

Sapienza, Università di Roma
a.a. 2008/09



Note per il corso di

Analisi

Parte seconda

versione 1.0 (January 27, 2009)

L.Lamberti e C.Mascia

Corsi di laurea in Fisica & Fisica ed Astrofisica



Licenza © 2008 Lamberto Lamberti & Corrado Mascia

Distribuzione Creative Commons

Tu sei libero di riprodurre, stampare, inoltrare via mail, fotocopiare, distribuire questa opera alle seguenti condizioni:

- * **Attribuzione:** devi attribuire la paternità dell'opera nei modi indicati dall'autore o da chi ti ha dato l'opera in licenza,
- * **Non commerciale:** non puoi usare quest'opera per fini commerciali,
- * **Non opere derivate:** Non puoi alterare o trasformare quest'opera, né usarla per crearne un'altra.

(Licenza Creative Commons *Attribuzione - Non commerciale - Non opere derivate 3.0*)

Testo completo: <http://creativecommons.org/licenses/by-nc-nd/3.0/>)

Indice

Capitolo 1. L'integrale	1
1. L'area di un sottografico e la definizione di integrale	1
2. Istruzioni per l'uso	11
3. Il Teorema della media integrale	18
4. Il Teorema fondamentale del calcolo integrale	20
Capitolo 2. Zoologia dell'integrazione	25
1. Metodo di sostituzione	26
2. Integrazione per parti	31
3. Integrazione di funzioni razionali	35
Capitolo 3. I numeri complessi	43
1. Nascita dei numeri complessi e loro infanzia	43
2. Successioni, serie e continuità nei complessi	47
3. L'esponenziale complesso	52
Capitolo 4. Equazioni differenziali	55
1. Equazioni lineari del I ordine a coefficiente costante	55
2. Equazioni lineari del II ordine a coefficienti costanti omogenee	59
3. Equazioni lineari del II ordine a coefficienti costanti non omogenee	66
Capitolo 5. Lo spazio reale multidimensionale	71
1. Serve una struttura metrica	71
2. Quattro salti in \mathbb{R}^d : successioni di punti	76
3. Informazioni di base sulla topologia di \mathbb{R}^d	81
4. <i>Natura non facit saltus</i> : le curve	85
Capitolo 6. Funzioni di più variabili	97
1. Questioni elementari	97
2. Le derivate parziali	105
3. Integrali curvilinei	110

CAPITOLO 1

L'integrale

Il problema della misurazione delle lunghezze di segmenti ci ha condotti per un sentiero particolarmente interessante: dai numeri naturali, ai relativi, ai razionali ed infine ai numeri reali. E se ci viene la voglia di misurare aree di regioni del piano? Nel caso di un rettangolo, l'area è nota: è il prodotto delle lunghezze dei lati. Nel caso di una regione che sia unione finita di rettangoli che si toccano al più lungo il perimetro, basta sommare le aree dei singoli rettangoli. Ma per regioni più generali?

1. L'area di un sottografico e la definizione di integrale

Data una funzione $f : [a, b] \rightarrow [0, +\infty)$, l'insieme

$$S_f = \{(x, y) \mid a \leq x \leq b, \quad 0 \leq y \leq f(x)\}$$

si dice **sottografico di f** nell'intervallo $[a, b]$. Come definire e/o calcolare l'area di un sottografico di funzione? L'area del sottografico delle funzioni costanti è elementare: se $f(x) = C \geq 0$, il sottografico S_f di f in $[a, b]$ è un rettangolo, la cui area è, da sempre, $\mathcal{A}(S_f) = (b - a)C$. Con poco impegno, possiamo definire una classe di funzioni il cui sottografico ha un'area facile da calcolare.

DEFINIZIONE 1.1. *Dato l'intervallo $[a, b]$, un insieme $P = \{x_0, x_1, \dots, x_n\}$ tale che $a = x_0 < x_1 < \dots < x_n = b$ è una **partizione** di $[a, b]$. L'ampiezza della partizione P è il numero $|P| := \max\{x_i - x_{i-1} : i = 1, \dots, n\}$.*

*Una funzione $f : [a, b] \rightarrow \mathbb{R}$ è una **funzione costante a tratti** (o **funzione a scala**) se esiste una partizione $P = \{x_0, x_1, \dots, x_n\}$ di $[a, b]$ tale che f è costante su ogni intervallo $[x_{i-1}, x_i]$ per $i = 1, \dots, n - 1$ e su $[x_{n-1}, x_n]$, ossia*

$$f(x) = \begin{cases} \alpha_i & x \in [x_{i-1}, x_i) \\ \alpha_n & x \in [x_{n-1}, x_n] \end{cases} \quad i = 1, \dots, n - 1$$

dove $\alpha_1, \dots, \alpha_n$ sono numeri reali.

Dato che per una funzione f costante a tratti e non negativa il sottografico è un'unione finita di rettangoli, l'area di S_f è data dalla somma delle aree di questi

rettangoli:

$$\mathcal{A}(S_f) = \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) = \sum_{i=1}^n \alpha_i(x_i - x_{i-1})$$

dove ξ_i è un qualsiasi punto di $[x_{i-1}, x_i]$.

Data una funzione $f : [a, b] \rightarrow \mathbb{R}$, non negativa, possiamo considerare approssimazioni del sottografico S_f date da sottografici di funzioni costanti a tratti g ed h , con $g \leq f \leq h$ in $[a, b]$ (vedi Figura 1(a)) e utilizzare le aree dei sottografici approssimanti come approssimazioni dell'area di S_f . Per poter stimare per eccesso la funzione f con una funzione costante a tratti non negativa occorre che la funzione f sia *superiormente limitata* (vedi Figura 1(b)). Per la stima dal basso con funzioni costanti a tratti, serve anche che f sia inferiormente limitata, ipotesi che nel nostro presente caso è automaticamente soddisfatta dato che f è non negativa. L'ipotesi di limitatezza della funzione f è fondamentale qui, così come in tutto il Capitolo e in tutta la definizione dell'integrale definito.

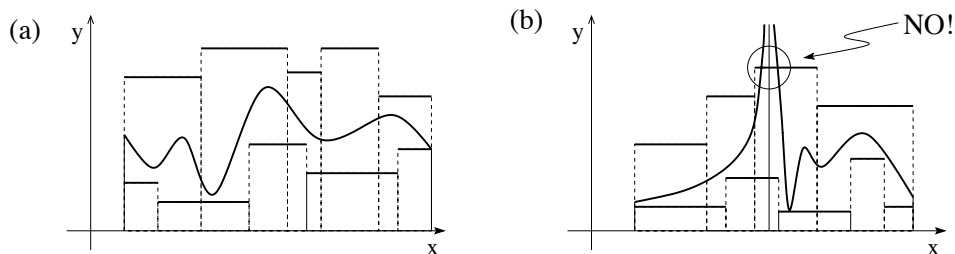


FIGURA 1. (a) Una funzione approssimata per eccesso e per difetto con due funzioni costanti a tratti. (b) Un tentativo (fallito) di approssimazione per eccesso di una funzione non limitata superiormente.

Data una partizione $P = \{x_0, x_1, \dots, x_n\}$ ci sono due funzioni \underline{f} e \bar{f} , costanti a tratti negli intervalli definiti da P e tali che $\underline{f} \leq f \leq \bar{f}$, che danno la migliore stima per difetto e la migliore per eccesso. Tali funzioni sono definite da

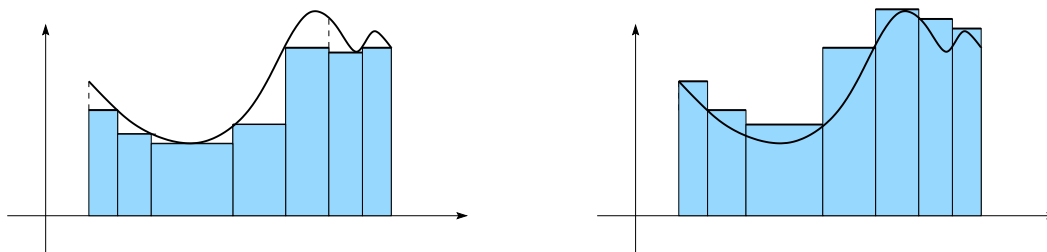
$$\begin{aligned} \underline{f}(x) &= \alpha_i := \inf_{[x_{i-1}, x_i]} f(x) & \forall x \in [x_{i-1}, x_i], \\ \bar{f}(x) &= \beta_i := \sup_{[x_{i-1}, x_i]} f(x) & \forall x \in [x_{i-1}, x_i]. \end{aligned}$$

In definitiva, data una partizione P dell'intervallo $[a, b]$, (tutte le volte che $\mathcal{A}(S_f)$ ha senso) si ha

$$\underline{\mathcal{A}}(S_f; P) \leq \mathcal{A}(S_f) \leq \bar{\mathcal{A}}(S_f; P).$$

dove

$$\underline{\mathcal{A}}(S_f; P) := \mathcal{A}(S_{\underline{f}}) = \sum_{i=1}^n \alpha_i(x_i - x_{i-1}) \quad \bar{\mathcal{A}}(S_f; P) := \mathcal{A}(S_{\bar{f}}) = \sum_{i=1}^n \beta_i(x_i - x_{i-1}).$$

FIGURA 2. Due approssimazioni per S_f tramite le funzioni \underline{f} e \bar{f} .

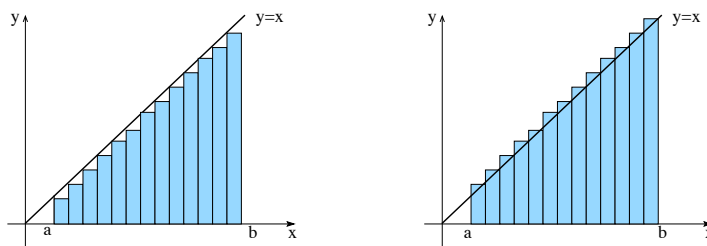
Dopo, si procede a migliorare l'approssimazione tramite la scelta di una partizione con un'ampiezza più piccola. Una possibilità è scegliere una successione di partizioni P_n con ampiezza $|P_n|$ che tenda a zero per $n \rightarrow \infty$. In questo modo, passando al limite nelle due successioni numeriche $\underline{A}(S_f; P_n)$ e $\bar{A}(S_f; P_n)$ si dovrebbe (incrociando le dita) ottenere proprio l'area richiesta.

ESEMPIO 1.2. Sia $f(x) = x$ per $x \in [a, b]$ con $0 \leq a < b$. Il sottografico della funzione f nell'intervallo $[a, b]$ ha una forma familiare: nel caso in cui $a = 0$, si tratta di un triangolo, e nel caso di $a > 0$ si tratta di un trapezio. In entrambi i casi, la geometria elementare fornisce una formula per il calcolo dell'area, che è data da:

$$\frac{(b-a)(b+a)}{2} = \frac{b^2}{2} - \frac{a^2}{2}.$$

Cosa succede se si calcola l'area attraverso il procedimento di approssimazione per eccesso e per difetto proposto in precedenza?

Dividiamo l'intervallo $[a, b]$ in n parti di uguale lunghezza tramite la partizione $P_n = \{x_k = a + kh : k = 0, \dots, n\}$ dove $h = (b-a)/n$. Dato che $f(x) = x$ è crescente,

FIGURA 3. La funzione $f(x) = x$ e le approssimazioni determinate dai punti $a + \frac{k(b-a)}{n}$ con $k = 0, \dots, n$.

$$\alpha_k = \inf_{[x_{k-1}, x_k]} x = x_{k-1} = a + (k-1)h, \quad \beta_k = \sup_{[x_{k-1}, x_k]} x = x_k = a + kh,$$

e $x_k - x_{k-1} = h$, valgono

$$\underline{A}(S_f; P_n) = ah + \cdots + (a + (n-1)h)h = h \sum_{k=0}^{n-1} (a + kh) = hna + h^2 \sum_{k=0}^{n-1} k,$$

$$\overline{A}(S_f; P_n) = (a+h)h + \cdots + (a+nh)h = h \sum_{k=1}^n (a + kh) = hna + h^2 \sum_{k=1}^n k$$

Tenendo conto della formula $\sum_{k=1}^n k = \frac{n(n+1)}{2}$ (dimostratela!) e di $h = \frac{b-a}{n}$

$$\underline{A}(S_f; P_n) = hna + h^2 \frac{n(n-1)}{2} = (b-a)a + (b-a)^2 \frac{n(n-1)}{2n^2}$$

$$\overline{A}(S_f; P_n) = hna + h^2 \frac{n(n+1)}{2} = (b-a)a + (b-a)^2 \frac{n(n+1)}{2n^2}.$$

Per $n \rightarrow +\infty$, le due quantità tendono allo stesso limite:

$$\lim_{n \rightarrow +\infty} \underline{A}(S_f; P_n) = \lim_{n \rightarrow +\infty} \overline{A}(S_f; P_n) = (b-a)a + \frac{(b-a)^2}{2} = \frac{b^2}{2} - \frac{a^2}{2},$$

che concorda con le formule note dalla geometria elementare.

ESEMPIO 1.3. Passiamo a considerare $f(x) = x^2$ in $[a, b]$ con $0 \leq a < b$. Qual è l'espressione dell'area per il sottografico di x^2 ?

Con la stessa partizione P_n di prima. Dato che x^2 è crescente su $[a, b]$ per $a \geq 0$,

$$\alpha_i = \inf_{[x_{k-1}, x_k]} x^2 = (a + (k-1)h)^2, \quad \beta_i = \sup_{[x_{k-1}, x_k]} x^2 = (a + kh)^2,$$

da cui segue

$$\underline{A}(S_f; P_n) = a^2h + (a+h)^2h + (a+2h)^2h + \cdots + (a+(n-1)h)^2h$$

$$\overline{A}(S_f; P_n) = (a+h)^2h + (a+2h)^2h + \cdots + (a+nh)^2h$$

Svolti i quadrati e tenuto presente che $h = \frac{b-a}{n}$, si ottiene

$$\underline{A}(S_f; P_n) = h \left\{ na^2 + 2ah \sum_{k=0}^{n-1} k + h^2 \sum_{k=0}^{n-1} k^2 \right\}$$

$$= (b-a) \left\{ a^2 + \frac{2a(b-a)}{n^2} \sum_{k=0}^{n-1} k + \frac{(b-a)^2}{n^3} \sum_{k=0}^{n-1} k^2 \right\}$$

$$\overline{A}(S_f; P_n) = h \left\{ na^2 + 2ah \sum_{k=1}^n k + h^2 \sum_{k=1}^n k^2 \right\}$$

$$= (b-a) \left\{ a^2 + \frac{2a(b-a)}{n^2} \sum_{k=1}^n k + \frac{(b-a)^2}{n^3} \sum_{k=1}^n k^2 \right\}$$

Utilizzando le formule $\sum_{k=1}^n k = \frac{1}{2}n(n+1)$ e $\sum_{k=1}^n k^2 = \frac{1}{6}n(n+1)(2n+1)$ (dimostratele!), si ottiene

$$\begin{aligned}\underline{\mathcal{A}}(S_f; P_n) &= (b-a) \left\{ a^2 + a(b-a) \frac{n-1}{n} + (b-a)^2 \frac{(n-1)(2n-1)}{6n^2} \right\} \\ \overline{\mathcal{A}}(S_f; P_n) &= (b-a) \left\{ a^2 + a(b-a) \frac{n+1}{n} + (b-a)^2 \frac{(n+1)(2n+1)}{6n^2} \right\}\end{aligned}$$

Quindi, passando al limite per $n \rightarrow \infty$,

$$\lim_{n \rightarrow +\infty} \underline{\mathcal{A}}(S_f; P_n) = \lim_{n \rightarrow +\infty} \overline{\mathcal{A}}(S_f; P_n) = \frac{b^3}{3} - \frac{a^3}{3} =: \mathcal{A}(S_f).$$

che è il valore dell'area cercato.

ESEMPIO 1.4. Sperimentiamo la tecnica per una funzione più bizzarra:

$$\text{funzione di Dirichlet: } D(x) := \begin{cases} 0 & x \in [0, 1] \setminus \mathbb{Q}, \\ 1 & x \in [0, 1] \cap \mathbb{Q}, \end{cases}$$

Data una qualsiasi partizione $P = \{a = x_0 < x_1 < \dots < x_n = b\}$, in ognuno dei sottointervalli $[x_{k-1}, x_k]$ cadono sia numeri razionali che numeri irrazionali, quindi

$$\alpha_i = \inf_{[x_{k-1}, x_k]} D(x) = 0, \quad \beta_i = \sup_{[x_{k-1}, x_k]} D(x) = 1.$$

Pertanto, per ogni partizione P , $\underline{\mathcal{A}}(S_D; P) = 0$ e $\overline{\mathcal{A}}(S_D; P) = 1$ e quindi

$$\sup_P \underline{\mathcal{A}}(S_D; P) = 0 < 1 = \inf_P \overline{\mathcal{A}}(S_D; P).$$

Anche scegliendo partizioni P con ampiezza sempre più piccola, le stime per difetto e quelle per eccesso restano sempre ben lontane le une dalle altre. L'interpretazione che diamo di questa situazione è che *esistono sotto insiemi del piano a cui non è possibile associare un'area*, ossia, per alcuni insiemi S , l'espressione $\mathcal{A}(S)$ non ha senso!

ESEMPIO 1.5. Ancora un esempio: $f(x) = e^x$ in $[a, b]$ con $a < b$. Scegliamo ancora una volta la partizione $P_n = \{x_k = a + kh : k = 0, \dots, n\}$ dove $h = (b-a)/n$. La funzione e^x è crescente su $[a, b]$ quindi

$$\alpha_i = \inf_{[x_{k-1}, x_k]} e^x = e^{x_{k-1}} = e^a e^{(k-1)h}, \quad \beta_i = \sup_{[x_{k-1}, x_k]} e^x = e^{x_k} = e^a e^{kh}.$$

Perciò

$$\begin{aligned}\underline{\mathcal{A}}(S_f; P_n) &= e^a h [1 + e^h + e^{2h} + \dots + e^{(n-1)h}] = e^a h [1 + e^h + (e^h)^2 + \dots + (e^h)^{n-1}] \\ \overline{\mathcal{A}}(S_f; P_n) &= e^a h [e^h + e^{2h} + e^{3h} + \dots + e^{nh}] = e^{a+h} h [1 + e^h + (e^h)^2 + \dots + (e^h)^{n-1}]\end{aligned}$$

Dato che $1 + e^h + (e^h)^2 + \dots + (e^h)^{n-1} = (e^{nh} - 1)/(e^h - 1)$ e $nh = b - a$,

$$\underline{A}(S_f; P_n) = \frac{e^a h (e^{b-a} - 1)}{e^h - 1} = \frac{h}{e^h - 1} (e^b - e^a)$$

$$\overline{A}(S_f; P_n) = \frac{e^{a+h} h (e^{b-a} - 1)}{e^h - 1} = \frac{h}{e^h - 1} e^h (e^b - e^a)$$

Quindi, passando al limite per $n \rightarrow \infty$,

$$\lim_{n \rightarrow +\infty} \underline{A}(S_f; P_n) = \lim_{n \rightarrow +\infty} \overline{A}(S_f; P_n) = e^b - e^a,$$

che è il valore cercato.

Tiriamo le fila di quello che abbiamo fatto fin qui. Considerando l'“area di una regione del piano” un concetto intuitivo, abbiamo considerato il caso di sottografici di funzioni non negative, proponendo un algoritmo per il calcolo dell'area: approssimare per difetto e per eccesso l'area richiesta tramite aree di sottografici di funzioni costanti a tratti e, facendo tendere l'ampiezza della partizione a 0, ottenere il valore dell'area cercata. Tramite un certo numero di esempi, ci siamo resi conto che il procedimento è ragionevole, ma che in alcune situazioni non porta a nessuna conclusione (come per la funzione di Dirichlet).

Mettendo da parte, per il momento, il problema dell'area e battezziamo le funzioni per cui il procedimento di sopra descritto converge *funzioni integrabili* e il valore limite delle approssimazioni *integrale (definito) della funzione*. Rispetto a quanto fatto in precedenza, c'è una differenza essenziale: *le funzioni che consideriamo possono avere segno qualsiasi*.

Sia $f : [a, b] \rightarrow \mathbb{R}$ limitata e $P = \{x_0, x_1, \dots, x_n\}$ una partizione di $[a, b]$. Poniamo

$$\alpha_i := \inf_{[x_{i-1}, x_i]} f(x) \quad \text{e} \quad \beta_i := \sup_{[x_{i-1}, x_i]} f(x).$$

DEFINIZIONE 1.6. Somme integrali. *Si chiamano somma integrale per difetto/per eccesso di f relativamente alla partizione P , i valori delle somme*

$$\underline{S}(f; P) := \sum_{i=1}^n \alpha_i (x_i - x_{i-1}) \quad \text{e} \quad \overline{S}(f; P) := \sum_{i=1}^n \beta_i (x_i - x_{i-1}).$$

La condizione di limitatezza della funzione f garantisce che, per ogni partizione P , le somme integrali per difetto e per eccesso sono sempre ben definite.

DEFINIZIONE 1.7. Integrale definito. *Una funzione $f : [a, b] \rightarrow \mathbb{R}$ limitata si dice integrabile (secondo Riemann) se vale l'uguaglianza*

$$(1) \quad \sup_P \underline{S}(f; P) = \inf_P \overline{S}(f; P),$$

dove l'estremo superiore e l'estremo inferiore sono presi nell'insieme di tutte le partizioni P dell'intervallo $[a, b]$.

Il valore comune è l'integrale definito di f in $[a, b]$ e si indica¹ con

$$\int_a^b f(x) dx.$$

La lettera usata per indicare la variabile di integrazione è indifferente: al posto di $\int_a^b f(x) dx$, si può scrivere $\int_a^b f(t) dt$ o $\int_a^b f(u) du$ (come per le sommatorie, in cui il nome dato all'indice è ininfluente).

OSSERVAZIONE 1.8. Nella definizione di integrale, non viene fatta nessuna richiesta di positività della funzione. Se la *funzione integranda* f è positiva in $[a, b]$ ed è integrabile, l'integrale da la definizione di area del sottografo di f in $[a, b]$

$$f \geq 0 \quad \Rightarrow \quad \mathcal{A}(\{(x, y) : a \leq x \leq b, 0 \leq y \leq f(x)\}) := \int_a^b f(x) dx.$$

Se f è negativa in tutto o in parte dell'intervallo, il significato dell'integrale non è più quello di un'area: *l'integrale è somma di termini positivi e negativi, gli uni e gli altri in corrispondenza delle zone in cui il grafico è sopra o sotto l'asse x* (vedi Figura 4).

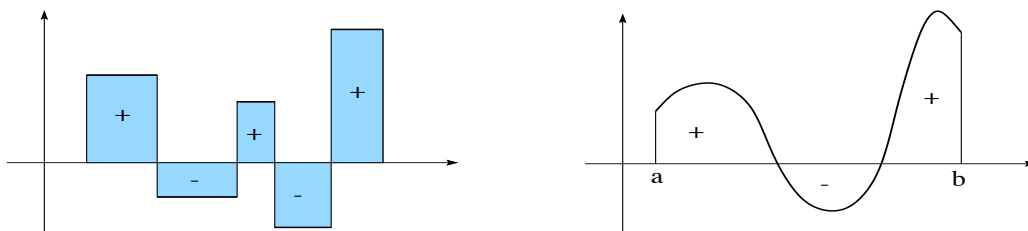


FIGURA 4. L'integrale non è l'area: (a) una funzione costante a tratti, (b) una funzione qualsiasi.

La condizione di integrabilità data dalla definizione è chiara e limpida da un punto di vista di rigore matematico. Meno chiaro è come rispondere concretamente alla domanda: quali classi di funzioni sono integrabili?

PROPOSIZIONE 1.9. Una funzione $f : [a, b] \rightarrow \mathbb{R}$ è integrabile in $[a, b]$ se e solo se

$$(2) \quad \forall \varepsilon > 0 \quad \exists P_\varepsilon \quad t.c. \quad \overline{S}(f; P_\varepsilon) - \underline{S}(f; P_\varepsilon) < \varepsilon.$$

¹Il simbolo dell'integrale è una variante del simbolo di somma indicato da una lunga S come si usava al tempo di Leibnitz. Il simbolo dx è l'erede della lunghezza dell'intervallo $x_i - x_{i-1}$. L'uso della "d" minuscola ricorda che le approssimazioni migliori si ottengono considerando partizioni con ampiezza piccola: $\Delta x := x_i - x_{i-1} \approx dx$.

La condizione espressa in (2) può essere riscritta come

$$\overline{S}(f; P_\varepsilon) - \underline{S}(f; P_\varepsilon) = \sum_{i=1}^n (\beta_i - \alpha_i) (x_i - x_{i-1}) < \varepsilon$$

con α_i e β_i definiti in precedenza. Dato che tale condizione è sufficiente, per dimostrare l'integrabilità di una funzione occorre stimare la differenza $\beta_i - \alpha_i$, cioè la variazione $|f(y) - f(x)|$ per $x, y \in [x_{i-1}, x_i]$ per partizioni con ampiezza piccola.

DIMOSTRAZIONE DELLA PROPOSIZIONE 1.9. La definizione di integrabilità e le proprietà dell'estremo superiore e dell'estremo inferiore indicano che per ogni $\varepsilon > 0$, esistono partizioni P'_ε e P''_ε tali che $\overline{S}(f; P'_\varepsilon) - \underline{S}(f; P''_\varepsilon) < \varepsilon$. Occorre ora dimostrare che la stessa stima vale per una scelta opportuna di una stessa partizione P_ε .

Passo 1. Dimostriamo prima di tutto che

$$(3) \quad \underline{S}(f; P_1) \leq \underline{S}(f; P_2) \quad \text{e} \quad \overline{S}(f; P_2) \leq \overline{S}(f; P_1) \quad \forall P_2 \subset P_1.$$

Dato che $P_2 \subset P_1$, si può costruire P_2 a partire da P_1 aggiungendo un numero finito di punti, basta studiare il caso in cui $P_2 = P_1 \cup \{\xi\}$ (il caso generale si ottiene iterando il procedimento). Inoltre consideriamo solo le somme per difetto, l'altra parte è analoga.

Supponiamo $P_1 = \{x_0, x_1, \dots, x_n\}$ e $\xi \in (x_{k-1}, x_k)$ per un opportuno $k \in \{1, \dots, n\}$. Le espressioni di $\underline{S}(f; P_1)$ e di $\underline{S}(f; P_2)$ coincidono in tutti i termini tranne in quelli relativi all'intervallo $[x_{k-1}, x_k]$ e quindi, indicando con

$$\alpha = \inf_{[x_{k-1}, x_k]} f(x), \quad \alpha' = \inf_{[x_{k-1}, \xi]} f(x), \quad \alpha'' = \inf_{[\xi, x_k]} f(x),$$

vale, dato che $\alpha \leq \alpha'$ e $\alpha \leq \alpha''$,

$$\begin{aligned} \underline{S}(f; P_2) - \underline{S}(f; P_1) &= \left[\alpha'(x_k - \xi) + \alpha''(\xi - x_{k-1}) \right] - \alpha(x_k - x_{k-1}) \\ &= \left[\alpha'(x_k - \xi) + \alpha''(\xi - x_{k-1}) \right] - \left[\alpha(x_k - \xi) + \alpha(\xi - x_{k-1}) \right] \\ &= (\alpha' - \alpha)(x_k - \xi) + (\alpha'' - \alpha)(\xi - x_{k-1}) \geq 0, \end{aligned}$$

Passo 2. Ora dimostriamo che

$$\underline{S}(f; P') \leq \overline{S}(f; P'') \quad \forall P', P'' \text{ partizioni.}$$

Infatti, per costruzione, $\underline{S}(f; P) \leq \overline{S}(f; P)$ per ogni partizione P . Quindi, se consideriamo la partizione $P = P' \cup P''$ e utilizziamo (3):

$$\underline{S}(f; P') \leq \underline{S}(f; P) \leq \overline{S}(f; P) \leq \overline{S}(f; P'').$$

Passo 3. Infine, dimostriamo la Proposizione. Per ogni partizione P , vale

$$0 \leq \inf_P \overline{S}(f; P) - \sup_P \underline{S}(f; P) \leq \overline{S}(f; P) - \underline{S}(f; P)$$

Se per ogni $\varepsilon > 0$ esiste una partizione P_ε che verifica (2), allora

$$0 \leq \inf_P \overline{S}(f; P) - \sup_P \underline{S}(f; P) < \varepsilon \quad \forall \varepsilon > 0,$$

cioè $\inf_P \overline{S}(f; P) = \sup_P \underline{S}(f; P)$, e f è integrabile.

Se invece supponiamo che la funzione f sia integrabile, per ogni $\varepsilon > 0$ esistono due partizioni P'_ε e P''_ε per cui vale $\overline{S}(f; P'_\varepsilon) - \underline{S}(f; P''_\varepsilon) < \varepsilon$. Scegliendo $P_\varepsilon = P'_\varepsilon \cup P''_\varepsilon$, grazie a $\underline{S}(f; P''_\varepsilon) \leq \underline{S}(f; P_\varepsilon)$ e $\overline{S}(f; P_\varepsilon) \leq \overline{S}(f; P'_\varepsilon)$,

$$\overline{S}(f; P_\varepsilon) - \underline{S}(f; P_\varepsilon) \leq \overline{S}(f; P'_\varepsilon) - \underline{S}(f; P''_\varepsilon) < \varepsilon,$$

cioè la condizione (2). □

Se si costruisce una successione di partizioni P_n per cui

$$\lim_{n \rightarrow +\infty} \underline{S}(f; P_n) = \lim_{n \rightarrow +\infty} \overline{S}(f; P_n) = \ell,$$

a maggior ragione si avrà che: per ogni $\varepsilon > 0$ esiste una partizione P_n per cui vale

$$0 \leq \overline{S}(f; P_n) - \underline{S}(f; P_n) = [\overline{S}(f; P_n) - \ell] + [\ell - \underline{S}(f; P_n)] < 2\varepsilon$$

cioè la funzione f è integrabile e l'integrale è il valore ℓ . In particolare, gli Esempi 1.2, 1.3, 1.5 indicano che le funzioni x , x^2 e e^x sono integrabili in $[a, b]$ e che valgono

$$\int_a^b x \, dx = \frac{b^2}{2} - \frac{a^2}{2}, \quad \int_a^b x^2 \, dx = \frac{b^3}{3} - \frac{a^3}{3}, \quad \int_a^b e^x \, dx = e^b - e^a.$$

L'Esempio 1.4 mostra che *esistono anche funzioni non integrabili!*

Il volume di solidi di rotazione. L'integrale è utile per il calcolo delle aree, ma anche, in alcune situazioni speciali, per il calcolo di volumi di solidi. Consideriamo una funzione $f : [a, b] \rightarrow [0, +\infty)$ continua e disegniamone il grafico Γ_f . Se si fa ruotare il grafico Γ_f attorno all'asse x , si ottiene la superficie laterale di un solido Σ . Qual'è il suo volume? Come per il calcolo delle aree, anche nel calcolo dei volumi bisogna partire da una formula nota. Qui, diamo per buona la regola che ci è stata insegnata da bambini: il volume di un cilindro di raggio di base r e altezza h è $\pi r^2 h$.

Ragionando in maniera simile a quanto fatto per la determinazione dell'area di un sottografo, approssimiamo il solido con l'unione di oggetti di cui conosciamo il volume. La scelta più ragionevole è l'unione di cilindri ottenuti tramite una rotazione di rettangoli con lati paralleli agli assi x e y come in Figura 5(b) Un'approssimazione di questo genere di Σ , si ottiene tramite la scelta di una partizione $P = \{a \equiv x_0 <$

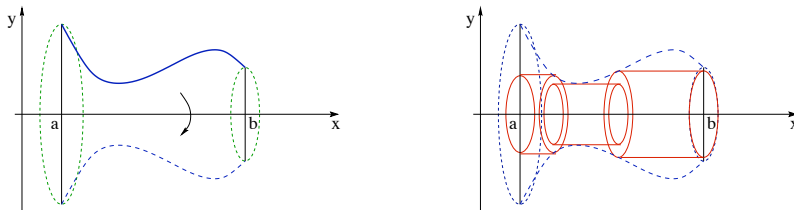


FIGURA 5. (a) Un solido di rotazione; (b) Una sua approssimazione con cilindri.

$x_1 < \dots < x_n = b$ dell'intervallo $[a, b]$ e di n punti ξ_1, \dots, ξ_n tali che $\xi_i \in [x_{i-1}, x_i]$. Quindi l'approssimazione del volume $\mathcal{V}(\Sigma)$ di Σ è data da

$$\mathcal{V}(\Sigma) \approx \sum_{i=1}^n \pi f^2(\xi_i)(x_i - x_{i-1}).$$

Quando l'ampiezza della partizione tende a zero, l'errore di approssimazione tende a zero e si ottiene la formula desiderata

$$\mathcal{V}(\Sigma) = \pi \int_a^b f^2(x) dx.$$

Collaudiamo la formula con un paio di esempi.

ESEMPIO 1.10. Il volume del cono. Un cono di altezza h e raggio di base r , si può ottenere tramite una rotazione del grafico della funzione

$$f(x) = \frac{rx}{h} \quad x \in [0, h].$$

Quindi

$$\mathcal{V}(\Sigma) = \pi \int_0^h \left(\frac{rx}{h}\right)^2 dx = \frac{\pi r^2}{h^2} \int_0^h x^2 dx = \frac{\pi r^2}{h^2} \frac{h^3}{3} = \frac{1}{3} \pi r^2 h,$$

che è la tradizionale formula del volume del cono.

ESEMPIO 1.11. Il volume della sfera. Una sfera di raggio r si ottiene con la rotazione del grafico di

$$f(x) = \sqrt{r^2 - x^2} \quad x \in [-r, r].$$

Applicando la formula del volume

$$\mathcal{V}(\Sigma) = \pi \int_{-r}^r \left(\sqrt{r^2 - x^2}\right)^2 dx = \pi \int_{-r}^r (r^2 - x^2) dx = \pi \left\{ 2r^3 - \frac{2}{3}r^3 \right\} = \frac{4}{3} \pi r^3.$$

Anche questa volta, giustamente, i conti tornano!

ESERCIZIO 1.12. Qual'è il volume del solido ottenuto facendo ruotare il grafico di e^x per $x \in [0, 1]$ attorno all'asse x ?

2. Istruzioni per l'uso

L'integrale di una funzione limitata è l'estremo superiore/inferiore di opportune sommatorie. Per le sommatorie valgono tre proprietà fondamentali: dati $a_1, \dots, a_n, b_1, \dots, b_n, \alpha, \beta \in \mathbb{R}$, si hanno

$$\text{linearità:} \quad \sum_{k=1}^n (\alpha a_k + \beta b_k) = \alpha \sum_{k=1}^n a_k + \beta \sum_{k=1}^n b_k$$

$$\text{additività:} \quad \sum_{k=1}^n a_k = \sum_{k=1}^m a_k + \sum_{k=m+1}^n a_k$$

$$\text{monotonìa:} \quad a_k \leq b_k \quad \Rightarrow \quad \sum_{k=1}^n a_k \leq \sum_{k=1}^n b_k.$$

Di conseguenza, analoghe proprietà vengono ereditate dagli integrali definiti.

Linearità. Una combinazione lineare di funzioni integrabile è integrabile: dati $c_1, c_2 \in \mathbb{R}$ e f, g integrabili in $[a, b]$, la funzione $c_1 f + c_2 g$ è integrabile e vale

$$(4) \quad \int_a^b [c_1 f(x) + c_2 g(x)] dx = c_1 \int_a^b f(x) dx + c_2 \int_a^b g(x) dx$$

Additività. Sia f integrabile in $[a, b]$ e sia $c \in (a, b)$. Allora f è integrabile in $[a, c]$ e in $[c, b]$ e vale

$$(5) \quad \int_a^b f(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx \quad \forall a < c < b.$$

In particolare, allora f è integrabile in ogni sottointervallo di $[a, b]$.

Monotonìa. Per ogni coppia di funzioni f, g integrabili in $[a, b]$,

$$(6) \quad f(x) \leq g(x) \quad \forall x \in [a, b] \quad \Rightarrow \quad \int_a^b f(x) dx \leq \int_a^b g(x) dx.$$

Per ora abbiamo definito $\int_a^b f(x) dx$ solo nel caso $a < b$. E' usanza diffusa definire l'integrale nel caso di $a = b$ o $a > b$, in modo che sia preservata la regola dell'additività. Scrivendo (5) con $c = a$, si ottiene

$$\int_a^b f(x) dx = \int_a^a f(x) dx + \int_a^b f(x) dx,$$

da cui segue

$$\int_a^a f(x) dx := 0,$$

coerente con l'eventuale interpretazione in termini di aree.

Scriviamo (5) per $b = a$, allora

$$0 = \int_a^a f(x) dx = \int_a^c f(x) dx + \int_c^a f(x) dx$$

da cui segue la definizione

$$\int_c^a f(x) dx := - \int_a^c f(x) dx \quad a < c,$$

dove il membro destro ha il significato definito nel paragrafo precedente.

Dalla proprietà di monotonia dell'integrale discende una proprietà che è, sostanzialmente, una "disuguaglianza triangolare per integrali":

PROPOSIZIONE 2.1. *Se la funzione limitata $f : [a, b] \rightarrow \mathbb{R}$ è integrabile in $[a, b]$, allora anche $|f|$ è integrabile in $[a, b]$ e vale*

$$(7) \quad \left| \int_a^b f(x) dx \right| \leq \int_a^b |f(x)| dx.$$

Dedichiamoci ora a dimostrare le proprietà dell'integrale. Le proprietà di linearità e di additività contengono due parti. Una prima parte concerne il fatto che dall'integrabilità di alcune funzioni se ne deduce l'integrabilità di certe altre. La seconda parte mette in relazione tra loro gli integrali definiti delle varie funzioni.

Linearità. Siano f e g integrabili in $[a, b]$ e $c_1, c_2 \in \mathbb{R}$. Data $P = \{x_0, x_1, \dots, x_n\}$, partizione di $[a, b]$, battezziamo

$$\alpha_i = \inf_{[x_{i-1}, x_i]} [c_1 f(x) + c_2 g(x)] \quad \text{e} \quad \beta_i = \sup_{[x_{i-1}, x_i]} [c_1 f(x) + c_2 g(x)],$$

e, analogamente,

$$\alpha_i^f = \inf_{[x_{i-1}, x_i]} f(x), \quad \beta_i^f = \sup_{[x_{i-1}, x_i]} f(x), \quad \alpha_i^g = \inf_{[x_{i-1}, x_i]} g(x), \quad \beta_i^g = \sup_{[x_{i-1}, x_i]} g(x).$$

Per ogni $x, y \in (x_{i-1}, x_i)$, si ha

$$\begin{aligned} [c_1 f(x) + c_2 g(x)] - [c_1 f(y) + c_2 g(y)] &= c_1 [f(x) - f(y)] + c_2 [g(x) - g(y)] \\ &\leq |c_1| (\beta_i^f - \alpha_i^f) + |c_2| (\beta_i^g - \alpha_i^g), \end{aligned}$$

quindi, passando all'estremo superiore in x e all'estremo inferiore in y , si deduce che

$$\beta_i - \alpha_i \leq |c_1| (\beta_i^f - \alpha_i^f) + |c_2| (\beta_i^g - \alpha_i^g).$$

Perciò, per ogni partizione P di $[a, b]$,

$$\begin{aligned} \overline{S}(c_1f + c_2g; P) - \underline{S}(c_1f + c_2g; P) &= \sum_{i=1}^n [\beta_i - \alpha_i] (x_i - x_{i-1}) \\ &\leq |c_1| \sum_{i=1}^n [\beta_i^f - \alpha_i^f] (x_i - x_{i-1}) + |c_2| \sum_{i=1}^n [\alpha_i^g - \beta_i^g] (x_i - x_{i-1}) \\ &= |c_1| [\overline{S}(f; P) - \underline{S}(f; P)] + |c_2| [\overline{S}(g; P) - \underline{S}(g; P)]. \end{aligned}$$

Dato che la funzione f è integrabile, grazie alla Proposizione 1.9, esiste una partizione P_1 tale che $\overline{S}(f; P_1) - \underline{S}(f; P_1) < \varepsilon$. Analogamente, esiste una partizione P_2 per cui $\overline{S}(g; P_2) - \underline{S}(g; P_2) < \varepsilon$. Scegliendo la partizione $P = P_1 \cup P_2$, entrambe le disequazioni sono soddisfatte e quindi

$$\overline{S}(c_1f + c_2g; P) - \underline{S}(c_1f + c_2g; P) < (|c_1| + |c_2|) \varepsilon.$$

Dalla Proposizione 1.9, segue che la funzione $c_1f + c_2g$ è integrabile in $[a, b]$.

Resta da dimostrare la formula (4). Data la solita partizione $P = \{x_0, \dots, x_n\}$, scegliamo $\xi_1, \xi_2, \dots, \xi_n$ con $\xi_i \in (x_{i-1}, x_i)$ per $i = 1, \dots, n$. Dato che

$$\underline{S}(f; P) \leq \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \leq \overline{S}(f; P), \quad \int_a^b f(x) dx \leq \overline{S}(f; P),$$

si ha

$$(8) \quad \left| \int_a^b f(x) dx - \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \right| \leq \overline{S}(f; P) - \underline{S}(f; P)$$

Aggiungendo e sottraendo i termini opportuni, si ricava

$$\begin{aligned} &\left| \int_a^b [c_1f + c_2g] dx - c_1 \int_a^b f dx - c_2 \int_a^b g dx \right| \\ &\leq \left| \int_a^b [c_1f + c_2g] dx - \sum_{i=1}^n [c_1f(\xi_i) + c_2g(\xi_i)] (x_i - x_{i-1}) \right| \\ &\quad + |c_1| \left| \int_a^b f dx - \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}) \right| + |c_2| \left| \int_a^b g dx - \sum_{i=1}^n g(\xi_i)(x_i - x_{i-1}) \right|. \end{aligned}$$

Utilizzando (8), si deduce la disuguaglianza

$$\begin{aligned} &\left| \int_a^b [c_1f + c_2g] dx - c_1 \int_a^b f dx - c_2 \int_a^b g dx \right| \\ &\leq \overline{S}(c_1f + c_2g; P) - \underline{S}(c_1f + c_2g; P) \\ &\quad + |c_1| [\overline{S}(f; P) - \underline{S}(f; P)] + |c_2| [\overline{S}(g; P) - \underline{S}(g; P)]. \end{aligned}$$

Scegliendo un partizione P per cui

$$\begin{aligned}\overline{S}(f; P) - \underline{S}(f; P) &< \varepsilon, & \overline{S}(g; P) - \underline{S}(g; P) &< \varepsilon, \\ \overline{S}(c_1 f + c_2 g; P) - \underline{S}(c_1 f + c_2 g; P) &< \varepsilon,\end{aligned}$$

si ottiene

$$\left| \int_a^b [c_1 f + c_2 g] dx - c_1 \int_a^b f dx - c_2 \int_a^b g dx \right| \leq (1 + |c_1| + |c_2|) \varepsilon.$$

Dato che ε è arbitrariamente piccolo, non c'è scampo: (4) è dimostrata.

La linearità garantisce, in particolare, che la somma di funzioni integrabili è integrabile. Lo stesso vale per il prodotto, come enunciato nell'esercizio che segue (ma non è vero che l'integrale del prodotto è il prodotto degli integrali!).

ESERCIZIO 2.2. Se f e g sono integrabili in $[a, b]$, anche il prodotto fg lo è.

Soluzione. Le funzioni f e g sono limitate, quindi esistono $M_f, M_g > 0$ tali che $|f(x)| \leq M_f$ e $|g(x)| \leq M_g$ per ogni x . Data la partizione $P = \{x_0, x_1, \dots, x_n\}$ indichiamo con

$$\alpha_i^{fg} = \inf_{[x_{i-1}, x_i]} f(x)g(x) \quad \text{e} \quad \beta_i^{fg} = \sup_{[x_{i-1}, x_i]} f(x)g(x),$$

con significati analoghi per $\alpha_i^f, \alpha_i^g, \beta_i^f$ e β_i^g . Dati $x, y \in (x_{i-1}, x_i)$, si ha

$$f(x)g(x) - f(y)g(y) = [f(x) - f(y)]g(x) + f(y)[g(x) - g(y)] \leq M_g(\beta_i^f - \alpha_i^f) + M_f(\beta_i^g - \alpha_i^g).$$

Prendendo l'estremo superiore in x e l'estremo superiore in y , si ottiene

$$\beta_i^{fg} - \alpha_i^{fg} \leq M_g(\beta_i^f - \alpha_i^f) + M_f(\beta_i^g - \alpha_i^g).$$

Perciò

$$\overline{S}(fg; P) - \underline{S}(fg; P) \leq M_g [\overline{S}(f; P) - \underline{S}(f; P)] + M_f [\overline{S}(g; P) - \underline{S}(g; P)].$$

Dato che il termine a secondo membro può essere reso arbitrariamente piccolo per P opportuno, la conclusione segue dalla solita Proposizione 1.9.

Additività. Se si pensa all'idea geometrica di partenza, la proprietà di additività sembra abbastanza naturale: per calcolare l'area possiamo dividere la regione in due parti e sommare i valori delle aree delle due sottoregioni. Ma qui la cosa è diversa: prima di tutto abbiamo una definizione analitica da rispettare e ogni affermazione deve discendere rigorosamente da quella definizione. In più c'è un particolare non banale: chi garantisce che se una funzione è integrabile in $[a, b]$, allora è anche integrabile in $[a, c]$ e $[c, b]$ per $c \in (a, b)$? Seppur ragionevole, quest'affermazione è tutta da verificare.

Per comodità, introduciamo la funzione $\chi_E : \mathbb{R} \rightarrow \mathbb{R}$

$$\text{funzione caratteristica di } E: \quad \chi_E(x) := \begin{cases} 1 & x \in E, \\ 0 & x \notin E. \end{cases}$$

Se l'insieme E è un intervallo, la funzione χ_E è integrabile. Quindi, se $f : [a, b] \rightarrow \mathbb{R}$ è integrabile in $[a, b]$ e $c \in [a, b]$, allora sono integrabili anche le funzioni prodotto $f\chi_{[a,c]}(x)$ e $f\chi_{[c,b]}(x)$ e vale l'uguaglianza

$$f(x) = f(x)\chi_{[a,c]}(x) + f(x)\chi_{[c,b]}(x),$$

Inoltre, dato che moltiplicare per una funzione caratteristica si traduce nel “troncare a zero” la funzione f fuori dall'insieme corrispondente, si hanno

$$\int_a^b f(x)\chi_{[a,c]}(x) dx = \int_a^c f(x) dx \quad \text{e} \quad \int_a^b f(x)\chi_{[c,b]}(x) dx = \int_c^b f(x) dx.$$

Quindi, grazie alla linearità,

$$\int_a^b f(x) dx = \int_a^b f(x)\chi_{[a,c]}(x) dx + \int_a^b f(x)\chi_{[c,b]}(x) dx = \int_a^c f(x) dx + \int_c^b f(x) dx.$$

Monotonia. Dimostrare la monotonia dell'integrale è particolarmente facile. Se $f(x) \leq g(x)$ per ogni $x \in [a, b]$, allora

$$\inf_{x \in [x_{i-1}, x_i]} f(x) \leq \inf_{x \in [x_{i-1}, x_i]} g(x) \quad \forall [x_{i-1}, x_i] \subset [a, b].$$

Quindi, per ogni partizione P di $[a, b]$, vale $\underline{S}(f; P) \leq \underline{S}(g; P)$, e passando all'estremo superiore si ottiene la conclusione.

DIMOSTRAZIONE DELLA PROPOSIZIONE 2.1. La parte più complicata della dimostrazione sta nel verificare che effettivamente la funzione $|f|$ sia integrabile.

Il problema è sempre lo stesso: data $P = \{x_0, \dots, x_n\}$ partizione di $[a, b]$ e posti $A_k = \inf_{[x_{k-1}, x_k]} |f(x)|$, $B_k = \sup_{[x_{k-1}, x_k]} |f(x)|$, quale stima possiamo recuperare per $B_k - A_k$?

Siano $\alpha_k = \inf_{[x_{k-1}, x_k]} f(x)$, $\beta_k = \sup_{[x_{k-1}, x_k]} f(x)$, dati x, y in (x_{k-1}, x_k) , si ha

$$||f(x)| - |f(y)|| \leq |f(x) - f(y)| \leq \beta_k - \alpha_k,$$

e prendendo l'estremo superiore in x e l'inferiore in y si ottiene $B_k - A_k \leq \beta_k - \alpha_k$.

Con questa stima alla mano, si deduce che

$$\overline{S}(|f|; P) - \underline{S}(|f|; P) \leq \sum_{k=1}^n (\beta_k - \alpha_k)(x_k - x_{k-1}) = \overline{S}(f; P) - \underline{S}(f; P).$$

e grazie alla Proposizione 1.9 si arriva alla conclusione.

La stima finale segue da $\pm f \leq |f|$, che, grazie alla monotonia dell'integrale, implica

$$\pm \int_a^b f(x) dx \leq \int_a^b |f(x)| dx,$$

da cui segue la conclusione. \square

ESERCIZIO 2.3. *Date f e g integrabili, dimostrare che anche $f_+(x) := \max\{f(x), 0\}$, $f_-(x) := -\min\{f(x), 0\}$, $\max\{f(x), g(x)\}$ e $\min\{f(x), g(x)\}$ sono integrabili.*

Una volta dato senso al concetto di integrale e determinate le proprietà principali, bisogna dedicarsi a determinare un certo numero di funzioni che siano effettivamente integrabili, altrimenti l'oggetto appena definito risulterebbe sostanzialmente inutile. Come si è detto, per dimostrare l'integrabilità di f , basta mostrare

$$\forall \varepsilon > 0 \quad \exists P_\varepsilon \text{ partizione tale che } \overline{S}(f; P_\varepsilon) - \underline{S}(f; P_\varepsilon) \leq \varepsilon.$$

Dunque seguiremo questa strategia: fissata una partizione P , mostreremo che la differenza $\overline{S}(f; P) - \underline{S}(f; P)$ può essere resa arbitrariamente piccola, a patto di scegliere una partizione P la cui ampiezza $|P|$ sia sufficientemente piccola.

La prima classe che consideriamo è quella delle funzioni monotone in un intervallo chiuso e limitato. Tali funzioni sono sempre limitate (perché?).

TEOREMA 2.4. *Sia $f : [a, b] \rightarrow \mathbb{R}$ monotona, allora è integrabile in $[a, b]$.*

DIMOSTRAZIONE. Supponiamo che f sia una funzione crescente. Data una partizione $P = \{x_0, \dots, x_n\}$, gli estremi inferiori e superiori di f in $[x_{i-1}, x_i]$ sono

$$\alpha_i = \inf_{[x_{i-1}, x_i]} f(x) = f(x_{i-1}) \quad \text{e} \quad \beta_i = \sup_{[x_{i-1}, x_i]} f(x) = f(x_i).$$

Le somme integrali per difetto e per eccesso sono date da

$$\underline{S}(f; P) = \sum_{i=1}^n f(x_{i-1})(x_i - x_{i-1}) \quad \text{e} \quad \overline{S}(f; P) = \sum_{i=1}^n f(x_i)(x_i - x_{i-1}).$$

e la loro differenza è stimata da

$$\overline{S}(f; P) - \underline{S}(f; P) = \sum_{i=1}^n [f(x_i) - f(x_{i-1})] (x_i - x_{i-1})$$

Indicando con $|P|$ l'ampiezza di P , cioè il massimo delle lunghezze $x_{i-1} - x_i$,

$$\overline{S}(f; P) - \underline{S}(f; P) \leq |P| \sum_{i=1}^n [f(x_i) - f(x_{i-1})] = |P| [f(b) - f(a)].$$

Per ogni $\varepsilon > 0$, è possibile scegliere $|P|$ sufficientemente piccola, in modo che la differenza $\overline{S}(f; P) - \underline{S}(f; P)$ sia minore di ε , pertanto, grazie (ancora una volta!) alla Proposizione 1.9, la funzione è integrabile. \square

A partire dalle funzioni monotone e grazie al fatto che combinazioni lineari di funzioni integrabili sono integrabili, è possibile costruire una classe ancora più ampia di funzioni. Ad esempio, sia f una funzione definita in $[a, b]$ tale che, per qualche $c \in [a, b]$, la funzione f è crescente in $[a, c]$ e decrescente in $[c, b]$. Allora la funzione f si può riscrivere come differenza di due funzioni crescenti (vedi Figura 6):

$$f = f_1 - f_2$$

dove

$$f_1(x) = \begin{cases} f(x) & x \in [a, c] \\ f(c) & x \in [c, b] \end{cases} \quad \text{e} \quad f_2(x) = \begin{cases} 0 & x \in [a, c] \\ f(c) - f(x) & x \in [c, b] \end{cases}$$

Quindi, anche funzioni con un cambio di monotonia sono integrabili. Con una co-

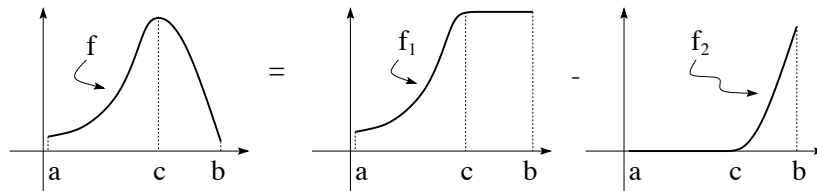


FIGURA 6. Una funzione f con un cambio di monotonia, decomposta come differenza delle funzioni crescenti f_1 e f_2 .

struzione analoga, si mostra che *tutte le funzioni con un numero finito di cambi di monotonia sono integrabili*. Ad esempio, tutti i polinomi e, più in generale, tutte le funzioni razionali sono integrabili.

ESERCIZIO 2.5. *Conoscete una funzione definita in un intervallo limitato che abbia un numero infinito di cambi di monotonia?*

La seconda classe che consideriamo è quella delle *funzioni lipschitziane*, cioè delle funzioni $f : [a, b] \rightarrow \mathbb{R}$ tali che esista $L > 0$ per cui

$$\exists L > 0 \quad \text{t.c.} \quad |f(x) - f(y)| \leq L|x - y| \quad \forall x, y \in [a, b].$$

TEOREMA 2.6. *Sia $f : [a, b] \rightarrow \mathbb{R}$ lipschitziana, allora è integrabile in $[a, b]$.*

DIMOSTRAZIONE. Fissata la beneamata partizione $P = \{x_0, x_1, \dots, x_n\}$,

$$\bar{S}(f; P) - \underline{S}(f; P) = \sum_{i=1}^n (\beta_i - \alpha_i)(x_i - x_{i-1}),$$

con il solito significato per α_i e β_i . Dato che f è lipschitziana, essa è anche continua in $[a, b]$, quindi ammette massimo e minimo in ogni intervallo $[x_{i-1}, x_i]$ e vale

$$\alpha_i = \inf_{(x_{i-1}, x_i)} f = \min_{[x_{i-1}, x_i]} f(x) = f(\eta_i) \quad \text{e} \quad \beta_i = \sup_{(x_{i-1}, x_i)} f = \max_{[x_{i-1}, x_i]} f(x) = f(\xi_i)$$

con η_i e ξ_i , rispettivamente, un punto di massimo ed uno di minimo della funzione f in $[x_{i-1}, x_i]$. Sostituendo nella relazione precedente, si ottiene

$$\overline{S}(f; P) - \underline{S}(f; P) = \sum_{i=1}^n (f(\xi_i) - f(\eta_i))(x_i - x_{i-1}) \leq L \sum_{i=1}^n |\xi_i - \eta_i|(x_i - x_{i-1}).$$

Dato che $\xi_i, \eta_i \in [x_{i-1}, x_i]$, la differenza $|\xi_i - \eta_i|$ è minore o uguale dell'ampiezza $|P|$ della partizione. Quindi

$$\overline{S}(f; P) - \underline{S}(f; P) \leq L|P| \sum_{i=1}^n (x_i - x_{i-1}) \leq L(b-a)|P|.$$

Scegliendo P tale che $|P| < \varepsilon/L(b-a)$, la differenza $\overline{S}(f; P) - \underline{S}(f; P)$ è strettamente minore di ε e, grazie alla Proposizione 1.9, segue la conclusione. \square

Più in generale si può dimostrare il fondamentale

TEOREMA 2.7. *Sia $f : [a, b] \rightarrow \mathbb{R}$ continua, allora f è integrabile in $[a, b]$.*

La dimostrazione passa per il concetto di *uniforme continuità*. Non daremo qui ulteriori dettagli sulla questione.

ESERCIZIO 2.8. *Dimostrare che la composizione di una funzione lipschitziana con una funzione integrabile dà luogo ad una funzione integrabile.*

3. Il Teorema della media integrale

Tutte le funzioni $f : [a, b] \rightarrow \mathbb{R}$ integrabili sono, per definizione, funzioni limitate. In particolare questo significa che data una funzione integrabile, il suo integrale definito può sempre essere stimato in maniera “rude”: per le proprietà di monotonia dell'integrale, se $m \leq f(x) \leq M$ per ogni $x \in [a, b]$,

$$m(b-a) = \int_a^b m \, dx \leq \int_a^b f(x) \, dx \leq \int_a^b M \, dx = M(b-a),$$

Questa formula è intuitivamente ovvia: se pensiamo ad una funzione non negativa e all'integrale come area, le quantità $M(b-a)$ e $m(b-a)$ rappresentano le aree di una rettangolo circoscritto ed inscritto nel sottografico di f (vedi Figura 7).

La formula precedente si può riscrivere come $m \leq \langle f \rangle \leq M$ dove $\langle f \rangle$, detto **media integrale**² di f in $[a, b]$, è definito da

$$(9) \quad \langle f \rangle := \frac{1}{b-a} \int_a^b f(x) \, dx.$$

²La *media aritmetica* dei numeri f_1, f_2, \dots, f_n è, per definizione $(f_1 + f_2 + \dots + f_n)/n$. La media integrale di una funzione in un intervallo è, in un certo senso, limite di medie aritmetiche di valori assunti dalla funzione. Infatti, data la partizione di $[a, b]$ in sottointervalli di lunghezza $\Delta x_i = (b-a)/n$

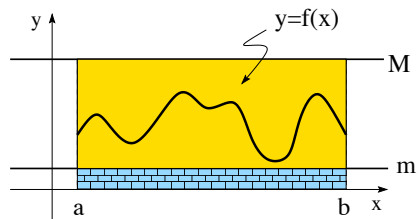


FIGURA 7. Significato geometrico della stima $m(b-a) \leq \int_a^b f(x) dx \leq M(b-a)$.

Quindi, espresso a parole, *la media integrale di una funzione f è sempre compresa tra un qualsiasi minorante ed un qualsiasi maggiorante di f .*

Se f è continua in un intervallo $[a, b]$, si può dire qualcosa di più. Per il teorema di Weierstrass, esistono due punti $\xi, \eta \in [a, b]$ tali che

$$\min_{[a,b]} f(x) = f(\xi) \leq f(x) \leq f(\eta) = \max_{[a,b]} f(x) \quad \forall x \in [a, b].$$

Quindi $\langle f \rangle$, definita in (9) è compresa tra il massimo ed il minimo della funzione f in $[a, b]$. Grazie al Teorema del valore intermedio, si può concludere che *per una funzione f continua in $[a, b]$, la media integrale fa sempre parte dell'insieme immagine $f([a, b])$.*

TEOREMA 3.1. Teorema della Media Integrale. *Sia $f : [a, b] \rightarrow \mathbb{R}$ continua. Allora esiste $\xi \in [a, b]$ tale che $f(\xi) = \langle f \rangle$ con $\langle f \rangle$ definito in (9).*

Nel caso di una funzione non negativa f , il teorema equivale ad affermare che esiste un rettangolo di base $[a, b]$ ed una altezza $f(\xi)$ opportuna con la stessa area del sottografico di f in $[a, b]$.

Non è particolarmente sconvolgente (ma più avanti servirà) osservare che la formula (9) vale anche nel caso $b < a$, infatti

$$\frac{1}{b-a} \int_a^b f(x) dx = \frac{1}{a-b} \int_b^a f(x) dx,$$

e a quest'ultimo termine si può applicare il Teorema della Media Integrale.

per ogni i e scelti $\xi_i \in (x_{i-1}, x_i)$, la media aritmetica di $f(\xi_1), \dots, f(\xi_n)$ è pari a

$$\frac{f(\xi_1) + \dots + f(\xi_n)}{n} = \frac{1}{b-a} \sum_{i=1}^n f(\xi_i)(x_i - x_{i-1}).$$

Per $n \rightarrow \infty$, il termine a destra converge proprio alla media integrale di f in $[a, b]$.

CONTROESEMPIO 3.2. Se la funzione f non è continua in tutto $[a, b]$, non è detto che valga la conclusione del Teorema 3.1. Ad esempio, si consideri la funzione

$$f(x) = \operatorname{sgn} x = \begin{cases} -1 & x < 0, \\ 0 & x = 0, \\ +1 & x > 0, \end{cases}$$

nell'intervallo $[-1, 2]$. Allora f è integrabile (per almeno due motivi... quali?)

$$\mu = \frac{1}{2 - (-1)} \int_{-1}^2 \operatorname{sgn} x \, dx = \frac{1}{3} (-1 + 2) = \frac{1}{3},$$

che non fa parte dell'immagine della funzione $\operatorname{sgn} x$.

ESERCIZIO 3.3. Siano $f \in C([a, b])$ e $p \in C([a, b])$ tale che $p(x) > 0$ per ogni x . Dimostrare che esiste $\xi \in [a, b]$ tale che

$$f(\xi) = \frac{\int_a^b f(x)p(x) \, dx}{\int_a^b p(x) \, dx}.$$

Soluzione. Se m e M indicano il minimo ed il massimo di f in $[a, b]$, grazie al fatto che $p(x) > 0$, si ha $mp(x) \leq f(x)p(x) \leq Mp(x)$ per ogni x . Integrate in $[a, b]$, mescolate e concludete...

4. Il Teorema fondamentale del calcolo integrale

Una volta fissata la funzione f , l'integrale definito è una funzione degli estremi di integrazione a e b . Per studiare questa dipendenza, supponiamo l'estremo inferiore fissato al valore a e indichiamo l'estremo superiore (variabile) con x : consideriamo, quindi, la funzione integrale

$$(10) \quad \phi(x) = \int_a^x f(t) \, dt.$$

La formula (10) può essere utilizzata per “generare nuove funzioni” a partire da una funzione integrabile f . Ad esempio, si può definire³

$$\phi(x) := \int_1^x \frac{1}{t} \, dt \quad x > 0,$$

che ha perfettamente senso, dato che la funzione $\frac{1}{t}$ è una funzione continua in ogni intervallo $[a, b]$ con $0 < a \leq b$.

PROPOSIZIONE 4.1. Se f è integrabile in $[a, b]$, la funzione ϕ , definita in (10), è lipschitziana.

³Tale funzione, come vedremo, si dimostra essere uguale al logaritmo naturale e la formula precedente può quindi essere scelta come definizione analitica di $\ln x$.

DIMOSTRAZIONE. Dato che f è integrabile, essa è anche limitata. Sia $M > 0$ tale che $|f(t)| \leq M$ per ogni $t \in [a, b]$, allora,

$$\text{se } x < y, \quad |\phi(x) - \phi(y)| = \left| \int_x^y f(t) dt \right| \leq \int_x^y |f(t)| dt \leq M(y - x);$$

$$\text{se } x > y, \quad |\phi(x) - \phi(y)| = \left| - \int_y^x f(t) dt \right| \leq \int_y^x |f(t)| dt \leq M(x - y).$$

Quindi

$$|f(t)| \leq M \quad \implies \quad |\phi(x) - \phi(y)| \leq M|x - y|.$$

In particolare, se f è integrabile, la funzione ϕ è una funzione continua. \square

ESERCIZIO 4.2. Consideriamo la funzione $f(x) = \operatorname{sgn} x$. Sapete riconoscere la funzione ϕ data in (10) in questo caso?

ESERCIZIO 4.3. Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ integrabile in ogni intervallo $[a, b]$ di \mathbb{R} e sia

$$\phi(x) := \int_0^x f(t) dt$$

(i) Dimostrare che, se $f(x) \geq 0$ per ogni $x \in \mathbb{R}$, allora ϕ è non decrescente.

(ii) Disegnare qualitativamente il grafico di ϕ nel caso in cui $xf(x) > 0$ per ogni $x \neq 0$.

Una volta stabilito che la funzione integrale ϕ è sempre lipschitziana, è naturale domandarsi se essa sia anche derivabile. Consideriamone il rapporto incrementale

$$\frac{\phi(x+h) - \phi(x)}{h} = \frac{1}{h} \left[\int_\alpha^{x+h} f(t) dt - \int_\alpha^x f(t) dt \right] = \frac{1}{h} \int_x^{x+h} f(t) dt$$

(l'ultima uguaglianza è conseguenza delle proprietà di additività dell'integrale). Supponendo che la funzione integranda f sia continua, è possibile applicare il Teorema della Media Integrale, Teorema 3.1, e riscrivere il rapporto incrementale come

$$\frac{\phi(x+h) - \phi(x)}{h} = f(\xi)$$

con ξ compreso tra x e $x+h$. Passando al limite $h \rightarrow 0$, dato che $\xi \rightarrow x$, si ha

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

Abbiamo quindi dimostrato il seguente risultato.

TEOREMA 4.4. Sia $f \in C([a, b])$, $\alpha \in [a, b]$ e $\phi(x) = \int_\alpha^x f(t) dt$ per $x \in [a, b]$. Allora ϕ è derivabile in $[a, b]$ e $\phi'(x) = f(x)$.

Una prima conseguenza (pratica) notevole è che, dato che siamo in grado di calcolarne la derivata, possiamo dedurre molte proprietà qualitative importanti anche per una funzione che non sia espressa direttamente tramite funzioni elementari, ma come integrale di una funzione elementare.

ESEMPIO 4.5. Consideriamo la funzione

$$\text{funzione degli errori: } \operatorname{Erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2} dt,$$

Dal Teorema 4.4 deduciamo che $D(\operatorname{Erf}(x)) = \frac{2}{\sqrt{\pi}} e^{-x^2} > 0$, quindi $\operatorname{Erf}(x)$ è strettamente crescente. Con una tecnologia più avanzata di quella che ci è disponibile a questo livello, è possibile dimostrare che $\operatorname{Erf}(\pm\infty) = \pm 1$.

Il Teorema 4.4 risolve anche un problema interessante:

data f , trovare una funzione F che risolva l'equazione $F' = f$.

L'equazione $F' = f$ è un'equazione differenziale in cui il dato è la funzione f e l'incognita è la funzione F . Una soluzione F di questa equazione si dice **primitiva** di f . Il Teorema 4.4 afferma che se $f \in C([a, b])$ il problema $F' = f$ ammette almeno una soluzione (data dalla funzione integrale ϕ), cioè esiste sempre almeno una primitiva. Da questo punto di vista *si può interpretare l'operazione di integrazione come l'operazione inversa della derivazione*. Tale operazione di inversione è univocamente definita? In altri termini, data una funzione f , quante primitive esistono?

TEOREMA 4.6. *Sia $f : [a, b] \rightarrow \mathbb{R}$ e siano F e G due sue primitive. Allora esiste $c \in \mathbb{R}$ tale che*

$$F(x) - G(x) = c \quad \text{per ogni } x \in [a, b].$$

DIMOSTRAZIONE. La derivata della funzione differenza $F - G$ è nulla:

$$(F(x) - G(x))' = F'(x) - G'(x) = f(x) - f(x) = 0.$$

Per quanto già visto, la differenza $F - G$ deve essere costante. □

Quindi, se $f \in C([a, b])$, l'equazione $F' = f$ è completamente risolta: tutte le soluzioni sono della forma

$$\int_{\alpha}^x f(t) dt + c \quad c \in \mathbb{R}.$$

La classe delle primitive della funzione f si indica con

$$\int f(x) dx,$$

e si chiama **integrale indefinito** di f . Si noti bene che l'integrale indefinito di una funzione indica una classe di funzioni, e non una singola funzione.

Ad eterna memoria, sintetizziamo i due risultati enunciati in un'unico Teorema.

TEOREMA 4.7. Teorema Fondamentale del Calcolo Integrale. *Data $f \in C([a, b])$, le soluzioni dell'equazione differenziale $F' = f$ sono tutte e sole della forma*

$$F(x) = \int_{\alpha}^x f(t) dt + c, \quad \text{con } \alpha \in [a, b] \quad \text{e } c \in \mathbb{R}.$$

Se si cerca una primitiva F di una funzione f con la richiesta aggiuntiva che la funzione F valga in un punto assegnato x_0 un valore dato y_0 , cioè se si vuole risolvere

$$(11) \quad \text{dati } f \in C([a, b]), x_0 \in [a, b], y_0 \in \mathbb{R}, \text{ trovare } F \text{ tale che } \begin{cases} F'(x) = f(x), \\ F(x_0) = y_0, \end{cases}$$

la soluzione F esiste, è unica ed è data da

$$(12) \quad F(x) = y_0 + \int_{x_0}^x f(t) dt.$$

Il problema (11) rientra nella classe dei *problemi di Cauchy* per equazioni differenziali.

OSSERVAZIONE 4.8. Il problema di Cauchy può essere interpretato in termini di moto di un particella. Se $f(x)$ è la velocità della particella all'istante x e y_0 la sua posizione all'istante iniziale x_0 , la soluzione $F = F(x)$, definita in (12), del problema di Cauchy (11), rappresenta la posizione della particella all'istante x .

Primitive e calcolo degli integrali definiti. Il Teorema fondamentale del calcolo ha una conseguenza interessante che riguarda il calcolo esplicito di integrali definiti. Supponiamo di voler calcolare

$$(13) \quad \int_a^b f(t) dt.$$

e supponiamo di conoscere già (per altre vie) una primitiva della funzione f , cioè una funzione F tale che $F' = f$. Sappiamo che anche la funzione integrale definita in (10) è una primitiva di f e, quindi, per il Teorema Fondamentale del Calcolo, differisce da F per una costante, cioè $\phi(x) = F(x) + c$ per qualche $c \in \mathbb{R}$. La costante c può essere determinata, calcolando in $x = a$:

$$0 = \phi(a) = F(a) + c \quad \implies \quad c = -F(a).$$

Si deduce quindi che $\phi(x) = F(x) - F(a)$ e quindi

$$\int_a^b f(t) dt = \phi(b) = F(b) - F(a).$$

Quindi, se si conosce una primitiva F della funzione f , l'integrale definito di f in $[a, b]$ è uguale alla differenza dei valori della primitiva in b e in a , cioè

$$F' = f \quad \Longrightarrow \quad \int_a^b f(t) dt = F(b) - F(a).$$

La differenza $F(b) - F(a)$ si indica anche con $F(x)\Big|_a^b$, o $[F(x)]_a^b$.

ESEMPIO 4.9. Consideriamo la funzione $f(x) = x^2$. Dato che $D(x^3) = 3x^2$, si ha

$$D\left(\frac{1}{3}x^3\right) = x^2,$$

quindi una primitiva di x^2 è $x^3/3$. Otteniamo perciò

$$\int_a^b x^2 dx = \frac{1}{3}x^3\Big|_a^b = \frac{b^3}{3} - \frac{a^3}{3},$$

che è proprio la formula calcolata ad inizio Capitolo. Più in generale, dato che

$$D\left(\frac{1}{n+1}x^{n+1}\right) = x^n,$$

vale la formula

$$\int_a^b x^n dx = \frac{1}{n+1}x^{n+1}\Big|_a^b = \frac{1}{n+1}(b^{n+1} - a^{n+1}).$$

CAPITOLO 2

Zoologia dell'integrazione

Per iniziare, ricordiamo i fatti principali che abbiamo visto sugli integrali indefiniti.

- Data una funzione f , si dice che F è una **primitiva** di f se $F' = f$; la famiglia di primitive di una funzione f si indica con il simbolo

$$\int f(x) dx.$$

Se la funzione f è considerata in un intervallo, tutte le primitive di f sono uguali a meno di una costante additiva, cioè, data F primitiva di f ,

$$\int f(x) dx = F(x) + C \quad C \in \mathbb{R}.$$

- Se la funzione F è una primitiva della funzione f , allora vale

$$F' = f \quad \implies \quad \int_a^b f(x) dx = F(b) - F(a).$$

In questo Capitolo ci poniamo il problema di determinare *esplicitamente* primitive F di una funzione data f , per opportune classi di funzioni. Cosa si intende qui con “primitive esplicite”? A partire dalle operazioni elementari (addizione, sottrazione, moltiplicazione, divisione) e dalle funzioni trigonometriche ed esponenziali, formando inverse e composte di queste funzioni, è possibile costruire una classe molto ampia di funzioni che possiamo descrivere come “funzioni esplicite”.

Per quanto riguarda l'operazione di derivazione, la derivata di una funzione esplicita è essa stessa una funzione esplicita. Al contrario, per l'integrazione la situazione è differente: *non è vero che tutti gli integrali delle funzioni esplicite si possano scrivere in termini di funzioni esplicite* (ad esempio, non è possibile esprimere in forma “esplicita”, le primitive di e^{-x^2}). Questo risultato può suonare sorprendente, ma è un fatto della vita. Prendere o lasciare.

Una prima classe di funzioni che sono esplicitamente integrabili si determina a partire dal *teorema fondamentale del calcolo*, che, come noto, afferma

$$F'(x) = f(x) \quad \implies \quad \int f(x) dx = F(x) + \text{costante}.$$

In particolare, questa proprietà indica che *ad ogni regola di derivazione corrisponde una regola di integrazione*. Ad esempio

$$D(x^{\alpha+1}) = (\alpha + 1)x^\alpha \quad \Rightarrow \quad \int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha + 1} + \text{costante} \quad \forall \alpha \neq -1.$$

Allo stesso modo si ottengono altre formule elementari:

$$\begin{aligned} \int \sin x dx &= -\cos x + C, & \int \cos x dx &= \sin x + C, \\ \int e^x dx &= e^x + C, & \int \frac{dx}{x} &= \ln |x| + C, \\ \int \frac{dx}{1+x^2} &= \arctan x + C, & \int a^x dx &= \frac{a^x}{\ln a} + C, \\ \int \frac{dx}{\cos^2 x} &= \tan x + C, & \int \frac{dx}{\sin^2 x} &= -\cot x + C, \\ \int \frac{dx}{\sqrt{1-x^2}} &= \arcsin x + C, & \int \frac{dx}{\sqrt{1-x^2}} &= -\arccos x + C, \\ \int \sinh x dx &= \cosh x + C & \int \cosh x dx &= \sinh x + C. \end{aligned}$$

Inoltre, grazie alla linearità dell'integrale, anche combinazioni lineari di funzioni di cui si conosce la primitiva, possono essere integrate esplicitamente. Ad esempio,

$$\int (1 + 2x + 3e^x) dx = \int 1 dx + 2 \int x dx + 3 \int e^x dx = x + x^2 + 3e^x + C.$$

Tramite queste formule è possibile calcolare il valore di certi integrali definiti, senza bisogno di passare per le approssimazioni con somme integrali per eccesso e per difetto.

Passiamo ora a sviluppare i due metodi principali da affiancare alle formule di integrazione elementari: l'integrazione per sostituzione e l'integrazione per parti. Entrambi, in sostanza, discendono da formule di derivazione: il primo discende dalla derivazione di funzione composta, il secondo dalla derivazione della funzione prodotto.

1. Metodo di sostituzione

Il *metodo di sostituzione* consiste nell'introduzione di una nuova variabile, cioè, moralmente, nel "cambiare punto di vista" e osservare lo stesso oggetto da un'altra posizione. La formula di derivazione di funzione composta assicura che

$$(14) \quad (F(\phi(u)))' = F'(\phi(u))\phi'(u).$$

Tale formula, letta in termini di integrazione, diviene

$$(15) \quad \int F'(\phi(x))\phi'(x) dx = F(\phi(x)) + \text{costante}$$

Ad esempio,

$$\int 2x \cos(x^2) dx = \sin(x^2) + \text{costante},$$

dove $F(s) = \sin s$ e $\phi(x) = x^2$. Vediamo un altro esempio. Calcoliamo

$$\int \frac{\ln x}{x} dx.$$

Dato che $D(\ln x) = 1/x$, qui $F(s) = s$ e $\phi(x) = \ln x$. Quindi

$$\int \frac{\ln x}{x} dx = \int \ln x (\ln x)' dx = \frac{1}{2} (\ln x)^2 + \text{costante}.$$

Ma *come individuare una decomposizione della funzione integranda come in (15)?* Occorre esercizio ed esperienza (anche una certa dose di intuizione non guasta!).

Vediamo la formula di sostituzione per integrali definiti: integriamo la formula (14) nell'intervallo $[\alpha, \beta]$,

$$F(\phi(\beta)) - F(\phi(\alpha)) = \int_{\alpha}^{\beta} (F(\phi(u)))' du = \int_{\alpha}^{\beta} F'(\phi(u))\phi'(u) du.$$

Ponendo $a = \phi(\alpha)$ e $b = \phi(\beta)$,

$$F(\phi(\beta)) - F(\phi(\alpha)) = F(b) - F(a) = \int_a^b F'(x) dx.$$

Quindi otteniamo la formula

$$\int_{\alpha}^{\beta} F'(\phi(u))\phi'(u) du = \int_a^b F'(x) dx,$$

che, chiamando $f = F'$, può essere riscritta come

$$(16) \quad \int_{\alpha}^{\beta} f(\phi(u))\phi'(u) du = \int_a^b f(x) dx \quad \text{con} \quad \begin{cases} a = \phi(\alpha) \\ b = \phi(\beta) \end{cases}$$

Questa formula esprime come si trasforma l'espressione dell'integrale cambiando la variabile di integrazione. Se, per calcolare $\int_a^b f(x) dx$, decidiamo di porre $x = \phi(u)$, dobbiamo sostituire formalmente dx con $\phi'(u) du$ e cambiare gli estremi compatibilmente con la formula che collega x con u , cioè $x = \phi(u)$. Ad esempio, per calcolare

$$\int_0^1 \frac{e^x}{1 + e^{2x}} dx,$$

poniamo $x = \ln u$. Dato che $\phi(u) = \ln u$, dobbiamo sostituire dx con $\frac{1}{u} du$:

$$\int_0^1 \frac{e^x}{1 + e^{2x}} dx = \int_{\alpha}^{\beta} \frac{u}{1 + u^2} \frac{1}{u} du.$$

Rimangono da calcolare α e β che sono soluzione di $0 = \phi(\alpha) = \ln \alpha$ e $1 = \phi(\beta) = \ln \beta$. Invertendo la funzione ϕ otteniamo

$$\alpha = \phi^{-1}(0) = e^0 = 1, \quad \beta = \phi^{-1}(1) = e^1 = e.$$

In definitiva

$$\int_0^1 \frac{e^x}{1+e^{2x}} dx = \int_1^e \frac{u}{1+u^2} \frac{1}{u} du = \int_1^e \frac{du}{1+u^2} = \arctan e - \frac{\pi}{4}.$$

Nel caso degli integrali indefiniti la formula di sostituzione prende la forma (qui F è una primitiva di f , $F' = f$)

$$(17) \quad \int f(\phi(u))\phi'(u) du = \int f(x) dx = F(x) + C = F(\phi(u)) + C$$

Anche in questo caso si può usare, come regola mnemonica, la relazione

$$x = \phi(u) \quad \Rightarrow \quad " \phi'(u) du = dx ".$$

L'uso delle virgolette “ ” sta a ricordare che non è stato dato senso ai simboli du e dx e che la regola suscritta è solo formale¹. Nell'uso di questa formula bisogna ricordarsi di tornare alla fine alla variabile u , sostituendo x con $\phi(u)$.

ESEMPIO 1.1. Sia ϕ una funzione derivabile. Calcoliamo

$$\int \frac{\phi'(u)}{\phi(u)} du.$$

Ponendo $x = \phi(u)$, si ha $dx = \phi'(u) du$, quindi

$$\int \frac{\phi'(u)}{\phi(u)} du = \int \frac{dx}{x} = \ln |x| + C = \ln |\phi(u)| + C.$$

Ad esempio,

$$\int \frac{dx}{x \ln x} = \ln |\ln x| + C, \quad \int \tan x dx = - \int \frac{-\sin x}{\cos x} dx = - \ln |\cos x| + C.$$

ESEMPIO 1.2. Allo stesso modo, ponendo $x = \phi(u)$ (e quindi $dx = \phi'(u) du$),

$$\int [\phi(u)]^\alpha \phi'(u) du = \int x^\alpha dx = \frac{x^{\alpha+1}}{\alpha+1} + C = \frac{1}{\alpha+1} [\phi(u)]^{\alpha+1} + C \quad \alpha \neq -1.$$

Ad esempio,

$$\int \sin^k x \cos x dx = \frac{1}{k+1} \sin^{k+1} x + C.$$

¹Se si pensa all'origine del simbolo dx negli integrali come limite della lunghezza Δx di una data partizione, la sostituzione da dx a du , con il relativo termine moltiplicativo $\phi'(u)$, indica che, nel cambio di variabile, bisogna cambiare opportunamente anche la lunghezza dell'intervallo della partizione, coerentemente con la trasformazione utilizzata.

La formula di sostituzione è sempre conveniente nel caso di funzioni composte di cui l'ultima sia lineare: ponendo $x = au + b$

$$\int f(au + b) du = \frac{1}{a} \int f(x) dx.$$

Anche se l'integrale di destra non fosse risolvibile, l'espressione a secondo membro è comunque più semplice.

Spesso ci si trova a lavorare con espressioni della forma

$$\int h(\phi(u)) du,$$

dove l'integrando è una funzione composta $h(\phi(u))$, senza il fattore moltiplicativo $\phi'(u)$. E' possibile applicare la sostituzione $x = \phi(u)$? Se la funzione ϕ è invertibile, con inversa $u = \psi(x)$, è possibile sostituire a du l'oggetto $\psi'(x) dx$, ottenendo

$$\int h(\phi(u)) du = \int h(x) \psi'(x) dx.$$

Per giustificare in modo rigoroso questa formula, chiamiamo $f(u) := h(\phi(u))$

$$\int h(\phi(u)) du = \int f(u) du = \int f(\psi(x)) \psi'(x) dx = \int h(x) \psi'(x) dx.$$

Nel caso di integrali definiti occorre cambiare gli estremi di integrazione coerentemente con la nuova variabile introdotta:

$$\int_a^b h(\phi(u)) du = \int_{\phi(a)}^{\phi(b)} h(x) \psi'(x) dx.$$

ESEMPIO 1.3. Calcoliamo l'integrale indefinito

$$\int (1 + e^x)^2 dx.$$

Ponendo $t = 1 + e^x$, si ha $x = \ln(t - 1)$ e $dx = \frac{1}{t-1} dt$ e, di conseguenza,

$$\int (1 + e^x)^2 dx = \int \frac{t^2}{t-1} dt.$$

Dato che $\frac{t^2}{t-1} = t + 1 + \frac{1}{t-1}$,

$$\begin{aligned} \int (1 + e^x)^2 dx &= \int \left(t + 1 + \frac{1}{t-1} \right) dt = \frac{1}{2}t^2 + t + \ln|t-1| + C \\ &= \frac{1}{2}(1 + e^x)^2 + 1 + e^x + x + C = \frac{1}{2}e^{2x} + 2e^x + x + C. \end{aligned}$$

Si sarebbe anche potuto procedere utilizzando la decomposizione $(1 + e^x)^2 = 1 + 2e^x + e^{2x}$ e integrando a partire dalle formule elementari.

ESEMPIO 1.4. Calcoliamo l'integrale definito

$$\int_0^{\pi/2} \cos^2 x \, dx.$$

Dato che $\cos(2x) = 2 \cos^2 x - 1$

$$\int_0^{\pi/2} \cos^2 x \, dx = \frac{1}{2} \int_0^{\pi/2} (\cos(2x) + 1) \, dx = \frac{1}{2} \int_0^{\pi/2} \cos(2x) \, dx + \frac{\pi}{4}.$$

Poniamo nell'integrale $t = 2x$:

$$\int_0^{\pi/2} \cos(2x) \, dx = \frac{1}{2} \int_0^{\pi} \cos t \, dt = \frac{1}{2} \sin t \Big|_0^{\pi} = 0.$$

Quindi il valore dell'integrale è $\pi/4$.

ESERCIZIO 1.5. Fissato $a > 0$, calcolare gli integrali (indefinito e definito)

$$\int x \sqrt{a^2 - x^2} \, dx \quad \text{e} \quad \int_0^a x \sqrt{a^2 - x^2} \, dx.$$

ESEMPIO 1.6. Dati $A, B, C \in \mathbb{R}$ tali che $\Delta := B^2 - 4AC < 0$, come risolvere

$$(18) \quad \int \frac{dx}{Ax^2 + Bx + C} \quad ?$$

La condizione $\Delta < 0$, implica che il polinomio è irriducibile (non ha radici reali). In questa classe di integrali rientra un integrale che conosciamo:

$$(19) \quad \int \frac{dx}{x^2 + 1} = \arctan x + C.$$

L'integrale (18) può essere risolto con una sostituzione opportuna che lo riconduce a (19). Invece di dare direttamente la soluzione, proviamo a ricostruire passo passo come possa essere ottenuta.

Il grafico della funzione $\frac{1}{Ax^2+Bx+C}$ è qualitativamente simile a quello della funzione $\frac{1}{x^2+1}$ (Fig.1). È possibile con traslazioni e dilatazioni trasformare il grafico della prima

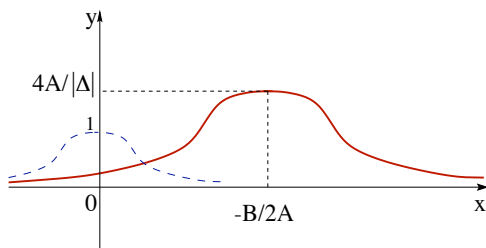


FIGURA 1. Il grafico della funzione $\frac{1}{Ax^2+Bx+C}$ (e, tratteggiato, quello di $\frac{1}{x^2+1}$).

funzione in quello della seconda? Bisogna prima di tutto correggere due “difetti”

evidenti di $f(x) = \frac{1}{Ax^2+Bx+C}$: il valore massimo è $f(-B/2A) = 4A/|\Delta|$ e non 1, l'asse di simmetria è $x = -B/2A$ e non $x = 0$. Per il primo problema, basta utilizzare la linearità dell'integrale: la funzione f può essere riscritta come

$$f(x) = \frac{4A}{|\Delta|} g(x) \quad \text{dove} \quad g(x) := \frac{1}{\frac{4A^2}{|\Delta|} x^2 + \frac{4AB}{|\Delta|} x + \frac{4AC}{|\Delta|}},$$

quindi

$$\int \frac{dx}{Ax^2 + Bx + C} = \frac{4A}{|\Delta|} \int \frac{dx}{\frac{4A^2}{|\Delta|} x^2 + \frac{4AB}{|\Delta|} x + \frac{4AC}{|\Delta|}},$$

dove la funzione all'interno dell'ultimo integrale vale 1 in $x = -B/2A$.

Per fare in modo che l'asse di simmetria sia in $x = 0$, bisogna traslare il grafico. Questo corrisponde ad introdurre una nuova variabile u , legata ad x dalla relazione

$$u = x + \frac{B}{2A}.$$

Con questa scelta, si ha $\frac{4A^2}{|\Delta|} x^2 + \frac{4AB}{|\Delta|} x + \frac{4AC}{|\Delta|} = \frac{4A^2}{|\Delta|} u^2 + 1$, quindi

$$\int \frac{dx}{Ax^2 + Bx + C} = \frac{4A}{|\Delta|} \int \frac{du}{\frac{4A^2}{|\Delta|} u^2 + 1} = \frac{4A}{|\Delta|} \int \frac{du}{\left(\frac{2A}{|\Delta|^{1/2}} u\right)^2 + 1}.$$

Resta da modificare ancora una volta la variabile u , attraverso la posizione

$$v := \frac{2A}{|\Delta|^{1/2}} u,$$

che consiste nel dilatare/comprimere (dipende dalla grandezza di $\frac{2A}{|\Delta|^{1/2}}$) la variabile u per un fattore opportuno. Così facendo si arriva alla conclusione:

$$\int \frac{dx}{Ax^2 + Bx + C} = \frac{4A}{|\Delta|} \int \frac{du}{\left(\frac{2A}{|\Delta|^{1/2}} u\right)^2 + 1} = \frac{2}{|\Delta|^{1/2}} \int \frac{dv}{v^2 + 1} = \frac{2}{|\Delta|^{1/2}} \arctan v + \text{costante}.$$

Per avere l'espressione della primitiva in x (e non in v), basta seguire a ritroso le definizioni di v e di u . In definitiva: (qui $|\Delta| = 4AC - B^2$)

$$\int \frac{dx}{Ax^2 + Bx + C} = \frac{2}{\sqrt{|\Delta|}} \arctan \left(\frac{2Ax + B}{\sqrt{|\Delta|}} \right) + \text{costante}.$$

2. Integrazione per parti

Il metodo di *integrazione per parti* emerge dalla formula di derivazione del prodotto: $(fg)' = f'g + fg'$. Integrando tale formula, si ottiene

$$f(x)g(x) = \int g(x) f'(x) dx + \int g'(x) f(x) dx$$

da cui la formula di integrazione per parti

$$(20) \quad \int g'(x) f(x) dx = f(x)g(x) - \int g(x) f'(x) dx.$$

Questa formula è nota come **integrazione per parti**. Il metodo è vantaggioso se per il termine $g f'$ si conosce un metodo di integrazione.

Per gli integrali definiti, la formula (20) diviene

$$(21) \quad \int_a^b g'(x) f(x) dx = f(b)g(b) - f(a)g(a) - \int_a^b g(x) f'(x) dx.$$

ESEMPIO 2.1. Ecco un primo esempio di applicazione della formula (20)

$$\int x e^x dx = \int x (e^x)' dx = x e^x - \int (x)' e^x dx = x e^x - \int e^x dx = (x - 1)e^x + C.$$

Anche nel caso della funzione $x^2 e^x$ si può procedere in modo analogo, applicando due volte l'integrazione per parti,

$$\begin{aligned} \int x^2 e^x dx &= \int x^2 (e^x)' dx = x^2 e^x - 2 \int x e^x dx = x^2 e^x - 2 \left[x e^x - \int e^x dx \right] \\ &= x^2 e^x - 2 [(x - 1)e^x + C] = (x^2 - 2x + 2)e^x + C. \end{aligned}$$

È possibile risolvere allo stesso modo un qualsiasi integrale del tipo

$$\int p(x) e^x dx \quad p \text{ polinomio di grado } n,$$

infatti basta iterare n volte l'uso della formula di integrazione per parti

$$\begin{aligned} \int p(x) e^x dx &= p(x) e^x - \int p'(x) e^x dx = [p(x) - p'(x)] e^x + \int p''(x) e^x dx \\ &= \dots = [p(x) - p'(x) + \dots + (-1)^n p^{(n)}(x)] e^x + C \end{aligned}$$

Osservando che $e^{\alpha x} = \frac{1}{\alpha} (e^{\alpha x})'$, si possono determinare le primitive anche di funzioni del tipo $p(x) e^{\alpha x}$ con p polinomio. Ad esempio,

$$\begin{aligned} \int x^2 e^{3x} dx &= \frac{1}{3} x^2 e^{3x} - \frac{2}{3} \int x e^{3x} dx \\ &= \frac{1}{3} x^2 e^{3x} - \frac{2}{9} \left(x e^{3x} - \int e^{3x} dx \right) = \left(\frac{1}{3} x^2 - \frac{2}{9} x + \frac{2}{27} \right) e^{3x} + C. \end{aligned}$$

ESEMPIO 2.2. Anche funzioni del tipo prodotto di polinomio e di $\sin x$ o $\cos x$ possono essere risolte integrando per parti. I casi più semplici sono

$$\begin{aligned} \int x \sin x dx &= \int x (-\cos x)' dx = -x \cos x + \int \cos x dx = \sin x - x \cos x + C; \\ \int x \cos x dx &= \int x (\sin x)' dx = x \sin x - \int \sin x dx = \cos x + x \sin x + C. \end{aligned}$$

Iterando il procedimento un certo numero di volte, si calcolano gli integrali

$$\int p(x) \sin(ax) dx, \quad \int p(x) \cos(ax) dx \quad (a \in \mathbb{R}).$$

ESEMPIO 2.3. Proponiamoci di determinare tutte le primitive di $x^2 \sin(2x)$. Applichiamo l'integrazione per parti:

$$\begin{aligned} \int x^2 \sin(2x) dx &= \int x^2 \left(-\frac{1}{2} \cos(2x) \right)' dx = -\frac{1}{2} x^2 \cos(2x) + \int x \cos(2x) dx \\ &= -\frac{1}{2} x^2 \cos(2x) + \frac{1}{2} \int x (\sin(2x))' dx = -\frac{1}{2} x^2 \cos(2x) + \frac{1}{2} x \sin(2x) - \frac{1}{2} \int \sin(2x) dx \\ &= -\frac{1}{2} x^2 \cos(2x) + \frac{1}{2} x \sin(2x) + \frac{1}{4} \cos(2x) + C. \end{aligned}$$

ESEMPIO 2.4. Sempre tramite l'integrazione per parti, si risolvono anche

$$\int p(x) \ln x dx \quad p \text{ polinomio.}$$

Calcoliamo l'integrale di $\ln x$:

$$\begin{aligned} \int \ln x dx &= \int 1 \cdot \ln x dx = \int (x)' \ln x dx = x \ln x - \int x (\ln x)' dx \\ &= x \ln x - \int x \cdot \frac{1}{x} dx = x \ln x - x + C. \end{aligned}$$

Analogamente,

$$\int x \ln x dx = \int \left(\frac{1}{2} x^2 \right)' \ln x dx = \frac{1}{2} x^2 \ln x - \frac{1}{2} \int x dx = \frac{1}{2} x^2 \ln x - \frac{1}{4} x^2 + C.$$

In generale, dato $k \in \mathbb{N}$,

$$\int x^k \ln x dx = \frac{x^{k+1}}{k+1} \ln x - \frac{1}{k+1} \int x^k dx = \frac{x^{k+1}}{k+1} \left(\ln x - \frac{1}{k+1} \right) + C.$$

ESERCIZIO 2.5. Calcolare gli integrali indefiniti

$$\int \arctan x dx, \quad \int x \arctan x dx.$$

ESEMPIO 2.6. Qui usiamo l'integrazione per parti in un modo leggermente diverso: iterando l'applicazione di (20) torniamo all'integrale originale, ottenendo in questo modo un'equazione per la primitiva. In questo modo risolviamo integrali della forma

$$\int e^{ax} \sin(bx) dx, \quad \int e^{ax} \cos(bx) dx.$$

Ad esempio,

$$\begin{aligned}\int e^{2x} \sin(3x) dx &= \frac{1}{3} \int e^{2x} (-\cos(3x))' dx = -\frac{1}{3} \cos(3x)e^{2x} + \frac{2}{3} \int e^{2x} \cos(3x) dx \\ &= -\frac{1}{3} \cos(3x)e^{2x} + \frac{2}{9} \int e^{2x} (\sin(3x))' dx \\ &= \frac{1}{9} (2 \sin(3x) - 3 \cos(3x)) e^{2x} - \frac{4}{9} \int e^{2x} \sin(3x) dx.\end{aligned}$$

Guardando il primo e l'ultimo termine, si ha

$$\int e^{2x} \sin(3x) dx = \frac{1}{9} (2 \sin(3x) - 3 \cos(3x)) e^{2x} - \frac{4}{9} \int e^{2x} \sin(3x) dx,$$

da cui, esplicitando rispetto all'integrale richiesto,

$$\int e^{2x} \sin(3x) dx = \frac{1}{13} (2 \sin(3x) - 3 \cos(3x)) e^{2x} + C.$$

In generale si ottengono le formule (verificare!)

$$\begin{aligned}\int e^{ax} \sin(bx) dx &= \frac{1}{a^2 + b^2} (a \sin(bx) - b \cos(bx)) e^{ax} + C, \\ \int e^{ax} \cos(bx) dx &= \frac{1}{a^2 + b^2} (a \cos(bx) + b \sin(bx)) e^{ax} + C.\end{aligned}$$

Formule ricorsive. Alcune famiglie di integrali (dipendenti da un parametro discreto $n \in \mathbb{N}$), possono essere risolte in modo iterativo, cioè si risolve l'integrale per $n = 1$, e poi si mostra come l'integrale al passo n -esimo si possa ricondurre al calcolo dell'integrale $(n - 1)$ -esimo. Vediamo un paio di esempi. Calcoliamo

$$I_n = \int \sin^{2n} x dx, \quad n \in \mathbb{N}.$$

Allo stesso modo si può calcolare $\int \cos^{2n} x dx$. Calcoliamo I_1 :

$$\begin{aligned}\int \sin^2 x dx &= \int \sin x \cdot \sin x dx = \int \sin x (-\cos x)' dx = -\sin x \cos x + \int \cos^2 x dx \\ &= -\sin x \cos x + \int (1 - \sin^2 x) dx = x - \sin x \cos x - \int \sin^2 x dx.\end{aligned}$$

Abbiamo ottenuto una relazione del tipo $I_1 = x - \sin x \cos x - I_1$, quindi

$$I_1 = \int \sin^2 x dx = \frac{1}{2} (x - \sin x \cos x) + C.$$

Per $n \in \mathbb{N}$,

$$\begin{aligned} I_{n+1} &= \int \sin^{2n+1} x \sin x \, dx = \int \sin^{2n+1} x (-\cos x)' \, dx \\ &= -\sin^{2n+1} x \cos x + (2n+1) \int \sin^{2n} x \cos^2 x \, dx \\ &= -\sin^{2n+1} x \cos x + (2n+1) \int \sin^{2n} x (1 - \sin^2 x) \, dx \\ &= -\sin^{2n+1} x \cos x + (2n+1)I_n - (2n+1)I_{n+1}. \end{aligned}$$

Quindi $I_{n+1} = -\sin^{2n+1} x \cos x + (2n+1)I_n - (2n+1)I_{n+1}$, da cui si deduce

$$I_{n+1} = \frac{1}{2n+2} \{ (2n+1)I_n - \sin^{2n+1} x \cos x \} + C.$$

3. Integrazione di funzioni razionali

Affrontiamo ora il problema di integrare funzioni razionali:

$$\int \frac{P(x)}{Q(x)} \, dx \quad P, Q \text{ polinomi.}$$

Le primitive di una qualsiasi funzione razionale in termini di funzioni esplicite sono note, ma non è in queste Note che troverete i dettagli della questione.

In concreto è possibile completare il calcolo a patto di saper fattorizzare il polinomio a denominatore Q nel prodotto di termini irriducibili, cioè polinomi di primo grado (con molteplicità opportuna) e polinomi di secondo grado irriducibili (con molteplicità opportuna). In questo Paragrafo vedremo come si integrino funzioni razionali nel caso in cui il polinomio Q sia di grado al più due, o sia fattorizzabile in termini di polinomi di grado 1, cioè sia riscrivibile nella forma

$$Q(x) = a(x-x_1)^{k_1} \cdots (x-x_n)^{k_n} \quad a, x_1, \dots, x_n \in \mathbb{R}, \quad a \neq 0, \quad k_1, \dots, k_n \in \mathbb{N}.$$

Denominatore Q di grado 1. Sia $Q(x) = a(x-x_0)$ con $a, x_0 \in \mathbb{R}$ e $a \neq 0$. Se P è un polinomio di grado $p \geq 1$, tramite l'algoritmo di divisione dei polinomi, si determinano un polinomio P_1 di grado $p-1$ e una costante $r \in \mathbb{R}$ tali che

$$\frac{P(x)}{Q(x)} = P_1(x) + \frac{r}{a(x-x_0)}.$$

Quindi l'integrale si può decomporre nella somma di due integrali

$$\int \frac{P(x)}{Q(x)} \, dx = \int P_1(x) \, dx + \frac{r}{a} \int \frac{dx}{x-x_0}.$$

Il polinomio P_1 è integrabile esplicitamente, grazie alla formula

$$\int x^k \, dx = \frac{x^{k+1}}{k+1} + C.$$

Anche l'altro integrale è risolvibile esplicitamente:

$$\frac{r}{a} \int \frac{dx}{x - x_0} = \frac{r}{a} \int \frac{(x - x_0)'}{x - x_0} dx = \frac{r}{a} \ln |x - x_0| + C.$$

Vediamo un esempio. Calcoliamo

$$\int \frac{x^5 + 1}{x - 2} dx.$$

Dato che

$$\frac{x^5 + 1}{x - 2} = x^4 + 2x^3 + 4x^2 + 8x + 16 + \frac{33}{x - 2},$$

si ha

$$\begin{aligned} \int \frac{x^5 + 1}{x - 2} dx &= \int \left(x^4 + 2x^3 + 4x^2 + 8x + 16 + \frac{33}{x - 2} \right) dx \\ &= \frac{1}{5}x^5 + \frac{1}{2}x^4 + \frac{4}{3}x^3 + 4x^2 + 16x + 33 \ln |x - 2| + C. \end{aligned}$$

Denominatore Q di grado 2. Supponiamo che Q sia un polinomio di grado 2. In questo caso Q è scrivibile nella forma

$$Q(x) = a(x^2 + 2bx + c) \quad a, b, c \in \mathbb{R}, \quad a \neq 0.$$

Se il polinomio a numeratore P ha grado $p \geq 2$, allora è possibile applicare l'algoritmo di divisione di polinomi e riscrivere la funzione razionale come somma

$$(22) \quad \frac{P(x)}{Q(x)} = P_1(x) + \frac{R(x)}{Q(x)},$$

dove P_1 è un polinomio di grado $p - 2$ e R è un polinomio di grado minore o uguale a 1. L'integrale della funzione razionale è la somma di due integrali

$$\int \frac{P(x)}{Q(x)} dx = \int P_1(x) dx + \int \frac{R(x)}{Q(x)} dx.$$

Il primo dei due integrali è risolvibile esplicitamente per via elementare. Consideriamo il secondo. Supponiamo che il resto R sia di grado 1 e scriviamolo nella forma $R(x) = \alpha(x + \beta)$ con $\alpha \neq 0$ e $\beta \in \mathbb{R}$. Si tratta di calcolare

$$\int \frac{\alpha(x + \beta)}{a(x^2 + 2bx + c)} dx = \frac{\alpha}{a} \int \frac{x + \beta}{x^2 + 2bx + c} dx.$$

Come primo passo, "costruiamo" a numeratore la derivata del denominatore. Moltiplichiamo e dividiamo per due e, successivamente, sommiamo e sottraiamo $2b$

$$\begin{aligned} \int \frac{\alpha(x + \beta)}{a(x^2 + 2bx + c)} dx &= \frac{\alpha}{2a} \int \frac{2x + 2\beta}{x^2 + 2bx + c} dx \\ &= \frac{\alpha}{2a} \int \frac{(2x + 2b) + 2(\beta - b)}{x^2 + 2bx + c} dx = \dots \end{aligned}$$

L'integrale finale può essere riscritto come somma dei due integrali di cui il primo è della forma $\int \frac{\phi'}{\phi} dx$; quindi

$$\begin{aligned} \dots &= \frac{\alpha}{2a} \int \frac{(x^2 + 2bx + c)'}{x^2 + 2bx + c} dx + \frac{2\alpha(\beta - b)}{2a} \int \frac{dx}{x^2 + 2bx + c} \\ &= \frac{\alpha}{2a} \ln |x^2 + 2bx + c| + \frac{2\alpha(\beta - b)}{2a} \int \frac{dx}{x^2 + 2bx + c}. \end{aligned}$$

Rimane quindi da risolvere l'integrale

$$(23) \quad \int \frac{dx}{x^2 + 2bx + c}.$$

Nel caso in cui R in (22) sia di grado 0 ci si riconduce direttamente a questa situazione. La risoluzione dell'integrale (23) varia a seconda di quante radici reali abbia il denominatore, cioè a seconda che sia $b^2 > c$, $b^2 = c$ o $b^2 < c$. Trattiamo i tre casi separatamente. Ci ricondurremo (sostanzialmente) ai seguenti integrali elementari

$$\begin{aligned} \text{Caso I : } \quad b^2 > c &\quad \longrightarrow \quad \int \frac{dx}{x} = \ln |x| + C, \\ \text{Caso II : } \quad b^2 = c &\quad \longrightarrow \quad \int \frac{dx}{x^2} = -\frac{1}{x} + C, \\ \text{Caso III : } \quad b^2 < c &\quad \longrightarrow \quad \int \frac{dx}{1 + x^2} = \arctan x + C. \end{aligned}$$

Caso I: $b^2 > c$. In questo caso il denominatore ha due radici reali

$$x^2 + 2bx + c = 0 \quad \Longleftrightarrow \quad x = -b \pm \sqrt{b^2 - c}.$$

Indicando le radici con x_1 e x_2 , il polinomio si fattorizza:

$$x^2 + 2bx + c = (x - x_1)(x - x_2).$$

Decomponiamo la funzione integranda nella forma

$$\frac{1}{x^2 + 2bx + c} = \frac{A_1}{x - x_1} + \frac{A_2}{x - x_2},$$

dove $A_1, A_2 \in \mathbb{R}$ sono due costanti da determinare. La somma delle due frazioni a secondo membro è uguale a

$$\frac{(A_1 + A_2)x - (A_1x_2 + A_2x_1)}{x^2 + 2bx + c},$$

e quindi A_1 e A_2 devono essere tali che $(A_1 + A_2)x - (A_1x_2 + A_2x_1) = 1$. Dato che due polinomi coincidono se e solo se coincidono i loro coefficienti, le costanti A_1, A_2 sono le soluzioni del sistema lineare (il cui determinante è $x_1 - x_2$ che, nel caso $b^2 > c$, è diverso da zero)

$$A_1 + A_2 = 0, \quad A_1x_2 + A_2x_1 = -1.$$

Individuati i valori di A_1 e A_2 , l'integrale è risolto:

$$\begin{aligned}\int \frac{dx}{x^2 + 2bx + c} &= A_1 \int \frac{dx}{x - x_1} + A_2 \int \frac{dx}{x - x_2} \\ &= A_1 \ln |x - x_1| + A_2 \ln |x - x_2| + C.\end{aligned}$$

ESERCIZIO 3.1. Calcolare

$$\int \frac{x^3}{x^2 - x - 2} dx.$$

Caso II: $b^2 = c$. In questa situazione, si tratta di risolvere

$$\int \frac{dx}{x^2 + 2bx + b^2}.$$

Questo integrale è immediato, infatti

$$\int \frac{dx}{x^2 + 2bx + b^2} = \int \frac{dx}{(x + b)^2} = -\frac{1}{x + b} + C.$$

ESERCIZIO 3.2. Calcolare

$$\int \frac{x(x + 3)}{(x - 1)^2} dx.$$

Caso III: $b^2 < c$. Questo caso è già stato considerato nell'Esempio 1.6. Ritroviamo qui la stessa soluzione senza fare ricorso al grafico della funzione integranda.

Dato che il polinomio $x^2 + 2bx + c$ è irriducibile, l'obiettivo è di ricondursi, con un opportuno cambiamento di variabili, all'integrale elementare

$$\int \frac{dx}{1 + x^2} = \arctan x + C.$$

Chiamiamo $\nu := \frac{1}{\sqrt{c-b^2}} > 0$ e riscriviamo in maniera opportuna il denominatore

$$x^2 + 2bx + c = x^2 + 2bx + b^2 + (c - b^2) = (x + b)^2 + \frac{1}{\nu^2} = \frac{1}{\nu^2} \{[\nu(x + b)]^2 + 1\}.$$

Ponendo $t = \nu(x + b)$,

$$\int \frac{dx}{x^2 + 2bx + c} = \nu \int \frac{dt}{1 + t^2} = \nu \arctan t + C = \nu \arctan(\nu(x + b)) + C.$$

Dalla definizione di ν si deduce che

$$\int \frac{dx}{x^2 + 2bx + c} = \frac{1}{\sqrt{c - b^2}} \arctan \left(\frac{x + b}{\sqrt{c - b^2}} \right) + C.$$

ESERCIZIO 3.3. Calcolare

$$\int \frac{3x - 2}{x^2 - 2x + 2} dx.$$

Denominatore Q con sole radici reali. Consideriamo prima di tutto un caso semplice: il polinomio Q ha radici reali *distinte*, cioè

$$Q(x) = a(x - x_1) \cdots (x - x_n) \quad x_1, \dots, x_n \in \mathbb{R} \quad \text{con } x_i \neq x_j \quad \text{se } i \neq j.$$

Se $p \geq n$, il primo passaggio è sempre lo stesso: si usa l'algoritmo della divisione di polinomi per riscrivere la funzione razionale nella forma

$$\frac{P(x)}{Q(x)} = P_1(x) + \frac{R(x)}{Q(x)}.$$

dove P_1 è un polinomio di grado $p - n$, e R è un polinomio di grado minore di n . L'integrale si decompone in

$$\int \frac{P(x)}{Q(x)} dx = \int P_1(x) dx + \int \frac{R(x)}{Q(x)} dx.$$

A questo punto sfruttiamo la fattorizzazione di Q per riscrivere la funzione razionale R/Q come somma di funzioni razionali più semplici. Cerchiamo n costanti A_1, \dots, A_n tali che

$$\frac{R(x)}{a(x - x_1) \cdots (x - x_n)} = \frac{1}{a} \left(\frac{A_1}{x - x_1} + \cdots + \frac{A_n}{x - x_n} \right).$$

Per determinare le costanti A_1, \dots, A_n si può imporre l'uguaglianza dei due membri ottenendo un sistema lineare. Equivalentemente si può moltiplicare per $x - x_1$ entrambi i membri

$$\frac{R(x)}{a(x - x_2) \cdots (x - x_n)} = \frac{1}{a} \left\{ A_1 + \frac{A_2(x - x_1)}{x - x_2} + \cdots + \frac{A_n(x - x_1)}{x - x_n} \right\}.$$

e successivamente porre $x = x_1$, ottenendo il valore di A_1

$$A_1 = \frac{R(x_1)}{(x_1 - x_2) \cdots (x_1 - x_n)}$$

Analogamente per A_2, \dots, A_n . Determinate le costanti A_i , si calcola l'integrale:

$$\begin{aligned} \int \frac{R(x)}{Q(x)} dx &= \frac{1}{a} \int \left(\frac{A_1}{x - x_1} + \cdots + \frac{A_n}{x - x_n} \right) dx \\ &= \frac{1}{a} (A_1 \ln |x - x_1| + \cdots + A_n \ln |x - x_n|) + C. \end{aligned}$$

Per digerire la tecnica, calcoliamo

$$\int \frac{dx}{(x + 1)(x + 2)(x + 3)}.$$

Dato che il grado del numeratore è minore del grado del denominatore, non occorre applicare l'algoritmo di divisione di polinomi. Passiamo subito alla decomposizione:

cerchiamo $A_1, A_2, A_3 \in \mathbb{R}$ tali che

$$\frac{1}{(x+1)(x+2)(x+3)} = \frac{A_1}{x+1} + \frac{A_2}{x+2} + \frac{A_3}{x+3}.$$

Moltiplichiamo per $x+1$ e calcoliamo in $x = -1$

$$\frac{1}{(x+2)(x+3)} = A_1 + \frac{A_2(x+1)}{x+2} + \frac{A_3(x+1)}{x+3} \quad \Rightarrow \quad A_1 = \frac{1}{2}.$$

Analogamente

$$\frac{1}{(x+1)(x+3)} = \frac{A_1(x+2)}{x+1} + A_2 + \frac{A_3(x+2)}{x+3} \quad \Rightarrow \quad A_2 = -1.$$

$$\frac{1}{(x+1)(x+2)} = \frac{A_1(x+3)}{x+1} + \frac{A_2(x+3)}{x+2} + A_3 \quad \Rightarrow \quad A_3 = \frac{1}{2}.$$

Quindi

$$\begin{aligned} \int \frac{dx}{(x+1)(x+2)(x+3)} &= \frac{1}{2} \int \left(\frac{1}{x+1} - \frac{2}{x+2} + \frac{1}{x+3} \right) dx \\ &= \frac{1}{2} \ln \left| \frac{(x+1)(x+3)}{(x+2)^2} \right| + C. \end{aligned}$$

Passiamo al caso generale: il denominatore Q si decompone come

$$Q(x) = a(x-x_1)^{k_1} \cdots (x-x_n)^{k_n} \quad a, x_1, \dots, x_n \in \mathbb{R}, \quad k_1, \dots, k_n \in \mathbb{N}.$$

Anche in questo caso, dopo aver applicato (se necessario) l'algoritmo di divisione di polinomi, si deve risolvere un integrale della forma

$$\int \frac{R(x)}{Q(x)} dx,$$

dove R è un polinomio di grado strettamente minore di quello di Q . In questo caso si cerca, analogamente a quanto fatto nel caso di Q di secondo grado con due radici coincidenti, una decomposizione della forma

$$\begin{aligned} \frac{R(x)}{a(x-x_1)^{k_1} \cdots (x-x_n)^{k_n}} &= \frac{1}{a} \left\{ \frac{A_1^1}{x-x_1} + \frac{A_1^2}{(x-x_1)^2} + \cdots + \frac{A_1^{k_1}}{(x-x_1)^{k_1}} + \cdots \right. \\ &\quad \left. + \frac{A_n^1}{x-x_n} + \frac{A_n^2}{(x-x_n)^2} + \cdots + \frac{A_n^{k_n}}{(x-x_n)^{k_n}} \right\}, \end{aligned}$$

dove le costanti $A_i^j \in \mathbb{R}$ con $i = 1, \dots, n$ e $j = 1, \dots, k_i$ sono da determinare. Una volta determinate queste costanti, l'integrale è risolto dato che

$$\int \frac{A dx}{(x-\bar{x})^k} = \begin{cases} A \ln |x-\bar{x}| + C & k = 1, \\ -\frac{A}{k-1} \cdot \frac{1}{(x-\bar{x})^{k-1}} + C & k > 1. \end{cases}$$

ESEMPIO 3.4. Ad esempio, consideriamo l'integrale

$$\int \frac{dx}{x^2(x+1)^2}.$$

In questo caso cerchiamo le costanti $A, B, C, D \in \mathbb{R}$ tali che

$$\frac{1}{x^2(x+1)^2} = \frac{A}{x} + \frac{B}{x^2} + \frac{C}{x+1} + \frac{D}{(x+1)^2}.$$

Imponendo l'uguaglianza, si ottengono

$$A = -2, \quad B = 1, \quad C = 2, \quad D = 1.$$

Perciò

$$\int \frac{dx}{x^2(x+1)^2} = 2 \ln \left| \frac{x+1}{x} \right| - \frac{1}{x} - \frac{1}{x+1} + C.$$

Altre classi di funzioni. Vediamo qualche altra classe di funzioni che si possono ricondurre, tramite un cambio di variabile, all'integrale di funzioni razionali.

ESEMPIO 3.5. Supponiamo di voler calcolare

$$\int R(\sin x, \cos x) dx$$

dove R è una funzione razionale dei suoi argomenti. Dalle relazioni

$$(24) \quad \sin x = \frac{2t}{1+t^2} \quad \text{e} \quad \cos x = \frac{1-t^2}{1+t^2} \quad \text{dove} \quad t = \tan\left(\frac{x}{2}\right),$$

ponendo $t = \tan(x/2)$ o, equivalentemente, $x = 2 \arctan t$, dato che $dx = 2/(1+t^2)dt$, l'integrale si trasforma nell'integrale di una funzione razionale

$$\int R\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2}\right) \frac{2}{1+t^2} dt,$$

Ad esempio,

$$\int \frac{dx}{\sin x} = \int \frac{1+t^2}{2t} \frac{2}{1+t^2} dt = \int \frac{dt}{t} = \ln |t| + C = \ln |\tan(x/2)| + C.$$

$$\int \frac{dx}{\cos x} = \int \frac{2 dt}{1-t^2} = \int \left(\frac{1}{1+t} + \frac{1}{1-t} \right) dt = \ln \left| \frac{1+\tan(x/2)}{1-\tan(x/2)} \right| + C.$$

Non sempre la sostituzione (24) è conveniente. Ad esempio, consideriamo

$$\int \frac{\sin x}{\cos^{100} x} dx.$$

Tramite (24), l'integrale si trasforma in

$$\int \frac{\sin x}{\cos^{100} x} dx = \int \frac{4t(1+t^2)^{98}}{(1-t^2)^{100}} dt,$$

con un polinomio di grado duecento a denominatore! Che fare? Sarebbe invece stato molto più semplice porre $s = \cos x$, da cui

$$\int \frac{\sin x}{\cos^{100} x} dx = - \int \frac{ds}{s^{100}} = \frac{1}{99s^{99}} + C = \frac{1}{99 \cos^{99} x} + C.$$

ESEMPIO 3.6. Abbiamo un problema: calcolare l'area della regione di piano

$$\Omega = \left\{ (x, y) \in \mathbb{R} : \frac{x^2}{a^2} + \frac{y^2}{b^2} \leq 1 \right\} \quad a, b > 0.$$

L'area $|\Omega|$ di Ω è pari al valore dell'integrale definito

$$|\Omega| = 4b \int_0^a \sqrt{1 - (x^2/a^2)} dx.$$

Introduciamo la variabile t definita da $x = a \cos t$, da cui $dx = -a \sin t dt$:

$$|\Omega| = 4ab \int_0^{\pi/2} \sqrt{1 - \cos^2 t} \sin t dt = 4ab \int_0^{\pi/2} \sin^2 t dt = 2ab [t - \sin t \cos t]_0^{\pi/2} = \pi ab.$$

Quindi l'area della regione delimitata dall'ellissi di semiassi a e b è πab .

Allo stesso modo è possibile integrare funzioni del tipo

$$R(x, \sqrt{1 - (x^2/a^2)})$$

con R funzione razionale dei suoi argomenti. Infatti

$$\int R(x, \sqrt{1 - (x^2/a^2)}) dx = -a \int R(a \cos t, \sin t) \sin t dt.$$

dove $x = a \cos t$, e il secondo membro è una funzione razionale in $\sin t$ e $\cos t$.

CAPITOLO 3

I numeri complessi

Fino ad adesso abbiamo studiato funzioni reali di variabile reale. Vari problemi matematici suggeriscono un'estensione dell'insieme dei numeri reali, particolarmente utile anche per molte applicazioni. Questa estensione dei numeri reali porta ad un nuovo concetto di numero: il *numero complesso*.

1. Nascita dei numeri complessi e loro infanzia

Il problema di partenza è il seguente: *in \mathbb{R} esistono polinomi di grado maggiore di 1 che non hanno radici (reali)*, ad esempio, $x^2 + 1$. Si potrebbe rimanere lì a guardare, risponderci che è un fatto della vita, e passare ad un altro problema. Per chi invece desidera cercare di andare oltre, l'idea plausibile è di costruire un nuovo insieme che estenda l'insieme dei numeri reali e che contenga anche il “numero” i , *unità immaginaria*, soluzione dell'equazione $x^2 + 1 = 0$. Dedichiamoci alla costruzione di questo nuovo insieme, che indichiamo con il simbolo \mathbb{C} .

DEFINIZIONE 1.1. *Si definisce insieme dei numeri complessi \mathbb{C} l'insieme*

$$\mathbb{C} := \{z = a + ib : a, b \in \mathbb{R}\}$$

con la regola di calcolo seguente: i numeri complessi si sommano, sottraggono, moltiplicano, dividono tenendo conto della proprietà $i^2 = -1$.

La regola di calcolo di somma e moltiplicazione per numeri complessi va letta, in concreto, come segue: dati $z_1 = a + ib$ e $z_2 = c + id$, allora

$$z_1 + z_2 = (a + ib) + (c + id) = (a + c) + i(b + d),$$

$$z_1 \cdot z_2 = (a + ib)(c + id) = ac + i(ad + bc) + i^2bd = (ac - bd) + i(ad + bc).$$

Tutte le proprietà di somma e moltiplicazione (commutativa, associativa, distributiva) valgono anche in \mathbb{C} .

Piccolo glossario per i numeri complessi

$z = a + ib$: numero complesso

$\mathbb{R} = \{z \in \mathbb{C} : b = 0\}$: numeri reali

$i\mathbb{R} = \{z \in \mathbb{C} : a = 0\}$: immaginari puri

i : unità immaginaria
 $\operatorname{Re} z = a$: parte reale di z
 $\operatorname{Im} z = b$: parte immaginaria di z

L'insieme \mathbb{C} può essere rappresentato geometricamente come un piano

$$\mathbb{C} = \{(a, b) : a, b \in \mathbb{R}\} = \mathbb{R}^2.$$

La struttura di \mathbb{C} è comunque più ricca di quella del piano dato che in \mathbb{C} è *definita l'operazione di prodotto*.

I numeri complessi si devono poter rappresentare sempre nella forma $a+ib$. Nel caso del rapporto di due numeri, può non essere evidente che questo sia possibile. Vediamo un esempio particolare. Supponiamo di voler scrivere nella forma $a+ib$ con $a, b \in \mathbb{R}$ il numero complesso $z = \frac{1+i}{1-i}$. Moltiplicando e dividendo per $1+i$, e ragionando con le stesse regole algebriche note su \mathbb{R} insieme alla relazione $i^2 = -1$, si ottiene

$$\frac{1+i}{1-i} = \frac{1+i}{1-i} \cdot \frac{1+i}{1+i} = \frac{(1+i)^2}{(1-i)(1+i)} = \frac{1+2i+i^2}{1-i^2} = i.$$

In generale, per scrivere un rapporto di numeri complessi nella forma $a+ib$, procediamo alla stessa maniera. Siano $a, b, c, d \in \mathbb{R}$ con $(c, d) \neq (0, 0)$, allora

$$\frac{a+ib}{c+id} = \frac{a+ib}{c+id} \cdot \frac{c-id}{c-id} = \frac{(ac+bd) + i(bc-ad)}{c^2+d^2}.$$

In particolare l'inverso del numero complesso $c+id$ si scrive come

$$\frac{1}{c+id} = \frac{1}{c+id} \cdot \frac{c-id}{c-id} = \frac{c-id}{c^2+d^2} = \frac{c}{c^2+d^2} - i \frac{d}{c^2+d^2},$$

che mostra che *ogni numero complesso z diverso da 0 è invertibile, cioè esiste un numero z^{-1} tale che $z \cdot z^{-1} = z^{-1} \cdot z = 1$* .

Dato $c+id \in \mathbb{C}$, il numero complesso $c-id \in \mathbb{C}$, che si ottiene cambiando di segno la parte immaginaria ha un ruolo determinante in questa costruzione ed è detto **complesso coniugato**. Il complesso coniugato di z si indica con \bar{z} .

ESERCIZIO 1.2. Verificare la validità di:

$$\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2 \quad \overline{z_1 - z_2} = \bar{z}_1 - \bar{z}_2 \quad \overline{z_1 \cdot z_2} = \bar{z}_1 \cdot \bar{z}_2 \quad \overline{\frac{z_1}{z_2}} = \frac{\bar{z}_1}{\bar{z}_2}.$$

Rappresentazione polare. Un numero complesso $z = x+iy$ si può rappresentare come un punto P di coordinate (x, y) nel piano cartesiano e, in quanto tale, può essere individuato anche dalle coordinate polari (r, θ) :

$$x = r \cos \theta, \quad y = r \sin \theta.$$

In questo caso $r = \sqrt{x^2 + y^2}$ è la distanza del punto $P = (x, y)$ dall'origine O e θ

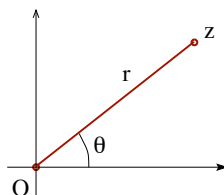


FIGURA 1. Rappresentazione polare di un numero complesso.

misura l'angolo tra il semiasse positivo x e il segmento di estremi 0 e z . Il numero complesso $z = x + iy$ si può, quindi, scrivere come

$$(25) \quad z = r(\cos \theta + i \sin \theta).$$

La quantità r è il **modulo** di z ; l'angolo θ si dice **argomento** di z e si indica con $\arg z$. Per via della periodicità delle funzioni \sin e \cos , l'argomento non è individuato univocamente. Si parla di **argomento principale** se $\theta \in [0, 2\pi)$ e si indica con $\text{Arg } z$ (Attenzione! c'è chi parla di "argomento principale" se $\theta \in (-\pi, \pi]$). Il complesso coniugato \bar{z} di z corrisponde al numero complesso di stesso modulo di z e argomento $-\theta$.

Tramite la rappresentazione polare è facile calcolare il prodotto di numeri complessi:

$$\begin{aligned} zz' &= r(\cos \theta + i \sin \theta) r'(\cos \theta' + i \sin \theta') \\ &= rr'[(\cos \theta \cos \theta' - \sin \theta \sin \theta') + i(\cos \theta \sin \theta' + \sin \theta \cos \theta')] \\ &= rr'[\cos(\theta + \theta') + i \sin(\theta + \theta')]. \end{aligned}$$

Questa formula permette di dare un'interpretazione geometrica del prodotto di due numeri complessi: il risultato è un numero complesso il cui modulo è pari al prodotto dei moduli e l'argomento è la somma degli argomenti. Ad esempio, che succede moltiplicando un generico numero complesso z per l'unità immaginaria i ? Dato che il modulo di i è 1 e il suo argomento è $\frac{\pi}{2}$, il numero iz si trova alla stessa distanza dall'origine di z e ha argomento aumentato di $\frac{\pi}{2}$ rispetto a quello di z . In definitiva, il numero iz si ottiene ruotando il punto z in senso antiorario di un angolo retto.

Dalla formula per il prodotto, segue la formula per la potenza n -esima di z

$$z^n = [r(\cos \theta + i \sin \theta)]^n = r^n [\cos(n\theta) + i \sin(n\theta)].$$

Il modulo nei complessi. Il modulo del numero complesso $a + ib$ è il modulo del punto $(a, b) \in \mathbb{R}^2$, cioè

$$\text{modulo di } z = a + ib : \quad |z| := \sqrt{a^2 + b^2}$$

Il modulo di un numero complesso è sempre un numero reale non negativo e gode delle stesse proprietà di cui gode il modulo reale

PROPOSIZIONE 1.3. *Il modulo complesso gode delle seguenti proprietà:*

- (i) $|z| \geq 0$ per ogni $z \in \mathbb{C}$ e $|z| = 0$ se e solo se $z = 0$;
- (ii) $|z_1 z_2| = |z_1| |z_2|$ per ogni $z_1, z_2 \in \mathbb{C}$;
- (iii) **diseguaglianza triangolare:** $|z_1 + z_2| \leq |z_1| + |z_2|$ per ogni $z_1, z_2 \in \mathbb{C}$.

DIMOSTRAZIONE. La prima proprietà è evidente (basta osservare che una somma di quadrati è nulla se e solo se sono nulli entrambi gli addendi).

La seconda si verifica direttamente, se $z_1 = x_1 + iy_1$ e $z_2 = x_2 + iy_2$, allora

$$\begin{aligned} |z_1 z_2| &= |(x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1)| = \sqrt{(x_1 x_2 - y_1 y_2)^2 + (x_1 y_2 + x_2 y_1)^2} \\ &= \sqrt{x_1^2 x_2^2 + y_1^2 y_2^2 + x_1^2 y_2^2 + x_2^2 y_1^2} \end{aligned}$$

$$|z_1| |z_2| = \sqrt{x_1^2 + y_1^2} \sqrt{x_2^2 + y_2^2} = \sqrt{(x_1^2 + y_1^2)(x_2^2 + y_2^2)} = \sqrt{x_1^2 x_2^2 + x_1^2 y_2^2 + y_1^2 x_2^2 + y_1^2 y_2^2}.$$

Resta da dimostrare la diseguaglianza triangolare. Per la definizione di modulo, la diseguaglianza triangolare si riscrive come

$$(26) \quad \sqrt{(x_1 + x_2)^2 + (y_1 + y_2)^2} \leq \sqrt{x_1^2 + y_1^2} + \sqrt{x_2^2 + y_2^2}$$

Elevando al quadrato

$$(x_1 + x_2)^2 + (y_1 + y_2)^2 \leq x_1^2 + y_1^2 + 2\sqrt{(x_1^2 + y_1^2)(x_2^2 + y_2^2)} + x_2^2 + y_2^2$$

Sviluppando i quadrati dei due binomi a sinistra e semplificando,

$$x_1 x_2 + y_1 y_2 \leq \sqrt{(x_1^2 + y_1^2)(x_2^2 + y_2^2)}$$

Se il termine a sinistra è negativo, la disuguaglianza è verificata; altrimenti, elevando di nuovo al quadrato,

$$x_1^2 x_2^2 + 2x_1 x_2 y_1 y_2 + y_1^2 y_2^2 \leq (x_1^2 + y_1^2)(x_2^2 + y_2^2) = x_1^2 x_2^2 + x_1^2 y_2^2 + y_1^2 x_2^2 + y_1^2 y_2^2.$$

Semplificando si ottiene $2x_1 x_2 y_1 y_2 \leq x_1^2 y_2^2 + y_1^2 x_2^2$, che si può riscrivere come

$$(x_1 y_2 - x_2 y_1)^2 = x_1^2 y_2^2 - 2x_1 x_2 y_1 y_2 + y_1^2 x_2^2 \geq 0,$$

che è sempre vera. □

ESERCIZIO 1.4. *Verificare che, per ogni $z \in \mathbb{C}$, valgono (nella seconda $z \neq 0$)*

$$z \cdot \bar{z} = |z|^2 \quad \text{e} \quad z^{-1} = \frac{\bar{z}}{|z|^2}$$

Il fatto che in \mathbb{C} sia definito il modulo (con le stesse proprietà note in \mathbb{R}) è di fondamentale importanza: con quest'oggetto è possibile estendere ai complessi tutto quello che si è visto nei reali a partire dalla nozione di distanza.

Nel piano complesso \mathbb{C} , la distanza tra due numeri complessi z e z' è data da $|z - z'|$, dove $|\cdot|$ rappresenta il modulo complesso.

Grazie alla presenza della distanza, è possibile introdurre in \mathbb{C} il concetto di intorno.

DEFINIZIONE 1.5. Dato $z_0 \in \mathbb{C}$ e $r > 0$, l'intorno di z_0 di raggio r (indicato con $I(z_0; r)$) è l'insieme dei punti $z \in \mathbb{C}$ che distano da z_0 meno di r :

$$I(z_0; r) := \{z \in \mathbb{C} : |z - z_0| < r\}.$$

Da qui limiti e tutto il resto...

Il Teorema fondamentale dell'algebra. Fin qui abbiamo scoperto che l'insieme \mathbb{C} gode delle stesse proprietà di \mathbb{R} per quanto riguarda le operazioni di somma e prodotto e anche per quanto riguarda la struttura metrica. Ma in \mathbb{C} , c'è qualcosa di più...

TEOREMA 1.6 (Teorema fondamentale dell'algebra). *Ogni polinomio $p(z) = a_0 + a_1z + \dots + a_nz^n$ a coefficienti complessi $a_i \in \mathbb{C}$ di grado $n \geq 1$ ammette almeno uno zero $z^* \in \mathbb{C}$, cioè esiste $z^* \in \mathbb{C}$ tale che $p(z^*) = 0$.*

Conseguenza di questo Teorema è che un polinomio di grado n in \mathbb{C} ha esattamente n zeri (contati con la loro molteplicità). Ad esempio, il polinomio a coefficienti reali

$$p(z) = z^2 + 2bz + c \quad b, c \in \mathbb{R},$$

ammette in \mathbb{C} sempre due zeri: se $b^2 > c$, gli zeri sono reali e sono $-b \pm \sqrt{b^2 - c}$, se $b^2 = c$ c'è un unico zero, $z = -b$, con molteplicità due, nel caso $b^2 < c$, gli zeri sono due numeri complessi coniugati, $-b \pm i\sqrt{c - b^2}$.

2. Successioni, serie e continuità nei complessi

Una volta definito l'insieme dei numeri complessi \mathbb{C} e introdotta la nozione di modulo, si può ripetere buona parte della teoria sviluppata in \mathbb{R} : successioni, serie, limiti, continuità, derivabilità... Ripercorriamo ora, rapidamente, quello che abbiamo visto nel caso reale, vedendo come sia possibile definire gli stessi oggetti nel caso complesso.

Successioni complesse. Una funzione che associa ad ogni numero naturale $n \in \mathbb{N}$ un valore $z_n \in \mathbb{C}$ è una **successione complessa**. Una successione complessa z_n può essere pensata come una famiglia di valori ordinati in base al loro indice n : z_0, z_1, z_2, \dots

DEFINIZIONE 2.1. Una successione complessa z_n è **convergente a $\ell \in \mathbb{C}$** se

$$(27) \quad \forall \varepsilon > 0 \quad \exists N \in \mathbb{N} \quad \text{tale che} \quad |z_n - \ell| < \varepsilon \quad \forall n > N.$$

Con una certa dose di fantasia, si scrive $\lim_{n \rightarrow +\infty} z_n = \ell$.

Una successione z_n è **infinitesima** se converge a 0 per $n \rightarrow +\infty$.

Il modulo che compare in (27) è il modulo del numero complesso $z_n - \ell$, e l'essere infinitesima significa semplicemente che comunque si fissi un intorno dell'origine 0 (misurato dal valore di ε), tutti gli elementi della successione tranne al più un numero finito (controllato dall'indice N) si trovano nell'intorno fissato.

Dato che il modulo complesso di un numero reale coincide con l'usuale definizione di modulo di un numero reale, la Definizione 2.1 è, a tutti gli effetti, una estensione di quella già vista per le successioni reali.

OSSERVAZIONE 2.2. A guardar bene, la (27) dice che la successione di numeri reali non negativi $d_n := |z_n - \ell|$ tende a zero per $n \rightarrow +\infty$. Quindi una successione complessa z_n converge ad ℓ se e solo se la distanza di z_n da ℓ è infinitesima per $n \rightarrow +\infty$.

ESEMPIO 2.3. Un esempio interessante di successione complessa è

$$z_n = q^n \quad (q \in \mathbb{C}).$$

Dato che $|z_n| = |q^n| = |q|^n$, il comportamento della successione nel caso $|q| < 1$ è determinato da quello della successione reale $a_n = |q|^n$

$$|q| < 1 \quad \Rightarrow \quad \lim_{n \rightarrow +\infty} q^n = 0.$$

Cosa succede nel caso $|q| \geq 1$? Ad esempio, come si comportano le successioni complesse $z_n = (2i)^n$ e $w_n = i^n$ per $n \rightarrow \infty$?

Ad una successione complessa $z_n = x_n + iy_n$ sono associate in modo naturale le due successioni reali x_n e y_n della sua parte reale e immaginaria, rispettivamente. E' possibile dimostrare che

$$(28) \quad z_n = x_n + iy_n \quad \text{è convergente in } \mathbb{C} \quad \Longleftrightarrow \quad x_n, y_n \quad \text{sono convergenti in } \mathbb{R}.$$

Quindi trattare successioni complesse può essere visto come trattare contemporaneamente una coppia di successioni reali.

ESERCIZIO 2.4. (i) *Dimostrare che, per ogni $z \in \mathbb{C}$, vale*

$$\max\{|\operatorname{Re} z|, |\operatorname{Im} z|\} \leq |z| \leq |\operatorname{Re} z| + |\operatorname{Im} z|.$$

(ii) *Utilizzare (i) per dimostrare l'equivalenza (28).*

La differenza sostanziale tra \mathbb{C} e \mathbb{R} sta nel fatto che *in \mathbb{C} non è definito un ordinamento*. Questo vuol dire che, per successioni complesse, non ha senso parlare di monotonia, né di divergenza a $+\infty$ e $-\infty$. Resta ben definito il concetto di divergenza in modulo: *la successione z_n diverge ad ∞ (o diverge in modulo), se la successione di numeri reali positivi $|z_n|$ diverge a $+\infty$.*

Serie complesse. Allo stesso modo, si possono considerare serie a termini complessi. Una serie complessa

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

è **convergente** se, in \mathbb{C} , esiste il limite della successione complessa $s_n := \sum_{k=0}^n z_k$ (successione delle somme parziali). Tale limite è detto **somma della serie**.

ESEMPIO 2.5. Come nel caso reale, è possibile considerare la serie geometrica, definita dalla successione $z_n = q^n$ con $q \in \mathbb{C}$. La successione delle somme parziali è

$$s_n = 1 + q + q^2 + \cdots + q^{n-1} + q^n = \frac{1 - q^{n+1}}{1 - q}.$$

Quindi se $|q| < 1$, la serie converge e la sua somma è

$$(29) \quad \sum_{n=0}^{\infty} q^n = \frac{1}{1 - q} \quad q \in \mathbb{C} \quad \text{tale che} \quad |q| < 1.$$

Otteniamo quindi la stessa formula già vista nel caso reale.

Anche la nozione di convergenza assoluta si estende al caso complesso.

DEFINIZIONE 2.6. *La serie complessa $\sum_{n=0}^{\infty} z_n$ è assolutamente convergente se la serie reale $\sum_{n=0}^{\infty} |z_n|$ è convergente.*

Dato che la serie $\sum_{n=0}^{\infty} |z_n|$ è una serie a termini reali, è possibile, almeno in alcuni casi, ricondurre la convergenza di una serie complessa a quella di una serie reale, tornando quindi in un ambito più familiare.

Inoltre, vale sempre l'implicazione

$$\text{convergenza assoluta} \quad \Rightarrow \quad \text{convergenza (semplice)}.$$

Infatti, se la serie $\sum_{n=0}^{\infty} |z_n|$ è convergente, allora lo sono anche le serie $\sum_{n=0}^{\infty} |\operatorname{Re} z_n|$ e $\sum_{n=0}^{\infty} |\operatorname{Im} z_n|$. Dato che per serie reali la convergenza assoluta implica la convergenza semplice, la precedente affermazione implica che le serie reali di termini generici $\operatorname{Re} z_n$ e $\operatorname{Im} z_n$ sono semplicemente convergenti. Per concludere, basta osservare che

$$\operatorname{Re} s_n = \operatorname{Re} \sum_{k=0}^n z_k = \sum_{k=0}^n \operatorname{Re} z_k \quad \text{e} \quad \operatorname{Im} s_n = \operatorname{Im} \sum_{k=0}^n z_k = \sum_{k=0}^n \operatorname{Im} z_k,$$

e ricordare (28)...

Per serie convergenti, vale la seguente versione della disuguaglianza triangolare

$$\left| \sum_{n=0}^{\infty} z_n \right| \leq \sum_{n=0}^{\infty} |z_n|.$$

ESEMPIO 2.7. Utilizziamo la convergenza assoluta per dimostrare la convergenza di una serie complessa fondamentale:

$$(30) \quad \sum_{n=0}^{\infty} \frac{z^n}{n!} \quad z \in \mathbb{C}.$$

Vi ricorda qualcosa? La convergenza di questa serie discende direttamente dalla convergenza della corrispondente serie nei reali. Infatti, studiamo la convergenza assoluta e consideriamo la serie

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

Questa serie è convergente (la sua somma è $e^{|z|}$), quindi la serie (30) è convergente assolutamente e, di conseguenza, semplicemente.

Funzioni complesse. Una funzione complessa di variabile complessa è:

$$f : I \subset \mathbb{C} \rightarrow \mathbb{C}$$

dove I è il sottoinsieme di \mathbb{C} in cui è definita la funzione f .

Se, per $\zeta \in \mathbb{C}$, esiste $r > 0$ tale che l'intorno di ζ di raggio r , escluso il punto ζ stesso, è contenuto in I , allora si può dare senso alla nozione di limite di f per $z \rightarrow \zeta$ utilizzando la stessa definizione già data nel caso reale.

DEFINIZIONE 2.8. *La funzione f tende ad $\ell \in \mathbb{C}$ per $z \rightarrow \zeta \in I$, e si scrive $\lim_{z \rightarrow \zeta} f(z) = \ell$, se*

$$\forall \varepsilon > 0, \quad \exists \delta > 0 \quad \text{tale che} \quad |f(z) - \ell| < \varepsilon \quad \forall z \in I, \quad 0 < |z - \zeta| < \delta.$$

Se (come faremo sempre nel seguito) si suppone che per ogni $\zeta \in I$ esista $r > 0$ tale che l'intorno di ζ di raggio r sia interamente contenuto in I , in altre parole, se si suppone che ogni punto $\zeta \in I$ sia interno ad I , allora è possibile estendere la definizione di *continuità* al caso di funzioni complesse di variabile complessa.

DEFINIZIONE 2.9. *La funzione f è continua in $\zeta \in I$ se $\lim_{z \rightarrow \zeta} f(z) = f(\zeta)$, cioè se*

$$\forall \varepsilon > 0, \quad \exists \delta > 0 \quad \text{tale che} \quad |f(z) - f(\zeta)| < \varepsilon \quad \forall z \in I, \quad |z - \zeta| < \delta.$$

Né più né meno della definizione di continuità già vista nel caso reale, con l'unica accortezza che i moduli che compaiono qui sono moduli di numeri complessi.

Ad esempio, la funzione $f(z) = 2iz - 1$ è continua in ζ per ogni $\zeta \in \mathbb{C}$. Infatti

$$|f(z) - f(\zeta)| = |(2iz - 1) - (2i\zeta - 1)| = |2i||z - \zeta| = 2|z - \zeta|,$$

quindi $|f(z) - f(\zeta)| < \varepsilon$ per tutti valori z per cui $|z - \zeta| < \delta = \varepsilon/2$.

ESEMPIO 2.10. Consideriamo un altro esempio: $f(z) = \bar{z}$, cioè la funzione che associa ad un numero complesso $z = x + iy$, il suo complesso coniugato $x - iy$. Fissato $\zeta = \xi + i\eta$, se $|z - \zeta| \equiv \sqrt{(x - \xi)^2 + (y - \eta)^2} < \delta$,

$$\begin{aligned} |f(z) - f(\zeta)| &= |(x - iy) - (\xi - i\eta)| = |(x - \xi) - i(y - \eta)| \\ &= \sqrt{(x - \xi)^2 + (y - \eta)^2} < \delta. \end{aligned}$$

La condizione $|f(z) - f(\zeta)| < \varepsilon$ è soddisfatta per $\delta = \varepsilon$.

Quali funzioni complesse sono continue? Ragioniamo come già fatto in \mathbb{R} . Somma, prodotto di funzioni continue sono funzioni continue. Quindi, dal fatto che le costanti e la funzione z sono funzioni continue, possiamo dedurre che tutti i polinomi in \mathbb{C} sono funzioni continue. Analogamente per le funzioni razionali.

Infine è possibile definire la derivata di funzioni complesse (qui è fondamentale la possibilità di dividere per un numero complesso).

DEFINIZIONE 2.11. Una funzione f è derivabile in $\zeta \in I$ (o olomorfa in $\zeta \in I$) se esiste finito il limite

$$(31) \quad f'(\zeta) = \lim_{z \rightarrow \zeta} \frac{f(z) - f(\zeta)}{z - \zeta}.$$

Ad esempio, per calcolare la derivata della funzione $f(z) = z^n$ ($n \in \mathbb{N}$), basta riscrivere il rapporto incrementale nella forma

$$\frac{f(z) - f(\zeta)}{z - \zeta} = \frac{z^n - \zeta^n}{z - \zeta} = z^{n-1} + z^{n-2}\zeta + \dots + \zeta^{n-1},$$

e passare al limite per $z \rightarrow \zeta$, ottenendo

$$(z^n)' = nz^{n-1} \quad \forall n \in \mathbb{N}.$$

Anche in quest'ambito è vera l'implicazione:

$$\text{derivabilità} \quad \Rightarrow \quad \text{continuità}.$$

Perché? Basta ragionare come in \mathbb{R} :

$$\lim_{z \rightarrow \zeta} [f(z) - f(\zeta)] = \lim_{z \rightarrow \zeta} \frac{f(z) - f(\zeta)}{z - \zeta} \cdot (z - \zeta) = f'(\zeta) \cdot 0 = 0$$

Nel caso reale, a partire dall'operazione di derivazione abbiamo dedotto informazioni relative alla monotonia delle funzioni. E' possibile fare lo stesso nel caso delle funzioni

complesse? Assolutamente no! Dato che in \mathbb{C} non è definito l'ordinamento, non ha senso parlare di funzioni crescenti/decrescenti. Analogamente concetti come massimo, minimo, convessità e concavità sono privi di senso per funzioni $f : I \subset \mathbb{C} \rightarrow \mathbb{C}$.

3. L'esponenziale complesso

Una funzione particolarmente importante è la funzione esponenziale. Dato che la serie considerata nell'Esempio 2.7 è convergente per ogni $z \in \mathbb{C}$, è più che ragionevole utilizzarla per definire la funzione esponenziale anche nei complessi:

$$(32) \quad e^z := \sum_{n=0}^{\infty} \frac{z^n}{n!} \quad z \in \mathbb{C}.$$

Come si può immaginare/sperare, anche per (32), vale la proprietà

$$(33) \quad e^{w+z} = e^w e^z \quad \forall w, z \in \mathbb{C}.$$

La dimostrazione si basa sull'uso dei prodotti di serie. Niente dettagli in questa sede.

La funzione esponenziale è continua? È derivabile? L'uno e l'altro. Dimostriamo prima di tutto la continuità in 0. Dato che $e^0 = 1$, bisogna stimare $|e^z - 1|$:

$$|e^z - 1| = \left| \sum_{n=0}^{\infty} \frac{z^{n+1}}{(n+1)!} \right| = \left| z \sum_{n=0}^{\infty} \frac{z^n}{(n+1)!} \right| \leq |z| \sum_{n=0}^{\infty} \frac{|z|^n}{(n+1)!} \leq |z| \sum_{n=0}^{\infty} \frac{|z|^n}{n!} = |z|e^{|z|},$$

quindi, dato che $e^{|z|} \leq e$ per $|z| \leq 1$,

$$|e^z - 1| \leq e|z| \quad \forall |z| \leq 1,$$

da cui si deduce la continuità in $\zeta = 0$. Se $\zeta \in \mathbb{C}$, allora

$$\lim_{z \rightarrow \zeta} e^z = \lim_{z \rightarrow \zeta} e^{\zeta + (z - \zeta)} = e^\zeta \lim_{z \rightarrow \zeta} e^{z - \zeta} = e^\zeta \lim_{h \rightarrow 0} e^h = e^\zeta,$$

quindi la funzione è continua in ζ .

Anche per la derivabilità, consideriamo prima $\zeta = 0$: si ha

$$\frac{e^z - 1}{z} - 1 = \sum_{n=1}^{\infty} \frac{z^n}{(n+1)!} = z \sum_{n=0}^{\infty} \frac{z^n}{(n+2)!},$$

quindi, per $|z| \leq 1$,

$$\left| \frac{e^z - 1}{z} - 1 \right| \leq |z| \sum_{n=0}^{\infty} \frac{|z|^n}{(n+2)!} \leq |z| \sum_{n=0}^{\infty} \frac{|z|^n}{n!} = |z|e^{|z|} \leq e|z|,$$

che mostra che la derivata di e^z in 0 è 1.

Per ζ qualsiasi,

$$\lim_{z \rightarrow \zeta} \frac{e^z - e^\zeta}{z - \zeta} = \lim_{z \rightarrow \zeta} \frac{e^{\zeta + (z - \zeta)} - e^\zeta}{z - \zeta} = e^\zeta \lim_{z \rightarrow \zeta} \frac{e^{z - \zeta} - 1}{z - \zeta} = e^\zeta \lim_{h \rightarrow 0} \frac{e^h - 1}{h} = e^\zeta.$$

Quindi

$$(e^z)' = e^z \quad \forall z \in \mathbb{C}.$$

Fissato $\lambda \in \mathbb{C}$, per la funzione $f(z) = e^{\lambda z}$, vale

$$\frac{d(e^{\lambda z})}{dz} = e^{\lambda z} \frac{d(\lambda z)}{dz} = \lambda e^{\lambda z}.$$

Per derivare la funzione f , basta moltiplicare per λ la funzione stessa: in altri termini, $f(z) = e^{\lambda z}$ verifica l'equazione (differenziale) $f' = \lambda f$. Lo stesso è vero per le funzioni $A e^{\lambda z}$ con $A \in \mathbb{C}$.

Quale equazione è soddisfatta dalla derivata seconda di $f(z) = A e^{\lambda z}$? Facile:

$$f''(z) = \frac{d(A \lambda e^{\lambda z})}{dz} = A \lambda \frac{d(e^{\lambda z})}{dz} = A \lambda^2 e^{\lambda z} = \lambda^2 f,$$

quindi $f'' - \lambda^2 f = 0$.

Il fatto che le funzioni del tipo $A e^{\lambda z}$ soddisfino le semplicissime equazioni $f' - \lambda f = 0$ e $f'' - \lambda^2 f = 0$ fa di questa classe il mattone fondamentale per la costruzione delle soluzioni di equazioni differenziali lineari a coefficienti costanti, come vedremo tra un paio di fogli.

Formula di Eulero. Prima di concludere, deduciamo una formula estremamente interessante. Se calcoliamo l'esponenziale e^z in $z = ix \in i\mathbb{R}$, otteniamo

$$e^{ix} = \sum_{n=0}^{\infty} \frac{i^n x^n}{n!}.$$

Dato che

$$i^0 = i^4 = \dots = 1, \quad i^1 = i^5 = \dots = i, \quad i^2 = i^6 = \dots = -1, \quad i^3 = i^7 = \dots = -i,$$

possiamo scrivere

$$e^{ix} = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} + i \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!}.$$

Dalle formule di Taylor per le funzioni $\cos x$ e $\sin x$, segue la **formula di Eulero**:

$$(34) \quad e^{ix} = \cos x + i \sin x \quad \forall x \in \mathbb{R}.$$

Chi è dunque l'esponenziale del numero complesso $z = x + iy$? Una sintesi opportuna di esponenziale reale e funzioni trigonometriche: applicando (34),

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

In particolare $\operatorname{Re} e^z = e^x \cos y$ e $\operatorname{Im} e^z = e^x \sin y$.

La formula di Eulero, interessante di per sé, è estremamente comoda per rappresentare i numeri complessi quando si usino le coordinate polari (ρ, θ) :

$$z = \rho (\cos \theta + i \sin \theta) = \rho e^{i\theta}$$

Le formule per il prodotto e le potenze diventano, facendo uso della rappresentazione esponenziale, conseguenze immediate di (33). Ad esempio,

$$z_1 \cdot z_2 = (\rho_1 e^{i\theta_1}) \cdot (\rho_2 e^{i\theta_2}) = \rho_1 \rho_2 e^{i(\theta_1 + \theta_2)}.$$

Da (34) si deduce, in particolare, che l'esponenziale complesso non è una funzione iniettiva e quindi nemmeno invertibile. La definizione della funzione logaritmo nei numeri complessi richiede quindi una particolare attenzione. Qui non approfondiremo la questione. Anzi, non la toccheremo proprio.

CAPITOLO 4

Equazioni differenziali

Il problema più semplice che abbiamo affrontato e che rientra nella categoria delle equazioni differenziali è la determinazione di primitive

data $f(x)$ trovare tutte le funzioni $F(x)$ tali che $F'(x) = f(x)$.

In generale, un'equazione differenziale è un oggetto del tipo:

$$\mathcal{F}(x, y, y', y'', \dots) = 0,$$

cioè è una relazione che collega una funzione incognita y con la variabile indipendente x e con le derivate y', y'', \dots . Il caso della ricerca delle primitive corrisponde alla scelta $\mathcal{F} = y - F(x)$. Una funzione y , sufficientemente regolare, che verifica

$$\mathcal{F}(x, y(x), y'(x), y''(x), \dots) = 0,$$

per ogni scelta della variabile x , è detta **soluzione** (o **integrale**) dell'equazione.

In quel che segue, ci limitiamo a considerare alcuni tipi specifici di equazioni differenziali: le equazioni lineari. Il significato, in quest'ambito, del vocabolo “lineare” sarà più chiaro tra qualche pagina.

1. Equazioni lineari del I ordine a coefficiente costante

Equazioni omogenee. Per iniziare, consideriamo un'equazione differenziale particolarmente semplice

$$(35) \quad y' + ay = 0 \quad a \in \mathbb{R}.$$

L'equazione (35) è *del primo ordine, lineare, a coefficiente costanti, omogenea*:

- del *primo ordine* perché nell'equazione compare solo la derivata prima;
- *lineare* perché la dipendenza da y, y' è lineare;
- a *coefficiente costante* perché il coefficienti di a è indipendente da t ;
- *omogenea*, perché la funzione $y = 0$ è una soluzione dell'equazione.

OSSERVAZIONE 1.1. *Il termine “lineare”.* Supponiamo che y_1 e y_2 siano due soluzioni dell'equazione (35). Dati $\alpha, \beta \in \mathbb{R}$, consideriamo $w(t) := \alpha y_1(t) + \beta y_2(t)$, combinazione lineare di y_1 e y_2 . Allora:

$$w' + aw = \alpha y_1' + \beta y_2' + \alpha a y_1 + \beta a y_2 = \alpha(y_1' + a y_1) + \beta(y_2' + a y_2) = 0.$$

Quindi, dal fatto che l'equazione è lineare e omogenea discende la proprietà seguente: *ogni combinazione lineare di soluzioni di (35) è essa stessa una soluzione di (35).*

Passiamo al problema di trovare tutte le soluzioni dell'equazione (35). Dato che tale equazione è risolta da tutte le funzioni della forma Ae^{-at} con $A \in \mathbb{R}$, poniamo $y(t) = A(t)e^{-at}$ e cerchiamo tutte le funzioni A per cui y è soluzioni. Dato che $y' = (A' - aA)e^{-at}$, l'equazione (35) equivale a

$$(A' - aA)e^{-at} + aAe^{-at} = 0 \quad \iff \quad A' = 0.$$

Dato che la derivata di A è identicamente nulla, A è costante. Ricapitolando:

$$y'(t) + ay(t) = 0 \quad \forall t \in \mathbb{R} \quad \iff \quad y = Ae^{-at} \quad \text{per qualche } A \in \mathbb{R}.$$

Se la soluzione cercata deve soddisfare, oltre all'equazione differenziale (35), la condizione iniziale $y(t_0) = y_0$ per t_0, y_0 assegnati, si parla di **problema di Cauchy** per (35):

$$(36) \quad \text{trovare } y = y(t) \text{ tale che } \begin{cases} y' + ay = 0, \\ y(t_0) = y_0 \end{cases}$$

Determinare la soluzione è più che facile: dato che conosciamo tutte le soluzioni dell'equazione differenziale, basta determinare quali di queste verificano il vincolo imposto dalla *condizione iniziale* $y(t_0) = y_0$. Se $y(t) = Ae^{-at}$, sostituendo

$$y_0 = y(t_0) = Ae^{-at_0} \quad \Rightarrow \quad A = y(t_0)e^{at_0} = y_0e^{at_0},$$

quindi c'è un'unica soluzione per il problema di Cauchy (36) ed è $y(t) = y_0e^{-a(t-t_0)}$.

Equazioni non omogenee. Data una funzione f integrabile, consideriamo

$$(37) \quad y' + ay = f(t) \quad a \in \mathbb{R}, \quad f : \mathbb{R} \rightarrow \mathbb{R}.$$

L'equazione (37) è *del primo ordine, lineare, a coefficienti costanti, non omogenea*. La presenza del termine noto $f = f(t)$ fa in modo che $y = 0$ non sia soluzione (tranne nel caso $f \equiv 0$). Per cercare le soluzioni di questa equazione utilizziamo il metodo di *variazione delle costanti*: cerchiamo soluzioni della forma

$$y(t) = A(t)e^{-at}$$

dove $A = A(t)$ è una funzione da determinare. Dato che $y' = (A' - aA)e^{-at}$, sostituendo nell'equazione differenziale:

$$f = y' + ay = (A' - aA)e^{-at} + aAe^{-at} = A'e^{-at} \quad \iff \quad A'(t) = f(t)e^{at}.$$

Se la funzione f è integrabile in ogni intervallo di \mathbb{R} , la funzione A è:

$$A(t) = \int_0^t f(s)e^{as} ds + C \quad C \in \mathbb{R}$$

e quindi anche la funzione y è determinata:

$$y(t) = A(t)e^{-at} = Ce^{-at} + \int_0^t f(s)e^{a(s-t)} ds \quad C \in \mathbb{R}.$$

La soluzione dell'equazione non omogenea è quindi della forma

$$y(t) = y_0(t) + \bar{y}(t) \quad \text{dove} \quad \begin{cases} y_0(t) = Ce^{-at} & \text{soluzione generale di (35)} \\ \bar{y}(t) = \int_0^t f(s)e^{a(s-t)} ds & \text{soluzione particolare di (37)}. \end{cases}$$

L'aver scelto in \bar{y} come primo estremo di integrazione $s = 0$ è del tutto arbitrario, un qualsiasi altro numero fissato sarebbe andato altrettanto bene. Ad esempio, per trovare la soluzione del problema di Cauchy

$$(38) \quad \text{trovare } y = y(t) \text{ tale che} \quad \begin{cases} y' + ay = f(t), \\ y(t_0) = y_0 \end{cases}$$

è più comodo scegliere come primo estremo di integrazione t_0

$$y(t) = Ce^{-at} + \int_{t_0}^t f(s)e^{a(s-t)} ds \quad C \in \mathbb{R},$$

di modo che

$$y_0 = y(t_0) = Ce^{-at_0} + \int_{t_0}^{t_0} f(s)e^{a(s-t_0)} ds = Ce^{-at_0} \quad \Rightarrow \quad C = y_0 e^{at_0},$$

e la soluzione richiesta è

$$y(t) = y_0 e^{-a(t-t_0)} + \int_{t_0}^t f(s)e^{a(s-t)} ds.$$

Sistemi disaccoppiati di equazioni lineari. Il caso di un'unica funzione incognita è evidentemente il caso più semplice. In generale, è utile considerare situazioni in cui sono presenti più funzioni incognite: y_1, y_2, \dots . Risolvere sistemi nel caso generale è complicato e non è questa la sede in cui affrontare la questione. Esiste però una situazione facile in cui possibile ricondursi al caso studiato in precedenza: assegnate le costanti a, b, c , cerchiamo le soluzioni y_1, y_2 del sistema

$$\begin{cases} y_1' + ay_1 = 0, \\ y_2' + by_1 + cy_2 = 0. \end{cases}$$

La peculiarità di questo sistema è che la prima equazione $y_1' + ay_1 = 0$ è disaccoppiata dalla seconda, cioè è una equazione per la sola incognita y_1 (y_2 non compare). Qui si può risolvere la prima equazione e, successivamente, risolvere la seconda dopo aver inserito l'espressione esplicita per y_1 . Procediamo

$$y_1' + ay_1 = 0 \quad \Rightarrow \quad y_1 = Ae^{-at} \quad A \in \mathbb{R},$$

quindi l'equazione per y_2 diviene $y_2' + cy_2 = -Abe^{-at}$, cioè un'equazione non omogenea. Perciò, se $c \neq a$,

$$y_2(t) = Be^{-ct} - Ab \int_0^t e^{-as+c(s-t)} ds = Be^{-at} + \frac{Ab}{c-a} (e^{-at} - e^{-ct}) \quad A, B \in \mathbb{R}.$$

Invece, se $c = a$,

$$y_2(t) = Be^{-at} - Ab \int_0^t e^{-at} ds = (B - Abt) e^{-at} \quad A, B \in \mathbb{R}.$$

Si faccia attenzione alle costanti in gioco: a, b, c sono dati del problema, mentre A, B sono costanti arbitrarie (ogni scelta di A, B corrisponde ad una soluzione).

Equazioni lineari del I ordine a coefficiente variabile. Torniamo, ora, al caso di una singola equazione differenziale, e consideriamo il caso

$$(39) \quad y' + a(t)y = b(t) \quad t \in \mathbb{R},$$

dove a e b sono funzioni reali continue e la funzione incognita $y = y(x)$ è anch'essa una funzione reale. L'equazione differenziale (39) si dice *equazione differenziale lineare del primo ordine a coefficiente variabile*.

Come sempre, il problema di Cauchy consiste nel trovare la soluzione di

$$(40) \quad y' + a(t)y = b(t) \quad y(t_0) = y_0,$$

dove t_0 e y_0 sono assegnati.

Ecco come procedere per determinare l'integrale generale di (39). Indicata con A una primitiva della funzione a , moltiplichiamo l'equazione per la funzione $e^{A(t)}$, ottenendo

$$e^{A(t)} y'(t) + e^{A(t)} a(t)y(t) = e^{A(t)} b(t).$$

Il termine a primo membro è la derivata della funzione $e^{A(t)}y(t)$, infatti

$$(e^{A(t)}y(t))' = e^{A(t)}y'(t) + (e^{A(t)})'y(t) = e^{A(t)}y'(t) + e^{A(t)}a(t)y(t).$$

Quindi l'equazione prende la forma

$$(e^{A(t)}y(t))' = e^{A(t)}b(t).$$

Calcoliamo l'integrale indefinito del secondo membro

$$\int e^{A(t)}b(t) dt = F(t) + C.$$

Dato che la derivata di $e^{A(t)}y(t)$ è la funzione $e^{A(t)}b(t)$, deve valere

$$e^{A(t)}y(t) = F(t) + C.$$

Esplicitando rispetto a y , otteniamo la soluzione generale di (39):

$$(41) \quad y(t) = (F(t) + C)e^{-A(t)} \quad \text{dove} \quad \begin{cases} A(t) = \int a(t) dt, \\ F(t) = \int e^{A(t)} b(t) dt, \\ C \in \mathbb{R}. \end{cases}$$

ESERCIZIO 1.2. Risolvere l'equazione differenziale $y' + y \cos t = \frac{1}{2} \sin(2t)$.

Nel caso in cui si debba risolvere un problema di Cauchy, quindi se è assegnata anche la condizione iniziale $y(t_0) = y_0$, basta imporre che la generica soluzione data da (41) soddisfi la condizione data.

Una strada alternativa, che utilizza l'uso di integrali definiti (anziché indefiniti), segue lo stesso schema visto in precedenza:

– si calcola A , primitiva di a , e si moltiplica l'equazione per $e^{A(t)}$ in modo da ottenere

$$(y e^{A(t)})' = b(t) e^{A(t)}.$$

– si integra in (t_0, t) e, per il teorema fondamentale del calcolo integrale,

$$y(t) e^{A(t)} - y(t_0) e^{A(t_0)} = \int_{t_0}^t b(t) e^{A(t)} dt,$$

– si usa la condizione $y(t_0) = y_0$ e si esplicita $y(t)$,

$$y(t) = y_0 e^{A(t_0) - A(t)} + e^{-A(t)} \int_{t_0}^t b(t) e^{A(t)} dt,$$

Se $A(t)$ è scelta come $\int_{t_0}^t a(s) ds$, la formula può essere riscritta come

$$y(t) = y_0 \exp\left(\int_{t_0}^t a(s) ds\right) + \int_{t_0}^t b(t) \exp\left(-\int_t^t a(s) ds\right) dt.$$

ESERCIZIO 1.3. Risolvere il problema di Cauchy

$$ty' - y + t \ln t = 0 \quad y(1) = 1.$$

2. Equazioni lineari del II ordine a coefficienti costanti omogenee

Consideriamo una particella, cioè un punto in cui supponiamo concentrata la massa m , e supponiamo che si muova lungo una retta. La posizione della particella al tempo t è individuata da una funzione $y = y(t)$, che ne rappresenta la coordinata rispetto ad un sistema di riferimento fissato. La funzione $y'(t)$ è la *velocità* della particella, mentre $y''(t)$ ne è l'*accelerazione*.

In meccanica, il moto della particella è causato dalla presenza di forze, descritte qui da una funzione F , che agiscono sulla particella stessa. La *legge di Newton* afferma

che: la massa m moltiplicata per l'accelerazione è uguale alla forza risultante che agisce sulla particella:

$$(42) \quad my'' = F.$$

Modellizzare l'azione di una certa forza F sulla particella corrisponde ad esprimere F in termini della posizione y , della velocità y' , dell'istante t . Di conseguenza, la legge di Newton prende la forma di una equazione differenziale

$$(43) \quad my'' = F(t, y, y').$$

Ad esempio, consideriamo una particella che si sposti lungo l'asse y , su cui agisca una forza elastica F_e , diretta verso l'origine e proporzionale alla distanza dall'origine. In altre parole, supponiamo $F_e = -ky$ con $k > 0$ (k è il coefficiente di proporzionalità e misura la grandezza delle forza elastica). La legge di Newton dice che il moto della particella è descritto dall'equazione

$$(44) \quad my'' = -ky \quad (\text{equazione dell'oscillatore armonico})$$

Se sulla particella agisce anche la forza d'attrito F_a , che supponiamo proporzionale alla velocità y' , ma diretta in senso contrario, cioè $F_a = -ry'$ con $r \geq 0$, otteniamo l'equazione

$$(45) \quad my'' = -ky - ry' \quad (\text{equazione dell'oscillatore con attrito}).$$

Questo è l'esempio tipo di equazione di cui vogliamo determinare le soluzioni.

Studiamo, perciò, l'equazione differenziale *del secondo ordine, lineare, a coefficienti costanti, omogenea*:

$$(46) \quad y'' + 2by' + cy = 0 \quad b, c \in \mathbb{R}.$$

L'equazione è:

- del *secondo ordine* perché compaiono derivate fino al secondo ordine;
- *lineare* perché la dipendenza da y , y' e y'' è lineare;
- a *coefficienti costanti* perché i coefficienti di b e c non dipendono dalla variabile t ;
- *omogenea*, perché la funzione $y = 0$ è una soluzione dell'equazione.

Un esempio di equazione di questo genere è $y'' + y = 0$. Per ogni $A, B \in \mathbb{R}$, la funzione $y = A \cos t + B \sin t$ ne è soluzione (verificatelo!). Questo fatto è generale: date y_1 e y_2 soluzioni di (46), *tutte le combinazioni lineari di y_1 e y_2 sono soluzioni della stessa equazione*.

DEFINIZIONE 2.1. *Due funzioni y_1 e y_2 sono linearmente dipendenti se una è proporzionale all'altra, cioè se esistono due costanti α, β , non entrambe nulle, tali che*

$\alpha y_1(t) + \beta y_2(t) = 0$ per ogni t . Si dicono **linearmente indipendenti** se non sono **linearmente dipendenti**.

ESEMPIO 2.2. Ad esempio, $y_1 = \sin t$ e $y_2 = \cos t$ sono soluzioni linearmente indipendenti di $y'' + y = 0$. Infatti, se, per qualche $\alpha, \beta \in \mathbb{R}$, si ha $\alpha \sin t + \beta \cos t = 0$ per ogni t , allora

$$\begin{aligned} t = 0 & \Rightarrow \alpha \sin 0 + \beta \cos 0 = \beta = 0, \\ t = \frac{\pi}{2} & \Rightarrow \alpha \sin \frac{\pi}{2} + \beta \cos \frac{\pi}{2} = \alpha = 0. \end{aligned}$$

Quindi $\alpha = \beta = 0$.

Il risultato che segue (di cui omettiamo la dimostrazione) descrive la struttura dell'insieme delle soluzioni di (46).

TEOREMA 2.3. Se y_1 e y_2 sono soluzioni linearmente indipendenti di (46), tutte le soluzioni di (46) sono della forma

$$y(t) = c_1 y_1(t) + c_2 y_2(t) \quad c_1, c_2 \in \mathbb{R}.$$

Quindi, per conoscere tutte le soluzioni dell'equazione differenziale lineare (46) basta individuare una coppia di soluzioni linearmente indipendenti. Tutte le altre soluzioni si ottengono per combinazione lineare. Ad esempio, dato che $\sin t$ e $\cos t$ sono linearmente indipendenti, tutte le soluzioni di $y'' + y = 0$ sono del tipo $A \sin t + B \cos t$ con $A, B \in \mathbb{R}$.

Costruzione delle soluzioni linearmente indipendenti. L'obiettivo ora è determinare una coppia di soluzioni linearmente indipendenti dell'equazione (46). A questo scopo procediamo in una maniera che può sembrare strana, ma molto efficace:

- consideriamo la stessa equazione differenziale per funzioni a valori complessi, cioè cerchiamo $z = z(t) \in \mathbb{C}$ tali che $z'' + 2bz' + cz = 0$;
- determiniamo due soluzioni complesse z_1 e z_2 opportune;
- scegliamo $y_1 = \operatorname{Re} z_1$ e $y_2 = \operatorname{Re} z_2$.

Consideriamo quindi l'equazione

$$(47) \quad z'' + 2bz' + cz = 0 \quad b, c \in \mathbb{R},$$

dove $z : \mathbb{R} \rightarrow \mathbb{C}$ e cerchiamo una soluzione della forma $z = e^{\lambda t}$ dove $\lambda \in \mathbb{C}$ è da determinare. Dato che $z' = \lambda e^{\lambda t}$ e $z'' = \lambda^2 e^{\lambda t}$, λ deve essere tale che

$$\lambda^2 e^{\lambda t} + 2b\lambda e^{\lambda t} + ce^{\lambda t} = 0 \quad \forall t \in \mathbb{R}.$$

Dato che $e^{\lambda t} \neq 0$, λ deve essere soluzione dell'equazione di secondo grado complessa

$$(48) \quad \text{polinomio caratteristico:} \quad \lambda^2 + 2b\lambda + c = 0.$$

Studiamo tre casi a seconda del tipo di radici del polinomio caratteristico.

Caso I. $b^2 > c$. Il polinomio caratteristico (48) ha due radici reali

$$\lambda_1 = -b + \sqrt{b^2 - c}, \quad \lambda_2 = -b - \sqrt{b^2 - c}.$$

Otteniamo in corrispondenza due soluzioni

$$z_1(t) = e^{\lambda_1 t} \quad z_2(t) = e^{\lambda_2 t}.$$

Prendendo la parte reale otteniamo le soluzioni y_1 e y_2 cercate, ma... z_1 e z_2 sono funzioni a valori reali! Quindi in realtà y_1 e y_2 coincidono con z_1 e z_2 , rispettivamente:

$$y_1(t) = \operatorname{Re} z_1(t) = e^{\lambda_1 t} \quad y_2(t) = \operatorname{Re} z_2(t) = e^{\lambda_2 t}.$$

Per verificare che y_1, y_2 sono indipendenti, supponiamo che esistano $\alpha, \beta \in \mathbb{R}$ tali che

$$\alpha e^{\lambda_1 t} + \beta e^{\lambda_2 t} = 0.$$

Derivando e calcolando entrambe le relazioni per $t = 0$, si ottiene un sistema lineare omogeneo per le incognite α, β

$$\alpha + \beta = 0, \quad \alpha \lambda_1 + \beta \lambda_2 = 0.$$

Dato che il determinante di questo sistema è $\lambda_2 - \lambda_1 = 2\sqrt{b^2 - c} > 0$, l'unica soluzione è $\alpha = \beta = 0$.

In conclusione, nel caso $b^2 > c$, tutte le soluzioni di (46) sono della forma

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}.$$

ESERCIZIO 2.4. *Determinare le soluzioni dell'equazione $y'' - 3y' + 2y = 0$.*

Caso II. $b^2 = c$. In questo caso il polinomio caratteristico è il quadrato di un binomio

$$\lambda^2 + 2b\lambda + b^2 = (\lambda + b)^2,$$

ed ha un'unica soluzione $\lambda = -b$. In corrispondenza troviamo un'unica soluzione $z_1(t) = e^{-bt}$ dell'equazione differenziale. Per determinare una seconda soluzione (indipendente da z_1), consideriamo $z_2(t) = te^{-bt}$. Dato che

$$z_2'(t) = (1 - bt)e^{-bt} \quad z_2''(t) = (-2b + b^2 t)e^{-bt},$$

si ha

$$z_2''(t) + 2bz_2'(t) + b^2 z_2(t) = [(-2b + b^2 t) + 2b(1 - bt) + b^2 t]e^{-bt} = 0.$$

Quindi anche z_2 è soluzione dell'equazione.

OSSERVAZIONE 2.5. La scelta di z_2 è suggerita da un procedimento di limite, immaginando il caso di due radici coincidenti come limite di un caso con due radici distinte. Se $\lambda_1 \neq \lambda_2$, l'espressione $\frac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1}$, è soluzione di (47). Per $\lambda_2 \rightarrow \lambda_1$ otteniamo

$$\lim_{\lambda_2 \rightarrow \lambda_1} \frac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 - \lambda_1} = \lim_{\lambda_2 \rightarrow \lambda_1} \frac{e^{\lambda_2 t} - e^{\lambda_1 t}}{\lambda_2 t - \lambda_1 t} \cdot t = te^{\lambda_1 t},$$

che è la soluzione determinata in precedenza.

Come nel caso precedente, le due soluzioni complesse $z_1(t) = e^{-bt}$ e $z_2(t) = te^{-bt}$ sono a valori reali. Inoltre si tratta di soluzioni indipendenti. Infatti, siano $\alpha, \beta \in \mathbb{R}$ tali che $\alpha e^{-bt} + \beta te^{-bt} = 0$, per ogni t . Dividendo per e^{-bt} , si deduce $\alpha + \beta t = 0$, cioè $\alpha = \beta = 0$.

In conclusione, tutte le soluzioni di (46) nel caso $b^2 = c$, sono del tipo

$$y(t) = c_1 e^{-bt} + c_2 t e^{-bt} = (c_1 + c_2 t) e^{-bt} \quad c_1, c_2 \in \mathbb{R}.$$

ESERCIZIO 2.6. *Determinare tutte le soluzioni dell'equazione $y'' + 4y' + 4y = 0$.*

Caso III. $b^2 < c$. Qui, finalmente!, è necessario l'uso dei complessi. Se $b^2 < c$, il polinomio caratteristico $\lambda^2 + 2b\lambda + c$ ha due radici complesse coniugate

$$\lambda^2 + 2b\lambda + c = 0 \quad \iff \quad \lambda \in \{\lambda_1, \lambda_2\} \quad \text{dove} \quad \begin{cases} \lambda_1 = -b + i\sqrt{c - b^2} \\ \lambda_2 = -b - i\sqrt{c - b^2} \end{cases}$$

Nel seguito, indichiamo con $\nu = \sqrt{c - b^2} > 0$. L'equazione differenziale complessa (47) ha due soluzioni indipendenti date da

$$\begin{aligned} z_1(t) &= e^{\lambda_1 t} = e^{-bt} \left\{ \cos(\nu t) + i \sin(\nu t) \right\}, \\ z_2(t) &= e^{\lambda_2 t} = e^{-bt} \left\{ \cos(\nu t) - i \sin(\nu t) \right\} \end{aligned}$$

La parte reale di queste due (distinte) soluzioni è però la stessa: $e^{-bt} \cos(\nu t)$. Occorre quindi sostituire la soluzione z_2 con un'altra soluzione \tilde{z}_2 diversa. Dato che $e^{\lambda_1 t}$ è soluzione di (47), anche tutte le funzioni della forma $Ae^{\lambda_1 t}$ sono soluzioni della stessa equazione per ogni scelta di $A \in \mathbb{C}$. Scegliamo $A = -i$, ossia poniamo $\tilde{z}_2(t) = -i e^{\lambda_1 t} = -i z_1(t)$, e scegliamo come soluzioni dell'equazione (47) le funzioni

$$z_1(t) = e^{-bt} \left\{ \cos(\nu t) + i \sin(\nu t) \right\}, \quad \tilde{z}_2(t) = e^{-bt} \left\{ \sin(\nu t) - i \cos(\nu t) \right\}.$$

Le soluzioni linearmente indipendenti di (46) sono date dalle parti reali di z_1 e \tilde{z}_2 :

$$\begin{aligned} y_1(t) &= \operatorname{Re} z_1(t) = e^{-bt} \cos(\nu t), \\ y_2(t) &= \operatorname{Re} \tilde{z}_2(t) = e^{-bt} \sin(\nu t). \end{aligned}$$

Quindi, nel caso $b^2 < c$, tutte le soluzioni di (46) sono della forma

$$y(t) = e^{-bt} \left(c_1 \cos(\nu t) + c_2 \sin(\nu t) \right) \quad \nu := \sqrt{c - b^2}.$$

Dato che $-b = \operatorname{Re} \lambda_1$ e $\nu = \operatorname{Im} \lambda_1$, se preferite la le soluzioni si possono scrivere così:

$$y(t) = e^{(\operatorname{Re} \lambda) t} \left(c_1 \cos[(\operatorname{Im} \lambda) t] + c_2 \sin[(\operatorname{Im} \lambda) t] \right) \quad \text{con } \lambda \text{ tale che } \lambda^2 + 2b\lambda + c = 0$$

ESERCIZIO 2.7. *Dimostrare che, se $\nu \neq 0$, le funzioni $e^{-bt} \cos(\nu t)$ e $e^{-bt} \sin(\nu t)$ sono linearmente indipendenti.*

ESERCIZIO 2.8. *Determinare tutte le soluzioni di $y'' + y' + y = 0$.*

Analizziamo le soluzioni nel caso $b, c \geq 0$. Nel modello di oscillatore con attrito, b e c sono positivi dato che rappresentano

$$b = \frac{r}{2m} \quad c = \frac{k}{m},$$

dove m è la massa, r misura l'attrito e k misura la forza elastica.

Nei Casi I e II, in cui $b^2 \geq c$, ossia $r^2 \geq 4km$ la soluzione è data o dalla sovrapposizione di due funzioni esponenziali

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} \quad \text{con} \quad \begin{cases} \lambda_1 = -b + \sqrt{b^2 - c} < 0, \\ \lambda_2 = -b - \sqrt{b^2 - c} < 0. \end{cases}$$

oppure dal prodotto di una funzione lineare con un'esponenziale

$$y(t) = (c_1 + c_2 t) e^{-bt}.$$

In entrambi i casi, la soluzione tende asintoticamente a zero per $t \rightarrow +\infty$:

$$\lim_{t \rightarrow +\infty} y(t) = 0.$$

Fisicamente questo corrisponde al fatto che l'effetto di attrito è talmente forte da impedire alla forza elastica di generare un moto oscillatorio.

Al contrario, nel Caso III, $b^2 < c$, cioè per $r^2 < 4km$, le soluzioni sono

$$y(t) = e^{-bt} (c_1 \cos(\nu t) + c_2 \sin(\nu t)) \quad \nu = \sqrt{c - b^2}.$$

Se $b > 0$ (presenza di attrito), le soluzioni sono il prodotto di un termine oscillatorio e di un termine esponenziale che regola l'ampiezza delle oscillazioni e che tende a 0 per $t \rightarrow +\infty$. In questo caso si parla di *oscillazioni armoniche smorzate*. Il valore $\nu = \sqrt{c - b^2}$ è la *frequenza* delle oscillazioni.

Nel caso in cui $b = 0$ (assenza di attrito), le soluzioni sono date da una combinazione lineare di $\sin(\sqrt{c} t)$ e $\cos(\sqrt{c} t)$, e quindi danno luogo a oscillazioni periodiche *non smorzate*.

Problema di Cauchy. Per individuare un'unica soluzione occorre fissare le *costanti di integrazione* c_1 e c_2 . Questo fatto è sensato rispetto al problema fisico: l'evoluzione del moto dipende dalle *condizioni iniziali*, cioè da posizione e velocità ad un istante t_0 . Il **problema di Cauchy** per l'equazione (46) è il problema di determinare la soluzione di

$$(49) \quad \begin{cases} y'' + 2by' + cy = 0, \\ y(t_0) = y_0, \quad y'(t_0) = y_1. \end{cases}$$

dove $t_0, y_0, y_1 \in \mathbb{R}$ sono costanti assegnate. Il problema è quindi di trovare, tra tutte le soluzioni dell'equazione $y'' + 2by' + cy = 0$, l'unica che soddisfi le condizioni iniziali $y(t_0) = y_0$ e $y'(t_0) = y_1$.

ESEMPIO 2.9. Cerchiamo la soluzione del problema di Cauchy

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

Dato che le soluzioni di $y'' + y = 0$ sono tutte e sole combinazioni lineari di $\sin t$ e $\cos t$, consideriamo una generica soluzione e imponiamo che soddisfi le condizioni iniziali. In questo modo otteniamo un sistema lineare per le costanti c_1, c_2 :

$$y(x) = c_1 \cos t + c_2 \sin t, \quad y(0) = 1, \quad y'(0) = 1,$$

implica

$$\begin{cases} 1 \cdot c_1 + 0 \cdot c_2 = 1 \\ 0 \cdot c_1 + 1 \cdot c_2 = 1. \end{cases}$$

Quindi $c_1 = c_2 = 1$ e l'unica soluzione del problema è $y(t) = \cos t + \sin t$.

In generale, se si sono determinate due soluzioni linearmente indipendenti y_1 e y_2 di (46), è sempre possibile risolvere il problema di Cauchy imponendo le condizioni iniziali date in (49): sia

$$y(x) = c_1 y_1(t) + c_2 y_2(t), \quad y(t_0) = y_0, \quad y'(t_0) = y_1,$$

allora

$$\begin{cases} y_1(0) c_1 + y_2(0) c_2 = y_0 \\ y_1'(0) c_1 + y_2'(0) c_2 = y_1. \end{cases}$$

Il fatto che le soluzioni siano linearmente indipendenti garantisce che il determinante del sistema sia sempre non nullo e che il sistema ammetta un'unica soluzione.

ESERCIZIO 2.10. *Determinare le soluzioni dei seguenti problemi di Cauchy*

$$\begin{cases} y'' - y' - 2y = 0 \\ y(0) = 1, \quad y'(0) = -2, \end{cases} \quad \begin{cases} y'' + y' - 6y = 0 \\ y(0) = -1, \quad y'(0) = 3, \end{cases}$$

3. Equazioni lineari del II ordine a coefficienti costanti non omogenee

Torniamo all'esempio dell'oscillatore con attrito. Se supponiamo che sulla particella agisca anche una forza esterna f che sia funzione solo del tempo, cioè $f = f(t)$, l'equazione del moto diviene $my'' + ry' + ky = f(t)$. Nel caso in cui la funzione f sia nulla, si parla di *moto libero*, se invece $f \neq 0$, si parla di *moto forzato*.

Prendiamo il toro per le corna e consideriamo l'equazione

$$(50) \quad y'' + 2by' + cy = f(t) \quad b, c \in \mathbb{R},$$

dove f è una funzione assegnata. L'equazione differenziale (50) è del secondo ordine, lineare, a coefficienti costanti, *non omogenea*.

Se w e v sono due soluzioni, allora la funzione differenza $u := w - v$ soddisfa l'equazione omogenea (46), come si può verificare con una semplice sostituzione. Viceversa, se u è una soluzione dell'equazione omogenea (46) e v dell'equazione non omogenea (50), la somma $w := u + v$ è soluzione dell'equazione (50). Quindi *la soluzione generale dell'equazione non omogenea (50) è*

$$y(t) = y_0(t) + \bar{y}(t) \quad \text{dove} \quad \begin{cases} y_0 & \text{soluzione generale di (46)} \\ \bar{y} & \text{soluzione particolare di (50)}. \end{cases}$$

Dato che abbiamo già una strategia generale per trovare tutte le soluzioni dell'equazione omogenea (46), per risolvere il problema nel caso non omogeneo, basta trovare una *singola* soluzione di (50).

Principio di sovrapposizione. Supponiamo di decomporre la funzione f come somma di due funzioni f_1 e f_2 , cioè supponiamo che sia

$$f(t) = f_1(t) + f_2(t).$$

Se y_1 è una soluzione di $y'' + 2by' + cy = f_1(t)$ e y_2 è una soluzione di $y'' + 2by' + cy = f_2(t)$, allora la funzione somma $y(t) := y_1(t) + y_2(t)$ è soluzione di (50). Questo fatto è noto come **principio di sovrapposizione**. Sostanzialmente significa che se abbiamo una forza “complicata” che, però, può essere decomposta come somma di singole parti “elementari”, possiamo ottenere l'evoluzione del fenomeno completo sommando le evoluzioni relative ad ogni singola parte. In termini matematici possiamo dire che la soluzione generale di

$$y'' + 2by' + cy = f_1(t) + \cdots + f_n(t)$$

è data da

$$y(t) = y_0(t) + \bar{y}_1(t) + \cdots + \bar{y}_n(t)$$

dove

$$y_0 \quad \text{soluzione generale di} \quad y'' + 2by' + cy = 0$$

$$\bar{y}_i \quad \text{soluzione particolare di} \quad y'' + 2by' + cy = f_i(t).$$

L'obiettivo è quindi quello di determinare una soluzione particolare di (50) nel caso di f particolarmente semplici e poi determinare la soluzione particolare per funzioni più complicate che si decompongano come somma di funzioni "semplici".

Forzanti periodiche. $f(t) = Ce^{i\omega t}$ Il caso più importante è quello di forzante periodica, cioè del tipo

$$C \cos(\omega t) \quad \text{o} \quad C \sin(\omega t) \quad (C, \omega \in \mathbb{R}).$$

Invece di lavorare con queste funzioni trigonometriche, è estremamente più semplice ed elegante lavorare nell'ambito complesso. Poniamo $f(t) = Ce^{i\omega t}$ e studiamo

$$(51) \quad z'' + 2bz' + cz = Ce^{i\omega t} \quad C, \omega \in \mathbb{R},$$

in cui l'incognita è la funzione $z = z(t)$ a valori in \mathbb{C} . Prima di risolvere questa equazione, poniamoci il problema del "ritorno": *una volta trovata una soluzione z a valori complessi, come determiniamo la soluzione reale y che cerchiamo?* Indichiamo la funzione z , soluzione di (51), nella forma

$$z(t) = y_1(t) + iy_2(t),$$

dove $y_1 = \operatorname{Re} z$ e $y_2 = \operatorname{Im} z$. Inserendo in (51), abbiamo

$$y_1'' + 2by_1' + cy_1 + i(y_2'' + 2by_2' + cy_2) = C \cos(\omega t) + iC \sin(\omega t).$$

Quindi, uguagliando la parte reale e la parte immaginaria,

$$y_1'' + 2by_1' + cy_1 = C \cos(\omega t)$$

$$y_2'' + 2by_2' + cy_2 = C \sin(\omega t),$$

cioè y_1 , parte reale di z , verifica l'equazione con termine forzante $C \cos(\omega t)$ e y_2 , parte immaginaria di z , con termine forzante $C \sin(\omega t)$. A seconda della funzione f che stiamo considerando, sceglieremo la parte reale o la parte immaginaria della soluzione complessa.

Con la tranquillità di poter tornare nell'ambito reale, riprendiamo lo studio dell'equazione (51) con l'obiettivo di trovarne una soluzione particolare. È sensato immaginare che esista una soluzione con la stessa frequenza di oscillazione di f :

$$z(t) = \sigma e^{i\omega t} \quad \sigma \in \mathbb{C}.$$

Il numero complesso σ è da determinare imponendo la condizione che z sia soluzione dell'equazione cercata.

ESEMPIO 3.1. Partiamo da un problema specifico: trovare una soluzione di

$$y'' - y = \cos t.$$

Dato che $\cos t = \operatorname{Re} e^{it}$, cerchiamo $\sigma \in \mathbb{C}$ tale che $z(t) = \sigma e^{it}$ sia soluzione di $z'' - z = e^{it}$. Dato che $z'' = -\sigma e^{it}$, sostituendo, otteniamo $(-\sigma - \sigma)e^{it} = e^{it}$ da cui $\sigma = -\frac{1}{2}$. Quindi la soluzione complessa z è $-\frac{1}{2}e^{it}$. La soluzione particolare \bar{y} dell'equazione richiesta è data dalla parte reale di z , cioè

$$\bar{y}(t) = \operatorname{Re} z(t) = -\frac{1}{2} \cos t.$$

Nel caso generale, dato che $z'(t) = i\omega\sigma e^{i\omega t}$ e $z''(t) = -\omega^2\sigma e^{i\omega t}$, sostituendo ed eliminando il termine $e^{i\omega t}$, otteniamo

$$-\omega^2\sigma + 2bi\omega\sigma + c\sigma = C \quad \Rightarrow \quad (-\omega^2 + 2bi\omega + c)\sigma = C.$$

Quindi, se $c - \omega^2 + 2bi\omega \neq 0$,

$$\sigma = \frac{C}{c - \omega^2 + 2bi\omega} = C \frac{c - \omega^2 - 2bi\omega}{(c - \omega^2)^2 + 4b^2\omega^2}.$$

Questo valore di σ può essere riscritto nella forma esponenziale

$$\sigma = C\alpha e^{-i\omega\delta} \quad \text{dove} \quad \begin{cases} \alpha = |\sigma| = \frac{1}{\sqrt{(c - \omega^2)^2 + 4b^2\omega^2}} \\ \cos(\omega\delta) = (c - \omega^2)\alpha, \quad \sin(\omega\delta) = -2b\omega\alpha \end{cases}$$

La soluzione complessa z è quindi

$$z(t) = C\alpha e^{i\omega(t-\delta)}.$$

di cui scegliere la parte reale o la parte immaginaria a seconda della funzione f data.

ESERCIZIO 3.2. Determinare una soluzione particolare di $y'' + 2y' + 2y = 2 \sin(2t)$.

Soluzione. Sostituendo $z(t) = \sigma e^{2it}$ in $z'' + 2z' + 2z = 2e^{2it}$, segue $-4\sigma + 4i\sigma + 2\sigma = 2$ da cui $\sigma = \frac{1}{2i-1} = -\frac{1+2i}{5}$. Quindi

$$z(t) = -\frac{1+2i}{5} [\cos(2t) + i \sin(2t)]$$

La soluzione particolare è data dalla parte immaginaria di z

$$y(t) = \operatorname{Im} z(t) = -\frac{1}{5} [2 \cos(2t) + \sin(2t)].$$

Attenzione! Il metodo appena esposto funziona se $c - \omega^2 + 2bi\omega \neq 0$. Rimane quindi da capire il caso in cui

$$c = \omega^2, \quad b\omega = 0.$$

Per cominciare, supponiamo $\omega \neq 0$. La seconda condizione implica $b = 0$ e quindi

$$z'' + cz = Ce^{i\omega t} \quad C, \omega \in \mathbb{R}.$$

Cercando soluzioni $z(t) = \sigma e^{i\omega t}$ con $\sigma \in \mathbb{C}$, otteniamo l'equazione $(c - \omega^2)\sigma = C$. Se $\omega^2 = c$ la relazione precedente dice che non esiste nessuna soluzione della forma $z(t) = \sigma e^{i\omega t}$! Non ci perdiamo d'animo e, in analogia con quanto visto per le equazioni omogenee, cerchiamo una soluzione della forma

$$z(t) = \sigma t e^{i\omega t} \quad \sigma \in \mathbb{C}.$$

Dato che $z'(t) = \sigma(1 + i\omega t)e^{i\omega t}$ e $z''(t) = \sigma(2i\omega - \omega^2 t)e^{i\omega t}$, sostituendo e semplificando il termine $e^{i\omega t}$,

$$2\sigma i\omega = C \quad \implies \quad \sigma = \frac{C}{2i\omega} = -C \frac{i}{2\omega}$$

avendo tenuto conto di $\omega^2 = c$.

ESERCIZIO 3.3. Determinare una soluzione particolare di $y'' + y = \sin t$.

Soluzione. Studiamo l'equazione complessa $z'' + z = e^{it}$ e cerchiamone soluzioni della forma $z(t) = \sigma t e^{it}$. Sostituendo nell'equazione $z'(t) = \sigma(1 + it)e^{it}$ e $z''(t) = \sigma(2i - t)e^{it}$,

$$\sigma(2i - t)e^{it} + \sigma t e^{it} = e^{it} \quad \implies \quad 2\sigma i = 1.$$

Quindi la soluzione complessa è

$$z(t) = -\frac{it}{2} e^{it} = -\frac{it}{2}(\cos t + i \sin t) = \frac{t}{2} \sin t - i \frac{t}{2} \cos t.$$

La soluzione desiderata è la parte immaginaria di z : $y(t) = \text{Im } z(t) = -\frac{1}{2} t \cos t$.

Resta escluso un solo caso: quello in cui $c - \omega^2 + 2bi\omega = 0$ e $\omega = 0$. Allora $c = 0$ e l'equazione diventa $y'' + 2by' = C$. Ecco le soluzioni cercate in questi casi

$$b \neq 0 : \quad \bar{y} = \frac{Ct}{2b} \quad \text{e} \quad b = 0 : \quad \bar{y} = \frac{Ct^2}{2}.$$

Verificate!

Risonanza. Se $c - \omega^2 + 2bi\omega \neq 0$, una soluzione particolare dell'equazione complessa è

$$z(t) = C\alpha e^{i\omega(t-\delta)} \quad \text{dove} \quad \begin{cases} \alpha = |\sigma| = \frac{1}{\sqrt{(c - \omega^2)^2 + 4b^2\omega^2}} \\ \cos(\omega\delta) = (c - \omega^2)\alpha, \quad \sin(\omega\delta) = -2b\omega\alpha \end{cases}$$

Il fattore α corrisponde ad un fattore di "distorsione" e il fattore $\omega\delta$ ad un fattore di "spostamento di fase" rispetto al segnale esterno originale $Ce^{i\omega t}$.

Il fattore α descrive quanto il "segnale esterno" $Ce^{i\omega t}$ viene amplificato dal sistema. Immaginiamo b, c fissati, e supponiamo di far variare la frequenza ω del segnale. Il grafico di

$$(52) \quad \phi(\omega) = \frac{1}{\sqrt{(c - \omega^2)^2 + 4b^2\omega^2}}$$

è detto *curva di risonanza*.

Nel caso $b, c \neq 0$, ϕ è definita dappertutto, pari e tende a zero per $\omega \rightarrow \pm\infty$ con ordine 2:

$$\lim_{\omega \rightarrow \pm\infty} \frac{1/\sqrt{(c-\omega^2)^2 + 4b^2\omega^2}}{1/\omega^2} = \lim_{\omega \rightarrow \pm\infty} \left[\left(1 - \frac{c}{\omega^2}\right)^2 + \frac{4b^2}{\omega^2} \right]^{-1/2} = 1.$$

La derivata prima di ϕ è

$$\phi'(\omega) = -\frac{4\omega[\omega^2 - (c - 2b^2)]}{[(c - \omega^2)^2 + 4b^2\omega^2]^{3/2}}.$$

– Se $2b^2 \geq c$, la funzione ha un andamento particolarmente semplice: decresce per $\omega > 0$ e cresce per $\omega < 0$, il suo valore massimo è assunto in $\omega = 0$

$$\max_{\omega \in \mathbb{R}} \phi(\omega) = \phi(0) = \frac{1}{|c|}.$$

– Se $2b^2 < c$, la funzione ha due punti di massimo simmetrici nei valori

$$\omega = \pm\sqrt{c - 2b^2},$$

il massimo della funzione è dato da

$$\max_{\omega \in \mathbb{R}} \phi(\omega) = \phi(\pm\sqrt{c - 2b^2}) = \frac{1}{2|b|\sqrt{c - b^2}}.$$

Il fatto significativo è che questo valore massimo tende a $+\infty$ per $b \rightarrow 0$, cioè se l'effetto dell'attrito è piccolo, per una scelta di ω particolare (vicino a $\sqrt{c - b^2}$), il sistema determinato dall'equazione differenziale risponde al segnale esterno $e^{i\omega t}$ con una distorsione estremamente grande. Questo fenomeno è detto *fenomeno della risonanza* e appare in natura in moltissime situazioni differenti.

CAPITOLO 5

Lo spazio reale multidimensionale

Lo spazio reale d -dimensionale, indicato con \mathbb{R}^d (dove $d \in \mathbb{N}$), è l'insieme costituito da d -ple ordinate, dette **punti** (o **vettori**), e indicate, nel seguito, con la notazione

$$P = (x_1, \dots, x_d)$$

dove x_1, \dots, x_d indicano le coordinate del punto P . Per indicare un secondo punto verrà spesso utilizzata la lettera Q (con corrispondenti coordinate date da (y_1, \dots, y_d)). Nei casi, particolarmente significativi, planare, i.e. $d = 2$, e spaziale, i.e. $d = 3$, le corrispondenti coordinate saranno indicate con

$$P = (x, y), \quad P = (x, y, z),$$

a seconda della situazione considerata.

L'insieme \mathbb{R}^d è uno spazio vettoriale su \mathbb{R} grazie alla presenza delle operazioni di somma e prodotto per uno scalare: dati $P = (x_1, \dots, x_d)$, $Q = (y_1, \dots, y_d) \in \mathbb{R}^d$ e $\lambda \in \mathbb{R}$,

$$P + Q = (x_1 + y_1, \dots, x_d + y_d), \quad \lambda P = (\lambda x_1, \dots, \lambda x_d)$$

Il vettore nullo $(0, \dots, 0)$ viene indicato con la lettera O .

La differenza fondamentale tra il caso uni-dimensionale, $d = 1$, e il caso multidimensionale, $d > 1$, sta nel fatto che, mentre a differenza di \mathbb{R} , nello spazio \mathbb{R}^d , $d > 1$, non è definita una nozione di ordine e simboli del tipo $<$, \leq non hanno senso in relazione a vettori di dimensione superiore ad uno¹.

1. Serve una struttura metrica

Il modulo è la distanza di un punto dall'origine del sistema di riferimento. Una volta definita tale distanza, grazie alla struttura di spazio vettoriale, è possibile definire la distanza tra due punti qualsiasi dello spazio. In altri termini, dato il modulo, lo spazio considerato viene dotato di una **struttura metrica**. Da tale struttura discende, a catena, il concetto di limite con annessi e connessi.

Come definire il modulo? Supponiamo di avere scelto come misura delle distanze nel caso unidimensionale, quella definita attraverso il valore assoluto: i punti di \mathbb{R} di

¹In alcuni contesti specifici vengono introdotte relazioni di ordine (parziale) nel caso di \mathbb{R}^d , $d > 1$, ma nessuna di queste sarà considerata in queste Note.

coordinate x e y distano, per definizione, $|x - y|$. Passiamo al caso planare: $d = 2$. Il Teorema di Pitagora afferma che, dato un triangolo rettangolo in \mathbb{R}^2 di vertici $O = (0, 0)$, $P = (x, y)$ e $Q = (x, 0)$ con $x, y > 0$ fissati, vale la relazione

$$\ell(OP)^2 = \ell(OQ)^2 + \ell(QP)^2$$

dove $\ell(AB)$ indica la lunghezza del segmento AB . La lunghezza di segmenti unidimensionali è nota ed è data dal modulo unidimensionale della differenza delle coordinate: nel caso in considerazione $\ell(OQ) = |x - 0|$ e $\ell(QP) = |y - 0|$. Quindi $\ell(OP)^2 = x^2 + y^2$, cioè

$$\ell(OP) = \sqrt{x^2 + y^2}.$$

Cosa cambia nel caso tridimensionale? Fissiamo il punto $P = (x, y, z)$ e consideriamo il triangolo rettangolo di vertici $O = (0, 0, 0)$, $P = (x, y, z)$ e $Q = (x, y, 0)$. Il calcolo del caso bidimensionale, mostra che $\ell(OQ)^2 = x^2 + y^2$, quindi, sempre dando fiducia al Teorema di Pitagora, l'espressione della distanza di P da O è data da

$$\ell(OP) = \sqrt{\ell(OQ)^2 + \ell(QP)^2} = \sqrt{x^2 + y^2 + z^2}.$$

A questo punto, non sorprende la **definizione** seguente.

DEFINIZIONE 1.1. Dato $P = (x_1, \dots, x_d) \in \mathbb{R}^d$ si chiama **norma** (o modulo, o valore assoluto) di P il numero reale

$$(53) \quad |P|_d := \sqrt{x_1^2 + \dots + x_d^2} = \left(\sum_{k=1}^d x_k^2 \right)^{1/2}.$$

Nel caso unidimensionale $d = 1$, la relazione (53) diviene $|x|_1 = \sqrt{x^2}$ che coincide con la usuale definizione di modulo di un numero reale.

OSSERVAZIONE 1.2. Nello spazio \mathbb{R}^d è possibile definire anche il **prodotto scalare**: dati $P = (x_1, \dots, x_d)$ e $Q = (y_1, \dots, y_d)$, si pone $P \cdot Q := x_1 y_1 + \dots + x_d y_d$. La norma di $P \in \mathbb{R}^d$ può essere, allora, riscritta nella forma

$$(54) \quad |P|_d = (P \cdot P)^{1/2}.$$

Il prodotto scalare permette di dare senso al concetto di **ortogonalità** e, per questo motivo, si dice che lo spazio \mathbb{R}^d , considerato con il prodotto scalare \cdot , è dotato di una **struttura euclidea**. Dato che la norma $|\cdot|_d$ definita in (53) discende, attraverso (54), dall'introduzione del prodotto scalare, essa viene anche detta **norma euclidea**.

Per fini pratici, occorre prima di tutto convincersi del fatto che la norma $|\cdot|_d$ gode di proprietà analoghe a quella utilizzate in lungo e in largo nel caso unidimensionale.

PROPOSIZIONE 1.3. (*Proprietà della norma*) Per la norma $|\cdot|_d$ valgono:

- i. (*positività*) $|P|_d \geq 0$ per ogni $P \in \mathbb{R}^d$ e $|P|_d = 0$ se e solo se $P = O$;
- ii. (*omogeneità*) $|\lambda P|_d = |\lambda| |P|_d$ per ogni $\lambda \in \mathbb{R}$ e $P \in \mathbb{R}^d$;
- iii. (*disuguaglianza triangolare*) per ogni $P, Q \in \mathbb{R}^d$,

$$(55) \quad |P + Q|_d \leq |P|_d + |Q|_d.$$

Premettiamo alla dimostrazione di queste proprietà il risultato seguente.

LEMMA 1.4 (Disuguaglianza di Cauchy–Schwarz). Per ogni $P, Q \in \mathbb{R}^d$, vale

$$(56) \quad |P \cdot Q| \leq |P|_d |Q|_d \quad \forall P, Q \in \mathbb{R}^d.$$

DIMOSTRAZIONE DEL LEMMA 1.4. Dato $t \in \mathbb{R}$, si ha $|P + tQ|_d^2 \geq 0$. Inoltre

$$|P + tQ|_d^2 = (P + tQ) \cdot (P + tQ) = |P|_d^2 + 2P \cdot Qt + |Q|_d^2 t^2.$$

Dato che il polinomio di secondo grado $|P|_d^2 + 2P \cdot Qt + |Q|_d^2 t^2$ è sempre non negativo, il suo discriminante $(P \cdot Q)^2 - |P|_d^2 |Q|_d^2$ deve essere minore od uguale a zero; ne segue la disuguaglianza (56). \square

DIMOSTRAZIONE DELLA PROPOSIZIONE 1.3. Le dimostrazioni di **i.** e **ii.** non riservano particolari sorprese e sono lasciate per esercizio al lettore. La proprietà **iii.** segue dalla disuguaglianza di Cauchy–Schwarz. Infatti, utilizzando (56), si ottiene

$$|P + Q|_d^2 = |P|_d^2 + 2P \cdot Q + |Q|_d^2 \leq |P|_d^2 + 2|P \cdot Q| + |Q|_d^2 \leq (|P|_d + |Q|_d)^2.$$

Applicando la radice quadrata al primo e all'ultimo membro, si ottiene (55). \square

La norma euclidea $|\cdot|_d$ gode di un'altra proprietà analoga a quella già vista nel caso unidimensionale.

COROLLARIO 1.5. Per ogni $P, Q \in \mathbb{R}^d$, vale la disuguaglianza

$$(57) \quad ||P|_d - |Q|_d| \leq |P - Q|_d.$$

DIMOSTRAZIONE. La disequazione (57) equivale a

$$|P|_d - |P - Q|_d \leq |Q|_d \leq |P|_d + |P - Q|_d.$$

Entrambe le disuguaglianze sono conseguenza delle proprietà della norma. Infatti:

$$|P|_d - |P - Q|_d = |Q + (P - Q)|_d - |P - Q|_d \leq |Q|_d + |P - Q|_d - |P - Q|_d = |Q|_d$$

$$|Q|_d = |P + (Q - P)|_d \leq |P|_d + |Q - P|_d = |P|_d + |P - Q|_d.$$

La dimostrazione è completa. \square

In prima battuta, lo spazio vettoriale \mathbb{R}^d può essere pensato come uno spazio *isotropo*, cioè uno spazio in cui tutti i punti siano uguali tra di loro. In particolare, non c'è nessun motivo per preferire il punto origine O ad altri punti del piano. Si potrebbe decidere, ad esempio, di cambiare sistema di riferimento applicando una traslazione allo spazio \mathbb{R}^d , cioè si può pensare di scegliere un nuovo punto origine P^0 e di individuare tutti i punti di \mathbb{R}^d attraverso il vettore $P - P^0$ anziché tramite il vettore $P = P - O$. Questa idea suggerisce di utilizzare la traslazione e la norma per introdurre una nozione di distanza tra punti qualsiasi di \mathbb{R}^d .

DEFINIZIONE 1.6. *Dati $P, Q \in \mathbb{R}^d$ si chiama distanza di P da Q il numero reale*

$$(58) \quad d(P, Q) := |P - Q|_a.$$

Dalla definizione (58) e dalla Proposizione 1.3, discendono alcune proprietà per la distanza (la cui dimostrazione è lasciata per indispensabile esercizio).

PROPOSIZIONE 1.7. *(Proprietà della distanza) Per la distanza definita in (58) valgono le seguenti proprietà:*

- i. *(positività) $d(P, Q) \geq 0$ per ogni P, Q e $d(P, Q) = 0$ se e solo se $P = Q$;*
- ii. *(simmetria) $d(P, Q) = d(Q, P)$ per ogni $P, Q \in \mathbb{R}^d$;*
- iii. *(disuguaglianza triangolare) per ogni $P, Q, R \in \mathbb{R}^d$,*

$$(59) \quad d(P, Q) \leq d(P, R) + d(R, Q).$$

OSSERVAZIONE 1.8. Conseguenza immediata della definizione (58), è che la distanza d è *invariante per traslazioni*, cioè per ogni $P, Q \in \mathbb{R}^d$ e per ogni $h \in \mathbb{R}^d$ vale

$$d(P + h, Q + h) = d(P, Q).$$

La distanza in \mathbb{R}^d gode anche di altre proprietà di invarianza. Sia A una trasformazione lineare di \mathbb{R}^n in sè che sia *unitaria*², cioè tale che $A^{-1} = A^t$. Da (54) segue

$$d(AP, AQ) = AP \cdot AQ = P \cdot A^tAQ = P \cdot Q = d(P, Q).$$

Si noti che mentre l'invarianza per traslazioni segue solamente dalla relazione che intercorre tra distanza e norma (si veda la formula (58)), l'invarianza per rotazioni discende dalla specifica scelta operata nella definizione di norma. Con norme diverse, tale invarianza potrebbe non valere.

²Un esempio particolarmente interessante di trasformazione unitaria, nel caso $d = 2$, è dato da

$$A_\theta := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix} \quad \theta \in \mathbb{R}$$

La matrice A_θ descrive la trasformazione del piano \mathbb{R}^2 in sè data da una rotazione in senso antiorario di angolo θ attorno al punto O .

Una volta introdotte le nozioni di norma e distanza, si possono definire, senza difficoltà, i concetti di insiemi limitati³ e di intorni di un punto.

DEFINIZIONE 1.9. *Un insieme $A \subset \mathbb{R}^d$ è limitato se esiste $M > 0$ tale che*

$$|P|_d \leq M \quad \forall P \in A.$$

Dato $P_0 \in \mathbb{R}^d$ e $r > 0$, si chiama intorno aperto di centro P_0 e raggio r l'insieme

$$I_r(P_0) := \{P \in \mathbb{R}^d : |P - P_0|_d < r\};$$

si chiama intorno chiuso di centro P_0 e raggio r l'insieme

$$\bar{I}_r(P_0) := \{P \in \mathbb{R}^d : |P - P_0|_d \leq r\}.$$

La condizione di limitatezza può essere formulata anche come segue: un insieme $A \subset \mathbb{R}^d$ è limitato se e solo se è contenuto in un intorno di centro O (e raggio M sufficientemente grande).

Altre norme. Oltre alla norma euclidea, definita in (53), si possono considerare altre espressioni che continuano a soddisfare le stesse proprietà e che pertanto meritano anchesse il nome di norme. Consideriamo, ad esempio, il caso di \mathbb{R}^2 (analoghe definizioni si possono dare nel caso generale). Dato $p \geq 1$, consideriamo

$$|(x, y)|_{2,p} := (|x|^p + |y|^p)^{1/p}.$$

oppure

$$|(x, y)|_{2,\infty} := \max\{|x|, |y|\}.$$

E' possibile dimostrare che sia ogni norma del tipo $|\cdot|_{2,p}$, $p \in [1, \infty]$, verifica le stesse proprietà espone nella Proposizione 1.3. Le definizioni di intorno di un punto descrivono insiemi di tipo diverso.

Sussiste, al riguardo di tali varie norme diverse, il seguente risultato.

PROPOSIZIONE 1.10. *Dato $p \in [1, +\infty]$, esistono due costanti m_p e M_p tali che*

$$(60) \quad m_p |(x, y)|_{2,2} \leq |(x, y)|_{2,p} \leq M_p |(x, y)|_{2,2} \quad \forall (x, y) \in \mathbb{R}^2.$$

DIMOSTRAZIONE. La disuguaglianza di destra è valida con $M_p = 2^{1/p}$, infatti

$$|(x, y)|_{2,p} = (|x|^p + |y|^p)^{1/p} \leq 2^{1/p} \max\{|x|, |y|\} \leq 2^{1/p} (|x|^2 + |y|^2)^{1/2}.$$

Per l'altra, si può procedere in maniera analoga:

$$|(x, y)|_{2,2} = (|x|^2 + |y|^2)^{1/2} \leq 2^{1/2} \max\{|x|, |y|\} \leq 2^{1/2} (|x|^p + |y|^p)^{1/p},$$

ottenendo la disuguaglianza di sinistra con $m_p := 2^{-1/2}$. □

³Dato che in \mathbb{R}^d , $d > 1$, non sono definite nozioni di ordinamento, non ha senso parlare di insiemi superiormente/inferiormente limitati.

Dalle disuguaglianze (60) segue che:

- gli insiemi limitati rispetto alla norma euclidea e quelli rispetto alla norma $|\cdot|_{2,p}$ coincidono;
- anche se gli intorni relativi a norme diverse sono diversi, all'interno di ogni intorno relativo ad una norma, è sempre contenuto un intorno (se necessario di raggio più piccolo) relativo all'altra norma. Questa proprietà è particolarmente rilevante nell'ambito delle proprietà di convergenza (che vedremo tra poche righe).

Versioni analoghe della Proposizione 1.10 valgono per qualsiasi norma in \mathbb{R}^d che verifichi le proprietà elencate nella Proposizione 1.3.

2. Quattro salti in \mathbb{R}^d : successioni di punti

Una successione di punti in \mathbb{R}^d è un'applicazione da \mathbb{N} a \mathbb{R}^d che associa ad ogni numero naturale n un punto P^n di \mathbb{R}^d . Tramite la definizione di distanza è possibile definire il concetto di limite di una successione $P^n \in \mathbb{R}^d$ utilizzando la stessa definizione del caso unidimensionale.

DEFINIZIONE 2.1. *Siano $P^n, \bar{P} \in \mathbb{R}^d$ per ogni $n \in \mathbb{N}$. La successione P^n converge a \bar{P} per $n \rightarrow +\infty$ e si scrive $\lim_{n \rightarrow +\infty} P^n = \bar{P}$ se*

$$(61) \quad \forall \varepsilon > 0, \quad \exists n_\varepsilon \in \mathbb{N}, \quad \forall n \geq n_\varepsilon, \quad \text{vale } |P^n - \bar{P}|_d < \varepsilon.$$

Che cosa vuol dire, in concreto, che la successione P^n converge a \bar{P} per $n \rightarrow +\infty$? La definizione (61) va letta a partire dalla fine: alcuni punti della successione P^n distano meno di ε dal punto limite \bar{P} . Quindi, la convergenza di P^n a \bar{P} indica una proprietà di "vicinanza" di (alcuni) punti della successione al corrispondente punto limite. Occorre precisare, però, **quanto** siano vicini tali punti al punto limite e **quali** siano i punti a godere di tale proprietà. È a questo fine che vengono introdotti i (misteriosi) valori ε e n_ε . La "vicinanza" è quantificata dal parametro ε , che è **qualsiasi**. L'arbitrarietà di ε fa in modo che la definizione sia interessante qualsiasi sia la scala significativa considerata, indipendentemente dal gusto e dalla preferenza dell'utente di turno. Più il valore di ε è piccolo e più la condizione $|P^n - \bar{P}|_d < \varepsilon$ diviene restrittiva e, di conseguenza, soddisfatta da un minor numero di punti della successione. Il valore n_ε permette di individuare punti della successione per cui la relazione è soddisfatta. La richiesta, nella definizione di limite, è che la distanza di P_n da \bar{P} sia minore di ε tutte le volte che si scelgano indici n sufficientemente grandi. In definitiva, fissata una qualsiasi soglia di errore $\varepsilon > 0$, i termini della successione P_n sono ben approssimati da \bar{P} per tutte le scelte di n maggiori di n_ε .

OSSERVAZIONE 2.2. La nozione di convergenza si basa sulla definizione di distanza. Disponendo di metriche diverse potrebbe quindi accadere che una stessa successione $\{P^n\}$ sia convergente rispetto a certe metriche e non convergente rispetto ad altre. Il risultato della Proposizione 1.10 consente di riconoscere che se una successione $\{P^n\}$ è convergente rispetto alla metrica euclidea lo è anche rispetto alle metriche $|\cdot|_{2,p}$. Per questo motivo, metriche che verifichino stime del tipo (60) si dicono essere **equivalenti** alla metrica euclidea.

Come nel caso del limite di successioni reali, anche il limite di successioni di punti di \mathbb{R}^d è lineare: se P^n e Q^n sono due successioni convergenti, rispettivamente, a P e Q ,

$$\lim_{n \rightarrow +\infty} \alpha P^n + \beta Q^n = \alpha P + \beta Q \quad \text{per ogni } \alpha, \beta \in \mathbb{R}.$$

La dimostrazione è conseguenza immediata della disuguaglianza

$$|(\alpha P^n + \beta Q^n) - (\alpha P + \beta Q)|_d \leq |\alpha| |P^n - P|_d + |\beta| |Q^n - Q|_d.$$

Inoltre, come nel caso reale, il limite di una successione, quando esiste, è unico.

Una successione di punti P^n è individuata univocamente dalle successioni reali delle sue coordinate. Che relazione intercorre tra la convergenza della successione P^n e quella delle successioni delle coordinate?

PROPOSIZIONE 2.3. *Sia $P^n = (x_1^n, \dots, x_d^n)$ e $\bar{P} = (\bar{x}_1, \dots, \bar{x}_d)$. Allora la successione P^n converge a \bar{P} se e solo se*

$$\lim_{n \rightarrow +\infty} x_j^n = \bar{x}_j \quad \text{per ogni } j = 1, \dots, d.$$

DIMOSTRAZIONE. Dalla definizione di distanza in \mathbb{R}^d , segue

$$|x_j^n - \bar{x}_j| \leq |P^n - \bar{P}|_d = \left(\sum_{k=1}^d (x_k^n - \bar{x}_k)^2 \right)^{1/2}$$

per ogni $j \in \{1, \dots, d\}$. Di conseguenza, se $|x_k^n - \bar{x}_k|$ tende a 0 per ogni k , allora anche $|P^n - \bar{P}|_d$ tende a zero. Viceversa, se $|P^n - \bar{P}|_d$ tende a zero, anche la successione reale $|x_j^n - \bar{x}_j|$ tende a zero per $n \rightarrow +\infty$ per ogni j . \square

La Proposizione 2.3 indica che determinare il limite di una successione di punti in \mathbb{R}^d equivale a determinare i limiti di d successioni reali. Ad esempio, per le successioni $P_n := (1/n^2, e^{1/n})$ e $Q_n := (n \sin(1/n), \sin n/n)$ valgono i limiti $\lim_{n \rightarrow +\infty} P_n = (0, 1)$ e $\lim_{n \rightarrow \infty} Q_n = (1, 0)$.

Come nel caso uni-dimensionale, ogni successione convergente è anche limitata e, di conseguenza, una successione non limitata non può essere convergente.

PROPOSIZIONE 2.4. Sia $\{P^n\}$ una successione convergente. Allora la successione $\{P^n\}$ è limitata.

DIMOSTRAZIONE. Sia \bar{P} il limite della successione P^n . Per definizione, esiste n_1 tale che $|P^n - \bar{P}|_d < 1$ per ogni $n \geq n_1$. Vale quindi la stima

$$|P^n|_d \leq |\bar{P}|_d + |P^n - \bar{P}|_d < |\bar{P}|_d + 1 \quad \forall n \geq n_1.$$

Per dimostrare la limitatezza della successione, basta controllare la norma dei termini P^0, \dots, P^{n_1-1} . Quindi, scegliendo

$$M := \max\{|P^0|_d, \dots, |P^{n_1-1}|_d, |\bar{P}|_d + 1\},$$

si ha che $|P^n| \leq M$ per ogni $n \in \mathbb{N}$. □

ESERCIZIO 2.5. Siano P^n e Q^n due successioni in \mathbb{R}^d che convergono, rispettivamente, a \bar{P} e \bar{Q} . Dimostrare che la successione $a_n := P^n \cdot Q^n$ converge a $\bar{P} \cdot \bar{Q}$.

La limitatezza non è una condizione sufficiente per la convergenza: si pensi all'esempio facile della successione $a_n = (-1)^n$. Occorre, quindi, qualche ipotesi aggiuntiva oltre alla limitatezza per essere certi della convergenza della successione in questione. Nel caso uni-dimensionale, ad esempio, vale il *criterio di monotonia*: una successione di numeri reali limitata e monotona è convergente. Nel caso di successioni vettoriali, dato che in \mathbb{R}^n non è possibile introdurre un ordinamento totale se $n \geq 2$, non esiste un analogo di questo criterio.

L'ipotesi di limitatezza, dunque, non implica la convergenza della successione assegnata. E' vero però una proprietà più debole che necessita la definizione seguente.

DEFINIZIONE 2.6. Data la successione di punti $\{P^n\}$ e una successione $\{n_k\}$ di numeri naturali, strettamente crescente, la successione $\{P^{n_k}\}$ si dice essere una **sottosuccessione** della successione $\{P^n\}$.

Il termine sottosuccessione può essere chiarito tramite alcuni esempi. Data la successione $\{n\} = \{0, 1, 2, 3, \dots\}$ dei numeri naturali:

– le successioni

$$\{1, 3, 5, 7, \dots, 2k+1, \dots\}, \quad \{1, 10, 100, 1000, \dots, 10^k, \dots\}$$

sono sottosuccessioni della successione assegnata;

– le successioni

$$\{1, 2, 3, \pi, 4, \dots\}, \quad \{1, 3, 2, 4, \dots\}, \quad \{1, 2, 2, 3, 4, \dots\}$$

non sono sottosuccessioni della successione dei naturali $\{n\}$.

Dalla definizione di convergenza, segue immediatamente che se una successione di punti $\{P^n\}$ converge a \bar{P} allora tutte le sottosuccessioni di $\{P^n\}$ convergono allo stesso punto limite \bar{P} .

TEOREMA 2.7 (*Teorema di Bolzano–Weierstrass*). *Da ogni successione limitata $\{P^n\}$ è possibile estrarre almeno una sottosuccessione $\{P^{n_k}\}$ convergente.*

DIMOSTRAZIONE. La dimostrazione è divisa in due passi: prima si considera il caso di \mathbb{R} e poi il caso generale.

1. **Caso uni-dimensionale.** Dimostriamo prima di tutto il Teorema nel caso uni-dimensionale. Sia $\{x^n\}$ una successione reale limitata, cioè interamente contenuta nell'intervallo $I_0 := [a, b]$ per qualche $a, b \in \mathbb{R}$, $a < b$. Consideriamo i sotto-intervalli I_0^\pm definiti da

$$I_0^- := \left[a, \frac{a+b}{2} \right], \quad I_0^+ := \left[\frac{a+b}{2}, b \right].$$

Dato che x^n è composta da un numero infinito di elementi, almeno uno dei due sotto-intervalli I_0^\pm contiene infiniti elementi della successione x^n . Indichiamo con I_1 , quello dei due sotto-intervalli per cui è verificata tale proprietà e con a_1, b_1 i suoi estremi, $I = [a_1, b_1]$ e ripetiamo il ragionamento precedente. Consideriamo i sotto-intervalli I_1^\pm definiti da

$$I_1^- := \left[a_1, \frac{a_1 + b_1}{2} \right], \quad I_1^+ := \left[\frac{a_1 + b_1}{2}, b_1 \right].$$

Dato che un numero infinito di elementi $\{x^n\}$ si trovano in I_1 , almeno uno dei due sotto-intervalli I_1^\pm contiene infiniti elementi della successione x^n . Indichiamo con I_2 tale sotto-intervallo.

Iterando il procedimento, si costruisce una successione di intervalli chiusi e limitati I_n , ciascuno contenente un numero infinito di elementi della successione $\{x^n\}$. Per l'assioma degli intervalli incapsulati, esiste $\ell \in \mathbb{R}$ contenuto in I_n per ogni n . Costruiamo ora una sottosuccessione $\{x^{n_k}\}$ di $\{x^n\}$ come segue:

- scegliamo n_1 in modo che $x^{n_1} \in I_1$;
- scegliamo $n_2 > n_1$ in modo che $x^{n_2} \in I_2$;
- in generale, scegliamo $n_{k+1} > n_k$ in modo che $x^{n_{k+1}} \in I_{k+1}$.

Per costruzione, $|x^{n_k} - \ell| < (b - a)/2^k$, e, di conseguenza, la sottosuccessione x^{n_k} converge ad ℓ .

2. **Caso multi-dimensionale.** Consideriamo il caso di successioni in \mathbb{R}^2 (il caso generale è analogo). Sia $\{P^n\} = \{(x^n, y^n)\}$ una successione limitata in \mathbb{R}^2 . Ne segue che anche le due successioni di numeri reali delle coordinate $\{x^n\}$ e $\{y^n\}$ sono limitate. Applicando alla successione $\{x^n\}$ il teorema di Bolzano–Weierstrass uni-dimensionale, si deduce che esiste una sotto-successione $\{x^{n_k}\}$ convergente.

La sottosuccessione $\{y^{n_k}\}$, estratta dalla successione $\{y^n\}$ con gli stessi indici della sotto-successione $\{x^{n_k}\}$, è limitata. Applicando nuovamente il teorema di Bolzano–Weierstrass unidimensionale, si deduce l'esistenza di una sotto-sotto-successione $\{y_{n_{k_j}}\}$

convergente. La successione $P^{n_{k_j}} := (x_{n_{k_j}}, y_{n_{k_j}})$, avendo entrambe le successioni delle coordinate convergenti, è una successione convergente. \square

Il criterio di Cauchy. Tramite il Teorema di Bolzano–Weierstrass si dimostra una condizione necessaria e sufficiente per la convergenza di successioni di punti: il criterio di Cauchy.

TEOREMA 2.8 (Criterio di Cauchy). *Una successione P^n in \mathbb{R}^d è convergente se e solo se vale la condizione seguente*

$$(62) \quad \forall \varepsilon > 0, \quad \exists n_\varepsilon \in \mathbb{N}, \quad \forall m, n \geq n_\varepsilon, \quad \text{vale} \quad |P^m - P^n|_d < \varepsilon.$$

DIMOSTRAZIONE. Se la successione P^n converge a \bar{P} , vale la stima

$$|P^m - P^n|_d \leq |P^m - \bar{P}|_d + |P^n - \bar{P}|_d,$$

da cui segue che la condizione (62).

Viceversa, se la successione $\{P^n\}$ è di Cauchy, allora è anche limitata. Infatti, scegliendo $\varepsilon = 1$, si deduce, da (62), che, per un opportuno $n_1 \in \mathbb{N}$ vale

$$|P^n|_d \leq |P^{n_1}|_d + |P^n - P^{n_1}|_d < |P^{n_1}|_d + 1$$

per ogni $n \geq n_1$. Posto $M := \max\{|P^0|_d, |P^1|_d, \dots, |P^{n_1-1}|_d, |P^{n_1}|_d + 1\}$, si ha

$$|P^n|_d \leq M \quad \forall n \in \mathbb{N}.$$

Grazie al Teorema di Bolzano–Weierstrass, si deduce che esiste una sottosuccessione $\{P^{n_k}\}$ convergente ad un opportuno punto limite \bar{P} . Per dimostrare che tutta la successione P^n converge a \bar{P} , basta utilizzare la stima

$$|P^n - \bar{P}|_d \leq |P^n - P^{n_k}|_d + |P^{n_k} - \bar{P}|_d$$

e lavorare con accortezza con ε e n_ε . I dettagli sono lasciati al volenteroso lettore. \square

ESERCIZIO 2.9. Sia $\{P^n\}$ tale che, per qualche $\theta \in [0, 1)$ valga la condizione

$$|P^{n+1} - P^n|_d \leq \theta |P^n - P^{n-1}|_d \quad \forall n \in \mathbb{N} \setminus \{0\}.$$

Dimostrare che la successione P^n converge ad un $\bar{P} \in \mathbb{R}^d$.

La condizione (62) non necessita la conoscenza esplicita del limite \bar{P} . Tale fatto permette di utilizzare il criterio di Cauchy come mattone fondamentale per una costruzione rigorosa dell'insieme dei numeri reali a partire dai numeri razionali (o, più in generale, per una costruzione dell'insieme \mathbb{R}^d a partire da \mathbb{Q}^d). L'idea di base sta nel fatto che un numero reale è individuato da una successione (razionale) che lo approssimi con un errore piccolo quanto si vuole. Ad esempio, per approssimare/definire il numero (irrazionale) π si può considerare la successione di numeri (razionali)

$$a_0 = 3 \quad a_1 = 3,1 \quad a_2 = 3,14 \quad a_3 = 3,141 \quad a_4 = 3,1415 \quad a_5 = 3,14159 \quad \dots,$$

cioè l'elemento a_n è quello che si ottiene considerando i primi n decimali del numero π .

Da questo punto di vista, un numero è individuato da una **successione di Cauchy** in \mathbb{Q} , cioè da una successione a_n tale che

$$\forall \varepsilon \in \mathbb{Q}, \quad \varepsilon > 0, \quad \exists n_\varepsilon \in \mathbb{N}, \quad \text{tale che} \quad |a_m - a_n| < \varepsilon \quad \forall m, n \geq n_\varepsilon.$$

L'insieme delle successioni di Cauchy

$$S = \{ \{x_n\} : \text{successione di Cauchy in } \mathbb{Q} \}.$$

consta, però, di troppi elementi: esistono successioni diverse che rappresentano lo stesso numero (basta pensare a due diverse successioni di razionali che tendano a zero). Occorre quindi raggruppare gli elementi di S in classi di equivalenza, in modo che ogni singola classe rappresenti un singolo numero reale. La relazione d'equivalenza si definisce così: date $\{x_n\}$ e $\{y_n\}$ in S

$$\{x_n\} \sim \{y_n\} \quad \text{se} \quad \lim_{n \rightarrow +\infty} (x_n - y_n) = 0.$$

E' facile verificare che si tratta di una relazione di equivalenza (riflessività e simmetria sono banali, la transitività discende dalla disuguaglianza triangolare). A questo punto il gioco è fatto: l'insieme dei numeri reali è il quoziente $\mathbb{R} := S / \sim$.

3. Informazioni di base sulla topologia di \mathbb{R}^d

Partiamo da un problema prototipo: data una funzione $f : A \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$, determinare una soluzione dell'equazione

$$(63) \quad f(P) = 0.$$

Supponiamo di essere in grado di dimostrare l'esistenza di P^n tali che $f(P^n)$ non sia nullo, ma verifichi la stima

$$|f(P^n)| \leq \frac{1}{n}.$$

Il limite \bar{P} della successione P^n è un buon candidato ad essere soluzione di (63), ma...

- ...la successione P^n è convergente ad un limite P ?
- ...il limite P appartiene all'insieme A ?
- ...è vero che $f(P) = 0$?

Ciascuna di queste domande porta all'introduzione di un concetto importante: la compattezza, la chiusura, la continuità. A seguire, vediamo i primi due, mentre il terzo sarà esplorato in lungo e in largo più avanti.

DEFINIZIONE 3.1. *Un insieme $C \subset \mathbb{R}^d$ è chiuso (rispetto all'operazione di limite) se per ogni successione $P^n \in A$ convergente il punto limite \bar{P} appartiene all'insieme A .*

ESEMPIO 3.2. Il prodotto cartesiano $[a, b] \times [c, d]$ è un insieme chiuso di \mathbb{R}^2 , per ogni scelta di $a, b, c, d \in \mathbb{R}$. Infatti, sia $P^n = (x^n, y^n) \in [a, b] \times [c, d]$ una successione convergente al limite $\bar{P} = (\bar{x}, \bar{y})$. Dato che

$$a \leq x^n \leq b, \quad c \leq y^n \leq d, \quad \forall n \in \mathbb{N},$$

passando al limite per $n \rightarrow +\infty$, si deduce che $(\bar{x}, \bar{y}) \in [a, b] \times [c, d]$.

ESEMPIO 3.3. L'insieme dei numeri naturali \mathbb{N} in \mathbb{R} è un insieme chiuso. Infatti, se x_n è una successione di numeri naturali convergente ad \bar{x} , si ha che, per qualche \bar{n} , vale la stima $|x_n - \bar{x}| < 1/2$ per ogni $n \geq \bar{n}$. Se $m, n \geq \bar{n}$, si ha

$$|x_n - x_m| \leq |x_n - \bar{x}| + |x_m - \bar{x}| < \frac{1}{2} + \frac{1}{2} = 1,$$

e, di conseguenza, $x_n = x_m$. Quindi la successione è costante per $n \geq \bar{n}$ e, pertanto, convergente.

ESERCIZIO 3.4. Sia C un insieme finito di punti. Dimostrare che C è chiuso.

In generale, una successione contenuta in un insieme chiuso potrebbe non essere convergente. Occorre, quindi, una condizione che garantisca, quanto meno, la convergenza di una qualche sottosuccessione della successione data. Eccola.

DEFINIZIONE 3.5. *Un insieme $K \subset \mathbb{R}^d$ è compatto (per successioni)⁴ se da ogni successione $P^n \in K$ è sempre possibile estrarre una sottosuccessione P^{n_k} convergente a qualche $\bar{P} \in K$.*

ESEMPIO 3.6. Il prodotto cartesiano $[a, b] \times [c, d]$ è un insieme compatto di \mathbb{R}^2 , per ogni scelta di $a, b, c, d \in \mathbb{R}$. Infatti, sia $P^n = (x^n, y^n) \in [a, b] \times [c, d]$. Dato che la successione P^n è limitata, essa ammette, per il Teorema di Bolzano–Weierstrass una sottosuccessione P^{n_k} convergente ad un qualche punto \bar{P} e, dato che $[a, b] \times [c, d]$ è chiuso, il punto \bar{P} appartiene ad $[a, b] \times [c, d]$.

Tutti gli insiemi compatti sono anche chiusi. Infatti, se $P^n \in K$ è una successione convergente a \bar{P} , ogni sottosuccessione P^{n_k} converge allo stesso limite \bar{P} . Per definizione, quindi, \bar{P} appartiene ad K . Il viceversa non è vero: esistono insiemi chiusi che non sono compatti.

ESEMPIO 3.7. L'insieme dei numeri naturali \mathbb{N} in \mathbb{R} non è compatto. Infatti, tutte le sottosuccessioni di $a_n = n$ sono illimitate e quindi non convergenti.

La limitatezza impedisce alle successioni di "scappare" all'infinito e, insieme alla chiusura, garantisce l'esistenza di sottosuccessioni convergenti.

⁴Esistono anche altre definizioni di compattezza, utilizzate in ambiti più generali, che comunque, nel caso di \mathbb{R}^d coincidono con la compattezza per successioni.

TEOREMA 3.8 (*Insiemi compatti di \mathbb{R}^d*). *Un sottoinsieme K di \mathbb{R}^d è compatto (per successioni) se e solo se è chiuso e limitato.*

DIMOSTRAZIONE. Sia K compatto e, di conseguenza, chiuso. Supponiamo, per assurdo, che K non sia limitato. Allora, per ogni $k \in \mathbb{N}$ esiste P^{n_k} tale che $|P^{n_k}|_d \geq n$ e, di conseguenza,

$$(64) \quad \lim_{k \rightarrow +\infty} |P^{n_k}|_d = +\infty.$$

Dato che K è compatto, esiste una sottosuccessione di P^{n_k} , che indichiamo per semplicità (e con abuso di notazione...) sempre con P^{n_k} convergente a qualche $\bar{P} \in K$. In particolare, dato che

$$||P^{n_k}|_d - |\bar{P}|_d| \leq |P^{n_k} - \bar{P}|_d,$$

la successione reale $|P^{n_k}|_d$ converge a $|\bar{P}|_d$, in contraddizione con (64).

Viceversa, sia A chiuso e limitato e sia P^n una successione in K . Per il Teorema di Bolzano–Weierstrass, Teorema 2.7, è possibile estrarre una sottosuccessione P^{n_k} convergente ad un punto \bar{P} . Dato che l'insieme K è chiuso, il punto limite appartiene ad K ed è quindi verificata la proprietà di compattezza. \square

Parallelamente alla nozione di insieme chiuso, si introduce anche la nozione di insieme aperto.

DEFINIZIONE 3.9. *Sia $S \subseteq \mathbb{R}^d$. Un punto $P \in \mathbb{R}^d$ è interno ad S se esiste $r > 0$ tale che $I_r(P) = \{P \in \mathbb{R}^d : |P - P_0|_d < r\} \subseteq S$. Un punto $P \in \mathbb{R}^d$ è esterno ad S se è interno all'insieme complementare $\mathbb{R}^d \setminus S$, cioè se esiste $r > 0$ tale che $S \cap I_r(P) = \emptyset$.*

Un insieme $A \subseteq \mathbb{R}^d$ si dice aperto se i suoi punti sono tutti interni.

ESEMPIO 3.10. Dati $a, b, c \in \mathbb{R}$ l'insieme $A := \{ax + by + c > 0\} \subset \mathbb{R}^2$ è aperto. Infatti, dato $P_0 = (x_0, y_0) \in A$, si ha $I_r(P_0) \subseteq A$ per ogni scelta di r strettamente minore della distanza P_0 dalla retta di equazione $ax + by + c = 0$.

ESEMPIO 3.11. Dati $a, b, c \in \mathbb{R}$ l'insieme $C := \{(x, y) \in \mathbb{R}^2 : ax + by + c \geq 0\}$ non è aperto. Infatti, tutti i punti della forma $P_0 = (x_0, y_0)$ con $ax_0 + by_0 + c = 0$ appartengono a C , ma ogni intorno di tali punti contiene sempre punti non appartenenti a C . Analogo discorso vale per l'insieme $C_2 := \{(x, y) \in \mathbb{R}^2 : ax + by + c = 0\}$.

Che relazione intercorre tra insiemi aperti e insiemi chiusi?

PROPOSIZIONE 3.12. *Un insieme A è aperto se e solo se il suo complementare $C := \mathbb{R}^d \setminus A$ è chiuso.*

DIMOSTRAZIONE. Supponiamo A aperto e dimostriamo che $C := \mathbb{R}^d \setminus A$ è chiuso. Sia $\{P^n\}$ una successione di punti in C convergente a \bar{P} . Se, per assurdo, \bar{P} appartenesse ad A , esisterebbe $r > 0$ tale che $I_r(\bar{P}) \subseteq A$. In particolare, per ogni n si avrebbe $|P^n - \bar{P}| \geq r > 0$ in contraddizione con il fatto che la successione P^n converge a \bar{P} .

Supponiamo $C = \mathbb{R}^d \setminus A$ chiuso e dimostriamo che A è aperto. Se, per assurdo, A non fosse aperto, esisterebbe $\bar{P} \in A$ non interno ad A . Quindi, per ogni $n \in \mathbb{N}$, esisterebbe $P^n \in C$ tale che $P^n \in I_{1/n}(\bar{P})$. La successione P^n sarebbe allora convergente al punto \bar{P} , in contraddizione con l'ipotesi. \square

Gli insiemi aperti e gli insiemi chiusi sono chiusi rispetto all'operazione di unione e di intersezione: cioè valgono le affermazioni seguenti

- se A_1, A_2 sono aperti, allora anche $A_1 \cup A_2, A_1 \cap A_2$ sono aperti;
- se C_1, C_2 sono chiusi, allora anche $C_1 \cup C_2, C_1 \cap C_2$ sono chiusi.

OSSERVAZIONE 3.13. Nel caso di unione o intersezione di un numero infinito di aperti/chiusi, la situazione è un po' più delicata. Valgono le implicazioni seguenti:

- se $A_n, n \in \mathbb{N}$, è una successione di aperti, allora l'unione $\bigcup_{n \in \mathbb{N}} A_n$ è un aperto;
- se $C_n, n \in \mathbb{N}$, è una successione di chiusi, allora l'intersezione $\bigcap_{n \in \mathbb{N}} C_n$ è un chiuso.

Non è difficile costruire esempi di successioni di aperti la cui intersezione sia un chiuso e successioni di chiusi la cui unione sia un aperto. Provateci.

Per ogni insieme $S \subset \mathbb{R}^d$ è possibile distinguere tre categorie di punti: i punti interni, i punti esterni e... gli altri, cioè i punti che non sono né interni né esterni ad S .

DEFINIZIONE 3.14. Un punto $P \in \mathbb{R}^d$ è un **punto di frontiera** dell'insieme S se ogni intorno di P contiene punti di S e punti non appartenenti ad S : per ogni $r > 0$,

$$S \cap I_r(P_0) \neq \emptyset, \quad (\mathbb{R}^d \setminus S) \cap I_r(P_0) \neq \emptyset.$$

L'insieme dei punti di frontiera di S si chiama **frontiera** di S e si denota con ∂S .

Si chiama **chiusura** dell'insieme S l'insieme $\bar{S} := S \cup \partial S$.

La chiusura \bar{S} di un insieme S è il più piccolo insieme chiuso contenente S .

ESEMPIO 3.15. Per gli insiemi

$$A := \{(x, y) : ax + by + c < 0\}, \quad C := \{(x, y) : ax + by + c \leq 0\}.$$

con $a, b, c \in \mathbb{R}$, si ha $\partial A = \partial C = \{(x, y) : ax + by + c = 0\}$. A voi la verifica.

ESEMPIO 3.16. Per gli insiemi

$$A := \{(x, y) : x^2 + y^2 < r^2\}, \quad C := \{(x, y) : x^2 + y^2 \leq r^2\},$$

con $r > 0$, si ha $\partial A = \partial C = \{(x, y) : x^2 + y^2 = r^2\}$.

ESERCIZIO 3.17. Determinare la frontiera e la chiusura dei seguenti insiemi

$$\begin{aligned} S_1 &= \{(x, y) : 0 < x^2 + y^2 \leq 1\}, & S_2 &= \{(x, y) : x + y + 1 < 0, x + 1 < 0\}, \\ S_3 &= \{(0, 0), (1, 0), (0, 1)\}, & S_4 &= \{(x, y) : x, y \geq 0, x + y \leq 0\} \end{aligned}$$

ESERCIZIO 3.18. Dimostrare che la frontiera di un insieme S e del suo complementare $T := \mathbb{R}^d \setminus S$ coincidono, cioè $\partial S = \partial T$.

Come espresso dal seguente risultato (la cui dimostrazione è lasciata come esercizio), gli insiemi aperti e gli insiemi chiusi hanno una relazione particolare con la loro frontiera.

PROPOSIZIONE 3.19. *Un insieme A è aperto se e solo se $A \cap \partial A = \emptyset$; un insieme C è chiuso se e solo se $\partial C \subseteq C$.*

4. Natura non facit saltus: le curve

Una successione di punti $P^n \in \mathbb{R}^d$ può essere interpretata come una sequenza di posizioni nello spazio d -dimensionale di una particella in movimento che viene osservata ad istanti successivi $t_0 < t_1 < \dots < t_n < \dots$: il punto P^n è la posizione all'istante t_n . Una maniera altrettanto (o più?) interessante per descrivere il moto di un punto è quella di considerare che la variabile temporale non vari in un insieme discreto, ma in un insieme continuo⁵. Questo corrisponde, in concreto a considerare funzioni del tipo

$$\phi : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}^d$$

dove I è un intervallo di \mathbb{R} , che associano ad uno scalare $t \in I$ il vettore $\phi(t) = (\phi_1(t), \dots, \phi_d(t))$. La funzione ϕ_i che associa a t la i -esima componente di $\phi(t)$ si chiama i -esima componente di ϕ .

Tenendo conto della definizione di modulo in \mathbb{R}^d , è possibile, senza fatica, introdurre il concetto di continuità anche in questo contesto.

DEFINIZIONE 4.1. *Sia I un intervallo di \mathbb{R} . Una funzione $\phi : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}^d$ è continua in $t_0 \in I$ se*

$$(65) \quad \forall \varepsilon > 0, \quad \exists \delta > 0 \quad \text{tale che} \quad |\phi(t) - \phi(t_0)|_d < \varepsilon \quad \forall t \in I, |t - t_0| < \delta.$$

Una funzione $\phi : I \subseteq \mathbb{R} \longrightarrow \mathbb{R}^d$ è continua in I se è continua in tutti i punti di I . L'insieme delle curve $\phi : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^d$ continue in I si indica con il simbolo $C(I; \mathbb{R}^d)$.

⁵La locuzione latina *Natura non facit saltus*, formulata per la prima volta da Carl von Linné, nella *Philosophia Botanica*, è stata utilizzata da Leibniz come assioma per negare l'esistenza di quantità discrete indivisibili ed esprimere che i processi naturali variano in maniera continua.

Dalla definizione di modulo e seguendo una strategia analoga a quella utilizzata nella Proposizione 2.3, si dimostra che la continuità della funzione ϕ equivale alla continuità di tutte le sue componenti ϕ_i .

Una sottoclasse delle funzioni continue è data dalle cosiddette *funzioni lipschitziane*: una funzione $\phi \in C(I; \mathbb{R}^d)$ si dice **lipschitziana** se esiste $L > 0$ tale che

$$(66) \quad |\phi(t) - \phi(\tau)|_d \leq L|t - \tau| \quad \forall t, \tau \in I.$$

In questo caso, la condizione di continuità (65) è verificata scegliendo $\delta = \varepsilon/L$. Come si vede, il valore di δ è indipendente dal punto t_0 di continuità. Sempre ragionando come nella Proposizione 2.3, si può mostrare che una funzione ϕ è lipschitziana se e solo se tutte le sue componenti ϕ_i sono lipschitziane.

Nella logica della descrizione di un moto nello spazio, è ragionevole restringere l'attenzione a funzioni ϕ le cui componenti ϕ_1, \dots, ϕ_d sono derivabili. In questo modo, per ogni valore $t \in I$ è ben definito il **vettore velocità** $\phi'(t) = (\phi'_1(t), \dots, \phi'_d(t))$. L'insieme delle curve $\phi : I \subset \mathbb{R} \rightarrow \mathbb{R}^d$ con derivata ϕ' continua, si indica con il simbolo $C^1(I; \mathbb{R}^d)$.

DEFINIZIONE 4.2. *Sia $I \subseteq \mathbb{R}$ un intervallo di estremi a e b . Una curva parametrizzata regolare è un'applicazione $\phi \in C^1(I; \mathbb{R}^d)$ tale che*

$$(67) \quad \phi'(t) \neq 0 \quad \forall t \in I.$$

Una curva parametrizzata regolare a tratti è un'applicazione $\phi \in C(I; \mathbb{R}^d)$ tale che esiste un insieme $S = \{a < t_0 < t_1 < \dots < t_N < b\} \subset I$ per cui vale

$$\phi \text{ è regolare a tratti in } [a, t_0] \cap I, [t_1, t_2], \dots, [t_{N-1}, t_N], [t_N, b] \cap I$$

In entrambi i casi, l'insieme immagine $\phi(I)$ è detto supporto della curva.

Da un punto di vista cinematico, la funzione ϕ che definisce la curva è la **legge oraria** del moto in questione, mentre l'insieme immagine $\phi(I)$ ne è la **traiettoria**. Si noti che la stessa traiettoria può essere percorsa con leggi orarie diverse e che, parallelamente, curve parametrizzate diverse possono avere lo stesso supporto.

La condizione $\phi'(t_0) \neq 0$ equivale a richiedere che il moto del punto descritto dalla legge oraria ϕ non abbia mai velocità nulla; dal punto di vista geometrico, essa garantisce l'esistenza in ogni punto della curva di una retta tangente, la cui equazione parametrica è $P = \phi(t_0) + \phi'(t_0)h$ con $h \in \mathbb{R}$.

Segmenti e poligonali. Consideriamo due punti distinti $P_0, P_1 \in \mathbb{R}^d$. Come descrivere un moto da P_0 a P_1 lungo una linea retta (cioè per la via più rapida)? In

altre parole, come parametrizzare il segmento $\overline{P_0P_1}$? E' immediato riconoscere che la definizione seguente

$$\phi(t) = P_0 + t(P_1 - P_0) \quad t \in [0, 1],$$

risolve la questione. Come si nota facilmente, si ha $\phi'(t) = P_1 - P_0 \neq 0$ per ogni t e, di conseguenza, la curva ϕ è effettivamente regolare.

Ovviamente, esistono (infinite) altre maniere di parametrizzare un segmento.

ESEMPIO 4.3. Il segmento nel piano di estremi $(0, 1)$ e $(1, 0)$ è parametrizzato da

$$\begin{cases} x = x(t) = t, \\ y = y(t) = 1 - t \end{cases} \quad t \in [0, 1].$$

Lo stesso segmento può essere parametrizzato in tante maniere diverse, ad esempio da

$$\begin{cases} x = x(\tau) = 1 - \cos(\tau), \\ y = y(\tau) = \cos(\tau) \end{cases} \quad \tau \in [0, \pi/2].$$

Le due rappresentazioni parametriche offerte differiscono dal punto di vista cinematico. Calcolando i corrispondenti vettori velocità si ha

$$\begin{aligned} (x', y')(t) = (1, -1) &\Rightarrow \sqrt{x'(t)^2 + y'(t)^2} = \sqrt{2}, \\ (x', y')(\tau) = \sin(\tau)(1, -1) &\Rightarrow \sqrt{x'(\tau)^2 + y'(\tau)^2} = \sqrt{2} \sin(\tau). \end{aligned}$$

Quindi, nel primo caso, il moto è rettilineo uniforme (i.e. la velocità è costante), nel secondo, il moto è ancora rettilineo, ma la velocità cresce al crescere di $\tau \in [0, \pi/2]$.

Consideriamo tre punti distinti $P_0, P_1, P_2 \in \mathbb{R}^d$. Come parametrizzare il percorso da P_0 a P_2 passando per P_1 , ottenuto percorrendo i segmenti $\overline{P_0P_1}$ e $\overline{P_1P_2}$? Basta operare piccole modifiche alla parametrizzazione vista per il caso di un singolo segmento:

$$\phi(t) = \begin{cases} P_0 + t(P_1 - P_0) & t \in [0, 1], \\ P_1 + (t - 1)(P_2 - P_1) & t \in (1, 2] \end{cases}$$

In questo caso,

$$\phi'(t) = \begin{cases} P_1 - P_0 & t \in [0, 1), \\ P_2 - P_1 & t \in (1, 2] \end{cases}$$

quindi, in generale, la parametrizzazione è solo regolare a tratti. In questo caso l'insieme S della Definizione 4.2 è composto da un singolo elemento: $S = \{1\}$.

Passiamo al caso generale di $n + 1$ punti.

DEFINIZIONE 4.4. *Assegnati $n + 1$ punti $P_0, P_1, P_2, \dots, P_n$ si dice poligonale S da essi determinata l'insieme unione degli n segmenti $\overline{P_0P_1}, \dots, \overline{P_{n-1}P_n}$. I punti $P_0, P_1, P_2, \dots, P_n$ sono detti vertici della poligonale e i due punti P_0 e P_n si dicono estremi della poligonale S .*

Parametrizzare una poligonale Π non è così proibitivo: basta porre

$$\phi(t) = P_k + (t - k)(P_{k+1} - P_k) \quad \forall t \in [k, k + 1], \quad k \in [0, n - 1].$$

Anche, in questo caso, in generale, la parametrizzazione è solo regolare a tratti con $S = \{1, 2, \dots, n\}$.

Altri esempi di curve. Il caso più noto è quello dei grafici di funzioni. Data $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ di classe C^1 , il grafico della funzione f si parametrizza banalmente ponendo

$$\phi(t) = (t, f(t)) \quad t \in I.$$

Dato che $\phi'(t) = (1, f'(t))$, la curva è regolare.

Non serve un grande sforzo per trovare qualche esempio di curva che non sia un grafico di funzione. Dati $P_0 = (x_0, y_0) \in \mathbb{R}^2$ ed $r > 0$, la circonferenza di centro P_0 e raggio r si rappresenta con

$$\begin{cases} x = x(t) = x_0 + r \cos(t), \\ y = y(t) = x_0 + r \sin(t) \end{cases} \quad t \in [0, 2\pi).$$

In maniera simile, dati $P_0 = (x_0, y_0) \in \mathbb{R}^2$ ed $a, b > 0$, l'ellissi di semiassi a e b , centrata in P_0 e simmetrica rispetto agli assi coordinati, si rappresenta parametricamente con

$$\begin{cases} x = x(t) = x_0 + a \cos(t), \\ y = y(t) = x_0 + b \sin(t) \end{cases} \quad t \in [0, 2\pi).$$

Una curva più originale è quella data da $\phi(t) = (t^2, t^3)$ con $t \in [-1, 1]$. Il grafico presenta una *cuspid*e nel punto $(0, 0)$. Si noti che, dato che $\phi'(t) = (2t, 3t^2)$, la condizione (67) non è soddisfatta in $t = 0$. La curva in esame, quindi, non è regolare. Si tratta, comunque, di una curva regolare a tratti, con $S = \{0\}$.

Tanto per gradire, vediamo anche un esempio di curva nello spazio tridimensionale. Dati $a > 0$ e $r > 0$, la curva parametrizzata data da

$$\begin{cases} x = x(t) = r \cos(t), \\ y = y(t) = r \sin(t), \\ z = z(t) = at \end{cases} \quad t \in \mathbb{R}.$$

ha come supporto un'elica circolare che si avvolge attorno all'asse z . A seconda della scelta di a , l'elica è più o meno ripida. Dato che $\phi'(t) = (-r \sin t, r \cos t, a) \neq 0$ per ogni t , la curva è regolare.

Lunghezza di curve. Il concetto di lunghezza è intuitivo: immaginando una curva come un oggetto deformabile, basta “raddrizzare” tale oggetto e confrontarlo con l’unità di misura prescelta. Ma, come nel caso degli integrali definiti, occorre dare una definizione rigorosa di lunghezza di una curva che, auspicabilmente, ne permetta anche il calcolo esplicito, almeno in un certo numero di situazioni semplici. Come procedere?

Nella costruzione dell’integrale, abbiamo ricondotto il calcolo di un’area con un bordo curvilineo a quello di insiemi dati da unioni di rettangoli, passando poi al limite in maniera opportuna. Nel caso di curve, l’idea naturale è di utilizzare come approssimanti le poligonal. In effetti, scegliendo un numero finito di punti della curva e considerando la poligonale determinata da tali punti, si ottiene una nuova curva la cui lunghezza è minore o uguale della lunghezza della curva originale in esame. La lunghezza della curva è data dall’estremo superiore delle lunghezze di tali poligonal approssimanti. Vediamo di formalizzare precisamente quanto descritto.

Consideriamo, una curva parametrizzata regolare ϕ definita in $I = [a, b] \subset \mathbb{R}$. Data la partizione $\pi = \{a = t_0 < t_1 < \dots < t_{n-1} < t_n = b\}$, in corrispondenza dei quali saranno assegnati sulla curva i punti $P_i = \phi(t_i)$. I punti P_0, P_1, \dots, P_n definiscono la poligonale $S(\pi)$ la cui lunghezza $\ell(\phi, \pi)$ è data da

$$\ell(\phi, \pi) = \sum_{k=0}^{n-1} |P_{k+1} - P_k|_d = \sum_{k=0}^{n-1} |\phi(t_{k+1}) - \phi(t_k)|_d.$$

DEFINIZIONE 4.5. La lunghezza ℓ di una curva parametrizzata regolare a tratti $\phi : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^d$ è data da

$$\ell(\phi) := \sup\{\ell(\phi, \pi) : \pi \text{ partizione di } [a, b]\}.$$

Le curve regolari hanno tutte lunghezza finita. Infatti, dato che la funzione ϕ appartiene a $C^1([a, b]; \mathbb{R}^d)$, le componenti ϕ_i sono di classe C^1 ed hanno quindi derivata limitata. Sia $M > 0$ tale che $|\phi'_i(t)| \leq M$ per ogni $t \in [a, b]$ e per ogni $i \in \{1, \dots, d\}$. Dati $t, \tau \in [a, b]$, applicando il Teorema di Lagrange, si deduce

$$|\phi(t) - \phi(\tau)|_d = \left(\sum_{k=1}^d (\phi_k(t) - \phi_k(\tau))^2 \right)^{1/2} \leq M\sqrt{d}|t - \tau|.$$

Data la partizione $\pi = \{a = t_0 < t_1 < \dots < t_{n-1} < t_n = b\}$, si ha

$$\ell(\phi, \pi) = \sum_{k=0}^{n-1} |\phi(t_{k+1}) - \phi(t_k)|_d \leq M\sqrt{d} \sum_{k=0}^{n-1} (t_{k+1} - t_k) = M\sqrt{d}(b - a).$$

L’insieme $\{\ell(\pi) : \pi \text{ partizione di } [a, b]\}$ è quindi superiormente limitato e, di conseguenza, ammette estremo superiore finito. Una curva regolare a tratti può essere

pensata come unione finita di curve regolari e, pertanto, ammette anch'essa lunghezza finita.

La definizione di lunghezza come estremo superiore è pienamente soddisfacente dal punto di vista teorico, ma lascia qualche perplessità dal punto di vista operativo: come calcolare in concreto la lunghezza di una curva?

TEOREMA 4.6. *Sia $\phi : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^d$ una curva parametrizzata regolare a tratti. Vale l'uguaglianza*

$$(68) \quad \ell(\phi) = \int_a^b |\phi'(t)|_a dt.$$

OSSERVAZIONE 4.7. Dato che la funzione ϕ è regolare a tratti e si ha

$$||\phi'(t)|_a - |\phi'(t_0)|_a| \leq |\phi'(t) - \phi'(t_0)|_a,$$

è possibile decomporre l'intervallo di integrazione $[a, b]$ in unione finita di sotto-intervalli chiusi in cui la funzione $|\phi'(\cdot)|_a$ è continua. Quindi, la funzione $|\phi'(\cdot)|_a$ è integrabile (secondo Riemann) in $[a, b]$ e l'integrale in (68) è ben definito.

La dimostrazione della relazione (68) si basa sull'approssimazione

$$\phi(\tau) - \phi(t) \approx \phi'(t) (\tau - t) \quad \tau \rightarrow t,$$

che consiste nel sostituire il segmento di estremi $\phi(t)$ e $\phi(t+h)$ con un segmento tangente al supporto della curva in $\phi(t)$ e di lunghezza opportuna. Il seguente Lemma si propone di stimare con precisione l'errore commesso in tale approssimazione, in un caso leggermente più semplice.

LEMMA 4.8. *Sia $I \subseteq \mathbb{R}$ un intervallo e sia $\phi \in C^1(I; \mathbb{R}^d)$ tale che ϕ' è lipschitziana in I . Allora, esiste $C > 0$ tale che*

$$(69) \quad |\phi(\tau) - \phi(t) - \phi'(t) (\tau - t)|_a \leq C |\tau - t|^2$$

per ogni $t, \tau \in I$.

DIMOSTRAZIONE. Per ogni $i \in \{1, \dots, d\}$, dal Teorema di Lagrange segue

$$|\phi_i(\tau) - \phi_i(t) - \phi'_i(t) (\tau - t)| = |(\phi_i(\sigma_i) - \phi'_i(t)) (\tau - t)|$$

per qualche σ_i compreso tra τ e t . Utilizzando l'ipotesi di lipschitzianità di ϕ' , si deduce che, per qualche $L > 0$, vale

$$(70) \quad |\phi_i(\tau) - \phi_i(t) - \phi'_i(t) (\tau - t)| \leq L |\sigma_i - t| |\tau - t| \leq L |\tau - t|^2.$$

Utilizzando (70), si ottiene

$$|\phi(\tau) - \phi(t) - \phi'(t) (\tau - t)|_a \leq L \sqrt{d} |\tau - t|^2.$$

La dimostrazione è completa. □

Grazie al Lemma 4.8, è possibile dimostrare il Teorema 4.6 supponendo, in aggiunta, ϕ' lipschitziana.

DIMOSTRAZIONE (SEMPLIFICATA) DEL TEOREMA 4.6. Data una partizione π dell'intervallo I , si ha, utilizzando (57) e (69),

$$\begin{aligned}
 (71) \quad \left| \ell(\phi, \pi) - \sum_{k=0}^{n-1} |\phi'(t_k)|_a (t_{k+1} - t_k) \right| &\leq \sum_{k=0}^{n-1} \left| |\phi(t_{k+1}) - \phi(t_k)|_a - |\phi'(t_k)|_a (t_{k+1} - t_k) \right| \\
 &\leq \sum_{k=0}^{n-1} \left| \phi(t_{k+1}) - \phi(t_k) - \phi'(t_k)(t_{k+1} - t_k) \right|_a \\
 &\leq C \sum_{k=0}^{n-1} (t_{k+1} - t_k)^2 \leq C(b-a)|\pi|,
 \end{aligned}$$

dove $|\pi| := \max_{k \in \{0, n-1\}} (t_{k+1} - t_k)$. Quando l'ampiezza $|\pi|$ della partizione tende a zero, la

lunghezza $\ell(\phi, \pi)$ converge alla lunghezza $\ell(\pi)$, mentre la sommatoria $\sum_{k=0}^{n-1} |\phi'(t_k)|_a (t_{k+1} - t_k)$ tende all'integrale della funzione $|\phi'(t)|_a$ rispetto a t nell'intervallo $[a, b]$. Grazie alla stima (71), ne segue la formula (68). \square

ESEMPIO 4.9. Come prima verifica/applicazione della formula (68), consideriamo la circonferenza parametrizzata da $\phi(t) = (\cos t, \sin t)$ con $t \in [0, 2\pi]$. In questo caso, si ha $\phi'(t) = (-\sin t, \cos t)$ e, di conseguenza, $|\phi'(t)|_a = 1$ per ogni t . Perciò, la lunghezza richiesta è

$$\ell = \int_0^{2\pi} 1 \, dt = 2\pi,$$

coerentemente con quanto noto dalla geometria elementare.

ESEMPIO 4.10. Calcoliamo la lunghezza della curva $\phi(t) = (t^2, t^3)$ con $t \in [-1, 1]$. Dato che $\phi'(t) = (2t, 3t^2)$, la lunghezza della curva è

$$L = \int_{-1}^1 \sqrt{4t^2 + 9t^4} \, dt = \frac{1}{9} \int_4^{13} \tau^{1/2} \, d\tau = \frac{2}{27} (13\sqrt{13} - 8).$$

avendo posto $\tau = 4 + 9t^2$.

ESEMPIO 4.11. Dati $a > 0$ e $r > 0$, consideriamo l'arco di elica descritto dalla parametrizzazione $\phi(t) = (r \cos t, r \sin t, at)$ con $t \in [t_0, t_1]$, $t_0 < t_1$. Dato che $\phi'(t) = (-r \sin t, r \cos t, a) \neq 0$ per ogni t , la lunghezza della curva è data da

$$\ell = \int_{t_0}^{t_1} |\phi'(t)|_a \, dt = \int_{t_0}^{t_1} \sqrt{r^2 + a^2} \, dt = \sqrt{r^2 + a^2} (t_1 - t_0).$$

dato che $|\phi'(t)|_a = \sqrt{r^2 + a^2}$.

Nel caso di parametrizzazione della forma $\phi(t) = (t, f(t))$ con $t \in [a, b]$, dato che $\phi(t) = (1, f'(t))$, l'espressione per la lunghezza della curva prende la forma

$$\ell = \int_a^b \sqrt{1 + (f'(t))^2} dt,$$

che, quindi, esprime la formula per la lunghezza di un grafico di funzione.

ESEMPIO 4.12 (Lunghezza di un arco di parabola). Sia $f(t) = t^2$ con $t \in [0, a]$ con $a > 0$. La lunghezza del grafico della funzione f è pari a

$$\begin{aligned} \ell &= \int_0^a \sqrt{1 + (f'(t))^2} dt = \int_0^a \sqrt{1 + 4t^2} dt = \frac{1}{2} \int_0^{\operatorname{settsinh}(2a)} \cosh^2 \tau dt \\ &= \frac{1}{2} \left(\operatorname{settsinh}(2a) + 2a \sqrt{1 + 4a^2} \right), \end{aligned}$$

avendo posto $\tau = \operatorname{settsinh}(2t)$.

ESEMPIO 4.13 (Lunghezza di un arco di curva esponenziale). Sia $f(t) = e^t$ con $t \in [0, a]$ con $a > 0$. In questo caso, la lunghezza del grafico della funzione f è data da

$$\begin{aligned} \ell &= \int_0^a \sqrt{1 + (f'(t))^2} dt = \int_0^a \sqrt{1 + e^{2t}} dt = \int_{\sqrt{2}}^{\sqrt{1+e^{2a}}} \left(1 + \frac{1}{\tau^2 - 1} \right) dt \\ &= \tau + \frac{1}{2} \ln \left(\frac{\tau - 1}{\tau + 1} \right) \Big|_{\sqrt{2}}^{\sqrt{1+e^{2a}}} = \sqrt{1 + e^{2a}} + a - \sqrt{2} + \ln \left(\frac{\sqrt{2} + 1}{\sqrt{1 + e^{2a}} + 1} \right) \end{aligned}$$

avendo posto $\tau = \sqrt{1 + e^{2t}}$ ed avendo applicato qualche trattamento cosmetico...

Cambiamenti di parametro. La lunghezza della curva è un concetto *intrinseco*, cioè non dipende dalla particolare scelta della parametrizzazione: uno stesso tragitto, anche se percorso con velocità diverse, indica sul contachilometri la stessa lunghezza di strada. Formalizziamo precisamente la questione.

DEFINIZIONE 4.14. Due curve regolari $\phi \in C^1([a, b]; \mathbb{R}^d)$ e $\psi \in C^1([\alpha, \beta]; \mathbb{R}^d)$ si dicono **equivalenti** se esiste una funzione reale di variabile reale $g \in C^1([\alpha, \beta]; [a, b])$ biunivoca tale che $g'(\tau) \neq 0$ per ogni τ e

$$(72) \quad \psi(\tau) = \phi(g(\tau)) \quad \forall \tau \in [\alpha, \beta].$$

In concreto, la funzione g consiste in una *riparametrizzazione* della curva: lo stesso supporto viene descritto sia dal parametro t che dal parametro τ e la funzione g definisce la corrispondenza tra i due parametri. In termini cinematici, le due funzioni ϕ e ψ corrispondono a due leggi orarie distinte e la funzione g mette in corrispondenza istanti relativi alla stessa posizione.

Derivando rispetto a τ la relazione (72), si deduce

$$\frac{d\psi}{d\tau} = \frac{d\phi}{dt} \frac{dg}{d\tau},$$

che descrive la relazione che intercorre tra i vettori velocità delle due curve. Utilizzando la formula 68 per la curva ψ ed introducendo la variabile $t = g(\tau)$, si ottiene

$$\ell(\psi) = \int_{\alpha}^{\beta} \left| \frac{d\psi}{d\tau} \right|_a d\tau = \int_{g(\alpha)}^{g(\beta)} \left| \frac{d\phi}{dt} \right|_a \frac{|g'|}{g'} dt$$

Dato che $g' \neq 0$ per ogni τ , la funzione g' o è sempre strettamente positiva o è sempre strettamente negativa. Nel primo caso, si ha $g(\alpha) = a$ e $g(\beta) = b$, quindi

$$\ell(\psi) = \int_a^b \left| \frac{d\phi}{dt} \right|_a dt = \ell(\phi).$$

Se $g' < 0$, si ha $g(\alpha) = b$ e $g(\beta) = a$, quindi

$$\ell(\psi) = \int_b^a \left| \frac{d\phi}{dt} \right|_a \frac{-g'}{g'} dt = - \int_b^a \left| \frac{d\phi}{dt} \right|_a dt = \int_a^b \left| \frac{d\phi}{dt} \right|_a dt = \ell(\phi).$$

Quindi, le formule precedenti mostrano che *curve equivalenti hanno la stessa lunghezza*. Oltre a dare ulteriore conforto e coerenza alla definizione di lunghezza espressa in (68), la precedente affermazione permette di scegliere la parametrizzazione che si preferisce nel calcolo della lunghezza di una curva.

Gli insiemi connessi. Una proprietà significativa degli insiemi di \mathbb{R}^d è quella di essere *connessi*, cioè composti da un singolo pezzo. Sinonimi del vocabolo “connesso” sono i termini “congiunto”, “allacciato”, perciò una maniera per formalizzare tale concetto è di richiedere che ogni coppia di punti dell’insieme possa essere collegata da un percorso interamente contenuto nell’insieme stesso. A seconda del tipo di percorsi ammissibili per realizzare tale collegamento tra punti, si utilizzano terminologie diverse.

La versione più semplice di tutte è quella che considera come unici percorsi ammissibili i segmenti.

DEFINIZIONE 4.15. *Un insieme $S \subset \mathbb{R}^d$ è convesso se, per ogni $P, Q \in S$, si ha*

$$tP + (1-t)Q \in S \quad \forall t \in [0, 1],$$

cioè se il segmento di estremi P e Q è interamente contenuto in S .

ESEMPIO 4.16. Ecco un piccolo elenco di insiemi convessi “facili”:

- gli intervalli in \mathbb{R} ;
- i sopragrafici di funzioni convesse, cioè gli insiemi della forma $S = \{(x, y) : y \geq f(x)\} \subset \mathbb{R}^2$, per $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ convessa;
- i rettangoli nel piano, cioè insiemi della forma $[a, b] \times [c, d] \subset \mathbb{R}^2$;

- i parallelepipedi nello spazio, cioè insiemi della forma $[a, b] \times [c, d] \times [e, f] \subset \mathbb{R}^3$;
- gli intorni di un punto, cioè insiemi della forma $\{P \in \mathbb{R}^d : |P - P_0|_d < r\}$ con $P_0 \in \mathbb{R}^d$ e $r > 0$.

OSSERVAZIONE 4.17. Se $S_1 \subset \mathbb{R}^{d_1}$ e $S_2 \subset \mathbb{R}^{d_2}$ sono due insiemi convessi, anche l'insieme $S := S_1 \times S_2 \subset \mathbb{R}^{d_1+d_2}$ è convesso. Infatti, siano $P = (x_P, y_P), Q = (x_Q, y_Q) \in \mathbb{R}^{d_1} \times \mathbb{R}^{d_2}$ due punti dell'insieme $S := S_1 \times S_2$. Dato che i punti x_P, x_Q appartengono ad S_1 , il punto $t x_P + (1-t) x_Q$ appartiene ad S_1 per ogni $t \in [0, 1]$. Analogamente, dato che i punti y_P, y_Q appartengono ad S_2 , anche il punto $t y_P + (1-t) y_Q$ è in S_2 . Di conseguenza, il punto $t P + (1-t) Q = (t x_P + (1-t) x_Q, t y_P + (1-t) y_Q)$ appartiene a $S_1 \times S_2$ per ogni $t \in [0, 1]$.

ESERCIZIO 4.18. Siano A e B due insiemi convessi di \mathbb{R}^n . L'insieme $A \cup B$ è convesso? L'insieme $A \cap B$ è convesso?

La convessità è una richiesta molto forte. Tanto forte che basta apportare piccole modifiche ad un insieme convesso per ottenere un insieme non convesso. Ad esempio, se $S \subseteq \mathbb{R}^d$ è un insieme convesso e P_0 è interno ad S , l'insieme $S \setminus \{P_0\}$ non è convesso. Infatti, se $r > 0$ è tale che $I_r(P_0) \subset S$ e $v \in \mathbb{R}^d$ è un qualsiasi vettore tale che $|v|_d < r$, i due punti $P_{\pm} := P_0 \pm v$ appartengono ad $S \setminus \{P_0\}$, ma, dato che $t P_- + (1-t) P_+ = P_0 + (1-2t)v$ coincide con P_0 per $t = 1/2$, il segmento che li congiunge non è interamente contenuto nello stesso insieme. In generale, tutto le volte che si crea un buco in un insieme convesso eliminando un sottoinsieme interno si ottiene un insieme non convesso.

Una condizione più flessibile della convessità è espressa dalla prossima definizione.

DEFINIZIONE 4.19. Un insieme S si dice **connesso per poligonalità** se per ogni coppia $P, Q \in S$ esiste una poligonale di estremi P e Q interamente contenuta in S .

Tutti gli insiemi convessi sono anche connessi per poligonalità. In questo caso, infatti, basta scegliere come poligonale il segmento che congiunge i due punti P e Q dell'insieme. Esistono anche tantissimi insiemi che sono connessi per poligonalità, ma che non sono convessi. Ad esempio, le corone circolari nel piano, i ferri di cavallo, e tutti gli insiemi ottenibili a partire da un insieme connesso con modificazioni elasto-plastiche, ma senza fratture...

ESEMPIO 4.20. Sia I un intervallo di \mathbb{R} e sia $f : I \rightarrow \mathbb{R}$ una funzione continua. Il sopragrafico $S := \{(x, y) \in \mathbb{R}^2 : y \geq f(x)\}$ è un insieme connesso per poligonalità. Infatti, dati $(x_1, y_1), (x_2, y_2) \in S$ con $x_1 \leq x_2$, sia $M := \max_{x \in [x_1, x_2]} f(x)$, che esiste finito per il Teorema di Weierstrass. Allora, la poligonale di vertici $P_0 = (x_1, y_1), P_1 = (x_1, M), P_2 = (x_2, M)$ e $P_3 = (x_2, y_2)$ è interamente contenuta in S .

Allo stesso modo, si mostra che il sopragrafico di una funzione superiormente limitata è connesso per poligonalità.

ESEMPIO 4.21. Un insieme convesso a cui siano stati sottratti un punto interno non è più convesso, ma è connesso per poligoni. Infatti, sia S l'insieme convesso di partenza, sia P_0 interno ad S . Se P, Q sono elementi di S , ci sono due eventualità: o il segmento di estremi P e Q non contiene il punto P_0 , o P_0 appartiene a tale segmento. Nel primo caso non c'è nulla da aggiungere. Nel secondo, occorre costruire una poligonale che connetta P e Q , aggirando il punto P_0 . Volete provvedere voi?

In generale, un insieme convesso a cui siano stati sottratti un numero finito di punti interni è un insieme connesso per poligoni.

ESERCIZIO 4.22. Trovare un insieme $S \subseteq \mathbb{R}^d$ convesso tale che esista $P_0 \in S$ (non interno ad S) per cui $S \setminus \{P_0\}$ non sia connesso per poligoni.

ESERCIZIO 4.23. Siano $A, B \subseteq \mathbb{R}^d$ due insiemi connessi per poligoni. L'insieme $A \cup B$ è convesso? L'insieme $A \cap B$ è convesso?

Esistono numerosi insiemi fatti di un solo pezzo che non sono connessi per poligoni: basta pensare ad un arco S di circonferenza. E' evidente che comunque si prendano due punti $P, Q \in S$ non esiste alcuna poligonale di estremi P e Q contenuta in S , per il semplice motivo che non esistono poligoni contenute in S . Per inserire anche questo tipo di situazione, si introduce un'ulteriore nuova definizione di connesso.

DEFINIZIONE 4.24. Un insieme S si dice **connesso per archi** se per ogni coppia $P, Q \in S$ esiste una curva $\phi : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^d$ regolare a tratti tale che

$$\phi(a) = P, \quad \phi(b) = Q, \quad \phi(t) \in S \quad \forall t \in [a, b].$$

Dato che le poligoni sono particolari scelte di curve regolari a tratti, ogni insieme connesso per poligoni è anche connesso per archi. Tutti gli archi di curva (che non siano poligoni) sono insiemi connessi per archi (ovviamente), ma non connessi per poligoni.

CAPITOLO 6

Funzioni di più variabili

1. Questioni elementari

Le funzioni sono algoritmi che fanno corrispondere ad ogni punto di un insieme (l'insieme di definizione) un altro punto di un altro insieme (il codominio). In quel che segue, ci interesseremo al caso in cui l'insieme di definizione I sia un sottoinsieme del dominio \mathbb{R}^d e il codominio sia l'insieme \mathbb{R}^p dove d, p sono numeri interi positivi opportuni. In notazione, consideriamo funzioni della forma seguente

$$f : I \subseteq \mathbb{R}^d \rightarrow \mathbb{R}^p.$$

Si tratta di una regola che associa alla variabile indipendente (o *input*) $P \in I$ la variabile dipendente (o *output*) $Q \in \mathbb{R}^p$. Passiamo in rassegna i tipi fondamentali di funzioni di questo genere.

Funzioni reali di una variabile reale: $d = p = 1$. Sono funzioni $f : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$, già stato considerate in lungo e in largo nei Capitoli precedenti.

Le curve: $d = 1, p > 1$. Si tratta di funzioni del tipo $\phi : I \subseteq \mathbb{R} \rightarrow \mathbb{R}^p$, di cui si è discusso nel Capitolo precedente.

Funzioni reali di più variabili reali: $d > 1, p = 1$. Si tratta di funzioni

$$f : I \subseteq \mathbb{R}^d \rightarrow \mathbb{R}.$$

Nel caso $d = 2$, si tratta di funzioni che fanno corrispondere ai punti $(x, y) \in \mathbb{R}^2$ un numero reale $z = f(x, y) \in \mathbb{R}$; nel caso $d = 3$, sono funzioni che associano ai punti $(x, y, z) \in \mathbb{R}^3$ un numero reale $w \in \mathbb{R}$.

In generale, una funzione f da \mathbb{R}^d in \mathbb{R}^p equivale all'assegnazione di p funzioni, indicate con f_1, \dots, f_p , da \mathbb{R}^d in \mathbb{R} che associano al punto $P \in \mathbb{R}^d$ la i -esima coordinata del punto immagine $f(P) = (f(P)_1, \dots, f(P)_p)$ con $i \in \{1, \dots, p\}$:

$$f_i(P) := f(P)_i \quad P \in I \subseteq \mathbb{R}^d.$$

Per questo motivo, le funzioni reali di più variabili reale sono particolarmente interessanti da analizzare e, in quel che segue, ci dedicheremo con particolare attenzione proprio a questa classe.

Le trasformazioni di \mathbb{R}^d in sé: $d = p \geq 1$. Nel caso $d = 2$, si tratta di trasformazioni del piano, cioè di funzioni che fanno corrispondere alla coppia $(x, y) \in \mathbb{R}^2$ una nuova coppia $(u, v) \in \mathbb{R}^2$. Una trasformazione del piano in sé può essere vista come una coppia di funzioni da $I \subseteq \mathbb{R}^2$ in \mathbb{R} :

$$(x, y) \in I \mapsto u = u(x, y) \in \mathbb{R}, \quad (x, y) \in I \mapsto v = v(x, y) \in \mathbb{R}.$$

In generale, si tratta di trasformazione che associano ad un punto $P = (x_1, \dots, x_d)$ un nuovo punto $Q = (y_1, \dots, y_d)$. Un'ampia classe di trasformazioni di \mathbb{R}^d è quella delle **trasformazioni lineari**, cioè funzioni della forma $f(P) := AP$, dove A è una matrice di dimensione $d \times d$.

Assegnata una funzione specifica, come procedere per riconoscerne le proprietà più significative? Nel caso di funzioni reali di una variabile reale, una strategia abbastanza generale consiste nel rappresentare l'andamento qualitativo della funzione disegnandone il grafico. Cosa cambia quando la dimensione aumenta? Per ora, accontentiamoci del caso di funzioni reali di più variabile reale.

Il primo passo è quello di determinare l'insieme di definizione della funzione, qualora non sia specificato esplicitamente. Una funzione è generalmente assegnata indicando un procedimento di calcolo che può essere applicabile ad alcuni punti di \mathbb{R}^d e non ad altri. Qualora non sia specificato diversamente, l'insieme dei punti P per cui il procedimento relativo alla funzione è ben definito si considera essere l'insieme di definizione della funzione. Ad esempio, consideriamo le funzioni

$$f(x, y) = \frac{1}{x + y}, \quad g(x, y) = \ln(1 - x^2 - y^2), \quad h(x, y) = \sqrt{3x^2 + 2y^2 - 5}.$$

La funzione f è definita in \mathbb{R}^2 privato della retta di equazione $x + y = 0$; la funzione g ha per insieme di definizione I il disco aperto di centro l'origine e raggio 1: $I := \{(x, y) : x^2 + y^2 < 1\}$; la funzione h è definita in tutto \mathbb{R}^2 privato dell'interno dell'ellisse di equazione $3x^2 + 2y^2 = 5$.

Una volta noti i punti per cui la funzione ha senso, si potrebbe essere interessati ad analizzare, in maniera qualitativa, l'andamento della funzione. Generalizzando la definizione già vista nel caso di funzioni di una variabile reale, si chiama **grafico della funzione f** l'insieme

$$\Gamma_f := \{(P, Q) \in \mathbb{R}^d \times \mathbb{R} : P \in I, Q = f(P)\}.$$

Nel caso $d = 2$, il grafico di una funzione $z = f(x, y)$ ha, quasi sempre l'aspetto di una superficie dello spazio. Se $d \geq 3$, il grafico è un sottoinsieme di uno spazio con dimensione maggiore o uguale a 4 e, di conseguenza, la sua visualizzazione grafica risulta sensibilmente più complicata... In quel che segue, ci concentriamo principalmente sul caso di funzioni da \mathbb{R}^2 in \mathbb{R} .

Profili altimetrici. Il grafico Γ_f di una funzione f da \mathbb{R}^2 in \mathbb{R} può essere immaginato come una superficie (di montagna o di pianura, a seconda della funzione) su cui si può liberamente passeggiare, trovandosi ad altezza $z = f(x, y)$ non appena le coordinate (cartesiane) della posizione siano date dalla coppia (x, y) .

Uno specifico cammino sul grafico Γ_f è quindi individuato da una legge oraria che associa alla variabile tempo t una corrispondente coppia di coordinate $(x(t), y(t))$. In corrispondenza, all'istante t , ci si trovare ad altezza $h(t) := f(x(t), y(t))$. Qualcosa suona familiare... la coppia $(x(t), y(t))$ è una curva a valori nell'insieme di definizione della funzione f e l'altezza h non è altro che la composizione tra la funzione $t \mapsto \phi(t) := (x(t), y(t))$ e la funzione $(x, y) \mapsto z = f(x, y)$. Per coerenza con il punto di vista escursionistico, chiamiamo nel seguito la funzione composta $h = f \circ \phi$ il **profilo altimetrico** di f relativo alla curva ϕ . Un profilo altimetrico fornisce sempre un'informazione parziale relativa alla struttura di un grafico di funzione, dato che si riferisce ad uno specifico cammino tra gli infiniti cammini possibili. Comunque, utilizzando l'informazione relativa ad un certo numero di profili altimetrici, si può sperare di ottenere un'idea qualitativa della struttura globale del grafico di una funzione.

ESEMPIO 1.1. Consideriamo la funzione $f(x, y) = x - y$, definita in tutto \mathbb{R}^2 . Proviamo ad esplorare il grafico di tale funzione considerando una curva della forma $\phi(t) = (t, 0)$ con $t \in \mathbb{R}$. Si tratta quindi di un'esplorazione che consiste nel tenere costante la variabile y e spostarsi nella direzione determinata dall'asse x . Il corrispondente profilo altimetrico è

$$h(t) = f(x(t), y(t)) = x(t) - y(t) = t.$$

Al crescere del tempo (e quindi al crescere della variabile x), l'altezza corrispondente cresce in maniera lineare (la funzione h è un polinomio di grado 1). In altre parole, l'intersezione tra il grafico della funzione f e il piano (x, z) nello spazio \mathbb{R}^3 è dato da una retta.

Se consideriamo la curva $\psi(\tau) = (0, \tau)$ con $\tau \in \mathbb{R}$, il profilo altimetrico è

$$h(\tau) = f(x(\tau), y(\tau)) = x(\tau) - y(\tau) = -\tau.$$

Anche in questo caso il profilo altimetrico è dato da un polinomio di grado 1. A differenza del caso precedente, al crescere di τ l'altezza diminuisce.

In generale, considerando una curva della forma della forma $\phi(t) = tv$ con $t \in \mathbb{R}$, dove $v = (v_1, v_2) \in \mathbb{R}^2$ è un vettore fissato, si ottiene un profilo altimetrico dato da un polinomio di primo grado

$$h(t) = f(x(t), y(t)) = v_1 t - v_2 t + 1 = (v_1 - v_2) t.$$

Tutte le sezioni del grafico della funzione con piani verticali, danno luogo a rette e, quindi, il grafico della funzione è un piano nello spazio.

ESEMPIO 1.2. Consideriamo la funzione $f(x, y) = x^2 + y^2$, definita in tutto \mathbb{R}^2 . Dato un vettore $v = (v_1, v_2) \in \mathbb{R}^2$ di norma unitaria, consideriamo la curva $\phi(t) = tv$ con $t \in \mathbb{R}$. Il corrispondente profilo altimetrico è

$$h(t) = f(x(t), y(t)) = v_1^2 t^2 + v_2^2 t^2 = (v_1^2 + v_2^2) t^2 = t^2$$

Si tratta cioè di una parabola con concavità rivolta verso l'alto e minimo nel punto $t = 0$. Come si vede, il profilo altimetrico non varia al variare della scelta del vettore unitario v , indice del fatto che il grafico della funzione considerata è *invariante per rotazioni* attorno all'asse z . Il grafico completo si ottiene quindi facendo ruotare il grafico della parabola $z = t^2$ attorno all'asse z ed ha quindi la forma di una scodella (a sezione parabolica) rivolta verso l'alto.

ESEMPIO 1.3. Piccola (ma non tanto...) variante dell'esempio precedente: studiamo la funzione $f(x, y) = x^2 - y^2$, definita in tutto \mathbb{R}^2 . Di nuovo, scegliamo come legge oraria dell'esplorazione la curva $\phi(t) = tv$ con $t \in \mathbb{R}$, dove $v = (v_1, v_2) \in \mathbb{R}^2$, $|v|_d = 1$. Si ha

$$h(t) = f(x(t), y(t)) = v_1^2 t^2 - v_2^2 t^2 = (v_1^2 - v_2^2) t^2.$$

In questo caso, a seconda della scelta di v il comportamento del profilo altimetrico varia. Precisamente:

- se $|v_1| > |v_2|$, il profilo è quello di una parabola con concavità verso l'alto e minimo in $t = 0$;
- se $|v_1| = |v_2|$, il profilo è identicamente nullo;
- se $|v_1| < |v_2|$, il profilo è quello di una parabola con concavità verso il basso e massimo in $t = 0$.

Inoltre, la convessità è massima per $v = (\pm 1, 0)$ e minima per $v = (0, \pm 1)$. Il grafico della funzione f ha la forma di una sella.

ESEMPIO 1.4. Studiamo la funzione

$$f(x, y) = \frac{xy}{x^2 + y^2} \quad (x, y) \neq (0, 0).$$

Consideriamo come curve di percorrenza della superficie, le rette per l'origine, espresse dalle funzioni $\phi(t) = tv$ con $t \in \mathbb{R}$, dove $v = (v_1, v_2) \in \mathbb{R}^2$, $|v|_d = 1$. Il profilo altimetrico è

$$h(t) = f(v_1 t, v_2 t) = \frac{v_1 v_2}{v_1^2 + v_2^2} = v_1 v_2.$$

L'altezza di ogni singolo profilo è costante, con altezza dipendente dalla retta scelta. Il grafico è rappresentato in Figura 1.

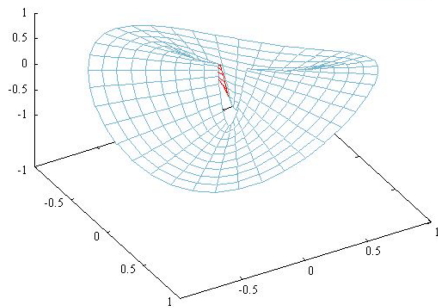


FIGURA 1. Il grafico della funzione $f(x, y) = \frac{xy}{x^2 + y^2}$.

I profili altimetrici possono essere considerati anche per funzioni di d variabili con $d \geq 2$. Date una funzione $f : I \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$ e una curva $\phi : J \subset \mathbb{R} \rightarrow \mathbb{R}^d$ tale che $\phi(J) \subset I$, la funzione composta $f \circ \phi : J \subseteq \mathbb{R} \rightarrow \mathbb{R}$ si chiama **profilo altimetrico** della funzione f lungo la curva ϕ .

Curve di livello. Tra i profili altimetrici di una funzione data, ne esistono alcuni particolarmente interessanti: sono quelli costanti. Una maniera alternativa per rappresentare una funzione $f = f(x, y)$ è quello di disegnare nel piano (x, y) gli insiemi di livello $\gamma_1 := \{(x, y) : f(x, y) = 1\}$, $\gamma_2 := \{(x, y) : f(x, y) = 2\}$, $\gamma_3 := \{(x, y) : f(x, y) = 3\}$,... e, in generale, $\gamma_c := \{(x, y) : f(x, y) = c\}$ con $c \in \mathbb{R}$. Tale metodo, detto **metodo delle linee di livello**, è quello utilizzato nelle carte geografiche¹.

Determinare insiemi di livello di una funzione è, in qualche modo, un procedimento inverso a quello dei profili altimetrici: data la costante c , si cercano tutti i percorsi possibili che diano come profilo altimetrico il valore costante c .

ESEMPIO 1.5. Consideriamo, di nuovo, la funzione $f(x, y) = x - y$. Dato il livello $c \in \mathbb{R}$, l'insieme $\gamma_c := \{(x, y) : f(x, y) = c\}$ è determinato da

$$x - y = c \quad \iff \quad y = x - c.$$

Si tratta quindi di rette parallele alla bisettrice del primo e terzo quadrante. Il fatto che variazioni di livello c uguali corrispondano a traslazioni uguali, si traduce nel fatto che la pendenza della superficie è sempre la stessa.

ESEMPIO 1.6. Gli insiemi di livello della funzione $f(x, y) = x^2 + y^2$ sono particolarmente semplici: data $c \geq 0$, si ha $\gamma_c := \{(x, y) : x^2 + y^2 = c\}$, si tratta cioè di

¹Per arricchire la rappresentazione delle linee di livello, in topografia, si utilizzano anche scale cromatiche che rappresentano i rilievi: toni di marrone via via più deciso per le catene montuose, toni di blu sempre più intenso per le profondità degli oceani e così via...

circonferenze centrate nell'origine e di raggio \sqrt{c} . Per $c < 0$, l'insieme γ_c è vuoto. Dalla struttura delle curve di livello si riconosce l'invarianza del grafico rispetto a rotazioni attorno all'asse z .

A differenza del caso precedente, a variazioni uguali del livello c non corrispondono variazioni uguali degli insiemi di livello: il raggio della circonferenza γ_c è \sqrt{c} , quindi il raggio cresce più rapidamente in prossimità del punto $(0, 0)$ e sempre più lentamente per valori di c grandi. Nella rappresentazione per curve di livello (che in concreto viene effettuata disegnando un numero finito di curve relative a livelli a distanza costante), si vedrà una minore concentrazione di curve vicino al punto $(0, 0)$. In corrispondenza, il grafico risulta essere meno pendente vicino all'origine e più pendente man mano che ci si allontana dal punto $(0, 0)$.

ESEMPIO 1.7. Consideriamo la funzione $f(x, y) = x^2 - y^2$. L'insieme di livello relativa al livello $c = 0$ è particolarmente semplice: dato che

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

si ha $\gamma_0 = \{y = x\} \cup \{y = -x\}$, cioè l'insieme di livello è dato dalle due bisettrici del piano (x, y) . Scegliendo $c \neq 0$, si ha $\gamma_c = \{x^2 - y^2 = c\}$, il cui grafico nel piano rappresenta una iperbole equilatera con asintoti dati dalle bisettrici $y = \pm x$. Se $c > 0$, l'insieme γ_c è contenuto nella regione $\{|y| < |x|\}$, mentre se $c < 0$, l'insieme γ_c è contenuto nella regione $\{|x| < |y|\}$.

Nel caso di funzioni di più variabili, la definizione di insieme di livello è del tutto analoga. In genere, dato un livello c , l'insieme di livello $\gamma_c = \{P \in \mathbb{R}^d : f(P) = c\}$ è un'unione di superfici di dimensione $d - 1$.

Funzioni con simmetrie. Disegnare il grafico di una funzione che possiede qualche forma di simmetria è, come al solito, più facile. Partiamo da un caso semplice semplice. Data la funzione reale di una variabile reale $F : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$, consideriamo la funzione di due variabili reali:

$$f(x, y) = F(x).$$

La funzione f è definita in $I \times \mathbb{R}$, infatti, dato che non c'è dipendenza esplicita dalla variabile y , il valore della seconda coordinata può essere qualsiasi. Inoltre, sempre per l'indipendenza della funzione f dalla variabile y , il grafico Γ_f risulta essere invariante rispetto a traslazioni nella direzione dell'asse y :

$$(73) \quad \Gamma_f + \lambda \mathbf{j} = \Gamma_f \quad \forall \lambda \in \mathbb{R}.$$

dove $\mathbf{j} = (0, 1)$. La proprietà (73) indica che per disegnare il grafico della funzione f basta disegnare il grafico della funzione F nel piano (x, z) e, poi, traslare parallelamente tale grafico nella direzione dell'asse y (vedi Figura 2). Il caso di funzioni della

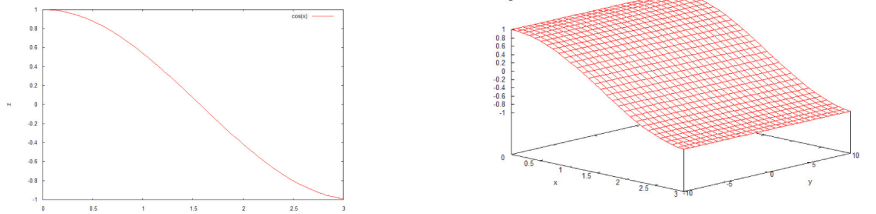


FIGURA 2. Il grafico della funzione $F(x) = \cos x$ e quella della funzione $f(x, y) = \cos x$.

forma $f(x, y) = F(y)$ è del tutto analogo, con l'unica differenza che l'invarianza è nella direzione dell'asse x .

Leggermente più in generale, data una funzione $F : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$ e date due costanti $a, b \in \mathbb{R}$, non entrambe nulle, si può considerare la funzione di due variabili reali:

$$f(x, y) = F(ax + by).$$

Le curve di livello di una funzione di questo genere sono (unioni di) rette della forma $ax + by = \text{costante}$. Di conseguenza, il grafico della funzione gode dell'invarianza seguente

$$(74) \quad \Gamma_f + \lambda v = \Gamma_f \quad \forall \lambda \in \mathbb{R}.$$

dove $v = (-b, a)$. Per disegnare il grafico della funzione f , basta disegnare il grafico della funzione F in uno qualsiasi dei piani verticali ortogonali alla direzione di invarianza e poi traslare parallelamente nella direzione di v . Sperimentate a vostro gradimento.

Consideriamo un secondo tipo di simmetria. Data $F : I \subseteq \mathbb{R} \rightarrow \mathbb{R}$, studiamo la funzione di due variabili reali definita nel modo seguente

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

Gli insiemi di livello di una funzione di questo genere sono (unioni di) circonferenze centrate in $(0, 0)$. Pertanto il grafico della funzione f risulta essere invariante rispetto a rotazioni attorno all'asse z . Formalmente, definendo

$$A_\theta := \begin{pmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{pmatrix}$$

la matrice che descrive la rotazione di angolo θ in senso antiorario del piano (x, y) , il grafico della funzione f gode della proprietà

$$(75) \quad A_\theta \Gamma_f = \Gamma_f \quad \forall \theta \in [0, 2\pi).$$

Concretamente, la proprietà (75) indica che per disegnare il grafico della funzione f , basta disegnare il grafico della funzione F nel semipiano (x, z) con $x \geq 0$ e, successivamente, far ruotare tale grafico attorno all'asse z (vedi Figura 3).

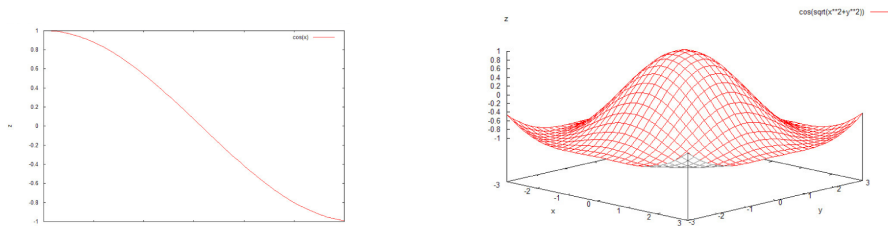


FIGURA 3. Il grafico delle funzioni $F(x) = \cos x$ e $f(x, y) = \cos(\sqrt{x^2 + y^2})$.

Operazioni elementari su grafici. A partire da grafici noti, è possibile dedurre la forma di altri grafici che si ottengono dai primi attraverso un certo numero di operazioni elementari. Qui, consideriamo il caso di dilatazioni/compressioni (o omotetie) e traslazioni.

Supponiamo di conoscere il grafico della funzione $f = f(x, y)$. Dato $\lambda > 0$, il grafico della funzione

$$g(x, y) = f(\lambda x, y)$$

si ottiene a partire dal grafico della funzione f applicando una compressione, nel caso $\lambda > 1$, o una dilatazione, nel caso $0 < \lambda < 1$, nella direzione dell'asse x .

ESEMPIO 1.8. A mo' d'esempio, si considerino i grafici delle funzioni

$$g(x, y) = 4x^2 + y^2, \quad h(x, y) = \frac{x^2}{4} + y^2,$$

che corrispondono, rispettivamente, ad una compressione e una dilatazione del grafico della funzione $f(x, y) = x^2 + y^2$, con fattori $\lambda = 2$ e $\lambda = 1/2$. Gli insiemi di livello delle funzioni g ed h relativi al livello $c = 1$, ad esempio, sono dati da ellissi di equazioni

$$\frac{x^2}{(1/2)^2} + y^2 = 1, \quad \frac{x^2}{2^2} + y^2 = 1,$$

quindi ellissi con assi di simmetria gli assi coordinati x e y , e semiassi di lunghezza $1/2$ e 1 , nel caso della funzione g , e di lunghezza 2 e 1 , nel caso della funzione f .

Considerazioni analoghe possono essere fatte per funzioni della forma

$$g(x, y) = f(\lambda_1 x, \lambda_2 y)$$

con $\lambda_1, \lambda_2 > 0$ dati. I fattori λ_1 e λ_2 danno luogo a dilatazioni e/o compressioni nelle direzioni degli assi principali, e gli insiemi di livello della funzione g si ottengono a partire da quelli della funzione f applicando le corrispondenti deformazioni.

Continuando a supporre il grafico della funzione f noto, consideriamo una funzione del tipo

$$g(x, y) = f(x + x_0, y + y_0)$$

con $(x_0, y_0) \in \mathbb{R}^2$ dato. Gli insiemi di livello della funzione g sono dati da

$$\begin{aligned} \gamma_c^g &= \{(x, y) : g(x, y) = c\} = \{(x, y) : f(x + x_0, y + y_0) = c\} \\ &= \{(\xi, \eta) : f(\xi, \eta) = c\} - (x_0, y_0) = \gamma_c^f - (x_0, y_0). \end{aligned}$$

avendo posto $(\xi, \eta) = (x + x_0, y + y_0)$. Le curve di livello della funzione g si ottengono per traslazione delle curve di livello della funzione f e, di conseguenza, il grafico della funzione g si ricava da quella della funzione f applicando un'analogha traslazione.

I casi considerati fin qui corrispondono a trasformazioni applicate alla variabile indipendente (x, y) . Analogamente, si può considerare il caso di trasformazioni applicate alla variabile dipendente. Al volenteroso lettore il compito di meditare sui casi

$$g(x, y) = \lambda f(x, y), \quad g(x, y) = f(x, y) + c$$

supponendo il grafico della funzione f noto e $\lambda > 0$, $c \in \mathbb{R}$ assegnati.

2. Le derivate parziali

Un possibile approccio allo studio di funzioni reale di più variabili reali è quello di analizzare i profili altimetrici relativi a percorsi diversi e cercare di “incollare” le informazioni parziali ottenute in questo modo. Il vantaggio di questo punto di vista è che studiare uno specifico profilo altimetrico, vuol dire, in concreto, studiare una funzione reale di variabile reale ed è quindi applicabile tutta la teoria e la tecnica vista nei Capitoli precedenti.

Ad esempio, data una funzione $f : I \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ ed un punto $P_0 = (x_0, y_0)$ interno ad I , può essere interessante studiare il comportamento della funzione f lungo le rette passanti per tale punto, ovvero analizzare i profili altimetrici

$$h(t) = f(x_0 + t v_1, y_0 + t v_2)$$

con $v = (v_1, v_2) \in \mathbb{R}^2$, $|v|_a = 1$, assegnato. Ad esempio, la monotonia lungo la retta per P_0 e di direzione v è determinata dal segno della derivata della funzione h . A tale derivata, ottenuta considerando la funzione lungo una retta fissata, si dà un nome ben preciso.

DEFINIZIONE 2.1. Sia $f : I \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$ e sia P_0 un punto interno ad I . Dato $v \in \mathbb{R}^d$, $|v|_d = 1$, si chiama derivata direzionale della funzione f nel punto P_0 rispetto alla direzione v il limite seguente

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

Scegliendo $v = \mathbf{e}_i$ per qualche i , indicando con $\mathbf{e}_i = (0, \dots, 0, 1, 0, \dots, 0)$ uno dei vettori della base canonica, la derivata direzionale si chiama derivata parziale della funzione f nel punto P_0 rispetto a x_i . Le derivate parziali si indicano, a seconda dei gusti, con i simboli

$$\frac{\partial f}{\partial x_i}, \quad f_{x_i}, \quad D_{x_i} f, \quad \partial_{x_i} f.$$

Si chiama vettore gradiente (o semplicemente gradiente) della funzione f , il vettore composto dalle derivate parziali prime

$$\nabla f := (f_{x_1}, \dots, f_{x_n}).$$

Il simbolo ∇ si legge *nabla*.

Il calcolo delle derivate parziali non comporta nessuna difficoltà aggiuntiva rispetto al calcolo delle derivate di funzioni di una variabile. Consideriamo infatti il caso di una funzione di due variabili $f = f(x, y)$. La derivata parziale rispetto alla variabile x è la derivata del profilo altimetrico ottenuto fissando la variabile y e facendo variare la sola variabile x . Pertanto, derivare rispetto ad x corrisponde a derivare la funzione $F(x) = f(x, y)$ dove y è considerato costante. Esempi:

$$\frac{\partial}{\partial x}(x + y^2) = 1, \quad \frac{\partial}{\partial x}(x y^2) = y^2, \quad \frac{\partial}{\partial x}(e^{x y^2}) = y^2 e^{x y^2}.$$

Analogamente:

$$\frac{\partial}{\partial y}(x + y^2) = 2y, \quad \frac{\partial}{\partial x}(x y^2) = 2xy, \quad \frac{\partial}{\partial x}(e^{x y^2}) = 2xy e^{x y^2}.$$

Il calcolo delle derivate direzionali a partire dalla definizione è meno immediato (vedremo più avanti una regola di calcolo, valida per una ampia classe di funzioni, che rende tale operazione molto più semplice). Ad esempio, calcoliamo le derivate direzionali della funzione $f(x, y) = x + y^2$. Dato $v \in \mathbb{R}^2$ con $|v|_2 = 1$, si ha

$$\frac{f(x + t v_1, y + t v_2) - f(x, y)}{t} = \frac{x + t v_1 + (y + t v_2)^2 - x - y^2}{t} = v_1 + 2 v_2 y + t v_2^2,$$

e, passando al limite per $t \rightarrow 0$, si ottiene

$$\frac{df}{dv} = v_1 + 2 v_2 y.$$

Nei casi particolari $v = (1, 0)$ e $v = (0, 1)$, si ottengono nuovamente le derivate parziali rispetto ad x e rispetto ad y .

Analogamente, per $g(x, y) = x y^2$, si ha

$$\begin{aligned} \frac{g(x + t v_1, y + t v_2) - g(x, y)}{t} &= \frac{(x + t v_1)(y + t v_2)^2 - x y^2}{t} \\ &= v_1 y^2 + 2 v_2 x y + (v_2^2 x + 2 v_1 v_2 y)t + v_1 v_2^2 t^2, \end{aligned}$$

da cui, passando al limite per $t \rightarrow 0$, segue

$$\frac{df}{dv} = v_1 y^2 + 2 v_2 x y.$$

Una conseguenza immediata della definizione di derivata direzionale e derivata direzionale è la seguente condizione necessaria per punti estremo relativo.

PROPOSIZIONE 2.2. *Sia $f : I \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$ e sia $P_0 \in I$, interno ad I . Se la funzione f ha un massimo (o un minimo) locale in P_0 , cioè se per qualche $r > 0$ si ha*

$$f(P) \leq (\geq) f(P_0) \quad \forall P \in I, |P - P_0|_d < r,$$

allora tutte le derivate direzionali di f in P_0 , qualora esistano, sono nulle.

DIMOSTRAZIONE. La dimostrazione è particolarmente semplice. Fissata una direzione $v \in \mathbb{R}^2$, $|v|_d = 1$, il profilo altimetrico $h(t) = f(P_0 + t v)$ descrive una funzione reale di variabile reale che ha, per ipotesi, un massimo (minimo) locale in $t = 0$. Di conseguenza, la derivata prima $h'(0)$, se esiste, è nulla. Dato che, per definizione, $h'(0)$ coincide con $\frac{df}{dv}(P_0)$, la proposizione è dimostrata. \square

Operativamente, la precedente Proposizione indica che, nella ricerca di massimi e minimi di una funzione data, occorre considerare come candidati possibili tutti i punti in cui le derivate direzionali, o semplicemente le derivate parziali (più semplici da calcolare!), siano nulle. Torneremo sulla questione tra qualche Capitolo.

Fissata una direzione $v \in \mathbb{R}^d$, la derivata direzionale rispetto alla direzione v è funzione del punto di derivazione e, di conseguenza, è essa stessa, qualora esista una funzione reale di più variabili reali. E' quindi possibile definire le derivate direzionali seconde, cioè le derivate direzionali di una derivata direzionale. Nel seguito, considereremo solamente il caso delle derivate parziali seconde, cioè delle derivate direzionali rispetto agli assi principali ed utilizzeremo notazioni coerenti con quelle utilizzate per le derivate parziali prime. Ad esempio, per le derivate seconde si usa comunemente scrivere

$$\frac{\partial^2 f}{\partial x_i \partial x_j}, \quad f_{x_i x_j}, \quad D_{x_i x_j}^2 f, \quad \partial_{x_i x_j} f.$$

Ad esempio, per le funzioni $x + y^2$ e $x y^2$, valgono, rispettivamente, le formule seguenti

$$\begin{aligned} \frac{\partial^2}{\partial x^2}(x + y^2) &= 0, & \frac{\partial^2}{\partial y \partial x}(x + y^2) &= 0, & \frac{\partial^2}{\partial x \partial y}(x + y^2) &= 0, & \frac{\partial^2}{\partial y^2}(x + y^2) &= 2, \\ \frac{\partial^2}{\partial x^2}(x y^2) &= 0, & \frac{\partial^2}{\partial y \partial x}(x y^2) &= 2 y, & \frac{\partial^2}{\partial x \partial y}(x y^2) &= 2 y, & \frac{\partial^2}{\partial y^2}(x y^2) &= 2 x. \end{aligned}$$

A voi, il gusto di calcolare tutte le derivate terze delle stesse due funzioni. Una funzione f da \mathbb{R}^d in \mathbb{R} possiede d derivate parziali prime, d^2 derivate seconde e, in generale, d^k derivate parziali k -esime.

DEFINIZIONE 2.3. *La matrice con elementi dati dalle derivate seconde della funzione $f : I \subseteq \mathbb{R}^d \rightarrow \mathbb{R}$*

$$d^2 f = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_d} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{\partial x_2^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_d \partial x_1} & \frac{\partial^2 f}{\partial x_d \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_d^2} \end{pmatrix}$$

si chiama *matrice hessiana* (o semplicemente *hessiano*) della funzione f .

Ad esempio, si hanno

$$d^2(x + y^2) = \begin{pmatrix} 0 & 0 \\ 0 & 2 \end{pmatrix} \quad \text{e} \quad d^2(x y^2) = \begin{pmatrix} 0 & 2 y \\ 2 y & 2 x \end{pmatrix}.$$

Come vedremo più avanti, la matrice hessiana dà informazioni relative alla curvatura del grafico della funzione f .

Guardare ad una funzione di più variabili studiandone profili altimetrici e/o sezioni con piani verticali fornisce informazioni rilevanti, ma resta, purtroppo, un approccio parziale. Ecco due esempi indicativi a riguardo.

ESEMPIO 2.4 (Una retta non basta!). Consideriamo la funzione

$$f(x, y) = \frac{xy}{x^2 + y^2} \quad (x, y) \neq (0, 0).$$

Dato che la funzione non è definita nell'origine, è interessante studiare cosa avvenga in prossimità di tale punto. Consideriamo come profilo altimetrico, quello determinato dall'asse x : data $\phi(t) = (t, 0)$, studiamo

$$h(t) = f(t, 0) = 0 \quad \forall t \neq 0.$$

La funzione è identicamente nulla lungo l'asse delle x e, quindi, tende a 0 lungo tale direzione. Cosa succede se si sceglie una direzione diversa? Dato $v = (v_1, v_2) \in \mathbb{R}^2$, $|v|_2 = 1$, lungo $\phi(t) = tv$, si ha

$$h(t) = f(tv_1, tv_2) = v_1 v_2 \quad \forall t \neq 0,$$

come già visto nell'Esempio 1.4. Lungo la direzione determinata dal vettore unitario v , la funzione f tende al valore $v_1 v_2$. Direzioni diverse danno luogo a limiti diversi.

ESEMPIO 2.5 (Tutte le rette non bastano!). Come secondo esempio, consideriamo la funzione

$$f(x, y) = \frac{x^2 y}{x^4 + y^2} \quad (x, y) \neq (0, 0).$$

Dato $v = (v_1, v_2) \in \mathbb{R}^2$, $|v|_2 = 1$, lungo $\phi(t) = tv$, il profilo altimetrico è dato da

$$h(t) = f(tv_1, tv_2) = \frac{t v_1^2 v_2}{t^2 v_1^4 + v_2^2}. \quad \forall t \neq 0,$$

Per $t \rightarrow 0$, si ha, per ogni scelta del vettore unitario v ,

$$\lim_{t \rightarrow 0} \frac{t v_1^2 v_2}{t^2 v_1^4 + v_2^2} = 0.$$

Lungo tutte le rette per l'origine, la funzione si avvicina al valore 0. Cosa succede lungo altri cammini che passano per l'origine? Consideriamo, ad esempio, la parabola $\phi(t) = (t, t^2)$: si ha

$$h(t) = f(t, t^2) = \frac{t^4}{t^4 + t^4} = \frac{1}{2}. \quad \forall t \neq 0;$$

e lungo le parabole la parabola $\phi(t) = (t, at^2)$, $a \in \mathbb{R}$

$$h(t) = f(t, at^2) = \frac{a^2 t^4}{t^4 + a^2 t^4} = \frac{a^2}{1 + a^2}. \quad \forall t \neq 0,$$

Quindi, percorsi diversi dai percorsi rettilinei portano a valori limite diversi.

La morale è semplice: una funzione di più variabili non può essere semplicemente considerata come l'incollamento delle sue sezioni planari o dei suoi profili altimetrici considerati separatamente. Una visione globale è indispensabile per avere una comprensione globale della funzione considerata. Tale punto di vista globale verrà analizzato a partire dal prossimo Capitolo.

3. Integrali curvilinei

Data una funzione $f : I \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ a valori positivi, cioè tale che $f(x, y) \geq 0$ per ogni (x, y) e data una curva $\phi : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^2$ tale che $\phi([a, b]) \subset I$, è ben definito il profilo altimetrico $f \circ \phi$. Consideriamo il sottoinsieme di $\mathbb{R}^2 \times \mathbb{R}$, delimitato dall'immagine del profilo $f \circ \phi$ e dal piano $z = 0$, cioè

$$\Sigma_{f,\phi} := \{(x, y, z) \in \mathbb{R}^3 : (x, y) = \phi(t), 0 \leq z \leq f(\phi(t)), t \in [a, b]\}.$$

L'insieme $\Sigma_{f,\phi}$ può essere immaginato come un nastro in \mathbb{R}^3 posato verticalmente sul piano $z = 0$: il supporto della curva ϕ determina la base d'appoggio del nastro ed il profilo altimetrico $f \circ \phi$ ne descrive lo spessore. Qual è l'area di tale insieme?

L'allenamento fatto nella costruzione dell'integrale di Riemann e ripreso successivamente nel calcolo delle lunghezze di curve suggerisce la possibile strategia... Data una partizione $\pi = \{a = t_0 < t_1 < \dots < t_n = b\}$ dell'intervallo $[a, b]$, consideriamo come approssimante dell'area cercata il valore:

$$\sum_{k=0}^{n-1} f(\phi(t_k)) |\phi'(t_k)|_2 (t_{k+1} - t_k).$$

Scegliendo partizioni π con ampiezza infinitesima, si ottiene formalmente l'integrale della funzione $f(\phi(t)) |\phi'(t)|_2$ nell'intervallo $[a, b]$. Dimentichiamo la richiesta di segno e passando a dimensione qualsiasi, si ottiene, in maniera naturale, la seguente definizione.

DEFINIZIONE 3.1. *Data una funzione $f : I \subset \mathbb{R}^d \rightarrow \mathbb{R}$ e una curva parametrizzata regolare a tratti $\phi : [a, b] \subset \mathbb{R} \rightarrow \mathbb{R}^d$ tale che $\phi([a, b]) \subset I$, si chiama integrale curvilineo di f lungo ϕ il valore dell'integrale definito (qualora esista)*

$$(76) \quad \int_{\phi} f ds := \int_a^b f(\phi(t)) |\phi'(t)|_d dt.$$

La richiesta che la funzione ϕ sia regolare a tratti, indica che la funzione $|\phi'(t)|_d$ è integrabile secondo Riemann, quindi, affinché l'integrale curvilineo (76) sia ben definito, è sufficiente che il profilo altimetrico $f \circ \phi$ determini una funzione integrabile secondo Riemann.

ESEMPIO 3.2. Consideriamo le funzioni $f(x, y) = x^2 + 4y^2$ e $\phi : [0, \pi/2] \rightarrow \mathbb{R}^2$ definita da $\phi(t) = (\cos t, \sin t)$, e calcoliamo l'integrale curvilineo di f lungo ϕ . Utilizzando la formula (76), si ottiene

$$\begin{aligned} \int_{\phi} f ds &= \int_0^{\pi/2} (\cos^2 t + 4 \sin^2 t) \sqrt{\sin^2 t + \cos^2 t} dt \\ &= \int_0^{\pi/2} (1 + 3 \sin^2 t) dt = \frac{1}{2} \pi + \frac{3}{4} \pi = \frac{5}{4} \pi. \end{aligned}$$

ESEMPIO 3.3. Consideriamo un caso in dimensione più alta. Sia $f(x, y, z) = y \sin z$ e consideriamo la curva ϕ definita da

$$\phi(t) = (\cos t, \sin t, t) \quad t \in [0, \pi].$$

Si ha

$$\int_{\phi} f ds = \int_0^{\pi} \sin t \sin t \sqrt{\sin^2 t + \cos^2 t + 1} dt = \sqrt{2} \int_0^{\pi} \sin^2 t dt = \frac{\sqrt{2}}{2} \pi.$$

Niente di speciale...

Non fatevi ingannare dai due esempi precedenti: in genere, il calcolo esplicito di un integrale curvilineo è difficile se non infattibile...

La costruzione che abbiamo visto dell'integrale curvilineo lascia sospettare che, cambiando parametrizzazione della curva, il valore dell'integrale curvilineo non cambi. La dimostrazione di tale proprietà è una semplice modifica di quella già vista nel caso della lunghezza di curve. A voi ripercorrerla implementando le modifiche del caso.

Come conseguenza del fatto che l'integrale curvilineo è, in realtà, un integrale di una funzione reale di variabile reale, esso gode delle stesse proprietà viste per l'integrale usuale. Eccole in rapida rassegna.

Linearità. Data una curva ϕ , due funzioni f e g e due costanti $\alpha, \beta \in \mathbb{R}$, se le funzioni f e g sono integrabili lungo ϕ , anche la funzione $\alpha f + \beta g$ è integrabile lungo ϕ e si ha

$$\int_{\phi} (\alpha f + \beta g) ds = \alpha \int_{\phi} f ds + \beta \int_{\phi} g ds$$

Monotonia. Data una curva ϕ e due funzioni f e g , integrabili lungo ϕ e tali che $f \leq g$,

$$\int_{\phi} f ds \leq \int_{\phi} g ds.$$

Additività. Data una curva $\phi : [a, b] \rightarrow \mathbb{R}^d$ ed una funzione f , integrabile lungo ϕ , per ogni $c \in (a, b)$, la funzione f risulta integrabile anche lungo ϕ_- e ϕ_+ , dove ϕ_{\pm} indicano le restrizioni di ϕ a $[a, c]$ e $[c, b]$ e si ha

$$\int_{\phi} f ds = \int_{\phi_-} f ds + \int_{\phi_+} f ds.$$