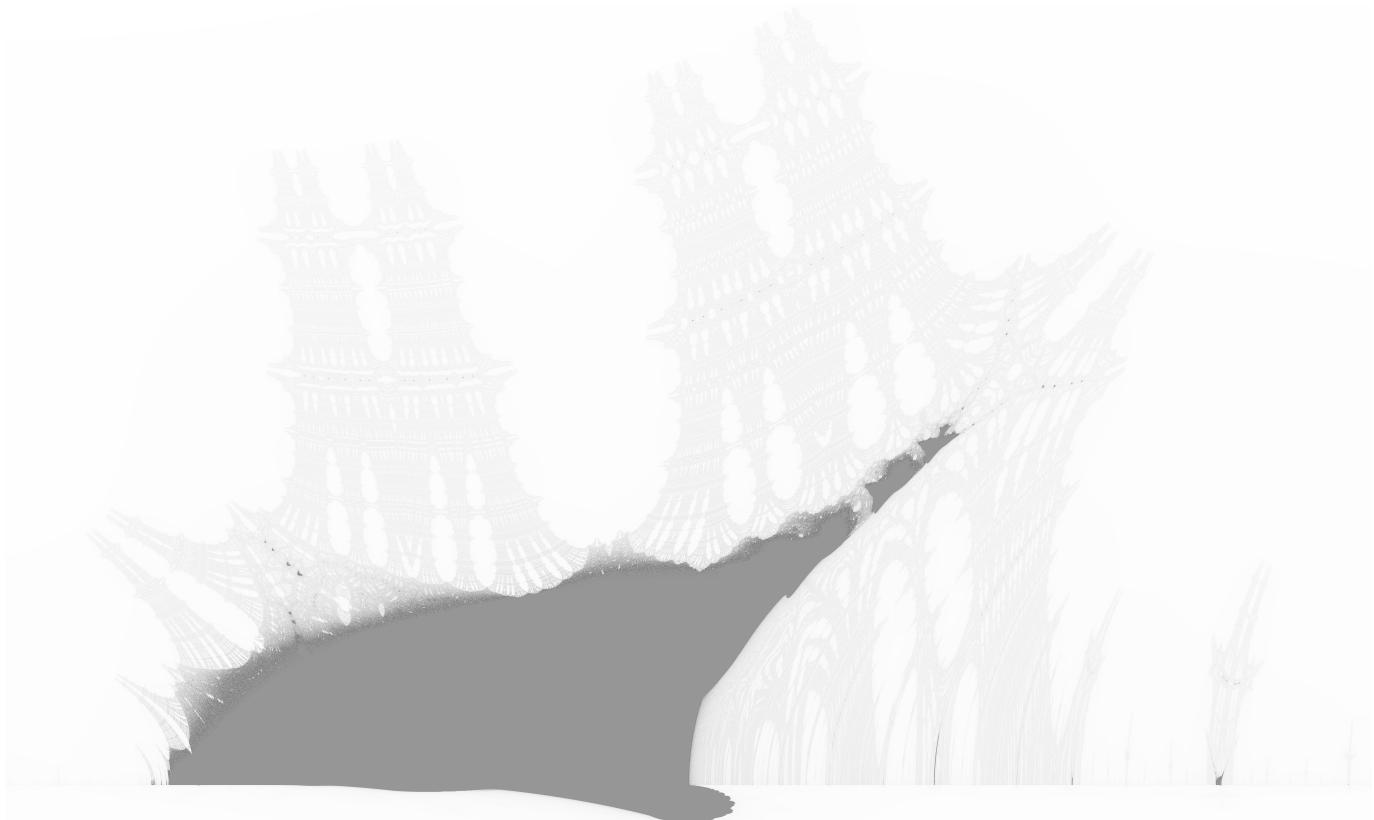


Appunti di Complementi di AM 2

Francesco Sermi



Introduzione

Queste note sono pensate come un riferimento per lo studio del corso di *Complementi di Analisi Matematica* tenuto dal professore Valentino Magnani. Esse fanno per lo più riferimento al programma seguito dal professore durante il primo semestre dell'anno 2023/2024, pertanto non garantisco che queste note riescano a coprire gli anni successivi.

In ogni caso, all'interno del documento sono riportate molte dimostrazioni non fatte dal professore che ho preso da altre fonti, mentre la soluzione proposta per la soluzione degli esercizi è 100% fatta da me. Nel capitolo sulla misura di Lebesgue, tuttavia, ho riportato tutte le (poche) dimostrazioni fatte dal professore e ne ho aggiunte alcune, tuttavia molte proposizioni e teoremi sono volutamente sprovvisti di dimostrazioni poiché abbastanza lunghe e complesse, tuttavia in futuro potrei pensare di integrarle al massimo nelle appendici come approfondimento.

Nella parte finale del documento (che si può riconoscere dalla presenza di una grafica che ho inserito sulla Γ) sono presenti delle appendici in cui affronto degli argomenti che il professore non è riuscito a "toccare" durante il corso, ma che mi sono risultate utili per comprendere altri corsi (nel momento in cui scrivevo queste note stavo seguendo il corso di *Analisi complessa e teoria delle superfici* in cui si faceva uso di qualche argomento presente nelle appendici) oppure interessanti. Prima di augurarvi il meglio, vorrei spendere qualche parola per ringraziare tutti coloro che sono intervenuti nella correzione di questo documento. In particolar modo ringrazio Francesco Fabiano Angelo Antonacci per l'aiuto prestatomi nel correggere certe dimostrazioni, Stefano Nordio e Davide Geria per avermi aiutato nell'individuare alcuni errori di scrittura e/o mie dimenticanze. Una menzione d'onore va a Beatrice, la mia ragazza, che mi ha dovuto sostenere e sopportare durante la stesura di questo documento che mi ha occupato un bel po' della mia esistenza e ai miei coinvolti Matteo, Anna e Alessandro che mi sopportano durante le mie farneticazioni fra spazi vettoriali.

Per segnalarmi eventuali errori nel documento (siano essi di tipo oppure formali), contattatemi presso

francesco291104 (dot) gmail.com

Buono studio!

Le grafiche presenti in questo libro sono disponibili nel repository di questo documento e sono tutte fatte da me, dunque, se utilizzate, mi farebbe estremamente piacere riceverne il merito (in pratica ve le fornisco senza problemi sotto una forma di licenza *Creative Commons "morale"*).

Notazioni

Riporto all'inizio del libro le notazioni adottate all'interno di questo documento:

\lim limite

$\frac{\partial f}{\partial x_i}, \partial_{e_i} f, \partial_{x_i} f, f_{x_i}$ derivata parziale rispetto a x_i

$\frac{\partial f}{\partial v}, \partial_v f$ derivata direzionale rispetto a v

S^n sfera unitaria centrata nell'origine in \mathbb{R}^{n+1}

∇ gradiente

$\operatorname{div} \cdot A$ divergenza del campo vettoriale A

$\operatorname{rot} A$ rotore del campo vettoriale A

\underline{v} vettore (alcune volte senza il trattino)

$\|\cdot\|, |\cdot|$ norma

$\langle \cdot, \cdot \rangle$ prodotto scalare/hermitiano

$a \times b, a \wedge b$ prodotto vettoriale

$\operatorname{Int}(A)$ parte interna di A

\bar{A} chiusura di A

$\operatorname{Fr} A$ frontiera di A

$B(x_0, r)$ palla aperta di raggio r centrata in x_0

$\mathbb{B}(x_0, r)$ palla chiusa di raggio r centrata in x_0

$\|A\|$ norma di Hilbert-Schmidt

$o(f)$ o-piccolo di f

$O(f)$ O-grande di f

Σ_n gruppo (rispetto all'operazione di composizione di funzioni) delle permutazioni $\sigma : \{1, \dots, n\} \rightarrow \{1, \dots, n\}$ bigettive.

ι_+ segnatura positiva di un prodotto scalare/hermitiano

ι_- segnatura negativa di un prodotto scalare/hermitiano

ι_0 segnatura nulla di un prodotto scalare/hermitiano

$(\iota_+, \iota_-, \iota_0)$ segnatura di un prodotto scalare/hermitiano

$I(x_0, \delta) = (x_0 - \delta, x_0 + \delta) = B(x_0, \delta) \subset \mathbb{R}$ con $x_0 \in \mathbb{R}, \delta > 0$ intervallo centrato in x_0 di ampiezza 2δ

m_n misura di Lebesgue

$\mathfrak{M}(X)$ classe degli insiemi misurabili per Lebesgue

$\mathcal{P}(X)$ insieme delle parti di X

\bigtimes_i prodotto cartesiano multiplo

\sum_i sommatoria

\sim relazione d'equivalenza

X/\sim insieme quoziante di X rispetto alla relazione d'equivalenza

Indice

1 Spazi euclidei	7
1.1 Spazio euclideo e prodotto scalare	7
1.2 Il prodotto vettoriale	10
1.3 Topologia di \mathbb{R}^n	15
2 Successioni e funzioni continue	19
2.1 Funzioni continue	20
2.2 Accenni ai limiti in più variabili	25
3 Calcolo differenziale	28
3.1 Funzione differenziabile	28
3.1.1 Interpretazione geometrica delle derivate prime	36
3.1.2 Legame fra differenziabilità e continuità	38
3.2 Derivate di ordine successivo	39
3.3 Funzioni polinomiali omogenee e polinomio di Taylor	42
3.4 Massimi e minimi in più variabili	48
4 Curve e superfici regolari	51
4.1 Teorema della funzione inversa	51
4.2 Teorema delle funzioni implicite o del Dini	52
4.3 Punti critici su curve e superfici regolari	56
5 Integrali curvilinei e potenziali	59
5.1 Curve e integrali curvilinei	59
5.2 Integrali curvilinei su forme differenziali e campi vettoriali	62
5.2.1 Forme differenziali esatte	65
5.2.2 Omotopia e curve omotope	72
5.2.3 Campi vettoriali radiali	75
6 Cenni alla misura di Lebesgue	76
6.1 La misura esterna	76
6.1.1 Insiemi trascurabili	79
6.2 Funzioni misurabili secondo Lebesgue	81
6.2.1 L'integrale di Lebesgue	83
6.2.2 Introduzione agli spazi \mathcal{L}^p	86

6.3	Sezionamenti di insiemi in \mathbb{R}^n	87
6.4	Teorema del cambiamento di variabili	91
6.4.1	Coordinate polari	93
6.4.2	Volumi dei solidi di rotazione	95
6.4.3	Invarianza per traslazioni e omogeneità della misura di Lebesgue	97
6.4.4	Coordinate sferiche	99
6.5	Misura di Hausdroff 2-dimensionale	99
6.5.1	Area dei grafici	102
6.6	Teorema di Gauss-Green	103
6.6.1	Forme differenziali d'area	106
6.7	Flusso e teorema della divergenza	108
6.7.1	Frontiera regolare di un aperto	110
6.7.2	Legge di Gauss	112
6.7.3	Equazione di continuità in fluidodinamica	112
6.8	Teorema del rotore	113
7	Altri esercizi proposti	115
A	Notazione multi-indice	118
B	Dimostrazione del teorema della funzione inversa	121
B.1	Norma matriciale indotta	121
B.2	Qualche risultato preparatorio	125
C	Il prodotto esterno	131
C.1	Forme algebriche esterne	131
Bibliografia		133

CAPITOLO 1

Spazi euclidei

In questo capitolo, ci soffermeremo su alcuni risultati che possono essere ottenuti considerando uno spazio euclideo e sullo studio della sua topologia, siccome molti concetti dell'Analisi 1 devono essere concettualmente rivisti per poter essere generalizzati a spazi di dimensione diversa da 1. Nell'utilizzo a dimensioni maggiore di 1 siamo quasi "obbligati" a fare riferimento ad alcuni strumenti presi dal corso di Geometria per sviluppare correttamente la teoria, pertanto per una trattazione più accurata rimando ai libri più celebri riguardo a Geometria come [1] oppure [2].

1.1 Spazio euclideo e prodotto scalare

Partiamo ricordando al lettore la definizione di prodotto scalare/hermitiano e di **spazio euclideo**:

Definizione 1.1 (prodotto scalare/hermitiano). dato V spazio vettoriale sul campo \mathbb{R} (\mathbb{C}) si definisce prodotto scalare (hermitiano) un'applicazione $\varphi : V \times V \rightarrow \mathbb{R}$ ($\varphi : V \times V \rightarrow \mathbb{C}$) che gode delle seguenti proprietà:

- (1) lineare nella prima componente, ovvero

$$\forall \underline{v}, \underline{w}, \underline{z}, \forall \alpha, \beta \in \mathbb{R}, \varphi(\alpha \underline{v} + \beta \underline{w}, \underline{z}) = \alpha \varphi(\underline{v}, \underline{z}) + \beta \varphi(\underline{w}, \underline{z})$$

- (2) lineare nella seconda componente, ovvero

$$\forall \underline{v}, \underline{w}, \underline{z}, \forall \alpha, \beta \in \mathbb{R}, \varphi(\underline{v}, \alpha \underline{w} + \beta \underline{z}) = \alpha \varphi(\underline{v}, \underline{w}) + \beta \varphi(\underline{v}, \underline{z})$$

- (3) simmetrica, ovvero

$$\forall \underline{v}, \underline{w} \in V, \varphi(\underline{v}, \underline{w}) = \varphi(\underline{w}, \underline{v})$$

(sostituendo \mathbb{C} al posto di \mathbb{R} si ottengono le proprietà che rendono identificano un prodotto hermitiano)

Definizione 1.2 (spazio euclideo). uno spazio vettoriale V munito di un prodotto scalare $\varphi : V \times V \rightarrow \mathbb{R}$ (analogamente nel caso di un prodotto hermitiano) definito positivo ($\forall \underline{v} \neq \underline{0}, \varphi(\underline{v}, \underline{v}) > 0$) si dice **spazio euclideo**

Oss.: Quando un prodotto scalare/hermitiano è definito positivo, diciamo che è **coercivo**. Durante questo documento capiterà spesso di riferirci a questa proprietà con questo termine

Introduciamo qualche notazione a noi utile:

- E_n : spazio affine euclideo di dimensione finita n ;
- \mathbb{E}_n : spazio vettoriale euclideo, ottenuto fissando un'origine in E_n ;

Sappiamo, dal corso di Geometria, che, fissando una base \mathcal{B} di \mathbb{E}_n , le coordinate di un generico vettore $\underline{x} \in \mathbb{E}_n$ sono univoche e definendo la funzione $\varphi : \mathbb{E}_n \rightarrow \mathbb{R}^n$ tale che $\underline{x} \in \mathbb{E}_n \mapsto [\underline{x}]_{\mathcal{B}}$, dove con $[\underline{x}]_{\mathcal{B}}$ indichiamo il vettore delle coordinate del vettore \underline{x} rispetto alla base \mathcal{B} nello spazio \mathbb{E}_n .

Per comodità, noi vorremmo che questa base fosse anche ortonormale per semplificare la trattazione (l'esistenza è garantita, ovviamente, dal teorema di Lagrange).

Definizione 1.3 (base ortonormale). sia $\mathcal{B} = \{\underline{v}_1, \dots, \underline{v}_n\}$, diciamo che essa è una base ortonormale se

$$\langle \underline{v}_i, \underline{v}_j \rangle = \delta_{ij} = \begin{cases} 1 & i = j \\ 0 & i \neq j \end{cases}$$

Oss.: Se con la norma indotta dal prodotto scalare lo spazio è completo, allora \mathbb{E}_n è uno spazio di Hilbert.

Indichiamo adesso fissata la base ortonormale $(\underline{e}_1, \dots, \underline{e}_n)$. Osserviamo che, dati $\underline{v}, \underline{w}$, allora

$$\langle \underline{v}, \underline{w} \rangle = \left\langle \sum_{j=1}^n x_j \underline{v}_j, \sum_{i=1}^n y_i \underline{v}_i \right\rangle = \sum_{j=1}^n \sum_{i=1}^n x_j y_i \langle \underline{v}_j, \underline{v}_i \rangle = \sum_{j=1}^n \sum_{i=1}^n x_j y_i \delta_{ij} = \sum_{i=1}^n x_i y_i$$

Definizione 1.4 (norma). Sia V uno spazio vettoriale reale (o complesso). Si definisce norma un'applicazione $|\cdot| : V \rightarrow \mathbb{R}$ ($|\cdot| : V \rightarrow \mathbb{C}$) che verifica le seguenti condizioni:

- ① $|\underline{v}| \geq 0 \quad \forall \underline{v} \in V$
- ② $|\underline{v}| = 0 \iff \underline{v} = \underline{0}$
- ③ $|\lambda \underline{v}| = |\lambda| |\underline{v}| \quad \forall \lambda \in \mathbb{R}, \forall \underline{v} \in V$
- ④ $|\underline{v} + \underline{w}| \leq |\underline{v}| + |\underline{w}| \quad \forall \underline{v}, \underline{w} \in V$

(analogamente per uno spazio definito su \mathbb{C} sostituendo \mathbb{C} a \mathbb{R})

Notazione: la norma di un vettore verrà indicata all'interno di questo documento, per comodità, sia con $|\cdot|$ e sia con $\|\cdot\|$

Oss.: Osserviamo che se (V, φ) è uno spazio euclideo allora $|\cdot| = \sqrt{\varphi(\cdot, \cdot)}$ è una norma: il prodotto scalare induce una norma su V . Mostreremo più avanti il punto ④ (ovvero la proprietà meno banale fra quelle) usando la norma indotta dal prodotto scalare

Sapendo che $\mathbb{E}_n \simeq \mathbb{R}^n$ (quando quest'ultimo è munito di un prodotto scalare definito positivo naturalmente) possiamo effettuare tutte le dimostrazioni in \mathbb{R}^n e queste saranno naturalmente valide in tutti gli spazi euclidei di dimensione finita n .

Teorema 1.1. Sia (V, φ) uno spazio euclideo. Allora

$$\forall \underline{v}, \underline{w} \in V, \langle \underline{v}, \underline{w} \rangle = |\underline{v}| |\underline{w}| \cos \hat{\theta}$$

dove $\hat{\theta}$ è l'angolo convesso fra i due vettori \underline{v} e \underline{w}

Dimostrazione. consideriamo due vettori $\underline{v} \in \mathbb{E}_2$ con $|\underline{v}| = |\underline{w}| = 1$. Prendiamo per semplicità $\underline{v} = \underline{e}_1$ e si osserva che

$$\langle \underline{e}_1, \underline{w} \rangle = \cos \hat{\theta}$$

Questo segue banalmente dall'interpretazione geometrica del prodotto scalare canonico.

Per estendere la validità di questo risultato a tutti i $\underline{v} \neq \underline{e}_1$ si osserva che $\exists R \in \text{SO}(\mathbb{E}_2) : R\underline{e}_1 = \underline{v}$ e dunque, considerando $R^{-1}(\underline{w})$ (l'esistenza di un'inversa è garantita dal fatto che $R \in \text{SO}(2)$) e sappiamo che:

$$\langle \underline{v}, \text{Id}(\underline{w}) \rangle = \langle R\underline{e}_1, (R \circ R^{-1})\underline{w} \rangle \stackrel{R \text{ è isometria}}{=} \langle \underline{e}_1, R^{-1}(\underline{w}) \rangle = \cos \hat{\theta}$$

A questo punto, dati due vettori qualunque \underline{v} e \underline{w} non di norma unitaria, possiamo utilizzare il ragionamento precedente osservando che:

$$\left\langle \frac{\underline{v}}{|\underline{v}|}, \frac{\underline{w}}{|\underline{w}|} \right\rangle = \cos \hat{\theta}$$

ma allora, si osserva che:

$$\frac{1}{|\underline{v}|} \cdot \frac{1}{|\underline{w}|} \langle \underline{v}, \underline{w} \rangle = \cos \hat{\theta} \implies \langle \underline{v}, \underline{w} \rangle = |\underline{v}| |\underline{w}| \cos \hat{\theta}$$

dunque la tesi (in \mathbb{E}_2). Per generalizzare questo concetto a qualunque spazio, noi sappiamo che possiamo considerare il piano $\pi = \text{Span}(\underline{v}, \underline{w})$ e considerare il loro angolo $\hat{\theta}$ convesso giacente in questo piano. \square

Proposizione 1.1 (Disuguaglianza di Cauchy-Schwarz).

$$\forall \underline{v}, \underline{w} \in \mathbb{E}_n, |\langle \underline{v}, \underline{w} \rangle| \leq |\underline{v}| |\underline{w}|$$

Dimostrazione. Consideriamo $\lambda \in \mathbb{R}$ e sappiamo che, per coercività del prodotto scalare, che:

$$\forall \underline{v}, \underline{w} \in \mathbb{E}_n, \langle \underline{v} + \lambda \underline{w}, \underline{v} + \lambda \underline{w} \rangle > 0$$

ma per bilinearità del prodotto scalare abbiamo che

$$\langle \underline{v} + \lambda \underline{w}, \underline{v} + \lambda \underline{w} \rangle = \langle \underline{v}, \underline{v} \rangle + 2\lambda \langle \underline{v}, \underline{w} \rangle + \lambda^2 \langle \underline{w}, \underline{w} \rangle > 0$$

dunque l'equazione in λ

$$\langle \underline{v}, \underline{v} \rangle + 2\lambda \langle \underline{v}, \underline{w} \rangle + \lambda^2 \langle \underline{w}, \underline{w} \rangle = 0$$

non deve avere soluzione, il che implica che

$$\Delta = 4\langle \underline{v}, \underline{w} \rangle^2 - 4\langle \underline{v}, \underline{v} \rangle \langle \underline{w}, \underline{w} \rangle < 0 \implies \langle \underline{v}, \underline{w} \rangle^2 < \langle \underline{v}, \underline{v} \rangle \langle \underline{w}, \underline{w} \rangle \implies |\langle \underline{v}, \underline{w} \rangle| < |\underline{v}| |\underline{w}|$$

\square

Proposizione 1.2 (Disuguaglianza triangolare). $\forall \underline{v}, \underline{w} \in \mathbb{E}_n$,

$$\textcircled{1} \quad |\underline{v} + \underline{w}| \leq |\underline{v}| + |\underline{w}|$$

$$\textcircled{2} \quad |\underline{v}| - |\underline{w}| \leq |\underline{v} - \underline{w}|$$

Dimostrazione. osserviamo che

$$|\underline{v} + \underline{w}|^2 = \langle \underline{v} + \underline{w}, \underline{v} + \underline{w} \rangle = \langle \underline{v}, \underline{v} \rangle + 2\langle \underline{v}, \underline{w} \rangle + \langle \underline{w}, \underline{w} \rangle = |\underline{v}|^2 + 2\langle \underline{v}, \underline{w} \rangle + |\underline{w}|^2$$

Osserviamo che possiamo utilizzare la disegualanza di Cauchy-Schwarz per minorare l'ultima espressione ottenuta, da cui si deduce che

$$|\underline{v} + \underline{w}|^2 \leq |\underline{v}|^2 + 2|\underline{v}||\underline{w}| + |\underline{w}|^2$$

in cui riconosciamo che $|\underline{v}|^2 + 2|\underline{v}||\underline{w}| + |\underline{w}|^2 = (|\underline{v}| + |\underline{w}|)^2$, dunque

$$|\underline{v} + \underline{w}|^2 \leq (|\underline{v}| + |\underline{w}|)^2 \implies |\underline{v} + \underline{w}| \leq |\underline{v}| + |\underline{w}|$$

e si ottiene la tesi del punto ①.

Il punto ② si ottiene come corollario del primo punto osservando che

$$|\underline{v}| = |\underline{v} + \underline{w} - \underline{w}| \leq |\underline{v} - \underline{w}| + |\underline{w}| \implies |\underline{v}| - |\underline{w}| \leq |\underline{v} - \underline{w}|$$

ma ragionando in maniera identica sul vettore \underline{w} si osserva che

$$|\underline{w}| = |\underline{w} - \underline{v} + \underline{v}| \leq |\underline{w} - \underline{v}| + |\underline{v}| \implies |\underline{w}| - |\underline{v}| \leq |\underline{w} - \underline{v}| = |\underline{v} - \underline{w}|$$

dunque possiamo concludere che

$$|\underline{v}| - |\underline{w}| \leq |\underline{v} - \underline{w}|$$

ottenendo la tesi □

Oss. Come avevo detto in una osservazione, ogni prodotto scalare definito positivo induce sempre una norma. La dimostrazione qua sopra non fa uso di nessuna proprietà specifiche del prodotto scalare canonico di \mathbb{R}^n , dunque può essere usata per ogni norma indotta dal prodotto scalare definito positivo di qualunque spazio euclideo

1.2 Il prodotto vettoriale

In \mathbb{R}^3 (ma in generale in qualunque spazio euclideo di dimensione 3) è anche possibile definire l'operazione di prodotto vettoriale, molto utile per trattare (come vedremo più avanti) l'orientazione delle superfici.

Fissando una base ortonormale $\{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ di \mathbb{E}_3 , definendo questa operazione $\times : V \times V \rightarrow V$ assegnando i prodotti elementari secondo l'invarianza per permutazioni cicliche, ponendo che

$$\underline{e}_1 \times \underline{e}_2 = \underline{e}_3$$

e, per invarianza per permutazioni cicliche, dovremo avere che

$$\begin{aligned} \underline{e}_3 \times \underline{e}_1 &= \underline{e}_2 \\ \underline{e}_2 \times \underline{e}_3 &= \underline{e}_1 \end{aligned}$$

Le permutazioni non cicliche invece fanno variare il segno dunque avremo, per il vettore \underline{e}_1

$$\begin{cases} \underline{e}_1 \times \underline{e}_2 = \underline{e}_3 \\ \underline{e}_1 \times \underline{e}_3 = -\underline{e}_2 \\ \underline{e}_2 \times \underline{e}_3 = \underline{e}_1 \end{cases} \quad (1.1)$$

Le regole che abbiamo visto possono essere anche facilmente ottenute tramite la cosiddetta *regola della mano destra*: indicando la direzione del primo vettore con il pollice (in questo caso \underline{e}_1) e con l'indice il secondo vettore (in questo caso \underline{e}_2), ottenendo sul pollice la direzione del terzo vettore. Vogliamo inoltre che questo prodotto vettoriale sia *bilineare*.

Dati adesso due vettori \underline{v} e $\underline{w} \in E_3$ abbiamo che $\underline{v}, \underline{w} \in \text{Span}(\underline{e}_1, \underline{e}_2, \underline{e}_3)$ dunque avremo che:

$$\underline{v} = x_1 \underline{e}_1 + x_2 \underline{e}_2 + x_3 \underline{e}_3 \quad \underline{w} = y_1 \underline{e}_1 + y_2 \underline{e}_2 + y_3 \underline{e}_3$$

e studiamo quali saranno le coordinate del prodotto vettoriale fra questi due:

$$\begin{aligned} \underline{v} \times \underline{w} &= (x_1 \underline{e}_1 + x_2 \underline{e}_2 + x_3 \underline{e}_3) \times (y_1 \underline{e}_1 + y_2 \underline{e}_2 + y_3 \underline{e}_3) = \\ &= (x_1 y_2 - x_2 y_1)(\underline{e}_1 \times \underline{e}_2) + (x_1 y_3 - x_3 y_1)(\underline{e}_1 \times \underline{e}_3) + (x_2 y_3 - x_3 y_2)(\underline{e}_2 \times \underline{e}_3) \end{aligned}$$

dunque definiamo una matrice C di questa forma:

$$C = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{pmatrix} \quad (1.2)$$

e ne consideriamo i minori:

$$M_{ij}(C) = \det \begin{pmatrix} x_i & y_i \\ x_j & y_j \end{pmatrix} = x_i y_j - x_j y_i \quad (1.3)$$

dunque otteniamo una nuova formula:

$$\underline{v} \times \underline{w} = M_{12}(C)(\underline{e}_1 \times \underline{e}_2) + M_{23}(C)(\underline{e}_2 \times \underline{e}_3) + M_{13}(C)(\underline{e}_1 \times \underline{e}_3) = M_{12}(C)\underline{e}_3 + M_{23}(C)\underline{e}_1 - M_{13}(C)\underline{e}_2$$

usando le proprietà del determinante, sappiamo che $M_{13}(C) = -M_{31}(C)$ dunque

$$\underline{v} \times \underline{w} = M_{12}(C)\underline{e}_3 + M_{23}(C)\underline{e}_1 + M_{31}(C)\underline{e}_2$$

abbiamo dunque dimostrato la seguente proposizione

Proposizione 1.3. dati i vettori $\underline{v}, \underline{w} \in E_3$ allora $\exists x_1, \dots, x_3, y_1, \dots, y_3 \in \mathbb{R}$ tali che $\underline{v} = x_1 \underline{e}_1 + x_2 \underline{e}_2 + x_3 \underline{e}_3$ e $\underline{w} = y_1 \underline{e}_1 + y_2 \underline{e}_2 + y_3 \underline{e}_3$. Posta la matrice C tale che

$$C = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{pmatrix}$$

allora

$$\underline{v} \times \underline{w} = M_{23}(C)\underline{e}_1 + M_{31}(C)\underline{e}_2 + M_{12}(C)\underline{e}_3 \quad (1.4)$$

Oss.: si noti come i numeri siano tutti disposti secondo permutazioni cicliche: questo può essere un buon trucco per ricordarsela.

Corollario 1.1.1 (matrice del prodotto vettoriale). siano $\underline{v}, \underline{w} \in E_3$ e siano x_1, \dots, x_3 e y_1, \dots, y_3 le rispettive coordinate rispetto alla base ortonormale $\mathcal{B} = \{\underline{e}_1, \underline{e}_2, \underline{e}_3\}$ di E_3 . Allora

$$\underline{v} \times \underline{w} = \det \begin{pmatrix} x_1 & y_1 & \underline{e}_1 \\ x_2 & y_2 & \underline{e}_2 \\ x_3 & y_3 & \underline{e}_3 \end{pmatrix} \quad (1.5)$$

Dimostrazione. si osservi che, procedendo con uno sviluppo di Laplace lungo la terza colonna abbiamo che:

$$\det \begin{pmatrix} x_1 & y_1 & \underline{e}_1 \\ x_2 & y_2 & \underline{e}_2 \\ x_3 & y_3 & \underline{e}_3 \end{pmatrix} = \underline{e}_1 M_{23}(C) - \underline{e}_2 M_{13}(C) + \underline{e}_3 M_{12}(C) = M_{23}(C)\underline{e}_1 + M_{31}(C)\underline{e}_2 + M_{12}(C)\underline{e}_3$$

□

Proposizione 1.4 (matrice del prodotto misto). siano dati i vettori $\underline{v} = x_1\underline{e}_1 + x_2\underline{e}_2 + x_3\underline{e}_3$ e $\underline{w} = y_1\underline{e}_1 + y_2\underline{e}_2 + y_3\underline{e}_3$ e $\underline{z} = \xi_1\underline{e}_1 + \xi_2\underline{e}_2 + \xi_3\underline{e}_3$. Allora

$$\langle \underline{v} \times \underline{w}, \underline{z} \rangle = \det \begin{pmatrix} x_1 & y_1 & \xi_1 \\ x_2 & y_2 & \xi_2 \\ x_3 & y_3 & \xi_3 \end{pmatrix} \quad (1.6)$$

Dimostrazione. sappiamo che $\underline{v} \times \underline{w} = M_{23}(C)\underline{e}_1 + M_{31}(C)\underline{e}_2 + M_{12}(C)\underline{e}_3$ dunque $\langle \underline{v} \times \underline{w}, \underline{z} \rangle = M_{23}(C)\xi_1 + M_{31}(C)\xi_2 + M_{12}(C)\xi_3$. D'altra parte definendo la matrice

$$M = \begin{pmatrix} x_1 & y_1 & \xi_1 \\ x_2 & y_2 & \xi_2 \\ x_3 & y_3 & \xi_3 \end{pmatrix}$$

allora si osserva che

$$\langle \underline{v} \times \underline{w}, \underline{z} \rangle = \text{cof}_{13}(M)\xi_1 + \text{cof}_{23}(M)\xi_2 + \text{cof}_{33}(M)\xi_3 = \det(M)$$

dunque la tesi. □

Corollario 1.1.2 ($\underline{v} \times \underline{w} \perp \text{Span}(\underline{v}, \underline{w})$). Dati due vettori $\underline{v}, \underline{w} \in \mathbb{E}_3$ allora

$$\underline{v} \times \underline{w} \perp \text{Span}(\underline{v}, \underline{w})$$

Dimostrazione. dalla proposizione 1.4 sappiamo che

$$\langle \underline{v} \times \underline{w}, \underline{v} \rangle = \det \begin{pmatrix} x_1 & y_1 & x_1 \\ x_2 & y_2 & x_2 \\ x_3 & y_3 & x_3 \end{pmatrix} = 0$$

per proprietà del determinante. Similmente per il vettore \underline{w} . Dunque abbiamo che $\underline{v} \times \underline{w} \perp \underline{v}$ e $\underline{v} \times \underline{w} \perp \underline{w}$ che implica che $\underline{v} \times \underline{w}$ è perpendicolare a qualunque combinazione lineare dei vettori \underline{v} e \underline{w} , quindi $\underline{v} \times \underline{w} \perp \text{Span}(\underline{v}, \underline{w})$ ovvero la tesi. □

Corollario 1.1.3 (norma del prodotto vettoriale). Siano $\underline{v}, \underline{w} \in \mathbb{E}_3$ con $\underline{v} = x_1\underline{e}_1 + x_2\underline{e}_2 + x_3\underline{e}_3$ e $\underline{w} = y_1\underline{e}_1 + y_2\underline{e}_2 + y_3\underline{e}_3$. Allora, ponendo

$$C = \begin{pmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{pmatrix}$$

allora

$$|\underline{v} \times \underline{w}| = \sqrt{M_{23}(C)^2 + M_{12}(C)^2 + M_{31}(C)^2}$$

Dimostrazione. Osserviamo, per la proposizione 1.4, che

$$\begin{aligned}\langle \underline{v} \times \underline{w}, \underline{v} \times \underline{w} \rangle &= \det(\underline{v} \ \underline{w} \ \underline{v} \times \underline{w}) = M_{23}(C)(\underline{v} \times \underline{w})_1 + M_{31}(C)(\underline{v} \times \underline{w})_2 + M_{12}(C)(\underline{v} \times \underline{w})_3 = \\ &= M_{23}(C)^2 + M_{31}(C)^2 + M_{12}(C)^2\end{aligned}$$

dunque $|\underline{v} \times \underline{w}| = \sqrt{M_{23}(C)^2 + M_{31}(C)^2 + M_{12}(C)^2}$

□

Proposizione 1.5. Sia $R \in \text{SO}(3)$ e siano dati $\underline{v}, \underline{w} \in \mathbb{E}_3$. Allora

$$R(\underline{v} \times \underline{w}) = R\underline{v} \times R\underline{w} \quad (1.7)$$

Dimostrazione. Sia $\underline{u} \in \mathbb{E}_3$ e consideriamo il seguente prodotto scalare:

$$\begin{aligned}\langle R\underline{v} \times R\underline{w}, R\underline{u} \rangle &= \det(R\underline{v} \ R\underline{w} \ R\underline{u}) \\ &= \det(R(\underline{v} \ \underline{w} \ \underline{u}))\end{aligned}$$

Usando il teorema di Binet abbiamo che

$$\langle R\underline{v} \times R\underline{w}, R\underline{u} \rangle = \det(R(\underline{v} \ \underline{w} \ \underline{u})) = \det(R)\det(\underline{v} \ \underline{w} \ \underline{u}) = \det(\underline{v} \ \underline{w} \ \underline{u})$$

dove abbiamo usato l'ipotesi che $R \in \text{SO}(3)$ dunque $\det(R) = 1$ e siccome sappiamo che R è un'isometria allora

$$\langle R(\underline{v} \times \underline{w}), R\underline{u} \rangle = \det(\underline{v} \ \underline{w} \ \underline{u}) = \langle \underline{v} \times \underline{w}, \underline{u} \rangle = \langle R(\underline{v} \times \underline{w}), R\underline{u} \rangle$$

ma allora $\langle R\underline{v} \times R\underline{w}, R\underline{u} \rangle = \langle R(\underline{v} \times \underline{w}), R\underline{u} \rangle \implies \langle R(\underline{v} \times \underline{w}) - (R\underline{v} \times R\underline{w}), R\underline{u} \rangle = 0$. Ma allora, per la coercività del prodotto scalare, abbiamo necessariamente che

$$R(\underline{v} \times \underline{w}) = R\underline{v} \times R\underline{w}$$

□

Proposizione 1.6. Se $\underline{v}, \underline{w} \in \mathbb{R}^3$, $\underline{v} \perp \underline{w}$ e $|\underline{v}| = |\underline{w}| = 1$ allora

$$|\underline{v} \times \underline{w}| = 1$$

Dimostrazione. Dati \underline{v} e \underline{w} con $\underline{v} \perp \underline{w}$, allora $\{\underline{v}, \underline{w}, \underline{v} \times \underline{w}\}$ è una base. Siccome appartengono a \mathbb{R}^3 e sono perpendicolari, allora sappiamo che $\exists R \in \text{SO}(3) : R\underline{v} = \underline{e}_1, R\underline{w} = \underline{e}_2$ e $R(\underline{v} \times \underline{w}) = \underline{e}_3$ dunque

$$|\underline{v} \times \underline{w}| = 1 = |R\underline{e}_1 \times R\underline{e}_2| = |R(\underline{e}_1 \times \underline{e}_2)| = |\underline{v} \times \underline{w}| \implies |\underline{v} \times \underline{w}| = 1$$

□

Corollario 1.1.4 (modulo del prodotto vettoriale di vettori perpendicolari). siano $\underline{v}, \underline{w} \in \mathbb{E}_3$ e $\underline{v} \perp \underline{w}$ allora:

$$|\underline{v} \times \underline{w}| = |\underline{v}| |\underline{w}|$$

Dimostrazione. siccome $\underline{v} \perp \underline{w}$ allora $\frac{\underline{v}}{|\underline{v}|} \perp \frac{\underline{w}}{|\underline{w}|}$. Ma allora possiamo applicare a questi due vettori la proposizione precedente, dunque:

$$\left| \frac{\underline{v}}{|\underline{v}|} \times \frac{\underline{w}}{|\underline{w}|} \right| = 1 = \frac{1}{|\underline{v}| |\underline{w}|} |\underline{v} \times \underline{w}| \implies |\underline{v} \times \underline{w}| = |\underline{v}| |\underline{w}|$$

□

Proposizione 1.7 (area del parallelogramma). Siano $\underline{v}, \underline{w} \in \mathbb{E}_3$. Allora

$$|\underline{v} \times \underline{w}| = |\underline{v}| |\underline{w}| \sin \hat{\theta}$$

dove $\hat{\theta}$ è l'angolo convesso fra i due vettori

Dimostrazione. possiamo "ortogonalizzare" il vettore \underline{v} usando il procedimento di Grand-Schmit, dunque

$$\begin{aligned} |\underline{v} \times \underline{w}|^2 &= \left| \left(\underline{v} - \frac{\langle \underline{v}, \underline{w} \rangle}{|\underline{w}|^2} \underline{w} \right) \times \underline{w} \right|^2 = \left| \underline{v} - \frac{\langle \underline{v}, \underline{w} \rangle}{|\underline{w}|^2} \underline{w} \right|^2 |\underline{w}|^2 = \langle \underline{v} - \frac{\langle \underline{v}, \underline{w} \rangle}{|\underline{w}|^2} \underline{w}, \underline{v} - \frac{\langle \underline{v}, \underline{w} \rangle}{|\underline{w}|^2} \underline{w} \rangle |\underline{w}|^2 = \\ &= |\underline{w}|^2 \left(\langle \underline{v}, \underline{v} \rangle - 2 \langle \underline{v}, \frac{\langle \underline{v}, \underline{w} \rangle}{|\underline{w}|^2} \underline{w} \rangle + \frac{(\langle \underline{v}, \underline{w} \rangle)^2}{|\underline{w}|^2} \langle \underline{w}, \underline{w} \rangle \right) = \\ &= |\underline{w}|^2 \left(\langle \underline{v}, \underline{v} \rangle - 2 \frac{\langle \underline{v}, \underline{w} \rangle}{|\underline{w}|^2} \langle \underline{v}, \underline{w} \rangle + \frac{(\langle \underline{v}, \underline{w} \rangle)^2}{|\underline{w}|^2} \langle \underline{w}, \underline{w} \rangle \right) = \\ &= |\underline{w}|^2 \left(\langle \underline{v}, \underline{v} \rangle - \frac{(\langle \underline{v}, \underline{w} \rangle)^2}{|\underline{w}|^2} \right) \end{aligned}$$

Sappiamo adesso che $\langle \underline{v}, \underline{w} \rangle = |\underline{v}| |\underline{w}| \cos \hat{\theta}$, dunque

$$|\underline{w}|^2 \left(|\underline{v}|^2 - |\underline{v}|^2 |\underline{w}|^2 \cos^2 \hat{\theta} \right) = |\underline{v}|^2 |\underline{w}|^2 (1 - \cos^2 \hat{\theta}) = |\underline{v}|^2 |\underline{w}|^2 \sin^2 \hat{\theta}$$

ma allora se ne conclude che

$$|\underline{v} \times \underline{w}|^2 = |\underline{v}|^2 |\underline{w}|^2 \sin^2 \hat{\theta} \implies |\underline{v} \times \underline{w}| = |\underline{v}| |\underline{w}| \sin \hat{\theta}$$

e questo conclude la dimostrazione. \square

Oss. Unendo quest'ultima proposizione con il corollario 1.1.3 possiamo facilmente vedere che

$$|\underline{v} \times \underline{w}| = |\underline{v}| |\underline{w}| \sin \hat{\theta} = \sqrt{M_{23}(C)^2 + M_{12}(C)^2 + M_{31}(C)^2}$$

Definizione 1.5 (base positivamente orientata). diremo che tre vettori $\underline{v}, \underline{w}, \underline{z} \in \mathbb{E}_3$ linearmente indipendenti sono una base positivamente orientata se

$$\det(\underline{v} \ \underline{w} \ \underline{z}) > 0$$

Oss. La base canonica di \mathbb{R}^3 è positivamente orientata, siccome

$$\det \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \det(I) = 1$$

Proposizione 1.8. Se $\underline{v}, \underline{w} \in \mathbb{R}^3$ sono linearmente indipendenti, allora l'insieme $\mathcal{B} = \{\underline{v}, \underline{w}, \underline{v} \times \underline{w}\}$ è una base positivamente orientata

Dimostrazione. banalmente

$$\det(\underline{v} \ \underline{w} \ \underline{v} \times \underline{w}) = \langle \underline{v} \times \underline{w}, \underline{v} \times \underline{w} \rangle = |\underline{v} \times \underline{w}|^2 > 0$$

\square

Oss.: Una base positivamente orientata si può sempre rappresentare con la regola della mano destra: il vettore \underline{v} è rappresentato dall'indice, il vettore \underline{w} è rappresentato dal dito medio e il pollice invece rappresenta $\underline{v} \times \underline{w}$.

Infine, per concludere per adesso la lista dei teoremi sul prodotto vettoriale, abbiamo che:

Proposizione 1.9. Una base ortonormale $\{\underline{v}, \underline{w}, \underline{v} \times \underline{w}\}$ è positivamente orientata se e solo se $\underline{z} = \underline{v} \times \underline{w}$

Dimostrazione.

\Leftarrow : già mostrato per la proposizione 1.8

\Rightarrow : supponiamo di avere la seguente base ortonormale positivamente orientata $\{\underline{i}, \underline{j}, \underline{k}\}$, allora abbiamo che questa segue la regola della mano destra, dunque $\underline{i} \times \underline{j} = \alpha \underline{k}$ con $\alpha \in \mathbb{R}$. Ma allora, per la proposizione 1.6, abbiamo che $|\alpha \underline{k}| = 1$ (siccome $\underline{v} \perp \underline{w}$) dunque $\underline{i} \times \underline{j} = \underline{k}$. \square

1.3 Topologia di \mathbb{R}^n

Come visto dal corso di Analisi Matematica del primo anno, possiamo pensare \mathbb{R}^n come uno spazio metrico se consideriamo la distanza indotta dal prodotto scalare. Non introdurremo in questa sezione "tutta" la topologia necessaria per l'Analisi siccome sarebbe, secondo la mia opinione, troppo dispersivo siccome rallenterebbe lo studio; dunque mi sono focalizzato nel mostrare una buona parte delle definizioni e nel fornire degli esercizi per impraticarsi. In ogni caso, per coloro che sono interessati ad una trattazione più specifica nella materia (pur conservando un approccio rivolto verso l'applicazione all'Analisi matematica) consiglio di sfogliare l'eccellente capitolo presente in [3] o nel [4] oppure un buon libro di Topologia, di cui consiglio [5] oppure [6].

Possiamo dunque andare a definire le **palle**:

Definizione 1.6 (palle aperte e chiuse). Sia $x \in \mathbb{R}^n$ e fissiamo $r > 0$. Allora definiamo la palla aperta di raggio r nella seguente maniera

$$B(x, r) = \{z \in \mathbb{R}^n : |z - x| < r\}$$

e la palla chiusa di raggio r nella seguente maniera

$$\mathbb{B}(x, r) = \{z \in \mathbb{R}^n : |z - x| \leq r\}$$

con centro x e raggio r .

Ma cosa vuol dire palla aperta? In generale possiamo definire la nozione di insieme aperto nella seguente maniera:

Definizione 1.7 (insieme aperto). $A \subset \mathbb{R}^n$ è un insieme aperto se $\forall y \in A \exists r > 0 : B(y, r_y) \subset A$

Esercizio 1.1. Mostrare che $B(x, r)$ è un aperto $\forall x \forall r$

Svolgimento. Sia $x \in \mathbb{R}^n$ e sia fissato $r > 0$. Allora scelto $y \in B(x, r)$ sappiamo che preso $\varepsilon = r - |x - y| > 0$ allora possiamo definire la palla $B(y, \varepsilon)$ e prendere $z \in B(y, \varepsilon)$ e osservare che

$$|z - x| \leq |z - y| + |y - x| < r - |x - y| + |x - y| = r$$

dunque $|z - x| < r$ il che implica che $B(y, \varepsilon) \subset B(x, r)$ \square

Esercizio 1.2. Siano $\{A_j : j \in J\}$ una famiglia di aperti (non necessariamente numerabile), allora $\bigcup_{j \in J} A_j$ è un insieme aperto.

Svolgimento. si osserva che, siccome A_j è un insieme aperto $\forall j$, allora $\forall a \in A_j \exists r > 0 : B(a, r) \subset A_j$, ma allora $B(a, r) \subset \bigcup_{j \in J} A_j$, dunque $\forall a \in \bigcup_{j \in J} A_j \exists r > 0 : B(a, r) \subset \bigcup_{j \in J} A_j$ \square

Esercizio 1.3. Siano $A_1, \dots, A_k \subseteq \mathbb{R}^n$ aperti, allora $\bigcap_i^k A_i$ è un aperto

Dimostrazione. Sia $x \in \bigcap_i^k A_i \implies \forall i \in \{1, \dots, n\}, x \in A_i$ e, siccome per ipotesi sono aperti, avremo che $\forall i \exists r_i(x) : B(x, r_i(x)) \subseteq A_i$, dunque preso $r = \min_i \{r_i(x) : i \in \{1, \dots, k\}\}$ avremo che $\forall i \in \{1, \dots, k\}, B(x, r) \subseteq A_i \implies B(x, r) \subseteq \bigcup_i^k A_i$, dunque $\forall x \in \bigcap_i^k A_i \exists r > 0 : B(x, r) \subseteq \bigcap_i^k A_i$, dunque è aperto per definizione. \square

Esercizio 1.4. Sia $A = \{(x, y) \in \mathbb{R}^2 : 1 < x^2 + y^2 < 4\} = B(0, 2) \setminus \mathbb{B}(0, 1)$, provare che A è aperto

Dimostrazione. Fissiamo $x \in A$ e prendiamo come $r = \min\{4 - |x|, |x| - 1\}$ avremo che

$$|y| \leq |y - x + x| \leq |y - x| + |x| < |x| + r \stackrel{r \leq 4 - |x|^2}{\leq} |x| + 4 - |x| = 4$$

A questo punto osserviamo che:

$$|y| \geq -|x - y| + |x| \geq |x| - r > 1$$

dunque in conclusione abbiamo che $\forall y \in B(x, r), 1 < |y| < 4$. Avremo quindi che A è aperto. \square

Esercizio 1.5. Provare che $A = \{(x, y, z) : -1 < x < 2\} = (-1, 2) \times \mathbb{R}^2$ ed è un insieme aperto in \mathbb{R}^3

Dimostrazione.

Mostriamo che $A \subseteq (-1, 2) \times \mathbb{R}^2$: osserviamo che, per definizione di A , $-1 < x < 2$ e, non essendo altri vincoli, $y, z \in \mathbb{R}^2$; pertanto $A \subseteq (-1, 2) \times \mathbb{R}^2$.

Mostriamo che $(-1, 2) \times \mathbb{R}^2 \subseteq A$: osserviamo che, per definizione di prodotto cartesiano, preso $(x, y, z) \in \mathbb{R}^3$ allora $x \in (-1, 2) \implies -1 < x < 2$ dunque $(-1, 2) \times \mathbb{R}^2 \subseteq A$.

Dimostriamo che è aperto: preso $x' = (x, y, z) \in A$ allora possiamo considerare $r = \min\{x + 1, 2 - x\}$ avremo che preso $y \in B(x', r)$ allora, posto $y = (y_x, y_y, y_z)$ e $x' = (x, x'_y, x'_z)$, avremo che

$$|x - y_x| \leq |x' - y|^2 < r \implies |y_x - x| < r \implies y_x \in (x - r, x + r)$$

ma $x + r \leq x + 2 - x = 2 \implies x + r \leq 2$ e $x - r \geq x - x + 1 = 1$, dunque $y_x \in (-1, 2)$. \square

Esercizio 1.6. Provare che \mathbb{R}^n è aperto

Dimostrazione. Sia $x \in \mathbb{R}^n$, allora preso $r > 0$ e la palla $B(x, r)$ avremo che, per definizione, $\forall y \in B(x, r), y \in \mathbb{R}^n \implies B(x, r) \subseteq \mathbb{R}^n$. \square

Un concetto chiave sono i cosiddetti *punti di frontiera*, definiti come

Definizione 1.8. Sia $x \in \mathbb{R}^n$ e sia $A \subseteq \mathbb{R}^n$, diremo che x è un punto di frontiera per A se $\forall r > 0, B(x, r) \cap A \neq \emptyset$ e $B(x, r) \cap A^c \neq \emptyset$.

Definiamo

$$\text{Fr}(A) = \{x \in \mathbb{R}^n : x \text{di frontiera}\}$$

Esercizio 1.7. Provare che $\text{Fr } B(x, r) = \{y \in \mathbb{R}^n : |x - y| = r\}$ e $\text{Fr } \mathbb{B}(x, r) = \{y \in \mathbb{R}^n : |x - y| = r\}$

Dimostrazione. Mostriamo che $\text{Fr } B(x, r) = \{y \in \mathbb{R}^n : |x - y| = r\}$: possiamo osservare che $\forall z \in \{y \in \mathbb{R}^n : |x - y| = r\}, B(z, \varepsilon) \cap B(x, r) \neq \emptyset$, infatti possiamo considerare innanzitutto il segmento che congiunge il punto y e x

$$y = z + t(x - z), t \in [0, 1] \implies |x - y| = |(1 - t)(x - z)| \implies |x - y| = |1 - t|r$$

la condizione che $|x - y| < r$ si ottiene ponendo che $|x - y| < r$, ovvero $|1 - t|r < r \implies |1 - t| < 1$. Considerando a questo punto il segmento che congiunge z con y , avremo che

$$|z - y| < \varepsilon \implies |z - y| = |z - (z + t(x - z))| = |t(x - z)| = t|x - z| = tr < \varepsilon \implies t < \frac{\varepsilon}{r}$$

pertanto, rimettendo insieme le condizioni trovate, dobbiamo risolvere il seguente sistema di disequazioni:

$$\begin{cases} |1 - t| < 1 \implies 0 < t < 2 \\ t < \frac{\varepsilon}{r} \end{cases}$$

che è sempre vero $\forall t \in (0, \frac{\varepsilon}{r})$. Possiamo fare lo stesso per mostrare che $B(z, \varepsilon) \cap (B(x, r))^c$, dunque $\{y \in \mathbb{R}^n : |x - y| = r\} \subseteq \text{Fr } B(x, r)$. Mostriamo adesso che vale il viceversa mostrando che se $z \notin \{y \in \mathbb{R}^n : |x - y| = r\} \implies z \notin \text{Fr } B(x, r)$ (il che equivale a mostrare che $\text{Fr } B(x, r) \subseteq \{y \in \mathbb{R}^n : |x - y| = r\}$), siccome sono possibili due casi:

- $|x - z| < r \implies z \in B(x, r)$;
- $|x - z| > r \implies z \in (B(x, r))^c$

e non è possibile che $z \in \text{Fr } B(x, r)$ siccome se per assurdo $\forall r' > 0, B(z, r') \cap B(x, r) \neq \emptyset \wedge B(z, r') \cap (B(x, r))^c$ sappiamo che, dall'apertura di $B(x, r)$, $\exists r'' > 0 : B(z, r'') \subseteq A$ ma allora

$$(B(x, r))^c \cap B(z, r'') \subseteq B(x, r)$$

il che è un assurdo siccome $(B(x, r))^c \cap B(x, r) = \emptyset$.

Per la sfera chiusa è analoga. \square

Definizione 1.9 (punti di aderenza). Sia $A \neq \emptyset$ e $A \subseteq \mathbb{R}^n$, diremo che $x \in \mathbb{R}^n$ è un punto di aderenza o di chiusura di A se $\forall r > 0, B(x, r) \cap A \neq \emptyset$.

Inoltre indicheremo con \bar{A} l'insieme dei punti di aderenza di A .

Oss.: Naturalmente $A \subseteq \bar{A}$, siccome $\forall r > 0 : B(x, r) \cap A \supseteq \{x\} \neq \emptyset$

Definizione 1.10 (insieme chiuso). Un insieme F è chiuso se contiene tutti i suoi punti di aderenza, ovvero $F = \bar{F}$.

Esercizio 1.8. Provare che $\text{Fr } A = \bar{A} \cap \bar{A}^c$

Dimostrazione. Sia $x \in \text{Fr } A \implies \forall r > 0, B(x, r) \cap A \neq \emptyset \wedge B(x, r) \cap A^c \neq \emptyset$, ma allora $\forall r > 0, B(x, r) \cap A \neq \emptyset \implies x$ è un punto di aderenza di A , dunque $x \in \bar{A}$, e $\forall r > 0, B(x, r) \cap A^c \neq \emptyset \implies x$ è un punto di aderenza di A^c , pertanto $x \in \bar{A} \cap \bar{A}^c$. Sia $x \in \bar{A} \cap \bar{A}^c$, allora, per definizione, avremo che $\forall r > 0 : B(x, r) \cap A \neq \emptyset$ e $\forall r > 0 : B(x, r) \cap A^c \neq \emptyset$, dunque $x \in \text{Fr } A$. \square

Esercizio 1.9. $A \subseteq \mathbb{R}^n$ è chiuso $\iff \text{Fr } A \subseteq A$

Dimostrazione.

\Rightarrow : osserviamo che se A è chiuso, allora $A = \bar{A}$, pertanto siccome $\text{Fr } A \subseteq \bar{A} \implies \text{Fr } A \subseteq A$. Si ha necessariamente che $\text{Fr } A \subseteq \bar{A}$ siccome $\text{Fr } A = \bar{A} \cap \bar{A}^c$.

\Leftarrow : se $\text{Fr } A \subseteq A$ allora preso $x \in \bar{A}$ avremo due possibilità:

- $x \in A$;
- $x \in \text{Fr } A \subseteq A \implies x \in A$

dunque $\bar{A} \subseteq A \implies A = \bar{A}$. □

Esercizio 1.10. *Provare che $\bar{A} = A \cup \text{Fr } A$*

Dimostrazione.

Mostriamo che vale $\bar{A} \supseteq A \cup \text{Fr } A$: è banale vedere che $A \subseteq \bar{A}$ e, in virtù di quanto visto in un esercizio precedente, siccome $\text{Fr } A = \bar{A} \cap \bar{A}^c \implies \text{Fr } A \subseteq \bar{A} \implies A \cup \text{Fr } A \subseteq \bar{A}$.

L'inclusione opposta si ottiene prendendo $x \in \bar{A}$ e ragionando come sopra. □

Definizione 1.11 (punti interni). Sia $A \subseteq \mathbb{R}^n$, diremo che x è un punto interno di A se $\exists r > 0 : B(x, r) \subseteq A$

Oss.: A aperto $\iff A = \text{Int } A$

Proposizione 1.10. F è chiuso $\iff F^c$ è aperto

Dimostrazione.

\Rightarrow : se F è chiuso e $z \in F^c \implies z \in \text{Int}(F^c)$. Se per assurdo $z \notin \text{Int}(F^c)$ allora $\forall r > 0 : B(z, r) \not\subseteq F^c \implies B(z, r) \cap F \neq \emptyset \implies z \in \bar{F} = F \implies z \in F$, che contraddice l'ipotesi iniziale.

\Leftarrow : se F^c è aperto, allora preso $z \in \bar{F}$ se, per assurdo, avessimo che $z \notin F \implies z \in F^c$ allora $\exists \varepsilon > 0 : B(z, \varepsilon) \subseteq F^c$ (per l'ipotesi di apertura), pertanto $B(z, \varepsilon) \cap F = \emptyset \implies z \notin \bar{F}$, il che contraddice l'ipotesi iniziale. □

CAPITOLO 2

Successioni e funzioni continue

In questo capitolo andremo a definire il concetto di successione e vedere il forte collegamento presente fra esse e il concetto di continuità in più variabili.

Ricordiamo al lettore che una successione a valori in un insieme A è una funzione $x_k : \mathbb{N} \rightarrow A$. Vediamo adesso come si definisce il concetto di *successione convergente* in più variabili

Definizione 2.1 (convergenza di una successione). diremo che $\{x_k\} \subset \mathbb{R}^n$ è una successione convergente a $z \in \mathbb{R}^n$ se $\lim_{k \rightarrow +\infty} |x_k - z| = 0$.

Oss.: Naturalmente se $\underline{x}_k = (x_{k_1}, x_{k_2}, \dots, x_{k_n}) \in \mathbb{R}^n$ converge a $\underline{z} = (z_1, z_2, \dots, z_n)$ $\iff \lim_{k \rightarrow \infty} x_{k_i} = z_i \forall i \in \{1, \dots, n\}$

Esempio 2.1. Consideriamo la successione $\underline{x}_k = (e^{-k} + 1, (-1)^k)$. Per il precedente teorema questa successione è convergente.

Proposizione 2.1 (Unicità del limite). Il limite di una successione è unico. Se $\underline{x}_k \in \mathbb{R}^n \forall k$ converge a $\underline{z} \in \mathbb{R}^n$ allora z è unico.

Dimostrazione. Supponiamo per assurdo che \underline{x}_k converga a \underline{z} e \underline{y} con $\underline{z} \neq \underline{y}$. Allora

$$|\underline{z} - \underline{y}| \leq |\underline{x}_k - \underline{z}| + |\underline{x}_k - \underline{y}| \xrightarrow{k \rightarrow +\infty} 0 \implies z = y$$

□

Proposizione 2.2. Se $\underline{x}_k \rightarrow \underline{x}$ allora $|\underline{x}_k| \rightarrow |\underline{x}|$

Dimostrazione. si osserva che, dalla proposizione 1.2, si ha che

$$|\underline{x}_k| - |\underline{x}| \leq |\underline{x}_k - \underline{x}| \xrightarrow{k \rightarrow +\infty} 0$$

□

Mostriamo adesso una banale proposizione, le cui conseguenze non sono così "scontate".

Proposizione 2.3 (spazio vettoriale delle successioni convergenti). Se $\underline{x}_k \rightarrow \underline{x}$ e $\underline{y}_k \rightarrow \underline{y}$ allora $\forall \lambda, \mu \in \mathbb{R}, \lambda \underline{x}_k + \mu \underline{y}_k \rightarrow \lambda \underline{x} + \mu \underline{y}$

Dimostrazione. osserviamo che

$$|\lambda \underline{x}_k + \mu \underline{y}_k - \lambda \underline{x} - \mu \underline{y}| = |\lambda(\underline{x}_k - \underline{x}) + \mu(\underline{y}_k - \underline{y})|$$

Per la proposizione 1.2 abbiamo che

$$|\lambda(\underline{x}_k - \underline{x}) + \mu(\underline{y}_k - \underline{y})| \leq |\lambda||\underline{x}_k - \underline{x}| + |\mu||\underline{y}_k - \underline{y}| \xrightarrow{k \rightarrow +\infty} 0$$

dunque la tesi \square

Oss.: l'importanza di questa dimostrazione sta nel fatto che questo teorema dimostra che l'insieme delle successioni convergenti in \mathbb{R}^n forma uno spazio vettoriale che è chiuso rispetto all'addizione $+_{\mathbb{R}}$ e prendendo come prodotto per scalare $*_{\mathbb{R}}$

Adesso andiamo a mostrare una proprietà che segue direttamente dalla topologia di \mathbb{R}^n (in spazi topologici qualunque non è sempre valido)

Proposizione 2.4 (caratterizzazione degli insiemi chiusi). Un insieme $A \subset \mathbb{R}^n$ è chiuso se e solo se $\forall \underline{x}_k \rightarrow \underline{z}, \underline{x}_k \in A \forall k$ allora $\underline{z} \in A$. Formalmente

$$(A \subset \mathbb{R}^n \text{ è chiuso}) \iff (\forall \underline{x}_k \rightarrow \underline{z}, \underline{x}_k \in A \forall k \implies \underline{z} \in A)$$

Dimostrazione.

\Rightarrow : Se A è chiuso, considerando una generica $\underline{x}_k \rightarrow \underline{z}, \underline{x}_k \in A \forall k$, allora $\forall r > 0$ dato che $|\underline{x}_k - \underline{z}| \rightarrow 0 \exists k_r \in \mathbb{N}$ tale che $|\underline{x}_k - \underline{z}| < r \forall k \geq k_r \implies \underline{x}_k \in B(\underline{z}, r) \implies B(\underline{z}, r) \cap A \neq \emptyset \implies \underline{z} \in \bar{A}$. Siccome A è chiuso allora $\bar{A} = A$ dunque $\underline{z} \in A$.

\Leftarrow : Sia $\underline{w} \in \bar{A} \implies B(\underline{w}, \frac{1}{k}) \cap A \neq \emptyset \forall k \geq 1$ quindi esiste una successione \underline{x}_k (potremmo prendere per esempio $\underline{w} - \frac{1}{k}$) tale che $|\underline{x}_k - \underline{w}| < \frac{1}{k} \implies \underline{x}_k \xrightarrow{k \rightarrow +\infty} \underline{w} \implies \underline{w} \in A$ per la seconda proprietà, ma allora segue che $\bar{A} = A$, dunque A è chiuso per definizione. \square

Questa proprietà è molto importante, siccome garantisce l'equivalenza fra **chiusura sequenziale** e **chiusura** (che negli spazi topologici non è generalmente garantito) in \mathbb{R}^n .

In \mathbb{R}^n , dato un insieme $A \subset \mathbb{R}^n$, se vogliamo verificare che sia chiuso sarà dunque necessario verificare che ogni successione \underline{x}_k tali che $\forall k, \underline{x}_k \in A$ avremo che $\underline{x}_k \rightarrow \underline{x} \in A$, ovvero che ogni successione convergente in A converga ad un punto che appartiene ad A .

2.1 Funzioni continue

Definizione 2.2 (funzioni continue). Una funzione $f : A \rightarrow \mathbb{R}$ con $A \subset \mathbb{R}^m$ si dice continua in $\underline{z} \in A$ se

$$\forall \underline{x}_k \in A, \underline{x}_k \rightarrow \underline{z}, \lim_{k \rightarrow +\infty} f(\underline{x}_k) = f(\underline{z})$$

Oss.: Diremo che $f : A \rightarrow \mathbb{R}$ con $A \subset \mathbb{R}^n$ è continua in A se è continua $\forall \underline{x} \in A$.

Andiamo a ricordare un concetto molto importante preso dal precedente capitolo:

Definizione 2.3 (apertura relativa). Fissato $A \subset \mathbb{R}^n$ diremo che $S \subset A$ è aperto in A se $\exists \Omega \subset \mathbb{R}^n$ aperto tale che

$$S = A \cap \Omega$$

Dopo questa definizione siamo pronti per enunciare una serie di teoremi sulle funzioni continue che le caratterizzano in \mathbb{R}^n .

Teorema 2.1 (teorema C1). Sia $f : A \rightarrow \mathbb{R}^m$ con $A \subset \mathbb{R}^n$. Allora f è continua $\iff \forall U \subset \mathbb{R}^m$ aperto $f^{-1}(U)$ è aperto in A

Dimostrazione.

\Rightarrow : se f è continua, supponiamo per assurdo che esista $U \subset \mathbb{R}^m$ aperto tale che $f^{-1}(U)$ non è aperto in A . Allora $\exists z \in f^{-1}(U)$ tale che $\forall k \geq 1, B(z, \frac{1}{k}) \notin f^{-1}(U)$. Ma questo implica che $\exists \underline{x}_k \in A \cap B(z, \frac{1}{k})$ tale che $\underline{x}_k \notin f^{-1}(U)$. Ma allora

$$|\underline{x}_k - \underline{w}| < \frac{1}{k} \rightarrow 0 \implies f(\underline{x}_k) \rightarrow f(z)$$

ma allora $f(\underline{x}_k) \in U^c$ (se appartenesse a U allora $\underline{x}_k \in f^{-1}(U)$ il che contraddice come abbiamo definito la successione).

Ma allora $f(z) \notin U$: infatti, U^c è chiuso dunque, per la proposizione 2.4, abbiamo che ogni successione a valori in U^c converge ad un punto appartenente in U^c , ma ciò contraddice l'ipotesi iniziale che $z \in f^{-1}(U)$.

\Leftarrow : sappiamo che la preimmagine di aperti è aperta in A allora sia $\underline{x}_k \rightarrow \underline{z}$ con $\underline{x}_k \in A$ e $\underline{z} \in A$: se supponiamo, per assurdo, $f(\underline{x}_k) \not\rightarrow f(\underline{z})$ allora $\exists \alpha : \mathbb{N} \rightarrow \mathbb{N}$ tale che $\alpha(k) \xrightarrow{k \rightarrow +\infty} +\infty$ ed esiste $\sigma > 0 : |f(\underline{x}_{\alpha(k)}) - f(\underline{z})| \geq \sigma$. D'altra parte $f^{-1}(B(f(z), \sigma))$ è aperto in A (siccome preimmagine di una balla, che è aperta), dunque esiste $\delta > 0$ tale che $A \cap B(z, \delta) \subset f^{-1}(B(f(z), \sigma))$ ed esiste $k_0 \geq 1$ tale che $\underline{x}_k \in A \cap B(z, \delta) \subseteq f^{-1}(B(f(z), \sigma)) \forall k \geq k_0 \implies |f(\underline{x}_k) - f(\underline{z})| \leq \sigma$ e questo contraddice la precedente disequazione $|f(\underline{x}_{\alpha(k)}) - f(\underline{z})| \geq \sigma$ per k sufficientemente grande. \square

Teorema 2.2 (teorema della permanenza del segno). Sia $f : A \rightarrow \mathbb{R}$ continua in $z \in A$. Se $f(z) > 0$ allora $\exists \delta > 0$ tale che $\forall x \in A \cap B(z, \delta), f(x) > 0$

Dimostrazione. Se neghiamo la tesi allora otteniamo che $\forall \delta > 0, \exists x \in A \cap B(z, \delta) : f(x) < 0$. Ma allora deduciamo che $\exists \underline{x}_k \in A \cap B(z, \frac{1}{k}) : f(\underline{x}_k) < 0$, ma $\underline{x}_k \rightarrow z$ e $f(\underline{x}_k) \rightarrow f(z) \leq 0$ che rappresenta una contraddizione. \square

Teorema 2.3 (teorema C2). Siano $A \subset \mathbb{R}^n, f, g : A \rightarrow \mathbb{R}^m$ e $h : B \rightarrow \mathbb{R}^k$ con $f(A) \subset B \subset \mathbb{R}^m$. Se f e g sono continue in $z \in A$ e h è continua in $f(z) \in B$ allora valgono le seguenti proprietà:

- ① $\forall \lambda, \mu \in \mathbb{R}$ abbiamo che $\lambda f + \mu g$ è continua in $z \in A$;
- ② Se $m = 1$, allora $f g$ è continua in z ;
- ③ Se $m = 1$ e $g(z) \neq 0$ allora $\exists \delta > 0$ per cui f/g è ben definita su $A \cap B(z, \delta)$ ed è continua in z ;
- ④ la composizione $h \circ f$ è continua in z

Dimostrazione. Per mostrare la ① osserviamo che, prendendo $x_k \rightarrow z$, abbiamo che

$$\lambda f(x_k) + \mu g(x_k) \rightarrow \lambda f(z) + \mu g(z)$$

Per mostrare la ② osserviamo che

$$x_k \rightarrow z \implies f(x_k)g(x_k) \rightarrow f(z)g(z)$$

Per mostrare la ③ si osserva che possiamo supporre che $g(z) > 0$ senza perdita di generalità da cui segue che

$$\exists \delta > 0 : \forall x \in A \cap B(z, \delta), g(x) > 0 \implies \frac{f}{g} : B(z, \delta) \rightarrow \mathbb{R}$$

che ci permette di concludere che $\frac{f}{g}$ sia ben definita e $\lim_{k \rightarrow +\infty} \frac{f(x_k)}{g(x_k)} = \frac{f(z)}{g(z)}$ con $x_k \rightarrow z$.

La ④ segue banalmente dal fatto che presa $x_k \in A, x_k \rightarrow z$ allora

$$\lim_{k \rightarrow +\infty} h(f(x_k)) = h(\lim_{k \rightarrow +\infty} f(x_k)) = h(f(z))$$

per la continuità di h in $f(z) \in f(A)$ e di f in $z \in A$. \square

Proposizione 2.5. Sia $f : A \rightarrow \mathbb{R}^m$, dove $f(x) = (f_1(x), f_2(x), \dots, f_m(x))$ e $f_j(x) : A \rightarrow \mathbb{R} \forall j \in \{1, \dots, m\}$ con $A \subset \mathbb{R}^n$. Allora abbiamo che

$$f \text{ è continua in } p \in A \iff f_j \text{ è continua in } p \in A \forall j \in \{1, \dots, m\}$$

Dimostrazione. Basta osservare che $f(x_k) \rightarrow f(p) \iff f_j(x_k) \rightarrow f_j(p) \forall j \in \{1, \dots, m\}$ \square

Oss.: f è continua in $A \iff f_j$ è continua in $A \forall j \in \{1, \dots, m\}$.

Esempio 2.2. La funzione $|\cdot| : \mathbb{R}^n \rightarrow [0; +\infty)$ è continua siccome abbiamo mostrato che se $x_k \xrightarrow{k \rightarrow +\infty} x \implies |x_k| \xrightarrow{k \rightarrow +\infty} |x|$

Esempio 2.3. La funzione $f(x, y, z) = x^2 - 3y + \sqrt{|xz|}, f : \mathbb{R}^3 \rightarrow \mathbb{R}$ è continua in quanto somma di funzioni continue: $\sqrt{|xz|}$ è continua in quanto composizione della funzione $\sqrt{t}, t \geq 0$ con $|\cdot|$ e xz che sono tutte continue.

Esempio 2.4. $f : \mathbb{R}^2 \setminus \{0\} \rightarrow \mathbb{R}$ con

$$f(x, y) = \frac{xy^3}{x^2 + y^4}$$

è continua in quanto quoziente di funzioni continue.

Vogliamo adesso individuare delle *buone* proprietà che le funzioni continue possiedono. Per fare ciò dobbiamo definire qualche altro concetto topologico:

Definizione 2.4. Un insieme $E \subseteq \mathbb{R}^n$ è limitato se $\exists R > 0 : E \subseteq B(0, R)$

Oss.: Un insieme è illimitato $\iff \exists \{x_k\} \in E \ \forall k : |x_k| \rightarrow +\infty$

Oss.: Ogni palla aperta o chiusa è limitata

Oss.: Se B è limitato e $A \subseteq B \implies A$ è limitato.

Esempio 2.5.

$$A = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 + 1 < z < 10\}$$

è limitato. Infatti preso $(x, y, z) \in A \implies |(x, y, z)|^2 = x^2 + y^2 + z^2 = x^2 + y^2 + z^2 + 1 - 1 < z + z^2 - 1 < 10 + 100 - 1 = 109 \implies A \subseteq B(0, \sqrt{109})$ dunque è limitato.

Da questo esempio intuiamo che per verificare la limitatezza di un insieme non è necessario disegnarlo (e spesso è anche molto difficile).

Esempio 2.6.

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

osserviamo che è illimitato siccome possiamo prendere $x_k = (\frac{1}{2k}, k)$ e osservare che $|\frac{1}{2k} \cdot k| = \frac{1}{2} < 1$ dunque $\forall k, x_k = (\frac{1}{2k}, k) \in A$ ma per $k \rightarrow +\infty, |x_k| = \sqrt{\frac{1}{4k^2} + k^2} \geq |k| \xrightarrow{k \rightarrow +\infty} +\infty$. Questo insieme, nel piano cartesiano, corrisponde ai punti che stanno sotto alla figura $|xy| = 1$

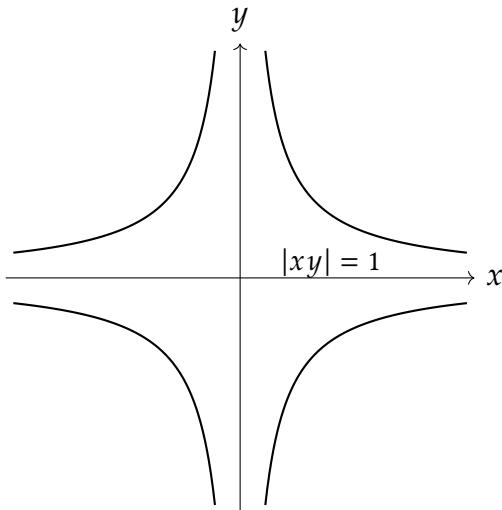


Figura 2.1: Immagine dell'insieme A

Definizione 2.5 (insiemi compatti). Un insieme $E \subseteq \mathbb{R}^n$ è compatto se $\forall \{x_k\} \subset E, \exists x \in E \exists \alpha : \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che

$$x_{\alpha(k)} \rightarrow x \in E$$

Teorema 2.4 (teorema di Heine-Borel). Dato $E \subseteq \mathbb{R}^n$, allora

$$E \subseteq \mathbb{R}^n \text{ è compatto} \iff E \text{ è chiuso e limitato.}$$

Dimostrazione.

\Rightarrow : sia $x_k \rightarrow x \implies$ (per l'ipotesi di compattezza) $\exists y \in E \exists \alpha : \mathbb{N} \rightarrow \mathbb{N} x_{\alpha(k)} \rightarrow y \in E \implies x = y \in E$ per unicità del limite. Se E non fosse limitato $\implies \exists \{x_k\} \subset E : |x_k| \rightarrow +\infty$ e per la compattezza $\exists \beta : \mathbb{N} \rightarrow \mathbb{N}$ strettamente crescente tale che $x_{\beta(k)} \rightarrow x \in E$ ma per unicità del limite $|x_{\beta(k)}| \rightarrow |x_k| \rightarrow +\infty$ e questo contraddice il limite $+\infty$.

\Leftarrow : Se $\{x_k\} \subset E$ e $x_k = (x_{k_1}, \dots, x_{k_n}) \implies x_{k_j}$ sono successioni reali limitate tali che $\exists x_{\alpha_1(k)_1} \rightarrow x_1, \exists x_{\alpha_2(k)_2} \rightarrow x_2, \dots \implies \exists x_{\alpha_1 \dots \alpha_n(k)} \rightarrow x = (x_1, \dots, x_n)$ e $x \in E$ per la chiusura di E . Il fatto di poter estrarre delle sottosuccessioni convergenti è garantito dal teorema di Bolzano-Weierstrass. \square

A questo punto ricordiamo la definizione di massimo e minimo locale o globale

Definizione 2.6 (massimo globale). Sia $A \subseteq \mathbb{R}^n$ e sia $f : A \rightarrow \mathbb{R}$. Diremo che $z \in A$ è un punto di massimo assoluto, o globale, su A se

$$\forall x \in A, f(z) \geq f(x)$$

La definizione di minimo globale è analoga (basta sostituire il \geq con \leq).

Diamo adesso la definizione di massimo e minimo locale:

Definizione 2.7 (massimo locale). Sia $A \subseteq \mathbb{R}^n$ e sia $f : A \rightarrow \mathbb{R}$. Diremo che $z \in A$ è un punto di massimo locale su A se

$$\exists \delta > 0 : \forall x \in A \cap B(z, \delta), f(z) \geq f(x)$$

analogamente si ottiene la definizione di minimo locale.

Data una funzione $f : A \rightarrow \mathbb{R}$ con $A \subset \mathbb{R}^n$, introdurremo la seguente notazione

$$\max_A f \quad \min_A f$$

per indicare, rispettivamente, il massimo e il minimo assoluto.

Enunciamo il seguente teorema, valido in \mathbb{R}^n , che garantisce l'esistenza del massimo e del minimo di una funzione continua quando *mappa* un insieme compatto ad un altro.

Teorema 2.5 (teorema di Weierstrass). Sia $A \subset \mathbb{R}^n$ un insieme compatto e sia $f : A \rightarrow \mathbb{R}$ una funzione continua su A . Allora $\exists \max_A f, \min_A f$.

Dimostrazione. Per caratterizzazione del $\sup f(A)$ sappiamo che esiste una successione y_k tale che $y_k \in f(A) \forall k \geq 1$ e $y_k \xrightarrow{k \rightarrow +\infty} \sup f(A)$, ovvero esiste una successione a valori in $f(A)$ che tende all'estremo superiore di tale insieme. Siccome $\forall k \geq 1, y_k \in f(A)$ allora $\exists x_k \in A : f(x_k) = y_k \forall k \geq 1$. Ma allora noi sappiamo che, per al compattezza di A , che esiste una sottosuccessione $x_{\alpha(k)}$ tale che $x_{\alpha(k)} \rightarrow z \in A$, da cui si deduce che

$$f(z) = \lim_{k \rightarrow +\infty} f(x_{\alpha(k)}) = \lim_{k \rightarrow +\infty} y_{\alpha(k)} = \sup_A f$$

dove la prima uguaglianza segue dalla continuità di f e la terza uguaglianza segue dal fatto che $y_{\alpha(k)}$ è una sottosuccessione estratta di y_k , ma siccome y_k converge al $\sup_A f$ allora anche $y_{\alpha(k)}$ vi convergerà. \square

Un'altra buona proprietà delle funzioni continue è quello di preservare la connessione per archi di un insieme

Definizione 2.8 (insieme connesso per archi). Un insieme $C \subseteq \mathbb{R}^n$ è connesso per archi se $\forall x, y \in C \exists \gamma : [0, 1] \rightarrow C$ continua tale che $\gamma(0) = x$ e $\gamma(1) = y$.

Esempio 2.7. Una palla è un insieme connesso per archi

Definizione 2.9 (insiemi convessi). $C \subseteq \mathbb{R}^n$ è convesso $\iff \forall x, y \in C$ si ha che $[x, y] = \{x + t(y - x) : t \in [0, 1]\} \subset C$.

Oss.: In maniera spicciola, un insieme è convesso se, presi due punti, il segmento che li congiunge è costituito da punti tutti appartenenti all'insieme C . Le palle aperte e chiuse rimangono un esempio di insiemi convessi.

Oss.: Dalla continuità del segmento segue banalmente che se E convesso $\implies E$ connesso per archi. Il viceversa naturalmente non è valido: si potrebbe pensare all'insieme

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

che corrisponde ad una corona circolare. E' connesso per archi siccome presi due punti possiamo trovare una curva che li "collega" tuttavia non è detto che, presi i due punti, sia possibile individuare un segmento che non passi nella zona che non fa parte di A .

Dimostriamo adesso che la continuità preserva la connessione per archi di un insieme:

Teorema 2.6 (teorema C4). Sia $f : C \rightarrow \mathbb{R}^m$ continua e sia C connesso per archi. Allora $f(C) \subset \mathbb{R}^m$ è connesso per archi

Dimostrazione. Siano $y, z \in f(C) \implies \exists x, u \in C$ tali che $f(x) = y$ e $f(u) = z$, pertanto $\exists \gamma : [0, 1] \rightarrow C$ grazie alla connettività per archi di C per cui $\gamma(0) = x$ e $\gamma(1) = u$. Ma allora la curva $f \circ \gamma : [0, 1] \rightarrow f(C)$ è continua (siccome composizione di funzioni continue) e $(f \circ \gamma)(0) = y$ e $(f \circ \gamma)(1) = z$, dunque $f(C)$ è connesso per archi. \square

Corollario 2.6.1. $f : C \rightarrow \mathbb{R}$ continua, C è connesso per archi $\implies f(C) \subset \mathbb{R}$ è un intervallo.

Dimostrazione. Dal precedente teorema segue, naturalmente, che $f(C)$ è connesso per archi. Mostriamo che è un intervallo: dati $t, s \in f(C)$, $\exists \Gamma : [0, 1] \rightarrow f(C) \subset \mathbb{R}$ continua tale che $\Gamma(0) = t$ e $\Gamma(1) = s$. Dunque $[t, s] \subseteq \Gamma([0, 1]) \subseteq f(C) \implies f(C)$ è un intervallo. \square

Da questo corollario è anche possibile anche dimostrare la seguente proposizione

Proposizione 2.6. Sia A connesso per archi e sia $f : A \rightarrow \mathbb{R}$ continua. Allora

- ① Se $\exists \max_A f, \min_A f$ allora $f(A) = [\min_A f, \max_A f]$
- ② Se $\nexists \max_A f, \exists \min_A f$ allora $f(A) = [\min_A f, \sup_A f]$

Esercizio 2.1. Mostrare la proposizione precedente

Dimostrazione. Per il corollario sappiamo che $f(A)$ è un intervallo.

Se il minimo e il massimo della funzione sono ben definiti, allora l'immagine è banalmente contenuta fra il minimo e il massimo la cui esistenza è garantita dal teorema di Weierstrass.

Per quanto riguarda il secondo caso, possiamo supporre, senza perdere di generalità, che $\nexists \max_A f$. Naturalmente, per caratterizzazione del sup abbiamo che $\forall \varepsilon > 0, \sup_A f - \varepsilon \in f(A)$, dunque $\exists x \in A : f(x) = \sup_A f - \varepsilon$, dunque $[\min_A f, f(x)] \subseteq [\min_A f, \sup_A f]$, dunque la tesi. \square

2.2 Accenni ai limiti in più variabili

Definizione 2.10 (definizione di limite). Sia data $f : A \rightarrow \mathbb{R}^k, A \subseteq \mathbb{R}^n$ e sia $x_0 \in \text{Fr}A$. Allora diciamo che f tende a $v \in \mathbb{R}^k$ per x che tende a x_0 e scriviamo che

$$\lim_{x \rightarrow x_0} f(x) = v \iff \forall \varepsilon > 0 \exists \delta > 0 : \forall x \in B(x_0, \delta) \cap A \setminus \{x_0\}, |f(x) - v| < \varepsilon$$

Diamo adesso la definizione di limite di $f(x)$ che tende all'infinito:

Definizione 2.11. Sia data $f : A \rightarrow \mathbb{R}^k, A \subseteq \mathbb{R}^n$ illimitato. Diremo che f tende all'infinito per x che tende all'infinito e scriveremo che

$$\lim_{|x| \rightarrow \infty} f(x) = \infty \iff \forall M > 0, \exists N > 0 : \forall x \in A, |x| > N, |f(x)| > M$$

Oss.: In generale tutte queste definizioni che stiamo dando funzionano in qualunque spazio metrico (X, d) sostituendo al posto dell'usuale distanza euclidea definita in \mathbb{R}^n la distanza d .

Potremo scrivere anche le definizioni per gli altri due casi che ci restano, ma è banale e lo lasciamo come semplice esercizio teorico al lettore.

Teorema 2.7 (caratterizzazione del limite per successioni). Sia $f : A \rightarrow \mathbb{R}^m$, $A \subseteq \mathbb{R}^n$, $q \in \text{Fr}A$, $v \in \mathbb{R}^m$. Abbiamo che

$$\lim_{x \rightarrow q} f(x) = v \iff (\forall \{p_k\} \subseteq A \setminus \{q\}, p_k \rightarrow q) \implies \lim_{k \rightarrow +\infty} f(p_k) = v$$

Dimostrazione.

\Rightarrow : si osserva che, presa una qualunque successione p_k che soddisfa le ipotesi della proposizione sulla destra, avremo che il limite $\lim_{k \rightarrow +\infty} f(p_k)$ è un limite dove avviene la composizione della successione p_k con la funzione f e sono verificate tutte le ipotesi riguardo al cambio di variabile nel limite, dunque $\lim_{k \rightarrow +\infty} f(p_k) = \lim_{x \rightarrow q} f(x) = v$.

\Leftarrow : osserviamo che se procedessimo per assurdo, negando la definizione di limite, avremo che $\exists \varepsilon > 0, \exists x \in B(q, \delta) \cap A \setminus \{q\} : |f(x) - v| \geq \varepsilon$. Ma allora, restringendosi a degli intorni di q via via sempre più piccoli, come $V_k = (q - \frac{1}{k}, q + \frac{1}{k})$ esisterebbe $x_k \in A \cap V_k : |f(x_k) - v| > \varepsilon$. Ma questo è un assurdo, siccome $x_k \rightarrow q$ e ma $f(x_k) \not\rightarrow v$ contraddicendo la nostra ipotesi iniziale. \square

Da questo teorema importantissimo segue questo semplice corollario (la cui dimostrazione è omessa siccome segue direttamente dalla definizione di funzione continua)

Corollario 2.7.1. Sia $\Omega \subseteq \mathbb{R}^n$ aperto e $f : \Omega \rightarrow \mathbb{R}^m$. Allora valgono le seguenti proprietà:

- (1) f è continua in $x_0 \in \Omega \iff \lim_{x \rightarrow q} f(x) = f(q)$;
- (2) f è continua in $\Omega \iff \forall q \in \Omega, \lim_{x \rightarrow q} f(x) = f(q)$

Teorema 2.8 (teorema del confronto). Siano $h, g, f : A \rightarrow \mathbb{R}$, $x_0 \in \text{Fr}A$, $l \in \mathbb{R}$ e supponiamo che $\exists \delta > 0$ con

$$h(x) \leq f(x) \leq g(x) \quad \forall x \in A \cap B(x_0, \delta) \setminus \{x_0\}$$

Se

$$\lim_{x \rightarrow x_0} h(x) = \lim_{x \rightarrow x_0} g(x) = l \implies \lim_{x \rightarrow x_0} f(x) = l$$

Dimostrazione. La dimostrazione è analoga a quella di Analisi 1 con qualche accorgimento sugli insiemi di definizione \square

Prima di procedere sulla parte del calcolo differenziale, diamo una breve introduzione ai simboli di Landau (i cosiddetti o-piccoli e O-grandì) che semplificano notevolmente il calcolo in più variabili:

Definizione 2.12 (o-piccolo). Siano $x_0 \in \text{Fr}A$, $A \subseteq \mathbb{R}^n$ e $f : A \rightarrow \mathbb{R}^m$ e $g : A \rightarrow \mathbb{R}^k$. Diremo che f è un o-piccolo di g per x che tende a x_0 , in simboli

$$f = o(g) \text{ per } x \rightarrow x_0$$

se

$$\forall \varepsilon > 0 \exists \delta > 0 : |f(x)| \leq \varepsilon |g(x)| \quad \forall x \in A \cap B(x_0, \delta) \setminus \{x_0\}$$

Definizione 2.13 (O-grande). Siano $x_0 \in \text{Fr}A$, $A \subseteq \mathbb{R}^n$ e $f : A \rightarrow \mathbb{R}^m$ e $g : A \rightarrow \mathbb{R}^k$. Diremo che f è un O-grande di g per x che tende a x_0 , in simboli

$$f = O(g) \text{ per } x \rightarrow x_0$$

se

$$\exists M > 0, \delta > 0 : |f(x)| \leq M |g(x)| \quad \forall x \in A \cap B(x_0, \delta) \setminus \{x_0\}$$

Esercizio 2.2. Se f è un o -piccolo di g allora f è O -grande di g

Dimostrazione. Segue direttamente dalla definizione, siccome se $f = o(g)$ allora $\forall \varepsilon > 0 \exists \delta > 0 : |f(x)| \leq \varepsilon |g(x)| \forall x \in A \cap B(x_0, \delta) \setminus \{x_0\}$, dunque fissando $M = \varepsilon > 0$ allora sappiamo che esiste un $\delta(\varepsilon)$ per cui

$$|f(x)| \leq M|g(x)| \forall x \in A \cap B(x_0, \delta) \setminus \{x_0\}$$

dunque risulta che f è un O -grande di g , ovvero la tesi. \square

Proposizione 2.7. Siano $A \subseteq \mathbb{R}^n$, $x_0 \in A$ e $f : A \rightarrow \mathbb{R}^m$ e $\alpha > 0$. Allora abbiamo

$$f(x) = o(|x - x_0|^\alpha) \text{ per } x \rightarrow x_0 \iff \lim_{x \rightarrow x_0} \frac{f(x)}{|x - x_0|^\alpha} = 0$$

Dimostrazione. la dimostrazione è banale, siccome

$$\begin{aligned} \forall \varepsilon > 0 \exists \delta > 0 : |f(x)| \leq \varepsilon |x - x_0|^\alpha \forall x \in A \cap B(x_0, \delta) \setminus \{x_0\} &\iff \forall \varepsilon > 0 \exists \delta > 0 : \\ \frac{|f(x)|}{|x - x_0|^\alpha} \leq \varepsilon \forall x \in A \cap B(x_0, \delta) \setminus \{x_0\} &\iff \lim_{x \rightarrow x_0} \frac{f(x)}{|x - x_0|^\alpha} = 0 \end{aligned}$$

\square

Proposizione 2.8. Siano $A \subseteq \mathbb{R}^n$, $x_0 \in A$, $f : A \rightarrow \mathbb{R}$ e $\alpha > 0$. Allora abbiamo che

$$f(x) = O(|x - x_0|^\alpha) \text{ per } x \rightarrow x_0 \iff \exists M, \delta > 0 : \frac{f(x)}{|x - x_0|^\alpha} \leq M \forall x \in A \cap B(x_0, \delta) \setminus \{x_0\}$$

Dimostrazione. Analoga alla dimostrazione fatta sopra \square

L'ultima proposizione si può anche riscrivere dicendo che se $f = O(|x - x_0|^\alpha)$ allora abbiamo

necessariamente, per $x \rightarrow x_0$, che $\sup_{A \cap B(x_0, \delta) \setminus \{x_0\}} \left| \frac{f(x)}{|x - x_0|^\alpha} \right| < +\infty$.

Osserviamo che

$$\lim_{x \rightarrow x_0} f(x) = l \iff f(x) = l + o(1) \text{ per } x \rightarrow x_0$$

e quindi potremmo (con tanta buona volontà) andare a riscrivere tutta la teoria appena fatta sui limiti in più variabili tramite i simboli di Landau. Ciò non verrà fatto, ma li useremo per introdurre, come avevo accennato, il calcolo differenziale.

CAPITOLO 3

Calcolo differenziale

In questo capitolo andremo a generalizzare il concetto di derivata alle funzioni vettoriali (sia nell'insieme di definizione che nell'insieme di arrivo). Tutto ciò avverrà passando, naturalmente, per la definizione di funzione differenziabile e, successivamente, andando ad enunciare i teoremi riguardo all'esistenza e l'unicità del differenziale, soffermandosi in particolare modo nell'individuare delle espressioni per il differenziale, il gradiente e la matrice jacobiana e, per poi passare, alle derivate seconde, necessarie per lo studio dei massimi e dei minimi tramite lo studio della matrice hessiana.

3.1 Funzione differenziabile

Definizione 3.1 (funzione differenziabile in un punto e matrice jacobiana). Sia $\Omega \subseteq \mathbb{R}^n$ aperto e $x_0 \in \Omega$. Diremo che $f : \Omega \rightarrow \mathbb{R}^m$ è differenziabile in x_0 se $\exists L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ lineare tale che

$$\frac{f(x_0 + h) - f(x_0) - L(h)}{|h|} \xrightarrow{h \rightarrow 0} 0$$

dove l'applicazione $L = df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}^m$ è detto il differenziale di f in x_0 . La matrice $Df(x_0) \in \mathbb{R}^{m \times n}$ che rappresenta $df(x_0)$ rispetto alla base canonica è detta **matrice jacobiana** di f in x_0 .

Oss.: Naturalmente diremo che f è differenziabile in Ω se è differenziabile $\forall x \in \Omega$.

Come avevo anticipato alla fine dello scorso capitolo, possiamo rendere molto semplice la trattazione del calcolo differenziabile tramite i simboli di Landau. Mostriamo per esempio la seguente proposizione:

Proposizione 3.1. $f : \Omega \rightarrow \mathbb{R}^m$ con $\Omega \subseteq \mathbb{R}^n$ aperto. Diremo che f è differenziabile in $x_0 \in \Omega \iff \exists L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ lineare tale che $f(x_0 + h) = f(x_0) + L(h) + o(h)$ per $h \rightarrow 0 \iff$ (cambio di variabile) $\exists L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ lineare tale che $f(x) = f(x_0) + L(x - x_0) + o(x - x_0)$ per $x \rightarrow x_0$.

Dimostrazione. Se f è differenziabile in $x_0 \in \Omega$, sappiamo che $\exists L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ tale che

$$\frac{f(x_0 + h) - f(x_0) - L(h)}{|h|} \rightarrow 0 \implies f(x_0 + h) - f(x_0) - L(h) = o(|h|) \text{ per } h \rightarrow 0$$

che è equivalente (per cambiamento di variabile) a:

$$f(x) = f(x_0) + L(x - x_0) + o(|x - x_0|) \text{ per } x \rightarrow x_0$$

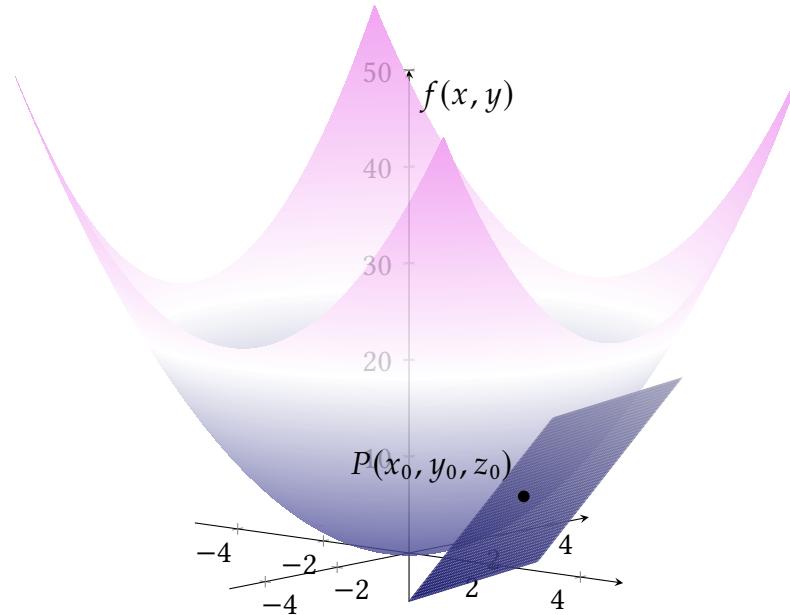


Figura 3.1: Nozione di differenziabilità da un punto di vista geometrico: la funzione $x^2 + y^2$ è differenziabile nel punto $x_0 = (1, 2, 5)$ e, dunque, possiamo approssimare la funzione in un intorno di x_0 alla funzione lineare L tale che $L(x, y) = 5 + 2(x - 1) + 4(y - 2)$

Il viceversa è analogo, siccome se

$$f(x_0 + h) - f(x_0) - L(h) = o(|h|) \implies \frac{f(x_0 + h) - f(x_0) - L(h)}{|h|} \rightarrow 0 \implies \\ f \text{ è differenziabile in } x_0$$

□

Ma in sostanza che cosa vuol dire essere differenziabili? Come forse alcuni avranno potuto capire dalla definizione, essere differenziabili vuol dire che è possibile approssimare la funzione, in quel punto $x_0 \in \Omega$ in cui è differenziabile, ad una applicazione lineare affine del tipo

$$a(x) = f(x_0) + df(x_0)(x - x_0)$$

che sarebbe, in maniera impropria, ciò che, in gergo da fisici, diciamo essere un'approssimazione del primo ordine.

Dopo questa definizione siamo persino pronti a definire il concetto di derivata direzionale, nozione strettamente connessa a quella di differenziabilità.

Definizione 3.2 (derivata direzionale). Sia $\Omega \subseteq \mathbb{R}^n$ aperto, $x_0 \in \Omega$, $f : \Omega \rightarrow \mathbb{R}^n$ e sia $v \in \mathbb{R}^n \setminus \{\underline{0}\}$. Definiamo la derivata direzionale rispetto a v il limite (se esiste!)

$$\partial_v f(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + tv) - f(x_0)}{t}$$

Nel caso in cui $v = e_i$ la derivata direzionale è detta derivata parziale rispetto a x_i .

Oss.: La derivata parziale rispetto a x_i è, di fatto, la derivata di una funzione di una sola variabile. Pertanto queste posso essere svolte tenendo "costanti" le altre variabili x_j con $j \neq i$ e calcolare la derivata come usualmente si faceva ad Analisi 1

Ricollegandoci all'osservazione fatta qua sopra, infatti, possiamo osservare che:

$$\partial_{e_1} f(x_0) = \lim_{t \rightarrow 0} \frac{f(x_0 + te_1) - f(x_0)}{t} = \lim_{t \rightarrow 0} \frac{f(x_1 + t, x_2, \dots) - f(x_1, x_2, \dots)}{t}$$

Siamo adesso pronti ad enunciare il seguente risultato, che ci dà qualche informazione su come sia il differenziale:

Teorema 3.1 (teorema DF1). Se $f : \Omega \rightarrow \mathbb{R}$ è differenziabile in x_0 allora $\exists \partial_v f \forall v \in \mathbb{R}^n \setminus \{\underline{0}\}$ e vale che

$$df(x_0)(v) = \partial_v f(x_0)$$

Dimostrazione. Siccome f è differenziabile in $x_0 \in \Omega$ allora $\exists L : \mathbb{R}^n \rightarrow \mathbb{R}$ tale che $f(x_0 + tv) = f(x_0) + L(tv) + o(|tv|)$ dove $|tv| \rightarrow 0$ per $v \in \mathbb{R}^n \setminus \{\underline{0}\}$. Ma questo allora implica che

$$\begin{aligned} \frac{f(x_0 + tv) - f(x_0) - L(|tv|)}{t} \rightarrow 0 &\implies \frac{f(x_0 + tv) - f(x_0) - tL(v)}{t} \rightarrow 0 \implies \\ \implies \lim_{t \rightarrow 0} \left(\frac{f(x_0 + tv) - f(x_0)}{t} - L(v) \right) &\rightarrow 0 \implies \exists \partial_v f(x_0) = L(v) = df(x_0)(v) \end{aligned}$$

□

Teorema 3.2 (rappresentazione e unicità del differenziale). Se $f : \Omega \rightarrow \mathbb{R}$ è differenziabile in $x_0 \in \Omega$ allora il differenziabile di f nel punto $x_0 \in \Omega$ è unico e

$$df(x_0)(\xi_1, \dots, \xi_n) = \sum_{j=1}^n \xi_j \frac{\partial f}{\partial x_j}(x_0)$$

Dimostrazione. Preso $v \in \mathbb{R}^n \setminus \{\underline{0}\}$, allora

$$v = \sum \xi_i e_i \simeq (\xi_1, \dots, \xi_n) \implies df(x_0)(v) = \sum_{j=1}^n \xi_j df(x_0)(e_j) = \sum_{j=1}^n \xi_j \partial_{e_j} f(x_0)$$

dove abbiamo solamente sfruttato la linearità della funzione differenziabile e il fatto che $df(x_0)(e_j) = \frac{\partial f}{\partial x_j}(x_0)$. L'unicità del differenziabile deriva dal fatto che se supponiamo per assurdo che esista $L' \neq df(x_0)$ lineare, tali che $L(v) = df(x_0)(v) \forall v \in \mathbb{R}^n \setminus \{\underline{0}\}$ allora preso un $v \in \mathbb{R}^n \setminus \{\underline{0}\}$ avremo che:

$$L(v) = \sum_j^n \xi_j L(e_j) = \sum_j^n \xi_j df(x_0)(e_j) = df(x_0)(v)$$

dove tutte queste uguaglianze sono ottenute sfruttando la linearità di L e $df(x_0)$. L'unica possibilità (siccome questa è una relazione valida per ogni vettore) è che $L(e_j) = df(x_0)(e_j) \forall j$ ma allora queste due applicazioni lineari coincidono, giungendo ad un assurdo. Quindi il differenziale $df(x_0)$ è unico. □

Possiamo dunque definire il concetto di gradiente:

Definizione 3.3 (gradiente di una funzione). Sia $f : \Omega \rightarrow \mathbb{R}$ con $\Omega \subseteq \mathbb{R}^n$, $x_0 \in \Omega$ e f differenziabile in x_0 . Allora definiamo il gradiente

$$\nabla f(x_0) = \left(\frac{\partial f}{\partial x_1}(x_0), \dots, \frac{\partial f}{\partial x_n}(x_0) \right) = \sum_{i=1}^n \frac{\partial f}{\partial x_i}(x_0) e_i \in \mathbb{R}^n$$

Oss.: Il gradiente di f nel punto x_0 è di fatto il vettore che ha come componenti le derivate parziali di f

Dalla formula precedente è possibile ricavare che

$$df(x_0)(v) = \langle \nabla f(x_0), v \rangle$$

Oss.: Da qua possiamo vedere il grandissimo problema di definire il gradiente: per definire le derivate parziali, di fatto, abbiamo solamente fatto riferimento alla natura da "spazio metrico" di \mathbb{R}^n , quindi in generale potremmo andare a definire la derivata su una classe maggiore di insiemi rispetto che a \mathbb{R} e \mathbb{C} . Per definire il gradiente, però, abbiamo richiesto il prodotto scalare, che è una condizione decisamente più forte che alla semplice distanza.

Cerchiamo adesso di capire che cosa rappresenta il gradiente. Il gradiente possiamo vederlo come la direzione in cui la funzione ha la massima pendenza¹. Per vedere questo possiamo considerare il vettore $v = \frac{\nabla f(x_0)}{|\nabla f(x_0)|}$ e considerando il vettore $w \in \mathbb{R}^n$ tale che $|w| = 1$ allora

$$\partial_w f(x_0) = \langle \nabla f(x_0), w \rangle \leq |\nabla f(x_0)| |w| = |\nabla f(x_0)|$$

per la diseguaglianza di Cauchy-Schwarz. Tuttavia si osserva che

$$|\nabla f(x_0)| = \langle \nabla f(x_0), v \rangle = \langle \nabla f(x_0), \nabla f(x_0) \rangle \frac{1}{|\nabla f(x_0)|} = \partial_v f(x_0)$$

dunque

$$\partial_w f(x_0) \leq \partial_v f(x_0)$$

Per poter ampliare la nostra teoria ed esprimere correttamente il differenziale è necessario riprendere la definizione di spazio duale dal corso di Geometria:

Definizione 3.4 (spazio duale). Sia V uno spazio vettoriale sul campo \mathbb{K} , definiamo lo spazio duale V^* come lo spazio vettoriale dei funzionali lineari $f : V \rightarrow \mathbb{K}$. In simboli

$$V^* = \{f \mid f : V \rightarrow \mathbb{K}\}$$

Introduciamo, a questo punto, le funzioni proiezioni $\pi_i : \mathbb{R}^n \rightarrow \mathbb{R}$ tali che $x = (x_1, \dots, x_n) \mapsto x_i$, ovvero le funzioni che associano l' i -esima componente del vettore $x \in \mathbb{R}^n$.

Osserviamo naturalmente che

$$|\pi_i(x)| = |x_i| \leq |x| \quad \forall x \in \mathbb{R}^n$$

e notiamo che le funzioni proiezioni sono lineari e se $L : \mathbb{R}^n \rightarrow \mathbb{R}$ è lineare, allora

$$L(x) = L\left(\sum_{i=1}^n x_i e_i\right) = \sum_{i=1}^n x_i L(e_i) = \sum_{i=1}^n \pi_i(x) L(e_i)$$

e, naturalmente, per come è definita L abbiamo che $L(e_i) = a_i \forall i$, quindi ogni funzione lineare da \mathbb{R}^n a \mathbb{R} è una combinazione lineare delle funzioni lineari π_i .

Passando in \mathbb{R}^n , osserviamo che queste funzioni proiezioni sono una base dello spazio duale $(\mathbb{R}^n)^*$ e definiamo $dx_i = \pi_i$.

¹questo è particolarmente comodo, per gli algoritmi locali di minimizzazione o massimizzazione, in cui si sfrutta la pendenza della funzione nel punto per trovare i massimi e i minimi locali.

Proposizione 3.2. Sia $f : \Omega \rightarrow \mathbb{R}$, f differenziabile in $x_0 \in \Omega$. Allora

$$df(x_0) = \sum_{i=1}^n \partial_{x_i} f(x_0) dx_i = \partial_{x_1} f(x_0) dx_1 + \dots + \partial_{x_n} f(x_0) dx_n \in (\mathbb{R}^n)^*$$

Dimostrazione. Dato $v = (\xi_1, \dots, \xi_n) = \sum_{i=1}^n \xi_i e_i$ allora

$$df(x_0)(v) \stackrel{\text{per linearità}}{=} \sum_{i=1}^n \xi_i df(x_0)(e_i) = \sum_{i=1}^n \pi_i(v) df(x_0)(e_i)$$

dove l'ultima uguaglianza si è ottenuta osservando che $\xi_i = \pi_i(v) = dx_j(v)$. A questo punto, ricordando che $df(x_0)(e_i) = \partial_{x_i} f(x_0)$ in virtù del teorema 3.1, otteniamo la tesi, siccome

$$\sum_{i=1}^n \partial_{x_i} f(x_0) dx_i(v) \implies df(x_0) = \sum_{i=1}^n \partial_{x_i} f(x_0) dx_i$$

□

Oss. Data $f : \Omega \rightarrow \mathbb{R}$, possiamo allora vedere il differenziale come una funzione che ad ogni punto $p \in \Omega$ associa $df(p) \in (\mathbb{R}^n)^*$, ovvero $df : \Omega \mapsto (\mathbb{R}^n)^*$

$$df(p) = \sum_{i=1}^n df(p) dx_i \quad \forall p \in \Omega$$

Alla luce dell'osservazione qua sopra possiamo dunque dare la seguente definizione:

Definizione 3.5 (1-forma differenziale). Diremo che $\omega : \Omega \rightarrow (\mathbb{R}^n)^*$ con $\Omega \subseteq \mathbb{R}^n$ aperto è una 1-forma differenziale su Ω . Inoltre, $\forall p \in \Omega$ avremo un'unica n -upla di coefficienti $(a_1(p), \dots, a_n(p))$ dipendenti da p tali che

$$\omega(x) = \sum_{i=1}^n a_i(x) dx_i$$

Oss. Quando abbiamo a che fare con ω 1-forma differenziale possiamo andare a definire delle funzioni $a_1(x), \dots, a_n(x) : \Omega \rightarrow \mathbb{R}$ tali che

$$\omega(x) = \sum_{j=0}^n a_j(x) dx_j$$

Definizione 3.6 (campo vettoriale). Sia $\Omega \subseteq \mathbb{R}^n$ aperto, diremo che F è un campo vettoriale se $F : \Omega \rightarrow \mathbb{R}^n$, ovvero $F(p) = \sum_{j=1}^n f_j(p) e_j = (f_1(p), \dots, f_n(p))$ ove $f_j : \Omega \rightarrow \mathbb{R}$ sono le componenti di F

Oss. Se $\omega : \Omega \rightarrow (\mathbb{R}^n)^*$ è una 1-forma differenziale e $F : \Omega \rightarrow \mathbb{R}^n$ è un campo vettoriale, allora $\langle \omega, F \rangle : \Omega \rightarrow \mathbb{R}$ tale che $p \mapsto \omega(p)F(p)$ è uno scalare, ovvero è indipendente dal fatto che utilizziamo le basi canoniche per rappresentare ω e F . In questo caso diremo che la 1-forma differenziale ω si è contratta su F

Enunciamo adesso il risultato più importante di questo capitolo:

Teorema 3.3. Sia $\Omega \subseteq \mathbb{R}^n$ aperto, $p_0 \in \Omega$ e sia $f : \Omega \rightarrow \mathbb{R}^m$ ($\iff f = (f_1, \dots, f_m)$ con $f_j : \Omega \rightarrow \mathbb{R} \forall j$). Allora

$$f \text{ è differenziabile in } p_0 \iff \text{ogni componente di } f_j \text{ è differenziabile in } p_0$$

In ogni caso avremo che

$$D(f(x_0)) = \begin{pmatrix} \nabla f_1(p_0) \\ \nabla f_2(p_0) \\ \vdots \\ \nabla f_m(p_0) \end{pmatrix} \in \mathbb{R}^{m \times n}$$

Dimostrazione.

\Rightarrow : se f è differenziabile allora possiamo definire la funzione $\pi_i \circ df(p_0) : \mathbb{R}^n \rightarrow \mathbb{R}$ e sappiamo che

$$f(x) = f(p_0) + df(p_0)(x - p_0) + o(x - p_0) \quad (3.1)$$

e osserviamo, ricordando che $x_i \leq |x|$ con dove x_i rappresenta l' i -esima, che

$$\begin{aligned} \frac{|\pi_i \circ o(x - p_0)|}{|x - p_0|} &\leq \frac{|o(x - p_0)|}{|x - p_0|} \rightarrow 0 \\ \implies \pi_i \circ o(x - p_0) &= o(x - p_0) \end{aligned}$$

dove l'ultima affermazione segue, banalmente, dal teorema del confronto.

Componendo π_i all'applicazione lineare affine otteniamo che

$$\pi_i \circ f(x) = f_i(p_0) + (df(p_0))_i(x - p_0) + o(x - p_0) \text{ per } x \rightarrow p_0$$

Per l'unicità del differenziale di f_i (che ricordiamo essere una funzione $\Omega \rightarrow \mathbb{R}$, dunque sappiamo, per il teorema precedente, che il suo differenziale è unico) segue che

$$\pi_i \circ L(x) = df_i(p_0)(x) = \langle \nabla f_i(p_0), x \rangle$$

il che naturalmente implica che

$$\begin{aligned} L(x) &= \sum_{i=1}^m L_i(x) e_i = \sum_{i=1}^m (\pi_i \circ L)(x) e_i = \sum_{i=1}^m \langle \nabla f_i(x_0), x \rangle e_i = \begin{pmatrix} \partial_{x_1} f_1(p_0) & \dots & \partial_{x_n} f_1(p_0) \\ \vdots & \ddots & \vdots \\ \partial_{x_1} f_m(p_0) & \dots & \partial_{x_n} f_m(p_0) \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \\ &= Df(x_0)x \end{aligned}$$

e questo mostra la tesi.

\Leftarrow : se f_i è differenziabile in $p_0 \forall i \in 1, \dots, n$ definiamo $df_i(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}$ e osserviamo che

$$\begin{aligned} f(p_0 + h) - f(p_0) &= \sum_{i=1}^m (f_i(p_0 + h) - f_i(p_0)) e_i = \sum_{i=1}^m (df_i(p_0)(h) + o(h)) e_i = \\ &= \sum_{i=1}^m df_i(p_0)(h) e_i + \sum_{i=1}^m o_i(h) e_i = L(h) + o(h) \implies f \text{ è differenziale in } p_0 \end{aligned}$$

dove, naturalmente, abbiamo che

$$L(h) = \sum_{i=1}^m \langle \nabla f_i(p_0), h \rangle e_i = Df(x_0)h$$

Abbiamo dunque ottenuto la tesi. □

Corollario 3.3.1. Il teorema precedente mostra anche come il differenziale sia unico anche nel caso di funzioni vettoriali.

Teorema 3.4 (del differenziale totale). Sia $f : \Omega \rightarrow \mathbb{R}$, $\Omega \subseteq \mathbb{R}^n$ aperto e $p_0 \in \Omega$. Se

$$\exists \partial_{x_j} f : B(p_0, r) \rightarrow \mathbb{R} \quad \forall j \in \{1, \dots, n\} \text{ con } B(p_0, r) \subseteq \Omega$$

e se

$$\partial_{x_j} f \text{ sono tutte continue in } p_0 \quad \forall j \in \{1, \dots, n\}$$

allora f è differenziale in p_0

Dimostrazione. Consideriamo $h = \sum_{j=1}^n h_j e_j = (h_1, \dots, h_n)$. Osserviamo che posto

$$L(h) = \sum_{j=1}^n \partial_j f(p_0) h_j$$

possiamo notare come

$$\begin{aligned} f(p_0 + h) - f(p_0) - L(h) &= \\ &= f(p_0 + h) - f\left(p_0 + \sum_{i=1}^{n-1} h_i e_i\right) + f\left(p_0 + \sum_{i=1}^{n-1} h_i e_i\right) - f\left(p_0 + \sum_{i=1}^{n-2} h_i e_i\right) + f\left(p_0 + \sum_{i=1}^{n-2} h_i e_i\right) - \dots \\ &\quad - f(p_0) - L(h) \end{aligned}$$

possiamo raggruppare a 2 a due i termini, notando che:

$$\begin{aligned} &\left(f(p_0 + h) - f\left(p_0 + \sum_{i=1}^{n-1} h_i e_i\right) \right) + \left(f\left(p_0 + \sum_{i=1}^{n-1} h_i e_i\right) - f\left(p_0 + \sum_{i=1}^{n-2} h_i e_i\right) \right) + \dots + \\ &(f(p_0 + h_1) - f(p_0)) - L(h) \end{aligned}$$

al raggruppamento j -esimo teniamo fisso la $n - j + 1$ componente dell'argomento della funzione sulla sinistra, pertanto possiamo applicare il teorema di Lagrange se definiamo una funzione dipendente esclusivamente dalla variabile h_j , dunque:

$$\begin{aligned} f(p_0 + h) - f(p_0) - L(h) &= \\ &= \partial_{x_n} f\left(p_0 + \sum_{j=1}^{n-1} h_j e_j + \theta_n h_n e_n\right) h_n + \partial_{x_{n-1}} f\left(p_0 + \sum_{j=1}^{n-2} h_j e_j + \theta_{n-1} h_{n-1} e_{n-1}\right) h_{n-1} + \dots - L(h) = \\ &= \left(\partial_{x_n} f\left(p_0 + \sum_{j=1}^{n-1} h_j e_j + \theta_n h_n e_n\right) - \partial_{x_n} f(p_0) \right) h_n + \left(\partial_{x_{n-1}} f\left(p_0 + \sum_{j=1}^{n-2} h_j e_j + \theta_{n-1} h_{n-1} e_{n-1}\right) - \right. \\ &\quad \left. - \partial_{x_{n-1}} f(p_0) \right) h_{n-1} + \dots = o(1)h_n + o(1)h_{n-1} + \dots + o(1)h_1 \quad \text{per } |h| \rightarrow 0. \end{aligned}$$

dove la penultima uguaglianza segue da come avevamo definito $L(h)$ e l'ultima dalla continuità delle derivate parziali che avevamo richiesto nelle ipotesi. Dunque, per $|h| \rightarrow 0$, osserviamo che $o(1)h_j \leq o(1)|h| \implies \frac{o(1)h_j}{|h|} \leq \frac{o(1)|h|}{|h|} = o(1) \implies o(1)h_j = o(h)$ e chiaramente $\frac{o(1)h_n + \dots + o(1)h_1}{|h|} \rightarrow 0$. Dunque

$$f(p_0 + h) - f(p_0) - L(h) = o(h)$$

ovvero la tesi. □

Prima di procedere con la prossima dimostrazione, introduciamo la norma di Hilbert-Schmidt

Definizione 3.7 (norma di Hilbert-Schmidt). Sia $A \in \mathbb{R}^{m \times n}$, definiamo

$$\|A\| = \sqrt{\sum a_{ij}^2} \text{ è la norma di Hilbert-Schmidt}$$

Oss.: $|Lx| \leq \|A\| |x| \forall x \in \mathbb{R}^n$

Teorema 3.5 (differenziale della composizione). Sia $f : \Omega \rightarrow \mathbb{R}^m$, $g : U \rightarrow \mathbb{R}^k$ con $\Omega \subseteq \mathbb{R}^n$ aperto, $U \subseteq \mathbb{R}^m$ aperto, $f(\Omega) \subseteq U$ e sia $p_0 \in \Omega$. Allora se f differenziabile in p_0 e g differenziabile in $f(p_0)$ allora

$$g \circ f : \Omega \rightarrow \mathbb{R}^k \text{ è differenziabile in } p_0$$

Inoltre, in tal caso, avremo che

- ① $d(g \circ f)(p_0) = dg(f(p_0)) \circ df(p_0)$
- ② $D(g \circ f)(p_0) = (Dg(f(p_0)))(Df(p_0))$

Dimostrazione. Osserviamo che, per l'ipotesi di differenziabilità di f , abbiamo che

$$f(x) = f(p_0) + df(p_0)(p - p_0) + o(|p - p_0|) \text{ per } p \rightarrow p_0 \in \Omega$$

e, siccome g è differenziabile in $f(p_0)$,abbiamo che

$$g(x) = g(f(p_0)) + dg(f(p_0))[f(x) - f(p_0)] + o(|f(x) - f(p_0)|)$$

ma allora

$$g(x) = g(f(p_0)) + dg(f(p_0))[df(p_0)(p - p_0) + o(|p - p_0|)] + o(|f(x) - f(p_0)|)$$

ovvero

$$g(x) = g(f(p_0)) + (dg(f(p_0)) \circ df(p_0))(x - p_0) + dg(f(x_0))o(|x - x_0|) + o(|f(x) - f(p_0)|)$$

Ma adesso osserviamo che

$$\left| \frac{dg(f(x_0))o(|x - x_0|)}{|x - x_0|} \right| \leq |Dg(f(x_0))| \frac{o(|x - x_0|)}{|x - x_0|} = o(1) \implies dg(f(x_0))o(|x - x_0|) = o(|x - x_0|)$$

e, inoltre, osserviamo che

$$o(|f(x) - f(x_0)|) = o(|df(x_0)(x - x_0) + o(|x - x_0|)|).$$

A questo punto fissiamo $0 < \varepsilon < 1$, da cui deduciamo che

$$o(|f(x) - f(x_0)|) = o(|df(x_0)(x - x_0) + o(|x - x_0|)|)$$

dunque

$$o(|f(x) - f(x_0)|) < \varepsilon |df(x_0)(x - x_0) + o(|x - x_0|)|$$

e consideriamo $|x - x_0| < \delta_\varepsilon$ tale che $|df(x_0)(x - x_0) + o(|x - x_0|)| < \sigma_\varepsilon$. Sappiamo allora, per definizione di o -piccolo, che $\exists \bar{\delta}_\varepsilon > 0 : |f(x) - f(x_0)| < \varepsilon |x - x_0|$, dunque prendiam $|x - x_0| <$

$\min(\delta_\varepsilon, \bar{\delta}_\varepsilon)$ affinché valgano entrambe le disuguaglianze dette sopra e, a questo punto, facendo qualche piccola minorazione possiamo dedurre che

$$\begin{aligned} o(|f(x) - f(x_0)|) &< \varepsilon |f(x) - f(x_0)| < \varepsilon |df(x_0)(x - x_0) + o(|x - x_0|)| < \varepsilon |df(x_0)(x - x_0)| + \\ &\varepsilon |x - x_0| < \varepsilon (|Df(x_0)| |x - x_0| + \varepsilon |x - x_0|) = \varepsilon (|Df(x_0)| + 1) |x - x_0| \implies \\ &\implies o(|f(x) - f(x_0)|) = o(|x - x_0|) \end{aligned}$$

Dunque si ottiene che

$$g(f(x)) = g(f(x_0)) + (dg(f(x_0)) \circ df(x_0))(x - x_0) + o(|x - x_0|) \text{ per } x \rightarrow x_0$$

Osserviamo infine che

$$\begin{aligned} d(g \circ f)(x_0)(v) &= D(g \circ f)(x_0)(v) = dg(f(x_0))(df(x_0)(v)) = Dg(f(x_0))Df(x_0)v \implies \\ &\implies D(g \circ f)(x_0) = Dg(f(x_0))Df(x_0) \end{aligned}$$

□

3.1.1 Interpretazione geometrica delle derivate prime

Nel caso delle funzioni a singola variabile, la derivata di $f : \Omega \rightarrow \mathbb{R}$ (dove $\Omega \subseteq \mathbb{R}$ aperto) nel punto $x_0 \in \Omega$ rappresenta il coefficiente angolare della retta tangente al grafico della funzione f nel punto $(x_0, f(x_0))$. In più variabili la differenziabilità che significato ha?

Mostriamo innanzitutto che una funzione a singola variabile a valori in \mathbb{R} derivabile è equivalente a dire che essa è differenziabile:

Proposizione 3.3 (differenzialità \iff derivabilità). Sia $f : \Omega \rightarrow \mathbb{R}$ con $\Omega \subseteq \mathbb{R}$ aperto e sia $x_0 \in \Omega$. Allora

$$f \text{ è differenziabile in } x_0 \iff f \text{ è derivabile in } x_0$$

Dimostrazione.

\Rightarrow : osserviamo che se f è derivabile in x_0 allora

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) \implies \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} = 0$$

ma per definizione di o -piccolo abbiamo che

$$f(x) - f(x_0) - f'(x_0)(x - x_0) = o(x - x_0) \implies f(x_0 + h) - f(x_0) - f'(x_0)h = o(h)$$

dunque² $f(x)$ è differenziabile siccome $L(h) = f'(x_0)h$ è una funzione lineare.

\Leftarrow : osserviamo che se f è differenziabile in x_0 allora

$$f(x) - f(x_0) - L(x - x_0) = o(x - x_0)$$

dove $L(x - x_0) = a(x_0)(x - x_0)$ (siccome le uniche funzioni lineari in \mathbb{R} sono le rette e, nel nostro caso, il coefficiente angolare dipenderà sicuramente dal punto x_0 considerato e questo motiva la notazione usata), ottenendo che

$$\frac{f(x) - f(x_0)}{x - x_0} - a(x_0) = 0 \text{ per } x \rightarrow x_0$$

dunque $\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = a(x_0) = f'(x_0)$ (per unicità del limite).

□

²il cambio di variabile l'ho fatto per tirare fuori la definizione di differenziabilità per come l'avevamo espressa

Dunque la differenziabilità è l'unica nozione che generalizza il concetto di derivabilità a dimensioni superiori a 1. Però ha ancora una valenza geometrica?

Per rispondere a questa domanda dobbiamo ricordare al lettore qualche concetto proveniente dal corso di Geometria: sappiamo che l'equazione di un piano in \mathbb{R}^3 è data dalla formula

$$ax + by + cz + d = 0 \quad (3.2)$$

L'equazione qua sopra può essere ottenuta in due modi: il primo consiste nel trovare una base del piano³ e, successivamente, ricavare l'equazione prendendo la matrice che ha come colonne i vettori della base più un vettore di coordinate generiche (x, y, z) che supponiamo appartenere al piano: la matrice avrà, conseguentemente, determinante nullo; ottenendo la formula attesa sopra attesa. Usufruendo, invece, del prodotto scalare possiamo osservare che, presa una base ortogonale del piano, allora possiamo decomporre lo spazio vettoriale \mathbb{R}^3 in due sottospazi: il piano (di dimensione 2) che indicheremo W e lo Span del vettore ortogonale al piano (di dimensione 1) W^\perp (riprendendo la notazione del corso di Geometria), dunque, siccome sappiamo dal corso di Geometria che ogni combinazione lineare dei vettori appartenenti a W sarà ortogonale a W^\perp , possiamo definire il nostro piano come l'insieme dei vettori il cui prodotto scalare con il vettore ortogonale è nullo. Questo ragionamento effettuato è facilmente estendibile a spazi euclidei muniti del prodotto scalare canonico di dimensione n per iperpiani di dimensione $n - 1$: infatti, presi $n - 1$ vettori linearmente ortogonali come base dell'iperpiano che indichiamo con W , allora W^\perp avrà dimensione 1 (grazie all'ipotesi di coercività del prodotto scalare) dunque $W^\perp = \text{Span}(v)$ con v vettore ortogonale; a questo punto possiamo ottenere la formula mostrata sopra osservando, come fatto prima, che ogni combinazione lineare dei vettori appartenenti a W sarà ortogonale a W^\perp e, quindi, indicare il piano come l'insieme dei vettori per cui il prodotto scalare con il vettore v è nullo. Vedendo questo procedimento in maniera esplicita, possiamo considerare un piano $\pi \in \mathbb{R}^3$ di dimensione pari a 2 da cui abbiamo che $\dim(\pi)^\perp = \dim \mathbb{R}^3 - \dim \pi + \dim \pi \cap (\mathbb{R}^n)^\perp = 1$ e, prendendo un generico vettore $v = (a, b, c)$ (ovvero la base del nostro $(\pi)^\perp$), se vogliamo trovare l'equazione di un piano affine, allora, fissato un punto $P_0 = (x_0, y_0, z_0) \in \pi$ preso $P = (x, y, z)$ appartenente al piano, allora per quanto detto, dobbiamo avere che

$$\vec{P_0P} \perp v$$

dunque

$$(x - x_0 \ y - y_0 \ z - z_0) \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} a \\ b \\ c \end{pmatrix} = 0 \implies a(x - x_0) + b(y - y_0) + c(z - z_0) = 0$$

da cui possiamo facilmente ottenere che

$$ax + by + cz + d = 0$$

ponendo $d = -ax_0 - by_0 - cz_0$. Dunque, ripensando al procedimento effettuato, possiamo ben motivare la seguente definizione

Definizione 3.8 (piano in uno spazio euclideo). Un piano nello spazio euclideo \mathbb{R}^3 (oppure un iperpiano di dimensione $n - 1$ in \mathbb{R}^{n-1}) si può vedere come il luogo geometrico dei punti perpendicolari ad una data direzione $v = (a, b, c) \neq 0_{\mathbb{R}^3}$ (equivalentemente, per l'iperpiano, $v = (x_1, \dots, x_n) \neq 0_{\mathbb{R}^n}$). Dunque, preso $p \in \mathbb{R}^3$ (per l'iperpiano $p \in \mathbb{R}^n$), possiamo trovare l'equazione

³più correttamente, si dovrebbe parlare di giacitura perché a priori lavoriamo con dei piani che dobbiamo supporre affini. Ricordiamo però dal corso di Geometria, che un sottospazio affine è uno spazio vettoriale "traslato" di un vettore e tale spazio vettoriale, che avrà dunque una sua base, è ciò che prima ho chiamato giacitura

cartesiana del piano passante per $p \in \mathbb{R}^3$ (per l'iperpiano $p \in \mathbb{R}^n$) perpendicolare a v tramite la seguente relazione:

$$\langle (x, y, z) - p, v \rangle \quad (3.3)$$

Sia adesso $f : \Omega \rightarrow \mathbb{R}$, con $\Omega \subseteq \mathbb{R}^m$ aperto con f differenziabile in $p_0 \in \Omega$. Allora osserviamo che

$$\begin{aligned} \pi &= \{(x, L_p(x)) : x \in \mathbb{R}^n\} = \{(x, t) \in \mathbb{R}^{n+1} : t = f(p) + \langle \nabla f(p), x - p \rangle\} = \\ &= \{(x, t) \in \mathbb{R}^{n+1} : \langle (-\nabla f(x_0), 1), (x - p, t - f(p)) \rangle = 0\} \end{aligned}$$

Dunque possiamo interpretare come la differenziabilità della funzione in un punto come quella funzione che descrive il piano tangente alla funzione nel punto.

3.1.2 Legame fra differenziabilità e continuità

Mostriamo adesso il legame fra differenziabilità e continuità. Per fare questo dobbiamo mostrare il seguente lemma:

Lemma 3.5.1. Ogni funzione lineare $L : \mathbb{R}^n \rightarrow \mathbb{R}^m$ continua e se $A = (a_{ij}) \in \mathbb{R}^{m \times n}$ è la sua matrice allora $|Lx| \leq \|A\| |x| \forall x \in \mathbb{R}^n$, dove

$$\|A\| = \sqrt{\sum a_{ij}^2}$$

è la norma di Hilbert-Schmidt

Dimostrazione. Dalla regola del prodotto matriciale abbiamo che

$$|Lx|^2 = \left| \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j \right) e_i \right|^2 = \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij} x_j \right)^2 \leq \sum_{i=1}^m \left(\sum_{j=1}^n a_{ij}^2 \right) \left(\sum_{j=1}^n x_j^2 \right) = \|A\|^2 |x|^2$$

e possiamo osservare che L è continua, infatti se $x_0 \in \mathbb{R}^n$

$$|Lx - Lx_0| = |L(x - x_0)| \leq \|A\| |x - x_0|$$

pertanto se $p_k \rightarrow x_0$ allora

$$|Lp_k - Lx_0| \leq \|A\| |p_k - x_0| \xrightarrow{k \rightarrow +\infty} 0$$

□

A questo punto possiamo dunque mostrare la seguente proposizione

Proposizione 3.4. Se $f : \Omega \rightarrow \mathbb{R}^m$, $p_0 \in \Omega$ differenziabile in p_0 , allora f è continua in p_0

Dimostrazione. Osserviamo che se f è differenziale in p_0 allora $\exists df(p_0) : \Omega \rightarrow \mathbb{R}^m$ tale che

$$f(p_0 + h) = f(p_0) + df(p_0)h + o(h) \xrightarrow{h \rightarrow 0} f(p_0)$$

in quanto $df(p_0)$ è lineare e quindi continua in \mathbb{R}^n e $o(h) \xrightarrow{h \rightarrow 0} 0$

□

3.2 Derivate di ordine successivo

Per trattare i massimi e i minimi in più variabili è necessario ricostruire il comportamento della funzione anche al "secondo ordine" (riprendendo la terminologia del polinomio di Taylor che verrà comunque ripreso più avanti), ovvero osservare le derivate seconde (che non abbiamo ancora introdotto, però la sostanza rimane a grandi linee quella nel caso ad una singola variabile): se nel caso delle funzioni a singola variabile non era necessario studiare la concavità della funzione (che ricordiamo essere il comportamento al secondo ordine della funzione a variabile singola), nel caso a più variabile è quasi inevitabile (a meno che non si voglia procedere per ragionamenti più "raffinati"). Andiamo dunque a vedere il concetto di derivate di ordine superiore al primo.

In generale, presa la funzione $f : \Omega \rightarrow \mathbb{R}$, $\Omega \subseteq \mathbb{R}^n$ aperto tale che $\exists \partial_{x_i} f : \Omega \rightarrow \mathbb{R}$ con $i \in \{1, \dots, n\}$, allora la derivata parziale rispetto a x_i o anche derivata parziale i -esima può ammettere un'ulteriore derivata parziale rispetto a x_j :

$$\partial_{x_j}(\partial_{x_i} f) = \frac{\partial^2 f}{\partial x_j \partial x_i} = (f_{x_i})_{x_j} = f_{x_i x_j}$$

(si faccia attenzione alla notazione, in cui viene **sempre** specificato l'ordine in cui sono fatte le derivate). Di fatto potremmo definire anche le derivate terze, quarte, ecc.

$$\partial_{x_3} \partial_{x_2} \partial_{x_1} f = f_{x_1 x_2 x_3} = \frac{\partial^3}{\partial x_3 \partial x_2 \partial x_1}$$

In generale, ove esistono, le derivate di ordine k possono essere denotate nella seguente maniera:

$$\partial_{x_{i_k}} \partial_{x_{i_{k-1}}} \dots \partial_{x_{i_1}} f = \frac{\partial^k f}{\partial x_{i_k} \dots \partial x_{i_1}} = f_{x_{i_1} \dots x_{i_k}}$$

tal derivata parziale iterata diremo essere la derivata parziale di ordine k o derivata parziale k -esima rispetto a x_{i_1}, \dots, x_{i_k}

Teorema 3.6 (di Schwarz). Sia $f : \Omega \rightarrow \mathbb{R}$, $x_0 \in \Omega \subseteq \mathbb{R}^n$ con Ω aperto e $\exists \partial_{x_i} f, \partial_{x_j} f, \partial_{x_i x_j} f, \partial_{x_j x_i} f$ continue in x_0 . Allora

$$\partial_{x_i x_j} f(x_0) = \partial_{x_j x_i} f(x_0)$$

Dimostrazione. Consideriamo la funzione

$$\varphi(\tau) = f(x_0 + \tau e_i + t e_j) - f(x_0 + \tau e_i)$$

dove t è sufficientemente piccolo e fissato, dunque:

$$\begin{aligned} \varphi(t) - \varphi(0) &= f(x_0 + t e_i + t e_j) - f(x_0 + t e_i) - (f(x_0 + t e_j) - f(x_0)) = \\ &= f(x_0 + t e_i + t e_j) - f(x_0 + t e_i) - f(x_0 + t e_j) + f(x_0) \end{aligned}$$

Adesso sappiamo, per il teorema di Lagrange, che

$$\varphi(t) - \varphi(0) = \varphi'(\theta t)t$$

con $0 < \theta < 1$. Possiamo "accoppare" i termini come

$$\varphi(t) - \varphi(0) = [(f(x_0 + t e_i + t e_j) - f(x_0 + t e_j)) - (f(x_0 + t e_i) - f(x_0))]t$$

e, per il teorema di Lagrange, definendo le funzioni $z(\tau) = f(x_0 + \tau e_i + t e_j)$ e $h(\tau) = f(x_0 + \tau e_i)$ allora avremo che

$$\varphi(t) - \varphi(0) = [(z(t) - z(0)) - (h(t) - h(0))]t = (\partial_{x_i} f(x_0 + \theta t e_i + t e_j) - \partial_{x_i} f(x_0 + \theta t e_i))t$$

Possiamo riapplicare nuovamente il teorema di Lagrange su questa differenza, osservando che

$$\begin{aligned} [\partial_{x_i} f(x_0 + \theta t e_i + t e_j) - \partial_{x_i} f(x_0 + \theta t e_i)]t &= [\partial_{x_i}(f(x_0 + \theta t e_i + t e_j) - f(x_0 + \theta t e_i))]t = \\ &= \partial_{x_i}(\partial_{x_j} f)(x_0 + \theta t e_i + t e_j)t^2 \end{aligned}$$

Ripetendo questo ragionamento invertendo gli indici, possiamo dunque concludere, per continuità delle derivate seconde nel punto x_0 , che:

$$\lim_{t \rightarrow 0} \frac{\varphi(t) - \varphi(0)}{t^2} = \partial_{x_i} \partial_{x_j} f(x_0) = \partial_{x_j} \partial_{x_i} f(x_0)$$

□

Oss.: Con degli argomenti più sofisticati è possibile indebolire ulteriormente le ipotesi richiedendo che le derivate parziali siano differenziabili in p_0 .

Teorema 3.7 (teorema di Schwarz di ordine superiore). Sia $f : \Omega \rightarrow \mathbb{R}$, Ω aperto e $\Omega \subseteq \mathbb{R}^n$. Supponiamo che esistano tutte le derivate di f fino all'ordine k e che siano continue in $\xi \in \Omega$. Allora per ogni scelta di indici $1 \leq i_1, \dots, i_k \leq n$ abbiamo

$$\frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}}(\xi) = \frac{\partial^k f}{\partial x_{\sigma(1)} \partial x_{\sigma(2)} \dots \partial x_{\sigma(k)}}(\xi)$$

dove $\sigma \in \Sigma_k$.

Dimostrazione. La dimostrazione è omessa siccome è un banale corollario del precedente teorema di Schwarz. □

E' adesso comodo introdurre i cosiddetti spazi funzionali $C^k(\Omega)$, con $k \in \mathbb{N} \setminus \{0\}$:

Definizione 3.9 (spazio C^k). Sia $\Omega \subseteq \mathbb{R}^n$ aperto e $k \in \mathbb{N} \setminus \{0\}$. Definiamo

$$C^k(\Omega) = \left\{ f : \Omega \rightarrow \mathbb{R} \mid \forall l \in \{1, \dots, k\}, \exists \partial_{x_{i_1}} \partial_{x_{i_2}} \dots \partial_{x_{i_l}} f : \Omega \rightarrow \mathbb{R} \text{ continua in } \Omega \text{ per ogni } i_1, \dots, i_l \in \{1, \dots, n\} \right\}$$

Se $f \in C^k(\Omega)$ allora diremo che f è una funzione di classe C^k .

Oss.: Possiamo estendere la definizione che abbiamo dato a $k = 0$ definendo lo spazio C^0 come

$$C^0(\Omega) = \{f : \Omega \rightarrow \mathbb{R}, f \text{ continua in } \Omega\}$$

osserviamo che $C^m(\Omega) \subseteq C^k(\Omega) \forall k, m : k \leq m$.

Definiamo infine $C^\infty(\Omega) = \bigcap_{k=1}^{+\infty} C^k(\Omega)$ lo spazio delle funzioni $C^{+\infty}$, ovvero lo spazio delle funzioni che hanno derivate parziali di ogni ordine.

Oss.: Se $f \in C^1(\Omega)$ allora è differenziabile ovunque in Ω dal teorema del differenziale totale

Oss.: Se $f \in C^k(\Omega)$ tutte le derivate parziali fino all'ordine $k - 1$ sono differenziabili (sempre per il teorema del differenziale totale)

Dunque possiamo generalizzare il teorema di Schwarz di ordine superiore che abbiamo visto prima

Teorema 3.8. Sia $f \in C^k(\Omega)$, $\Omega \subseteq \mathbb{R}^n$ aperto. Allora $\forall 1 \leq i_1, \dots, i_l \leq n$ e $2 \leq l \leq k$, definendo $\sigma \in \sum_l$, abbiamo che

$$\partial_{x_l} \partial_{x_{l-1}} \dots \partial_{x_1} f = \partial_{x_{\sigma(l)}} \partial_{x_{\sigma(l-1)}} \dots \partial_{x_{\sigma(1)}} f$$

Dimostrazione. Segue banalmente dal teorema precedente. \square

Questo teorema è molto importante siccome stabilisce che, per funzioni di classe C^k , l'ordine con cui deriviamo per ottenere una derivata parziale iterata non modifica la derivata.

Adesso abbiamo tutti gli strumenti per introdurre i differenziali di ordine superiore

Definizione 3.10 (differenziali di ordine superiore). Sia $f \in C^k(\Omega)$, $\Omega \subseteq \mathbb{R}^n$ aperto, $x_0 \in \Omega$ e preso $1 \leq l \leq k$ definiamo:

$$d^l f(x_0) = \sum_{i_1, \dots, i_l=1}^n \frac{\partial^l f}{\partial x_{i_1} \dots \partial x_{i_l}} h_{i_1} \dots h_{i_l}$$

posto $h = (h_1, \dots, h_n) \in \mathbb{R}^n$ allora diciamo che $d^l f(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}$ è detto differenziale di ordine l

Oss.: Sia $f \in C^k(\Omega)$, $x_0 \in \Omega$, $v \in \mathbb{R}^n$ con $v = (\xi_1, \dots, \xi_n) = \sum_{j=1}^n \xi_j e_j$. Allora abbiamo che

$$\begin{aligned} \frac{d}{dt} f(x_0 + tv) &= \sum_{i=1}^n \partial_{x_i} f(x_0 + tv) \xi_i \implies \frac{d^2}{dt^2} f(x_0 + tv) = \sum_{i_1, i_2=1}^n \partial_{x_{i_1}} \partial_{x_{i_2}} f(x_0 + tv) \xi_{i_1} \xi_{i_2} \stackrel{\text{per induz.}}{\implies} \\ \stackrel{\text{per induz.}}{\implies} \frac{d^k}{dt^k} f(x_0 + tv) &= d^k f(x_0 + tv) \implies \\ \implies \boxed{\frac{\partial^k}{\partial v^k} f(x_0) = d^l f(x_0)(v)} & \end{aligned} \quad (3.4)$$

Come per le derivate prime, possiamo "definire" anche una matrice delle derivate. Particolarmente importante, oltre alla matrice jacobiana, è la matrice hessiana, ovvero la matrice delle derivate seconde:

$$Hf(x_0) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(x_0) & \frac{\partial^2 f}{\partial x_2 \partial x_1}(x_0) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_1}(x_0) \\ \frac{\partial^2 f}{\partial x_1 \partial x_2}(x_0) & \frac{\partial^2 f}{\partial x_2^2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_n \partial x_2}(x_0) \\ \vdots & \vdots & \dots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(x_0) & \frac{\partial^2 f}{\partial x_n \partial x_2}(x_0) & \dots & \frac{\partial^2 f}{\partial x_n^2}(x_0) \end{pmatrix} \quad (3.5)$$

Per il teorema di Schwarz, abbiamo che per funzioni di cui esistono derivate seconde continue, allora

$$Hf(x_0) = Hf(x_0)^T$$

dunque la matrice hessiana è una matrice quadrata simmetrica. Inoltre, in virtù della definizione fatta prima sui differenziali di ordine successivo, avremo che

$$d^2 f(x_0)(v) = \langle v, Hf(x_0)(v) \rangle$$

e quindi è omogeneo di grado 2, ovvero

$$d^2 f(x_0)(tv) = t^2 d^2 f(x_0)(v) \quad \forall t \in \mathbb{R}, \forall v \in \mathbb{R}^n$$

In sintesi $df^2(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}$ è un polinomio omogeneo (di cui daremo la definizione formale nel prossimo paragrafo) di secondo grado in n variabili.

3.3 Funzioni polinomiali omogenee e polinomio di Taylor

Riprendiamo il concetto di forma quadratica dal corso di Geometria

Definizione 3.11 (forma quadratica). $g : \mathbb{R}^n \rightarrow \mathbb{R}$ è una forma quadratica se $\exists A = A^T \in \mathbb{R}^n$ tale che $g(x) = \langle x, Ax \rangle \iff g(x) = \sum_{i,j=1}^n a_{ij}x_i x_j$ con $A = (a_{ij})_{\substack{i=1, \dots, n \\ j=1, \dots, n}}$

Proposizione 3.5. Sia $q(x)$ una forma quadratica, allora

$$q(x) = \sum_{i=1}^n a_{ii}x_i^2 + 2 \sum_{1 \leq i < j \leq n} a_{ij}x_i x_j$$

pertanto possiamo scrivere il nostro polinomio $q(x)$ quadratico e omogeneo come

$$h(x) = \langle x, Bx \rangle$$

dove

$$B = \begin{pmatrix} b_{11} & \frac{b_{12}}{2} & \frac{b_{13}}{2} & \cdots & \frac{b_{1n}}{2} \\ \frac{b_{21}}{2} & b_{22} & \frac{b_{23}}{2} & \cdots & \frac{b_{2n}}{2} \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ \frac{b_{n1}}{2} & \frac{b_{n2}}{2} & \cdots & \cdots & b_{nn} \end{pmatrix}$$

Dimostrazione. In virtù della definizione precedente sappiamo che

$$q(x) = \sum_{i,j=1}^n b_{ij}x_i x_j = \sum_{i=1}^n a_{ii}x_i^2 + \sum_{1 \leq i < j \leq n} b_{ij}x_i x_j$$

con un semplice "riarrangiamento" dei termini presenti. Si vede facilmente che

$$B = \begin{pmatrix} b_{11} & \frac{b_{12}}{2} & \cdots & \frac{b_{1n}}{2} \\ \frac{b_{21}}{2} & b_{22} & \cdots & \frac{b_{2n}}{2} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{b_{n1}}{2} & \cdots & \cdots & b_{nn} \end{pmatrix}$$

dunque, utilizzando le semplici proprietà dei prodotti scalari, abbiamo che

$$q(x) = \langle x, Bx \rangle$$

□

Possiamo generalizzare questo concetto anche a polinomi di n variabili $P : \mathbb{R}^n \rightarrow \mathbb{R}$

Definizione 3.12 (polinomio di n variabili omogeneo). Sia $P : \mathbb{R}^n \rightarrow \mathbb{R}$ un polinomio in n variabili, quindi esiste una famiglia di numeri $c_{i_1, \dots, i_l} \in \mathbb{R}$, $1 \leq l \leq k$ tale che

$$P(x) = \sum_{l=1}^k \sum_{(i_1, \dots, i_l) \in \mathcal{I}_l} c_{i_1, \dots, i_l} x_{i_1} \dots x_{i_l}$$

dove $\mathcal{I}_l \subseteq \mathbb{N}^2$ è un sottoinsieme di indici, diremo che P è omogeneo di grado $k \geq 1$ se

$$\forall x \in \mathbb{R}^n \forall t \in \mathbb{R}, P(tx) = t^k P(x)$$

tutti i polinomi omogenei di grado k saranno quindi della forma

$$P(x) = \sum_{(i_1, \dots, i_k) \in \mathcal{I}_k} c_{i_1, \dots, i_k} x_{i_1} \dots x_{i_k}$$

Il differenziale $d^k f(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}$ di ordine k è un polinomio in n variabili omogeneo di ordine k . Abbiamo infatti per $t \in \mathbb{R}, x \in \mathbb{R}^n$

$$\begin{aligned} d^k f(x_0) &= \sum_{i_1, \dots, i_k=1}^n \frac{\partial^k f(x_0)}{\partial x_{i_1} \dots \partial x_{i_k}} (tx_{i_1})(tx_{i_2}) \dots (tx_{i_k}) = t^k \sum_{i_1, \dots, i_k=1}^n \frac{\partial^k f(x_0)}{\partial x_{i_1} \dots \partial x_{i_k}} x_1 \dots x_{i_k} = \\ &= t^k d^k f(x_0) \end{aligned}$$

Possiamo a questo punto definire il polinomio di Taylor

Definizione 3.13 (polinomio di Taylor di f in x_0 di ordine k). Sia $f \in C^k(\Omega), \Omega \subseteq \mathbb{R}^n$ aperto. Definiamo il polinomio di Taylor di f nel punto $x_0 \in \Omega$ di ordine k come quel polinomio di grado $\leq k$ tale che

$$P_k(f, x_0)(x) = \sum_{i=0}^k \frac{d^i f(x_0)(x - x_0)}{i!}, x \in \mathbb{R}^n$$

Andiamo adesso a studiare alcune proprietà che caratterizzano il polinomio di Taylor, partendo dalla seguente proposizione

Proposizione 3.6. Sia $f \in C^k(\Omega), x_0 \in \Omega$. Allora, per ogni $1 \leq i_1, \dots, i_l \leq n$ e $1 \leq l \leq k$,

$$\partial_{x_{i_1}} \dots \partial_{x_{i_l}} f(x_0) = \partial_{x_{i_1}} \dots \partial_{x_{i_l}} P_k(f, x_0)(x_0)$$

Dimostrazione. Consideriamo

$$\partial_{x_{i_1}} \dots \partial_{x_{i_l}} P_k(f, x_0) = \partial_{x_{i_1}} \dots \partial_{x_{i_l}} \sum_{p=0}^k \frac{1}{p!} \sum_{j_1, \dots, j_p=1}^n \partial_{x_{j_1}} \dots \partial_{x_{j_p}} f(x_0) (x_{j_1} - x_{0,j_1}) \dots (x_{j_p} - x_{0,j_p}).$$

Naturalmente abbiamo che, siccome stiamo derivando l volte, le prime l derivate vanno a 0, dunque i primi contributi in questa sommatoria saranno dati da $p \geq l$ e, siccome stiamo operando su sommatorie finite, possiamo sfruttare la linearità delle derivate parziali su somme finite:

$$\begin{aligned} &\sum_{p=l}^k \frac{1}{p!} \sum_{j_1, \dots, j_p=1}^n \partial_{x_{j_1}} \dots \partial_{x_{j_p}} f(x_0) \partial_{x_{i_1}} \dots \partial_{x_{i_l}} \left((x_{j_1} - x_{0,j_1}) \dots (x_{j_p} - x_{0,j_p}) \right) = \\ &= \frac{1}{l!} \sum_{j_1, \dots, j_l=1}^n \partial_{x_{j_1}} \dots \partial_{x_{j_l}} f(x_0) \partial_{x_{i_1}} \dots \partial_{x_{i_l}} \left((x_{j_1} - x_{0,j_1}) \dots (x_{j_l} - x_{0,j_l}) \right) \\ &+ \sum_{p=l+1}^k \frac{1}{p!} \sum_{j_1, \dots, j_p=1}^n \partial_{x_{j_1}} \dots \partial_{x_{j_p}} f(x_0) \partial_{x_{i_1}} \dots \partial_{x_{i_l}} \left((x_{j_1} - x_{0,j_1}) \dots (x_{j_p} - x_{0,j_p}) \right). \end{aligned}$$

Ora osserviamo che il secondo termine di questa addizione è nullo per $x \rightarrow x_0$ mentre per il primo termine, in virtù del fatto $f \in C^k(\Omega)$, abbiamo che le derivate, indipendentemente dall'ordine, saranno tutte uguali, dunque avremo che

$$\begin{aligned} \partial_{x_{i_1}} \dots \partial_{x_{i_l}} P_k(f, x_0) &= \frac{1}{l!} \sum_{j_1, \dots, j_l=1}^n \partial_{x_{j_1}} \dots \partial_{x_{j_l}} f(x_0) \partial_{x_{i_1}} \dots \partial_{x_{i_l}} \left((x_{j_1} - x_{0,j_1}) \dots (x_{j_l} - x_{0,j_l}) \right) + \\ &\quad \sum_{p=l+1}^k \frac{1}{p!} \sum_{j_1, \dots, j_p=1}^n \partial_{x_{j_1}} \dots \partial_{x_{j_p}} f(x_0) \partial_{x_{i_1}} \dots \partial_{x_{i_l}} \left((x_{j_1} - x_{0,j_1}) \dots (x_{j_p} - x_{0,j_p}) \right) \end{aligned}$$

Osserviamo che la sommatoria

$$\sum_{p=l+1}^n \frac{1}{p!} \sum_{j_1, \dots, j_l}^n \partial_{x_{j_1}} \dots \partial_{x_{j_p}} f(x_0) \partial_{x_{i_1}} \dots \partial_{x_{i_l}} \left((x_{j_1} - x_{0_{j_1}}) \dots (x_{j_p} - x_{0_{j_p}}) \right)$$

tende a 0 per $x \rightarrow x_0$ siccome se valutata in x_0 , per continuità dei termini, tende a 0. Per quanto riguarda il primo termine, osserviamo che della sommatoria sugli indici j_1, \dots, j_l sopravvive solamente i termini con indici che rappresentano una permutazione di (i_1, \dots, i_l) che, in virtù del teorema di Schwarz di ordine superiore, abbiamo che le derivate, indipendentemente dall'ordine in cui deriviamo, coincidono

$$\sum_{\substack{\sigma: \{i_1, \dots, i_l\} \rightarrow \{i_1, \dots, i_l\} \\ \sigma \text{ bigettiva}}} \partial_{x_{i_1}} \dots \partial_{x_{i_l}} f(x_0) = \partial_{x_{i_1}} \dots \partial_{x_{i_l}} f(x_0) l!$$

da cui segue la tesi, in quanto

$$\partial_{x_{i_1}} \dots \partial_{x_{i_l}} P_k(f, x_0) = \frac{1}{l!} \partial_{x_{i_1}} \dots \partial_{x_{i_l}} f(x_0) l! + o(1) \text{ per } x \rightarrow x_0$$

□

Teorema 3.9 (formula di Taylor). Se $f \in C^k(\Omega)$, $x_0 \in \Omega$, $\Omega \subseteq \mathbb{R}^n$ aperto, allora abbiamo che

$$f(x) = P_k(f, x_0) + o(|x - x_0|^k) = \sum_{j=0}^k \frac{1}{j!} d^j f(x_0)(x - x_0)^j + o(|x - x_0|^k) \text{ per } x \rightarrow x_0$$

Dimostrazione. Faremo questa dimostrazione per induzione su k . Per ora studiamo la funzione $g(x) = f(x) - P_k(f, x_0)(x)$ senza fissare un particolare valore di k (vogliamo innanzitutto studiare quali sono i passaggi comuni a tutti i valori di k). Osserviamo, grazie alla proposizione precedente, che $g(x_0) = 0$ e $\forall 1 \leq i_1, \dots, i_l \leq n$ e $\forall 1 \leq l \leq k$.

Facciamo il cambio di variabile $x = x_0 + h$ e consideriamo il caso in cui $k = 1$: proviamo a questo punto che $g(x_0 + h) = g(x_0 + h) - g(x_0)$ siccome $g(x_0) = 0$. Ma allora

$$\begin{aligned} g(x_0 + h) - g(x_0) &= \int_0^1 \frac{d}{dt} g(x_0 + th) dt = \int_0^1 \partial_h g(x_0 + th) dt = \int_0^1 dg(x_0 + th)(h) dt = \\ &= \int_0^1 \sum_{j=1}^n \partial_{x_j} g(x_0 + th) h_j dt \end{aligned}$$

dove queste eguaglianze derivano dalla semplice osservazione del seguente fatto:

$$\partial_h g(x_0 + th) = \lim_{\Delta t \rightarrow 0} \frac{g(x_0 + (t + \Delta t)h) - g(x_0 + th)}{\Delta t} = \frac{d}{dt} g(x_0 + th)$$

e le altre dalle proprietà del differenziale e del gradiente. Concludiamo che

$$g(x_0 + h) = \int_0^1 \langle \nabla g(x_0 + th), h \rangle dt$$

da cui segue, tramite la 1.1 (diseguaglianza di Cauchy-Schwarz), che

$$\begin{aligned} |g(x_0 + h)| &= \left| \int_0^1 \langle \nabla g(x_0 + th), h \rangle dt \right| \leq \left| \int_0^1 |\langle \nabla g(x_0 + th), h \rangle| dt \right| \leq \\ &\leq \left\| \int_0^1 \nabla g(x_0 + th) dt \right\| |h| \end{aligned}$$

e poiché $\nabla g(x_0) = 0$ allora

$$\int_0^1 |\nabla g(x_0 + th)| = o(1) \text{ per } h \rightarrow 0$$

dunque $g(x_0 + h) \leq o(h) \implies g(x_0 + h) = o(h)$, ovvero $g(x) = o(x - x_0)$. Abbiamo quindi mostrato il caso $k = 1$, ma adesso, supponendo che questo sia vero per induzione per $k - 1$, e sapendo che $\partial_{x_{i_1}} \dots \partial_{x_{i_k}} g(x_0) = 0 \forall i_1 \dots i_k \in \{1, \dots, n\}$, vogliamo adesso mostrare che $g(x_0 + h) = o(h^k)$.

Osserviamo che $\partial_{x_i} g$ soddisfa le ipotesi della proprietà induttiva per $k - 1$, infatti abbiamo che

$$\partial_{x_i} g(x_0) = 0$$

e naturalmente

$$\partial_{x_{i_1}} \dots \partial_{x_{i_{k-1}}} \partial_{x_i} g(x_0) = 0 \forall i_1, \dots, i_{k-1} \in \{1, \dots, n\}$$

dunque, per ipotesi induttiva, abbiamo che

$$\partial_{x_i} g = o(|h|^{k-1})$$

da cui segue che

$$\begin{aligned} g(x_0 + h) &= \int_0^1 \langle \nabla g(x_0 + th), h \rangle dt = \int_0^1 \sum_{j=1}^n \partial_{x_j} g(x_0 + th) h_j dt \stackrel{\text{ip.induttiva}}{=} \\ &\stackrel{\text{ip. induttiva}}{=} \int_0^1 \sum_{j=1}^n o_j(|th|^{k-1}) h_j dt = \int_0^1 \sum_{j=1}^n o_j(|h|^{k-1}) h_j dt \implies \\ &\implies |g(x_0 + h)| \leq |h| \int_0^1 |(o_1(|h|^{k-1}), \dots, o_n(|h|^{k-1}))| dt = o(|h|^k) \text{ per } h \rightarrow 0 \end{aligned}$$

ovvero la tesi. \square

Siamo adesso pronti a mostrare l'unicità del teorema di Taylor, tuttavia abbiamo bisogno di qualche lemma preliminare:

Lema 3.9.1. Se Q è un polinomio di grado $k \geq 1$ omogeneo, allora $\exists x_0 \in \mathbb{R}^n$ tale che $Q(x_0) \neq 0$

Dimostrazione. Procediamo con la contro-nominale⁴ per $k \geq 1$: se $\forall x \in \mathbb{R}^n, Q(x) = 0 \implies Q(x) \equiv 0 \implies Q$ non è un polinomio omogeneo di ordine k siccome il grado del polinomio coincide col il grado di omogeneità del polinomio.

Per $k = 0$ allora il polinomio è, effettivamente, omogeneo. \square

Corollario 3.9.1. Sia $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ è un polinomio omogeneo di ordine k e sia $x_0 \in \mathbb{R}^n : Q(x_0) = 0$. Allora

$$Q\left(\frac{x_0}{|x_0|}\right) = 0$$

Dimostrazione. Per ipotesi di omogeneità $Q\left(\frac{x_0}{|x_0|}\right) = \frac{1}{|x_0|^k} Q(x_0) = 0$. \square

Lema 3.9.2. Sia $P : \mathbb{R}^n \rightarrow \mathbb{R}$ un polinomio in n -variabili di ordine $k \geq 1$. Se $P(x) = o(|x|^k)$ per $x \rightarrow 0 \implies P = 0$

⁴ricordiamo che la dimostrazione per la contro-nominale consiste nel dimostrare che $\neg B \implies \neg A$, infatti $A \implies B = \neg A \wedge B = \neg(A \wedge \neg B) = \neg B \implies \neg A$

Dimostrazione. Procediamo per assurdo: supponiamo che $\exists P(x) = o(k) \neq 0$ e definiamo j_0 come il grado non nullo più piccolo del polinomio. Consideriamo:

$$P(x) = \sum_{|\alpha| \leq k} c_\alpha x^\alpha \neq 0, c_\alpha \in \mathbb{R} \implies P(x) = \sum_{|\alpha|=j_0} c_\alpha x^\alpha + o(|x|^{j_0}) = q(x) + o(|x|^{j_0})$$

per qualche $j_0 \in \{1, \dots, k\}$ e q è un polinomio omogeneo di grado j_0 non nullo, ovvero

$$q(\lambda x) = \lambda^{j_0} q(x) \quad \forall \lambda \in \mathbb{R} \quad \forall x \in \mathbb{R}$$

Consideriamo il seguente rapporto

$$\frac{P(x)}{|x|^k} = \frac{1}{|x|^{k-j_0}} \left(\sum_{|\alpha|=j_0} c_\alpha \left(\frac{x}{|x|} \right)^\alpha + o(1) \right) = q \left(\frac{x}{|x|} \right) + o(1)$$

Sappiamo allora, per il lemma precedente, che $\exists \xi \in \mathbb{S}^{n-1} : q(\xi) \neq 0$, dunque considerando la quantità $\frac{t\xi}{|t|}$ per $t \rightarrow 0$ abbiamo che

$$\frac{P(t\xi)}{|t|^k} = \frac{1}{|t|^{k-j_0}} \left(q \left(\frac{t}{|t|} \xi \right) + o(1) \right) = \frac{1}{|t|^{k-j_0}} \left(\left(\frac{t}{|t|} \right)^{j_0} q(\xi) + o(1) \right) \not\rightarrow 0$$

che contraddice l'ipotesi che $P(x) = o(x^k)$. Abbiamo dunque la tesi. \square

Teorema 3.10 (unicità del polinomio di Taylor). Sia $f : \Omega \rightarrow \mathbb{R}$, $f \in C^k(\Omega)$, $x_0 \in \Omega$, $\Omega \subseteq \mathbb{R}^n$ aperto. Se

$$f(x) = Q(x - x_0) + o(|x - x_0|^k) \text{ per } x \rightarrow x_0$$

con $Q : \mathbb{R}^n \rightarrow \mathbb{R}$ polinomio di grado k allora

$$Q(x - x_0) = P_k(f, x_0)$$

Dimostrazione. Per il teorema 3.9 abbiamo che

$$f(x) = P_k(f, x_0) + o(|x - x_0|^k)$$

adesso, per ipotesi, abbiamo che

$$Q(x - x_0) + o(|x - x_0|^k) = f(x) = P_k(f, x_0)(x) + o(|x - x_0|^k)$$

ma allora, facendo il cambio di variabile $h = x - x_0$, abbiamo che

$$R(h) = \sum_{j=0}^k \frac{1}{j!} d^j f(x_0)(h)$$

e

$$Q(h) - R(h) = o(|h|^k) \text{ per } h \rightarrow 0 \implies Q - R = 0 \implies Q = R$$

per il lemma precedente. \square

Per ricostruire il comportamento nei punti critici (che verranno definiti più avanti) di una funzione siamo interessati alla positività delle forme quadratiche: infatti, il comportamento della funzione intorno a quei punti è determinato dal comportamento al secondo ordine della funzione (naturalmente per funzioni per cui valgono le ipotesi del teorema di Schwarz, altrimenti bisogna attaccare il problema con dei ragionamenti più sofisticati).

Consideriamo $q : \mathbb{R}^n \rightarrow \mathbb{R}$ forma quadratica, allora sappiamo che $q(x) = \langle Ax, x \rangle$ con $A \in \mathbb{R}^{n \times n}$ simmetrica. Diamo adesso la definizione di positività o negatività della forma quadratica:

Definizione 3.14 (forma quadratica definita positiva o negativa). Sia $q : \mathbb{R}^n \rightarrow \mathbb{R}$ una forma quadratica con matrice A in base canonica, allora diremo che q o A è definita positiva se

$$\forall x \in \mathbb{R}^n, q(x) > 0.$$

Invece, diremo che è definita negativa se

$$\forall x \in \mathbb{R}^n, q(x) < 0.$$

Esiste anche la nozione di semi-positività e semi-negatività (nel caso in cui la restrizione ad un certo sottospazio di \mathbb{R}^n annulli la forma quadratica ma nel complementare dello spazio vettoriale è sempre definita positivo oppure sempre definito negativo):

Definizione 3.15 (forma quadratica semi-definita positiva o semi-definita negativa). Sia $q : \mathbb{R}^n \rightarrow \mathbb{R}$ forma quadratica rappresentata dalla matrice A in base canonica, allora diremo che è semi-definita positiva se

$$\forall x \in \mathbb{R}^n, q(x) \geq 0.$$

Invece, diremo che è semi-definita negativa se

$$\forall x \in \mathbb{R}^n, q(x) \leq 0$$

Siamo giunti al seguente teorema:

Teorema 3.11 (criterio di Sylvester). La matrice $A = A^T \in \mathbb{R}^{n \times n}$ è definita positiva se e solo se $\det [A]_k > 0 \forall k \in \{1, \dots, n\}$ dove

$$[A]_k = \begin{pmatrix} a_{11} & \dots & a_{1k} \\ \vdots & \ddots & \vdots \\ a_{k1} & \dots & a_{kk} \end{pmatrix}$$

è la sottomatrice principale di ordine k

Oss.: Se $A = A^T$ è definita negativa, allora $B = -A$ è definita positiva e, per il teorema di Sylvester, avremo che

$$\det [B]_k = \det (-[A]_k) \stackrel{\text{togliendo il } -\text{ a tutte le colonne}}{=} (-1)^k \det [A]_k > 0 \forall k \in \{1, \dots, n\}$$

Dimostrazione.

\Rightarrow : osservando che $\langle Ax, x \rangle = (Ax)^T \mathbf{Id}x = x^T A^T x = x^T Ax$, allora possiamo considerare il prodotto scalare che in base canonica è rappresentato dalla matrice A , dunque la positività del prodotto scalare $\varphi > 0$ implica che $\varphi_{V_k} > 0 \forall k \implies A_i$ congruente a $Id_k \forall k \implies \det [A]_k$ ha lo stesso segno di $\det Id_k = 1$ (infatti, presa una matrice A e la matrice del cambio di base $A \rightarrow B$ allora $\det B = \det B^T AB \stackrel{\text{Binet}}{=} \det B^T \det B \det A = \det B^2 \det A$ siccome il determinante di una matrice coincide con quello della sua trasposta).

\Leftarrow : possiamo procedere per induzione sulle dimensioni della matrice A . Il caso $n = 1$ è banale e mostriamo adesso che $n - 1 \implies n$: se consideriamo una matrice di ordine n tale che $\det [A]_k > 0 \forall k \in \{1, \dots, n+1\}$ allora sappiamo che la restrizione a $V_{n-1} = \text{Span}(\{e_1, \dots, e_{n-1}\})$ della matrice A ha determinante positivo e, per ipotesi induttiva, è definita positiva. L'aggiunta del vettore e_n può provocare che 3 casi possibili per la segnatura del prodotto scalare

$$(k, 1, 0)$$

$$(k, 0, 1)$$

$$(k + 1, 0, 0)$$

da cui notiamo che il primo caso con ι_- uguale a 1 è impossibile, siccome sarebbe congruente ad una matrice che ha -1 sulla diagonale, il che implicherebbe che il determinante sia negativo, contraddizione con l'ipotesi $\det [A]_k > 0 \forall k \in \{1, \dots, n+1\}$. Il caso con ι_0 è impossibile siccome implicherebbe che è congruente ad una matrice con un determinante nullo, il che è una contraddizione con l'ipotesi $\det [A]_k > 0 \forall k \in \{1, \dots, n+1\}$. \square

Corollario 3.11.1. Una matrice simmetrica $A = A^T$ è definita positiva se e solo se tutti i suoi autovalori sono positivi. Equivalentemente, una matrice simmetrica è definita negativa se e solo se tutti i suoi autovalori sono negativi

Dimostrazione. Senza perdita di generalità supponiamo che A sia definita positiva: il determinante è invariante per cambio di base e coincide con il prodotto degli autovalori, pertanto siccome per ogni sotto-matrice di ordine k avremo che il determinante è positivo, allora segue che gli autovalori sono tutti positivi. \square

Propongo anche un'altra dimostrazione di questa fatto, presente sulle note degli appunti del professore

Dimostrazione alternativa. Per il teorema spettrale sappiamo che esiste una base di autovettori ortonormali rispetto al prodotto scalare canonico. In questa base la matrice della forma quadratica è diagonale e la indichiamo con Λ e la matrice del cambio di base dalla canonica a questa di autovettori è $C \in O(n)$. Per $x \neq 0$ abbiamo che

$$q(x) = \langle Ax, x \rangle = \langle C^T \Lambda Cx, x \rangle = \langle \Lambda Cx, Cx \rangle = \langle \Lambda y, y \rangle$$

dunque se $\lambda_i > 0 \forall i \implies \langle \Lambda y, y \rangle > 0$. \square

3.4 Massimi e minimi in più variabili

Lo studio dei massimi e dei minimi è naturalmente cruciale in ogni scienza, ma non solo: molto spesso molti aspetti del quotidiano sono facilmente modellizzabili da funzioni di cui spesso è necessario trovare i valori di massimo o minimo. Così come nel caso a singola variabile, vogliamo individuare un procedimento analitico per la ricerca di essi e, pertanto, ci dilungheremo nella presentazione di qualche risultato teorico molto importante che ci aiuterà a fare questo:

Teorema 3.12 (teorema di Fermat). Siano $\Omega \subseteq \mathbb{R}^n$ aperto e $f : \Omega \rightarrow \mathbb{R}$. Supponiamo che $x_0 \in \Omega$ sia un punto di minimo (o massimo) locale e che f sia differenziabile in x_0 . Allora

$$\nabla f(x_0) = 0 \quad \text{o equivalentemente} \quad df(x_0) = 0$$

Dimostrazione. Possiamo supporre senza perdita di generalità che x_0 sia un punto di minimo locale e considerare la funzione $g(t)$ tale che

$$t \xrightarrow{g} f(x_0 + te_j) \text{ per } |t| < \sigma$$

ovvero $x_0 + te_j \in B(x_0, \sigma) \subseteq \Omega$, osservando che ha un minimo locale in $t = 0$

$$\frac{d}{dt} (f(x_0 + te_j)) \Big|_{t=0} = \partial_{x_j} f(x_0) = 0$$

dove l'ultima eguaglianza segue dal teorema di Fermat per funzioni di variabile singola. Possiamo ripetere lo stesso procedimento $\forall j \in \{1, \dots, n\}$ e dunque si ottiene la tesi. \square

Oss. In realtà l'ipotesi che la funzione sia differenziabile non è indispensabile, siccome è più che sufficiente richiedere l'esistenza delle derivate parziali.

Definizione 3.16 (punto critico). Diremo che $x_0 \in \Omega$ è un punto critico di $f : \Omega \rightarrow \mathbb{R}$ differenziabile in x_0 se $\nabla f(x_0) = 0$. Similmente diremo che $f(x_0)$ è un valore critico di f .

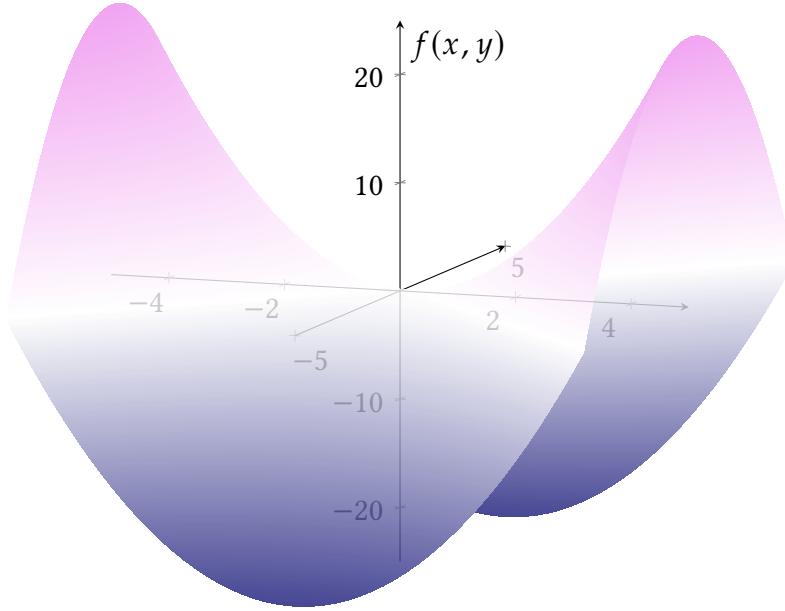


Figura 3.2: Grafico della funzione $f(x, y) = x^2 - y^2$: come è possibile vedere il punto $(0, 0)$ non è né un punto di minimo locale né un punto di massimo locale: si tratta di un *punto di sella*

Ora, si possono trovare bizzeffe di esempi in cui in un particolare punto si annulla il gradiente di una funzione ma non coincide né con un punto di massimo locale né con un punto di minimo locale: un esempio è la funzione $f(x, y) = x^2 - y^2$ che nel punto $(0, 0)$ ha gradiente nullo.

Enunciamo il seguente criterio

Teorema 3.13 (condizione sufficiente per il massimo/minimo locale). Sia $f : \Omega \rightarrow \mathbb{R}$, $f \in C^2(\Omega)$ con $\Omega \subseteq \mathbb{R}^n$ aperto e sia $x_0 \in \Omega$ un punto critico di $f(\nabla f(x_0) = 0)$. Se $Hf(x_0)$ è definita positiva, allora $x_0 \in \Omega$ è un punto di minimo locale per f ; altrimenti se $Hf(x_0)$ è definita negativa, allora x_0 è un punto di massimo locale per f .

Dimostrazione. Senza perdita di generalità, facciamo la dimostrazione supponendo che $Hf(x_0)$ sia definita positiva (il caso della matrice hessiana definita negativa è simile). Per il teorema della formula di Taylor 3.9 abbiamo che

$$f(x) = f(x_0) + \frac{1}{2} \langle Hf(x_0)(x - x_0), x - x_0 \rangle + o(|x - x_0|^2) \text{ per } x \rightarrow x_0$$

Consideriamo adesso $\lambda = \min_{v \in \mathbb{S}^{n-1}} \langle Hf(x_0)v, v \rangle > 0$ per l'ipotesi di positività di $Hf(x_0)$. Allora sappiamo che $\exists \delta > 0$ tale che $B(x_0, \delta) \subseteq \Omega$ e $|o(|x - x_0|^2)| < \frac{\lambda}{2}|x - x_0|^2$ per $0 < |x - x_0| < \delta$, dunque

$$o(|x - x_0|^2) \geq -|o(|x - x_0|^2)| \geq -\frac{\lambda}{2}|x - x_0|^2$$

dunque $\forall x \in B(x_0, \delta) \setminus \{x_0\}$

$$\begin{aligned} f(x) &= f(x_0) + \frac{1}{2} \langle Hf(x_0)(x - x_0), (x - x_0) \rangle + o(|x - x_0|^2) = \\ &= f(x_0) + \frac{1}{2} \langle Hf(x_0) \frac{x - x_0}{|x - x_0|}, \frac{x - x_0}{|x - x_0|} \rangle |x - x_0|^2 + o(|x - x_0|^2) \geq \\ &\geq f(x_0) + \frac{\lambda}{2}|x - x_0|^2 - \frac{\lambda}{2}|x - x_0|^2 = f(x_0) \implies f(x) \geq f(x_0) \end{aligned}$$

questo conclude la dimostrazione. □

Adesso vogliamo trovare un criterio con cui stabilire quando i punti critici sono di sella. Anche in questo caso tramite il comportamento al secondo *ordine* della funzione (ovvero dalla matrice hessiana) possiamo stabilirlo tramite il seguente teorema

Teorema 3.14 (dei punti di sella). Sia $f \in C^2(\Omega)$, $x_0 \in \Omega$, $\nabla f(x_0) = 0$, $\det Hf(x_0) \neq 0$ e $Hf(x_0)$ non è definita positiva e non è definita negativa

$$\implies x_0 \text{ non è né di massimo locale né di minimo locale}$$

Dimostrazione.

$$Hf(x_0) = A = A^T = C^T \Lambda C$$

dove $\Lambda = \begin{pmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & \dots & \dots & \lambda_n \end{pmatrix}$ e $\lambda_1, \dots, \lambda_n$ autovalori reali. Osserviamo che

$$\langle Ax, x \rangle = \langle C^T \Lambda Cx, Cx \rangle = \langle \Lambda Cx, Cx \rangle = \langle \Lambda Cx, Cx \rangle \stackrel{y=Cx}{=} \langle \Lambda y, y \rangle = \sum_{j=1}^n \lambda_j y_j^2$$

Osserviamo che, grazie al teorema spettrale, $\forall j \in \{1, \dots, n\} \exists v_j \in \mathbb{S}^{n-1} : e_j = Cv_j$ dove v_j è il j -esimo autovettore

$$Av_j = C^{-1} \Lambda C v_j = C^{-1} \Lambda e_j = C^{-1} (\lambda_j e_j) = \lambda (C^{-1} e_j) = \lambda_j v_j$$

dunque

$$\langle Hf(x_0)v_j, v_j \rangle = \lambda_j$$

Dalla formula di Taylor concludiamo che

$$f(x_0 + tv_j) = f(x_0) + \frac{1}{2} \langle Hf(x_0)tv_j, tv_j \rangle + o(|tv_j|^2) = f(x_0) + t^2 \left(\frac{\lambda_j}{2} + o(1) \right)$$

poiché $Hf(x_0)$ non è né definita positiva né definita negativa avremo che $\lambda_j > 0$ e $\lambda_j < 0$ dunque la precedente uguaglianza, unita al ragionamento compiuto nella dimostrazione del teorema precedente portano alla tesi. \square

Oss.: In questa dimostrazione comprendiamo come mai si chiamino *punti di sella*: nel caso $n = 2$ (ovvero in due dimensioni) con la seguente matrice hessiana

$$Hf(x_0) = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

vuol dire che la funzione, lungo una direzione, assume un massimo locale in quel punto e nell'altra un minimo. La funzione $f(x, y) = x^2 - y^2$ (Figura 3.4) nel punto $(0, 0)$ lungo la direzione x ha un punto minimo (basti osservare la restrizione della funzione a $y = 0$, per cui otteniamo $f(x, 0) = x^2$) e un punto di massimo (basti osservare la restrizione della funzione a $x = 0$, per cui otteniamo $f(0, y) = -y^2$). Il motivo della terminologia *punto di sella* deriva proprio del fatto che in due dimensione assume proprio la forma di una sella.

Il problema è capire come si comporta nei punti per cui è presente un autovalore nullo: purtroppo, quando l'hessiana è semi-definita abbiamo bisogno di procedere tramite dei ragionamenti più sofisticati che dipendono da caso a caso.

CAPITOLO 4

Curve e superfici regolari

In questo capitolo ci si sofferma nello studio delle curve e della superfici regolari, con particolare riguardo alla caratterizzazione delle curve tramite pre-immagini di funzioni implicite e lo studio dei massimi e i minimi di una funzione vincolata. Questo è particolarmente utile nello studio di alcuni tipi di problemi che possono capitare nella vita di tutti i giorni, in cui si è interessati a dei massimi o dei minimi vincolati ad una curva oppure ad una superficie.

4.1 Teorema della funzione inversa

Uno dei problemi molto ricorrenti nello studio della matematica e nelle scienze è il risolvere un'equazione del tipo

$$f(x) = 0$$

Nel corso di Geometria si è visto che nel caso di applicazioni lineari affini e invertibili $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$, ci si chiedeva per $x \in \mathbb{R}^n$ avevamo che

$$Lx = \underline{0} \implies Ax + B = \underline{0} \implies Ax = -B$$

e riducevamo il problema ad un sistema lineare non omogeneo che aveva soluzione unica se $\det A \neq 0$ e $B \in \text{Span}\{A^1, A^2, \dots, A^n\}$.

Nel caso di funzioni differenziabili è possibile dire qualcosa?

Dal caso affine in generale possiamo innanzitutto intuire che per f differenziabile bisogna assumere che $\det Df(x_0) \neq 0$, però cos'altro possiamo dire?

Innanzitutto, fissato $\Omega \subseteq \mathbb{R}^n$ e definiamo lo spazio di tutte le funzioni vettoriali di classe C^k e a valori in $A \subseteq \mathbb{R}^m$:

Definizione 4.1 (spazio delle funzioni vettoriali di classe C^k). Sia $\Omega \subseteq \mathbb{R}^n$, $A \subseteq \mathbb{R}^m$ e $k \in \mathbb{N}$. Allora definiamo lo spazio delle funzioni vettoriali di classe C^k come

$$C^k(\Omega, A) = \{f : \Omega \rightarrow A \text{ con } f = (f_1, \dots, f_m) \mid \forall j \in \{1, \dots, m\}, f_j \in C^k(\Omega)\}$$

Adesso siamo pronti ad enunciare il seguente teorema, di cui la dimostrazione è riportata nelle appendici, che risponde alla domanda che ci siamo posti all'inizio di questo paragrafo

Teorema 4.1 (teorema della funzione inversa). Sia $\Omega \subseteq \mathbb{R}^n$ aperto, $f \in C^k(\Omega, \mathbb{R}^n)$, $x_0 \in \Omega$ e $Df(x_0) \in \mathbb{R}^{n \times n}$ invertibile.

Allora $\exists \rho > 0 : \exists U \subseteq \mathbb{R}^n$ aperto tali che $f|_{B(x_0, \rho)} \in C^k(B(x_0, \rho), U)$ è invertibile e, posto $g = (f|_{B(x_0, \rho)})^{-1} : U \rightarrow B(x_0, \rho)$, abbiamo che $g \in C^k(U, B(x_0, \rho))$ con

$$Dg(y) = (Df(g(y)))^{-1} \forall y \in U$$

4.2 Teorema delle funzioni implicite o del Dini

Adesso il nostro argomento di interesse saranno le curve e le superfici: vogliamo sviluppare degli strumenti con cui trattare questo tipo di oggetti. In generale, possiamo procedere verso il nostro obiettivo con due *approcci* possibili:

- un approccio "parametrico", in cui parametrizziamo la superficie che vogliamo rappresentare in base a delle variabili che facciamo variare all'interno di un intervallo;
- un approccio più vicino a quello dell'Analisi, ovvero quello di pensarle come curve di livello di una funzione a più variabili

Un esempio di questo è sicuramente la circonferenza: col primo approccio una circonferenza di raggio uguale a 2 e col centro coincidente con l'origine è possibile parametrizzarla tramite le funzioni $2 \cos t$ e $2 \sin t$ facendo variare per $t \in [0; 2\pi]$

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 2 \cos t \\ 2 \sin t \end{pmatrix}$$

oppure visualizzandola come il luogo dei punti in cui la funzione $g(x, y) = x^2 + y^2 - 4$ si annulla e caratterizzare la nostra circonferenza Γ come

$$\Gamma : g^{-1}(0) \mapsto 0$$

quindi definire la circonferenza come la pre-immagine del valore 0 assunto dalla funzione g .

Il teorema della funzione inversa rende possibile anche la caratterizzazione delle funzioni implicite ad un livello *locale*: infatti ci dice che, in un intorno molto piccolo del punto $f(x_0)$, è possibile visualizzare la funzione implicita come un oggetto che vive in una dimensione più piccola. Andiamo ad enunciare questo teorema

Teorema 4.2 (teorema della funzione implicita rispetto a x). Sia $\Omega \subseteq \mathbb{R}^2$ aperto, $\underline{x}_0 = (x_{0_1}, x_{0_2}) \in \Omega$ e sia $f \in C^k(\Omega)$, $\partial_y f(x_0) \neq 0$

$$\implies \exists \delta, \sigma > 0 : f^{-1}(f(x_0)) \cap (x_{0_1} - \delta, x_{0_1} + \delta) \times (x_{0_2} - \sigma, x_{0_2} + \sigma) = \{(t, \varphi(t)) : |t - x_{0_1}| < \delta\}$$

ovvero esiste una funzione $\varphi : (x_{0_1} - \delta, x_{0_1} + \delta) \rightarrow (x_{0_2} - \sigma, x_{0_2} + \sigma)$ tale che $\varphi(x_{0_1}) = x_{0_2}$ e dunque

$$f(t, \varphi(t)) = f(x_0)$$

Inoltre

$$\varphi'(t) = -\frac{\partial_x f(t, \varphi(t))}{\partial_y f(t, \varphi(t))} \quad \forall t \in (x_{0_1} - \delta, x_{0_1} + \delta)$$

Osserviamo due cose molto importanti:

- ① viene richiesta un'ipotesi di **regolarità** sulla funzione f (questo risulta intuitivo, da un lato, siccome stiamo cercando di approssimare, in un intorno *piccolo* una funzione che vive, per esempio, in \mathbb{R}^2 con un x -grafico quindi deve essere regolare, dall'altra parte dobbiamo applicare il teorema della funzione inversa per esplicitare una variabile in funzione dell'altra localmente);

- ② è necessario conoscere almeno un valore della funzione f : nel caso del teorema è il punto $x_0 \in \Omega$, il quale svolge la funzione di una sorta di "ancora" per definire il comportamento locale della funzione

Dimostrazione. Presa la funzione $f(x, y)$ possiamo considerare la funzione $\tilde{f} : \Omega \rightarrow \mathbb{R}^2$ tale che

$$\tilde{f}(x, y) = (x, f(x, y))$$

ed osservare che la condizione di $\partial_y f(x_0) \neq 0$ ci consente di invertire localmente la funzione \tilde{f} siccome

$$D\tilde{f}(x_0, y_0) = \begin{pmatrix} 1 & 0 \\ \partial_x f(x_0, y_0) & \partial_y f(x_0, y_0) \end{pmatrix}$$

e, dunque, il determinante di questa funzione è pari a $\partial_y f(x_0, y_0)$ che, per ipotesi, è diverso da zero, il che ci consente di applicare il teorema della funzione inversa a \tilde{f} . Conseguentemente, esisterà un intorno U di $(x_0, f(x_0, y_0))$ tale che $\exists \tilde{g} : U \rightarrow \Omega$ tale che, ricordando che se $f : X \rightarrow Y$ e $g = f^{-1} : Y \rightarrow X$ allora $f(g(y)) = y \forall y \in Y$, allora $f(\tilde{g}(x, y)) = (x, y) \forall (x, y) \in U$. Ma, posto $\tilde{g}(x, y) = (g_1(x, y), g_2(x, y))$, possiamo dedurre che

$$\tilde{f}(\tilde{g}(x, y)) = \tilde{f}((g_1(x, y), g_2(x, y))) = (g_1(x, y), f(g_1(x, y), g_2(x, y)))$$

permettendoci di concludere che

$$\begin{aligned} \forall (x, y) \in U, (x, y) = \tilde{f}(\tilde{g}(x, y)) = (g_1(x, y), f(g_1(x, y), g_2(x, y))) &\implies \\ \implies \begin{cases} g_1(x, y) = x \\ f(g_1(x, y), g_2(x, y)) = f(x, g_2(x, y)) = y \end{cases} \end{aligned}$$

permettendoci, in conclusione, di giungere alla tesi definendo $\varphi(x) = g_2(x, f(x_0, y_0))$

$$f(x, \varphi(x)) = f(x, g_2(x, f(x_0, y_0))) = f(x_0, y_0)$$

Per dimostrare la formula della derivata basta derivare quest'ultima relazione rispetto ad x :

$$\begin{aligned} \partial_x f(x, \varphi(x)) = \partial_x f(x_0, y_0) &\implies \partial_x f(x, \varphi(x)) + \partial_y f(x, \varphi(x)) \partial_x \varphi(x) = 0 \implies \\ \implies \partial_x \varphi(x) &= -\frac{\partial_x f(x, g(x))}{\partial_y f(x, g(x))} \end{aligned}$$

□

Oss.: Possiamo riformulare il teorema della funzione implicita nei punti in cui $\partial_y f(x_0) = 0$ ma $\partial_x f(x_0) \neq 0$: infatti, la versione più generale del teorema richiede che $\nabla f \neq 0$. Ciò è possibile rifacendo la dimostrazione ma ridefinendo ad hoc la funzione \tilde{f} in maniera tale che il determinante della sua matrice jacobiana sia pari a $\partial_x f(x_0)$. Enunciamo qua sotto tutte le "versioni" alternative del teorema nel caso di $n = 2$ e $n = 3$

Teorema 4.3 (teorema della funzione implicita rispetto a y). Sia $\Omega \subseteq \mathbb{R}^2$ aperto, $\underline{x}_0 = (x_{0_1}, x_{0_2}) \in \Omega$ e $f \in C^1(\Omega)$ con $\partial_x f(x_0) \neq 0$. Allora

$$\exists \delta, \sigma > 0 : f^{-1}(f(x_0)) \cap (x_{0_1} - \sigma, x_{0_1} + \sigma) \times (x_{0_2} - \delta, x_{0_2} + \delta) = \{(\varphi(t), t) : |t - x_{0_2}| < \delta\}$$

ovvero $\exists \varphi : (x_{0_2} - \delta, x_{0_2} + \delta) \rightarrow (x_{0_1} - \sigma, x_{0_1} + \sigma)$ tale che $\varphi(x_{0_2}) = x_{0_1}$ e dunque

$$f(\varphi(t), t) = f(x_0)$$

Inoltre

$$\varphi'(t) = -\frac{\partial_y f(\varphi(t), t)}{\partial_x f(\varphi(t), t)} \quad \forall t \in (x_{0_2} - \delta, x_{0_2} + \delta)$$

Dimostrazione. Si riprende la dimostrazione fatta prima usando $\tilde{f}(x, y) = (f(x, y), y)$. \square

Con qualche piccolo accorgimento, la dimostrazione fatta qua sopra è valida anche per $\Omega \subseteq \mathbb{R}^n$

Definizione 4.2 (curva regolare). Diremo che $C \subseteq \mathbb{R}^2$ è una curva regolare se $\exists \Omega \subseteq \mathbb{R}^2$ aperto tale che $C = f^{-1}(f(x_0)) \subseteq \Omega$ dove $f \in C^k(\Omega)$ e $\nabla f(x) \neq (0, 0) \forall x \in C$.

Evidenziando la regolarità di C , diremo che C è una curva di classe C^k

Teorema 4.4 (funzione implicita rispetto a y e z). Sia $\Omega \subseteq \mathbb{R}^3$ aperto, $\underline{x}_0 = (x_{01}, x_{02}, x_{03}) \in \Omega$ e sia $f \in C^1(\Omega)$ tale che $\partial_{x_1} f(x_0) \neq 0$. Allora

$$\begin{aligned} \exists \delta, \sigma > 0 : f^{-1}(f(x_0)) \cap ((x_{01} - \sigma, x_{01} + \sigma) \times B((x_{02}, x_{03}), \delta)) = \\ &= \{(\varphi(x_2, x_3), x_2, x_3) : (x_2, x_3) \in B((x_{02}, x_{03}), \delta)\} \end{aligned}$$

ed esiste $\varphi : B((x_{02}, x_{03}), \delta) \rightarrow (x_{01} - \sigma, x_{01} + \sigma)$ con $\varphi(x_{02}, x_{03}) = x_{01}$ tale che

$$f(\varphi(x_2, x_3), x_2, x_3) = f(x_{01}, x_{02}, x_{03})$$

Inoltre

$$\partial_{x_j} \varphi(x_2, x_3) = -\frac{\partial_{x_j} f(\varphi(x_2, x_3), x_2, x_3)}{\partial_{x_1} f(\varphi(x_2, x_3), x_2, x_3)} \quad \forall j \in \{2, 3\} \quad \forall (x_2, x_3) \in B(x_0, \sigma)$$

Dimostrazione. Si modifica la dimostrazione fatta prima scomponendo la funzione $f(x_1, x_2, x_3) = f(x, y)$ con $x \in \mathbb{R}$, $y \in \mathbb{R}^2$, considerando la funzione $\tilde{f} = (f(x, y), x, y)$ imponendo che $D_y f \neq 0$. \square

Teorema 4.5 (teorema della funzione implicita rispetto a x e y). Sia $\Omega \subseteq \mathbb{R}^3$ aperto, $x_0 = (x_{01}, x_{02}, x_{03}) \in \Omega$ e sia $f \in C^1(\Omega)$ tale che $\partial_{x_3} f(x_0) \neq 0$. Allora

$$\begin{aligned} \exists \delta, \sigma > 0 : f^{-1}(f(x_0)) \cap (B((x_{01}, x_{02}), \delta) \times (x_{03} - \sigma, x_{03} + \sigma)) = \\ &= \{(x_1, x_2, \varphi(x_1, x_2)) : (x_1, x_2) \in B((x_{01}, x_{02}), \delta)\} \end{aligned}$$

ed esiste $\varphi : B((x_{01}, x_{02}), \delta) \rightarrow (x_{03} - \sigma, x_{03} + \sigma)$ con $\varphi \in C^1(B((x_{01}, x_{02}), \delta))$ e $\varphi(x_{01}, x_{02}) = x_{03}$ tale che

$$f(x_1, x_2, \varphi(x_1, x_2)) = f(x_{01}, x_{02}, x_{03})$$

Inoltre

$$\partial_{x_j} \varphi(x_1, x_2) = -\frac{\partial_{x_j} f(x_1, x_2, \varphi(x_1, x_2))}{\partial_{x_3} f(x_1, x_2, \varphi(x_1, x_2))} \quad \forall j \in \{1, 2\} \quad \forall (x_1, x_2) \in B((x_{01}, x_{02}, x_{03}), \sigma)$$

Definizione 4.3 (superficie regolare). Siano $S \subseteq \mathbb{R}^3$ è una superficie regolare se $\exists \Omega \subseteq \mathbb{R}^3$ aperto tale che $S = f^{-1}(f(x_0)) \subseteq \Omega$ dove $x_0 \in \Omega$ e $f \in C^k(\Omega)$ e $\nabla f \neq 0 \forall u \in S$. Volendo evidenziare la classe di regolarità di S diremo che S è una superficie regolare di classe C^k

Oss.: La definizione che abbiamo dato di curva/superficie regolare ci consente di affermare che le funzioni implicite, almeno localmente, sono tutte delle curve regolari: infatti, possiamo osservare che se

$$S = \{(\varphi(x_2, x_3), x_2, x_3) : (x_2, x_3) \in U\}$$

allora possiamo considerare la funzione

$$f(x, y, z) = \varphi(y, z) - x$$

tale che $f : \mathbb{R} \times U \rightarrow \mathbb{R}$ da cui osserviamo che $\nabla g|_S \neq 0$ siccome $\partial_x g = 1$ e $S = f^{-1}(0) = f^{-1}(\{(x, y, z) : x = \varphi(y, z)\})$

A questo punto diamo la nozione di spazio tangente, che va a generalizzare il concetto di "tangenza" a più dimensioni

Definizione 4.4 (spazio tangente). Sia $\Omega \subseteq \mathbb{R}^n$ aperto e sia $x_0 \in \Omega$ e sia $x_0 \in S \subseteq \mathbb{R}^n$ dove S è una curva regolare o una superficie regolare. Sia $f \in C^k(\Omega)$ tale che $S = f^{-1}(f(x_0))$ e $\forall x \in S, \nabla f \neq 0$. Definiamo

$$T_{x_0} S = \{v \in \mathbb{R}^n : df(x_0)(v) = 0\} = \ker df(x_0)$$

lo spazio tangente ad S in $x_0 \in S$. Diremo che f è una funzione definente per S .

Proposizione 4.1 (dimensione dello spazio tangente). Sia $f \in C^k(\Omega)$ con $\Omega \subseteq \mathbb{R}^n$ la funzione definente di una curva regolare o una superficie regolare S . Allora

$$\dim \ker df(x_0) = n - 1$$

Dimostrazione. Sappiamo che $df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}$, dunque per il teorema del rango, avremo che

$$\dim \ker df(x_0) = \dim \mathbb{R}^n - \dim \text{Im}(df(x_0)) = n - 1$$

siccome la dimensione del differenziale valutato in x_0 è sicuramente 1 per ipotesi essendo un'applicazione lineare fra \mathbb{R}^n e \mathbb{R} . \square

Osserviamo che se $\gamma \in C^1(I_{0,\varepsilon}, \Sigma)$ allora $f(\gamma(t)) = f(x_0)$ siccome definiamo le curve come $\Sigma = f^{-1}(f(x_0))$, assumendo inoltre che $\gamma(0) = x_0 \implies 0 = \frac{d}{dt}(f \circ \gamma)(t)|_{t=0} = \langle \nabla f(\gamma(0)), \dot{\gamma}(0) \rangle = df(\gamma(0))(\dot{\gamma}(0)) \implies \dot{\gamma}(0) \in T_{x_0}\Sigma$ come ci si aspetta dall'intuizione geometrica (e anche fisica, infatti, dal corso di Fisica 1, sappiamo che la velocità di un punto materiale è sempre tangente alla traiettoria).

Definizione 4.5 (spazio delle velocità). Sia $\Sigma \subseteq \mathbb{R}^n$ una curva o superficie regolare e sia $x_0 \in \Sigma$. Definiamo

$$V_{x_0} = \{v \in \mathbb{R}^n : \gamma \in C^1((- \varepsilon, \varepsilon), \Sigma), \dot{\gamma}(0) = v, \gamma(0) = x_0\}$$

Teorema 4.6 (della caratterizzazione del tangente). Sia $\Sigma \subseteq \mathbb{R}^n$ una curva regolare o una superficie regolare di classe C^k . Allora $\forall x_0 \in \Sigma$ abbiamo che

$$T_{x_0}\Sigma = V_{x_0}\Sigma$$

Dimostrazione. Mostriamo che $V_{x_0}\Sigma \subseteq T_{x_0}\Sigma$: come avevamo già visto in precedenza, se $\gamma \in C^1((- \varepsilon, \varepsilon), \Sigma)$ tale che $\gamma(0) = x_0$ e $g^{-1}(g(x_0)) = \Sigma \implies g(\gamma(t)) = \gamma(0) \implies \frac{d}{dt}g(\gamma(t)) = \frac{d}{dt}\gamma(0) = \langle \nabla g(\gamma(0)), \dot{\gamma}(0) \rangle = dg(\gamma(0))(\dot{\gamma}(0)) \implies (\forall \dot{\gamma}(0), \dot{\gamma}(0) \in V_{x_0}\Sigma \implies \dot{\gamma}(0) \in T_{x_0}\Sigma) \iff V_{x_0}\Sigma \subseteq T_{x_0}\Sigma$.

Mostriamo adesso che $T_{x_0}\Sigma \subseteq V_{x_0}\Sigma$: facciamo la dimostrazione per $n = 2$ e per $n = 3$. Nel caso $n = 2$ sappiamo, in virtù della funzione implicita, che $\exists \varphi : I \rightarrow J$ tale che

$$\gamma(t) = (t, \varphi(t)) \quad \text{o} \quad \gamma(t) = (\varphi'(t), t)$$

tale che $\forall t, \gamma(t) \in \Sigma, \gamma(t_0) = x_0$ e $\dot{\gamma}(t_0) \neq 0$. Definiamo, a questo punto, la funzione $\tilde{\gamma}(s) = \gamma(t_0 + s)$ da cui si deduce, banalmente, che $\tilde{\gamma}(0) = x_0$ e, fissato $\alpha \in \mathbb{R}$, possiamo definire anche

$$\tilde{\gamma}_\alpha \in C^1(I_{0,\varepsilon}, \Sigma)$$

tale che

$$\tilde{\gamma}_\alpha'(0) = \alpha \dot{\gamma}'(0) \neq 0$$

per ipotesi. Ma ciò implica che

$$\text{Span}(\dot{\gamma}(0)) \subset V_{x_0}\Sigma \subset T_{x_0}\Sigma$$

ma allora, siccome $\dim \text{Span}(\dot{\gamma}(0)) = 1$ e $\dim T_{x_0}\Sigma = 1$, che

$$\text{Span}(\dot{\gamma}(0)) = T_{x_0}\Sigma = V_{x_0}\Sigma$$

Nel caso $n = 3$ si procede similmente, sfruttando il fatto che il teorema della funzione implicita garantisce sempre la possibilità di esprimere localmente la superficie regolare con g funzione definendo come un (x, y) -grafico (oppure come un grafico rispetto a (y, z) p (x, z)), la dimostrazione è equivalente quindi supporremo senza perdita di generalità di scrivere localmente la nostra funzione come un (x, y) -grafico tramite la funzione φ :

$$\Phi(x, y) = (x, y, \varphi(x, y)).$$

Osserviamo a questo punto che, in maniera simile a quanto fatto prima, possiamo definire

$$\tilde{\Phi}_{\alpha, \beta}(t) = \Phi(x_{0_1} + \alpha t, x_{0_2} + \beta t)$$

tale che $\forall t, \tilde{\Phi}(t) \in \Sigma$ e osserviamo che

$$\tilde{\Phi}'(0) = \alpha \partial_x \Phi(x_{0_1}, x_{0_2}) + \beta \partial_y \Phi(x_{0_1}, x_{0_2}) \in V_{x_0}\Sigma$$

Ma abbiamo che

$$\text{Span}(\partial_x \Phi(x_{0_1}, x_{0_2}), \partial_y \Phi(x_{0_1}, x_{0_2})) \subset V_{x_0}\Sigma \subset T_{x_0}\Sigma$$

da cui deduciamo, siccome $\dim T_{x_0}\Sigma = 2$ e $\dim \text{Span}(\partial_x \Phi(x_{0_1}, x_{0_2}), \partial_y \Phi(x_{0_1}, x_{0_2})) = 2$

$$\text{Span}(\partial_x \Phi(x_{0_1}, x_{0_2}), \partial_y \Phi(x_{0_1}, x_{0_2})) = V_{x_0}\Sigma = T_{x_0}\Sigma$$

ovvero la tesi. □

Oss.: Con il solito procedimento è possibile dimostrare il caso n -dimensionale.

Oss.: Dal teorema precedente è possibile vedere che la definizione di spazio tangente è ben posta sebbene non dipende dalla funzione definente che abbiamo scelto. Considerando il caso banale della circonferenza centrata nell'origine di raggio 2 io potrei scegliere come funzione definente sia la funzione

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

e definire la circonferenza $\Gamma : f^{-1}(0)$ oppure considerare la funzione

$$g(x, y) = x^2 + y^2$$

e definire la circonferenza come $\Gamma : g^{-1}(4)$: il teorema precedente mi garantisce che in entrambi i casi lo spazio tangente è indipendente dalla funzione scelta

4.3 Punti critici su curve e superfici regolari

Adesso che abbiamo caratterizzato le curve e le superfici, siamo interessati a studiare i massimi e i minimi della funzione che la funzione assume se "vincolata" su curve o superfici.

Definizione 4.6 (punti critici su curve/superfici). Sia $S \subseteq \mathbb{R}^n$ con $n \in \{2, 3\}$ e sia $x_0 \in S$. Supponiamo che $\exists \Omega \subseteq \mathbb{R}^n$ aperto tale che $x_0 \in \Omega$ e $\Sigma = \Omega \cap S$ è una curva regolare in \mathbb{R}^2 o una superficie regolare in \mathbb{R}^3 (di classe C^1). Allora diremo che $x_0 \in S$ è un punto critico di $f \in C^1(\Omega)$ su Σ se

$$\forall (\gamma \in C^1((-\varepsilon, \varepsilon), \Sigma), \gamma(0) = x_0), \frac{d}{dt}(f \circ \gamma)|_{t=0} = \langle \nabla f(x_0), \dot{\gamma}(0) \rangle = 0$$

Teorema 4.7 (teorema dei punti critici). Sia $\Sigma \subseteq \mathbb{R}^n$ una curva regolare per $n = 2$ o per $n = 3$. Sia $f \in C^1(\Omega)$, $x_0 \in \Omega \subseteq \mathbb{R}^n$ aperto. Allora

$$x_0 \in \Sigma \text{ è un punto critico di } f \text{ su } \Sigma \iff \nabla f(x_0) \perp T_{x_0} \Sigma$$

Dimostrazione. Per il teorema 4.6, abbiamo che se x_0 è punto critico allora, per definizione di punto critico, $\forall (\gamma \in C^1((-\varepsilon, \varepsilon), \Sigma), \gamma(0) = x_0), \langle \nabla f(x_0), \dot{\gamma}(0) \rangle = 0 \iff \nabla f(x_0) \perp T_{x_0} \Sigma$. \square

Teorema 4.8 (di Fermat su Σ). Sia Σ una curva o una superficie regolare con $x_0 \in \Sigma$ e sia $f \in C^1(\Omega)$ e sia $x_0 \in \Omega$ tale che $f|_{\Omega \cap S}$ ha un minimo o un massimo su $\Omega \cap S$. Allora x_0 è un punto critico di f su Σ

Dimostrazione. Sia $\gamma : (-\varepsilon, \varepsilon) \rightarrow \Sigma$ derivabile tale che $\gamma(0) = x_0$. Poiché $f|_{\Omega \cap \Sigma}$ ha un minimo (o un massimo) in $x_0 \implies \forall \gamma \in C^1((-\varepsilon, \varepsilon), \Sigma), (f \circ \gamma)(t)$ ha un massimo o un minimo in $t = 0$. Quindi $\frac{d}{dt}(f \circ \gamma)(t)|_{t=0} = 0$ da cui la tesi \square

Oss.: Tutto questo che abbiamo fatto in cosa si traduce? In pratica, se abbiamo una funzione f che andiamo a vincolare su una curva/superficie regolare Σ allora ogni punto critico *libero* $x_0 \in \Omega$ tale che $x_0 \in S$ è un punto critico di f su Σ , mentre il viceversa non è così. Per esempio potremo prendere la funzione $f(x, y) = x^2 + y$ e la retta $\Sigma = \{(x, y) \in \mathbb{R}^2 : y = 0\} = \mathbb{R} \times \{0\}$ da cui scopriremo che il minimo della funzione f su Σ è il punto $(0, 0)$ in cui, tuttavia $\nabla f((0, 0)) = (0, 1)$.

Teorema 4.9 (dei moltiplicatori di Lagrange). Sia $f \in C^1(\Omega)$, $x_0 \in \Omega \cap S$, $\Omega \subseteq \mathbb{R}^n$ aperto, $n = \{2, 3\}$, $S \subseteq \mathbb{R}^n$, $g \in C^1(\Omega)$ tale che $\Omega \cap S = g^{-1}(g(x_0)) = \Sigma \subseteq \mathbb{R}^n$, $\nabla g(x) \neq 0 \forall x \in \Sigma$ (Σ è una curva/superficie regolare).

Allora $x_0 \in \Sigma$ è un punto critico di f su Σ se e solo se $\exists \lambda \in \mathbb{R}$ tale che $\nabla f(x_0) = \lambda \nabla g(x_0)$

Dimostrazione.

\Rightarrow : sia $x_0 \in \Sigma$ un punto critico di f su Σ , allora per il teorema 4.7 avremo che $\nabla f(x_0) \perp T_{x_0} \Sigma \implies \nabla f(x_0) \in (T_{x_0} \Sigma)^\perp$, tuttavia, siccome $\Sigma = g^{-1}(g(x_0))$, abbiamo che $(\nabla g(x_0))^\perp = T_{x_0} \Sigma$ da cui concludiamo che $\nabla f(x_0), \nabla g(x_0) \in (T_{x_0} \Sigma)^\perp$ e, siccome $\dim(T_{x_0} \Sigma)^\perp = 1$ visto che $g : \mathbb{R}^n \rightarrow \mathbb{R} \implies df(x_0) : \mathbb{R}^n \rightarrow \mathbb{R}$, concludiamo che

$$\nabla f(x_0) = \lambda \nabla g(x_0)$$

\Leftarrow : supponiamo che $\exists \alpha \in \mathbb{R}$ tale che

$$\nabla f(x_0) = \alpha \nabla g(x_0)$$

allora, procedendo in maniera simile a quanto detto prima, sappiamo che $\nabla g(x_0) \in (T_{x_0} \Sigma)^\perp$, ma allora, per proprietà degli spazi vettoriali, abbiamo che $\alpha \nabla g(x_0) = \nabla f(x_0) \in (T_{x_0} \Sigma)^\perp \implies \nabla f(x_0) \perp T_{x_0} \Sigma \implies x_0$ è un punto critico di f su Σ sempre per il teorema 4.7. \square

Oss.: Supponiamo di avere una funzione $f : D \rightarrow \mathbb{R}$ con $D \subseteq \mathbb{R}^n$ con D compatto. Allora sappiamo, per il teorema di Weierstrass (teorema 2.5), che $\exists \max_D f, \min_D f$, tuttavia se non vi sono dei punti critici *liberi* della funzione in $\text{Int}(D)$ avremo che $\max_D f, \min_D f$ si trovano sulla frontiera di D . Nel caso in cui questi siano superfici/curve regolari allora bisogna procedere con i moltiplicatori di Lagrange per trovare i punti su di essa e il sistema che otteniamo con i moltiplicatori **dove** avere una soluzione altrimenti si otterrebbe un assurdo con il teorema di Weierstrass.

Oss.: Una trattazione a N vincoli dei problemi dell'ottimizzazione fa uso della funzione $\Lambda : \mathbb{R}^n \rightarrow \mathbb{R}$ lagrangiana definita come

$$\Lambda(x, \lambda_1, \dots, \lambda_N) = f(x) - \lambda_1 g_1(x) - \lambda_2 g_2(x) - \dots - \lambda_N g_N(x)$$

dove $x \in \mathbb{R}^n$ e le funzioni $g_i(x) : \mathbb{R}^n \rightarrow \mathbb{R}$ sono le equazioni esplicitate dei nostri vincoli. La teoria per dimostrare che questa formulazione (come l'intuito geometrico ci fa intuire) è pressoché equivalente, tuttavia non verrà dimostrato.

Oss.: Il moltiplicatore di Lagrange è naturalmente unico e dipende, com'è lecito aspettarsi, da $\nabla f(x_0)$ e $\nabla g(x_0)$: questa unicità è immediata conseguenza del fatto che $\nabla g(p) \neq 0$, dunque è una base dello spazio $(T_{x_0} \Sigma)^\perp$ e sappiamo, per unicità delle coordinate rispetto a vettori linearmente indipendenti, che le coordinate rispetto ad una base sono uniche.

CAPITOLO 5

Integrali curvilinei e potenziali

In questo capitolo andremo a generalizzare il concetto di integrale di Riemann con il concetto di integrale curvilineo, il quale consente di valutare una funzione muovendosi lungo una qualunque direzione, e il concetto di potenziale, il quale ricopre una grandissima importanza in Fisica.

5.1 Curve e integrali curvilinei

Vogliamo estendere il concetto di integrale di Riemann: il primo passo è definire il concetto di *integrale curvilineo*, ovvero l'integrale dove ci muoviamo lungo una *direzione* data dalla curva γ . Per fare ciò, dobbiamo però richiedere un certo tipo di regolarità da questa curva per non avere problemi nell'integrazione.

Definizione 5.1 (curva C^1 a tratti). Sia $\gamma : [a, b] \rightarrow \mathbb{R}^n$. Diremo che γ è C^1 a tratti se $\exists k \in \mathbb{N} : \exists t_0, \dots, t_k \in [a, b], a = t_0 < \dots < t_k = b$ tali che $\gamma \in C^1([t_j, t_{j+1}]) \forall j \in \{0, \dots, k-1\}$ e $\gamma \in C^0([a, b])$.

Possiamo anche definire la famiglia di curve C^1 a tratti su $[a, b]$

Definizione 5.2 (famiglie delle curve C^1 a tratti). Indichiamo con $C_{\text{tratti}}^1([a, b], \mathbb{R}^n)$ la famiglia di tutte le curve a tratti su $[a, b]$

Se vogliamo specificare la suddivisione $a = t_0 < t_1 < \dots < t_k = b$ per la quale la curva $\gamma \in C^1([t_j, t_{j+1}], \mathbb{R}^n) \forall j \in \{1, \dots, k\}$ diremo che la curva γ è C^1 a tratti rispetto alla suddivisione (o partizione) $a = t_0 < \dots < t_k = b$.

Definizione 5.3 (lunghezza di una curva γ). Sia $\gamma : [a, b] \rightarrow \mathbb{R}^n$, definiamo la lunghezza di γ come

$$l(\gamma) = \sup \left\{ \sum_{j=1}^k |\gamma(t_j) - \gamma(t_{j-1})| : a = t_0 < t_1 < \dots < t_k = b, k \in \mathbb{N}^+ \right\}$$

Consideriamo adesso una curva $\gamma \in C_{\text{tratti}}^1([a, b], \mathbb{R}^n)$ rispetto alla suddivisione $a = t_0 < t_1 < \dots < t_k = b$ allora definiamo

$$\int_a^b |\dot{\gamma}(t)| dt = \sum_{j=1}^k \int_{t_{j-1}}^{t_j} |\dot{\gamma}(t)| dt$$

e possiamo verificare che l'integrale non dipende dalla suddivisione rispetto la quale γ è C^1 a tratti, infatti se esistesse un'altra suddivisione $a = \tau_0 < \tau_1 < \dots < \tau_p = b$ allora

$$\sum_{i=1}^k \int_{t_{i-1}}^{t_i} |\dot{\gamma}(t)| dt = \sum_{i=1}^p \int_{\tau_{i-1}}^{\tau_i} |\dot{\gamma}(t)| dt$$

Teorema 5.1. Se $\gamma \in C_{\text{tratti}}^1([a, b], \mathbb{R}^n)$ allora

$$l(\gamma) = \int_a^b |\dot{\gamma}(t)|$$

Dimostrazione. Possiamo innanzitutto osservare che, fissando una parametrizzazione $a = t_0 < t_1 < \dots < t_k = b$ per cui la curva $\gamma \in C_{\text{tratti}}^1([a, b])$, allora:

$$\sum_{j=1}^k |\gamma(t_j) - \gamma(t_{j-1})| = \sum_{i=1}^k \left| \int_{t_{i-1}}^{t_i} \dot{\gamma}(t) dt \right| \leq \sum_{i=1}^k \int_{t_{i-1}}^{t_i} |\dot{\gamma}(t)| dt = \int_a^b |\dot{\gamma}(t)| dt$$

dunque, siccome la parametrizzazione scelta è arbitraria, allora

$$l(\gamma) = \sup \left\{ \sum_{j=1}^k |\gamma(t_j) - \gamma(t_{j-1})| : a = t_0 < t_1 < \dots < t_k = b, k \in \mathbb{N}^+ \right\} \leq \int_a^b |\dot{\gamma}(t)| dt$$

Possiamo, per concludere la dimostrazione, osservare che $\dot{\gamma}$, per il teorema di Heine-Cantor, è uniformemente continua (siccome $\gamma \in C_{\text{tratti}}^1$, la sua derivata sarà continua $\forall x \in (t_{i-1}, t_i) \forall i$), ovvero

$$\forall \varepsilon > 0, \exists \delta > 0 : \forall x, y, |x - y| < \delta \implies |\dot{\gamma}(x) - \dot{\gamma}(y)| < \varepsilon$$

e possiamo allora prendere una suddivisione tale che $\max\{t_{j+1} - t_j\} < \delta$ in maniera tale che

$$|\gamma(x) - \gamma(y)| \leq \varepsilon \quad \forall x, y \in [t_j, t_{j+1}]$$

dunque

$$\gamma(t_{j+1}) - \gamma(t_j) = \int_{t_j}^{t_{j+1}} \dot{\gamma}(t) dt = \int_{t_j}^{t_{j+1}} (\dot{\gamma}(t) - \dot{\gamma}(s) + \dot{\gamma}(s)) dt = \int_{t_j}^{t_{j+1}} (\dot{\gamma}(t) - \dot{\gamma}(s)) dt + \dot{\gamma}(s)(t_{j+1} - t_j).$$

Esplicitiamo a questo punto il termine con $\dot{\gamma}(s)$ e passiamo ai valori assoluti:

$$\begin{aligned} |\dot{\gamma}(s)|(t_{j+1} - t_j) &\leq |\gamma(t_{j+1}) - \gamma(t_j) - \int_{t_j}^{t_{j+1}} (\dot{\gamma}(t) - \dot{\gamma}(s)) dt| \stackrel{\text{dis. triang}}{\leq} \\ &\stackrel{\text{dis. triang}}{\leq} |\gamma(t_{j+1}) - \gamma(t_j)| + \left| \int_{t_j}^{t_{j+1}} (\dot{\gamma}(t) - \dot{\gamma}(s)) dt \right| \stackrel{|\dot{\gamma}(t) - \dot{\gamma}(s)| \leq \varepsilon}{\leq} |\dot{\gamma}(t_{j+1}) - \dot{\gamma}(t_j)| + \varepsilon(t_{j+1} - t_j) \end{aligned}$$

dividendo per $t_{j+1} - t_j$ concludiamo che

$$|\dot{\gamma}(s)| \leq \frac{|\gamma(t_{j+1}) - \gamma(t_j)|}{t_{j+1} - t_j} + \varepsilon$$

e, integrando per $s \in [t_j, t_{j+1}]$, avremo che

$$\int_{t_j}^{t_{j+1}} |\dot{\gamma}(s)| ds \leq \int_{t_j}^{t_{j+1}} \frac{|\gamma(t_{j+1}) - \gamma(t_j)|}{t_{j+1} - t_j} ds + \int_{t_j}^{t_{j+1}} \varepsilon ds = |\gamma(t_{j+1}) - \gamma(t_j)| + \varepsilon(t_{j+1} - t_j)$$

sommando adesso sulle j avremo che

$$\int_a^b |\dot{\gamma}(s)| ds \leq \sum_{j=1}^n |\gamma(t_{j+1}) - \gamma(t_j)| + \varepsilon(b-a) \leq l(\gamma) + \varepsilon(b-a)$$

ma, siccome ε è arbitrario, per $\varepsilon \rightarrow 0$ ricaviamo che

$$\int_a^b |\dot{\gamma}(s)| ds \leq l(\gamma)$$

ma allora segue necessariamente che $l(\gamma) = \int_a^b |\dot{\gamma}(t)| dt$. \square

Oss.: Il risultato ottenuto è vero anche per curve lipschitziane, tuttavia va modificata leggermente la dimostrazione.

Oss.: $l(\gamma)$ misura quello che noi consideriamo come *lunghezza* della curva solo se γ è iniettiva, altrimenti misura la lunghezza dell'immagine della curva "con molteplicità".

Definizione 5.4. Sia $\gamma \in C_{\text{tratti}}^1([a, b], \mathbb{R}^n)$ e $f : \gamma([a, b]) \rightarrow \mathbb{R}$ continua, definiamo l'integrale di lunghezza o l'integrale curvilineo di prima specie di f lungo γ come

$$\int_{\gamma} f = \int_a^b f(\gamma(t)) |\dot{\gamma}(t)| dt$$

Teorema 5.2. Sia $\gamma \in C_{\text{tratti}}^1([a, b], \mathbb{R}^n)$ e $f \in C^1(\gamma([a, b]))$. Allora

$$\varphi : [c, d] \rightarrow [a, b]$$

di classe C^1 e bigettiva. Allora $\gamma \circ \varphi \in C_{\text{tratti}}^1([c, d], \mathbb{R}^n)$ e

$$\int_{\gamma} f = \int_{\gamma \circ \varphi} f$$

Dimostrazione. Siccome $\varphi : [c, d] \rightarrow [a, b]$ è bigettiva, allora dev'essere monotona. Facciamo la dimostrazione, senza perdita di generalità, nel caso in cui φ è strettamente crescente e possiamo definire $\tau_i = \varphi^{-1}(t_j)$ (che sarà strettamente crescente a sua volta grazie alla continuità della funzione φ), dunque

$$\gamma \circ \varphi \in C^1([\tau_i, \tau_{i+1}], \mathbb{R}^n)$$

(nel caso di φ decrescente allora avevamo che $\gamma \circ \varphi \in C^1([\tau_{i+1}, \tau_i], \mathbb{R}^n)$) e la composizione di due funzioni continue è ancora una funzione continua, dunque $\gamma \circ \varphi \in C^0([c, d], \mathbb{R}^n)$ e questo ci porta a concludere che

$$\gamma \circ \varphi \in C_{\text{tratti}}^1([c, d], \mathbb{R}^n)$$

Mostriamo adesso la tesi per i singoli intervalli nel caso in cui φ è strettamente crescente

$$\int_{t_i}^{t_{i+1}} f(\gamma(t)) |\dot{\gamma}(t)| dt \stackrel{t=\varphi(\tau)}{=} \int_{\varphi^{-1}(t_i)}^{\varphi^{-1}(t_{i+1})} f((\gamma \circ \varphi)(\tau)) |\dot{\gamma}(\varphi(\tau))| |\dot{\varphi}(\tau)| d\tau$$

e osserviamo che, siccome φ è strettamente crescente, allora $|\dot{\varphi}(\tau)| = |\dot{\varphi}(\tau)|$, dunque (con un leggero abuso di notazione nelle variabili di integrazioni e sugli estremi di integrazione)

$$\int_{\tau_i}^{\tau_{i+1}} f((\gamma \circ \varphi)(\tau)) |\dot{\gamma}(\varphi(\tau))| |\dot{\varphi}(\tau)| d\tau = \int_{\tau_i}^{\tau_{i+1}} f((\gamma \circ \varphi)(\tau)) |\gamma \circ \varphi|'(\tau) d\tau.$$

Sommendo in j otteniamo che

$$\int_{\gamma} f = \sum_{i=1}^k \int_{t_i}^{t_{i+1}} f(\gamma(t)) |\dot{\gamma}(t)| dt = \sum_{i=1}^l \int_{\tau_i}^{\tau_{i+1}} f((\gamma \circ \varphi)(\tau)) |\gamma \circ \varphi'(\tau)| d\tau = \int_{\gamma \circ \varphi} f$$

ovvero la tesi. \square

Corollario 5.2.1 (lunghezza di un arco). Sia $f(x) : I \rightarrow \mathbb{R}$ con $f \in C^1(I)$ allora la lunghezza dell'arco descritto da $f(x)$ è dato da

$$\int_I \sqrt{1 + f'(x)} dx$$

Dimostrazione. Dal teorema precedente, possiamo considerare la curva $\gamma : I \rightarrow \mathbb{R}^2$ tale che $\gamma = (t, f(t)) \forall t \in I$ e calcolare la lunghezza della curva con il teorema precedente:

$$l(\gamma) = \int_I |\dot{\gamma}(t)| dt = \int_I \sqrt{1 + (f'(t))^2} dt$$

ovvero la tesi. \square

5.2 Integrali curvilinei su forme differenziali e campi vettoriali

Abbiamo generalizzato l'integrale di Riemann all'integrale curvilineo come l'integrale di una funzione a valori reali lungo una direzione impartita dalla curva γ , vogliamo adesso generalizzarlo a funzioni a valori vettoriali. I candidati ideali per fare questo sono le forme differenziali e i campi vettoriali di cui richiediamo una certa regolarità

Definizione 5.5 (1-forma differenziale di classe C^k). Sia $\Omega \subseteq \mathbb{R}^n$ un insieme aperto, allora

$$\omega(x) = \sum_{j=1}^n a_j(x) dx_j$$

è una 1-forma differenziale di classe $C^k(\Omega)$, ovvero

$$\omega \in C^k(\Omega, (\mathbb{R}^n)^*) \iff a_j \in C^k(\Omega) \forall j \in \{1, \dots, n\}$$

Diamo adesso la nozione di *esattezza* di una forma differenziale

Definizione 5.6 (esattezza di una 1-forma differenziale). Diremo che ω è esatta se esiste una primitiva $V \in C^1(\Omega)$ tale che

$$dV = \omega$$

ovvero se $\omega = \sum_j a_j dx_j$ allora $a_j = \frac{\partial V}{\partial x_j}$ su $\Omega \forall j \in \{1, \dots, n\}$

Equivalentemente diamo la nozione di *chiusura* di una forma differenziale

Definizione 5.7. Sia $\omega = \sum_{j=1}^n a_j dx_j \in C^k(\Omega, (\mathbb{R}^n)^*)$, ovvero una 1-forma differenziale di classe $C^k(\Omega)$, allora diremo che è chiusa se abbiamo che

$$\frac{\partial a_j}{\partial x_i} - \frac{\partial a_i}{\partial x_j} = 0 \quad \forall i, j : 1 \leq i < j \leq n$$

Equivalentemente, possiamo dare delle definizioni che sono "duali" nel caso, però, di campi vettoriali:

Definizione 5.8 (campo vettoriale di classe C^k). Sia $\Omega \subseteq \mathbb{R}^n$ un insieme aperto, allora

$$F(x) = \sum_{j=1}^n f_j(x)e_j$$

è un campo vettoriale di classe C^k , ovvero

$$F \in C^k(\Omega, \mathbb{R}^n) \iff f_j \in C^k(\Omega) \forall j \in \{1, \dots, n\}$$

Nel caso di campi vettoriali la nozione di *esattezza* e di *chiusura* è sostituita dal concetto di *campo conservativo* (che è un concetto che ogni fisico ha incontrato nel corso di Fisica 1 nel caso del moto in campo centrale per esempio) e dal concetto di *campo irrotazionale*

Definizione 5.9 (campo conservativo). Diremo che F è un campo conservativo se esiste un potenziale $V \in C^1(\Omega)$ tale che

$$F = \nabla V$$

ovvero se $F = (f_1, \dots, f_n)$ e abbiamo che $f_j = \frac{\partial V}{\partial x_j}$ su $\Omega \forall j \in \{1, \dots, n\}$

Definizione 5.10 (campo irrotazionale). Sia $\Omega \subseteq \mathbb{R}^n$ aperto e sia $F = \sum_{i=1}^n f_i e_i \in C^k(\Omega, \mathbb{R}^n)$ un campo vettoriale, diremo che F è un campo irrotazionale se

$$\frac{\partial f_i}{\partial x_j} - \frac{\partial f_j}{\partial x_i} = 0 \quad \forall i, j : 1 \leq i < j \leq n$$

Possiamo andare a definire il concetto di integrale curvilineo di seconda specie di ω (oppure di F) lungo una curva γ come segue

Definizione 5.11 (integrale curvilineo di seconda specie di ω lungo una curva γ). Sia $\Omega \subseteq \mathbb{R}^n$, $\omega \in C^0(\Omega, (\mathbb{R}^n)^*)$ una 1-forma differenziale e sia $\gamma \in C^1_{\text{tratti}}([a, b], \Omega)$ una curva. Definiamo l'integrale curvilineo di seconda specie di ω lungo una curva γ come

$$\int_{\gamma} \omega = \sum_{i=1}^k \int_{t_{j-1}}^{t_j} \omega(\gamma(t))(\dot{\gamma}(t)) dt$$

Definizione 5.12 (integrale curvilineo di seconda specie di F lungo una curva γ). Sia $\Omega \subseteq \mathbb{R}^n$, $F \in C^0(\Omega, \mathbb{R}^n)$ un campo vettoriale e sia $\gamma \in C^1_{\text{tratti}}([a, b], \Omega)$ una curva. Definiamo l'integrale curvilineo di seconda specie di F lungo una curva γ come

$$\int_{\gamma} F = \sum_{i=1}^k \int_{t_{j-1}}^{t_j} \langle F(\gamma(t)), \dot{\gamma}(t) \rangle dt$$

Oss.: Dalla definizione di una forma differenziale sappiamo che

$$\omega(\gamma(t))(\dot{\gamma}(t)) = \sum_{i=1}^n a_i(\gamma(t))dx_i(\gamma(t)) = \sum_{i=1}^n a_i(\gamma(t))\gamma_i(t)$$

Oss.: Questa *dualità* presente fra i campi vettoriali e le 1-forme differenziali deriva dal fatto che, di fatto, questi sono la stessa cosa: infatti, noi potremmo pensare di considerare $F \in C^0(\Omega, \mathbb{R}^n)$, $F = (f_1, \dots, f_n)$ e costruire una 1-forma del tipo $\omega_F = f_1 dx_1 + f_2 dx_2 + \dots + f_n dx_n$ come 1-forma differenziale associata ad F e osservare che $\forall \gamma \in C_{\text{tratti}}^1([a, b], \Omega)$ abbiamo che

$$\int_{\gamma} F = \int_{\gamma} \omega_F$$

Viceversa, presa $\omega = \sum_{i=1}^n a_i dx_i$ se definiamo $F_{\omega} = (a_1, a_2, \dots, a_n) \implies$

$$\int_{\gamma} \omega = \int_{\gamma} F_{\omega}$$

Tali eguaglianze seguono immediatamente dalle definizioni e giustificano questa doppia formulazione che abbiamo dato.

Osservando la questione da un punto di vista più elevato, in presenza di metrica riemanniana, l'applicazione dell'operazione di contrazione fornisce un modo per far corrispondere ad ogni campo vettoriale una 1-forma differenziale *contraendo* il campo vettoriale con la metrica. Quindi da questo breve accenno senza formule, possiamo intuire che campi vettoriali e 1-forme differenziali siano, in sostanza, due lati della stessa moneta.

Con le forme differenziali (e dunque anche con i campi vettoriali) possiamo giungere ad un risultato equivalente a quello raggiunto con l'integrale valutato su una composizione di curve, così come avevamo fatto nel caso di funzioni a valori reali

Teorema 5.3 (di invarianza). Sia $\gamma \in C_{\text{tratti}}^1([a, b], \Omega)$, sia $\omega \in C^0(\Omega, (\mathbb{R}^n)^*)$ e sia $\varphi : [c, d] \rightarrow [a, b]$ di classe C^1 e invertibile. Ponendo $\sigma(\varphi) = 1$ se φ è strettamente crescente e $\sigma(\varphi) = -1$ se φ è strettamente decrescente. Allora abbiamo che

$$\int_{\gamma} \omega = \sigma \int_{\gamma \circ \varphi} \omega$$

Dimostrazione. La dimostrazione è abbastanza simile a quella fatta nel caso di funzioni a valori reali, infatti osserviamo che

$$\int_{\gamma} \omega = \sum_{j=1}^k \int_{t_{j-1}}^{t_j} \omega(\gamma(t))(\dot{\gamma}(t)) dt \stackrel{\tau_j = \varphi^{-1}(t_j)}{=} \sum_{j=1}^k \int_{\tau_{i-1}}^{\tau_i} \omega((\gamma \circ \varphi)(\tau))(\dot{\gamma}(\varphi(\tau))(\dot{\varphi}(\tau))) d\tau$$

Supponiamo che φ sia strettamente decrescente, dunque anche φ^{-1} è strettamente decrescente, quindi $\tau_j < \tau_{j-1}$ e, conseguentemente, $a = \tau_k < \tau_{k-1} < \dots < \tau_0 = b$ pertanto

$$\begin{aligned} \int_{\gamma} \omega &= \sum_{j=1}^k \int_{\tau_{i-1}}^{\tau_i} \omega((\gamma \circ \varphi)(\tau))(\dot{\gamma}(\varphi(\tau))(\dot{\varphi}(\tau))) d\tau = \sum_{j=1}^k - \int_{\tau_i}^{\tau_{i-1}} \omega((\gamma \circ \varphi)(\tau))(\dot{\gamma}(\varphi(\tau))(\dot{\varphi}(\tau))) d\tau = \\ &= \sigma \sum_{j=1}^k \int_{\tau_i}^{\tau_{i-1}} \omega((\gamma \circ \varphi)(\tau))(\dot{\gamma}(\varphi(\tau))(\dot{\varphi}(\tau))) d\tau = \sigma \int_{\gamma \circ \varphi} \omega \end{aligned}$$

ovvero la tesi. Nel caso in cui φ fosse strettamente crescente e, dunque, $\sigma = 1$ la sommatoria già dà l'integrale curvilineo della composizione \square

5.2.1 Forme differenziali esatte

Proposizione 5.1. Siano $\Omega \subseteq \mathbb{R}^n$ un insieme aperto e $\omega \in C^0(\Omega, (\mathbb{R}^n)^*)$ tale che ω sia esatta con primitiva $P \in C^1(\Omega)$. Allora $\forall \gamma \in C_{\text{tratti}}^1([a, b], \Omega)$ abbiamo che

$$\int_{\gamma} \omega = \int_{\gamma} dP = P(\gamma(b)) - P(\gamma(a))$$

Equivalentemente, nel caso vettoriale, sia sempre $\Omega \subseteq \mathbb{R}^n$ un insieme aperto e sia $F \in C^0(\Omega, \mathbb{R}^n)$ con potenziale $V \in C^1(\Omega)$.

Allora $\forall \gamma \in C_{\text{tratti}}^1([a, b], \Omega)$ abbiamo che

$$\int_{\gamma} F = \int_{\gamma} \nabla V = V(\gamma(b)) - V(\gamma(a))$$

Dimostrazione. E' quasi un conto siccome

$$\begin{aligned} \int_{\gamma} \omega &= \int_{\gamma} dP = \sum_{j=1}^k \int_{\gamma|_{[t_{j-1}, t_j]}} dP = \sum_{j=1}^k \int_{t_{j-1}}^{t_j} dP(\gamma(t))(\dot{\gamma}(t))dt = \sum_{j=1}^k \int_{t_{j-1}}^{t_j} \sum_{i=1}^n \frac{\partial P}{\partial x_i}(\gamma(t))\dot{\gamma}_i(t)dt = \\ &= \sum_{j=1}^k \int_{t_{j-1}}^{t_j} \frac{d}{dt}((P \circ \gamma)(t))dt = \sum_{j=1}^k (P \circ \gamma)(t_j) - (P \circ \gamma)(t_{j-1}) = (P \circ \gamma)(b) - (P \circ \gamma)(a) \end{aligned}$$

ovvero la tesi. Per quanto riguarda il campo vettoriale, osserviamo che

$$\int_{\gamma} F = \int_{\gamma} \nabla V = \sum_{j=1}^k \int_{t_{j-1}}^{t_j} \sum_{i=1}^n \frac{\partial V}{\partial x_i}(\gamma(t))\dot{\gamma}_i(t)$$

e con ragionamenti simili a quanto fatto prima possiamo arrivare alla tesi. \square

Corollario 5.3.1. Siano $\Omega \subseteq \mathbb{R}^n$ un insieme aperto e $\omega \in C^0(\Omega, (\mathbb{R}^n)^*)$ tale che ω sia esatta. Allora $\forall \gamma \in C_{\text{tratti}}^1([a, b], \Omega)$ chiusa abbiamo che

$$\oint_{\gamma} \omega = 0$$

Equivalentemente, sia sempre $\Omega \subseteq \mathbb{R}^n$ un insieme aperto e $F \in C^0(\Omega, \mathbb{R}^n)$ tale che F sia un campo conservativo. Allora $\forall \gamma \in C_{\text{tratti}}^1([a, b], \Omega)$ chiusa abbiamo che

$$\oint_{\gamma} F = 0$$

Dimostrazione. Applicando la proposizione precedente, se P è una primitiva di ω , sappiamo che

$$\oint_{\gamma} \omega = P(\gamma(b)) - P(\gamma(a)) = P(\gamma(a)) - P(\gamma(a)) = 0$$

Procediamo allo stesso modo per i campi vettoriali. \square

Esercizio 5.1 (integrali in coordinate polari). *Spesso, a seconda della simmetria di determinate funzioni, è molto comodo passare in coordinate polari. Consideriamo una curva $\gamma(t)$ della forma*

$$\gamma(t) = (\rho(t) \cos \varphi(t), \rho \sin \varphi(t))$$

dove, naturalmente, richiediamo un certo tipo di regolarità alla nostra curva per non avere problemi nell'integrazione, pertanto richiediamo che

$$\begin{aligned}\rho &\in C^1([a, b], (0, +\infty)) \\ \varphi &\in C^1([a, b], [0, 2\pi])\end{aligned}$$

Dimostrare la seguente formula per il calcolo della lunghezza della curva γ :

$$l(\gamma) = \int_a^b ((\dot{\rho})^2 + (\rho \dot{\varphi})^2)^{\frac{1}{2}} dt$$

e la seguente formula dell'integrale di una funzione $f \in C^0(\gamma([a, b]), \mathbb{R})$ valutata lungo la funzione γ :

$$\int_{\gamma} f = \int_{\gamma} f(\rho(t) \cos \varphi(t), \rho(t) \sin \varphi(t)) \sqrt{\dot{\rho}^2 + \rho \dot{\varphi}^2} dt$$

Svolgimento.

Tramite i teoremi visti nella precedente sezione sappiamo che

$$\begin{aligned}l(\gamma) &= \int_{\gamma} dl = \int_a^b ((\dot{\rho} \cos \varphi - \rho \dot{\varphi} \sin \varphi)^2 + (\dot{\rho} \sin \varphi + \rho \dot{\varphi} \cos \varphi)^2)^{\frac{1}{2}} dt = \\ &= \int_a^b ((\dot{\rho})^2 + (\rho \dot{\varphi})^2)^{\frac{1}{2}} dt\end{aligned}$$

Se $f \in C^0(\gamma([a, b]), \mathbb{R})$ allora

$$\int_{\gamma} f = \int_a^b f(\gamma(t)) |\dot{\gamma}(t)| dt = \int_{\gamma} f(\rho(t) \cos \varphi(t), \rho(t) \sin \varphi(t)) \sqrt{\dot{\rho}^2 + \rho \dot{\varphi}^2} dt$$

□

Andiamo a mostrare l'additività dell'integrazione curvilinea di seconda specie rispetto alla "somma" di curve. Per la prossima serie di proposizioni faremo implicitamente uso del seguente lemma che, geometricamente, garantisce il fatto di poter individuare, presi due punti a e b di un sottoinsieme aperto e connesso per archi $\Omega \subseteq \mathbb{R}^n$, una poligonale C_{tratti}^1 che li unisce.

Lemma 5.3.1. Sia $\Omega \subseteq \mathbb{R}^n$ aperto e connesso per archi. Per ogni coppia $x, y \in \Omega$ esiste $\gamma : [a, b] \rightarrow \Omega$ tale che $\gamma(a) = x$, $\gamma(b) = y$ e $\gamma \in C_{\text{tratti}}^1$.

Dimostrazione. Dalla connessione per archi di Ω sappiamo che esiste $\gamma : [a, b] \rightarrow \Omega$ continua tale che $\gamma(a) = x$ e $\gamma(b) = y$.

Potremo a questo punto pensare di "ricoprire" con un numero finito di palle aperte¹ la linea $\gamma([a, b])$ osservando che $\forall x \in \Omega, \exists r > 0 : B_r(x) \subseteq \Omega$. Preso $a = x_0$ allora prendiamo $r_0 > 0 : B_{r_0} \subseteq \Omega$ (garantito dall'ipotesi che Ω sia aperto) e consideriamo l'insieme $B_{r_0}(x_0) \cap \gamma([a, b]) \neq \emptyset$ (se fosse vuoto allora avremmo un assurdo riguardo all'ipotesi che Ω sia connesso per archi). Preso $x_1 \in B_{r_0}(x_0) \cap \gamma([a, b])$ con $x_1 \neq x_0$, possiamo, a questo punto, iterare il procedimento descritto e ottenere il partizionamento di una curva C_{tratti}^1 fra a e b . □

¹il fatto di poter ricoprire con un numero finito di palle deriva dal fatto che la funzione γ è continua, dunque l'intervallo $[x, y] = \gamma([a, b])$ è un insieme compatto, pertanto è, per definizione di compattezza in senso topologico, possibile per ogni ricoprimento di insiemi aperti trovare un sottoricoprimento finito di aperti.

Proposizione 5.2. Sia $\Omega \subseteq \mathbb{R}^n$ un insieme aperto, $\gamma \in C_{\text{tratti}}^1([a, b], \Omega)$, $\Gamma \in C_{\text{tratti}}^1([c, d], \Omega)$ tali che $\gamma(b) = \Gamma(c)$. Allora, definendo

$$\alpha(t) = \begin{cases} \gamma(t) & \text{se } a \leq t \leq b \\ \Gamma(c + t - b) & \text{se } b \leq t \leq b + d - c \end{cases},$$

avremo che $\alpha \in C_{\text{tratti}}^1([a, b + d - c], \Omega)$ e se $\omega \in C^0([a, b + d - c], (\mathbb{R}^n)^*)$ allora

$$\int_{\alpha} \omega = \int_{\gamma} \omega + \int_{\Gamma} \omega$$

Equivalentemente, per un campo vettoriale $F \in C^0(\Omega, \mathbb{R}^n)$ allora

$$\int_{\alpha} F = \int_{\gamma} F + \int_{\Gamma} F$$

Dimostrazione. Il fatto che α sia C_{tratti}^0 deriva dal fatto che, per $0 \leq t \leq b$ $\alpha(t) = \gamma(t)$ e per $b \leq t \leq b + d - c$ $\alpha(t) = \Gamma(c + t - b)$ e per $t = b$, $\alpha(b) = \gamma(b) = \Gamma(c)$ (dove l'ultima eguaglianza segue per ipotesi), da cui avremo anche che $\alpha \in C_{\text{tratti}}^1([a, b + d - c], \Omega)$. Consideriamo un partizionamento della curva γ dato da $a = a_0 < a_1 < \dots < a_k$ e il partizionamento Γ dato da $c = b_0 < b_1 < \dots < b_n$

$$t_j = \begin{cases} a_j & \text{se } j \leq k \\ b_{j-k} - c + b & \text{se } k < j \leq k + n \end{cases}$$

dunque

$$\begin{aligned} \int_{\alpha} \omega &= \sum_{j=0}^k \int_{t_j}^{t_{j+1}} \omega(\alpha(t))(\dot{\alpha}(t))dt = \sum_{j=0}^k \int_{t_j}^{t_{j+1}} \omega(\alpha(t))(\dot{\alpha}(t))dt + \sum_{j=k}^n \int_{t_j}^{t_{j+1}} \omega(\alpha(t))(\dot{\alpha}(t))dt = \\ &= \sum_{j=0}^k \int_{a_j}^{a_{j+1}} \omega(\gamma(t))(\dot{\gamma}(t))dt + \sum_{j=k}^n \int_{b_{j-k} - c + b}^{b_{j+1} - c + b} \omega(\Gamma(t))(\dot{\Gamma}(t))dt \end{aligned}$$

possiamo effettuare un cambio di variabile nella seconda sommatoria con $j' = j - k$ (nella sommatoria si usa comunque j essendo gli indici della sommatoria delle variabili mute), pertanto

$$= \sum_{j=0}^k \int_{a_j}^{a_{j+1}} \omega(\gamma(t))(\dot{\gamma}(t))dt + \sum_{j=0}^n \int_{b_j + b - c}^{b_{j+1} + b - c} \omega(\Gamma(t))(\dot{\Gamma}(t))dt =$$

Possiamo a questo punto fare un cambio di variabile all'interno dell'integrale, prendendo la variabile $t' = t - b + c \implies dt' = dt$

$$\sum_{j=0}^k \int_{a_j}^{a_{j+1}} \omega(\gamma(t))(\dot{\gamma}(t))dt + \sum_{j=0}^n \int_{b_j}^{b_{j+1}} \omega(\Gamma(t'))(\dot{\Gamma}(t'))dt' = \int_{\gamma} \omega + \int_{\Gamma} \omega$$

e questo conclude la dimostrazione. \square

Oss.: Motivati da questa proprietà appena mostrata, si fa spesso uso della seguente notazione $\alpha = \gamma + \Gamma$, tuttavia bisogna naturalmente tenere di conto che non si tratta di una somma effettiva di curve.

Teorema 5.4 (teorema CF1). Sia $\Omega \subseteq \mathbb{R}^n$ aperto e connesso per archi. Sia $\omega \in C^0(\Omega, (\mathbb{R}^n)^*)$. Allora ω è esatta $\iff \forall \gamma \in C_{\text{tratti}}^1([a, b], \Omega)$ chiusa abbiamo che

$$\oint_{\gamma} \omega = 0$$

Equivalentemente, se $F \in C^0(\Omega, \mathbb{R}^n)$, allora F è conservativo $\iff \forall \gamma \in C_{\text{tratti}}^1([a, b], \Omega)$ chiusa abbiamo che

$$\oint_{\gamma} F = 0$$

Dimostrazione.

\Rightarrow : se ω è esatta allora, per il corollario 5.3.1, abbiamo che

$$\oint_{\gamma} \omega = 0$$

\Leftarrow : per quanto riguarda l'implicazione inversa, osserviamo fissando un punto $x_0 \in \Omega$ e consideriamo $x \in \Omega$, allora posto

$$V(x) = \int_{\gamma_x} \omega$$

dove $\gamma_x \in C_{\text{tratti}}^1([a, b], \Omega)$ e consideriamo una palla $B_r(x) \subseteq \Omega$ (la sua esistenza è assicurata dal fatto che Ω è un insieme aperto), pertanto possiamo considerare, per $0 < |h| < r$, un cammino che va da x a $x + h$. Vogliamo mostrare che $dV = \omega$, il che è equivalente a mostrare che

$$V(x + h) - V(x) - \langle \omega(x), h \rangle = o(|h|)$$

per $h \rightarrow 0$:

$$F(x + h) - F(x) = \int_{\gamma_x + [x, x+h]} \omega - \int_{\gamma_x} \omega \stackrel{\text{proposizione 5.2}}{=} \int_{[x, x+h]} \omega$$

parametrizzando il cammino $[x, x + h]$ con $x + th$ dove $t \in [0, 1]$ allora

$$F(x+h) - F(x) - \langle \omega(x), h \rangle = \int_0^1 \langle \omega(x+th), h \rangle dt - \int_0^1 \langle \omega(x), h \rangle dt = \int_0^1 \langle \omega(x+th) - \omega(x), h \rangle dt$$

Siccome $\omega \in C^0(\Omega, (\mathbb{R}^n)^*)$ allora sappiamo che

$$\forall \varepsilon > 0 \exists \delta > 0 : \forall x' \in B_\delta(x), |\omega(x') - \omega(x)| \leq \varepsilon$$

dunque, se fissiamo $\varepsilon > 0$, allora sappiamo che per $|h| < \delta$ abbiamo che

$$\begin{aligned} \int_0^1 \langle \omega(x+th) - \omega(x), h \rangle dt &\leq \int_0^1 |\omega(x+th) - \omega(x)| |h| dt \leq \varepsilon |h| \implies \\ \left| \int_0^1 \langle \omega(x+th) - \omega(x), h \rangle dt \right| &\leq \varepsilon |h| \implies \int_0^1 \langle \omega(x+th) - \omega(x), h \rangle dt = o(|h|) \end{aligned}$$

dunque $V(x)$ è differenziabile con $dV = \omega$ (per unicità del differenziale) e, dunque, $V \in C^1(\Omega)$. Mostriamo adesso che il valore di $V(x)$ non dipende dalla curva scelta, consideriamo, oltre a γ_x , la curva $\Gamma_x \in C_{\text{tratti}}^1([c, d], \Omega)$ tale che $\Gamma_x(c) = x_0$ e $\Gamma_x(d) = x$. Possiamo considerare la curva

$\tilde{\Gamma}(t) = \Gamma(c + d - t)$ e definendo la curva $\alpha(t)$ come nella proposizione 5.2, avremo che $\alpha_x = \gamma_x + \tilde{\Gamma}$ è una curva chiusa (che possiamo fare siccome $\gamma_x(b) = \tilde{\Gamma}(c) = x$).

Conseguentemente, avremo che $\alpha_x \in C^1_{\text{tratti}}([a, b + d - c], \Omega)$ ed è chiusa, dunque

$$\int_{\alpha_x} \omega = \int_{\gamma_x + \tilde{\Gamma}_x} \omega = \int_{\gamma_x} \omega + \int_{\tilde{\Gamma}_x} \omega = \int_{\gamma_x} \omega - \int_{\Gamma_x} \omega \stackrel{\text{per ip.}}{=} 0 \implies \int_{\gamma_x} \omega = \int_{\Gamma_x} \omega$$

dove si osserva che $\int_{\tilde{\Gamma}_x} \omega = -\int_{\Gamma_x} \omega$ per il teorema di invarianza (teorema 5.3) siccome $\tilde{\Gamma}_x = \Gamma_x \circ \varphi$ dove $\varphi(t) = c + d - t$. Abbiamo pertanto dimostrato che dunque la funzione $V(x)$ non dipende dalla curva C^1 a tratti scelta \square

Teorema 5.5 (teorema CF2). Sia $\Omega \subseteq \mathbb{R}^n$ aperto e connesso per archi. Sia $\omega \in C^0(\Omega, (\mathbb{R}^n)^*)$, allora ω è esatta $\iff \forall \gamma \in C^1_{\text{tratti}}([a, b], \Omega)$ e $\forall \Gamma \in C^1_{\text{tratti}}([c, d], \Omega)$ tali che $\gamma(a) = \Gamma(c)$ e $\gamma(b) = \Gamma(d)$,

$$\int_{\gamma} \omega = \int_{\Gamma} \omega$$

Dimostrazione.

\Rightarrow : è banale siccome se ω è chiusa allora possiamo prendere una curva $\gamma \in C^1_{\text{tratti}}([a, b], \Omega)$ e $\Gamma \in C^1_{\text{tratti}}([c, d], \Omega)$ e definiamo una curva $\alpha(t)$ come fatto nella proposizione 5.2, dunque avremo che

$$\oint_{\alpha} \omega = \int_{\gamma} \omega + \int_{\tilde{\Gamma}} \omega \stackrel{\text{corollario 5.3.1}}{=} 0 \implies \int_{\gamma} \omega = \int_{\Gamma} \omega$$

ovvero quello che volevamo mostrare.

\Leftarrow : consideriamo una curva chiusa $c(t) : [t_1, t_2] \rightarrow \Omega$ con $c \in C^1_{\text{tratti}}$ con $-\infty < t_1 < t_2 < +\infty$. Possiamo prendere $\tau = \frac{t_1+t_2}{2}$ e $c_1 = c|_{[t_1, \tau]}$ e $c_2 = c|_{[\tau, t_2]}$.

In analogia a quanto fatto nei precedenti teoremi possiamo definire una funzione $\tilde{c}_2 : [\tau, t_2] \rightarrow \Omega$ dove $\tilde{c}_2 = c_2(t_2 + \tau - t)$, pertanto avremo che $\tilde{c}_2(t_2) = c_2(t_2 + \tau - t_2) = c_2(t_2) = c(t_2) = c(t_1) = c_1(t_1)$ essendo c una curva chiusa. Inoltre $\tilde{c}_2(t_2) = c(\tau) = c_1(\tau) \implies \int_{c_1} \omega = \int_{\tilde{c}_2} \omega$, tuttavia osserviamo che, in virtù del teorema di invarianza (teorema 5.3), avremo che $\int_{\tilde{c}_2} \omega = -\int_{c_2} \omega$, siccome \tilde{c}_2 è decrescente, conseguentemente avremo che

$$\int_c \omega = \int_{c_1} \omega + \int_{c_2} \omega \stackrel{\text{per ipotesi}}{=} 0 = \int_{c_1} \omega - \int_{\tilde{c}_2} \omega = 0 \implies \int_{c_1} \omega = \int_{\tilde{c}_2} \omega.$$

e la tesi segue dal teorema 5.4, siccome esso garantisce adesso che ω sia esatta. \square

Proposizione 5.3 (le 1-forme differenziali esatte sono chiuse). Sia $\omega \in C^1(\Omega, (\mathbb{R}^n)^*)$ tale che ω sia esatta. Allora ω è chiusa.

Equivalentemente, se $F \in C^1(\Omega, \mathbb{R}^n)$ tale che F sia un campo conservativo. Allora F è irrotazionale.

Dimostrazione. Sia $P \in C^1(\Omega)$ una primitiva di ω con $\omega = \sum_{j=1}^n a_j dx_j$, $a_j = \frac{\partial P}{\partial x_j}$ su $\Omega \forall j \in \{1, \dots, n\}$. Poiché ω è C^1 avremo che $a_j \in C^1(\Omega) \forall j \in \{1, \dots, n\} \implies P \in C^2(\Omega, \mathbb{R}^n)$. Per il teorema di Schwarz allora avremo che

$$\frac{\partial^2 P}{\partial x_i \partial x_j} = \frac{\partial^2 P}{\partial x_j \partial x_i} \quad \forall i, j \in \{1, \dots, n\} \text{ con } i \neq j$$

ma allora

$$\begin{aligned} \frac{\partial a_i}{\partial x_j} = \frac{\partial^2 P}{\partial x_j \partial x_i} &\stackrel{\text{thm di Schwarz}}{=} \frac{\partial^2 P}{\partial x_i \partial x_j} = \frac{\partial a_j}{\partial x_i} \implies \frac{\partial a_i}{\partial x_j} - \frac{\partial a_j}{\partial x_i} = 0 \quad \forall 1 \leq i < j \leq n \implies \\ &\implies \omega \text{ è chiusa} \end{aligned}$$

ovvero la tesi. Analogamente si può fare la dimostrazione nel caso di campi vettoriali. \square

Oss.: L'ipotesi che $\omega \in C^1(\Omega, (\mathbb{R}^n)^*)$ è fondamentale: esistono infatti delle forme differenziali esatte che non sono chiuse. Come esempio, consideriamo la funzione definita su \mathbb{R}^2

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

osserviamo che $F(x, 0) = F(y, 0) = 0 \quad \forall x \in \mathbb{R}, \forall y \in \mathbb{R}$. Avremo dunque che

$$\partial_x F(0, 0) = \lim_{x \rightarrow 0} \frac{F(x, 0) - F(0, 0)}{x} = \lim_{x \rightarrow 0} \frac{0 - 0}{x} = 0$$

mentre, al di fuori dell'origine, avremo che

$$\partial_x F(x, y) = \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2}$$

dunque

$$\partial_x F = \begin{cases} \frac{y(x^4 + 4x^2y^2 - y^4)}{(x^2 + y^2)^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

La continuità della derivata segue dal fatto che

$$\lim_{(x,y) \rightarrow (0,0)} \partial_x F(x, y) = \partial_x F(0, 0) = 0$$

infatti, passando in coordinate polari, abbiamo che

$$\begin{aligned} \lim_{(x,y) \rightarrow (0,0)} \partial_x F(x, y) &= \lim_{\rho \rightarrow 0} \frac{\rho \sin \varphi \rho^4 (\cos^4 \varphi + 4 \cos^2 \varphi \sin^2 \varphi - \sin^4 \varphi)}{\rho^4} = \\ &= \lim_{\rho \rightarrow 0} \rho \sin \varphi (\cos 2\varphi + \sin^2 2\varphi) = 0. \end{aligned}$$

Analogamente si ricava lo stesso per quanto riguarda la derivata parziale rispetto a y

$$\partial_y F(x, y) = \begin{cases} \frac{x(x^4 - 4x^2y^2 - y^4)}{(x^2 + y^2)^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0). \end{cases}$$

Definiamo allora la 1-forma

$$\omega = X(x, y)dx + Y(x, y)dy$$

dove

$$\begin{aligned} X(x, y) &= \partial_x F(x, y) \\ Y(x, y) &= \partial_y F(x, y). \end{aligned}$$

Questa forma è, naturalmente, esatta, siccome ha come primitiva la funzione $F(x, y) + c$. Tale forma non risulta chiusa, però, infatti

$$\partial_y X(0, 0) = F_{xy}(0, 0) = \lim_{y \rightarrow 0} \frac{\partial_x F(0, y) - \partial_x F(0, 0)}{y} = \lim_{y \rightarrow 0} \frac{-\frac{y^5}{y^4} - 0}{y} = \lim_{y \rightarrow 0} (-1) = -1$$

mentre

$$\partial_x Y(0, 0) = F_{yx}(0, 0) = \lim_{x \rightarrow 0} \frac{\partial_y F(x, 0) - \partial_y F(0, 0)}{x} = \lim_{x \rightarrow 0} \frac{\frac{x^5}{x^4} - 0}{x} = 1.$$

Il fatto che la forma dell'esempio sia esatta su \mathbb{R}^2 ma non chiusa è legata al fatto che i suoi coefficienti X e Y , sebbene continui su \mathbb{R}^2 e di cui esiste la derivata $\partial_y X$ e $\partial_x Y$, non ammettono derivate continue.

Faremo uso adesso del seguente lemma

Lemma 5.5.1 (CF2). Sia $\Omega \subseteq \mathbb{R}^n$ un insieme aperto, connesso per archi e sia $f \in C^1(\Omega)$ tale che $\nabla f(x) = 0 \forall x \in \Omega$. Allora $f(x) = c \forall x \in \mathbb{R}^n$

Dimostrazione. Fissato $x_0 \in \Omega$ allora indichiamo con $c \in \mathbb{R}$ il valore assunto dalla funzione f in x_0 . Sia $x \in \Omega$ arbitrario e sappiamo sempre che $\exists \gamma \in C^1_{\text{tratti}}([a, b], \Omega)$ tale che $\gamma(a) = x_0$ e $\gamma(b) = x$. Conseguentemente avremo che

$$f(x) - f(x_0) = \sum_{j=1}^k f(\gamma(t_j)) - f(\gamma(t_{j-1}))$$

dove l'eguaglianza è ottenuta costruendo una semplice serie telescopica. Osserviamo adesso che $f(\gamma(t_j)) - f(\gamma(t_{j-1})) = \int_{t_{j-1}}^{t_j} \langle \nabla f(\gamma(t)), \dot{\gamma}(t) \rangle dt = 0$ in virtù dell'annullamento del gradiente. \square

Proposizione 5.4. Sia $\Omega \subseteq \mathbb{R}^n$ un insieme aperto e connesso per archi e sia $\omega \in C^0(\Omega, (\mathbb{R}^n)^*)$. Allora se ω è esatta con primitive $P_1, P_2 \in C^1(\Omega)$

$$\exists c \in \mathbb{R}^n : \forall x \in \Omega, P_2 - P_1 = c.$$

Equivalentemente, se $F \in C^0(\Omega, \mathbb{R}^b)$ è un campo conservativo con potenziali $V_1, V_2 \in C^1(\Omega)$

$$\exists c' \in \mathbb{R}^n : \forall x \in \Omega, V_2 - V_1 = c'.$$

Dimostrazione. Osserviamo che se $\omega = dP_1 = dP_2 \implies d(P_1 - P_2) = 0 \iff \nabla(P_1 - P_2) = 0 \implies P_1 - P_2 = c \in \mathbb{R} \forall x \in \Omega$ per il lemma precedente. Analogamente si può mostrare lo stesso per i campi vettoriali. \square

Oss.: In virtù della precedente proposizione, se $\Omega \subseteq \mathbb{R}^n$ è un insieme aperto e connesso per archi, allora abbiamo che le primitive di una 1-forma differenziali o i potenziali di un campo vettoriale sono univocamente determinati a meno di una costante additiva.

Oss.: Abbiamo visto che se Ω è un insieme aperto, allora

$$\omega \text{ esatta} \implies \omega \text{ chiusa}$$

Ma vale anche il viceversa?

La risposta è no, siccome basta considerare i seguenti insiemi $\Omega = \mathbb{R}^2 \setminus (0, 0)$ e $H = \mathbb{R}^2 \setminus \{(x, 0) :$

$x \in \mathbb{R}$ } e osservare che la 1-forma differenziale $\eta_0 \in C^\infty(\Omega, \mathbb{R}^2)$ definita come $\eta_0 = -\frac{y}{x^2+y^2}dx + \frac{x}{x^2+y^2}dy$ è chiusa in Ω e su H , infatti

$$\begin{aligned}\frac{\partial}{\partial y} \left(-\frac{y}{x^2+y^2} \right) &= -\frac{y^2-x^2}{(x^2+y^2)^2}, \quad \Rightarrow \quad \frac{\partial}{\partial x} \left(\frac{x}{x^2+y^2} \right) = \frac{\partial}{\partial y} \left(-\frac{y}{x^2+y^2} \right). \\ \frac{\partial}{\partial x} \left(\frac{x}{x^2+y^2} \right) &= \frac{y^2-x^2}{(x^2+y^2)^2},\end{aligned}$$

Vediamo se è esatta: osserviamo che, per le proposizioni viste in precedenza, per mostrare che ω non è esatta basta individuare un cammino chiuso per cui l'integrale curvilineo valutato sulla curva non è zero. In questo caso è abbastanza semplice da individuare², infatti basta considerare $\gamma(t) = (\cos t, \sin t)$ con $t \in [0, 2\pi]$ e notare che

$$\oint_{\gamma} \eta_0 = \int_0^{2\pi} ((-\sin t)(-\sin t) + \cos t \cos t) dt = \int_0^{2\pi} 1 dt = 2\pi \neq 0 \implies \eta_0 \text{ non è esatta}$$

Studiamo invece η_0 su altri insiemi come $\mathbb{R}^+ \times \mathbb{R}$. Prendiamo la funzione $\theta(x, y) = \arctan \frac{y}{x}$ detta *funzione argomento*. Osserviamo che

$$\begin{aligned}d\theta(x, y) &= \partial_x \arctan \left(\frac{y}{x} \right) dx + \partial_y \arctan \left(\frac{y}{x} \right) dy = -\frac{\frac{y}{x^2}}{1 + \frac{y^2}{x^2}} dx + \frac{\frac{1}{x}}{1 + \frac{y^2}{x^2}} dy = \eta_0 \implies \\ &\implies \theta(x, y) \text{ è una primitiva di } \eta_0 \text{ sull'aperto } \mathbb{R}^+ \times \mathbb{R}\end{aligned}$$

Per quanto riguarda la *funzione argomento* possiamo pensare di estenderla anche a H nella seguente maniera

$$\theta(x, y) = \begin{cases} \arccos \frac{x}{\sqrt{x^2+y^2}} & \text{se } y > 0 \\ -\pi & \text{se } \begin{cases} y = 0 \\ x < 0 \end{cases} \\ 2\pi - \arccos \frac{x}{\sqrt{x^2+y^2}} & \text{se } y < 0 \end{cases}$$

e possiamo verificare che $\theta \in C^\infty(H)$. Inoltre, calcolando le derivate di θ su H , vediamo che il differenziale di θ coincide con la 1-forma differenziale η_0 .

5.2.2 Omotopia e curve omotope

Vogliamo adesso studiare quali sono le condizioni per cui una 1-forma differenziale chiusa (o Equivalentemente un campo vettoriale irrotazionale è un campo conservativo). Per fare questo dobbiamo definire il concetto di omotopia, senza il quale non sarebbe possibile dare la nozione di insieme semplicemente connesso che ci darà una condizione necessaria e sufficiente affinché una 1-forma differenziale chiusa sia anche esatta

Definizione 5.13 (omotopia). Sia $E \subseteq \mathbb{R}^n$ e sia $\gamma \in C^0([a, b], E)$ e sia $\gamma(a) = \gamma(b)$, ovvero γ è chiusa. Diremo che γ è omotopa a $x_0 \in E$ in E se $\exists H : [0, 1] \times [a, b] \rightarrow E$ tale che

- (1) H è continua;
- (2) $H(0, \cdot) = \gamma$;

²il denominatore è della forma x^2+y^2 facendoci pensare di usare una curva con $x = \cos t$ e $y = \sin t$ con $t \in [0, 2\pi]$

$$\textcircled{3} \quad H(1, \cdot) = x_0$$

$$\textcircled{4} \quad H(t, a) = H(t, b) \quad \forall t \in [0, 1]$$

e chiameremo H con le proprietà $\textcircled{1}$, $\textcircled{2}$, $\textcircled{3}$ e $\textcircled{4}$ un'omotopia tra γ e x_0 relativamente ad E .

Il concetto di omotopia generalizza il concetto di *deformare* con continuità una curva. Nel nostro caso consideriamo una curva γ chiusa che andiamo a *strizzare* in maniera continua (ovvero senza strappi) fino a farla richiedere su un punto x_0

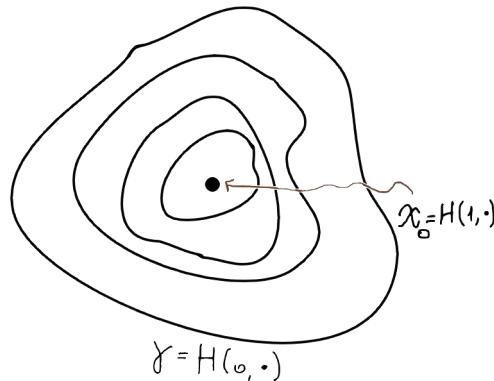


Figura 5.1: Rappresentazione di un'omotopia: la curva γ viene strizzata fino a diventare il punto x_0 .

Definizione 5.14 (insieme semplicemente connesso). Diremo che $E \subseteq \mathbb{R}^n$ è un insieme semplicemente connesso se

$\textcircled{1}$ è connesso per archi;

$\textcircled{2}$ $\forall \gamma \in C^0([a, b], E)$ chiusa ha un punto $x_0 \in E$ a cui è omotopa in E .

Oss.: Un esempio di insieme semplicemente connesso è $\Omega = \mathbb{R}^2 \setminus \{(x, y) \in \mathbb{R}^2 : y = 0, x > 0\}$: questo perché è possibile deformare con continuità la curva chiusa γ affinché scavalchi la parte dell'insieme esclusa. Un insieme non semplicemente connesso è $\mathbb{R}^2 \setminus \{(0, 0)\}$ siccome prendendo una curva sufficientemente vicina all'origine, per esempio, e un raggio sufficientemente grande la curva sarebbe costretta a passare per l'origine che è però esclusa. Per $n > 2$ avremo che $\mathbb{R}^n \setminus \{(0, 0)\}$ è semplicemente connesso siccome è possibile deformare la curva, sfruttando le dimensioni "in più" rispetto ad \mathbb{R}^2 , per scavalcare l'origine.

Oss.: Sia $\Omega = \mathbb{R} \setminus S$ e $S = q + \mathbb{R}v$. $S = \{q + tv : t \in \mathbb{R}\}$, $q, v \in \mathbb{R}^3$, ovvero lo spazio tridimensionale senza una retta. Allora Ω non è semplicemente connesso, siccome prendendo una circonferenza sufficientemente vicina alla retta e con un raggio maggiore della distanza dal punto alla retta, la curva sarebbe costretta a "toccare" la retta per poter essere omotopa ad un punto, pertanto non è un insieme semplicemente connesso.

Proposizione 5.5 (gli insiemi convessi sono semplicemente connessi). Sia $E \subseteq \mathbb{R}^n$, allora E è semplicemente connesso

Dimostrazione. Se E è convesso e $\gamma : [a, b] \rightarrow E$ è chiusa e continua scelgo $x_0 \in E$ e osservo che, preso

$$H(s, t) = sx_0 + \gamma(t)(1 - s) \quad 0 \leq s \leq 1,$$

per la convessità di E avremo che $H(s, t) \in E$. Avremo per $s = 0$ che $H(0, t) = \gamma(t)$ e per $s = 1$ che $H(1, t) = x_0$. Siccome H è continua, avremo che H è un'omotopia in E tra γ e x_0 . \square

Teorema 5.6 (lemma di Poincaré per gli aperti semplicemente connessi). Sia $\Omega \subseteq \mathbb{R}^n$ aperto e semplicemente connesso e sia $\omega \in C^1(\Omega, (\mathbb{R}^n)^*)$. Allora

$$\omega \text{ è chiusa} \iff \omega \text{ è esatta}.$$

Equivalentemente, se $F \in C^1(\Omega, \mathbb{R}^n)$ allora

$$F \text{ è irrotazionale} \iff F \text{ è conservativo}.$$

La dimostrazione proposta per l'implicazione verso destra (\Rightarrow) riporta qua sotto vale nel caso di insiemi stellati rispetto a x_0 , ovvero insiemi in cui esiste almeno un punto x_0 per cui ogni segmento fra x_0 e x appartenente all'insieme è contenuto in esso. Tuttavia sebbene insieme stellato \Rightarrow insieme semplicemente connesso, non vale il viceversa. Nelle appendici riporterò la dimostrazione per insiemi semplicemente connessi, dove necessito di usare alcuni strumenti della geometria differenziale.

Dimostrazione. la dimostrazione verrà fatta, senza perdere di generalità, nel caso delle 1-forme differenziali.

\Leftarrow : già dimostrato nella proposizione 5.3

\Rightarrow : fissiamo $x_0 \in \Omega$ e, preso $x \in \Omega$, consideriamo la funzione

$$F(x) = \int_{\gamma} \omega$$

dove $\gamma(t) : [0, 1] \mapsto \Omega$ è una curva con $\gamma(t) = (1-t)x_0 + tx$. Per definizione di integrale curvilineo di seconda specie avremo che

$$\int_0^1 \omega(\gamma(t)) \dot{\gamma}(t) dt = \int_0^1 \omega((1-t)x_0 + tx)(x - x_0) dt$$

ma possiamo derivare sotto il segno di integrale (la derivata di $\omega \circ \gamma$ esiste ed è continua siccome composizione di funzioni derivabili lungo una direzione x_j):

$$\begin{aligned} \partial_{x_j} F(x) &= \int_0^1 \partial_{x_j} \left(\sum_{i=1}^n a_i((1-t)x_0 + tx)(x_i - x_{0_i}) \right) dt = \int_0^1 \left(a_j((1-t)x_0 + tx) + \right. \\ &\quad \left. + \sum_{i=1}^n t \frac{\partial a_i}{\partial x_j} ((1-t)x_0 + tx)(x_i - x_{0_i}) \right) dt \end{aligned}$$

adesso osserviamo che la chiusura della 1-forma differenziale implica che $\forall 1 \leq i < j \leq n$

$$\frac{\partial a_i}{\partial x_j} = \frac{\partial a_j}{\partial x_i}.$$

Inoltre osserviamo che, tramite il teorema del differenziale della composizione, abbiamo che

$$\frac{\partial}{\partial t} a_j((1-t)x_0 + tx) = Da_j((1-t)x_0 + tx) D((1-t)x_0 + tx)(t) = \sum_{i=1}^n \frac{\partial a_j}{\partial x_i} (x_i - x_{0_i})$$

dunque, effettuando questa sostituzione nell'integrale di sopra, otterremo che

$$\begin{aligned}\partial_{x_j} F(x) &= \int_0^1 \partial_{x_j} \left(a_j((1-t)x_0 + tx) + \sum_{i=1}^n t \frac{\partial a_j}{\partial x_i} (((1-t)x_0 + tx)(x_i - x_{0i})) \right) dt = \\ &= \int_0^1 \left(a_j((1-t)x_0 + tx) + t \frac{\partial}{\partial t} [a_j((1-t)x_0 + tx)] \right) dt = \int_0^1 \frac{\partial}{\partial t} (ta_j((1-t)x_0 + tx)) dt = \\ &= a_j(x)\end{aligned}$$

dunque $dF = \omega$, il che conclude la dimostrazione. \square

5.2.3 Campi vettoriali radiali

Definizione 5.15 (campo vettoriale radiale). Consideriamo $h : (0, +\infty) \mapsto \mathbb{R}$ continua e $F : \mathbb{R}^n \setminus \{0\} \rightarrow \mathbb{R}^n$ con $F(x) = h(|x|)x$. Diremo allora che F è un campo vettoriale radiale.

Oss.: Siccome $\nabla|x| = \frac{x}{|x|} \forall x \neq 0$ allora $F(x) = |x|h(|x|)\frac{x}{|x|} \implies F(x) = |x|h(|x|)\nabla|x|$

Proposizione 5.6 (i campi radiali sono conservativi). Sia F un campo vettoriale radiale, allora F è conservativo

Dimostrazione. Possiamo pensare di definire

$$U(s) = \int_1^s th(t)dt \text{ con } s > 0$$

e posto $V(x) = U(|x|) \implies$

$$\implies \nabla V(x) = \sum_{j=1}^n \frac{\partial}{\partial x_j} U(|x|)e_j$$

ma osserviamo che $\frac{\partial U(|x|)}{\partial x_j} = \frac{\partial U(|x|)}{\partial |x|} \frac{\partial |x|}{\partial x_j} = U'(|x|)\frac{x_j}{|x|}$ dove osserviamo che $\frac{\partial |x|}{\partial x_j} = \frac{1}{\sqrt{\sum_{i=1}^n x_i^2}} \frac{\partial x_j}{\sqrt{\sum_{i=1}^n x_i^2}}$

$$\sum_{j=1}^n U'(|x|)\frac{x_j}{|x|}e_j = \frac{U'(|x|)}{|x|} \sum_{j=1}^n x_j e_j = \frac{U'(|x|)}{|x|}|x| = U'(|x|) = |x|h(|x|)$$

dove si è usato il teorema fondamentale del calcolo per calcolare la derivata di $U(|x|)$ rispetto a $|x|$ siccome $\frac{\partial}{\partial |x|} = \frac{\partial}{\partial |x|} \int_1^{|x|} th(t)dt = th(t)|_{t=|x|} = |x|h(|x|)$ \square

Oss.: Osserviamo che per $n = 2$ il nostro campo vettoriale è definito su $\mathbb{R}^2 \setminus \{(0, 0)\}$, che ricordiamo essere un insieme non semplicemente connesso. Detto questo, non è detto che non vi siano delle forme differenziali esatte: un esempio è il campo gravitazionale che è un campo conservativo.

Similmente, possiamo definire le 1-forme differenziali radiali nella seguente maniera

Definizione 5.16 (1-forma differenziale radiale). Sia $h(x) \in C^0((0, +\infty), \mathbb{R})$ e sia $\omega \in C^\infty(\mathbb{R}^n \setminus \{0\}, (\mathbb{R}^n)^*)$ è una 1-forma differenziale radiale se

$$\omega(x) = h(|x|) \sum_{j=1}^n x_j dx_j$$

Come prima, avremo che è esatta con primitiva $P(x) = U(|x|)$, dove $U(s) = \int_1^s th(t)dt$ e $dP(x) = \omega(x)$.

CAPITOLO 6

Cenni alla misura di Lebesgue

In questo capitolo affronteremo il concetto di misura e daremo alcuni accenni all'impostazione teorica data da Lebesgue nello sviluppo del suo concetto di misura. Alcuni approfondimenti inseriti sono stati presi da [3] oppure [7]. Per una trattazione completa e, naturalmente, rigorosa consiglio un testo molto famoso della materia, ovvero [8].

6.1 La misura esterna

Vogliamo adesso determinare una funzione di insiemi $m_n : \mathcal{P}(\mathbb{R}^n) \rightarrow [0; +\infty]$ che rappresenta il nostro concetto intuitivo di *volume* dell'insieme, detto *misura dell'insieme*.

L'intuizione ci porta a pensare che questa funzione goda delle seguenti proprietà:

- ① $m_n(\emptyset) = 0$
- ② Se $A, B \subseteq X$ tali che $A \subseteq B \implies m_n(A) \leq m_n(B)$
- ③ Se $A, B \subseteq X$ tali che $A \cap B = \emptyset \implies m_n(A \cup B) = m_n(A) + m_n(B)$

Prima di procedere ulteriormente sullo sviluppo della teoria della misura di Lebesgue, vorrei aprire una breve parentesi "qualitativa" su come trovare la misura di un insieme: l'approccio più *naive* possibile consiste nel coprire il nostro insieme con dei rettangoli e sommare i volumi di questi rettangoli. Questo approccio ci consente di ottenere quello che viene chiamata *misura esterna*, ovvero una funzione d'insieme che è ben definita su ogni insieme e che soddisfa tutte le buone proprietà che ci aspettiamo da una misura che generalizza il nostro concetto di misura tranne quella di additività della misura. Per essere una misura "vera e propria" (ovvero che goda anche di additività) si può mostrare che è necessario restringerci ad una classe di insiemi (che saranno gli insiemi *misurabili*) su cui la nostra misura esterna godrà di additività (questo risultato prende il nome di *teorema di Caratheodory*, di cui non daremo dimostrazione).

Definizione 6.1 (misura di un intervallo e di un n -intervallo). Per ogni intervallo $I \subseteq \mathbb{R}$ limitato definiamo

$$l(I) = \sup I - \inf I.$$

Sia adesso $I_1 \times \dots \times I_n \subseteq \mathbb{R}^n$ definiamo

$$v_n \left(\bigtimes_{i=1}^n I_i \right) = \prod_{i=1}^n l(I_i)$$

Definizione 6.2 (misura di Lebesgue di un insieme E). Sia $E \subseteq \mathbb{R}^n$, definiamo la misura di Lebesgue di E come

$$m_n(E) = \inf \left\{ \sum_{i=1}^{+\infty} v_n(J_i) : E \subseteq \bigcup_{i=1}^{+\infty} J_i \text{ e } J_i \subseteq \mathbb{R}^n \text{ è un } n\text{-intervallo} \right\} \quad (6.1)$$

Definizione 6.3. Diremo che $\mu^* : \mathcal{P}(X) \mapsto [0; +\infty]$ è una misura esterna se

$$\textcircled{1} \quad \mu^*(\emptyset) = 0;$$

$$\textcircled{2} \quad \mu^*(E) \leq \sum_{i=1}^{+\infty} \mu^*(E_i) \text{ se } E \subseteq \bigcup_{i=1}^{+\infty} E_i$$

Proposizione 6.1 (monotonia della misura esterna). Se μ^* è una misura esterna allora

$$A \subseteq B \implies \mu^*(A) \leq \mu^*(B)$$

Dimostrazione. Consideriamo $E = A$ e il seguente insieme

$$E_j = \begin{cases} B & j = 1 \\ \emptyset & j \geq 2 \end{cases}$$

osserviamo allora che $E \subseteq \bigcup_{i=1}^{+\infty} E_i$ allora

$$\mu^*(E) = \mu^*(A) \stackrel{\text{per la } \textcircled{2}}{\leq} \sum_{j=1}^{+\infty} \mu^*(E_j) = \mu^*(E_1) = \mu^*(B)$$

ovvero

$$\mu^*(A) \leq \mu^*(B)$$

□

Proposizione 6.2 (subadditività finita). Le misure esterne sono finitamente subadditive, ovvero preso E_1, \dots, E_k avremo che

$$\mu^* \left(\bigcup_{i=1}^k E_i \right) \leq \sum_{i=1}^k \mu^*(E_i)$$

Dimostrazione. Siano $E_1, \dots, E_k \subseteq \mathbb{R}^n$. Imponendo che $\forall j > k, E_j = \emptyset$, allora

$$\forall E \subseteq \bigcup_{i=1}^{+\infty} E_i, \mu^*(E) \leq \sum_{i=1}^{+\infty} \mu^*(E_i) = \sum_{i=1}^k \mu^*(E_i) + \sum_{i=k+1}^{+\infty} \mu^*(\emptyset) = \sum_{i=1}^k \mu^*(E_i) \implies \mu^*(E) \leq \sum_{i=1}^k \mu^*(E_i)$$

ma allora, siccome abbiamo banalmente che $\bigcup_{i=1}^k E_i \subseteq \bigcup_{i=1}^{+\infty} E_i$, allora

$$\mu^* \left(\bigcup_{i=1}^k E_i \right) \leq \sum_{i=1}^k \mu^*(E_i)$$

□

Vogliamo adesso mostrare che la funzione d'insieme m_n da noi definita è effettivamente una misura esterna, dunque, in virtù di quanto detto, possiamo individuare e restringere m_n ad una classe di insiemi per cui m_n rappresenta una vera e propria misura.

Proposizione 6.3 (misura di Lebesgue è una misura esterna). $m_n : \mathcal{P}(X) \rightarrow [0; +\infty]$ è una misura esterna

Dimostrazione. La dimostrazione consiste nel mostrare che $m_n : \mathcal{P}(X) \mapsto [0; +\infty]$ gode delle proprietà della misura esterna.

La ① segue banalmente, siccome l' n -intervallo nullo è un ricoprimento dell'insieme nullo, dunque $m_n(\emptyset) = 0$. Per la ②, supponiamo naturalmente che $\sum_{i=1}^{+\infty} m_n(E_i) < +\infty$ (altrimenti è banale) e prendiamo un insieme $E \subseteq \bigcup_{i=1}^{+\infty} E_i$, osservando che $\forall i \in \mathbb{N}, E_i \subseteq \bigcup_{k=1}^{+\infty} I_k^{(i)}$, pertanto

$$E \subseteq \bigcup_{i=1}^{+\infty} \bigcup_{k=1}^{+\infty} I_k^{(i)}$$

e questo vale per ogni ricoprimento di $\bigcup_{i=1}^{+\infty} E_i$, dunque anche per quello che "fornisce" la misura di $\bigcup_{i=1}^{+\infty} E_i$

$$m_n(E) \leq \inf \left\{ \sum_{i=1}^{+\infty} l(E_i) \right\} = m_n \left(\bigcup_{i=1}^{+\infty} E_i \right)$$

resta da mostrare che $m_n \left(\bigcup_{i=1}^{+\infty} E_i \right) \leq \sum_{i=1}^{+\infty} m_n(E_i)$. Questo può essere fatto fissando $\varepsilon > 0$ e osservando che $\forall i \in \mathbb{N}, \exists \bigcup_{k=1}^{+\infty} I_i^{(k)} \supseteq E_i : \sum_{k=1}^{+\infty} l(I_i^{(k)}) < m_n(E_i) + \frac{\varepsilon}{2^{i+1}}$, dunque

$$m_n \left(\bigcup_{i=1}^{+\infty} E_i \right) \leq \sum_{i=1}^{+\infty} \sum_{k=1}^{+\infty} l(I_i^{(k)}) < \sum_{i=1}^{+\infty} \left(m_n(E_i) + \frac{\varepsilon}{2^{i+1}} \right) = \varepsilon + \sum_{i=1}^{+\infty} m_n(E_i)$$

e, siccome ε era arbitrario, otteniamo la tesi facendo il limite. \square

Definizione 6.4 (insieme misurabile secondo Lebesgue). Diremo che $A \subseteq X$ è misurabile secondo Lebesgue se

$$\forall S \subseteq X \implies m_n(S) = m_n(A \cap S) + m_n(S \setminus A)$$

Per ogni insieme X definiamo con $\mathfrak{M}(X)$ la classe di tutti gli insiemi misurabili secondo Lebesgue.

Definizione 6.5 (σ -algebra). Diremo che $\mathcal{F} \subseteq \mathcal{P}(X)$ è una σ -algebra se valgono

$$\textcircled{1} \quad \emptyset, X \in \mathcal{F};$$

$$\textcircled{2} \quad A \in \mathcal{F} \implies (X \setminus A) \in \mathcal{F}$$

$$\textcircled{3} \quad \{A_j : j \geq 1\} \subseteq \mathcal{F} \implies \bigcup_{i=1}^{+\infty} A_i \in \mathcal{F}$$

Definizione 6.6 (misura). Diremo che $\mu : \mathcal{F} \mapsto [0; +\infty]$ è una misura se

- ① \mathcal{F} è una σ -algebra;
- ② $\mu(\emptyset) = 0$;
- ③ Se $(\{A_i : i \geq 1\} \subseteq \mathcal{F}) \wedge (\forall i \neq j, A_i \cap A_j = \emptyset)$ allora

$$\mu\left(\bigcup_{i=1}^{+\infty} A_i\right) = \sum_{i=1}^{+\infty} \mu(A_i)$$

Teorema 6.1. La classe dei misurabili $\mathfrak{M}(X)$ è una σ -algebra e $m_n : \mathfrak{M}(X) \rightarrow [0; +\infty]$ è una misura

Diremo allora che $m_n : \mathfrak{M}(X) \rightarrow [0; +\infty]$ è la misura di Lebesgue su X . Per lo più noi lavoreremo supponendo che $X = \mathbb{R}^n$.

Proposizione 6.4. La misura $m_n : \mathfrak{M}(X) \rightarrow [0; +\infty]$ è finitamente additiva

Dimostrazione. La dimostrazione è equivalente a quella della proposizione 6.2 sostituendo opportunamente alle minorazione le egualanze. \square

Oss. Ammettendo l'assioma della scelta è possibile mostrare che esistono degli insiemi che non sono misurabili (come ampiamente detto in precedenza). L'esempio più famoso è l'insieme di Vitali, che si può costruire dentro un qualunque insieme di misura positiva $E \in \mathfrak{M}(\mathbb{R}^n)$. Dalla definizione di misurabilità allora $\exists S_0 \subseteq \mathbb{R}^n$ tale che

$$m_n(S_0) \leq m_n(S_0 \cap E) + m_n(S_0 \setminus E)$$

dunque si perderebbe la σ -additività su insiemi disgiunti.

6.1.1 Insiemi trascurabili

Definizione 6.7 (insiemi trascurabili). Diremo che $Z \subseteq \mathbb{R}^n$ ha misura nulla oppure che è trascurabile se $m_n(Z) = 0$

Esercizio 6.1. Mostrare che un punto $x \in \mathbb{R}^n$ ha misura nulla

Svolgimento. Sia $x \in \mathbb{R}^n$ e consideriamo un n -intervallo $\bigtimes_{i=1}^n I_i$ dove $I_i = (x_i - \frac{\varepsilon}{2}, x_i + \frac{\varepsilon}{2})$ (x_i rappresenta la i -esima componente del punto x), dunque rappresenta un cubo di volume pari a ε^n per $i \in \{1, \dots, n\}$ con $\varepsilon > 0$. Allora possiamo osservare che

$$\{x\} \subseteq \bigcup_{i=1}^{+\infty} J_i$$

dove

$$J_i = \begin{cases} \bigtimes_{i=1}^n (x_i - \frac{\varepsilon}{2}, x_i + \frac{\varepsilon}{2}) \\ \emptyset \end{cases}$$

ma allora

$$m_n(\{x\}) \subseteq \bigcup_{i=1}^{+\infty} v_n(J_i) = \varepsilon^n$$

ma allora, siccome $\varepsilon > 0$ è un valore arbitrario, possiamo farne il limite per $\varepsilon \rightarrow 0$, dunque $m_n(\{x\}) = 0$. \square

Proposizione 6.5. Se $Z \subseteq \mathbb{R}^n$ è numerabile, allora

$$m_n(Z) = 0$$

Oss.: Non tutti gli insiemi a misura nulla sono numerabili: un esempio è l'insieme di Cantor.

Dimostrazione. Se Z è numerabile allora

$$Z = \bigcup_{i=1}^{+\infty} \{x_i\}$$

dove $x_i : \mathbb{N} \rightarrow Z$ dunque

$$m_n(Z) \leq \sum_{i=1}^{+\infty} m_n(\{x_i\}) = 0 \implies m_n(Z) = 0$$

□

Corollario 6.1.1. \mathbb{Q}^n è un insieme trascurabile

Dimostrazione. $\mathbb{Q}^n \cong \mathbb{N}$ dunque $m_n(\mathbb{Q}^n) = 0$

□

Corollario 6.1.2. Gli insiemi a misura nulla sono misurabili per Lebesgue.

Dimostrazione. In virtù della definizione 6.4 dobbiamo verificare che, preso Z insieme trascurabile, allora

$$\forall S \subseteq \mathbb{R}^n, m_n(S) = m_n(S \setminus Z) + m_n(S \cap Z)$$

Ora osserviamo che, siccome $S = (S \setminus Z) \cup (S \cap Z)$, avremo per la subadditività finita (di cui gode m_n siccome nella proposizione 6.3 abbiamo mostrare che è una misura esterna) che

$$m_n(S) \leq m_n(S \setminus Z) + m_n(S \cap Z)$$

Per la disegualanza opposta osserviamo che

$$m_n(S) \geq m_n(S \setminus Z) \text{ e } m_n(Z) \geq m_n(S \cap Z)$$

per monotonia della misura esterna (proposizione 6.1), ma siccome $0 = m_n(Z) \geq m_n(S \cap Z) \implies m_n(S \cap Z) = 0$ allora

$$m_n(S) \geq m_n(S \setminus Z) = m_n(S \setminus Z) + m_n(S \cap Z) \implies m_n(S) \geq m_n(S \setminus Z) + m_n(S \cap Z)$$

dunque $m_n(S) = m_n(S \setminus Z) + m_n(S \cap Z)$, ovvero la tesi.

□

Teorema 6.2 (gli aperti sono misurabili). Tutti gli aperti di \mathbb{R}^n sono misurabili secondo Lebesgue

Per dimostrare questo teorema abbiamo bisogno del seguente lemma:

Lemma 6.2.1. Ogni aperto di \mathbb{R}^n è unione numerabile di n -intervalli a due a due disgiunti

Dimostrazione. Sia $A \subseteq \mathbb{R}^n$ aperto non vuoto. Allora possiamo considerare la famiglia degli n -intervalli $I \subseteq A$ ad estremi razionali, ovvero:

$$\mathcal{V} = \left\{ I = \prod_{i=1}^n (a_i, b_i] : I \subseteq A, a_i, b_i \in \mathbb{Q} \quad \forall i \in \{1, \dots, n\} \right\}$$

Siccome ogni n -intervallo all'interno di \mathcal{V} può essere messo in corrispondenza biunivoca con la $2n$ -upla $(a_1, \dots, a_n, b_1, \dots, b_n) \in \mathbb{Q}^{2n}$ che è un insieme numerabile, allora anche \mathcal{V} è un insieme numerabile. E' banale l'inclusione $\bigcup_{I \in A} I \subseteq A$ (per definizione di \mathcal{V}), vogliamo adesso mostrare che vale anche l'inclusione opposta: sia $x \in A$ e, siccome A è aperto, avremo che

$$\forall y \in A, \exists \varepsilon > 0 : \{y = (y_1, \dots, y_n) \in \mathbb{R}^n : |y_i - x_i| < \varepsilon \ \forall i \in \{1, \dots, n\}\} \subseteq A$$

Siccome \mathbb{Q} è denso in \mathbb{R} avremo che $\forall i \in \{1, \dots, n\}, \exists \alpha_i, \beta_i \in \mathbb{Q}$ tali che

$$x_i - \varepsilon < \alpha_i < x_i < \beta_i < x_i + \varepsilon$$

Dunque avremo che l' n -intervallo $I' = \prod_{i=1}^n (\alpha_i, \beta_i]$ contiene il punto x , dunque $x \in \mathcal{V} \implies x \in \bigcup_{I \in \mathcal{V}} I \implies A \subseteq \bigcup_{I \in \mathcal{V}} I$. Concludiamo che $A = \bigcup_{I \in \mathcal{V}} I$.

Mostriamo adesso che possiamo prendere gli n -intervalli a due a due disgiunti. Consideriamo $A = \bigcup_{i=i}^{+\infty} I_i$ ma allora, posto $J_i = I_i \setminus \bigcup_{j=1}^{i-1} I_j$, possiamo vedere facilmente che

$$\bigcup_{i=1}^{+\infty} I_i = \bigcup_{i=1}^{+\infty} J_i = A$$

□

Siamo pronti adesso a dimostrare il teorema:

Dimostrazione (del teorema). Sia A un insieme aperto. Per il lemma precedente sappiamo che $A = \bigcup_{k=1}^{+\infty} I_i$ dove I_i è un n -intervallo, dunque $m_n(P \setminus A) = m_n(\emptyset) = 0$. Pertanto $P \setminus A \in \mathfrak{M}(\mathbb{R}^n)$ e $P \in \mathfrak{M}(\mathbb{R}^n)$ siccome unione di n -intervalli. □

Corollario 6.2.1. Tutti gli insiemi chiusi di \mathbb{R}^n sono misurabili secondo Lebesgue

Dimostrazione (del corollario). Siccome $m_n : \mathfrak{M}(\mathbb{R}^n) \rightarrow [0; +\infty]$ allora $\mathfrak{M}(\mathbb{R}^n)$ è una σ -algebra, dunque se $A \in \mathfrak{M}(\mathbb{R}^n) \implies A^c \in \mathfrak{M}(\mathbb{R}^n)$. □

Corollario 6.2.2. Tutte le unioni e le intersezioni numerabili di insiemi chiusi o aperti sono ancora misurabili secondo Lebesgue

Dimostrazione. Se $\{B_i : i \geq 1\} \subseteq \mathfrak{M}(\mathbb{R}^n) \implies \bigcup_{i=1}^{+\infty} B_i \in \mathfrak{M}(\mathbb{R}^n)$ perché $\mathfrak{M}(\mathbb{R}^n)$ è una σ -algebra, mentre per le intersezioni numerabili abbiamo che se $B_i \in \mathfrak{M}(\mathbb{R}^n) \implies B_i^c \in \mathfrak{M}(\mathbb{R}^n)$ (sempre per le proprietà delle σ -algebre) e, per le leggi di De Morgan, avremo che $\left(\bigcup_{i=1}^{+\infty} B_i^c\right)^c = \bigcap_{i=1}^{+\infty} B_i \in \mathfrak{M}(\mathbb{R}^n)$ □

6.2 Funzioni misurabili secondo Lebesgue

Andiamo adesso ad affrontare il concetto di funzione misurabile, il quale ci condurrà, naturalmente, verso il concetto di integrale di Lebesgue

Definizione 6.8 (funzioni misurabili secondo Lebesgue). Sia $E \in \mathfrak{M}(\mathbb{R}^n)$ misurabile e $f : E \rightarrow \bar{\mathbb{R}}$. Diremo che f è misurabile secondo Lebesgue se $\forall t \in \mathbb{R}$ l'insieme $f^{-1}([-∞, t)) \in \mathfrak{M}(\mathbb{R}^n)$, ovvero esso è misurabile.

Oss.: (**Topologia di $[-\infty, \infty]$**) Ogni aperto di $[-\infty, +\infty]$ è un'unione arbitraria di intervalli che possono essere sia aperti oppure del tipo $[-\infty, t)$ o $(t, +\infty]$.

Proposizione 6.6 (caratterizzazione delle funzioni misurabili). $E \in \mathfrak{M}(\mathbb{R}^n), f : E \rightarrow \bar{\mathbb{R}}$ è misurabile se e solo se $\forall O \subseteq \bar{\mathbb{R}}$ aperto di $[-\infty, +\infty]$ abbiamo che $f^{-1}(O) \in M(\mathbb{R}^n)$

Proposizione 6.7 (le funzioni continue sono misurabili). Se $E \in \mathcal{R}^n, f : E \rightarrow \mathbb{R}$ è continua, allora è misurabile

Dimostrazione. Se $O \subseteq [-\infty, +\infty]$ è un aperto contenuto in \mathbb{R} in realtà è un aperto di \mathbb{R} $\implies f^{-1}(O)$ è aperto in E in virtù del teorema 5.4, dunque avremo che $\exists \Omega \subseteq \mathbb{R}^n$ tale che

$$f^{-1}(O) = E \cap \Omega \in \mathfrak{M}(\mathbb{R}^n) \implies f \text{ è misurabile}$$

ovvero la tesi. \square

Proposizione 6.8. Se $f_n : E \rightarrow \bar{\mathbb{R}}$ è una successione di funzioni misurabili e $f : E \rightarrow \bar{\mathbb{R}}$. Se $f_n \xrightarrow{n \rightarrow +\infty} f \forall x \in E$ allora f è misurabile

Prendiamo in causa una funzione che causava molti problemi all'integrale di Riemann, ovvero la funzione di Dirichlet. Al momento, consideriamo una sua variante $\chi_k : [0, 1] \rightarrow \mathbb{R}$ definita come

$$\chi_k(x) = \begin{cases} 1 & \text{se } x \in \{q_0, \dots, q_k\} \\ 0 & \text{se } x \notin \{q_0, \dots, q_k\} \end{cases}$$

allora $\chi_k \xrightarrow{n \rightarrow +\infty} \chi$, ovvero la funzione di Dirichlet. Grazie alla precedente proposizione siamo in grado di concludere che χ è misurabile per Lebesgue, ma non per Riemann (sebbene $\forall k, \chi_k$ è Riemann-integrabile).

Esercizio 6.2. Mostrare la misurabilità di χ tramite la definizione di funzione misurabile secondo Lebesgue.

Svolgimento. Osserviamo che $\chi : [0, 1] \mapsto \{0, 1\}$. Mostriamo il seguente lemma

Lemma 6.2.2. Sia $f : A \mapsto B$ e presi $B_1, B_2 \subseteq B$ allora

$$f^{-1}(B_1 \cup B_2) = f^{-1}(B_1) \cup f^{-1}(B_2)$$

Dimostrazione. Mostriamo che $f^{-1}(B_1 \cup B_2) \subseteq f^{-1}(B_1) \cup f^{-1}(B_2)$: per definizione sappiamo che $f^{-1}(B_1 \cup B_2) = \{x \in A : f(x) \in B_1 \cup B_2\} \implies f(x) \in B_1 \vee f(x) \in B_2 \implies x \in f^{-1}(B_1) \vee x \in f^{-1}(B_2)$. Mostriamo che $f^{-1}(B_1 \cup B_2) \supseteq f^{-1}(B_1) \cup f^{-1}(B_2)$: supponiamo, per assurdo, che $\exists x \in f^{-1}(B_1) \cup f^{-1}(B_2) : x \notin f^{-1}(B_1 \cup B_2) \implies f(x) \in B_1 \vee f(x) \in B_2$ e $f(x) \notin B_1 \cup B_2$, il che è un assurdo. \square

Dunque, in virtù di questo lemma appena mostrato, avremo che

$$\chi^{-1}(\{0, 1\}) = \chi^{-1}(\{0\}) \cup \chi^{-1}(\{1\})$$

\square

e abbiamo che $\chi^{-1}(0) = \mathbb{Q} \cap [0, 1]$ che è un'unione numerabile di punti, ovvero i numeri razionali all'interno di questo intervallo che, in virtù del corollario 6.1.2, è misurabile.

L'intervallo restante, invece, sarebbe pari a $[0, 1] \setminus (\mathbb{Q} \cap [0, 1])$ che, tuttavia, è misurabile siccome $[0, 1]$ è misurabile, $\mathbb{Q} \cap [0, 1]$ è misurabile e le differenze sono misurabili secondo Lebesgue per le proprietà delle σ -algebre.

Proposizione 6.9. Se $f, g : E \rightarrow \bar{\mathbb{R}}$ sono misurabili, allora

- ① se $f + g$ è ben definita ovunque su E allora $f + g$ è misurabile;
- ② se $\lambda \in \mathbb{R}$ e λf è ben definita ovunque, allora λf è misurabile;
- ③ $E \mapsto \max\{f(x), g(x)\}$ e $E \mapsto \min\{f(x), g(x)\}$ sono misurabili;
- ④ se $f_j : E \rightarrow \mathbb{R}$ è misurabile $\forall j \in \mathbb{N} \implies x \in E \mapsto \sup_{j \in \mathbb{N}} f_j(x)$ è misurabile.

6.2.1 L'integrale di Lebesgue

Per una buona costruzione della misura di Lebesgue, è comodo definire le operazioni algebriche in $[0; +\infty]$. Definiamo il prodotto come

$$xy = \begin{cases} xy & \text{se } \max\{x, y\} < +\infty \\ +\infty & \text{se } \max\{x, y\} = +\infty \text{ e } \min\{x, y\} > 0 \\ 0 & \text{se } \max\{x, y\} = +\infty \text{ e } \min\{x, y\} = 0 \end{cases}$$

Il terzo caso, che sembra patologico, si può facilmente interpretare in maniera geometrica: avendo in mente la teoria dell'integrazione, il prodotto fra $0 \cdot +\infty$ corrisponde a quella di un rettangolo illimitato con base di lunghezza infinita e altezza nulla, dunque ha area "pari" a 0.

Introduciamo adesso la nozione di funzione semplice, concetto cardine per lo sviluppo dell'integrale di Lebesgue.

Definizione 6.9 (funzione semplice). $\varphi : E \rightarrow \mathbb{R}$ è una funzione semplice se $\varphi(E) = \{\lambda_1, \dots, \lambda_k\} \subset \mathbb{R}$ dove $\lambda_i \neq \lambda_j \ \forall i \neq j \in \{1, \dots, k\}$ dove $\varphi^{-1}(\lambda_i) = A_i \in \mathfrak{M}(\mathbb{R}^n) \ \forall i \in \{1, \dots, k\}$.

Definizione 6.10 (funzione caratteristica).

$$\mathbb{1}_{A_i} = \begin{cases} 1 & x \in A_i \\ 0 & x \notin A_i \end{cases}$$

Oss.: Segue banalmente che

$$\varphi(x) = \sum_{i=1}^k \lambda_i \mathbb{1}_{A_i}(x)$$

se φ è una funzione semplice, naturalmente per opportuni $\lambda_1, \dots, \lambda_k$ e $A_1, \dots, A_k \in \mathfrak{M}(\mathbb{R}^n)$.

Geometricamente non è difficile intuire che il volume sotto una funzione semplice (che assume valori pari a costanti che dipendono dalla regione di spazio) corrisponde ad un "solido" n -dimensionale con altezza λ_j con $j \in \{1, \dots, k\}$ e base pari alla "superficie" (naturalmente il termine usato qua è per rendere intuitiva la trattazione, ma non necessariamente è una superficie) del dominio. Nell'ipotesi di $\lambda_j \geq 0 \ \forall j \in \{1, \dots, k\}$ (per non avere problemi con volumi negativi) allora possiamo già definire il concetto di integrale

Definizione 6.11. Sia $\varphi : E \rightarrow \mathbb{R}$ una funzione semplice con $\lambda_j \geq 0 \ \forall j \in \{1, \dots, k\}$ definiamo il suo integrale come

$$I(\varphi) = \sum_{j=1}^n \lambda_j m_n(A_j) \tag{6.2}$$

Oss.: Si osservi che nella definizione 6.2 compare il prodotto $0 \cdot +\infty$ quando $m_n(A_i) = +\infty$ e $\lambda_i = 0$.

Definizione 6.12 (integrale di Lebesgue). Data $f : E \rightarrow [0; +\infty]$ misurabile allora

$$\int_E f(x)dx = \sup\{I(\varphi) : 0 \leq \varphi \leq f \text{ e } \varphi \text{ è semplice}\}$$

Introdurremo le seguenti notazioni per l'integrale di Lebesgue

$$\int_E f(x)dx = \int_E f(x_1, \dots, x_n)dx_1, \dots, dx_n = \int_E f = \int_E f(x)dm_n(x) = \int_E f dm_n$$

Oss.: Segue subito dalla definizione che se $f, g : E \rightarrow [0; +\infty]$ sono misurabili e $f \leq g$ su E allora

$$\int_E f \leq \int_E g$$

Oss.: Se $E \in \mathfrak{M}(\mathbb{R}^n) \implies \mathbb{1}_E$ è misurabile e vale

$$m_n(E) = \int_{\mathbb{R}^n} \mathbb{1}_E dm_n$$

Oss.: Se $\varphi : E \rightarrow [0; +\infty]$ è semplice allora $\int_E \varphi dm_n = I(\varphi) = \sum_{i=1}^k \lambda_i m_n(E_i)$

Proposizione 6.10. Sia $f : E \rightarrow \bar{\mathbb{R}}$ misurabile, allora anche $|f|$ è misurabile

Dimostrazione. $|f| = f^+ + f^-$ dove $f^+ = \max\{f, 0\}$ e $f^- = \max\{0, -f\}$. Sappiamo, dalla proposizione 6.9, che $\max\{f, 0\}$ e $\max\{-f, 0\}$ sono misurabili (dove abbiamo preso $g(x) = 0$) e la somma di due funzioni misurabili è misurabile. \square

Definizione 6.13 (integrabilità secondo Lebesgue). Diremo che f è integrabile secondo Lebesgue se

$$\int_E |f(x)|dx < +\infty$$

e, in tal caso, segue che

$$\int_E f(x)dx = \int_E f^+(x)dx - \int_E f^-(x)dx$$

Definizione 6.14 (proprietà quasi ovunque). Diremo che una proprietà o un'affermazione $P(x)$, con $x \in A$, vale quasi ovunque su A se l'insieme dei punti

$$Z = \{x \in A : P(x) \text{ è falsa}\}$$

ha misura nulla, ovvero $m_n(Z) = 0$

Proposizione 6.11. Se $f : E \rightarrow [0; +\infty]$ misurabile allora

$$\textcircled{1} \quad \int_E f dm_n = 0 \implies f = 0 \text{ quasi ovunque su } E;$$

$$\textcircled{2} \quad \int_E f dm_n < +\infty \implies f < +\infty \text{ quasi ovunque su } E$$

Oss.: La tesi del punto $\textcircled{1}$ ci dice che f è quasi ovunque nulla su E o equivalentemente che l'insieme $Z \subseteq E$ dei punti in cui f assume dei valori positivi ha misura nulla, dunque $m_n(Z) = 0$.

La tesi del punto $\textcircled{2}$ è analoga

Teorema 6.3. Siano $f, g : E \rightarrow \bar{\mathbb{R}}$ integrabili, ovvero $\int_E |f| dm_n < +\infty$ e $\int_E |g| dm_n < +\infty$, allora

(1) $\lambda, \tau \in \mathbb{R} \implies \lambda f + \tau g \in \mathbb{R}$ è quasi ovunque ben definita e sommabile. Inoltre $\int_E (\lambda f + \tau g) dm_n = \lambda \int_E f dm_n + \tau \int_E g dm_n$

(2) $|\int_E f dm_n| \leq \int_E |f| dm_n$

(3) $f \leq g$ quasi ovunque su $E \implies \int_E f \leq \int_E g$

Teorema 6.4 (integrale di Riemann 1-dimensionale con l'integrale di Lebesgue). Siano $-\infty \leq a < b \leq +\infty$, $I \subseteq \mathbb{R}$ intervallo tale che $a = \inf I$ e $b = \sup I$ e sia $f : I \rightarrow \mathbb{R}$ assolutamente integrabile secondo Riemann in senso generalizzato su I , ovvero

$$\int_a^b |f(t)| dt < +\infty$$

allora $f : I \rightarrow \mathbb{R}$ è integrabile (e quindi misurabile) secondo Lebesgue e

$$\int_a^b f(t) dt = \int_I f dm_n \quad (6.3)$$

Oss.: Osserviamo che l'equazione 6.3 è il primo procedimento operativo con cui calcolare l'integrale di Lebesgue

Oss.: Il teorema include anche il caso particolare in cui $I = [a, b]$.

Oss.: (Esistenza di funzioni integrabili secondo Riemann in senso generalizzato ma non secondo Lebesgue)

La funzione $f(x) = \frac{\sin x}{x}$ è integrabile (ma non assolutamente!) secondo Riemann in senso generalizzato nell'intervallo $[1, +\infty]$ ma non per Lebesgue.

Mostriamo che non è integrabile secondo Lebesgue mostrando che

$$\int_1^{+\infty} \left| \frac{\sin x}{x} \right| dx = +\infty$$

Osserviamo che possiamo "spaccare" il nostro integrale, per linearità rispetto agli ordini di integrazione, fra multipli interi di π :

$$\begin{aligned} \int_1^{+\infty} \left| \frac{\sin x}{x} \right| dx &= \int_1^\pi \frac{\sin x}{x} + \sum_{k=1}^{+\infty} \int_{k\pi}^{(k+1)\pi} \left| \frac{\sin x}{x} \right| dx \geq \sum_{k=1}^{+\infty} \int_{k\pi}^{(k+1)\pi} \left| \frac{\sin x}{(k+1)\pi} \right| = \\ &= \sum_{k=1}^{+\infty} \frac{1}{(k+1)\pi} \int_{k\pi}^{(k+1)\pi} |\sin x| dx = \sum_{k=1}^{+\infty} \frac{2}{(k+1)\pi} = \frac{2}{\pi} \sum_{k=1}^{+\infty} \frac{1}{k+1} \geq \frac{2}{\pi} \sum_{k=1}^{+\infty} \frac{1}{k} = +\infty \end{aligned}$$

Possiamo anche far vedere che la funzione $\frac{\sin x}{x}$ ammette integrale convergente in $(1, +\infty)$, infatti:

$$\int_1^{+\infty} \frac{\sin x}{x} dx = \int_1^\pi \frac{\sin x}{x} dx + \sum_{i=1}^{+\infty} \int_{k\pi}^{(k+1)\pi} \frac{\sin x}{x} dx = \int_1^\pi \frac{\sin x}{x} dx + \sum_{i=1}^{+\infty} (-1)^k \int_{k\pi}^{(k+1)\pi} \frac{|\sin x|}{x} dx$$

dove abbiamo riscritto il nostro integrale come una sommatoria del tipo $\sum_{i=1}^{+\infty} (-1)^k a_k$ dove $a_k =$

$\int_{k\pi}^{(k+1)\pi} \frac{|\sin x|}{x} dx$ siccome, per periodicità della funzione $\sin x$ avremo che

$$\int_{k\pi}^{(k+1)\pi} \sin x = \begin{cases} 1 & \text{se } k \text{ pari} \\ -1 & \text{se } k \text{ dispari} \end{cases}$$

dunque avremo una serie a segni alterni. Tuttavia

$$a_{k+1} = \int_{(k+1)\pi}^{(k+2)\pi} \frac{|\sin x|}{x} dx < \int_{(k+1)\pi}^{(k+2)\pi} \frac{|\sin x|}{(k+1)\pi} < \int_{k\pi}^{(k+1)\pi} \frac{|\sin x|}{x} dx = a_k < a_{k+1}$$

dunque converge per il criterio di Leibniz.

Teorema 6.5 (teorema della convergenza dominata di Lebesgue).

Se $\forall k, f_k : E \rightarrow \bar{\mathbb{R}}$ sono misurabili e $\exists g : E \rightarrow [0; +\infty]$ integrabile tale che $|f_k| \leq g$ e $f_k \xrightarrow{k \rightarrow +\infty} f$, allora

$$\int_E |f| dm_n < +\infty \text{ (ovvero è integrabile secondo Lebesgue)}$$

e

$$\lim_{k \rightarrow +\infty} \int_E |f_k - f| dm_n = 0 \quad (6.4)$$

Oss.: (Scambio del limite con l'integrale)

Prendiamo la tesi del teorema e osserviamo che

$$\lim_{k \rightarrow +\infty} \int_E |f_k - f| dm_n = 0 \implies \lim_{k \rightarrow +\infty} \int_E f_k dm_n = \int_E f dm_n = \int_E \lim_{k \rightarrow +\infty} f_k dm_n$$

ovvero il limite si scambia con l'integrale

6.2.2 Introduzione agli spazi \mathcal{L}^p

Definizione 6.15 (spazio vettoriale \mathcal{L}^p). Sia $1 \leq p < +\infty$, definiamo

$$\mathcal{L}^p(E) = \left\{ f : E \rightarrow \mathbb{R} \text{ mis.} : \int_E |f|^p dm_n < +\infty \right\}$$

e osserviamo che è uno spazio normato rispetto alla norma

$$\|f\|_{\mathcal{L}^p(E)} = \left(\int_E |f|^p dm_n \right)^{\frac{1}{p}}$$

In virtù di quanto detto prima, se $\|f\|_{\mathcal{L}^p(E)} = 0 \implies f = 0$ quasi ovunque su E . È possibile dimostrare che, su questi spazi,

- ① $\mathcal{L}^p(E)$ è uno spazio vettoriale di dimensione infinita quindi se $f, g \in \mathcal{L}^p(E), \lambda, \tau \in \mathbb{R} \implies \lambda f + \tau g \in \mathcal{L}^p(E)$ (abbastanza banale)
- ② $\|f + g\|_{\mathcal{L}^p(E)} \leq \|f\|_{\mathcal{L}^p(E)} + \|g\|_{\mathcal{L}^p(E)}$ (si fa con la diseguaglianza di Holder)

Per rendere completamente $\|\cdot\|_{\mathcal{L}^p(E)}$ una norma possiamo definire una relazione d'equivalenza fra gli elementi dello spazio vettoriale

$$f \sim g \iff f = g \text{ quasi ovunque su } E$$

dunque consideriamo $L^p(E) = \mathcal{L}^p(E)/\sim$ l'insieme quoziente della relazione, dove $\|f\|_{L^p(E)} = \|g\|_{\mathcal{L}^p(E)}$. Osserviamo che $g \sim f$ è ben definita siccome

$$\|f\|_{L^p(E)} = \left(\int_E |f|^p \right)^{\frac{1}{p}} = \left(\int_E |g|^p \right)^{\frac{1}{p}} = \|g\|_{\mathcal{L}^p(E)}$$

dunque $\|f\|_{L^p(E)} = 0 \implies [f]_\sim = 0$. Quindi $\|\cdot\|_{L^p(E)}$ è una norma su $L^p(E)$ ed è completa: il teorema di convergenza dominata ci assicura che le successioni di Cauchy a valori in questo spazio convergono, dunque

$\|f_k - f_m\|_{L^p(E)} \rightarrow 0 \implies \exists f \in L^p(E) : \|f - f_k\|_{L^p(E)} \rightarrow 0 \implies (L^p(E), \|\cdot\|_{L^p(E)})$ è uno spazio di Banach.

Nel caso di $n = 2$ possiamo anche definire un prodotto scalare reale $\varphi : L^2(E) \times L^2(E) \rightarrow \mathbb{R}$ tale che $(u, v) \mapsto \int_E uv dm_n$, osservando che la norma indotta dal prodotto scalare coincide con $L^2(E)$ data prima. La completezza di $L^2(E)$ e la presenza di un prodotto scalare rende tale spazio uno spazio di Hilbert.

6.3 Sezionamenti di insiemi in \mathbb{R}^n

Argomento su cui ci sofferemeremo parecchio è lo sviluppo di alcuni strumenti con cui facilitare il calcolo della misura degli insiemi. Uno degli strumenti più potenti è sicuramente il teorema di Tonelli, il quale ci assicura di poter calcolare la misura di un insieme parametrizzando le altre variabili in funzione delle altre.

Supponiamo che $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^h$ con $x \in \mathbb{R}^k, y \in \mathbb{R}^h$.

Teorema 6.6 (di Tonelli). Sia $E \in \mathfrak{M}(\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^h)$ e sia $f : E \rightarrow [0; +\infty]$ misurabile. Allora abbiamo che

- ① per quasi ovunque $x \in \mathbb{R}^k$, l'insieme E_x e la funzione $f(x, \cdot) : E_x \rightarrow [0; +\infty]$ sono misurabili
- ② la funzione $x \mapsto \int_{E_x} f(x, y) dm_h(y)$ è definita quasi ovunque e misurabile
- ③ $\int_E f dm_n = \int_{\mathbb{R}^k} \left(\int_{E_x} f(x, y) dm_h(y) \right) dm_k(x)$

Oss.: Il quasi ovunque all'interno del teorema ci consente di poter andare ad integrare $x \mapsto \int_{E_x} f(x, y) dy$ anche se $\exists Z \subseteq \mathbb{R}^k$ con $m_k(Z) = 0$ ove non è definita. Infatti, possiamo definire la funzione uguale a 0 su Z , dando comunque luogo ad una funzione misurabile in \mathbb{R}^n il cui valore è sempre indipendente dalla scelta, siccome l'insieme Z è a misura nulla.

Oss.: Il teorema di Tonelli può essere formulato anche invertendo l'ordine di integrazione, ovvero scambiando la x con la y . Più in generale il teorema si può formulare con un qualsiasi raggruppamento di coordinate che fattorizzi $\mathbb{R}^n = \mathbb{R}^k \times \mathbb{R}^h$.

Teorema 6.7 (di Tonelli "invertito"). Sia $E \in \mathfrak{M}(\mathbb{R}^n)$ e sia $f : E \rightarrow [0; +\infty]$ misurabile. Allora abbiamo

- ① per quasi ovunque $y \in \mathbb{R}^h$, E_y e $f(\cdot, y) : E_y \rightarrow [0; +\infty]$ sono misurabili
- ② la funzione $y \mapsto \int_{E_y} f(x, y) dm_k(x)$ è definita quasi ovunque e misurabile
- ③ $\int_E f dm_n = \int_{\mathbb{R}^h} \left(\int_{E_y} f(x, y) dm_k(x) \right) dm_h(y)$

Oss.: (Scambio dell'ordine di integrazione)

Se $f : E \rightarrow [0; +\infty]$ è misurabile allora combinando le due versioni del teorema di Tonelli concludiamo che

$$\int_E f dm_n = \int_{\mathbb{R}^k} \left(\int_{E_x} f(x, y) dm_h(y) \right) dm_k(x) = \int_{\mathbb{R}^h} \left(\int_{E_y} f(x, y) dm_k(x) \right) dm_h(y)$$

ovvero la formula di scambio dell'integrazione.

Nel caso particolare in cui $E = A \times B$ con $A \in \mathfrak{M}(\mathbb{R}^k)$ e $B \in \mathfrak{M}(\mathbb{R}^h)$ (dunque $E \in \mathfrak{M}(\mathbb{R}^n)$ che andrebbe mostrato, ma lo daremo per buono), allora presa $f : A \times B \rightarrow [0, +\infty]$ misurabile allora

$$\int_{A \times B} f dm_n = \int_A \left(\int_B f(x, y) dy \right) dx = \int_B \left(\int_A f(x, y) dx \right) dy$$

siccome $(A \times B)_x = B \forall x \in A$ e $(A \times B)_y = A \forall y \in B$.

Nel caso ancora più particolare in cui, oltre alle condizioni di prima, abbiamo che $f : A \times B \rightarrow [0; +\infty]$ misurabile con $f(x, y) = g(x)w(y)$ con $g : A \rightarrow [0; +\infty]$ e $w : B \rightarrow [0; +\infty]$ allora

$$\int_{A \times B} f(x, y) dm_n = \int_A g(x) dx \int_B w(y) dy$$

Andiamo a formulare un teorema simile a quello di Tonelli ma valido per le funzioni integrabili (secondo Lebesgue).

Teorema 6.8 (teorema di Fubini). Sia $E \in \mathfrak{M}(\mathbb{R}^n)$ e sia $f : E \rightarrow \bar{\mathbb{R}}$ integrabile. Allora abbiamo che

- ① per quasi ovunque $x \in \mathbb{R}^k$, E_x e $f(x, \cdot) : E_x \rightarrow \bar{\mathbb{R}}$ è sommabile;
- ② la funzione $x \mapsto \int_{E_x} f(x, y) dm_h(y)$ è definita quasi ovunque e sommabile;
- ③ $\int_E f dm_n = \int_{\mathbb{R}^k} \left(\int_{E_x} f(x, y) dm_h(y) \right) dm_k(x)$

Oss.: Valgono le medesime considerazioni fatte per il teorema di Tonelli: anche in questo caso, è possibile scambiare l'ordine di integrazione ed è possibile con le condizioni particolare di prima, ovvero $E = A \times B$ e $f : A \times B \rightarrow \bar{\mathbb{R}}$ con $f(x, y) = g(x)w(y)$, separare l'integrale come

$$\int_E f(x, y) dm_n = \int_A g(x) dm_k(x) \int_B w(y) dm_h(y)$$

se f è integrabile, $A \in \mathfrak{M}(\mathbb{R}^k)$, $B \in \mathfrak{M}(\mathbb{R}^h)$ e $g : A \rightarrow \bar{\mathbb{R}}$ e $W : b \rightarrow \bar{\mathbb{R}}$ misurabili.

Esempio 6.1. Consideriamo una funzione $f(x, y)$ integrabile sul dominio rettangolare $D = (0, 1)^2$ ($\iff \int_{(0,1)^2} |f(x, y)| dx dy < +\infty$) con la proprietà che $f(x, y) = -f(y, x)$. Mostrare che $\int_{(0,1)^2} f(x, y) dx dy = 0$.

Svolgimento.

$$\begin{aligned} \int_{(0,1)^2} f dm_2 &\stackrel{(0,1)^2 = (0,1) \times (0,1)}{=} \int_0^1 \left(\int_0^1 f(x, y) dy \right) dx = - \int_0^1 \left(\int_0^1 f(y, x) dy \right) dx = \\ &= - \int_{(0,1)^2} f dm_2 \implies \int_{(0,1)^2} f dm_2 = 0 \end{aligned}$$

□

Esempio 6.2. Consideriamo la funzione $f(x, y) = \frac{x^2 - y^2}{(x^2 + y^2)^2}$ che soddisfa $f(x, y) = -f(y, x)$. Mostrare che non è integrabile.

Svolgimento. Non sappiamo ancora se f sia integrabile in $E = (0, 1)^2$, ovvero se $\int_E |f| dm_2$ è finito o $+\infty$. Supponiamo per assurdo che lo sia, allora dovrebbe valere lo scambio dell'ordine di integrazione:

$$\int_E f(x, y) = \int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx \right) dy = \int_0^1 \left(\int_0^1 \frac{1}{x^2 + y^2} - 2 \frac{y^2}{(x^2 + y^2)^2} \right) dy$$

a questo punto facciamo il cambio di variabile $x = ty \implies dx = y dt$, pertanto

$$\begin{aligned} & \int_0^1 \left(\int_0^1 \frac{1}{x^2 + y^2} - 2 \frac{y^2}{(x^2 + y^2)^2} \right) dy = \int_0^1 \left(\int_0^{\frac{1}{y}} \frac{1}{y^2(1+t^2)} - 2 \frac{y^2}{y^4(1+t^2)} y dt \right) dy = \\ & = \int_0^1 \frac{1}{y} \left[\arctan \frac{1}{y} - 2 \int_0^1 \frac{1}{(1+t^2)} dt \right] dy \end{aligned}$$

Studiamo il secondo termine:

$$\int_0^{\frac{1}{y}} \frac{1}{(1+t^2)^2} dt = \int_0^{\frac{1}{y}} \frac{1}{(1+t^2)} dt - \int_0^{\frac{1}{y}} \frac{t^2}{(1+t^2)^2} dt = \arctan \frac{1}{y} - \int_0^{\frac{1}{y}} \frac{t^2}{(1+t^2)^2} dt$$

dunque

$$\int_0^{\frac{1}{y}} \frac{t^2}{(1+t^2)^2} dt = \int_0^{\frac{1}{y}} \frac{t}{2} \left(\frac{(1+t^2)^{-1}}{(-1)} \right) dt = \frac{t}{2(1+t^2)} \Big|_0^{\frac{1}{y}} + \int_0^{\frac{1}{y}} \frac{1}{2} \frac{1}{1+t^2} dt = -\frac{\frac{1}{y}}{(1+\frac{1}{y^2})} + \frac{1}{2} \arctan \frac{1}{y}$$

dunque

$$\int_E f dm_2 = - \int_0^1 \frac{1}{1+y^2} dy = - \int_0^1 \frac{1}{1+y^2} dy = -\frac{\pi}{4}.$$

Ora osserviamo che, sezionando rispetto a y costante prima, abbiamo che

$$\begin{aligned} & \int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dy \right) dx \stackrel{\text{raccolgo un } -}{=} - \int_0^1 \left(\int_0^1 \frac{y^2 - x^2}{(x^2 + y^2)^2} dy \right) dx \stackrel{\text{usiamo l'ip. } f(x,y) = -f(y,x)}{=} \\ & = - \int_0^1 \left(\int_0^1 \frac{x^2 - y^2}{(x^2 + y^2)^2} dx \right) dy = - \int_E f dm_2 = \frac{\pi}{4} \end{aligned}$$

Avevamo supposto che f fosse integrabile, ma questo porta ad un assurdo con il teorema di Fubini; pertanto concludiamo che la nostra funzione non fosse integrabile secondo Lebesgue. \square

Esercizio 6.3. Calcolare la misura di $E = \{(x, y) \in \mathbb{R}^2 : 0 < x \leq 1, 0 < |y| < \frac{1}{\sqrt{x}}\}$

Svolgimento. L'insieme E è misurabile in quanto è l'unione di un insieme aperto con l'insieme $S = \{(x, y) \in \mathbb{R}^2 : x = 1, |y| < 1\} = \{1\} \times (-1, 1)$ che ha misura nulla.

$$m_2(E) = \int_{\mathbb{R}} m_1(E_x) dx = \int_0^1 m_1 \left(\left(-\frac{1}{\sqrt{x}}, \frac{1}{\sqrt{x}} \right) \right) dx = 2 \int_0^1 \frac{1}{\sqrt{x}} dx = 4 \sqrt{x} \Big|_{x=0}^{x=1} = 4$$

dove abbiamo applicato il teorema di Tonelli alla funzione caratteristica $\mathbb{1}_E$, non negativa e misurabile.

Possiamo, pertanto, anche scambiare l'ordine di integrazione osservando che $|y| < \frac{1}{\sqrt{x}} \implies$

$x < \frac{1}{y^2}$: tuttavia devono essere vere contemporaneamente le due diseguaglianze $0 < x \leq 1$ e $0 < x < \frac{1}{y^2} \implies 0 < x < \min\{1, \frac{1}{y^2}\}$. Dunque

$$\begin{aligned} m_2(E) &= \int_{\mathbb{R}^2} \mathbb{1}_E dm_2 = \int_{\mathbb{R}} \left(\int_{\mathbb{R}} \mathbb{1}_E(x, y) dx \right) dy = \int_{\mathbb{R}} m_1(E_y) dy = \\ &= \int_{\mathbb{R}} m_1 \left(\left\{ x \in (0, 1) : |y| < \frac{1}{\sqrt{x}} \right\} \right) dy = 2 \int_0^{+\infty} m_1 \left(\left\{ x \in (0, 1) : 0 < x < \frac{1}{y^2} \right\} \right) dy = \\ &= 2 \int_0^{+\infty} \left(\int_0^{\min\{1, \frac{1}{y^2}\}} \mathbb{1}_E dx \right) dy = 2 \int_0^1 \left(\int_0^1 \mathbb{1}_E dx \right) dy + \\ &\quad + 2 \int_1^{+\infty} \left(\int_0^{\frac{1}{y^2}} \mathbb{1}_E dx \right) dy = 2 + 2 \int_1^{+\infty} \frac{1}{y^2} dy = 2 + 2 = 4 \end{aligned}$$

Dunque da questo esercizio capiamo bene che spesso sono convenienti alcuni sezionamenti rispetto ad altri \square

Esercizio 6.4. Consideriamo $f : \Omega \rightarrow \mathbb{R}$ dove $\Omega = \{(x, y) \in \mathbb{R}^2 : 1 < x < y < e^2, y > e\}$ e $f(x, y) = \frac{1}{x \log y}$. Stabilire se f è integrabile ed, in tal caso, calcolare $\int_{\Omega} f dm_2$

Svolgimento. Osserviamo innanzitutto che la funzione $f > 0 \forall (x, y) \in \Omega$, f è continua e quindi anche misurabile in Ω ed è pertanto ben definito $\int_{\Omega} \frac{1}{x \log y} dx dy$. Possiamo, pertanto, applicare il teorema di Tonelli rispetto ad uno dei possibili sezionamenti:

$$\int_{\Omega} \frac{1}{x \log y} dx dy = \int_{\mathbb{R}} \left(\int_{\Omega_y} \frac{1}{x \log y} dx \right) dy = \int_e^{e^2} \frac{1}{\log y} \left(\int_1^y \frac{1}{x} dx \right) dy = \int_e^{e^2} 1 dy = e(e - 1)$$

Scambiamo l'ordine di integrazione

$$\begin{aligned} \int_{\Omega} \frac{1}{x \log y} dx dy &= \int_1^{e^2} \frac{1}{x} \left(\int_{\Omega_x} \frac{1}{\log y} dy \right) dx = \int_1^e \frac{1}{x} \left(\int_e^{e^2} \frac{1}{\log y} dy \right) dx + \\ &\quad + \int_e^{e^2} \frac{1}{x} \left(\int_x^{e^2} \frac{1}{\log y} dy \right) dx = \int_e^{e^2} \frac{1}{\log y} dy + \int_e^{e^2} \left(\int_x^{e^2} \frac{1}{\log y} dx \right) dy \end{aligned}$$

tuttavia $\int_e^x \frac{1}{\log t} dt$ non ha primitiva esprimibile sotto forma delle note funzioni che conosciamo. \square

Esercizio 6.5. Sia $f : E \rightarrow \mathbb{R}$ e $f(x, y, z) = \sqrt{x^2 - z^2} \log y$ dove $E = \{(x, y, z) \in \mathbb{R}^3 : 0 \leq z \leq x \leq 1, 0 < y \leq x\}$

Svolgimento. Osserviamo che la funzione $f(x) \leq 0 \forall (x, y, z) \in E$. Questo si può facilmente vedere siccome $0 < y \leq x \leq 1 \implies 0 < y \leq 1$ e possiamo inoltre vedere che f è continua, dunque misurabile.

Per le proposizioni viste in precedenza sappiamo che $f^-(x, y, z)$ è ben definita e misurabile $\implies \int_E f^-(x, y, z) dx$ è ben definito, dunque possiamo applicarvi il teorema di Tonelli (ricordiammo infatti che tale teorema è applicabile esclusivamente il nostro dominio è scrivibile sotto forma di prodotti cartesiani):

$$\begin{aligned} \int_E f^-(x, y, z) dm_3 &= \int_0^1 \left(\int_{E_x} f dy dz \right) dx = \int_0^1 \left(\int_{E_x} f dy dz \right) dx = \\ &= \int_0^1 \left(\int_{E_x} \sqrt{x^2 - z^2} \log y dy dz \right) dx \end{aligned}$$

osserviamo che $E_x = \{(y, z) \in \mathbb{R}^2 : 0 \leq z \leq x, 0 < y \leq z\}$, dunque

$$\int_E f dm_3 = \int_0^1 \left(\int_0^x \left(\sqrt{x^2 - z^2} \log y dy \right) dz \right) dx = \int_0^1 \left(\int_0^x \left(\sqrt{x^2 - z^2} \int_0^z \log y dy \right) dz \right) dx$$

osserviamo che

$$\int_0^x \log y dy = y \log y \Big|_{y \rightarrow 0^+}^{y=x} - \int_0^x 1 dy = x \log x - x$$

$$\begin{aligned} \int_0^x (x^2 - z^2)^{\frac{1}{2}} dz &= x^2 \int_0^1 \sqrt{1 - t^2} dt = x^2 \int_0^{\frac{\pi}{2}} \sqrt{1 - \sin^2 \theta} \cos \theta d\theta = x^2 \int_0^{\frac{\pi}{2}} \cos^2 \theta d\theta = \\ &= x^2 \int_0^{\frac{\pi}{2}} \left(\frac{1 + \cos 2\theta}{2} \right) d\theta = x^2 \left(\frac{\pi}{4} + \frac{1}{2} \int_0^{\frac{\pi}{2}} \cos 2\theta d\theta \right) = \\ &= x^2 \frac{\pi}{4} \end{aligned}$$

dunque

$$\begin{aligned} \int_E f dm_3 &= \int_0^1 (x \log x - x) \frac{\pi}{4} x^2 dx = \frac{\pi}{4} \int_0^1 (x^3 \log x - x^3) dx = \\ &= \frac{\pi}{4} \left(\frac{x^4}{4} \log x \Big|_{x \rightarrow 0^+}^{x=1} - \int_0^1 \frac{x^3}{4} dx - \frac{1}{4} \right) = \frac{\pi}{4} \left(-\frac{1}{4} - \frac{1}{16} \right) = \\ &= -\frac{5\pi}{64} \end{aligned}$$

□

6.4 Teorema del cambiamento di variabili

Come nel caso degli integrali a singola variabile, vogliamo adesso effettuare dei cambi di variabile all'interno degli integrali. Infatti il passaggio a coordinate permette, spesso, di fruire più facilmente di alcune simmetrie di cui godono certi insiemi oppure certe funzioni (si pensi a voler calcolare il volume occupato da una sfera: in coordinate cartesiane può essere un incubo parametrizzare le variabili, tuttavia potremmo pensare di scrivere le coordinate dei punti $(x, y, z) \mapsto (\rho, \theta, \varphi)$). Per fare ciò abbiamo bisogno di una serie di risultati:

Teorema 6.9. Sia $v_1, \dots, v_n \in \mathbb{R}^n$ e sia $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $L(x) = \sum_{i=1}^n x_i v_i$. Allora

$$m_n(L([0, 1]^n)) = |\det [v_1 \ \dots \ v_n]| \tag{6.5}$$

Oss.: Naturalmente se $\{v_1, \dots, v_n\}$ sono una base ortonormale di \mathbb{R}^n allora

$$m_n(L([0, 1]^n)) = |\det [v_1, \dots, v_n]| = 1$$

ovvero le isometrie preservano la misura di Lebesgue. Questo, naturalmente, risulta intuitivamente ovvio dal punto di vista geometrico, siccome l'area di un rettangolo ruotato o traslato sicuramente risulta sempre la stessa.

Oss.: Supponendo che $l_1, \dots, l_n > 0$, allora

$$L([0, l_1] \times \dots \times [0, l_n]) = \left\{ v = \sum_{i=1}^n t_i v_i : 0 \leq t_i \leq x_i \right\} = \left\{ \sum_{i=1}^n \tau_i l_i v_i : 0 \leq \tau_i \leq 1 \right\} = T([0, 1]^n)$$

dove $T(x) = \sum_{i=1}^n \tau_i l_i v_i$, dunque

$$\begin{aligned} m_n(L([0, l_1] \times \dots \times [0, l_n])) &= m_n(T([0, 1]^n)) = |\det [l_1 v_1, l_2 v_2, \dots, l_n v_n]| = \\ &= |\det [v_1, \dots, v_n]| \prod_{i=1}^n l_i = |\det [v_1, \dots, v_n]| m_n([0, l_1] \times \dots \times [0, l_n]) \end{aligned}$$

Definizione 6.16 (jacobiano). Se $f : \Omega \rightarrow \mathbb{R}^n$, Ω aperto $\subseteq \mathbb{R}^n$ e $f \in C^1(\Omega, \mathbb{R}^n)$, $x \in \Omega$ allora

$$Jf(x) = |\det Df(x)|$$

detto jacobiano di f in x .

Oss.: Sia $L : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $L(e_i) = v_i \forall i \in \{1, \dots, n\} \implies A = [v_1, \dots, v_n] \in \mathbb{R}^{n \times n}$ è la matrice associata ad L allora

$$DL = A = [v_1, \dots, v_n]$$

e dunque per l'osservazione precedente avremo che

$$\begin{aligned} m_n(L([0, l_1] \times \dots \times [0, l_n])) &= |\det [v_1, \dots, v_n]| l_1 \dots l_n = DL m_n([0, l_1] \times \dots \times [0, l_n]) \implies \\ \implies DL &= \frac{m_n(L([0, l_1] \times \dots \times [0, l_n]))}{m_n([0, l_1] \times \dots \times [0, l_n])} \end{aligned}$$

dove in realtà Q può essere sostituito da un qualunque insieme $E \in \mathfrak{M}(\mathbb{R}^n)$ con $m_n(E) > 0$.

In maniera euristica possiamo dare una traccia di dimostrazione del teorema di cambio di variabile che vogliamo adesso utilizzare. Supponiamo di voler effettuare un cambio di variabili all'interno dell'integrale: la funzione $f(x) \in C^1$ intorno a x_0 può essere approssimata con un'applicazione lineare del tipo

$$f(x) = f(x_0) + Df(x_0)(x - x_0) \implies f(x_0 + I) \approx f(x_0) + Df(x_0)(I)$$

ove $I = [0, x_1] \times [0, x_n]$ per $x_1, \dots, x_n > 0$. Pertanto per quanto visto prima

$$m_n(f(x_0 + I)) = m_n(Df(x_0)(I)) = Jf(x_0) \prod_{i=1}^n x_i \implies df(x) = Jf(x) dx_1 \dots dx_n$$

Per alleggerire l'enunciato del teorema, diamo la definizione di diffeomorfismo che rende un poco più snello l'enunciato del teorema del cambiamento di variabile.

Definizione 6.17 (diffeomorfismo). Diremo che $\varphi : \Omega \rightarrow \mathcal{U}$ è un diffeomorfismo se

- ① φ è invertibile;
- ② $\varphi \in C^1(\Omega, \mathcal{U})$;
- ③ $\varphi^{-1} \in C^1(\mathcal{U}, \Omega)$

Possiamo dare due versioni di questo teorema: uno per le funzioni misurabili, un altro per le funzioni integrabili (come per il teorema di Tonelli e di Fubini).

Teorema 6.10 (cambio di variabili per funzioni misurabili). Sia $E \in \mathfrak{M}(\mathbb{R}^n)$, $f : E \rightarrow [0; +\infty]$ misurabile, $\Omega, \mathcal{U} \subseteq \mathbb{R}^n$ aperti, $T : \Omega \rightarrow \mathcal{U}$ diffeomorfismo. Allora

① $T^{-1}(E) \in \mathfrak{M}(\mathbb{R}^n)$ e $f \circ T : T^{-1}(E) \rightarrow [0; +\infty]$ è misurabile;

②

$$\int_E f(y) dy = \int_{T^{-1}(E)} f(T(x)) JT(x) dx$$

Teorema 6.11 (cambio di variabili per funzioni integrabili). Sia $E \in \mathfrak{M}(\mathbb{R}^n)$, $f : E \rightarrow \bar{\mathbb{R}}$ integrabile secondo Lebesgue, $\Omega, \mathcal{U} \subseteq \mathbb{R}^n$ aperti, $T : \Omega \rightarrow \mathcal{U}$ diffeomorfismo. Allora

① $T^{-1}(E) \in \mathfrak{M}(\mathbb{R}^n)$ e $f \circ T : T^{-1}(E) \rightarrow [0; +\infty]$ è misurabile;

②

$$\int_E f(y) dy = \int_{T^{-1}(E)} f(T(x)) JT(x) dx$$

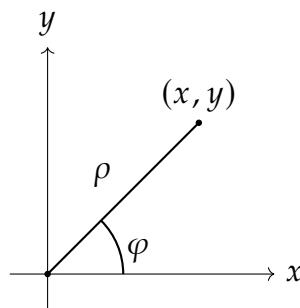
Andiamo adesso a discutere dei vari sistemi di coordinate a cui passare per rendere più *handy* il calcolo degli integrali.

6.4.1 Coordinate polari

Per identificare un punto, invece di identifierlo tramite la sua ascissa e la sua ordinata, possiamo pensare di individuarlo tramite la sua distanza dall'origine e dall'angolo che il segmento che lo congiunge con l'origine forma con l'asse delle ascisse. Dunque

$$\begin{cases} x = \rho \cos \varphi \\ y = \rho \sin \varphi \end{cases} \iff (x, y) = T(\rho, \varphi) \quad \rho > 0, 0 < \varphi < 2\pi$$

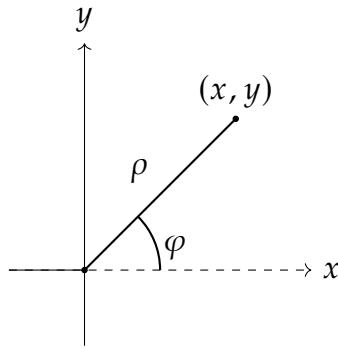
dunque potremmo



osserviamo che il jacobiano di questa trasformazione T che è in funzione delle variabili (ρ, φ) sarà pari a

$$DT(\rho, \varphi) = \det \begin{bmatrix} \cos \varphi & -\rho \sin \theta \\ \sin \varphi & \rho \cos \theta \end{bmatrix}$$

per garantire l'iniettività della nostra trasformazione, definiamo $T : \mathbb{R}^+ \times (0, 2\pi) \mapsto \mathbb{R}^2 \setminus ([0; +\infty) \times \{0\})$ pertanto verrà "generato" tutto il piano tranne l'asse delle ascisse positive, come si può vedere dall'immagine qua sotto:



e, per quanto riguarda la sua inversa, possiamo vedere che presi x e y possiamo risalire alla ρ tramite la semplice relazione $\rho = \sqrt{x^2 + y^2}$, mentre l'angolo φ , posto $\Omega_0 = \mathbb{R}^2 \setminus ([0; +\infty) \times \{0\})$, può essere ottenuto considerando banalmente la funzione argomento, che si può mostrare essere, posto $C^\infty(\Omega_0)$.

Oss.: Ci si chiede, inoltre, se il fatto che T non generi tutto lo spazio possa, in qualche maniera contribuire agli integrali; tuttavia si può mostrare che

$$m_2(\mathbb{R}^2 \setminus \Omega_0) = 0$$

questo si può facilmente vedere dal fatto che $\mathbb{R}^2 \setminus \Omega_0 = [0; \infty) \times \{0\}$ che corrisponde ad un insieme che, per come avevamo definito le operazioni algebriche in $[0; +\infty]$, corrisponde ad un insieme a misura nulla.

Dunque, il suo contributo all'integrale è nullo.

Per le coordinate polari, dunque, se $E \subseteq \mathbb{R}^2$ misurabile allora

$$\int_E f(x) dx = \int_{T^{-1}(E)} f(\rho \cos \varphi, \rho \sin \varphi) \rho d\rho d\varphi \quad (6.6)$$

Esercizio 6.6. Calcolare $\int_{-\infty}^{+\infty} e^{-t^2} dt$

Svolgimento. La funzione e^{-x^2} è naturalmente integrabile, siccome $\frac{e^{-x^2}}{\frac{1}{x^2}} \xrightarrow{x \rightarrow +\infty} 0$ per gli ordini di infinito, dunque $e^{-x^2} \ll \frac{1}{x^2} \implies \int_1^{+\infty} e^{-x^2} dx \ll \int_1^{+\infty} \frac{1}{x^2}$ mentre sull'intervallo $[0; 1]$ la funzione è limitata e continua. La parità della funzione garantisce lo stesso per l'intervallo $(-\infty; 0]$. Osserviamo adesso che, tramite l'integrabilità, siamo sicuri di poter applicare *Fubini* siccome l'intervallo $(-\infty; +\infty)$ è sicuramente misurabile in quanto aperto. Osserviamo che, applicando *Fubini*, su un insieme del tipo $\mathbb{R} \times \mathbb{R}$ allora

$$\int_{-\infty}^{+\infty} e^{-(x^2+y^2)} dx dy = \int_{-\infty}^{+\infty} e^{-y^2} dy \int_{-\infty}^{+\infty} e^{-x^2} dx = \left(\int_{-\infty}^{+\infty} e^{-x^2} dx \right)$$

dunque possiamo studiare direttamente l'integrale in due variabili dove possiamo passare alle coordinate polari. Osserviamo che

$$\begin{aligned} \int_{\mathbb{R}^2} e^{-x^2-y^2} dx dy &= \int_0^{+\infty} \int_0^{2\pi} e^{-\rho^2 \cos^2 \varphi - \rho^2 \sin^2 \varphi} \rho d\theta d\rho = \int_0^{+\infty} \rho d\rho \int_0^{2\pi} e^{-\rho^2} \rho d\rho = \\ &= -\pi \int_0^{+\infty} -2\rho e^{-\rho^2} d\rho = -\pi e^{-\rho^2} \Big|_0^{+\infty} = 0 + \pi = \pi \end{aligned}$$

dunque $\int_{-\infty}^{+\infty} e^{-x^2} dx = \sqrt{\pi}$

□

Esercizio 6.7. Mostrare che $\int_{(0,1)^2} \frac{|x^2-y^2|}{(x^2+y^2)^2} dx dy = +\infty$

Dimostrazione.

$$\begin{aligned} \int_{(0,1)^2} \frac{|x^2-y^2|}{(x^2+y^2)^2} dx dy &\geq \int_{B(0,1) \cap (0,1)^2} \frac{|x^2-y^2|}{(x^2+y^2)^2} dx dy = \int_0^1 \int_0^{2\pi} \frac{|\cos^2 \varphi - \sin^2 \varphi|}{\rho} d\rho d\varphi = \\ &= \int_0^1 \frac{1}{\rho} d\rho \int_0^{2\pi} |\cos \frac{\varphi}{2}| d\varphi = +\infty \end{aligned}$$

□

6.4.2 Volumi dei solidi di rotazione

Definiamo un solido di rotazione attorno all'asse z come l'insieme misurabile $E \subseteq \mathbb{R}^3$ descritto dal profilo $\rho : [h_1, h_2] \rightarrow [0; +\infty)$ continuo tale che abbiamo

$$E = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \leq \rho(z), h_1 \leq z \leq h_2\}$$

Per misurare tale insieme possiamo usare il teorema di Tonelli banalmente

$$\begin{aligned} m_3(E) &= \int_E \mathbb{1}_E dm_3 = \int_{h_1}^{h_2} \left(\int_{E_z} dx dy \right) = \int_{h_1}^{h_2} \left(\int_{E_z} \rho d\rho d\varphi \right) = \\ &= \int_{h_1}^{h_2} \left(\int_0^{\rho(z)} \rho d\rho \int_0^{2\pi} d\varphi \right) dz = \int_{h_1}^{h_2} \pi \rho^2(z) dz = \pi \int_{h_1}^{h_2} \rho^2(z) dz \end{aligned}$$

Esempio 6.3. Consideriamo un cono retto di altezza $H > 0$ e raggio $R > 0$. Definiamo $\rho(z)$ lineare e tale che $\rho(H) = 0, \rho(0) = R \implies \rho(z) = R - \frac{R}{H}z, \rho : [0, H] \rightarrow [0; +\infty)$, pertanto

$$m_3(C) = \pi \int_0^H \left(R - \frac{R}{H}z \right)^2 dz = \pi R^2 \int_0^H \left(1 - \frac{z}{H} \right) dz = \pi R^2 \left[\left(\frac{z}{H} - 1 \right)^3 - \frac{H}{3} \right]_{z=0}^{z=H} = \frac{\pi}{3} R^2 H$$

Abbiamo quindi mostrato il seguente teorema:

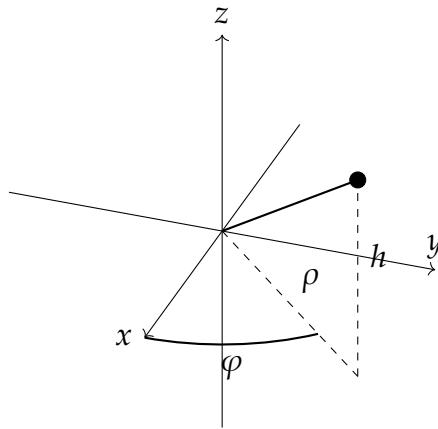
Teorema 6.12. Sia $E = \{(x, y, z) \in \mathbb{R}^3 : x^2 + y^2 \leq \rho(z), h_1 \leq z \leq h_2\}$ (a meno di scambiare una coordinata). Allora

$$m_3(E) = \pi \int_{h_1}^{h_2} \rho^2(z) dz$$

Adesso cerchiamo di estendere la definizione di coordinate polari a 3 dimensioni. Vogliamo che, nel caso di z costante, individuare un punto alla "vecchia" maniera 2-dimensionale: in pratica si vuole individuare una trasformazione per cui il punto è caratterizzato dalla sua "altezza" (ovvero la sua coordinata z) e, proiettando il punto sul piano (x, y), dalla lunghezza e dall'angolo formato con l'asse x del segmento che congiunge l'origine con il punto proiettato. Una possibile trasformazione "candidata" può essere la seguente

$$T(\rho, \varphi, h) = (\rho \cos \varphi, \rho \sin \varphi, h) = (x, y, z) \implies \begin{cases} x = \rho \cos \varphi \\ y = \rho \sin \varphi \\ z = h \end{cases}$$

e riporto una figura più chiarire meglio il concetto espresso sopra:



pertanto $T : \mathbb{R}^+ \times (0, 2\pi) \times \mathbb{R} \mapsto (\mathbb{R}^2 \setminus [0; +\infty) \times \{0\}) \times \mathbb{R}$. Dunque

$$JT(\rho, \varphi, h) = \det \begin{bmatrix} \cos \varphi & -\rho \sin \varphi & 0 \\ \sin \varphi & \rho \cos \varphi & 0 \\ 0 & 0 & 1 \end{bmatrix} = \rho$$

il che naturalmente implica

$$\int_E f(x, y, z) dx dy dz = \int_{T^{-1}(E)} f(\rho \cos \varphi, \rho \sin \varphi, h) \rho d\rho d\varphi dh$$

Possiamo adesso pensare di calcolare il volume di rotazione E ottenuto ruotando un insieme $A \subseteq (0, +\infty) \times \mathbb{R}$ attorno all'asse z : supponiamo per esempio che

$$E = \{(\rho, \varphi, h) : (\rho, \varphi) \in A, \varphi \in [0, 2\pi]\}$$

e calcoliamo la misura di tale insieme

$$m_3(E) = \int_E dx dy dz = \int_{T^{-1}(E)} \rho d\rho d\varphi dh = \int_0^{2\pi} \left(\int_{\Omega_\varphi} \rho d\rho dh \right) d\varphi = 2\pi \int_A \rho d\rho dh$$

Introducendo il concetto di baricentro, potremmo pensare di definire una grandezza che misura la distanza *media* dall'asse z che consente di calcolare il volume come una sorta di base \times altezza:

$$m_3(E) = 2\pi R_b m_2(A) = 2\pi \int_A \rho d\rho dh \implies R_b = \frac{\int_A \rho d\rho dh}{m_2(A)}$$

Abbiamo mostrato il seguente teorema

Teorema 6.13 (secondo teorema di Guldino).

$$m_3(A) = 2\pi R_b m_2(A)$$

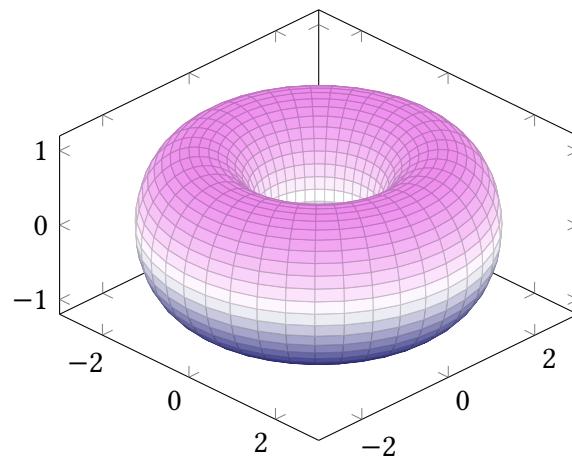
dove R_b è la distanza del baricentro rispetto all'asse intorno a cui ruota

Esercizio 6.8 (Volume del toro). *Individuare il volume di un toro, ottenuto dalla rotazione di una circonferenza con centro a distanza $R > 0$ dall'origine e con raggio $r > 0$ con $R > r$*

Svolgimento. Cercare di parametrizzare l'integrale non è difficile ma non necessario: abbiamo il secondo teorema di Guldino dalla nostra parte. Il baricentro, in questo caso, sarà sempre R e l'area dell'insieme ruotato sarà invece πr^2 dunque

$$m_3(T) = 2\pi^2 R r^2$$

□



6.4.3 Invarianza per traslazioni e omogeneità della misura di Lebesgue

Sia $E \in \mathfrak{M}(\mathbb{R}^n)$, $\lambda \in \mathbb{R} \setminus \{0\}$, $x \in \mathbb{R}^n \implies$

$$\textcircled{1} \quad m_n(x + E) = m_n(E)$$

$$\textcircled{2} \quad m_n(\lambda E) = |\lambda|^n m_n(E)$$

dove abbiamo definito

$$\begin{aligned} \lambda E &= \{\lambda y \in \mathbb{R}^n : y \in E\} \\ x + E &= \{x + y \in \mathbb{R}^n : y \in E\} \end{aligned}$$

Questo può essere facilmente visto come conseguenza di una banalissima trasformazione delle coordinate per mezzo della funzione

$$T(y) = x + y \text{ con } JT = 1 \implies \int_T(E) 1 dz = \int_{E+x} 1 dz = \int_E 1 dz \iff \textcircled{1}$$

nel caso della traslazione e nel caso della dilatazione poiché

$$H(y) = \lambda y \text{ con } JH = |\lambda|^n \implies \int_{\lambda E} 1 dz = \int_{T(E)} 1 dz = |\lambda|^n \int_E 1 dz \iff \textcircled{2}$$

Proponiamo un esercizio piuttosto complicato, che è un buon esercizio per gli integrali

Esercizio 6.9. Provare che $\int_{B(0,1)} \frac{1}{|x|^\alpha}$ converge $\iff 0 < \alpha < n$. Supporre per ipotesi che $\alpha > 0$

Svolgimento. Vogliamo determinare per quali valori di $\alpha > 0$ il seguente integrale converge:

$$\int_{B(0,1)} \frac{1}{|x|^\alpha} dx$$

possiamo pensare di definire una funzione del tipo

$$\mathbb{1}_{[0, \frac{1}{|x|^\alpha})}(t) = \begin{cases} 1 & \text{se } 0 \leq t < \frac{1}{|x|^\alpha} \\ 0 & \text{se } t \geq \frac{1}{|x|^\alpha} \end{cases}$$

ora osserviamo che $0 \leq |x| \leq 1$ e dalla condizione $0 < t < \frac{1}{|x|^\alpha}$ segue che $0 \leq x < \min\{1, \frac{1}{t^{\frac{1}{\alpha}}}\}$, dunque avremo che per $t \in [0; 1]$, $t^{\frac{1}{\alpha}} < 1 \implies \frac{1}{t^{\frac{1}{\alpha}}} \geq 1$ ma abbiamo il vincolo che $|x| <$

$\min(1, \frac{1}{t^{\frac{1}{\alpha}}}) = 1$. Per $t \geq 1$ avremo invece che $t^{\frac{1}{\alpha}} \geq 1 \implies \frac{1}{t^{\frac{1}{\alpha}}} \leq 1 \implies |x| < \frac{1}{t^{\frac{1}{\alpha}}}$ dunque la funzione caratteristica non si annullerà per $|x| < \frac{1}{t^{\frac{1}{\alpha}}}$

$$\begin{aligned} & \int_0^1 \left(\int_{B(0,1)} \mathbb{1}_{B\left(0, \frac{1}{t^{\frac{1}{\alpha}}}\right)}(x) dx \right) dt + \int_1^{+\infty} \int_{B(0,1)} \left(\mathbb{1}_{B\left(0, \frac{1}{t^{\frac{1}{\alpha}}}\right)}(x) dx \right) dt = \\ &= m_n(B(0,1)) + \int_1^{+\infty} m_n \left(B \left(0, \frac{1}{t^{\frac{1}{\alpha}}} \right) \right) \text{per omogeneità} \\ &\stackrel{\text{per omogeneità}}{=} m_n(B(0,1)) + m_n(B(0,1)) \int_1^{+\infty} \frac{1}{t^{\frac{n}{\alpha}}} = m_n(B(0,1)) \left(1 + \frac{t^{-\frac{n}{\alpha}+1}}{1 - \frac{n}{\alpha}} \Big|_1^{+\infty} \right) \end{aligned}$$

dove quest'ultima espressione converge solamente per $0 < \alpha < n$. \square

Esercizio 6.10. Provare che $\int_{\mathbb{R}^n \setminus B(0,1)} \frac{1}{|x|^\alpha} dx < +\infty \iff \alpha > n$.

Svolgimento. L'idea è simile a come svolto precedente: definiamo la funzione caratteristica definita come segue

$$\mathbb{1}_{[0; \frac{1}{|x|^\alpha})} = \begin{cases} 1 & \text{se } 0 \leq t \leq \frac{1}{|x|^\alpha} \\ 0 & \text{altrimenti} \end{cases}$$

e possiamo dunque considerare l'integrale della funzione $\mathbb{1}_{[0; \frac{1}{|x|^\alpha})} : [0; +\infty] \times (\mathbb{R}^n \setminus B(0,1))$, parametrizzando la $|x|$ in funzione della t osservando che quando $1 \leq t \leq +\infty$ siamo nella regione in cui la funzione caratteristica si annulla. Fra $[0, 1]$ la funzione $\mathbb{1}_{[0; \frac{1}{|x|^\alpha})}$ sarà 0 per $t \geq \frac{1}{|x|^\alpha} \implies |x| \geq t^{-\frac{1}{\alpha}}$.

Fissando la t integreremo la x fra $B(0, \frac{1}{t^{-\frac{1}{\alpha}}}) \setminus B(0, 1)$, dunque

$$\begin{aligned} & \int_{[0, +\infty] \times (\mathbb{R}^n \setminus B(0,1))} \mathbb{1}_{[0; \frac{1}{|x|^\alpha})} = \\ &= \int_0^1 \left(\int_{\mathbb{R}^n \setminus B(0,1)} \mathbb{1}_{[0; \frac{1}{t^{\frac{1}{\alpha}}})} dx \right) dt + \int_1^{+\infty} \left(\int_{\mathbb{R}^n \setminus B(0,1)} \mathbb{1}_{[0; \frac{1}{t^{\frac{1}{\alpha}}})} dx \right) dt \end{aligned}$$

dove l'ultimo integrale si annulla per quanto detto prima. Osserviamo che il primo integrale invece è pari a

$$\begin{aligned} & \int_0^1 m_n \left(B \left(0, \frac{1}{t^{\frac{1}{\alpha}}} \right) \setminus B(0, 1) \right) dt = \int_0^1 m_n \left(B \left(0, \frac{1}{t^{-\frac{1}{\alpha}}} \right) \right) - m_n(B(0,1)) dt \stackrel{\text{per omogeneità}}{=} \\ &\stackrel{\text{per omogeneità}}{=} m_n(B(0,1)) \int_0^1 t^{-\frac{n}{\alpha}} - 1 dt \end{aligned}$$

il quale converge se e solo se $-\frac{n}{\alpha} > -1 \implies \frac{n}{\alpha} < 1 \implies \alpha > n$. \square

Potremmo mostrare, in maniera più formale, la seguente proposizione

Proposizione 6.12. Se $f : E \rightarrow [0; +\infty]$ misurabile, allora

$$\int_E f(x) dx = m_{n+1}(\{(x, t) \in \mathbb{R}^{n+1} : x \in E, 0 \leq t \leq f(x)\})$$

Dimostrazione.

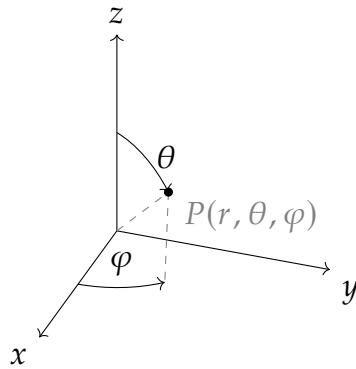
$$\int_E f(x) dx = \int_E \int_0^{f(x)} 1 dt dx = \int_0^{f(x)} \left(\int_E f(x) dx \right) dt = \int_0^{f(x)} m_n(\{x \in E : f(x) \geq t\}) dt$$

\square

Questo, applicato ai precedenti esercizi, rende molto facile il conto.

6.4.4 Coordinate sferiche

Possiamo definire un altro tipo di coordinate molto utili in tre dimensioni, ovvero le coordinate sferiche. In questo caso identifichiamo il singolo punto sempre tramite la sua distanza dall'origine ρ e tramite gli angoli che esso crea con l'asse x (che indicheremo sempre con φ) e l'angolo che crea con l'asse z (che indicheremo con θ). Graficamente



Esercizio 6.11. Studiare l'integrabilità di $f(x, y) = (x^6y^5 + x^3y^2 + 2x + 2y + 2)e^{-x^2-y^2}$ su \mathbb{R}^2 e, in tal caso, calcolarne l'integrale

Dimostrazione. Osserviamo che

$$\int_{\mathbb{R}^2} |x^n y^m| e^{-x^2-y^2} \ll \int_{\mathbb{R}^2} (x^2 + y^2)^{\frac{n+m}{2}} = \int_{\mathbb{R}^2} \rho^{n+m+1} e^{-\rho^2} d\rho d\theta < +\infty$$

adesso osserviamo che

$$\int_{\mathbb{R}^2} x^n y^m e^{-x^2-y^2} dx dy = \int_{\mathbb{R}} x^n e^{-x^2} dx \int_{\mathbb{R}} y^m e^{-y^2} dy = 0$$

siccome n e m sono sempre dispari (se fossero stati entrambi pari non sarebbe stato così), pertanto l'unico termine che sopravvive è

$$\int_{-\infty}^{+\infty} 2e^{-x^2-y^2} = 2\pi$$

□

6.5 Misura di Hausdroff 2–dimensionale

Possiamo anche dare la nozione di area per le superfici immerse in \mathbb{R}^3 : per esempio se fossimo interessati a studiare la superficie di una sfera o di un cilindro, la misura di Lebesgue m_3 darebbe che il valore di essa è 0, siccome essa misura il volume 3–dimensionale delle superfici, che in questo caso risulta essere pari a 0 siccome lo "spessore" della superficie occupata da un cilindro ha ovviamente volume nullo.

Definizione 6.18 (misura di Hausdroff 2–dimensionale). Sia $A \subseteq \mathbb{R}^3$, allora l'area di tale insieme è

$$\mathcal{H}^2(A) = \sup_{\delta>0} \inf \left\{ \sum_{j=0}^{+\infty} \frac{\pi}{4} (\text{diam } E_j)^2 \mid A \subseteq \bigcup_{j=0}^{+\infty} E_j \right\}$$

per tutti i possibili ricopimenti $\{E_i\}_i$ per cui $\sup_{i \geq 0} \text{diam } E_i < \delta$, dove $\text{diam } E = \sup_{x,y \in E} |x - y|$

Poniamo $\sigma = \mathcal{H}^2$ = misura 2–dimensionale di E .

Si può mostrare che la teoria dell'integrazione rispetto a σ si costruisce esattamente come con la misura di Lebesgue.

Teorema 6.14 (teorema A1).

- ① gli insiemi chiusi, la loro unione e intersezione (finita o numerabile) e i loro complementari (quindi anche gli insiemi aperti) sono σ –misurabili;
- ② le funzioni continue sono σ –misurabili

Definizione 6.19 (jacobiano di una funzione non iniettiva). Sia $L : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ lineare con matrice

$A \in \mathbb{R}^{3 \times 2}$, $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix}$ e posto $M_{ij}(A) = \det \begin{bmatrix} a_{i1} & a_{i2} \\ a_{j1} & a_{j2} \end{bmatrix}$ definiamo il jacobiano di L come

$$JL = \sqrt{M_{12}(A)^2 + M_{13}(A)^2 + M_{23}(A)^2}$$

usando il prodotto vettoriale è possibile mostrare che $JL = |Le_1 \times Le_2|$. Questo si può, naturalmente, applicare ad ogni funzione $\phi \in C^1(\Omega, \mathbb{R}^3)$, Ω aperto $\subseteq \mathbb{R}^2$ con ϕ dipende dalla variabile $u = (u_1, u_2) \in \Omega$. Fissando un punto $w \in \Omega$ possiamo definire

$$J\phi(w) = J(d\phi(w)) = |\partial_{e_1}\phi(w) \times \partial_{e_2}\phi(w)|$$

Teorema 6.15 (teorema H2). Sia $\Omega \subseteq \mathbb{R}^2$ aperto e $\Phi \in C^1(\Omega, \mathbb{R}^3)$ con Φ iniettiva, $E \in \mathfrak{M}(\mathbb{R}^2)$, $E \subseteq \Omega$ (ovvero E è misurabile secondo Lebesgue e contenuto in Ω), allora abbiamo

- ① $\Phi(E) \in \mathfrak{M}_2(\mathbb{R}^3)$ ovvero è σ –misurabile e

$$\sigma(\Phi(E)) = \int_E J\Phi(u) du$$

- ② se $f : \Phi(E) \rightarrow \bar{\mathbb{R}}$ σ –misurabile e $\int_{\Phi(E)} |f| d\sigma < +\infty \implies$

$$\int_{\Phi(E)} f d\sigma = \int_E f(\Phi(u)) J\Phi(u) du$$

che prende il nome di integrale di area.

naturalmente si ottiene questo banale corollario se f è continua (in quanto misurabile)

Corollario 6.15.1. Siano $\Omega \subseteq \mathbb{R}^2$ aperto e $\Phi \in C^1(\Omega, \mathbb{R}^3)$ con Φ iniettiva, $E \in \mathfrak{M}(\mathbb{R}^2)$, $E \subseteq \Omega$ e $f : \Phi(E) \rightarrow \mathbb{R}$ è continua con $\int_{\Phi(E)} |f| d\sigma < +\infty$ allora

$$\int_{\Phi(E)} f d\sigma = \int_E f(\Phi(u)) J\Phi(u) du$$

Esempio 6.4 (calcolo della superficie della sfera). Calcolare l'area della superficie di una sfera

Svolgimento. Sia data una superficie sferica di raggio $R > 0$ e centro $x \in \mathbb{R}^3$, allora

$$\Phi(\varphi, \theta) = x + R(\sin \theta \cos \varphi, \sin \theta \sin \varphi, \cos \theta)$$

al variare di $\varphi \in (0, 2\pi)$ e $\theta \in (0, \pi)$. Osserviamo che

$$\sigma(x + R\mathbb{S}^2 \setminus \Phi((0, 2\pi) \times (0, \pi))) = 0$$

dunque la parte di superficie non "generata" da Φ (per garantire l'iniettività della trasformazione) non contribuisce all'integrale.

Procediamo, dunque, nel calcolare il jacobiano di Φ :

$$\partial_\varphi \Phi = R [\sin \theta (-\sin \varphi), \sin \theta \cos \varphi, 0] \quad \partial_\theta \Phi = R [\cos \theta \cos \varphi, \cos \theta \sin \varphi, \sin \theta]$$

dunque

$$J\Phi(\varphi, \theta) = R^2 \sqrt{(\sin^2 \theta \cos \varphi)^2 + (\sin^2 \theta \sin \varphi)^2 + (\cos \theta \sin \theta)^2} = R^2 \sin \theta$$

pertanto

$$\sigma(x + R\mathbb{S}^2) = \int_{(0, 2\pi) \times (0, \pi)} J\Phi(\varphi, \theta) d\varphi d\theta = R^2 \int_0^{2\pi} \int_0^\pi \sin \theta d\theta d\varphi = 4\pi R^2$$

□

Esempio 6.5 (area del toro di raggi $0 < r < R$). Calcolare l'area del toro di raggi $0 < r < R$

Svolgimento. Possiamo parametrizzare la superficie di un toro nella seguente maniera:

$$\Phi(\varphi, \theta) = ((R + r \cos \theta) \cos \varphi \quad (R + r \cos \theta) \sin \varphi \quad r \sin \theta)$$

dunque

$$\partial_\varphi \Phi = (-\sin \varphi (R + r \cos \theta) \quad \cos \varphi (R + r \cos \theta) \quad 0)$$

$$\partial_\theta \Phi = (-r \sin \theta \cos \varphi \quad -r \sin \theta \sin \varphi \quad r \cos \theta)$$

i quali ci consentono di concludere che

$$\begin{aligned} J\Phi(\varphi, \theta) &= r(R + r \cos \theta) \implies \sigma(T) = \int_{(0, 2\pi) \times (0, \pi)} J\Phi d\varphi d\theta = r \int_0^{2\pi} \int_0^\pi (R + r \cos \theta) d\varphi d\theta = \\ &= 2\pi r \int_0^{2\pi} R + r \cos \theta d\theta = 2\pi r 2\pi R = 4\pi^2 r R \end{aligned}$$

□

Esercizio 6.12. Calcolare l'integrale di superficie

$$\int_T \frac{|xz|}{(x^2 + y^2)^{\frac{3}{2}}} d\sigma$$

dove $T = \varphi((0, 2\pi)^2)$ toro di raggi $0 < r < R$

Svolgimento. Dal precedente esercizio, sappiamo è possibile parametrizzare come

$$\Phi(\varphi, \theta) = [(R + r \cos \theta) \cos \varphi \quad (R + r \cos \theta) \sin \theta \quad r \sin \theta]$$

pertanto

$$\begin{aligned} & \int_{(0,2\pi)^2} \frac{(R + r \cos \theta) |\cos \varphi| |r \sin \theta| r(R + r \cos \theta)}{R + r \cos \theta} d\varphi d\theta = r^2 \int_{(0,2\pi)^2} \frac{|\cos \varphi| |\sin \theta|}{R + r \cos \theta} d\varphi d\theta = \\ & = r^2 \int_0^{2\pi} |\cos \varphi| d\varphi \int_0^{2\pi} \frac{|\sin \theta|}{R + r \cos \theta} d\theta = 4r \int_0^\pi \frac{\sin \theta}{R + r \cos \theta} d\theta - 4r \int_\pi^{2\pi} \frac{\sin \theta}{R + r \cos \theta} d\theta \\ & = 8r \log \frac{R + r}{R - r} \end{aligned}$$

□

Adesso abbiamo tutti gli strumenti per determinare l'area ottenuta facendo ruotare una curva di 2π attorno ad un asse: supponiamo di avere una curva $\gamma : [a, b] \rightarrow \mathbb{R}^2$ C^1 tratti tale che $\gamma_1(t) > 0 \forall t \in [a, b]$ e di voler determinare l'area della superficie Σ ottenuta ruotando γ attorno all'asse z .

Parametrizziamo la nostra superficie come

$$\Phi(\varphi, t) = (\gamma_1(t) \cos \varphi, \quad \gamma_1(t) \sin \varphi, \quad \gamma_2(t)) \quad \text{con } t \in (a, b), \varphi \in (0, 2\pi)$$

pertanto

$$D\Phi = \begin{pmatrix} -\gamma_1(t) \sin \varphi & \dot{\gamma}_1(t) \cos \varphi \\ \gamma_1(t) \cos \varphi & \dot{\gamma}_1(t) \sin \varphi \\ 0 & \dot{\gamma}(t) \end{pmatrix}$$

dunque

$$J\varphi = \sqrt{\gamma_1^2 \dot{\gamma}_1^2 + \dot{\gamma}_2^2 \dot{\gamma}_1^2 \sin^2 \varphi + \dot{\gamma}_2^2 \gamma_1^2 \cos^2 \varphi} = \gamma_1 \sqrt{\dot{\gamma}_1^2 + \dot{\gamma}_2^2} = \gamma_1 |\dot{\gamma}(t)|$$

pertanto

$$\sigma(\Sigma) = \int_{(a,b) \times (0,2\pi)} d\sigma = \int_0^{2\pi} \int_a^b \gamma_1(t) |\dot{\gamma}(t)| dt d\varphi = 2\pi \int_a^b \gamma_1(t) |\dot{\gamma}(t)| dt = 2\pi L \int_\gamma x dl$$

dove $L = \int_a^b |\dot{\gamma}(t)| dt$ è la lunghezza della curva. Dunque abbiamo mostrato il seguente teorema:

Teorema 6.16 (primo teorema di Guldino).

$$\sigma(\Sigma) = 2\pi d_\gamma l(\gamma)$$

dove poniamo $d_\gamma = \frac{\int_a^b \gamma_1(t) |\dot{\gamma}(t)| dt}{\int_a^b |\dot{\gamma}(t)| dt}$.

6.5.1 Area dei grafici

Consideriamo $u \in C^1(\Omega, \mathbb{R})$, $\Omega \subseteq \mathbb{R}^2$ aperto e la funzione grafico associata $\varphi(x, y) = (x, y, u(x, y)) \implies \varphi \in C^1(\Omega, \mathbb{R}^3)$ (in virtù del teorema del differenziale totale) ed è iniettiva. In virtù del teorema del Dini, possiamo definire una funzione $\Sigma = f(\Omega)$, $\sigma(\text{gr } u) = \sigma(\Sigma) = \int_\Omega J\varphi dx dy$. In virtù di quanto detto in precedenza, possiamo dunque definire il jacobiano di φ come

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ u_x & u_y \end{pmatrix} \implies (J\varphi)^2 = 1 + u_x^2 + u_y^2 \implies J\varphi = \sqrt{1 + |\nabla u|^2}$$

dunque possiamo concludere che

$$\textcircled{1} \quad \sigma(\Sigma) = \int_{\Omega} \sqrt{1 + |\nabla u|^2} dm_2$$

\textcircled{2} se $f : \Sigma \rightarrow \mathbb{R}$ continua allora

$$\int_{\Sigma} f d\sigma = \int_{\Omega} f(x, y, u(x, y)) \sqrt{1 + |\nabla u|^2} dx dy$$

Oss.: Possiamo dunque vedere $\sqrt{1 + |\nabla u|^2}$ come l'infinitesimo elemento d'area.

Oss.: Possiamo formulare questo teorema in altri modi analoghi, considerando dei grafici rispetto alle variabili (x, z) e (y, z) .

Esempio 6.6. Consideriamo $S = \{(x, y, z) \in \mathbb{R}^3 : z^2 + y^2 \leq 9, x - 2y + z = 1\}$. Determinare

$$\textcircled{1} \quad \sigma(S)$$

$$\textcircled{2} \quad \text{studiare l'integrabilità di } \frac{x+y+z}{(y^2+z^2)^{\frac{7}{8}}} \text{ su } S \text{ ed, in tal caso, calcolare}$$

$$\int_S \frac{x+y+z}{(y^2+z^2)^{\frac{7}{8}}} d\sigma$$

Svolgimento. Possiamo rappresentare la nostra superficie come

$$x = 1 + 2y - z$$

dunque definiamo la seguente funzione $\varphi(y, z) = (1 + 2y - z, y, z)$ da cui possiamo determinare il jacobiano:

$$J\varphi = \begin{pmatrix} 2 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \implies J\varphi = \sqrt{1 + 4 + 1} = \sqrt{6}$$

dunque

$$\sigma(S) = \int_S d\sigma = \int_{\mathbb{B}(0,3)} \sqrt{6} dm_2 = 9\pi \cdot \sqrt{6}$$

Per quanto riguarda l'integrabilità possiamo osservare che

$$\int_S \frac{|x+y+z|}{(y^2+z^2)^{\frac{7}{8}}} d\sigma = \int_S \frac{|1+3y|}{(y^2+z^2)^{\frac{7}{8}}} d\sigma \leq \int_S \frac{10}{(y^2+z^2)^{\frac{7}{8}}} d\sigma = 10\sqrt{6} \int_{\mathbb{B}(0,3)} \frac{1}{|(y,z)|^{\frac{7}{4}}} < +\infty$$

in virtù dell'esercizio in cui avevamo mostrato che $\int_{\mathbb{B}(0,r)} \frac{1}{|x|^\alpha} < +\infty \iff 0 < \alpha < n$. L'integrale è un banale conto \square

6.6 Teorema di Gauss-Green

Vogliamo adesso ricavare il teorema di Gauss-Green, il quale ci consente di ricavare una forma

Teorema 6.17 (curva semplice). $\varphi : [a, b] \rightarrow \mathbb{R}^2$ è semplice se $\gamma|_{[a,b]}$ e $\gamma'|_{[a,b]}$ sono iniettive.

Oss.: Una curva semplice può essere chiusa: se $\gamma(a) = \gamma(b)$ allora $\gamma|_{[a,b]}$ e $\gamma'|_{[a,b]}$.

Oss.: Se $\gamma : [a, b] \rightarrow \mathbb{R}^2$, $\gamma|_{[a,b]}$ è iniettiva allora non può essere semplice: infatti potrebbe esistere un punto $c \in (a, b)$ tale che $\gamma(b) = \gamma(c)$ allora $\gamma|_{[a,b]}$ e $\gamma'|_{[a,b]}$.

Definizione 6.20 (aperto regolare). Sia $\Omega \subseteq \mathbb{R}^2$ un insieme aperto, connesso per archi e limitato. Se supponiamo che $\text{Fr}\Omega$ sia un'unione di curve chiuse, semplici e C^1_{tratti} diremo che è un aperto regolare e denoteremo la sua frontiera con il simbolo $\partial\Omega$, detto *bordo di Ω* .

Definizione 6.21 (orientazione del bordo regolare). Sia Ω un insieme regolare e siano $\gamma_1, \dots, \gamma_N$ le curve semplici, chiuse e C^1_{tratti} le cui immagini costituiscono $\partial\Omega$. Se tutte le curve γ_i percorrono $\partial\Omega$ lasciando Ω alla sua sinistra, diremo che orientano positivamente Ω , altrimenti lo orientano negativamente.

Indicheremo con $\partial^+\Omega$ il bordo di Ω positivamente orientato, mentre con $\partial^-\Omega$ il bordo di Ω negativamente orientato.

Tramite questa definizione, possiamo dare un senso all'integrale curvilineo calcolato sul bordo positivamente orientato

Definizione 6.22 (integrale curvilineo sul bordo positivamente orientato). Sia $\Omega \subseteq \mathbb{R}^2$ un aperto regolare e sia $\partial^+\Omega$ il bordo positivamente orientato dalle curve $\gamma_1, \dots, \gamma_N$ C^1_{tratti} , semplici e chiusi; ove $\gamma_i : [a_i, b_i] \rightarrow \partial\Omega \forall i \in \{1, \dots, N\}$. Sia $\omega \in C^0(\mathcal{U}, (\mathbb{R}^2)^*)$ una 1-forma differenziale ove $\partial\Omega \subseteq \mathcal{U} \subseteq \mathbb{R}^2$.

Definiamo allora

$$\int_{\partial^+\Omega} = \sum_{i=1}^N \int_{\gamma_i} \omega = \sum_{i=1}^N \int_{a_i}^{b_i} \omega(\gamma(t))(\dot{\gamma}(t))dt$$

Se $\Gamma_i : [\alpha_i, \beta_i] \rightarrow \partial\Omega$ sono delle curve C^1_{tratti} , semplice e chiuse che orientano negativamente il bordo $\partial\Omega$ definiamo

$$\int_{\partial^-\Omega} = \sum_{i=1}^N \int_{\Gamma_i} \omega = \sum_{i=1}^N \int_{\alpha_i}^{\beta_i} \omega(\Gamma(t))(\dot{\Gamma}(t))dt$$

Oss.: Gli integrali $\int_{\partial^+\Omega} \omega$, $\int_{\partial^-\Omega} \omega$ non dipendono dalla scelta della parametrizzazione delle curve

chiuse, semplici e C^1_{tratti} che parametrizzano il bordo, ma solo dal loro verso di percorrenza. Questo è una proprietà dell'integrale curvilineo di seconda specie (teorema di invarianza).

Si può inoltre osservare che

$$\int_{\partial^+\Omega} \omega = - \int_{\partial^-\Omega} \omega$$

Teorema 6.18 (teorema di Gauss-Green nel piano). Sia $\Omega \subseteq \mathbb{R}^2$ un aperto regolare e $F \in C^1(\bar{\Omega}, \mathbb{R}^2)$.

Allora

$$\int_{\Omega} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dx dy = \int_{\partial^+\Omega} F_1 dx + F_2 dy$$

Dimostrazione euristica. Vediamo l'integrazione lungo il bordo $\partial\Omega$ come orientata " $dx \wedge dy$:

$$\begin{aligned} \int_{\Omega} \left(\frac{\partial F_2}{\partial y} - \frac{\partial F_1}{\partial x} \right) dx dy &= \int_{\Omega} \frac{\partial F_2}{\partial x} dx \wedge dy - \int_{\Omega} \frac{\partial F_1}{\partial y} dx \wedge dy = \\ &= \int_{\partial^+\Omega} F_2 dy + \int_{\partial^-\Omega} \frac{\partial F_1}{\partial y} dy \wedge dx = \int_{\partial^+\Omega} F_2 dy + \int_{\partial^+\Omega} F_1 dx \end{aligned}$$

dove la penultima eguaglianza si è ottenuta osservando che dobbiamo "invertire" $dx \wedge dy$ ma geometricamente percorrere un bordo positivamente o negativamente cambia solamente il segno, dunque $dx \wedge dy = -dy \wedge dx$ (rimando alle appendici per una trattazione più accurata del prodotto wedge). \square

Esempio 6.7. Sia $\Omega = (-1, 1) \times (1, 2)$ e sia $F = (x^2 - y, x^2 - y)$. Sia inoltre γ una curva semplice, chiusa e C^1 che percorre $\partial\Omega$ in senso orario. Determinare

$$\int_{\gamma} F$$

Dimostrazione. Il valore dell'integrale lungo il bordo è, ovviamente, indipendente dalla parametrizzazione. Conseguentemente potremmo pensare di applicare il teorema di Gauss-Green sulla curva semplice, chiusa e C^1 più "semplice" possibile, ovvero una curva di forma rettangolare. Pertanto

$$\begin{aligned} \int_{\gamma} F &= - \int_{\partial^+ \Omega} F = - \int_{\partial^+ \Omega} F_1 dx + F_2 dy = - \int_{\Omega} \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dx dy = \\ &= \int_{-1}^1 dx \int_1^2 (-1 - 2x) dy = \int_{-1}^1 (-1 - 2x) dx = -2 - 2 \int_{-1}^1 x dx = -2 \implies \int_{\gamma} F = -2 \end{aligned}$$

□

Esempio 6.8. Sia $\gamma(t) = \left(\cos t, \frac{\sin t}{\sqrt{2}} \right)$, $\gamma : [0, 2\pi] \rightarrow \mathbb{R}^2$. Calcolare $\int_{\gamma} \eta_0$, dove $\eta_0 = -\frac{y}{x^2+y^2} dx + \frac{x}{x^2+y^2} dy$.

Dimostrazione.

$$\int_0^{2\pi} \left(\frac{1}{\sqrt{2}} \frac{\sin^2 t}{\cos^2 t + \frac{1}{2} \sin^2 t} + \frac{1}{\sqrt{2}} \frac{\cos^2 t}{\cos^2 t + \frac{1}{2} \sin^2 t} \right) dt = \frac{1}{\sqrt{2}} \int_0^{2\pi} \frac{1}{\cos^2 t + \frac{1}{2} \sin^2 t} dt$$

Osserviamo che, nonostante questo integrale non sia risolvibile, abbiamo un grande problema: la funzione non è definita su

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

che è un insieme aperto, regolare e orientato positivamente da γ , tuttavia η_0 non è definita in $(0, 0)$ e non è possibile estenderla con continuità.

Possiamo però restringere l'insieme a

$$H = E \setminus \mathbb{B}(0, \frac{1}{2})$$

che è un aperto regolare, dove $\eta_0 \in C^1(\mathcal{U}, (\mathbb{R}^2)^*)$, $\bar{H} \subseteq U \subseteq \mathbb{R}^2$ e posso applicare Gauss-Green ad η_0 sull'aperto H : definiamo

$$F_1 = -\frac{y}{x^2 + y^2} \quad F_2 = \frac{x}{x^2 + y^2}$$

e osserviamo che, come avevamo visto già in precedenza, η_0 è chiusa $\implies \frac{\partial F_2}{\partial x} = \frac{\partial F_1}{\partial y}$, dunque

$$\int_H \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) = 0 = \int_{\partial^+ H} F_1 dx + F_2 dy \implies \int_{\partial^+ H} \eta_0 = 0 = \int_{\gamma} \eta_0 + \int_c \eta_0$$

dove γ è la curva che orienta positivamente il bordo, mentre la curva c orienta positivamente il bordo interno (dovuto alla rimozione della palla centrata in 0 e di raggio 2). Ma allora

$$\int_{\gamma} \eta_0 = - \int_c \eta_0$$

possiamo parametrizzare $\tilde{c}(t) = \frac{1}{2} (\cos t, \sin t)$ che si muove in senso antiorario, da cui ricaviamo che

$$\int_{\tilde{c}} \eta_0 = 2\pi \implies \int_{\gamma} \eta_0 = - \int_c \eta_0 = \int_{\tilde{c}} \eta_0$$

□

6.6.1 Forme differenziali d'area

Proposizione 6.13. Sia $H \subseteq \mathbb{R}^2$ un aperto regolare. Allora

$$m_2(H) = \int_H 1 dx dy = \int_{\partial^+ H} x dy = - \int_{\partial^+ H} y dx = \int_{\partial^+ H} \frac{xdy - ydx}{2}$$

Dimostrazione. Preso $F_1 = y$ e $F_2 = 0$ avremo che

$$\int_H 1 dx dy = - \int_{\partial^+ H} y dx$$

e preso $F_1 = 0$ e $F_2 = x$ avremo che

$$\int_H 1 dx dy = \int_{\partial^+ H} x dy$$

dunque

$$m_2(H) = \int_H 1 dx dy = \int_{\partial^+ H} \frac{xdy - ydx}{2}$$

□

Un'operazione molto utile che possiamo fare è il cosiddetto *pull-back* di una forma differenziale rispetto a delle coordinate: questo è particolarmente interessante per le forme differenziali di un'area, le quali rappresentano, in un certo senso, l'elemento infinitesimo di area.

Definizione 6.23 (pull-back di una 1-forma differenziale). Sia T una funzione differenziabile tale che $T : \mathcal{V} \rightarrow \mathcal{U}$ e ω una 1-forma differenziale tale che $\omega \in C^0(\mathcal{U}, (\mathbb{R}^n)^*) \implies$

$$T^* \omega \in C^0(\mathcal{V}, (\mathbb{R}^n)^*)$$

Oss.: Il nome antisonante deriva proprio dal fatto che sto tirando indietro la 1-forma differenziale da \mathcal{U} a \mathcal{V}

Un'interessante proprietà del pull-back è la seguente:

$$\int_{\gamma} \omega = \int_{T^{-1} \circ \gamma} T^* \omega$$

Esempio 6.9 (forma d'area in coordinate polari). Sia $T(\rho, \varphi) = (\rho \cos \varphi, \rho \sin \varphi)$ con $\rho > 0, 0 < \varphi < 2\pi$. Vogliamo determinare $T^* \omega$ dove, naturalmente, $\omega = \frac{xdy - ydx}{2}$. Il calcolo si ottiene, banalmente, sostituendo al posto di x e y le nuove "espressioni", calcolandone naturalmente il differenziale

$$\begin{aligned} \frac{xdy - ydx}{2} &= \frac{\rho \cos \varphi d(\rho \sin \varphi) - \rho \sin \varphi d(\rho \cos \varphi)}{2} = \\ &\frac{1}{2} (\rho \cos \varphi \sin \varphi d\rho + \rho^2 \cos^2 \varphi d\varphi - \rho \sin \varphi \cos \varphi d\rho + \rho^2 \sin^2 \varphi d\varphi) = \frac{1}{2} \rho^2 d\varphi \end{aligned}$$

dunque $T^* \omega = \frac{1}{2} \rho^2 d\varphi$ è la 1-forma differenziale d'area in coordinate polari, definita in $(0, +\infty) \times (0, 2\pi)$. Siamo adesso autorizzati a dire che

$$d\sigma = \frac{1}{2} \rho^2 d\varphi$$

Esempio 6.10 (area di una regione tramite la 1-forma d'area in coordinate polari). Consideriamo $E = \{(\rho \cos \varphi, \rho \sin \varphi) : \alpha \leq \varphi \leq \beta, 0 \leq \rho \leq r(\varphi)\}$ e consideriamo $r(\varphi) : [\alpha, \beta] \rightarrow [0; +\infty)$ continua con $0 \leq \alpha < \beta \leq 2\pi$. Calcolare la forma d'area

Svolgimento. Osserviamo che E è un aperto regolare, il cui bordo può essere ottenuto tramite due segmenti radiali γ_1, γ_3 e una curva regolare γ_2 che descrive il bordo con lunghezza variabile.

$$m_2(E) = \int_{\partial^+ E} \omega = \int_{\partial^+ E} \frac{-ydx + xdy}{2} = \int_{\gamma_1} \frac{-ydx + xdy}{2} + \int_{\gamma_2} \frac{-ydx + xdy}{2} + \int_{\gamma_3} \frac{-ydx + xdy}{2}$$

e possiamo mostrare che $\int_{\gamma_1} \frac{-ydx + xdy}{2} = 0$ siccome

$$\int_{\gamma_1} \frac{-ydx + xdy}{2} = \frac{1}{2} \int_{T^{-1} \circ \gamma_1} \rho^2 d\varphi = 0$$

siccome $\varphi = \text{costante}$. Similmente per γ_3 , da cui possiamo concludere che

$$m_2(E) = \frac{1}{2} \int_{T^{-1} \circ \gamma_2} \rho^2 d\varphi \stackrel{\Gamma = T^{-1} \circ \gamma}{=} \frac{1}{2} \int_{\Gamma} \rho^2 d\varphi$$

□

Esercizio 6.13. Scrivere in coordinate polari $x dy$ e $-y dx$

Dimostrazione. E' banale dalla formula del pull-back: se $\omega(x, y) = x dy$ allora

$$\omega(T(\rho, \varphi)) = \rho \cos \varphi d(\rho \sin \varphi) = \rho \sin \varphi \cos \varphi d\rho + \rho^2 \cos^2 \varphi d\varphi$$

mentre per $\omega(x, y) = -y dx$ allora

$$\omega(T(\rho, \varphi)) = -\rho \sin \varphi d\rho + \rho^2 \sin^2 \varphi d\varphi$$

□

Esercizio 6.14. Calcolare l'area dell'ellisse di equazione $\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$ con $a, b > 0$ scegliendo un'opportuna forma d'area

Svolgimento. Possiamo sfruttare il teorema di Gauss-Green: possiamo parametrizzare il bordo dell'ellisse come

$$\begin{cases} x = a \cos \varphi \\ y = b \sin \varphi \end{cases} \implies \frac{x^2}{a^2} + \frac{y^2}{b^2} = \cos^2 \varphi + \sin^2 \varphi = 1$$

Usando la forma d'area unito al teorema di Gauss-Green abbiamo che

$$\begin{aligned} \int_{\partial^+ \Omega} x dy &= \int_E dx dy = \int_{\gamma} x dy = \int_0^{2\pi} ab \cos^2 \varphi = ab \int_0^{2\pi} \frac{1 + \cos 2\varphi}{2} d\varphi = \\ &= ab \left(\int_0^{2\pi} \frac{1}{2} d\varphi + \frac{1}{2} \int_0^{2\pi} \cos 2\varphi d\varphi \right) = \pi ab \end{aligned}$$

□

Oss.: E' possibile ricavare la 1-forma d'area senza l'utilizzo del teorema di Gauss-Green nella seguente maniera: posto l'insieme

$$E = \{(\rho \cos \varphi, \rho \sin \varphi) : \varphi \in (\alpha, \beta) \text{ e } 0 < \rho < r(\varphi)\}$$

dove abbiamo naturalmente che $0 \leq \alpha < \beta \leq 2\pi$ e $r : [\alpha, \beta] \rightarrow [0; +\infty)$ è continua. Vediamo E come l'immagine della trasformazione $T(\rho, \varphi) = (\rho \cos \varphi, \rho \sin \varphi)$ dell'insieme

$$F = \{(\rho, \varphi) : 0 < \rho < r(\varphi), \alpha < \varphi < \beta\}$$

pertanto

$$\begin{aligned} m_2(E) &= \int_E dx dy = \int_{T^{-1}(E)} \rho d\rho d\varphi = \int_F \rho d\rho d\varphi = \int_\alpha^\beta \int_0^{r(\varphi)} \rho d\rho \\ \implies m_2(E) &= \frac{1}{2} \int_\alpha^\beta r(\varphi)^2 d\varphi \end{aligned}$$

Dunque, sempre ponendo $\Gamma = T^{-1} \circ \gamma$ avremo che

$$\int_\Gamma \omega = \int_\alpha^\beta \omega(\Gamma(\varphi))(\dot{\Gamma}(\varphi)) = \frac{1}{2} \int_\alpha^\beta r(\varphi) d\varphi$$

si osservi che la curva dà un'area positiva se si muove in senso antiorario (poiché l'angolo φ cresce).

Esercizio 6.15. Consideriamo $F(x, y) = (e^{x^2} - xy^2, e^{-\sin e^y} + x^2 y)$ e l'aperto regolare

$$\Omega = \{(x, y) \in \mathbb{R}^2 : y^2 - x^2 < 1, 0 < x < \sqrt{3}, y > 0\}$$

Sia Γ una curva C^1_{tratti} , chiusa e semplice che percorre $\partial\Omega$ in senso orario. Calcoliamo $\int_\Gamma F$

Svolgimento. Usiamo il teorema di Gauss-Green

$$\begin{aligned} \int_\Gamma F &= - \int_{\gamma_1 + \gamma_2 + \gamma_3 + \gamma_4} F = \int_{\partial^+ \Omega} F = - \int_\Omega \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y} \right) dx dy = \\ &= - \int_\Omega (2yx + 2yx) dx dy = -4 \int_\Omega yz dx dy = -4 \int_0^{\sqrt{3}} dx + \int_0^{\sqrt{1+x^2}} yx dx dy = \\ &= -4 \int_0^{\sqrt{3}} x \left(\frac{1+x^2}{2} \right) dx = - \left(\frac{x^2}{2} + \frac{x^4}{4} \right) \Big|_{x=0}^{x=\sqrt{3}} = -\frac{15}{2} \end{aligned}$$

□

6.7 Flusso e teorema della divergenza

Vogliamo definire il concetto di flusso attraverso una superficie, nozione che ha molte applicazioni in molte discipline scientifiche. Dobbiamo, tuttavia, stabilire un modo per orientare una superficie

Definizione 6.24 (superficie regolare orientata da ν). Sia $\Sigma \subseteq \mathbb{R}^3$ una superficie regolare, diremo che Σ è orientata da $\nu : \Sigma \mapsto \mathbb{S}^2$ è continua e

$$\nu(x) \perp T_x \Sigma \quad \forall x \in \Sigma.$$

Diremo che $\nu(x)$ è un campo normale a Σ

Definizione 6.25 (flusso di F attraverso Φ). Se $F \in C^0(\mathcal{U}, \mathbb{R}^3)$, $\mathcal{U} \subseteq \mathbb{R}^3$ aperto con $\Sigma \subseteq \mathcal{U}$ e $\int_{\Sigma} |F| d\sigma < +\infty$, definiamo il flusso di F attraverso ν come

$$\Phi(F, \Sigma, \nu) = \int_{\Sigma} \langle F, \nu \rangle d\sigma$$

Oss.: Esistono superfici regolari non orientabili: un esempio è il nastro di Möbius

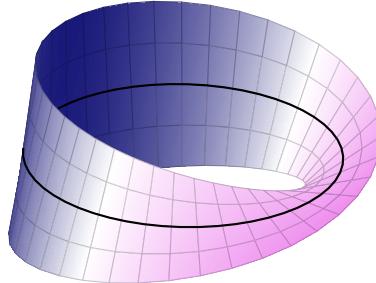


Figura 6.1: Nastro di Möbius. La circonferenza disegnata serve proprio per far vedere che un giro attorno a questo nastro inverte l'orientazione: se io prendessi un versore normale "puntato" verso l'esterno, allora dopo un giro questo sarebbe invertito, dunque ci sarebbe una discontinuità e ν non sarebbe continua, come riesco dalla definizione di superficie orientabile

Esempio 6.11 (esempio di calcolo del flusso attraverso una superficie).

Sia $T = \{(x, y, z) \in \mathbb{R}^3 : x, y, z > 0, x + \frac{y}{2} + \frac{z}{4} = 1\}$ e $F(x, y, z) = (y, x, \frac{z}{2} + y - 2)$. Calcolare il flusso attraverso questa superficie

Svolgimento. Osserviamo che, essendo il triangolo un piano "confinato" a $x, y, z > 0$, possiamo facilmente trovare il versore normale prendendo il vettore $(1, \frac{1}{2}, \frac{1}{4})$. Possiamo altrimenti porre $g(x, y, z) = x + \frac{y}{2} + \frac{z}{4}$ e definire il nostro triangolo tramite la funzione definente g : lo spazio tangente ad ogni punto x_0 di questa superficie sarà dato da $\nabla g(x_0) = (1, \frac{1}{2}, \frac{1}{4})$.

Splicitiamo la x :

$$x = 1 - \frac{y}{2} - \frac{z}{4} > 0 \text{ siccome } 0 < x = 1 - \frac{y}{2} - \frac{z}{4} = \varphi(y, z)$$

La normale a questo piano è dato naturalmente dal gradiente che corrisponde proprio a

$$\nabla g(x, y, z) = \begin{pmatrix} \partial_y \varphi & \partial_z \varphi \\ 1 & 0 \\ 0 & 1 \end{pmatrix} \implies |\nabla g| = \sqrt{(\partial_y \varphi)^2 + (\partial_z \varphi)^2 + 1} = \sqrt{1 + |\nabla \varphi|^2}$$

dunque il vettore normale è proprio

$$\nu = \frac{(1, \frac{1}{2}, \frac{1}{4})}{\sqrt{1 + (\frac{1}{2})^2 + (\frac{1}{4})^2}} \implies \Phi(F, \Sigma, \nu) = \int_H \langle F, \frac{\nabla g}{\sqrt{1 + |\nabla \varphi|^2}} \rangle \sqrt{1 + |\nabla \varphi|^2} dy dz$$

dove $H = \{(y, z) \in \mathbb{R}^2 : 0 < \frac{y}{2} + \frac{z}{4} < 1\}$. Dunque

$$\int_H \langle F, (1, \frac{1}{2}, \frac{1}{4}) \rangle dy dz = \int_0^{4-2y} y \int_0^{4-2y} dz = \frac{8}{3}$$

□

6.7.1 Frontiera regolare di un aperto

Definizione 6.26 (punto regolare di frontiera e derivata esterna).

Sia $\Omega \subseteq \mathbb{R}^3$ un insieme aperto. Diremo che $x \in \text{Fr } \Omega$ è un punto di frontiera regolare se

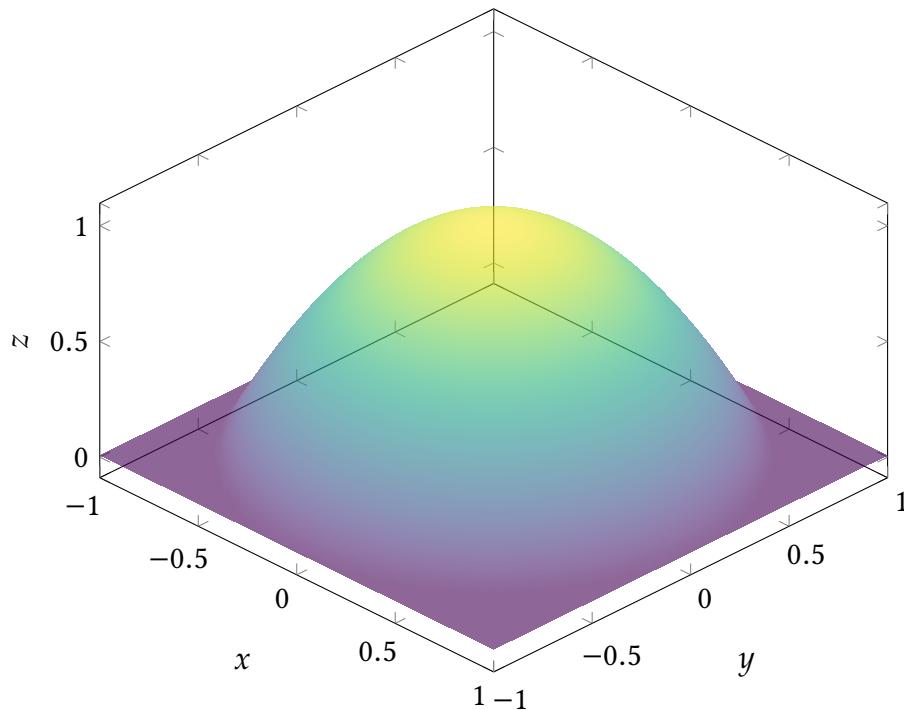
- ① $\exists O \subseteq \mathbb{R}^3$ aperto tale che $x \in O \cap \text{Fr } \Omega$
- ② $\exists g \in C^1(O)$ tale che $O \cap \Omega = \{y \in O : g(y) < 0\}$
- ③ $\nabla g(y) \neq 0 \quad \forall y \in O$

Denoteremo con $\partial_{\text{reg}} \Omega$ la frontiera regolare o bordo regolare di Ω . Diremo che g definisce una porzione di frontiera regolare $O \cap \text{Fr } \Omega$ e

$$\nu(y) = \frac{\nabla g(y)}{|\nabla g(y)|} \quad \forall y \in O \cap \text{Fr } \Omega$$

dove $\nu : \partial_{\text{reg}} \Omega \mapsto \mathbb{S}^2$ è detta normale esterna ad Ω in $x \in \partial_{\text{reg}} \Omega$.

Oss.: I punti regolari di frontiera sono quei punti che separano l'interno dall'esterno. Fare riferimento alla seguente immagine



Proposizione 6.14. Dato $\Omega \subseteq \mathbb{R}^3$ il sottoinsieme $\partial_{\text{reg}} \Omega$ di $\text{Fr } \Omega$ è un'unione al più numerabile di superfici regolari.

Idea della dimostrazione. Si osserva che $\partial_{\text{reg}} \Omega$ è un'aperto in $\text{Fr } \Omega$ rispetto la topologia (indotta) di sottospazio e che tale topologia ha una base numerabile di aperti, dunque per il teorema di Lindelöf ogni ricoprimento aperto ha un sottoricoprimento numerabile. Prendendo come aperti del ricoprimento di $\partial_{\text{reg}} \Omega$ gli $\{O_x : x \in \partial_{\text{reg}} \Omega\}$ dove $O_x \subseteq \mathbb{R}^3$ e $g : O_x \rightarrow \mathbb{R}$ definisce $\partial_{\text{reg}} \Omega$ intorno a x , dunque la tesi. \square

Proposizione 6.15. Se $\nu : \partial_{\text{reg}} \rightarrow \mathbb{S}^2$ è la normale esterna allora è continua

Dimostrazione.

$$\nu(y) = \frac{\nabla g(y)}{|\nabla g(y)|} \quad y \in O \cap \text{Fr } \Omega$$

è continua sulla porzione di frontiera $O \cap \text{Fr } \Omega$ siccome rapporto di funzioni continue \square

Presa una funzione $F : \mathbb{R}^3 \rightarrow \mathbb{R}^3$ possiamo allora definire il flusso, supponendo che $\int_{\partial_{\text{reg}}} |F| < +\infty$ (ovvero F è σ -integrale su ∂_{reg}), attraverso la frontiera regolare

$$\Phi(F, \partial_{\text{reg}} \Omega, \nu) = \int_{\partial_{\text{reg}} \Omega} \langle F, \nu \rangle d\sigma$$

Definizione 6.27 (divergenza). Sia $F \in C^1(\Omega, \mathbb{R}^3)$, $\Omega \subseteq \mathbb{R}^3$ aperto, allora definiamo la divergenza di F la funzione scalare

$$\text{div } F = \partial_x F_x + \partial_y F_y + \partial_z F_z$$

dove $F = (F_x, F_y, F_z)$

Teorema 6.19 (teorema della divergenza). Sia $\Omega \subseteq \mathbb{R}^3$ limitato e connesso per archi tale che

$$(1) \quad \sigma(\text{Fr } \Omega \setminus \partial_{\text{reg}}) = 0$$

$$(2) \quad \sigma(\partial_{\text{reg}}) < +\infty$$

allora se $F \in C^1(\mathcal{U}, \mathbb{R}^3)$, $\bar{\Omega} \subseteq \mathcal{U}$ dove $\mathcal{U} \subseteq \mathbb{R}^3$ è aperto vale la seguente formula

$$\int_{\Omega} \text{div } F dm_3 = \int_{\partial_{\text{reg}} \Omega} \langle F, \nu \rangle d\sigma = \Phi(F, \partial_{\text{reg}} \Omega, \nu)$$

Esempio 6.12 (calcolo del flusso tramite il teorema della divergenza). Calcolare $\Phi(F, \Sigma, \nu)$ dove $F(x, y, z) = (x, y, z)$ e $\Sigma = \{(x, y, z) \in \mathbb{R}^3 : 4x^2 + 4y^2 + z^2 = 4, x \leq 0\}$, dove questa è orientata dalla normale ν che è individuata dalla condizione $\nu(-1, 0, 0) = (-1, 0, 0)$.

Svolgimento. Osserviamo che sul semipiano $D = \Sigma \cap \{x = 0\}$ avremo che la normale $\eta = (1, 0, 0)$ mentre per quanto riguarda il resto della superficie dobbiamo essere un po' più originali.

$$\Phi(F, \partial \Omega, n) = \Phi(F, \Sigma, \nu) + \Phi(F, D, \eta) \stackrel{\text{thm divergenza}}{=} \int_{\Omega} \text{div } F dm_3 = 3 \int_{\Omega} dm_3 = 3m_3(\Omega)$$

dove $\Omega = \{(x, y, z) \in \mathbb{R}^3 | 4x^2 + 4y^2 + z^2 < 4, x < 0\}$. Possiamo fare questo integrale senza problemi in coordinate sferiche

$$m_3(\Omega) = \int_{\Omega} dx dy dz = 2 \int_{\Omega'} dx dy dz' = 2m_3(B(0, 1) \cap \{x < 0\}) = \frac{4}{3}\pi$$

dove, per agevolare il calcolo dell'integrale, abbiamo fatto il cambio di coordinate $(x', y', z') \mapsto (x', y', 2z') = (x, y, z)$ dunque il jacobiano di questa trasformazione era pari a 2 e poi abbiamo diviso per 2 per ottenere solo metà contributo della sfera (siccome la condizione $x < 0$ taglia metà della sfera). Dunque

$$\Phi(F, \Sigma, \nu) = 3m_3(\Omega) = 4\pi$$

\square

Mostriamo due esempi famosi del teorema della divergenza

6.7.2 Legge di Gauss

Sia $E = k_e \frac{q}{r^2} \hat{r}$ il campo elettrico generato da una carica puntiforme, allora

$$\begin{aligned}\operatorname{div} E &= \operatorname{div}(k_e \frac{q}{r^2} \hat{r}) = \operatorname{div}(k_e \frac{q}{r^3} r) = k_e q \sum_{i=1}^3 \partial_{x_i} \left(\frac{x_i}{|x|^3} \right) = k_e q \sum_{i=1}^3 \frac{|x|^3 - 3x_i^2 |x|}{|x|^6} = \\ &= k_e q \frac{3|x|^2 - 3|x|^2}{|x|^5} = 0\end{aligned}$$

E è ben definito $\Omega \setminus B(x_0, \varepsilon)$ e, dunque, possiamo applicare il teorema della divergenza su $\Omega \setminus B(x_0, \varepsilon)$, da cui deduciamo che

$$\begin{aligned}\Phi(E, \partial_{\text{reg}}(\Omega \setminus B(x_0, \varepsilon)), v) &= \int_{\partial\Omega} \langle E, v \rangle d\sigma + \int_{\partial B(x_0, \varepsilon)} \langle E, v \rangle d\sigma = \Phi(E, \partial\Omega, v) - k_e \frac{q}{\varepsilon^2} (4\pi\varepsilon^2) = \\ &= \int_{\Omega \setminus B(x_0, \varepsilon)} \operatorname{div} E dm_3 = 0 \implies \Phi(E, \partial\Omega, v) = 4\pi k_e q\end{aligned}$$

dove abbiamo usato il fatto che $\int_{B(x_0, \varepsilon)} \langle E, v \rangle = E 4\pi \varepsilon^2$ in virtù delle simmetrie radiali del campo elettrico.

6.7.3 Equazione di continuità in fluidodinamica

Sia Ω una regione di spazio di un fluido che sta "uscendo" da questo e indichiamo con $d\Sigma$ l'elemento infinitesimo di superficie, $d\sigma$ l'elemento infinitesimo di area e $v(x, y, z, t)$ la velocità della particella di un liquido nel punto (x, y, z) al tempo t . In un tempo infinitesimale dt , l'elemento infinitesimale di volume di liquido passante attraverso l'elemento di superficie $d\Sigma$ è pari a $|\vec{v} \cdot \nu| d\sigma$, dove \vec{v} è la velocità del liquido mentre ν è la normale alla superficie. Supponendo che il nostro liquido abbia una densità $\rho(x, y, z, t) \geq 0$ variabile, la quantità di fluido totale è data da

$$V = \int_{\Sigma} \rho(\vec{v} \cdot \nu) d\sigma = \Phi(\rho \vec{v}, \Sigma, \nu)$$

Per la conservazione della massa dobbiamo avere che

$$\frac{M_t - M_{t+dt}}{dt} = V$$

ma sapendo che $M_t = \int_{\Omega} \rho(x, y, z, t) dx dy dz$ (ovvero la massa di tutto il fluido è l'integrale in tutta la regione di spazio Ω in cui si trova il fluido al tempo t), allora

$$-\frac{\partial}{\partial t} \int_{\Omega} \rho(x, y, z, t) dx dy dz = \int_{\partial\Omega} \rho \vec{v} \cdot \nu d\sigma$$

dove il $-$ deriva dal fatto che la V è naturalmente positiva ma il bilancio di massa dove che il fluido esce dalla superficie diminuisce, dunque la derivata è negativa. Tramite la teoria di Lebesgue è possibile portare la derivata all'interno dell'integrale con ipotesi deboli, dunque

$$\int_{\Omega} -\frac{\partial \rho}{\partial t} dm_3 = \Phi(\rho \vec{v}, \partial\Omega, \nu) \stackrel{\text{thm divergenza}}{=} \int_{\Omega} \operatorname{div}(\rho \vec{v}) \implies \frac{\partial \rho}{\partial t} + \operatorname{div}(\rho \vec{v}) = 0$$

che prende il nome di equazione di continuità.

6.8 Teorema del rotore

Definizione 6.28 (curva regolare a tratti). Una curva $\gamma : [a, b] \rightarrow \mathbb{R}^n$ si dice regolare a tratti se $\gamma \in C^1_{\text{tratti}}$ e $\dot{\gamma}(t) \neq 0 \forall t \in (t_{i-1}, t_i) \forall i \in \{1, \dots, N\}$ dove $a = t_0 < t_1 < \dots < t_{N-1} < t_N = b$ è il partizionamento che rende $\gamma \in C^1_{\text{tratti}}$. Se $N = 1$, ovvero $t_0 = a$ e $t_1 = b$ diremo che γ è regolare.

Definizione 6.29 (dominio piano elementare). Un insieme $H \subseteq \mathbb{R}^2$ è detto dominio piano elementare se è limitato e la sua frontiera ∂H è una curva semplice, chiusa e regolare a tratti

Oss.: Il "bordo" di questo dominio può essere frastagliato in alcuni punti

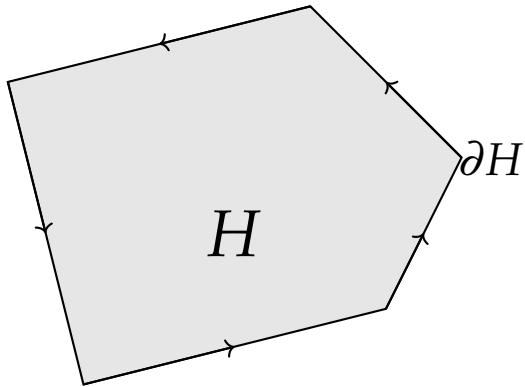


Figura 6.2: Esempio di dominio piano elementare. Come si vede il bordo è frastagliato, eppure è possibile parametrizzare il bordo con una curva semplice, chiusa e regolare a tratti

Tramite i domini elementari possiamo definire le superfici elementari

Definizione 6.30 (superfici). Sia $H \subseteq \mathbb{R}^2$ dominio piano elementare, $U \subseteq \mathbb{R}^2$ aperto con $\bar{H} \subseteq U$, $\Phi \in C^1(U, \mathbb{R}^3)$ con

- ① Φ iniettiva;
- ② $d\Phi : \mathbb{R}^2 \rightarrow \mathbb{R}^3$ iniettivo $\forall u \in H$;

allora $\Sigma = \Phi(H)$ è una superficie elementare con Φ parametrizzazione di Σ tale che

$$\partial\Sigma = \Phi(\partial H) \tag{6.7}$$

Oss.: La condizione $\Phi(\partial H) = \partial\Sigma$ serve per evitare delle situazioni come il grafico dell'otto: possiamo parametrizzare un 8 solo per $y > 0$ ma in $z = 0$, se invertiamo la funzione localmente, possiamo estenderla anche al semipiano inferiore ottenendo qualcosa che non è più un (x, y) -grafico.

Oss.: Possiamo mostrare che una superficie elementare è una superficie regolare, siccome Σ è localmente un grafico. Infatti

$$D\Phi(u) = \begin{pmatrix} \nabla\Phi_1(u) \\ \nabla\Phi_2(u) \\ \nabla\Phi_3(u) \end{pmatrix}$$

e supponiamo che $M_{12}(D\Phi) = \det \begin{pmatrix} \nabla\Phi_1 \\ \nabla\Phi_2 \end{pmatrix} \neq 0$: possiamo considerare $\Phi_{(1,2)} : U \rightarrow \mathbb{R}^2$ tale che $u \mapsto (\Phi_1(u), \Phi_2(u))$. Per ipotesi avremo che $\det D\Phi_{(1,2)}(u) = M_{12}(D\Phi(u)) \neq 0 \implies$ per il teorema della funzione inversa $\exists \delta > 0 : \exists V \subseteq \mathbb{R}^2 : \Phi_{(1,2)} : B(u, \delta) \rightarrow V$ invertibile con inversa $\Psi : V \rightarrow B(u, \delta) \in C^1(V, B(u, \delta))$. Ma allora $(\Phi \circ \Psi)(u) = (\Phi_1(\Psi(v)), \Phi_2(\Psi(v)), \Phi_3(\Psi(v))) =$

$(\Phi_{(1,2)}(\Psi(v)), \Phi_3(\Psi(v))) = (v, \Phi_3(\Psi(v)))$ siccome $\Psi = \Phi_{(1,2)}^{-1}$. A questo punto possiamo localmente vedere questa superficie elementare come quella che soddisfa l'equazione $z = \Phi_3(\Psi(v))$ e, ponendo $g = \Phi_3 \circ \Psi$, avremo la seguente equazione definente $z = g(v) \implies z - g(v) = 0 \implies \Sigma = f^{-1}(0)$ dove $f(v, z) = z - g(v)$.

CAPITOLO 7

Altri esercizi proposti

Esercizio 7.1. Sia $E = \{(x, y, z) \in \mathbb{R}^3 : z \geq 0, y \geq 1, y^{2\alpha}(z + x^2) \leq 1\}$. Determinare se è misurabile

Svolgimento. Osserviamo che la condizione $z + x^2 \leq y^{-2\alpha} \implies 0 \leq y^{-2\alpha}$ e $x^2 \leq y^{-2\alpha}$. Sezioniamo con Tonelli con la y

$$\int_E dx dy dz = \int_1^{+\infty} dy \int_{E_y} dx dz$$

dove E_y è un sezionamento a y costante. Possiamo allora sezionare con $x^2 \leq y^{-2\alpha} \implies -y^{-\alpha} < x < y^{-\alpha}$

$$\begin{aligned} \int_1^{+\infty} dy \int_{E_y} dx dz &= \int_1^{+\infty} dy \int_{-y^{-\alpha}}^{y^{-\alpha}} dx \int_0^{y^{-2\alpha}-x^2} dz = 2 \int_1^{+\infty} dy \int_0^{y^{-\alpha}} y^{-2\alpha} - x^2 dx = \\ &= \int_1^{+\infty} y^{-3\alpha} - \frac{x^3}{3} \Big|_0^{y^{-\alpha}} dy = \frac{4}{3} \frac{y^{-3\alpha+1}}{1-3\alpha} \Big|_1^{+\infty} \implies \alpha \neq \frac{1}{3} \end{aligned}$$

□

Esercizio 7.2. Siano $\gamma(t) = (t^2 e^t, t^2)$ e siano $c(t) = (te, t)$ con $\gamma, c : [0, 1] \rightarrow \mathbb{R}$. Calcolare l'area fra le due curve

Svolgimento. Possiamo usare la forma differenziale d'area per calcolare quest'area. **Attenzione:** usare una forma differenziale d'area rispetta ad un'altra fa cambiare il valore delle aree sotto le curve γ e $c(t)$, in maniera tale che le differenze e le somme diano lo stesso valore: la forma differenziale

$$\Gamma_c = \int_{\partial^+ E} \frac{xdy - ydx}{2} = 0$$

mentre

$$\Gamma_c = \int_{\partial^+ E} xdy = \frac{e}{2} \neq 0$$

Quindi quando si usano le forme differenziali e si sceglie quale usare, dobbiamo fare i conti tutti con la stessa!

□

Esercizio 7.3. Sia $\omega = \log(x^2 + y^2 + z^2)x dx + \log(x^2 + y^2 + z^2)y dy + \log(x^2 + y^2 + z^2)z dz$ definita su $\Omega = \mathbb{R}^3 \setminus \{(0, 0, 0)\}$. Mostrare che

1. ω è esatta
2. Sia $\gamma(t) : [0, \pi] \rightarrow \Omega$ con $\gamma(t) = (\sin(t), \cos(t), 0)$. Calcolare $\int_{\gamma} \omega$

have usually similar feelings; there is no scorn more profound, or on the whole more common, is work for second-rate minds. I can remember arguing this point once in o

^a Name and Nature of Poetry had denied very emphatically that he was a literary critic. 'I^{ts} had begun not so Talley imp^t cri
y criticism is the comparison w^mome respects a me deplorable e had said to b^gh dinner, and ne end, his repl here is no doul of weakness, f passed sixty, I^{ll} I may be told t ble, since the need now c ntroduction to is practically ne h your labour? ges and steam ts way very cor le achievement al defence. At study of math is worth while e extent egotist at it really nec his subject an uragement to o t and himself ria er the work which i.

couraging, but most people will find the second

ay. If then I find are vigorous mathem nce to carry on effective, erally m^ond, for good ato^t the only xt thⁿ descri^vsical kⁱ, the sa^s us of n^e es not^e se com^{it} is base^y lificul^l 'l be, b^e m^ost p^oics wi^z a defe^r s one^z its fail^r ofess^r, for e^m I the rig^t they de^s activⁱ irst c^e e one^h he one^{am} a law^{ter} because^s is a defenc^{al} minority: ings well is n^I told him th^{an} powers .. tall such atte^rly disappo^{'e}asant se[.]

APPENDICE A

Notazione multi-indice

La notazione multi-indice è una notazione matematica che consente di semplificare enormemente la scrittura di formule matematiche. Ciò è possibile *generalizzando* il concetto di indice che passa dall'essere un semplice "numero" ad una n -upla ordinata di indici. Consideriamo $\alpha \in \mathbb{N}^n$, ovvero $\alpha = (\alpha_1, \dots, \alpha_n)$ come nostro multi-indice, allora possiamo definire una serie di operazioni: la più semplice di tutti, data la natura di spazio vettoriale di cui gode \mathbb{N}^n , è sicuramente la somma dei multi-indici:

$$\forall \alpha, \beta \in \mathbb{N}^n, \alpha \pm \beta = (\alpha_1 \pm \beta_1, \dots, \alpha_n \pm \beta_n).$$

Per quanto riguarda i nostri scopi, è utile definire la *norma multi-indice*, il *fattoriale multi-indice* e il *coefficiente multinomiale*:

Definizione A.1 (norma multi-indice). Sia $\alpha \in \mathbb{N}^n$ allora definiamo la norma multi-indice come

$$|\alpha| = \sum_{i=1}^n \alpha_i$$

Definizione A.2 (fattoriale multi-indice). Sia $\alpha \in \mathbb{N}^n$, definiamo il fattoriale multi-indice di α come

$$\alpha! = \alpha_1! \alpha_2! \dots \alpha_n!$$

Definizione A.3 (coefficiente multinomiale). Sia $\alpha \in \mathbb{N}^n$ tale che $|\alpha| = k$, allora definiamo il coefficiente multinomiale come

$$\binom{k}{\alpha} = \frac{k!}{\alpha!} = \frac{k!}{\alpha_1! \alpha_2! \dots \alpha_n!}$$

Notazione: per distinguerlo dal normale coefficiente binomiale, nel caso del coefficiente multinomiale, metterò in grassetto il multi-indice.

Oss.: Potremmo pensare a introdurre un ordinamento parziale (ma non è di grande interesse) dicendo che $\alpha \leq \beta \iff \alpha_i \leq \beta_i \forall i \in \{1, \dots, n\}$. Si vede subito che, così come è stato definito, questo non può essere un ordinamento totale

Vediamo come questa notazione può agevolarci la scrittura del polinomio di Taylor e la scrittura dei polinomi in n -variabili. Innanzitutto definiamo

Definizione A.4 (potenza multi-indice). Sia $x \in \mathbb{R}^n$ allora

$$x^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}$$

Questa notazione facilita enormemente la scrittura dei polinomi in n -variabili, infatti

Definizione A.5 (polinomio in n -variabili). $P : \mathbb{R}^n \rightarrow \mathbb{R}$ è un polinomio in n -variabili di grado $k \geq 0$ se

$$\exists C = \{c_\alpha : |\alpha| \leq k, \alpha \in \mathbb{N}^n\} \subset \mathbb{R}$$

dove $0 \neq c_\beta \in C$ e $|\beta| = k$. Allora

$$P(x) = \sum_{|\alpha| \leq k} c_\alpha x^\alpha$$

Oltre a questo possiamo definire la derivata multi-indice

Definizione A.6 (derivata multi-indice). Sia $\alpha \in \mathbb{N}^n$ allora

$$\partial_x^\alpha = \partial_{x_1}^{\alpha_1} \dots \partial_{x_n}^{\alpha_n}$$

dove, naturalmente, poniamo che $\partial_{x_j}^0 = 1$

Riprendiamo il polinomio di Taylor: da quanto visto nella sezione 3.3 abbiamo visto che possiamo scrivere il polinomio di Taylor di una funzione $f : \mathbb{R}^n \rightarrow \mathbb{R}$ come

$$P_k(f, \xi) = \sum_{i_1, \dots, i_k=1}^k \partial_{x_{i_1}} \dots \partial_{x_{i_k}} f(\xi) (x_{i_1} - \xi_{i_1}) \dots (x_{i_k} - \xi_{i_k}).$$

Osserviamo che possiamo allora usare la potenza multi-indice per scrivere i termini $(x_{i_j} - \xi_{i_j})$, considerando $\alpha \in \mathbb{N}^n$ e osservando che $\alpha_{j_1}, \dots, \alpha_{j_m} \geq 1$ e $1 \leq j_1, \dots, j_m \leq n$ tali che $\alpha_{j_1} + \dots + \alpha_{j_m} = k$. In altre parole avremo solamente le derivate di ordine superiore

$$\partial_{x_{j_1}}^{\alpha_{j_1}} \partial_{x_{j_2}}^{\alpha_{j_2}} \dots \partial_{x_{j_m}}^{\alpha_{j_m}} f(\xi)$$

e $\alpha_i = 0$ se $i \neq j_1, \dots, j_m$.

Adesso ricordiamo che il teorema di Schwarz ci assicura che la derivata di una funzione di ordine C^k rispetto agli indici $\alpha_{j_1}, \dots, \alpha_{j_m}$ (rispetto, naturalmente, in ordine a x_{j_1}, \dots, x_{j_m}) è pari a quella ottenuta con una permutazione σ degli indici $\alpha_{j_1}, \dots, \alpha_{j_m}$, dunque dobbiamo moltiplicare i termini della serie di Taylor per un termine che ci "tronca" i termini uguali (si ricordi la dimostrazione della proposizione 3.6, dove compariva il termine $l!$). La possibilità di avere α_{j_1} derivate rispetto a x_{j_1} sulle k possibili è pari a $\binom{k}{\alpha_{j_1}}$ (non siamo interessati all'ordine, dunque usiamo per questo il coefficiente binomiale) e la stessa cosa vale per tutti gli altri termini, infatti le possibilità di avere α_{j_2} derivate rispetto a x_{j_2} sulle $k - \alpha_{j_1}$ rimanenti è pari a $\binom{k - \alpha_{j_1}}{\alpha_{j_2}}$ e così via. In generale, il numero di modi in cui possiamo distribuire le derivate è pari a

$$\begin{aligned} & \binom{k}{\alpha_{j_1}} \binom{k - \alpha_{j_1}}{\alpha_{j_2}} \dots \binom{k - \alpha_{j_1} - \dots - \alpha_{j_{m-1}}}{\alpha_{j_m}} = \\ & = \frac{k!}{\alpha_{j_1}! \cancel{(k - \alpha_{j_1})!} \alpha_{j_2}! \cancel{(k - \alpha_{j_1} - \alpha_{j_2})!} \alpha_{j_3}! \cancel{(k - \alpha_{j_1} - \alpha_{j_2} - \alpha_{j_3})!} \dots \cancel{\alpha_{j_m}!} 0!} = \\ & = \frac{k!}{\alpha_{j_1}! \alpha_{j_2}! \dots \alpha_{j_m}!} = \binom{k}{\alpha} \end{aligned}$$

Con gli argomenti precedenti in combinazione con il teorema di Schwarz di ordine superiore, possiamo dunque dimostrare il seguente lemma

Lemma A.0.1. Sia $f \in C^k(\Omega)$ e $\xi \in \Omega$, allora abbiamo

$$d^l f(\xi)(x - \xi) = \sum_{|\alpha|=l} \binom{l}{\alpha} \partial^\alpha f(\xi)(x - \xi)^\alpha.$$

Proposizione A.1. Se $f \in C^k(\Omega)$, $\Omega \subseteq \mathbb{R}^n$ aperto e $\xi \in \Omega$, allora abbiamo che

$$P_k(f, \xi)(x) = \sum_{|\alpha| \leq k} \frac{1}{\alpha!} \partial^\alpha f(\xi)(x - \xi)^\alpha$$

Dimostrazione. Sappiamo che per $0 \leq j \leq k$ abbiamo che

$$d^j f(\xi)(h) = \sum_{|\alpha|=j} \binom{j}{\alpha} \partial^\alpha f(\xi) h^\alpha$$

e ciò implica, per quanto visto nel capitolo 3, che

$$\begin{aligned} P_k(f, \xi) &= \sum_{j=0}^k \frac{1}{j!} d^j f(\xi)(x - \xi) = \sum_{j=0}^k \frac{1}{j!} \sum_{|\alpha|=j} \binom{j}{\alpha} \partial_x^\alpha f(\xi)(x - \xi)^\alpha \stackrel{\binom{j}{\alpha} = \frac{j!}{\alpha!}}{=} \\ &\stackrel{\binom{j}{\alpha} = \frac{j!}{\alpha!}}{=} \sum_{j=0}^k \sum_{|\alpha|=j} \frac{1}{\alpha!} \partial_x^\alpha f(\xi)(x - \xi)^\alpha = \sum_{0 \leq |\alpha| \leq k} \frac{1}{\alpha!} \partial_x^\alpha f(\xi)(x - \xi)^\alpha \end{aligned}$$

□

Teorema A.1 (formula di Taylor coi multi-indici). Se $f \in C^k(\Omega)$, $\xi \in \Omega$, $\Omega \subseteq \mathbb{R}^n$ aperto, allora abbiamo che

$$f(x) = \sum_{0 \leq |\alpha| \leq k} \frac{1}{\alpha!} \partial_x^\alpha f(\xi)(x - \xi)^\alpha + o(|x - \xi|^\alpha)$$

Dimostrazione. La dimostrazione segue dalla precedente proposizione e dalla formula di Taylor che abbiamo visto nel capitolo 3. □

APPENDICE B

Dimostrazione del teorema della funzione inversa

In questa appendice ci proponiamo di dimostrare in maniera rigorosa il teorema della funzione inversa, riportando tutti i teoremi precedentemente non enunciati necessari per effettuare la dimostrazione. Sarebbe anche interessante mostrare che il teorema del Dini implica il teorema della funzione inversa, mostrando così che i due teoremi sono equivalenti, tuttavia ciò non verrà fatto.

Teorema B.1 (teorema della funzione inversa). Sia $\Omega \subseteq \mathbb{R}^n$ aperto, $f \in C^k(\Omega, \mathbb{R}^n)$, $x_0 \in \Omega$ e $Df(x_0) \in \mathbb{R}^{n \times n}$ invertibile.

Allora $\exists \rho > 0 : \exists U \subseteq \mathbb{R}^n$ aperto tali che $f|_{B(x_0, \rho)} \in C^k(B(x_0, \rho), U)$ è invertibile e, posto $g = (f|_{B(x_0, \rho)})^{-1} : U \rightarrow B(x_0, \rho)$, abbiamo che $g \in C^k(U, B(x_0, \rho))$ con

$$Dg(y) = (Df(g(y)))^{-1} \forall y \in U$$

B.1 Norma matriciale indotta

Per dimostrare questo risultato, è conveniente passare ad una norma differente di quella di Hilbert-Schmidt. Sebbene tutte le norme, in uno spazio finito dimensionale sono equivalenti (e dunque inducono la stessa topologia) l'utilizzo di questa norma rende più semplice la trattazione di alcuni risultati di cui abbiamo bisogno. Iniziamo dando la definizione di spazio delle applicazioni lineari fra due spazi vettoriali, la quale dovrebbe già essere stata data nel corso di Geometria, che riporto per completezza

Definizione B.1 (spazio delle applicazioni lineari). Siano X, Y due spazi vettoriali su un campo \mathbb{K} . Definiamo

$$\mathcal{L}(X, Y) = \{f : X \mapsto Y \mid f \text{ è lineare}\}$$

Oss.: Se $A : \mathbb{R}^n \mapsto \mathbb{R}^m$ lineare, scriveremo che $A \in \mathcal{L}(\mathbb{R}^n)$

Oss.: $\mathcal{L}(X, Y)$ con X, Y spazi vettoriali su un campo \mathbb{K} è uno spazio vettoriale

Definizione B.2 (norma matriciale indotta). Sia $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$, allora definiamo la norma $\|A\|$ matriciale di A come

$$\|A\| = \sup_{x \in \mathbb{S}^{n-1}} |Ax| \tag{B.1}$$

dove $|\cdot| : \mathbb{R}^m \rightarrow \mathbb{R}$ è la usuale norma euclidea.

Oss.: Questa norma viene chiamata, in maniera suggestiva, *indotta* siccome non è altro che la norma vettoriale utilizzata in \mathbb{R}^n che viene "indotta" sullo spazio delle matrici considerando il sup della norma euclidea della matrice applicata sui vettori di norma 1.

Oss.: Si potrebbero definire le p -norme indotte, definite nell'usuale maniera riportata sopra, dove si utilizza come norma vettoriale la norma $\|\cdot\|_p : \mathbb{R}^n \rightarrow \mathbb{R}$ tale che

$$(x_1, \dots, x_n) \xrightarrow{\|\cdot\|_p} \left(\sum_{i=1}^n |x_i|^p \right)^{\frac{1}{p}}$$

Oss.: Osserviamo che

$$|Ax| \leq \|A\| |x| \quad (\text{B.2})$$

vale $\forall x \in \mathbb{R}^n$. Inoltre, se $\exists \lambda \in \mathbb{R} : \forall x \in \mathbb{R}^n, |Ax| \leq \lambda |x|$ allora $\|A\| \leq \lambda$

Esercizio B.1. *Mostrare la disegualanza (B.2).*

Svolgimento. Sia $x \in \mathbb{R}^n$. Allora sappiamo che

$$|Ax| = \left| A \left(|x| \frac{x}{|x|} \right) \right| = |x| \left| A \frac{x}{|x|} \right| \leq |x| \|A\|$$

siccome $\frac{x}{|x|}$ è un vettore di norma unitaria. \square

Per pura completezza, riporto la dimostrazione che due norme, nel caso finito-dimensionale, sono *equivalenti*

Definizione B.3 (equivalenza fra norme). Sia V un \mathbb{K} -spazio vettoriale e siano $\|\cdot\|_1, \|\cdot\|_2$ due norme definite sullo spazio vettoriale. Allora diremo che esse sono equivalenti se esistono due costanti c e C strettamente positive per cui

$$\forall x \in V, c\|x\|_1 \leq \|x\|_2 \leq C\|x\|_1. \quad (\text{B.3})$$

Oss.: Due norme equivalenti inducono la stessa topologia (impropriamente la stessa metrica), quindi strutture topologiche quali gli insiemi aperti, le funzioni continue e successioni convergenti rimarranno tali "cambiando" norma: questo potrà sembrare di poco conto e un'astrazione non necessaria, tuttavia, ripensando a come abbiamo costruito tutta la teoria dei limiti e delle successioni in \mathbb{R}^n , ci accorgiamo che la norma è l'oggetto su cui abbiamo basato la nostra nozione di convergenza. Dire che due norme inducono la stessa topologia vuol dire che presi due oggetti $x, y \in \mathbb{R}^n$ tali che $|x - y| \rightarrow 0$ (dove con $|\cdot| : \mathbb{R}^n \rightarrow \mathbb{R}$ intendo la norma euclidea) allora presa una qualunque altra norma $\|\cdot\|_\star : \mathbb{R}^n \rightarrow \mathbb{R}$ avremo che $\|x - y\|_\star \rightarrow 0$. Questo è un risultato non banale e che, in generale, non è vero per spazi vettoriali infinito-dimensionalni.

Proposizione B.1 (tutte le norme sono equivalenti a quella euclidea su $\mathbb{R}^n / \mathbb{C}^n$). Sia $n \in \mathbb{N}$. Tutte le norme su \mathbb{R}^n (o \mathbb{C}^n) sono equivalenti alla norma euclidea.

Dimostrazione. Senza perdita di generalità effettueremo la dimostrazione nel caso di \mathbb{R}^n . Sia $\|\cdot\|_\star$ una norma su \mathbb{R}^n e denotiamo con $\|x\|_2$ la norma euclidea. Siano $\{e_1, \dots, e_n\}$ una base canonica su \mathbb{R}^n , allora $\forall x \in \mathbb{R}^n, \exists x_1, x_2, \dots, x_n \in \mathbb{R}^n : x = x_1 e_1 + \dots + x_n e_n$; possiamo allora applicare ripetutamente la disegualanza triangolare sulla norma di x per minorarla

$$\|x\|_\star = \left\| \sum_{i=1}^n x_i e_i \right\|_\star \leq \sum_{i=1}^n \|x_i e_i\|_\star = \sum_{i=1}^n |x_i| \|e_i\|_\star \leq \sqrt{\sum_{i=1}^n |x_i|^2} \sqrt{\sum_{i=1}^n \|e_i\|_\star^2} = \|x\|_2 \sqrt{\sum_{i=1}^n \|e_i\|_\star^2}.$$

dove nell'ultima minorazione si è utilizzata la disuguaglianza di Cauchy-Schwarz.

Dunque se poniamo $A = \sqrt{\sum_{i=1}^n \|e_i\|_\star^2}$ otteniamo che $\|x\|_\star \leq A\|x\|_2 \forall x \in \mathbb{R}^n$, ovvero la norma euclidea domina la norma \star . A questo punto, per vedere che la norma \star domina quella euclidea, possiamo considerare la funzione

$$f(x) = \|x\|_\star$$

Osserviamo che

$$|f(x) - f(y)| = |||x||_\star - ||y||_\star| \leq \|x - y\|_\star \leq A\|x - y\|_2, \quad \forall x, y \in \mathbb{R}^n,$$

questo prova che la funzione f è lipschitziana, dunque continua rispetto alla norma euclidea. Consideriamo la sfera unitaria \mathbb{S}^{n-1} : tale insieme è infatti compatto per il teorema di Heine-Borel (teorema 2.4) in quanto limitato (banale) e chiuso (corollario del lemma B.4.1 dimostrato più avanti), conseguentemente, avremo per il teorema di Weierstrass (teorema 2.5) che la funzione f assumerà massimo e minimo. Sia $u_\star \in \mathbb{S}^{n-1}$ tale che

$$\min_{u \in \mathbb{S}^{n-1}} \|u\|_\star = \min_{u \in U} f(u) = f(u_\star) = \|u_\star\|_\star.$$

Indichiamo tale risultato con $B = \|u_\star\|_\star > 0$ è strettamente positivo siccome $\|v\|_\star = 0 \iff v = 0$ (una delle proprietà di cui gode la norma) e nel nostro caso stiamo considerando $u \in \mathbb{S}^{n-1} = u \in \mathbb{R}^n : \|u\|_2 = 1$, dunque $\forall u \in \mathbb{S}^{n-1}, u \neq 0$. Ma allora, per ogni x non nullo abbiamo che

$$\|x\|_\star = |||x||_2 \frac{x}{\|x\|_2}||_\star = \|x\|_2 \|\frac{x}{\|x\|_2}\|_\star \geq \|x\|_2 B$$

siccome $\frac{x}{\|x\|_2}$ è un vettore unitario rispetto alla norma euclidea e, per quanto visto prima, abbiamo che $\forall u \in \mathbb{S}^{n-1}, B = \|u_\star\|_\star \leq \|u\|_\star$, giustificando così la maggiorazione effettuata nell'ultimo passaggio. Concludendo, abbiamo che

$$B\|x\|_2 \leq \|x\|_\star \leq A\|x\|_2$$

e

$$\frac{1}{A}\|x\|_\star \leq \|x\|_2 \leq \frac{1}{B}\|x\|_\star$$

dunque $\|\cdot\|_\star$ e $\|\cdot\|_2$ sono equivalenti. \square

Corollario B.1.1 (tutte le norme sono equivalenti in $\mathbb{R}^n / \mathbb{C}^n$). Sia $n \in \mathbb{N}$. Tutte le norme sono equivalenti in $\mathbb{R}^n / \mathbb{C}^n$.

Dimostrazione. Come visto nella dimostrazione del teorema B.1, sappiamo che prese due norme $\|\cdot\|_\star$ e $\|\cdot\|_*$ avremo che $\exists A_\star, B_\star, A_*, B_* \in \mathbb{R}$ tali che

$$\begin{aligned} B_\star\|x\|_2 &\leq \|x\|_\star \leq A_\star\|x\|_2 & B_*\|x\|_2 &\leq \|x\|_* \leq A_*\|x\|_2 \\ \frac{1}{A_\star}\|x\|_\star &\leq \|x\|_2 \leq \frac{1}{B_\star}\|x\|_\star & \frac{1}{A_*}\|x\|_* &\leq \|x\|_2 \leq \frac{1}{B_*}\|x\|_* \end{aligned}$$

pertanto abbiamo che

$$\|x\|_* \leq A_*\|x\|_2 = \frac{A_*}{B_\star} B_\star\|x\|_2 \leq \frac{A_*}{B_\star} \|x\|_\star \implies \|x\|_* \leq \frac{A_*}{B_\star} \|x\|_\star$$

e similmente possiamo vedere che

$$\|x\|_* \geq B_*\|x\|_2 = B_* \frac{1}{A_\star} \|x\|_\star \implies \|x\|_* \geq \frac{B_*}{A_\star} \|x\|_\star.$$

Conseguentemente, si conclude che

$$\frac{B_*}{A_*} \|x\|_\star \leq \|x\|_* \leq \frac{A_*}{B_*} \|x\|_\star,$$

e

$$\frac{B_\star}{A_*} \|x\|_* \leq \|x\|_\star \leq \frac{A_\star}{B_*} \|x\|_*.$$

Dunque le norme $\|\cdot\|_\star$ e $\|\cdot\|_*$ sono equivalenti. \square

Corollario B.1.2 (tutte le norme sono equivalenti su spazi di dimensione finita). Sia V un \mathbb{R} -spazio (o un \mathbb{C} -spazio) di dimensione finita n . Allora tutte le norme sono equivalenti.

Dimostrazione. Senza perdita di generalità faremo la dimostrazione nel caso di \mathbb{R} -spazio vettoriale. Sia $(V, \|\cdot\|_V)$ uno spazio vettoriale normato reale di dimensione finita n e fissiamo una base $\mathcal{B} = \{v_1, \dots, v_n\}$ di tale spazio. Sappiamo allora che $\forall v \in V \exists \lambda \in \mathbb{R}^n : v = \sum_{i=1}^n b_i v_i$ con $\lambda = (\lambda_1, \dots, \lambda_n)$: possiamo dunque considerare l'isomorfismo lineare

$$\phi : \mathbb{R}^n \rightarrow V \text{ tale che } \lambda \in \mathbb{R}^n \mapsto \sum_{i=1}^n \lambda_i v_i \in V.$$

Grazie a questo isomorfismo possiamo far corrispondere alla norma di V una norma definita su \mathbb{R}^n definita come

$$\|\lambda\|_\star = \|\phi(\lambda)\|_V, \quad \forall \lambda \in \mathbb{R}^n.$$

Verifichiamo che essa è una norma su \mathbb{R}^n : abbiamo che $\phi(\lambda) = 0_V \iff \lambda = \underline{0}$ (siccome ϕ è un isomorfismo lineare) e per proprietà della norma $\|\cdot\|_V$ abbiamo che $\|v\|_V = 0 \iff v = 0$, dunque abbiamo effettivamente che $\|\lambda\|_\star = 0 \iff \lambda = \underline{0}$. Preso adesso uno scalare $c \in \mathbb{R}$ e $\lambda \in \mathbb{R}^n$, abbiamo che $\phi(c\lambda) = c\phi(\lambda)$ (sempre perché ϕ è un isomorfismo lineare) e, per proprietà di omogeneità della norma, abbiamo che $\|c\lambda\|_\star = \|\phi(c\lambda)\|_V = \|c\phi(\lambda)\|_V = |c| \|\phi(\lambda)\|_V = |c| \|\lambda\|_\star$, mentre per la disuguaglianza triangolare abbiamo che presi $\lambda_1, \lambda_2 \in \mathbb{R}^n$ abbiamo che

$$\|\lambda_1 + \lambda_2\|_\star = \|\phi(\lambda_1 + \lambda_2)\|_V \leq \|\phi(\lambda_1)\|_V + \|\phi(\lambda_2)\|_V = \|\lambda_1\|_\star + \|\lambda_2\|_\star,$$

dove nei passaggi abbiamo utilizzato che $\phi(\lambda_1 + \lambda_2) = \phi(\lambda_1) + \phi(\lambda_2)$ (sempre perché ϕ è un isomorfismo lineare) e la disuguaglianza triangolare di cui gode la norma $\|\cdot\|_V$. Dunque la norma definita come

$$\|\lambda\|_\star = \|\phi(\lambda)\|_V$$

è una norma su \mathbb{R}^n .

Siano adesso $\|\cdot\|_{b\star} = \|\phi(\cdot)\|_b$ e $\|\cdot\|_{\sharp\star} = \|\phi(\cdot)\|_\sharp$ due norme su \mathbb{R}^n . Per il corollario B.1.2 abbiamo che

$$\exists A_b, B_b, A_\sharp, B_\sharp \in \mathbb{R} : \begin{cases} A_b \|\lambda\|_{\sharp\star} \leq \|\lambda\|_{b\star} \leq B_b \|\lambda\|_{\sharp\star} & \forall \lambda \in \mathbb{R}^n. \\ A_\sharp \|\lambda\|_{b\star} \leq \|\lambda\|_{\sharp\star} \leq B_\sharp \|\lambda\|_{b\star} \end{cases}$$

Ma siccome ϕ è un isomorfismo suriettivo allora avremo che

$$\begin{cases} A_b \|v\|_\sharp \leq \|v\|_b \leq B_b \|v\|_\sharp & \forall v \in V, \\ A_\sharp \|v\|_b \leq \|v\|_\sharp \leq B_\sharp \|v\|_b \end{cases}$$

ovvero le norme $\|\cdot\|_b$ e $\|\cdot\|_\sharp$ sono equivalenti. \square

B.2 Qualche risultato preparatorio

Andiamo ad enunciare alcuni risultati di cui abbiamo bisogno per dimostrare quanto ci siamo prefissi all'inizio di questa appendice. Alcuni di essi dovrebbero essere già noti dal corso di Analisi Matematica:

Teorema B.2 (teorema di Banach-Caccioppoli). Sia X uno spazio metrico completo e sia $\varphi : X \rightarrow X$ una contrazione (ovvero una funzione L -lipschitziana con $L < 1$). Allora $\exists!x \in X : \varphi(x) = x$

Dimostrazione. Sia (X, d) uno spazio metrico completo e sia $x_0 \in X$. Definiamo una successione $\{x_n\}$ in maniera ricorsiva, imponendo che

$$x_{n+1} = \varphi(x_n) \quad (n = 0, 1, \dots).$$

La dimostrazione segue in due *step*: la prima consiste nel mostrare l'esistenza di un punto fisso, la seconda l'unicità di tale punto.

Esistenza: l'idea è quella di mostrare che questa successione da noi definita è una successione di Cauchy, da cui segue, data la completezza di X , che la successione converge a $x \in X$.

Siccome φ è una contrazione, allora $\exists c \in \mathbb{R}$ tale che $c < 1$ e $\forall x, y \in X, d(\varphi(x), \varphi(y)) \leq cd(x, y)$. Dunque avremo che

$$d(x_{n+1}, x_n) = d(\varphi(x_n), \varphi(x_{n-1})) \leq cd(x_n, x_{n-1})$$

Procedendo induttivamente (il caso $n = 1$ è naturalmente banale) avremo che

$$d(x_{n+1}, x_n) \leq cd(x_n, x_{n-1}) \leq c^n d(x_1, x_0)$$

mentre se prendiamo $n \neq m$ e supponiamo (senza perdere di generalità) che $n > m$ allora

$$\begin{aligned} d(x_n, x_m) &\stackrel{\text{dis. triang.}}{\leq} \sum_{i=m}^{n-1} d(x_i, x_{i+1}) \leq \sum_{i=m}^{n-1} c^i d(x_1, x_0) = d(x_1, x_0) \sum_{i=m}^{n-1} c^i = \\ &= d(x_1, x_0) \left(\frac{1 - c^{n+1}}{1 - c} - \frac{1 - c^{m+1}}{1 - c} \right) = \frac{c^m - c^n}{1 - c} d(x_1, x_0). \end{aligned}$$

Ma, siccome $c < 1$, se $m \rightarrow +\infty$ e $n > m$ allora quest'ultima quantità tende a 0, dunque risulta che $\{x_n\}$ è una successione di Cauchy. Conseguentemente, essendo X completo, tale successione converge ad un punto $x \in X$ e, usando la continuità di φ , abbiamo che

$$x_{k+1} = \varphi(x_k) \implies \lim_{k \rightarrow +\infty} x_{k+1} = \lim_{k \rightarrow +\infty} \varphi(x_k) \implies x = \varphi(x).$$

Unicità: supponiamo per assurdo che esistano due punti fissi, ovvero $\exists x, y \in X, x \neq y : \varphi(x) = x, \varphi(y) = y$. Allora avremo che

$$d(x, y) = d(\varphi(x), \varphi(y)) \leq cd(x, y) \implies d(x, y) \leq cd(x, y) \implies c > 1$$

il che è un assurdo siccome φ è una contrazione. \square

Teorema B.3 (criterio di Lipschitz). Sia $\Omega \subseteq \mathbb{R}^n$ convesso, $f : \Omega \rightarrow \mathbb{R}^m$ una funzione di classe $C^1(\Omega, \mathbb{R}^m)$ e supponiamo che si abbia $L = \sup_{x \in \Omega} \|Df(x)\| < +\infty$. Allora la funzione f è L -lipschitziana, ossia

$$\forall x, y \in \Omega, |f(x) - f(y)| \leq L|x - y|$$

Dimostrazione. Fissiamo $x, x' \in \Omega$, allora consideriamo la funzione $g(t) = f(tx + (1-t)y)$ con $t \in [0, 1]$, pertanto $g : [0, 1] \rightarrow \mathbb{R}^m$ ed è anch'essa di classe C^1 e si ha che

$$g'(t) = Df(tx + (1-t)y)(x - y)$$

e, dalla formula fondamentale del calcolo integrale, segue che

$$f(x) - f(y) = g(1) - g(0) = \int_0^1 g'(t)dt = \int_0^1 Df(tx + (1-t)y)(x - y)dt$$

da cui segue che

$$\begin{aligned} |f(x) - f(y)| &= \left| \int_0^1 Df(tx + (1-t)y)(x - y)dt \right| \leq \int_0^1 \|Df(tx + (1-t)y)\| |x - y| dt \leq \\ &\leq \int_0^1 L dt |x - y| = L|x - y| \end{aligned}$$

□

Lemma B.3.1. $\forall n > 0, \forall r > 0, \forall x \in \mathbb{R}^n, B_r(x) \subset \mathbb{R}^n$ sono convesse.

Dimostrazione. Fissiamo $n > 0, r > 0$ e $x \in \mathbb{R}^n$. Siano adesso $y, z \in B_r(x)$ e consideriamo il luogo dei punti descritti da

$$\alpha = ty + (1-t)z, \text{ con } t \in [0, 1].$$

Allora

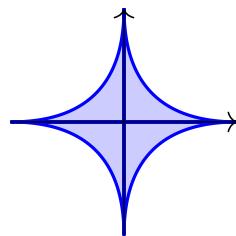
$$\begin{aligned} |\alpha - x| &= |ty + (1-t)z - x| \leq |ty + (1-t)z - tx - (1-t)x| \leq |ty - tx| + |(1-t)z - (1-t)x| = \\ &= t|y - x| + (1-t)|z - x| < tr + (1-t)r = r \end{aligned}$$

□

Oss.: Questa dimostrazione è valida, in realtà, per ogni spazio vettoriale munito di norma. Il fatto che le palle siano convesse non è valido per ogni spazio metrico: un esempio può essere dato da \mathbb{R}^2 munito della distanza

$$d((x_1, x_2), (y_1, y_2)) = \sqrt{|x_1 - y_1|} + \sqrt{|x_2 - y_2|}$$

dove la palla centrata in $(0, 0)$ e di raggio 1 assume la seguente forma



Nel prossimo teorema facciamo uso del seguente risultato che propongo come esercizio

Esercizio B.2. Sia $A \in \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$ e sia $B \in \mathcal{L}(\mathbb{R}^m, \mathbb{R}^k)$, allora

$$\|BA\| \leq \|B\| \|A\|.$$

Mostrare questo risultato sia per la norma matriciale indotta che per la norma di Hilbert-Schmidt

Svolgimento. Nel caso della norma matriciale indotta abbiamo che

$$\|BAx\| \leq \|B(Ax)\| \leq \|B\| |Ax| \leq \|B\| \|A\| |x| \implies \|BA\| \leq \|B\| \|A\|$$

dove abbiamo utilizzato ripetutamente la (B.2).

Per quanto riguarda la norma di Hilbert-Schmidt, osserviamo che, posto $C = BA$, abbiamo che

$$\begin{aligned} \|C\| &= \sqrt{\sum_{i,j} c_{ij}^2} = \sqrt{\sum_i^k \sum_j^n \left(\sum_{k=1}^n a_{ik} b_{kj} \right)^2} \leq \sqrt{\sum_i^k \sum_j^n \sum_{l=1}^m a_{il}^2 \sum_{l=1}^m b_{lj}^2} = \\ &= \sqrt{\sum_i^k \sum_l^m a_{il}^2 \sum_j^n \sum_l^m b_{lj}^2} = \sqrt{\sum_{i,l} a_{il}^2 \sum_{l,j} b_{lj}^2} = \|B\| \|A\| \end{aligned}$$

dove abbiamo utilizzato la diseguaglianza di Cauchy-Schwarz per giustificare la minorazione. Dunque

$$\|BA\| \leq \|B\| \|A\|.$$

□

Teorema B.4 (continuità dell'operazione di inversione di una matrice). Sia Ω l'insieme di tutte le applicazioni lineari invertibili a valori in \mathbb{R}^n . Allora

- ① Se $A \in \Omega$, $B \in \mathcal{L}(\mathbb{R}^n)$ e

$$\|B - A\| \cdot \|A^{-1}\| < 1 \quad (*)$$

allora $B \in \Omega$.

- ② Ω è un insieme aperto di $\mathcal{L}(\mathbb{R}^n)$ e la mappa $A \mapsto A^{-1}$ è continua su Ω

Dimostrazione. Mostriamo la ①. Osserviamo che prese $A \in \Omega$, $B \in \mathcal{L}(\mathbb{R}^n)$ che soddisfano la (*) e, per comodità, poniamo $\frac{1}{\alpha} = \|A^{-1}\|$ e $\|B - A\| = \beta$. Dunque la (*) diventa

$$\frac{\beta}{\alpha} < 1 \implies \alpha > \beta.$$

Osserviamo adesso che $\forall x \in \mathbb{R}^n$

$$\alpha|x| = \alpha|A^{-1}Ax| \leq \alpha\|A^{-1}\| |Ax| = |Ax| = |(A - B + B)x| \leq |(A - B)x| + |Bx| \leq \beta|x| + |Bx|$$

dunque

$$0 \leq (\alpha - \beta)|x| \leq |Bx| \quad (\star)$$

dove la prima eguaglianza segue dal fatto che $\alpha - \beta > 0$. Ma questo allora implica che $Bx \neq 0$ se $x \neq 0$: questo ci permette di concludere, per una serie di noti risultati visti dal corso di Geometria, che B è un isomorfismo, dunque è invertibile, quindi $B \in \Omega$.

Mostriamo la ②. Prendiamo $x = B^{-1}y$ e mettiamola all'interno della (\star)

$$(\alpha - \beta)|B^{-1}y| \leq |BB^{-1}y| = |y|$$

che una relazione valida $\forall y \in \mathbb{R}^n$ da cui possiamo dedurre che

$$|B^{-1}y| \leq \frac{1}{\alpha - \beta}|y| \implies \|B\| \leq \frac{1}{\alpha - \beta}.$$

Osservando che

$$B^{-1} - A^{-1} = B^{-1}(A - B)A^{-1}$$

abbiamo che

$$\|B^{-1} - A^{-1}\| \leq \|B^{-1}\| \|A - B\| \|A^{-1}\| \leq \frac{1}{\alpha - \beta} \frac{\beta}{\alpha} = \frac{\beta}{\alpha(\alpha - \beta)}$$

e per $\beta \rightarrow 0$ osserviamo che $\|B - A\| \rightarrow 0$, dunque la mappa che associa ad una matrice la sua inversa è continua. \square

Oss. Il passaggio in cui abbiamo affermato che $\|B\| \leq \frac{1}{\alpha - \beta}$ sarebbe fallito con la norma di Hilbert-Schmidt. Sarebbe invece stato corretto nel caso in cui io avessi affermato che $\|B\| \leq \frac{\sqrt{N}}{\alpha - \beta}$: il motivo di questo deriva dal fatto che in base ortonormale la norma di Hilbert-Schmidt altro non è che $\|B\| = \sqrt{\sum_{i=1}^n |Be_i|^2}$ dove $\{e_1, \dots, e_n\}$ è una qualunque base ortonormale dello spazio vettoriale su cui mi trovo e $|\cdot| : V \rightarrow \mathbb{K}$ è la norma su essa definita. Nel caso sopra se risultava che $|Bx| \leq \frac{1}{\alpha - \beta} |x| \forall x \in \mathbb{R}^n \implies |Be_i| < \frac{1}{\alpha - \beta} \implies \|B\| = \sqrt{\sum_{i=1}^n |Be_i|^2} \leq \sqrt{\sum_{i=1}^n \frac{1}{(\alpha - \beta)^2}} = \frac{\sqrt{N}}{\alpha - \beta}$. Siccome queste dimostrazioni risultano più naturali con la norma matriciale indotta, ho preferito passare a questa norma piuttosto che rimanere con quella di Hilbert-Schmidt, sebbene, in linea di principio, queste dimostrazioni si potrebbero "riadattare" siccome le due norme sono equivalenti.

Lemma B.4.1. $\forall \Omega \subseteq \mathbb{R}^n, \text{Fr } \Omega$ è un insieme chiuso.

Dimostrazione. La dimostrazione è immediata: sappiamo dall'esercizio 1.8 che la frontiera di Ω è definita come $\text{Fr } \Omega = \overline{\Omega} \cap \overline{\Omega^c}$. Conseguentemente, abbiamo che $\overline{\Omega}$ è un insieme chiuso, $\overline{\Omega^c}$ è chiuso e abbiamo che

$$(\overline{\Omega} \cap \overline{\Omega^c})^c = (\mathbb{R}^n \setminus \overline{\Omega}) \cup (\mathbb{R}^n \setminus \overline{\Omega^c})$$

ma, per quanto mostrato nella proposizione 1.10, abbiamo che $\mathbb{R}^n \setminus \overline{\Omega}$ e $\mathbb{R}^n \setminus \overline{\Omega^c}$ sono insiemi aperti e, per quanto visto nell'esercizio 1.2, abbiamo che l'unione di insiemi aperti è un insieme aperto. Ma allora il complementare di $\text{Fr } \Omega$ è aperto, dunque, sempre per la proposizione 1.10, avremo che è un insieme chiuso. \square

Lemma B.4.2. Siano dati $A \subseteq \mathbb{R}^n$ aperto e $B \subseteq \mathbb{R}^n$ chiuso. Allora $A \setminus B$ è aperto e $B \setminus A$ è chiuso.

Dimostrazione. Siccome B è chiuso, allora $B^c = \mathbb{R}^n \setminus B$ è aperto in virtù della proposizione 1.10, dunque

$$A \setminus B = A \cap B^c = A \cap (\mathbb{R}^n \setminus B)$$

tuttavia, per quanto visto nell'esercizio 1.3, abbiamo che l'intersezione di aperti è un insieme aperto. Per quanto riguarda $B \setminus A$ osserviamo che

$$(B \cap (\mathbb{R}^n \setminus A))^c = B^c \cup (\mathbb{R}^n \setminus A)^c = B^c \cup A$$

che è aperto come conseguenza dell'esercizio 1.2, dunque $B \cap (\mathbb{R}^n \setminus A)$ è chiuso sempre per la proposizione 1.3 ma

$$B \setminus A = B \cap (\mathbb{R}^n \setminus A) \implies B \setminus A \text{ è aperto.}$$

\square

Procediamo nella dimostrazione del teorema della funzione inversa

Dimostrazione (del teorema B.1). Denotiamo con $\mathbb{R}^{n \times n} \ni A = Df(x_0)$. L'idea è quello di utilizzare il teorema di Banach-Caccioppoli per mostrare che, preso un insieme $U \subset f(B(x_0, r))$ aperto sufficientemente piccolo

Siccome Ω è aperto e $f \in C^1(\Omega, \mathbb{R}^n) \implies Df : \Omega \rightarrow \mathbb{R}^n$ è continuo, allora possiamo scegliere $\rho > 0$ in maniera tale che

$$\overline{B_\rho(x_0)} \subset \Omega \text{ e } |Df(x) - Df(x_0)| < \frac{1}{2\|A\|} \forall x \in \overline{B_\rho(x_0)}.$$

Poniamo $r = \frac{\rho}{2\|A\|}$, $y_0 = f(x_0)$. Fissato $y \in B_r(y_0)$, consideriamo l'applicazione

$$T : \overline{B_\rho(x_0)} \mapsto \mathbb{R}^n \text{ tale che } T(x) = x + A(y - f(x))$$

e osserviamo che $T(x) = x \iff y = f(x)$. Allora, ricordando che $ADf(x_0) = \text{Id}$, avremo, per $x \in \overline{B_\rho(x_0)}$, che

$$\|DT(x)\| = \|\text{Id} - ADf(x)\| \leq \|A\| \|Df(x) - Df(x_0)\| \leq \|A\| \frac{1}{2\|A\|} = \frac{1}{2}$$

dunque, per il criterio di Lipschitz, avremo che la funzione T è $\frac{1}{2}$ -lipschitziana. Mostriamo adesso che $T(\overline{B_\rho(x_0)}) \subset \overline{B_\rho(x_0)}$: infatti, preso $x \in \overline{B_\rho(x_0)}$, abbiamo che

$$\begin{aligned} |T(x) - x_0| &\stackrel{\text{dis. triang.}}{\leq} |T(x) - T(x_0)| + |T(x_0) - x_0| \leq \frac{1}{2}|x - x_0| + |A(y - y_0)| \leq \frac{\rho}{2} + \|A\|r = \\ &= \frac{\rho}{2} + \|A\| \frac{r}{2\|A\|} = \rho. \end{aligned}$$

Dunque $T : \overline{B_\rho(x_0)} \mapsto \overline{B_\rho(x_0)}$ è una contrazione, $B_\rho(x_0)$ è uno spazio metrico completo (rispetto alla metrica indotta) e, quindi, per il teorema delle contrazioni esiste un'unica soluzione all'equazione

$$T(x) = x.$$

Ma prendendo arbitrariamente $y_0 \in B_\rho(y_0)$ possiamo trovare un unico $x \in B_\rho(x_0)$ che risolve l'equazione $f(x) = y$ pertanto possiamo definire g l'applicazione $y \in B_r(y_0) \mapsto x \in B_\rho(x_0)$

$$g : B_r(y_0) \subset \overline{B_\rho(x_0)} \mapsto \overline{B_\rho(x_0)}, \quad f(g(y)) = x.$$

Cerchiamo di capire, adesso, dove viene mandata la frontiera di $B_\rho(x_0)$. Consideriamo l'insieme $\Sigma = f(\text{Fr } B_\rho(x_0))$: dalla continuità di f e dal fatto che $\text{Fr } B_\rho(x_0)$ è compatto (è banalmente limitato, è chiuso in virtù del lemma B.4.1 e, dunque, è compatto per il teorema 2.4) deriva che Σ è compatto. Inoltre $y_0 \notin \Sigma$ siccome, per quanto mostrato in precedenza, $\exists! x \in \overline{B_\rho(x_0)}$ per cui $f(x) = y_0$ e tale punto è proprio $x_0 \neq \text{Fr } B_\rho(x_0)$. Dunque avremo che $V = B_r(y_0) \setminus \Sigma$ è un insieme aperto (in virtù del lemma B.4.2). Posto $U = g(V)$ risulta che $f|_U : U \rightarrow V$ è bigettiva e ha g come inversa. Inoltre, per costruzione, abbiamo evitato che U possieda tutti i punti di $\text{Fr } B_\rho(x_0)$ e, dunque, risulta che $U \subset B_\rho(x_0)$.

Vogliamo mostrare adesso che U è aperto: preso $x \in U$ poniamo $y = f(x)$ da cui, per come abbiamo definito U e V , si deduce che $y \in V$. Siccome quest'ultimo insieme è aperto, sappiamo che $\exists \varepsilon > 0 : B_\varepsilon(y) \subset V$ ed, essendo f continua, possiamo trovare $\delta > 0$ tale che

$$f(B_\delta(x)) \subset B_\varepsilon(y) \subset V,$$

ma allora, siccome $x \in U \subset B_\rho(x_0)$, a meno di rimpicciolire δ possiamo supporre che $B_\delta(x) \subset B_\rho(x_0)$. Risulta, dunque, che $f(B_\delta(x)) \subset V \implies B_\delta(x) = g(f(B_\delta(x))) \subset U$. Dunque U contiene

un intorno di ogni suo punto, dunque è aperto.

Mostriamo, infine, che g è differenziale in ogni suo punto $y \in B_r(y_0)$ e che il suo differenziale $Dg(y)$ è uguale a $Df(x)^{-1}$ con $x = g(y)$, ossia si vuole mostrare che

$$\lim_{y' \rightarrow y} \frac{g(y') - g(y) - Df(x)^{-1}(y' - y)}{|y' - y|} = 0.$$

Osserviamo, tuttavia, che $y \in B_r(y_0)$ dunque stiamo considerando la seguente quantità (che abbiamo riscritto sfruttando la differenziabilità della funzione f)

$$\begin{aligned} \frac{g(y') - g(y) - Df(x)^{-1}(f(x') - f(x))}{|y' - y|} &= \frac{x' - x - Df(x)^{-1}(Df(x)(x' - x) + o(|x' - x|))}{|y' - y|} = \\ &= \frac{Df(x)^{-1}(o(|x' - x|))}{|y' - y|} = \frac{Df(x)^{-1}(o(|x' - x|))}{|x' - x|} \frac{|x' - x|}{|y' - y|}. \end{aligned}$$

Per concludere, ci basta mostrare che se $y' \rightarrow y \implies x' \rightarrow x$, cosicché il primo termine dell'ultimo prodotto tende a zero per definizione di o -piccolo. Per fare questo, possiamo riutilizzare l'applicazione T definita in precedenza: ricordando che essa è $\frac{1}{2}$ -lipschitziana abbiamo che

$$\begin{aligned} \frac{1}{2}|x' - x| &\geq |T(x') - T(x)| = |x' - x + A(y - f(x))| \geq |x' - x| - |A(y - f(x'))| = \\ &= |x' - x| - ||A|| |y' - y| \end{aligned}$$

da cui possiamo dedurre che $|x' - x| \leq 2||A|| |y' - y|$, pertanto $x' \rightarrow x$ se $y' \rightarrow y$ e per $y' \rightarrow y$ il rapporto $\frac{|x' - x|}{|y' - y|}$ è limitato. Questo ci porta a concludere che

$$\frac{Df(x)^{-1}(o(|x' - x|))}{|x' - x|} \frac{|x' - x|}{|y' - y|} \rightarrow 0 \text{ per } y' \rightarrow y.$$

Questo ci permette di concludere che la funzione g è differenziale in y (e dall'arbitrarietà di tale punto segue che è differenziale per ogni punto in $B_r(y_0)$, dunque anche in $V \subset B_r(y_0)$). A questo punto, resta da mostrare che il differenziale $Dg(y)$ è una funzione continua, ma dato che $Dg(y) = Df(g(y))^{-1}$ si osserva che essendo g continua (in quanto differenziabile), essendo Df continua (per ipotesi) ed essendo continua anche la mappa che associa ad una matrice la sua inversa (in virtù del teorema B.4), la funzione Dg dev'essere continua. \square

APPENDICE C

Il prodotto esterno

Le forme differenziali, dal punto di vista storico, derivano dalla generalizzazione al caso multidimensionale di concetti quali il lavoro di un campo su un cammino e il flusso di un liquido attraverso una superficie. In fisica classica, esse sono particolarmente rilevanti per la piena comprensione del formalismo hamiltoniano della meccanica classica mentre in fisica moderna sono fondamentali per la teoria della relatività generale e per la teoria dei campi quantistici.

C.1 Forme algebriche esterne

Vogliamo introdurre definendo le forme algebriche esterne, strettamente collegate alle forme differenziali. Partiamo dalle 1-forme algebriche esterne

Definizione C.1. Sia $\omega : \mathbb{R}^n \rightarrow \mathbb{R}$, diremo che essa è 1-forma algebrica esterna se è una funzione lineare.

Bibliografia

- [1] Serge Lang. *Linear Algebra*. Undergraduate Texts in Mathematics. Springer New York, NY, third edition, 2010.
- [2] Edoardo Sernesi. *Geometria 1*. Programma di Matematica, Fisica, Elettronica. Bollati Boringhieri, second edition, 2000.
- [3] W. Rudin. *Principles of Mathematical Analysis*. McGraw Hill, third edition, 1976.
- [4] Fulvio Ricci Luigi Ambrosio, Carlo Mantegazza. *Complementi di Matematica*. Edizioni della Normale Pisa, seconda edition, 2023.
- [5] Marco Manetti. *Topologia*. UNITEXT. Springer Milano, seconda edition, 2014.
- [6] James Munkres. *Topology*. Pearson Education, second edition, 2013.
- [7] Terence Tao. *Analysis II*, volume 38 of *Texts and Readings in Mathematics*. Hindustan Book Agency, third edition, 2016.
- [8] Andrea Mennucci Luigi D'Ambrosio, Giuseppe Prato. *Introduction to Measure Theory and Integration*. Publications of the Scuola Normale Superiore. Edizioni della Normale Pisa, 2011.