy.$$

Con riferimento alla suddetta definizione, si fa notare che, essendo \hat{f} nulla fuori da A , il secondo integrale non dipende dal rettangolo R contenente A . Pertanto, la suddetta definizione è ben posta.

Teorema (di additività rispetto all'insieme di integrazione). *Supponiamo che una funzione $f(x, y)$ sia integrabile sia in un insieme A che in un insieme B , con $A \cap B = \emptyset$. Allora f è integrabile in $A \cup B$ e*

$$\iint_{A \cup B} f(x, y) dx dy = \iint_A f(x, y) dx dy + \iint_B f(x, y) dx dy.$$

Dimostrazione^{fac.} Fissiamo un rettangolo R contenente $A \cup B$ e consideriamo, rispettivamente, le estensioni standard $\hat{f}_{A \cup B}$, \hat{f}_A e \hat{f}_B di f relative agli insiemi $A \cup B$, A e B . Dal fatto che $A \cap B = \emptyset$ si deduce facilmente che $\hat{f}_{A \cup B} = \hat{f}_A + \hat{f}_B$. Quindi

$$\int_{A \cup B} f = \int_R \hat{f}_{A \cup B} = \int_R (\hat{f}_A + \hat{f}_B) = \int_R \hat{f}_A + \int_R \hat{f}_B = \int_A f + \int_B f,$$

e ciò prova la tesi. □

Sia $A \subseteq \mathbb{R}^2$ un insieme del tipo

$$A = \left\{ (x, y) : a \leq x \leq b, \varphi_1(x) \leq y \leq \varphi_2(x) \right\},$$

dove $\varphi_1, \varphi_2: [a, b] \rightarrow \mathbb{R}$ sono due funzioni continue. Si dice che l'insieme A *presenta il caso semplice rispetto all'asse y* , o che è *y -semplice*, perché 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 $f(x, y)$ sia una funzione integrabile in A . Dato un rettangolo $R = [a, b] \times [c, d]$ contenente A , per definizione l'integrale di f in A è

$$\iint_R \hat{f}(x, y) \, dx dy,$$

dove \hat{f} è l'estensione standard di f (relativa ad A). Dal Teorema di Fubini si ha

$$\iint_R \hat{f}(x, y) \, dx dy = \int_a^b dx \int_c^d \hat{f}(x, y) \, dy.$$

D'altra parte

$$\int_c^d \hat{f}(x, y) \, dy = \int_c^{\varphi_1(x)} \hat{f}(x, y) \, dy + \int_{\varphi_1(x)}^{\varphi_2(x)} \hat{f}(x, y) \, dy + \int_{\varphi_2(x)}^d \hat{f}(x, y) \, dy,$$

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 quando l'insieme di integrazione è y -semplice):

$$\iint_A f(x, y) \, dx dy = \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} f(x, y) \, dy.$$

Analogamente, se $A \subseteq \mathbb{R}^2$ è un insieme del tipo

$$A = \left\{ (x, y) : c \leq y \leq d, \psi_1(y) \leq x \leq \psi_2(y) \right\},$$

dove $\psi_1, \psi_2 : [c, d] \rightarrow \mathbb{R}$ sono due funzioni continue ed $f(x, y)$ è integrabile in A , si ha l'altra **formula di riduzione**, valida quando A è x -semplice:

$$\iint_A f(x, y) \, dx dy = \int_c^d dy \int_{\psi_1(y)}^{\psi_2(y)} f(x, y) \, dx.$$

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

$$\mu(A) = \iint_A dx dy.$$

Purtroppo, non tutti i sottoinsiemi limitati del piano sono misurabili. Si consideri, ad esempio, l'insieme A dei punti di \mathbb{R}^2 con entrambe le coordinate razionali comprese tra 0 e 1. Ossia

$$A = \left\{ (x, y) \in [0, 1] \times [0, 1] : x \in \mathbb{Q}, y \in \mathbb{Q} \right\}.$$

Si potrebbe provare che la funzione \hat{f} che vale 1 in A e 0 nel complementare di A è discontinua in tutti i punti dell'intero quadrato $Q = [0, 1] \times [0, 1]$, che ovviamente non è trascurabile. Pertanto \hat{f} non è integrabile e, di conseguenza, A non è misurabile (secondo Peano-Jordan). Lo è, però, secondo una più moderna teoria dell'integrazione dovuta al matematico francese Lebesgue (si legge "Lebeg"). È bene precisare che l'importanza della teoria di Lebesgue non è dovuta al fatto che ci permette di misurare insiemi strani (in fondo, chi se ne frega!?): sono le sue proprietà e i teoremi che ne conseguono che la rendono particolarmente utile, specialmente per le applicazioni alla Fisica e all'Ingegneria. In un certo senso la teoria dell'integrazione di Lebesgue sta a quella di Cauchy-Riemann come i numeri reali stanno ai razionali. I numeri razionali (gli unici noti al tempo di Pitagora) sono infatti sufficienti per misurare, con l'approssimazione che si desidera, tutte le grandezze fisiche che ci interessano, ma senza i numeri reali non ci sarebbero importanti risultati come il Teorema di Weierstrass, il Teorema di Rolle, ecc.

Sia A un sottoinsieme limitato di \mathbb{R}^2 . Consideriamo la cosiddetta *funzione caratteristica* di A . Ossia la funzione $\hat{f} : \mathbb{R}^2 \rightarrow \mathbb{R}$ che vale 1 in A e 0 fuori di A . Non è difficile verificare che l'insieme dei punti di discontinuità di \hat{f} coincide con ∂A . Si può pertanto concludere che 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.

128 - Martedì 01/06/10

Primo teorema della media per gli integrali doppi. *Sia $f : A \rightarrow \mathbb{R}$ una funzione integrabile in un insieme misurabile $A \subseteq \mathbb{R}^2$ di misura non nulla. Allora la media di f in*

A , ossia

$$\frac{1}{\mu(A)} \int_A f(p) d\mu,$$

è un numero compreso tra l'estremo inferiore e l'estremo superiore di f . In particolare, se f è continua ed A è connesso, allora (per il Teorema dei valori intermedi) esiste un punto $c \in A$ per il quale si ha

$$\int_A f(p) d\mu = f(c)\mu(A).$$

Dimostrazione. Denotiamo, rispettivamente, con m e M l'estremo inferiore e l'estremo superiore di $f(p)$ per $p \in A$. Si ha

$$m \leq f(p) \leq M, \quad \forall p \in A.$$

Quindi, per la proprietà di monotonia, risulta

$$\int_A m d\mu \leq \int_A f(p) d\mu \leq \int_A M d\mu.$$

Dividendo i tre membri della suddetta disuguaglianza per l'area

$$\mu(A) = \int_A d\mu$$

di A si ottiene la tesi. □

Secondo teorema della media per gli integrali doppi. Siano $f, g: A \rightarrow \mathbb{R}$ due funzioni integrabili in un insieme $A \subseteq \mathbb{R}^2$. Se g è positiva in A , allora (quando ha senso) la media ponderata di f in A (con peso g), ossia

$$\frac{\int_A f(p)g(p) d\mu}{\int_A g(p) d\mu},$$

è un numero compreso tra l'estremo inferiore e l'estremo superiore di f . Pertanto, se f è continua ed A è connesso, esiste un punto $c \in A$ per il quale si ha

$$\int_A f(p)g(p) d\mu = f(c) \int_A g(p) d\mu.$$

Dimostrazione. Denotiamo, rispettivamente, con m e M l'estremo inferiore e l'estremo superiore di $f(p)$ per $p \in A$. Dato che $g(p) > 0$ in A , risulta

$$mg(p) \leq f(p)g(p) \leq Mg(p), \quad \forall p \in A.$$

Quindi, dalla proprietà di monotonia, si ottiene

$$m \int_A g(p) d\mu \leq \int_A f(p)g(p) d\mu \leq M \int_A g(p) d\mu.$$

Dividendo (quando ha senso) i tre membri della precedente disuguaglianza per

$$\int_A g(p) d\mu$$

si ottiene la tesi. □

Si osservi che il secondo teorema della media si riduce al primo quando $g(p)$ è costante.

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 la media delle ordinate. Si ha pertanto

$$x_c = \frac{1}{\mu(A)} \iint_A x dx dy, \quad y_c = \frac{1}{\mu(A)} \iint_A y dx dy.$$

Si osservi che dal primo teorema della media segue

$$\inf_{(x,y) \in A} x \leq x_c \leq \sup_{(x,y) \in A} x \quad \text{e} \quad \inf_{(x,y) \in A} y \leq y_c \leq \sup_{(x,y) \in A} y.$$

Quindi, se A è contenuto in un rettangolo $[a, b] \times [c, d]$, allora $a \leq x_c \leq b$ e $c \leq y_c \leq d$.

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

$$m = \int_A \delta(x, y) dx dy$$

è la massa della piastra.

Dal secondo teorema della media segue che se la piastra A è contenuta in un rettangolo, allora anche il suo centro di massa sta nel rettangolo.

Determiniamo, per esempio, il centro di massa (geometrico) del semicerchio

$$A = \left\{ (x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2, y \geq 0 \right\}$$

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

$$y_c = \frac{2}{\pi r^2} \iint_A y dx dy = \frac{2}{\pi r^2} \int_{-r}^r dx \int_0^{\sqrt{r^2-x^2}} y dy = \frac{1}{\pi r^2} \int_{-r}^r (r^2 - x^2) dx = \frac{4r}{3\pi}.$$

Si osservi che $4r/3\pi$ è un numero tra 0 ed r (in accordo col teorema della media); anzi, è addirittura minore di $r/2$ (per quale ragione deve essere così?).

129 - Martedì 01/06/10

Teorema di Pappo^{sd} (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 uguale al 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

$$\frac{\pi r^2}{2} \frac{8r}{3} = \frac{4}{3}\pi r^3.$$

Si osservi che se si deriva la funzione $f(r) = \frac{4}{3}\pi r^3$ che dà il volume di una sfera di raggio r si ottiene $4\pi r^2$, cioè l'area della sfera di raggio r . Perché?

Cosa si ottiene se si deriva la funzione $f(r) = \pi r^2$? Dare un'interpretazione geometrica alla risposta.

Per quale motivo se si deriva la funzione $f(l) = l^2$, che rappresenta l'area di un quadrato di lato l , non si ottiene il perimetro del quadrato?

Esercizio. Determinare il volume del (toro) solido che si ottiene facendo ruotare intorno all'asse y un cerchio di raggio r col centro posto ad una distanza $l > r$ dall'asse di rotazione.

Il *momento d'inerzia* rispetto ad un punto $c \in \mathbb{R}^2$ di una piastra omogenea A di peso m è il numero

$$I = \iint_A d(p, c)^2 \delta \, d\mu,$$

dove $d(p, c)$ è la funzione distanza di un generico punto p dal punto di riferimento c e $\delta = m/\mu(A)$ è la densità superficiale della piastra.

Se la piastra non è omogenea il suddetto integrale dà ancora il momento d'inerzia della piastra, ma in tal caso la densità è una funzione $\delta(p)$ del generico punto $p \in A$. Talvolta l'espressione $\delta(p) \, d\mu$ si denota col simbolo dm , detto *elemento di massa*.

Analogamente, il *momento d'inerzia* rispetto ad una retta $\alpha \subseteq \mathbb{R}^2$ di una piastra A (non necessariamente omogenea) è il numero

$$I = \iint_A d(p, \alpha)^2 \, dm,$$

dove $d(p, \alpha)$ è la funzione distanza di un generico punto p dalla retta di riferimento α .

Ricordiamo che la matrice jacobiana in un punto p di una funzione φ si denota $\varphi'(p)$. Quindi, se φ è una funzione da \mathbb{R}^2 in \mathbb{R}^2 , $\det \varphi'(u, v)$ rappresenta il determinante della

matrice jacobiana di φ nel punto $p = (u, v)$, detto *jacobiano* di φ in (u, v) . Ovviamente $|\det \varphi'(u, v)|$ denota il valore assoluto dello jacobiano di φ in (u, v) .

Ricordiamo inoltre che un sottoinsieme $A \subseteq \mathbb{R}^k$ si dice compatto se è limitato e chiuso.

Teorema^{sd} (di cambiamento di variabili per gli integrali doppi). *Sia*

$$\varphi(u, v) = (\varphi_1(u, v), \varphi_2(u, v))$$

un'applicazione continua da un compatto $A \subseteq \mathbb{R}^2$ in \mathbb{R}^2 . Supponiamo che A e $\varphi(A)$ siano misurabili e che φ sia C^1 e iniettiva nell'interno $\overset{\circ}{A} = A \setminus \partial A$ di A . Allora, data una funzione $f(x, y)$ continua su $\varphi(A)$, risulta

$$\iint_{\varphi(A)} f(x, y) \, dx dy = \iint_A f(\varphi_1(u, v), \varphi_2(u, v)) |\det \varphi'(u, v)| \, du \, dv.$$

Coordinate polari (v. libro di testo).

A titolo di esempio, calcoliamo il momento d'inerzia (rispetto al centro) di un disco omogeneo di massa m e raggio r . Denotiamo con D 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

$$I = \iint_D (x^2 + y^2) \, dm,$$

dove $dm = \delta \, dx dy$ è l'elemento di massa. Data la simmetria circolare della funzione integranda $x^2 + y^2$ e del dominio di integrazione D , è conveniente individuare i punti di D mediante le coordinate polari ed esprimere la funzione f in tali coordinate. Ogni punto $p \in D$ è individuato da due numeri, ρ e θ , dette coordinate polari, e le coordinate cartesiane di p sono legate alle polari dalle seguenti due equazioni (di cambiamento di coordinate):

$$x = \rho \cos \theta, \quad y = \rho \sin \theta.$$

I punti di D si ottengono (tutti quanti) facendo variare ρ tra 0 e r e θ tra 0 e 2π ; cioè facendo variare la coppia di numeri (ρ, θ) nel rettangolo compatto $A = [0, r] \times [0, 2\pi]$ del piano $\rho\theta$. Abbiamo quindi definito un'applicazione $\varphi: A \rightarrow \mathbb{R}^2$ la cui immagine $\varphi(A)$ coincide col dominio d'integrazione

$$D = \left\{ (x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq r^2 \right\}.$$

Poiché, come si può facilmente verificare, 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

$$I = \delta \iint_D (x^2 + y^2) \, dx dy = \delta \iint_A \rho^2 |\rho| \, d\rho \, d\theta = \delta \iint_A \rho^3 \, d\rho \, d\theta.$$

Si osservi che le ipotesi del teorema di cambiamento di variabili sono soddisfatte. Infatti A è compatto, $D = \varphi(A)$, φ è continua in A , è C^1 nell'interno di A (è addirittura C^∞)

è iniettiva nell'interno di A (anche se non lo è nella frontiera). Concludendo, per il Teorema di Fubini, si ha

$$I = \delta \iint_A \rho^3 d\rho d\theta = \frac{m}{\pi r^2} \int_0^{2\pi} d\theta \int_0^r \rho^3 d\rho = \frac{1}{2} m r^2.$$

Esercizio. Calcolare l'area dell'ellisse

$$E = \left\{ (x, y) \in \mathbb{R}^2 : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1 \right\}$$

di semiassi a e b .

Suggerimento. Usare il seguente cambiamento di coordinate:

$$x = ta \cos \varphi, \quad y = tb \sin \varphi, \quad (t, \varphi) \in [0, 1] \times [0, 2\pi].$$

Esercizio. Determinare il baricentro del cerchio forato

$$A = \left\{ (x, y) \in \mathbb{R}^2 : x^2 + y^2 \leq 16, (x-1)^2 + y^2 \geq 1 \right\}.$$

Suggerimento. Usare la proprietà di additività dell'integrale rispetto all'insieme di integrazione.

130 - Venerdì 04/06/10

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 $n_1 n_2 n_3$ sottoparallelepipedi

$$Q_{ijk} = [x_{i-1}, x_i] \times [y_{j-1}, y_j] \times [z_{k-1}, z_k]$$

di volume

$$\mu(Q_{ijk}) = (x_i - x_{i-1})(y_j - y_{j-1})(z_k - z_{k-1}) = (\Delta x)_i (\Delta y)_j (\Delta z)_k.$$

In ogni parallelepipedo Q_{ijk} scegliamo un punto $q_{ijk} = (x_{ijk}, y_{ijk}, z_{ijk})$. L'insieme s dei punti q_{ijk} si dice una *scelta di punti* nella partizione $p = (p_1, p_2, p_3)$ di Q . Ogni parallelepipedo Q_{ijk} della partizione col punto q_{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 parallelepipedi individuati dalla partizione p .

Sia $f(x, y, z)$ una funzione definita in Q . Ad ogni partizione puntata $\alpha = (p, s)$ di Q possiamo associare il numero

$$S_f(\alpha) = \sum_{(i,j,k) \in \mathcal{K}} f(q_{ijk}) \mu(Q_{ijk}) = \sum_{i=1}^{n_1} \sum_{j=1}^{n_2} \sum_{k=1}^{n_3} f(x_{ijk}, y_{ijk}, z_{ijk}) (\Delta x)_i (\Delta y)_j (\Delta z)_k,$$

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 (secondo Cauchy-Riemann) 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 I è l'integrale triplo di f in Q se, fissato un "errore" $\varepsilon > 0$, esiste un $\delta > 0$ tale che, comunque si assegni una partizione puntata α con parametro di finezza $|\alpha|$ minore di δ , la somma $S_f(\alpha)$ sopra definita dista da I meno di ε . Se ciò accade, si scrive

$$\lim_{|\alpha| \rightarrow 0} S_f(\alpha) = I$$

e la funzione f si dice *integrabile* (in Q) secondo Cauchy-Riemann. Il numero I si chiama *integrale (triplo)* di $f(x, y, z)$ in Q e si denota con uno dei seguenti simboli:

$$\int_Q f, \quad \int_Q f \, d\mu, \quad \int_Q f(p) \, d\mu, \quad \int_Q f(x, y, z) \, dx dy dz,$$

$$\iiint_Q f, \quad \iiint_Q f \, d\mu, \quad \iiint_Q f(p) \, d\mu, \quad \iiint_Q f(x, y, z) \, dx dy dz.$$

È ovvio che il numero

$$I = \iiint_Q f(x, y, z) \, dx dy dz$$

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 $S_f(\alpha)$ per $|\alpha| \rightarrow 0$ non cambia).

Un sottoinsieme di \mathbb{R}^3 si dice *trascurabile* (in \mathbb{R}^3) se per ogni $\varepsilon > 0$ può essere ricoperto con una famiglia (al più) numerabile di parallelepipedi di volume totale minore di ε . Si potrebbe dimostrare che il grafico di una funzione continua di due variabili ($z = g(x, y)$, o $x = g(y, z)$, o $y = g(z, x)$) 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

$$\int_Q (\alpha f + \beta g) d\mu = \alpha \int_Q f d\mu + \beta \int_Q g d\mu,$$

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

$$\int_Q f d\mu \geq 0,$$

e da ciò segue (tenendo conto della linearità) la seguente proprietà dell'integrale triplo:

Proprietà di monotonia. *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*

$$\int_Q f d\mu \leq \int_Q g d\mu.$$

Il risultato che segue 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^{sd} (per gli integrali tripli). *Sia $f(x, y, z)$ una funzione reale definita in un parallelepipedo $Q = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$. Allora, quando ha senso, risulta*

$$\iiint_Q f(x, y, z) dx dy dz = \iint_R dx dy \int_{a_3}^{b_3} f(x, y, z) dz,$$

$$\iiint_Q f(x, y, z) dx dy dz = \int_{a_3}^{b_3} dz \iint_R f(x, y, z) dx dy,$$

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 (per gli integrali tripli) i ruoli delle variabili x , y e z possono essere permutati, ottenendo altre quattro formule di riduzione. In tutto sono sei: due se l'integrale è semplice rispetto a z (come nell'enunciato), due se è rispetto ad y e due se è rispetto ad x .

131 - Venerdì 04/06/10

Definizione. Dato un insieme A di \mathbb{R}^3 e data $f(x, y, z)$ definita (almeno) in A , la funzione

$$\hat{f}_A(x, y, z) = \begin{cases} f(x, y, z) & \text{se } (x, y, z) \in A \\ 0 & \text{se } (x, y, z) \notin A \end{cases}$$

si chiama *estensione standard di f (relativa ad A)*.

Spesso risulterà evidente dal contesto rispetto a quale insieme A si sta considerando l'estensione standard di una funzione f . In tal caso scriveremo \hat{f} al posto di \hat{f}_A .

Sia $f(x, y, z)$ 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 suddetta definizione è ben posta.

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 $A \cap B = \emptyset$. Allora f è integrabile in $A \cup B$ e*

$$\iiint_{A \cup B} f(x, y, z) \, dx dy dz = \iiint_A f(x, y, z) \, dx dy dz + \iiint_B f(x, y, z) \, dx dy dz.$$

Dimostrazione ^{fac.} Fissiamo un parallelepipedo Q contenente $A \cup B$ e consideriamo, rispettivamente, le estensioni standard $\hat{f}_{A \cup B}$, \hat{f}_A e \hat{f}_B di f relative agli insiemi $A \cup B$, A e B . Dal fatto che $A \cap B = \emptyset$ si deduce facilmente che $\hat{f}_{A \cup B} = \hat{f}_A + \hat{f}_B$. Quindi

$$\int_{A \cup B} f = \int_Q \hat{f}_{A \cup B} = \int_Q (\hat{f}_A + \hat{f}_B) = \int_Q \hat{f}_A + \int_Q \hat{f}_B = \int_A f + \int_B f,$$

e ciò prova la tesi. □

Sia $A \subseteq \mathbb{R}^3$ un insieme del tipo

$$A = \{(x, y, z) : (x, y) \in B, \varphi_1(x, y) \leq z \leq \varphi_2(x, y)\},$$

dove $\varphi_1, \varphi_2: B \rightarrow \mathbb{R}$ sono due funzioni continue definite in un sottoinsieme compatto B di \mathbb{R}^2 . Si dice che A presenta il caso semplice rispetto all'asse z , o che è z -semplice, perché ogni retta parallela a tale asse lo interseca in un intervallo (di estremi $\varphi_1(x, y)$ e $\varphi_2(x, y)$, per $(x, y) \in B$). Supponiamo che $f(x, y, z)$ sia una funzione integrabile in A . Per definizione, dato un parallelepipedo $Q = [a_1, b_1] \times [a_2, b_2] \times [a_3, b_3]$ contenente A , l'integrale di f in A è

$$\iiint_Q \hat{f}(x, y, z) \, dx dy dz.$$

Dal Teorema di Fubini si ha

$$\iiint_Q \hat{f}(x, y, z) \, dx dy dz = \iint_R dx dy \int_{a_3}^{b_3} \hat{f}(x, y, z) \, dz,$$

dove $R = [a_1, b_1] \times [a_2, b_2]$. Ossia,

$$\iiint_A f(x, y, z) \, dx dy dz = \iint_R g(x, y) \, dx dy,$$

dove $g: R \rightarrow \mathbb{R}$ è la funzione definita da

$$g(x, y) = \int_{a_3}^{b_3} \hat{f}(x, y, z) \, dz.$$

Osserviamo ora che, per la definizione di \hat{f} , la funzione $g(x, y)$ è nulla se $(x, y) \notin B$. Di conseguenza,

$$\iint_R g(x, y) \, dx dy = \iint_B g(x, y) \, dx dy + \iint_{R \setminus B} g(x, y) \, dx dy = \iint_B g(x, y) \, dx dy.$$

Tenendo conto che, dato $(x, y) \in B$, la funzione parziale $z \mapsto \hat{f}(x, y, z)$ è nulla se z non appartiene all'intervallo $[\varphi_1(x, y), \varphi_2(x, y)]$, si ha

$$g(x, y) = \int_{a_3}^{b_3} \hat{f}(x, y, z) \, dz = \int_{\varphi_1(x, y)}^{\varphi_2(x, y)} \hat{f}(x, y, z) \, dz = \int_{\varphi_1(x, y)}^{\varphi_2(x, y)} f(x, y, z) \, dz.$$

Si ottiene così la seguente importante **formula di riduzione**, detta anche **formula degli spaghetti** (paralleli all'asse z), valida per gli insiemi che presentano il caso semplice rispetto all'asse z :

$$\iiint_A f(x, y, z) \, dx dy dz = \iint_B dx dy \int_{\varphi_1(x, y)}^{\varphi_2(x, y)} f(x, y, z) \, dz,$$

dove B è la proiezione ortogonale di A sul piano xy e $\varphi_1, \varphi_2: B \rightarrow \mathbb{R}$ 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.

Il Teorema di Fubini in \mathbb{R}^3 ci dice che per calcolare un integrale triplo si può eseguire prima un integrale doppio e poi un integrale semplice. Da tale teorema si deduce 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 .

Riportiamo la formula delle fette perpendicolari all'asse z . Il compito di scrivere le altre due è lasciato per esercizio allo studente.

Sia $f(x, y, z)$ 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 col piano perpendicolare all'asse z e passante per il punto $(0, 0, z)$. Sia $[a, b]$ un qualunque intervallo contenente la proiezione ortogonale di A sull'asse z (la scelta più conveniente si ottiene prendendo $a = \inf\{z : (x, y, z) \in A\}$ e $b = \sup\{z : (x, y, z) \in A\}$). Allora vale la seguente **formula delle fette**:

$$\iiint_A f(x, y, z) \, dx dy dz = \int_a^b dz \iint_{A_z} f(x, y, z) \, dx dy.$$

Dimostrazione della formula delle fette. Siano, rispettivamente, R ed $[a, b]$ un rettangolo nel piano xy e un intervallo nell'asse z scelti in modo che il parallelepipedo $Q = R \times [a, b]$ contenga A . Per definizione, si ha

$$\iiint_A f(x, y, z) \, dx dy dz = \iiint_Q \hat{f}(x, y, z) \, dx dy dz,$$

e per il Teorema di Fubini risulta

$$\iiint_A f(x, y, z) \, dx dy dz = \int_a^b dz \iint_R \hat{f}(x, y, z) \, dx dy.$$

Poiché, fissato $z \in [a, b]$, la funzione $\hat{f}(x, y, z)$ è nulla fuori da A_z , si ottiene

$$\iint_R \hat{f}(x, y, z) \, dx dy = \iint_{A_z} f(x, y, z) \, dx dy.$$

Pertanto

$$\iiint_A f(x, y, z) \, dx dy dz = \int_a^b dz \iint_{A_z} f(x, y, z) \, dx dy,$$

e la formula è provata. □

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

$$\mu_3(A) = \iiint_A dx dy dz.$$

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 $b = \sup\{z : (x, y, z) \in A\}$).

Esercizio. Calcolare il volume di un cono circolare retto di altezza h e raggio di base r .

Esercizio. Calcolare il volume di una sfera di raggio r col metodo delle fette.

Esercizio. Calcolare il volume di una sfera di raggio r col metodo degli spaghetti.

Esercizio. Enunciare i due teoremi della media per gli integrali tripli.

Esercizio ^{fac.} Provare i due teoremi della media per gli integrali tripli.

Esercizio. Definire il concetto di centro di massa di un solido $A \subseteq \mathbb{R}^3$ di densità $\delta(x, y, z)$.

Esercizio. Determinare il centro di massa di una semisfera (omogenea) di raggio r (ossia, determinare la distanza del centro di massa dal centro della sfera).

Esercizio. Definire il concetto di momento d'inerzia rispetto ad una retta di un solido $A \subseteq \mathbb{R}^3$ di densità $\delta(x, y, z)$.

Esercizio. Calcolare il momento d'inerzia di una sfera (omogenea) di massa m e raggio r rispetto ad una retta passante per il centro.

Esercizio. Enunciare il teorema di cambiamento di variabili per gli integrali tripli.

Coordinate sferiche (v. libro di testo o gli appunti presi a lezione).

Esercizio. Calcolare il volume di una sfera mediante le coordinate sferiche.

Appendice (argomenti facoltativi)

Ricordiamo che una curva parametrica $\gamma: [a, b] \rightarrow \mathbb{R}^k$ si dice semplice se esistono al più due punti con la stessa immagine, e quando ciò accade tali punti sono soltanto gli estremi a e b dell'intervallo di definizione (in tal caso la curva si dice chiusa).

Un sottoinsieme di \mathbb{R}^2 si dice una *curva di Jordan* (si pronuncia “giordàn”, con l'accento tonico sull'ultima sillaba) se è il sostegno (cioè l'immagine) di una curva semplice e chiusa. Il più banale esempio di curva di Jordan è costituito da una circonferenza. Un altro semplice esempio è dato dalla frontiera di un rettangolo.

Enunciamo, senza dimostrazione, un famoso risultato topologico dovuto al matematico francese Camille Jordan (1838-1922). Risultato tanto intuitivo quanto non banale da provare (come molti teoremi di Topologia).

Teorema di Jordan ^{sd}. *Il complementare di una curva di Jordan è unione di due aperti connessi, disgiunti, la cui frontiera è la curva stessa. Uno dei due aperti, detto “insieme*

dei punti racchiusi dalla curva”, è limitato; l’altro, detto “insieme dei punti esterni alla curva”, è illimitato.

Una curva di Jordan si dice una *catena di Jordan* se è decomponibile nell’unione di un numero finito di archi regolari (la decomposizione, ovviamente, non è unica). Ad esempio, la frontiera di un triangolo è una catena di Jordan, così come lo è una circonferenza. Non tutte le curve di Jordan sono catene. Esistono infatti curve di Jordan così irregolari da non contenere archi regolari. Una di queste è la frontiera frastagliata della cosiddetta isola di Koch (un noto frattale).

Un sottoinsieme **compatto** X di \mathbb{R}^2 si dice una *placca piana* se la sua frontiera ∂X è una catena di Jordan. Più in generale, diremo che un insieme **compatto e connesso** $X \subseteq \mathbb{R}^2$ è una *lamina (piana)* se la sua frontiera ∂X è costituita da un numero finito di catene di Jordan a due a due disgiunte. La frontiera di una lamina si dice anche *bordo*.

Ovviamente una placca è anche una lamina, ma non viceversa. Ad esempio, una corona circolare è una insieme delimitato da due catene di Jordan (nella fattispecie, due circonferenze concentriche), e quindi è una lamina (con un buco), ma non è una placca; mentre i quadrati, i cerchi e i triangoli, essendo delimitati da una sola catena di Jordan, sono placche, oltre che lamine (senza buchi). Le lamine, insomma, non sono altro che placche con eventuali fori. I fori, però, devono essere fatti bene: devono essere delimitati da catene di Jordan a due a due disgiunte (non deve capitare che due fori abbiano punti di frontiera a comune). In una placca, un foro come l’isola di Koch non fa una lamina: la frontiera non è una catena di Jordan.

Si osservi che le placche sono insiemi connessi (è una conseguenza del Teorema di Jordan). Si potrebbe provare, ma non lo facciamo, che sono addirittura semplicemente connessi. Le lamine, invece, se hanno dei buchi, non sono insiemi semplicemente connessi (si ricorda, però, che per definizione sono insiemi connessi, e anche compatti).

Una singola catena di Jordan può essere orientata in due modi, a seconda del senso di percorrenza: orario o antiorario. Quindi, un insieme costituito da n catene di Jordan a due a due disgiunte (come, ad esempio, il bordo di una lamina) può essere orientato in 2^n modi (due per ogni curva).

Per convenzione, data una lamina $X \subseteq \mathbb{R}^2$, l’*orientazione indotta* da X sulla sua frontiera (detta anche *orientazione canonica del bordo*) si ottiene percorrendo ∂X in modo che X si trovi sul lato sinistro e il complementare di X sul lato destro. Per esempio, l’orientazione indotta da una corona circolare sulla sua frontiera è antioraria sulla circonferenza esterna e oraria su quella interna. In parole povere il bordo di una lamina X si percorre in senso antiorario lungo la curva di Jordan che racchiude X (in base al Teorema di Jordan) e in senso orario lungo la frontiera degli eventuali fori.

Per comprendere meglio la suddetta definizione (non ortodossa) di orientazione indotta, si pensi al concetto di riva sinistra (o destra) di un fiume. D’altra parte, bisogna accontentarsi dell’idea intuitiva, perché la definizione formale richiederebbe concetti topologici troppo avanzati per il livello del corso. Purtroppo, le definizioni informali dei concetti

non consentono dimostrazioni formali dei teoremi che utilizzano detti concetti (c'è poco da fare!).

Da ora in poi, per semplicità di linguaggio, diremo che una funzione $f: A \rightarrow \mathbb{R}^s$ è C^n (o C^∞) su un insieme A di \mathbb{R}^k (non necessariamente aperto) se è C^n (o C^∞) su un aperto U contenente A . Ad esempio, la funzione $f(x, y) = x^2 + y^2$ è C^∞ nel quadrato $Q = [0, 1] \times [0, 1]$ perché in realtà è C^∞ in tutto \mathbb{R}^2 , che è un aperto contenente Q .

Il seguente risultato rappresenta per gli integrali doppi quello che per gli integrali semplici è la formula fondamentale del calcolo integrale. Sotto opportune ipotesi, infatti, l'integrale doppio dipende soltanto da ciò che accade sulla frontiera dell'insieme di integrazione (così come un integrale semplice dipende soltanto dai valori assunti negli estremi dell'intervallo di integrazione da una primitiva della funzione integranda).

Teorema^{sd} (formule di Gauss-Green nel piano). *Siano $A(x, y)$ e $B(x, y)$ due funzioni di classe C^1 su una lamina $X \subseteq \mathbb{R}^2$. Allora*

$$\begin{aligned} \iint_X \frac{\partial B}{\partial x}(x, y) dx dy &= \int_{\partial X} B(x, y) dy, \\ \iint_X \frac{\partial A}{\partial y}(x, y) dx dy &= - \int_{\partial X} A(x, y) dx, \end{aligned}$$

dove l'orientazione di ∂X è quella indotta da X .

Il risultato che segue ha un'importante interpretazione fisica (che gli studenti avranno modo di incontrare, in una formulazione più generale, studiando elettromagnetismo) e si ottiene sommando le due formule di Gauss-Green.

Teorema (della circuitazione nel piano). *Siano $A(x, y)$ e $B(x, y)$ due funzioni di classe C^1 su una lamina $X \subseteq \mathbb{R}^2$. Allora*

$$\int_{\partial X} A dx + B dy = \iint_X \det \begin{pmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ A & B \end{pmatrix} dx dy,$$

dove l'orientazione di ∂X è quella indotta da X .

Osservazione. Dal teorema della circuitazione, ponendo $A = 0$ o $B = 0$, si ottengono, come casi particolari, le due formule di Gauss-Green.

Osservazione. Se si scelgono A e B in modo che la funzione

$$\det \begin{pmatrix} \frac{\partial}{\partial x} & \frac{\partial}{\partial y} \\ A & B \end{pmatrix} = \frac{\partial B}{\partial x} - \frac{\partial A}{\partial y}$$

valga 1, allora l'integrale curvilineo

$$\int_{\partial X} A dx + B dy$$

rappresenta l'area di X . I casi più importanti sono i seguenti:

- 1) $A = -y/2, B = x/2;$
- 2) $A = 0, B = x;$
- 3) $A = -y, B = 0.$

Tra i tre casi, il secondo (quello con $A = 0$ e $B = x$) è particolarmente significativo in termodinamica, dove al posto della x c'è p (la pressione) e al posto della y c'è v (il volume). In tal caso l'integrale curvilineo dà il lavoro compiuto in un ciclo termodinamico; lavoro che coincide con l'area racchiusa dal ciclo stesso (gli studenti avranno modo di incontrare tali concetti in altri corsi).

Esercizio. Scrivere le tre formule per calcolare l'area mediante un integrale curvilineo che si deducono dai suddetti casi particolari.

Esercizio. Calcolare l'area di un cerchio mediante un integrale curvilineo.

Esercizio. Calcolare l'area di un'ellisse mediante un integrale curvilineo.

Una *superficie parametrica* è un'applicazione continua $\varphi: X \rightarrow \mathbb{R}^3$ definita su una placca piana. L'applicazione φ si dice *semplice* se è iniettiva; si dice *regolare* se è di classe C^1 e la sua matrice jacobiana ha rango due in ogni punto. Un sottoinsieme Σ di \mathbb{R}^3 è una *placca (di superficie regolare)* se è il sostegno (cioè l'immagine) di una superficie parametrica semplice e regolare $\varphi: X \rightarrow \mathbb{R}^3$, detta *parametrizzazione* della placca. L'immagine $\varphi(\partial X)$ della frontiera di X si dice *bordo* di Σ e si denota $\partial\Sigma$.

Un'espressione differenziale (scalare) di grado due su un aperto U di \mathbb{R}^k è una funzione continua $\omega: U \times \mathbb{R}^k \times \mathbb{R}^k \rightarrow \mathbb{R}$. In altre parole, è una "legge" che ad ogni terna (p, v_1, v_2) di $U \times \mathbb{R}^k \times \mathbb{R}^k$ (ossia, ad ogni coppia di vettori applicati nello stesso punto) associa (in modo continuo) un numero reale $\omega(p, v_1, v_2)$.

Ovviamente il prodotto $f\omega$ di una funzione (continua) $f: U \rightarrow \mathbb{R}^3$ per un'espressione differenziale ω su U (di grado due) è ancora un'espressione differenziale su U (di grado due): ad ogni terna (p, v_1, v_2) di $U \times \mathbb{R}^k \times \mathbb{R}^k$ associa il numero $f(p)\omega(p, v_1, v_2)$.

Tra tutte le espressioni differenziali di grado due, la più importante è l'*elemento di area* $d\sigma$, ossia quell'espressione differenziale che ad ogni coppia di vettori $v_1, v_2 \in \mathbb{R}^3$ (applicati nello stesso punto) associa l'area del parallelogramma da questi individuato. In altre parole $d\sigma(p, v_1, v_2) = \|v_1 \wedge v_2\|$, dove $v_1 \wedge v_2$ denota il prodotto vettoriale di v_1 per v_2 (un'altra notazione usata per il prodotto vettoriale è $v_1 \times v_2$). Si osservi che sia l'elemento di lunghezza ds sia l'elemento d'area $d\sigma$ non dipendono dal punto di applicazione (ma soltanto dai vettori liberi).

Ricordiamo che, data una funzione $\varphi(u, v)$ da \mathbb{R}^2 in \mathbb{R}^3 , con

$$\frac{\partial\varphi}{\partial u}(u, v) \quad \text{e} \quad \frac{\partial\varphi}{\partial v}(u, v)$$

si denotano, rispettivamente, i vettori che si ottengono derivando (nel punto (u, v)) le curve

parametriche “ v costante” e “ u costante”. Cioè i vettori colonna della matrice jacobiana

$$\varphi'(u, v) = \begin{pmatrix} \frac{\partial \varphi_1}{\partial u} & \frac{\partial \varphi_1}{\partial v} \\ \frac{\partial \varphi_2}{\partial u} & \frac{\partial \varphi_2}{\partial v} \\ \frac{\partial \varphi_3}{\partial u} & \frac{\partial \varphi_3}{\partial v} \end{pmatrix}_{(u,v)}$$

Definizione. Sia $\varphi: X \rightarrow \mathbb{R}^3$ una superficie parametrica di classe C^1 ed ω un’espressione differenziale di grado due su un aperto U contenente l’immagine di φ . L’integrale di ω su φ si definisce nel seguente modo:

$$\int_{\varphi} \omega = \iint_X \omega\left(\varphi(u, v), \frac{\partial \varphi}{\partial u}(u, v), \frac{\partial \varphi}{\partial v}(u, v)\right) dudv.$$

In particolare, se l’espressione ω è l’elemento di area $d\sigma$, il suddetto integrale si chiama *area della superficie parametrica* φ .

Per giustificare la definizione di area si invita^{fac} lo studente a controllare che quando φ è lineare ed X è un rettangolo $[a, b] \times [c, d]$, l’integrale di $d\sigma$ su φ coincide con l’area del parallelogramma immagine di φ .

Teorema^{sd}. Se $\varphi_1: X_1 \rightarrow \Sigma$ e $\varphi_2: X_2 \rightarrow \Sigma$ sono due parametrizzazioni di una stessa placca di superficie ed f è una funzione continua definita su un aperto contenente Σ , allora

$$\int_{\varphi_1} f d\sigma = \int_{\varphi_2} f d\sigma.$$

In particolare l’area di Σ non dipende dalla parametrizzazione.

In base al suddetto teorema ha senso quindi definire l’area di una placca Σ come l’integrale dell’espressione differenziale $d\sigma$ esteso ad una qualunque parametrizzazione $\varphi: X \rightarrow \Sigma$ di Σ . Integrale che, vista l’indipendenza dalla parametrizzazione, potrà anche essere scritto nella forma

$$\int_{\Sigma} d\sigma.$$

Più in generale, data una funzione continua $g: \Sigma \rightarrow \mathbb{R}$, tramite la formula di cambiamento di variabili per gli integrali doppi, si potrebbe dimostrare che l’integrale lungo una parametrizzazione di Σ dell’espressione differenziale $\omega = g d\sigma$, ottenuta moltiplicando g per l’elemento di area $d\sigma$, non dipende dalla parametrizzazione scelta. Pertanto, analogamente al caso dell’area, col simbolo

$$\int_{\Sigma} g d\sigma$$

si intenderà l’integrale dell’espressione differenziale $g d\sigma$ lungo una qualunque parametrizzazione della placca Σ .

Definizione. Un sottoinsieme Σ di \mathbb{R}^3 si dice una *superficie a placche* (o *superficie generalmente regolare*) se è unione finita di placche “ben collegate”; ossia, tali che l’intersezione

di due placche qualunque o è vuota, o è un sol punto, o è un arco di curva (regolare) facente parte del bordo di entrambe.

Ad esempio, una superficie sferica è unione di due placche (emisfero nord e emisfero sud) che si intersecano lungo il comune bordo (l'equatore).

Un altro esempio è costituito dalla superficie di un cubo: è unione di sei placche (sei facce quadrate).

Il *bordo* $\partial\Sigma$ di una superficie a placche Σ è l'unione degli archi di curva che delimitano una sola placca (un arco di curva che delimita due placche non fa parte del bordo di Σ).

Una superficie a placche si dice *chiusa* se il suo bordo è vuoto (ad esempio, la superficie di un cubo).

Ovviamente, se Σ è una superficie a placche, costituita dalle placche $\Sigma_1, \Sigma_2, \dots, \Sigma_n$, e $g: \Sigma \rightarrow \mathbb{R}$ è una funzione continua, l'integrale dell'espressione $g d\sigma$ sulla superficie Σ si definisce nel seguente modo:

$$\int_{\Sigma} g d\sigma = \int_{\Sigma_1} g d\sigma + \int_{\Sigma_2} g d\sigma + \dots + \int_{\Sigma_n} g d\sigma.$$

In particolare, l'area di una superficie a placche è la somma delle aree delle placche che compongono la superficie.

Esercizio. Calcolare l'area di una sfera di raggio R parametrizzandola in coordinate sferiche ($x = R \sin \varphi \cos \theta$, $y = R \sin \varphi \sin \theta$, $z = R \cos \varphi$; $-\pi \leq \theta \leq \pi$, $0 \leq \varphi \leq \pi$).

Un'*orientazione* di una placca di superficie è un campo continuo di vettori normali alla superficie; ossia una legge che ad ogni punto della placca assegna, con continuità, un vettore di norma unitaria e normale alla superficie in quel punto (cioè normale allo spazio tangente). Si potrebbe provare che ogni placca di superficie ha soltanto due possibili orientazioni. Una placca si dice *orientata* se è stata scelta una sua orientazione. Intuitivamente ciò significa avere scelto come positiva una delle due facce della superficie, quella da cui dipartono i vettori normali, e che possiamo dipingere con un colore convenzionale (ad esempio rosso). L'altra faccia, che consideriamo negativa, può essere colorata di bianco. Il fatto che si possano dipingere le facce di una placca con due differenti colori (senza che questi vengano a contatto) costituisce l'aspetto intuitivo di un concetto topologico non facile da definire in termini elementari: il concetto di superficie orientabile. Le placche, appunto, sono superfici orientabili.

Una placca orientata Σ induce in modo canonico un'orientazione sul suo bordo $\partial\Sigma$. Poiché la definizione rigorosa di *orientazione indotta (sul bordo)* richiede nozioni topologiche che trascendono i limiti del corso, ci limitiamo a darne un'idea intuitiva. Occorre una convenzione per decidere come deve essere percorso il bordo $\partial\Sigma$ di una placca orientata Σ . Abbiamo visto che Σ , essendo orientata, ha una faccia positiva (che supponiamo rossa) e una faccia negativa (che supponiamo bianca). In sintesi questo è il criterio di percorrenza: il bordo va percorso stando sulla faccia positiva in modo che la superficie si trovi sul lato sinistro (come se il bordo fosse un fiume e la faccia positiva lambisse la riva sinistra).

Un'orientazione di una superficie a placche è una collezione di orientazioni delle singole placche, in modo che due placche confinanti inducano orientazioni discordi sul bordo a comune.

Non sempre è possibile orientare una superficie a placche; quando lo è, la superficie si dice *orientabile*. Un famoso esempio di superficie (a placche) non orientabile è il cosiddetto *nastro di Möbius*, che può essere ottenuto incollando opportunamente due lati opposti di una placca rettangolare con due lati opposti di un'altra placca rettangolare (delle stesse dimensioni). Si osservi, infatti, che le due placche si possono “assemblare” in due modi possibili: in un caso si ottiene un cilindro (che è una superficie orientabile), nell'altro si ottiene il nastro di Möbius (che è non orientabile). Se i due rettangoli sono orientati (da una parte rossi e dall'altra bianchi), uno dei due modi di incollarli induce sul un lato comune ai due rettangoli la stessa orientazione, e non orientazioni opposte. In questo modo la superficie costituita dalle due placche rettangolari non è orientabile (infatti i due colori, rosso e bianco, “si toccano” sulla stessa faccia).

Il *flusso* di un campo vettoriale $\mathbf{f}: \Sigma \rightarrow \mathbb{R}^3$ attraverso una placca orientata Σ è il numero

$$\int_{\Sigma} \mathbf{f} \cdot \mathbf{n} \, d\sigma,$$

dove $\mathbf{n}: \Sigma \rightarrow \mathbb{R}^3$ è il campo di versori normali che definisce l'orientazione della placca. In generale, il flusso di un campo vettoriale attraverso una superficie a placche orientata è la somma dei flussi attraverso le singole placche (ricordiamo che tutte le placche sono orientate).

Definizione. Sia $\mathbf{f}: U \rightarrow \mathbb{R}^3$ un campo vettoriale di classe C^1 su un aperto U di \mathbb{R}^3 . La *divergenza* di \mathbf{f} è la funzione (reale di tre variabili reali) definita da

$$\operatorname{div} \mathbf{f} = \frac{\partial f_1}{\partial x} + \frac{\partial f_2}{\partial y} + \frac{\partial f_3}{\partial z},$$

dove $f_1, f_2, f_3: U \rightarrow \mathbb{R}$ denotano le tre componenti di \mathbf{f} . Un campo vettoriale si dice *solenoidale* se la sua divergenza è nulla.

Teorema della divergenza^{sd} (o di Gauss). Sia $\mathbf{f}: A \rightarrow \mathbb{R}^3$ un campo vettoriale di classe C^1 su un sottoinsieme compatto A di \mathbb{R}^3 avente per frontiera una superficie a placche ∂A . Allora, denotato con \mathbf{n}_e il campo dei versori normali alla frontiera di A e diretto verso l'esterno, risulta

$$\int_A \operatorname{div} \mathbf{f} = \int_{\partial A} \mathbf{f} \cdot \mathbf{n}_e \, d\sigma.$$

In altre parole: l'integrale della divergenza di \mathbf{f} esteso ad A coincide con il flusso di \mathbf{f} uscente dalla superficie ∂A .

Definizione. Sia $\mathbf{f}: U \rightarrow \mathbb{R}^3$ un campo vettoriale di classe C^1 su un aperto U di \mathbb{R}^3 . Si chiama *rotore* di \mathbf{f} il campo vettoriale di componenti

$$\frac{\partial f_3}{\partial y} - \frac{\partial f_2}{\partial z}, \quad \frac{\partial f_1}{\partial z} - \frac{\partial f_3}{\partial x}, \quad \frac{\partial f_2}{\partial x} - \frac{\partial f_1}{\partial y},$$

dove $f_1, f_2, f_3: U \rightarrow \mathbb{R}$ denotano le tre componenti di \mathbf{f} .

Come regola mnemonica, si usa scrivere

$$\operatorname{rot} \mathbf{f} = \det \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ \frac{\partial}{\partial x} & \frac{\partial}{\partial y} & \frac{\partial}{\partial z} \\ f_1 & f_2 & f_3 \end{pmatrix}$$

Un campo vettoriale con rotore nullo si dice *irrotazionale*.

Teorema della circuitazione^{sd} (o di Stokes). *Sia $\Sigma \subseteq \mathbb{R}^3$ una placca orientata (o, più in generale, una superficie a placche orientata) e sia $\mathbf{f}: \Sigma \rightarrow \mathbb{R}^3$ un campo vettoriale di classe C^1 su Σ . Denotato con \mathbf{n} il campo di versori normali a Σ secondo l'orientazione scelta, si ha*

$$\int_{\Sigma} \operatorname{rot} \mathbf{f} \cdot \mathbf{n} \, d\sigma = \int_{\partial\Sigma} \mathbf{f} \cdot d\mathbf{p},$$

dove il secondo integrale va calcolato tenendo conto dell'orientazione indotta da Σ sul suo bordo. Ovvero, il flusso del rotore di \mathbf{f} attraverso la superficie orientata Σ è uguale alla circuitazione di \mathbf{f} lungo il bordo di Σ (secondo l'orientazione indotta).

Esercizio. Provare che se $f(x, y)$ è una funzione di classe C^∞ (in un aperto U di \mathbb{R}^2), allora anche ogni soluzione dell'equazione differenziale $y' = f(x, y)$ è di classe C^∞ .

Suggerimento. Si ricorda che una funzione $y(x)$ è di classe C^∞ se è C^n per ogni $n \in \mathbb{N}$ e che una funzione è C^n se (è derivabile e) la sua derivata è di classe C^{n-1} . Si osservi inoltre che se $y(x)$ è una funzione C^{n-1} , allora è C^{n-1} anche la funzione composta $f(x, y(x))$.