

Analisi Matematica I

Appunti delle lezioni tenute dal Prof. A. Fonda

Università di Trieste,

CdL Ingegneria Civile e Ambientale, Elettronica e Informatica

a.a. 2025/2026

I numeri naturali e il principio di induzione

Nel 1898 il matematico torinese Giuseppe Peano (1858–1932), nel suo articolo fondamentale *Arithmetices principia: nova methodo exposita* enunciò i seguenti assiomi per l'insieme $\mathbb{N}$ dei numeri naturali.

- a) Esiste un elemento, chiamato “zero”, indicato con 0.
- b) Ogni elemento n ha un ben determinato “successivo” n' .
- c) 0 non è il successivo di alcun elemento.
- d) Elementi diversi hanno successivi diversi.
- e) (**Principio di induzione**) Se S è un sottoinsieme di $\mathbb{N}$ tale che
 - i) $0 \in S$,
 - ii) $n \in S \Rightarrow n' \in S$,

allora $S = \mathbb{N}$.

È sottinteso che la condizione ii) deve valere per $n \in \mathbb{N}$ qualsiasi. Possiamo quindi leggerla in questo modo:

- ii) se per un certo n si ha che $n \in S$, ne consegue che anche $n' \in S$.

Si introducono i simboli $0' = 1$, $1' = 2$, $2' = 3$, ecc.

Da questi pochi assiomi, facendo uso della teoria degli insiemi, Peano ha mostrato come si possono ricavare tutte le proprietà dei numeri naturali. In particolare, si possono definire le operazioni di addizione e di moltiplicazione, ricavando l'uguaglianza

$$n' = n + 1.$$

Inoltre, scrivendo $m \leq n$ qualora esista un $p \in \mathbb{N}$ tale che $m + p = n$, si ottiene una relazione d'ordine. Supporremo qui ben note tutte le proprietà delle operazioni di addizione, moltiplicazione e della relazione d'ordine definite su $\mathbb{N}$.

Il principio di induzione può essere usato per definire una successione di oggetti

$$A_0, A_1, A_2, A_3, \dots$$

Si procede in questo modo (**definizione per ricorrenza**):

j) si definisce A_0 ;

jj) supponendo di aver definito A_n per un certo n , si definisce A_{n+1} .

In tal modo, se indichiamo con S l'insieme degli n per cui A_n è definita, si ha che S verifica *i*) e *ii*). Quindi S coincide con $\mathbb{N}$, ossia tutti gli A_n sono definiti.

Ad esempio, possiamo definire le “potenze” a^n ponendo, per $a \neq 0$,

j) $a^0 = 1$,

jj) $a^{n+1} = a \cdot a^n$.

Si vede in questo modo che $a^1 = a \cdot a^0 = a \cdot 1 = a$, $a^2 = a \cdot a^1 = a \cdot a$, e così via. Se $a = 0$, si pone $0^n = 0$ per ogni $n \geq 1$, mentre di solito resta non definito 0^0 . (Da ora in poi supporremo ben note le proprietà elementari delle potenze.)

Inoltre, definiamo il “fattoriale” $n!$ ponendo

j) $0! = 1$,

jj) $(n+1)! = (n+1) \cdot n!$.

Definiamo infine la somma (o sommatoria) di $\alpha_0, \alpha_1, \dots, \alpha_n$ utilizzando il simbolo

$$\sum_{k=0}^n \alpha_k.$$

Si pone $\sum_{k=0}^0 \alpha_k = \alpha_0$ e, una volta definito $\sum_{k=0}^n \alpha_k$ per un certo n , si pone

$$\sum_{k=0}^{n+1} \alpha_k = \sum_{k=0}^n \alpha_k + \alpha_{n+1}.$$

Nella notazione precedente compare un indice, qui indicato con k ; esso assume tutti i valori interi compresi tra 0 e n . Informalmente,

$$\sum_{k=0}^n \alpha_k = \alpha_0 + \alpha_1 + \alpha_2 + \dots + \alpha_n.$$

Il fatto che l'indice sia indicato con la lettera k è inessenziale; si può usare qualsiasi altra lettera o simbolo per denotarlo, come ad esempio

$$\sum_{j=0}^n \alpha_j, \quad \sum_{\ell=0}^n \alpha_\ell, \quad \sum_{m=0}^n \alpha_m, \quad \sum_{\star=0}^n \alpha_\star, \quad \dots$$

Si noti che la stessa somma potrebbe essere scritta come

$$\sum_{k=1}^{n+1} \alpha_{k-1}, \quad \sum_{k=2}^{n+2} \alpha_{k-2}, \quad \dots \quad \sum_{k=m}^{n+m} \alpha_{k-m},$$

e anche come

$$\sum_{k=0}^n \alpha_{n-k}.$$

Come si può vedere, ci sono molte varianti nell'uso del simbolo di sommatoria, e talvolta ne avremo bisogno nel seguito.

Il principio di induzione può inoltre essere usato per dimostrare una successione di proposizioni

$$P_0, P_1, P_2, P_3, \dots$$

Si procede in questo modo (**dimostrazione per induzione**):

j) si verifica P_0 ;

jj) supponendo vera P_n per un certo n , si verifica P_{n+1} .

Se indichiamo con S l'insieme degli n per cui P_n è dimostrata, si ha che S verifica *i*) e *ii*). Quindi S coincide con $\mathbb{N}$, ossia tutte le P_n sono dimostrate.

In questo modo si possono dimostrare le varie proprietà delle operazioni di addizione, moltiplicazione e delle potenze, nonché le proprietà della sommatoria, che supporremo da ora in poi note.

Esempio 1. Vogliamo dimostrare che, presi due numeri naturali a e n , si ha la seguente **disuguaglianza di Bernoulli**:

$$P_n : (1 + a)^n \geq 1 + na.$$

Vediamo che vale P_0 , essendo sicuramente $(1 + a)^0 \geq 1 + 0 \cdot a$. Supponiamo ora vera P_n per un certo n e verifichiamo P_{n+1} :

$$(1 + a)^{n+1} = (1 + a)^n(1 + a) \geq (1 + na)(1 + a) = 1 + (n + 1)a + na^2 \geq 1 + (n + 1)a,$$

per cui anche P_{n+1} è vera. Abbiamo così dimostrato che P_n è vera per ogni $n \in \mathbb{N}$.

Esempio 2. Dimostriamo la seguente uguaglianza:

$$P_n : a^{n+1} - b^{n+1} = (a - b) \left(\sum_{k=0}^n a^k b^{n-k} \right).$$

Vediamo P_0 :

$$a^{0+1} - b^{0+1} = (a - b)a^0b^{0-0};$$

essa è chiaramente vera. Supponiamo ora che P_n sia vera, per un certo $n \in \mathbb{N}$; allora

$$\begin{aligned}(a-b)\left(\sum_{k=0}^{n+1} a^k b^{n+1-k}\right) &= (a-b)\left(\sum_{k=0}^n a^k b^{n+1-k}\right) + (a-b)a^{n+1}b^0 \\ &= (a-b)\left(\sum_{k=0}^n a^k b^{n-k}\right)b + (a-b)a^{n+1} \\ &= (a^{n+1} - b^{n+1})b + (a-b)a^{n+1} \\ &= a^{n+2} - b^{n+2},\end{aligned}$$

per cui anche P_{n+1} è vera. Quindi, P_n è vera per ogni $n \in \mathbb{N}$.

In particolare, si ha:

$$\begin{aligned}a^2 - b^2 &= (a-b)(a+b), \\ a^3 - b^3 &= (a-b)(a^2 + ab + b^2), \\ a^4 - b^4 &= (a-b)(a^3 + a^2b + ab^2 + b^3), \\ a^5 - b^5 &= (a-b)(a^4 + a^3b + a^2b^2 + ab^3 + b^4), \\ &\dots\end{aligned}$$

Si noti anche la formula

$$\sum_{k=0}^n a^k = \frac{a^{n+1} - 1}{a - 1},$$

valida per $a \neq 1$, che si ottiene dalla precedente prendendo $b = 1$.

In alcuni casi potrebbe essere comodo iniziare la successione delle proposizioni, ad esempio, da P_1 invece che da P_0 , o da una qualsiasi altra di esse. Il principio di dimostrazione resta naturalmente lo stesso: se ne verifica la prima e si dimostra che da una qualsiasi di esse segue la successiva.

Altri esempi ed esercizi. Si possono dimostrare per induzione le seguenti formule:

$$\begin{aligned}1 + 2 + 3 + \dots + n &= \frac{n(n+1)}{2}, \\ 1^2 + 2^2 + 3^2 + \dots + n^2 &= \frac{n(n+1)(2n+1)}{6}, \\ 1^3 + 2^3 + 3^3 + \dots + n^3 &= \frac{n^2(n+1)^2}{4}.\end{aligned}$$

Si noti l'uguaglianza

$$1^3 + 2^3 + 3^3 + \dots + n^3 = (1 + 2 + 3 + \dots + n)^2.$$

Definiamo ora, per ogni coppia di numeri naturali n, k tali che $k \leq n$, i “coefficienti binomiali”

$$\binom{n}{k} = \frac{n!}{k!(n-k)!}.$$

Verifichiamo che, per $1 \leq k \leq n$, vale la formula

$$\binom{n}{k-1} + \binom{n}{k} = \binom{n+1}{k};$$

abbiamo infatti:

$$\begin{aligned} \binom{n}{k-1} + \binom{n}{k} &= \frac{n!}{(k-1)!(n-k+1)!} + \frac{n!}{k!(n-k)!} \\ &= \frac{n!k + n!(n-k+1)}{k!(n-k+1)!} \\ &= \frac{n!(n+1)}{k!(n-k+1)!} \\ &= \frac{(n+1)!}{k!((n+1)-k)!}. \end{aligned}$$

Dimostriamo ora che, per ogni $n \in \mathbb{N}$, vale la seguente **formula del binomio (di Newton)**:¹

$$P_n : \quad (a+b)^n = \sum_{k=0}^n \binom{n}{k} a^{n-k} b^k.$$

Iniziamo con il verificare che la formula vale per $n = 0$:

$$(a+b)^0 = \binom{0}{0} a^{0-0} b^0.$$

Per $n \geq 1$, procediamo per induzione. Vediamo che vale per $n = 1$:

$$(a+b)^1 = \binom{1}{0} a^{1-0} b^0 + \binom{1}{1} a^{1-1} b^1.$$

¹Anche in questa formula si supporrà che $a^0 = 1$, $b^0 = 1$ e $(a+b)^0 = 1$ anche nei casi in cui risultino del tipo 0^0 .

Ora, supponendo vera P_n , per un certo $n \geq 1$, vediamo che vale anche P_{n+1} :

$$\begin{aligned}
(a+b)^{n+1} &= (a+b)(a+b)^n \\
&= (a+b) \left(\sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \right) \\
&= a \left(\sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \right) + b \left(\sum_{k=0}^n \binom{n}{k} a^{n-k} b^k \right) \\
&= \sum_{k=0}^n \binom{n}{k} a^{n-k+1} b^k + \sum_{k=0}^n \binom{n}{k} a^{n-k} b^{k+1} \\
&= a^{n+1} + \sum_{k=1}^n \binom{n}{k} a^{n-k+1} b^k + \sum_{k=0}^{n-1} \binom{n}{k} a^{n-k} b^{k+1} + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^n \binom{n}{k} a^{n-k+1} b^k + \sum_{k=1}^n \binom{n}{k-1} a^{n-(k-1)} b^{(k-1)+1} + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^n \left[\binom{n}{k} + \binom{n}{k-1} \right] a^{n-k+1} b^k + b^{n+1} \\
&= a^{n+1} + \sum_{k=1}^n \binom{n+1}{k} a^{n-k+1} b^k + b^{n+1} \\
&= \sum_{k=0}^{n+1} \binom{n+1}{k} a^{n+1-k} b^k,
\end{aligned}$$

avendo usato la formula

$$\sum_{k=0}^{n-1} \alpha_k = \sum_{k=1}^n \alpha_{k-1}.$$

Abbiamo così dimostrato che P_n è vera per ogni $n \in \mathbb{N}$.

Ricordiamo che risulta talvolta utile rappresentare i coefficienti binomiali nel cosiddetto “triangolo di Tartaglia (o di Pascal)”

$$\begin{array}{ccccccc}
& & & \binom{0}{0} & & & \\
& & \binom{1}{0} & & \binom{1}{1} & & \\
& \binom{2}{0} & & \binom{2}{1} & & \binom{2}{2} & \\
& \binom{3}{0} & & \binom{3}{1} & & \binom{3}{2} & & \binom{3}{3} \\
& \binom{4}{0} & & \binom{4}{1} & & \binom{4}{2} & & \binom{4}{3} & & \binom{4}{4} \\
\binom{5}{0} & & \binom{5}{1} & & \binom{5}{2} & & \binom{5}{3} & & \binom{5}{4} & & \binom{5}{5} \\
\ldots & & \ldots & & \ldots & & \ldots & & \ldots & & \ldots
\end{array}$$

che possiamo scrivere esplicitamente così:

$$\begin{array}{ccccccc}
 & & & & 1 & & \\
 & & & & 1 & & 1 \\
 & & & 1 & 2 & 1 & \\
 & & 1 & 3 & 3 & 1 & \\
 & 1 & 4 & 6 & 4 & 1 & \\
 1 & 5 & 10 & 10 & 5 & 1 & \\
 \dots & \dots & \dots & \dots & \dots & \dots &
 \end{array}$$

Come casi particolari della formula del binomio, abbiamo quindi:

$$\begin{aligned}
 (a+b)^2 &= a^2 + 2ab + b^2, \\
 (a+b)^3 &= a^3 + 3a^2b + 3ab^2 + b^3, \\
 (a+b)^4 &= a^4 + 4a^3b + 6a^2b^2 + 4ab^3 + b^4, \\
 (a+b)^5 &= a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5, \\
 &\dots
 \end{aligned}$$

I numeri reali

Non ci soffermeremo sulle ragioni di carattere algebrico che portano, a partire dall'insieme dei numeri naturali

$$\mathbb{N} = \{0, 1, 2, 3, \dots\},$$

alla costruzione dell'insieme dei numeri interi relativi

$$\mathbb{Z} = \{\dots, -3, -2, -1, 0, 1, 2, 3, \dots\},$$

e dell'insieme dei numeri razionali

$$\mathbb{Q} = \left\{ \frac{m}{n} : m \in \mathbb{Z}, n \in \mathbb{N}, n \neq 0 \right\}.$$

Ci interessa però far notare che l'insieme dei numeri razionali non è sufficiente a trattare questioni geometriche elementari, quali ad esempio la misurazione della diagonale di un quadrato di lato 1.

Teorema. *Non esiste alcun numero razionale x tale che $x^2 = 2$.*

Dimostrazione. Per assurdo, supponiamo che esistano $m, n \in \mathbb{N}$ non nulli tali che

$$\left(\frac{m}{n}\right)^2 = 2,$$

ossia $m^2 = 2n^2$. Allora m deve essere pari, per cui esiste un $m_1 \in \mathbb{N}$ non nullo tale che $m = 2m_1$. Ne segue che $4m_1^2 = 2n^2$, ossia $2m_1^2 = n^2$. Pertanto anche n deve essere pari, per cui esiste un $n_1 \in \mathbb{N}$ non nullo tale che $2n_1 = n$. Quindi

$$\frac{m}{n} = \frac{m_1}{n_1} \quad \text{e} \quad \left(\frac{m_1}{n_1}\right)^2 = 2.$$

Possiamo ora ripetere lo stesso ragionamento quante volte vogliamo, continuando a dividere per 2 numeratore e denominatore:

$$\frac{m}{n} = \frac{m_1}{n_1} = \frac{m_2}{n_2} = \frac{m_3}{n_3} = \dots = \frac{m_k}{n_k} = \dots$$

dove m_k e n_k sono numeri naturali non nulli tali che $m = 2^k m_k$, $n = 2^k n_k$. Quindi, essendo $n_k \geq 1$, si ha che $n \geq 2^k$, per ogni numero naturale $k \geq 1$. In particolare, $n \geq 2^n$. Ma la disuguaglianza di Bernoulli ci dice che $2^n = (1+1)^n \geq 1+n$, e ne consegue che $n \geq 1+n$, il che è palesemente falso. ■

Si rende pertanto necessario estendere ulteriormente l'insieme dei numeri razionali.

È possibile costruire l'insieme dei numeri reali $\mathbb{R}$ a partire dai razionali. Essendo però tale costruzione piuttosto laboriosa, ci limiteremo qui ad enunciare le principali proprietà di $\mathbb{R}$.

1) È definita una “relazione d'ordine” $\leq$ con le seguenti proprietà:

per ogni scelta di x, y, z in $\mathbb{R}$,

- a) $x \leq x$,
- b) $[x \leq y \text{ e } y \leq x] \Rightarrow x = y$,
- c) $[x \leq y \text{ e } y \leq z] \Rightarrow x \leq z$;

inoltre, tale relazione d'ordine è “totale”:

- d) $x \leq y$ o $y \leq x$.

Se $x \leq y$, scriveremo anche $y \geq x$. Se $x \leq y$ e $y \neq x$, scriveremo $x < y$ oppure $y > x$.

2) È definita un'operazione di addizione $+$ con le seguenti proprietà:

per ogni scelta di x, y, z in $\mathbb{R}$,

- a) (associativa) $x + (y + z) = (x + y) + z$;
- b) esiste un “elemento neutro” 0 : si ha $x + 0 = x = 0 + x$;
- c) l'elemento x ha un “opposto” $-x$: si ha $x + (-x) = 0 = (-x) + x$;
- d) (commutativa) $x + y = y + x$;
- e) se $x \leq y$, allora $x + z \leq y + z$.

3) È definita un'operazione di moltiplicazione $\cdot$ con le seguenti proprietà:

per ogni scelta di x, y, z in $\mathbb{R}$,

- a) (associativa) $x \cdot (y \cdot z) = (x \cdot y) \cdot z$;
- b) esiste un “elemento neutro” 1 : si ha $x \cdot 1 = x = 1 \cdot x$;

- c) se $x \neq 0$, l'elemento x ha un "reciproco" x^{-1} : si ha $x \cdot x^{-1} = 1 = x^{-1} \cdot x$;
- d) (commutativa) $x \cdot y = y \cdot x$;
- e) se $x \leq y$ e $z \geq 0$, allora $x \cdot z \leq y \cdot z$;

e una proprietà che coinvolge entrambe le operazioni:

- f) (distributiva) $x \cdot (y + z) = (x \cdot y) + (x \cdot z)$;

4) (**Proprietà di separazione**) Dati due sottoinsiemi non vuoti A, B tali che

$$\forall a \in A \quad \forall b \in B \quad a \leq b,$$

esiste un elemento $c \in \mathbb{R}$ tale che

$$\forall a \in A \quad \forall b \in B \quad a \leq c \leq b.$$

Dalle proprietà elencate qui sopra si possono ricavare tutte le proprietà algebriche dei numeri reali, che supporremo già note.

Ritroviamo l'insieme $\mathbb{N}$ dei numeri naturali come sottoinsieme di $\mathbb{R}$: 0 e 1 sono gli elementi neutri di addizione e moltiplicazione, dopodiché si ha $2 = 1 + 1$, $3 = 2 + 1$ e così via, per ricorrenza.

Nel seguito, ometteremo quasi sempre il $\cdot$ nella moltiplicazione. Scriveremo, come è noto, $z = y - x$ se $z + x = y$, e $z = \frac{y}{x}$ se $zx = y$, con $x \neq 0$. In particolare, $x^{-1} = \frac{1}{x}$.

Le potenze a^n si definiscono come nella Sezione 1 per ogni $a \in \mathbb{R}$ e, se $a \neq 1$, continua a valere la formula per la somma delle potenze ivi dimostrata (Esempio 1 e sua generalizzazione). La disuguaglianza di Bernoulli risulta valida per ogni $a > -1$ e la formula del binomio di Newton continua a valere se a, b sono numeri reali qualsiasi.

Un sottoinsieme E di $\mathbb{R}$ si dice "limitato superiormente" se esiste un $\alpha \in \mathbb{R}$ tale che, per ogni $x \in E$, si ha $x \leq \alpha$; un tale α è allora una "limitazione superiore" di E . Se in più si ha che $\alpha \in E$, si dirà che α è il "massimo" di E e si scriverà $\alpha = \max E$.

Analogamente, E si dice "limitato inferiormente" se esiste un $\beta \in \mathbb{R}$ tale che, per ogni $x \in E$, si ha $x \geq \beta$; un tale β è allora una "limitazione inferiore" di E . Se in più si ha che $\beta \in E$, si dirà che β è il "minimo" di E e si scriverà $\beta = \min E$.

Diremo che E è "limitato" se è sia limitato superiormente che limitato inferiormente.

Teorema. *Se E è un sottoinsieme non vuoto di $\mathbb{R}$ limitato superiormente, l'insieme delle limitazioni superiori di E ha sempre un minimo.*

Dimostrazione. Sia B l'insieme delle limitazioni superiori di E . Allora

$$\forall a \in E \quad \forall b \in B \quad a \leq b,$$

e per la proprietà di separazione esiste un elemento $c \in \mathbb{R}$ tale che

$$\forall a \in E \quad \forall b \in B \quad a \leq c \leq b.$$

Ciò significa che c è una limitazione superiore di E , e quindi $c \in B$, ed è anche una limitazione inferiore di B . Pertanto, $c = \min B$. ■

Se E è limitato superiormente, la minima limitazione superiore di E si chiama “estremo superiore” di E : è un numero reale $s \in \mathbb{R}$ e si scrive $s = \sup E$. Esso è caratterizzato dalle seguenti proprietà:

- i) $\forall x \in E \quad x \leq s$,
- ii) $\forall s' < s \quad \exists x \in E : \quad x > s'$.

Se l'estremo superiore s appartiene ad E , si ha che $s = \max E$; succede spesso, però, che E , pur essendo limitato superiormente, non abbia un massimo. Talvolta le due proprietà si scrivono nella forma equivalente

- i) $\forall x \in E \quad x \leq s$,
- ii) $\forall \varepsilon > 0 \quad \exists x \in E : \quad x > s - \varepsilon$.

Nella seconda, si capisce che il numero $\varepsilon > 0$ può essere preso arbitrariamente piccolo.

Analogamente a quanto sopra, si può dimostrare il seguente

Teorema. *Se E è un sottoinsieme non vuoto di $\mathbb{R}$ limitato inferiormente, l'insieme delle limitazioni inferiori di E ha sempre un massimo.*

Se E è limitato inferiormente, la massima limitazione inferiore di E si chiama “estremo inferiore” di E : è un numero reale $i \in \mathbb{R}$ e si scrive $i = \inf E$. Esso è caratterizzato dalle seguenti proprietà:

- j) $\forall x \in E \quad x \geq i$,
- jj) $\forall i' > i \quad \exists x \in E : \quad x < i'$.

Se l'estremo inferiore i appartiene ad E , si ha che $i = \min E$; non è detto, però, che E , pur essendo limitato inferiormente, abbia un minimo. Le due proprietà si possono scrivere equivalentemente come

- j) $\forall x \in E \quad x \geq i$,
- jj) $\forall \varepsilon > 0 \quad \exists x \in E : \quad x < i + \varepsilon$.

Si noti che, definendo l'insieme

$$E^- = \{x \in \mathbb{R} : -x \in E\},$$

si ha che

$$E \text{ è limitato superiormente} \Leftrightarrow E^- \text{ è limitato inferiormente},$$

e in tal caso si ha che

$$\sup E = -\inf E^-,$$

mentre

$$E \text{ è limitato inferiormente} \Leftrightarrow E^- \text{ è limitato superiormente,}$$

e in tal caso si ha che

$$\inf E = -\sup E^-.$$

Nel caso in cui E non sia limitato superiormente, useremo la scrittura

$$\sup E = +\infty.$$

Teorema. $\sup \mathbb{N} = +\infty$.

Dimostrazione. Supponiamo per assurdo che $\mathbb{N}$ sia limitato superiormente, e sia $s = \sup \mathbb{N}$. Per le proprietà dell'estremo superiore, esiste un $n \in \mathbb{N}$ tale che $n > s - \frac{1}{2}$. Ma allora $n + 1 \in \mathbb{N}$ e

$$n + 1 > s - \frac{1}{2} + 1 > s,$$

in contraddizione col fatto che s è una limitazione superiore per $\mathbb{N}$. ■

Nel caso in cui E non sia limitato inferiormente, useremo la scrittura

$$\inf E = -\infty.$$

Ad esempio, si ha che $\inf \mathbb{Z} = -\infty$.

Ci sarà utile, anche in seguito, la seguente proprietà dei numeri reali.

Lemma. Se $0 \leq \alpha < \beta$, allora $\alpha^2 < \beta^2$.

Dimostrazione. Se $0 \leq \alpha < \beta$, si ha $\alpha^2 = \alpha\alpha \leq \alpha\beta < \beta\beta = \beta^2$. ■

Dimostreremo ora che esiste un numero reale $c > 0$ tale che $c^2 = 2$.

Definiamo gli insiemi

$$A = \{x \in \mathbb{R} : x \geq 0 \text{ e } x^2 < 2\},$$

$$B = \{x \in \mathbb{R} : x \geq 0 \text{ e } x^2 > 2\}.$$

Si può vedere che

$$\forall a \in A \quad \forall b \in B \quad a \leq b;$$

(altrimenti avremmo $0 \leq b < a$, quindi, per il Lemma, $b^2 < a^2$, mentre è $a^2 < 2$ e $b^2 > 2$, impossibile). Usando la proprietà di separazione, esiste un elemento $c \in \mathbb{R}$ tale che

$$\forall a \in A \quad \forall b \in B \quad a \leq c \leq b.$$

Si noti che, essendo $1 \in A$, sicuramente $c \geq 1$. Vogliamo ora mostrare che si ha proprio $c^2 = 2$.

Per assurdo, se $c^2 > 2$, allora, per $n \geq 1$,

$$\left(c - \frac{1}{n}\right)^2 = c^2 - \frac{2c}{n} + \frac{1}{n^2} \geq c^2 - \frac{2c}{n};$$

quindi, se $n > 2c/(c^2 - 2)$, essendo $c \geq 1$ e $n \geq 1$ si ha che

$$c - \frac{1}{n} \geq 0 \quad \text{e} \quad \left(c - \frac{1}{n}\right)^2 > 2,$$

per cui $c - \frac{1}{n} \in B$. Ma allora deve essere $c \leq c - \frac{1}{n}$, il che è impossibile.

Supponiamo ora, sempre per assurdo, che $c^2 < 2$. Allora, se $n \geq 1$,

$$\left(c + \frac{1}{n}\right)^2 = c^2 + \frac{2c}{n} + \frac{1}{n^2} \leq c^2 + \frac{2c}{n} + \frac{1}{n} = c^2 + \frac{2c+1}{n};$$

quindi, se $n > (2c+1)/(2-c^2)$, si ha che $(c + \frac{1}{n})^2 < 2$, e pertanto $c + \frac{1}{n} \in A$. Ma allora deve essere $c + \frac{1}{n} \leq c$, il che è impossibile.

Non potendo essere né $c^2 > 2$ né $c^2 < 2$, deve quindi essere $c^2 = 2$.

Il Lemma ci assicura inoltre che non ci possono essere altre soluzioni positive dell'equazione

$$x^2 = 2,$$

la quale pertanto ha esattamente due soluzioni, c e $-c$.

Lo stesso tipo di procedimento può essere usato per dimostrare che, qualunque sia il numero reale positivo r , esiste un unico numero reale positivo c tale che $c^2 = r$. Questo si chiama “radice quadrata” di r e si scrive $c = \sqrt{r}$. Si noti che l'equazione $x^2 = r$ ha due soluzioni: $x = \sqrt{r}$ e $x = -\sqrt{r}$. Si pone inoltre $\sqrt{0} = 0$, mentre la radice quadrata di un numero negativo resta non definita.

Studieremo ora la “densità” degli insiemi $\mathbb{Q}$ e $\mathbb{R} \setminus \mathbb{Q}$ nell'insieme dei numeri reali $\mathbb{R}$.

Teorema. *Dati due numeri reali α, β , con $\alpha < \beta$, esiste un numero razionale tra essi compreso.*

Dimostrazione. Consideriamo tre casi distinti.

Primo caso: $0 \leq \alpha < \beta$. Scegliamo $n \in \mathbb{N}$ tale che

$$n > \frac{1}{\beta - \alpha},$$

e sia $m \in \mathbb{N}$ il più grande numero naturale tale che

$$m < n\beta.$$

Quindi $\frac{m}{n} < \beta$, e resta da vedere che $\frac{m}{n} > \alpha$. Per assurdo, sia $\frac{m}{n} \leq \alpha$; allora

$$\frac{m+1}{n} \leq \alpha + \frac{1}{n} < \alpha + (\beta - \alpha) = \beta,$$

ossia $m+1 < n\beta$, in contraddizione col fatto che m è il più grande numero naturale minore di $n\beta$.

Secondo caso: $\alpha < 0 < \beta$. Basta scegliere il numero 0, che è razionale.

Terzo caso: $\alpha < \beta \leq 0$. Ci si può ricondurre al primo caso cambiando i segni: $0 \leq -\beta < -\alpha$, per cui esiste un razionale $\frac{m}{n}$ tale che $-\beta < \frac{m}{n} < -\alpha$. Allora $\alpha < -\frac{m}{n} < \beta$. ■

Teorema. *Dati due numeri reali α, β , con $\alpha < \beta$, esiste un numero irrazionale tra essi compreso.*

Dimostrazione. Per il teorema precedente, esiste un numero razionale $\frac{m}{n}$ tale che

$$\alpha + \sqrt{2} < \frac{m}{n} < \beta + \sqrt{2}.$$

Ne segue che

$$\alpha < \frac{m}{n} - \sqrt{2} < \beta,$$

con $\frac{m}{n} - \sqrt{2} \notin \mathbb{Q}$. ■

Scopriremo ora una sostanziale differenza tra gli insiemi $\mathbb{Q}$ e $\mathbb{R} \setminus \mathbb{Q}$. Consideriamo la seguente successione di numeri razionali non negativi:

$$\begin{array}{ccccccccccccccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & \dots \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\ \frac{0}{1} & \frac{1}{1} & \frac{1}{2} & \frac{2}{1} & \frac{1}{3} & \frac{2}{2} & \frac{3}{1} & \frac{1}{4} & \frac{2}{3} & \frac{3}{2} & \frac{4}{1} & \frac{1}{5} & \frac{2}{4} & \frac{3}{3} & \frac{4}{2} & \frac{5}{1} & \dots \end{array}$$

Come si vede, essa è costruita elencando i numeri razionali in cui la somma tra numeratore e denominatore è 1, poi 2, poi 3 e così via. Essa è sicuramente suriettiva, in quanto tutti i numeri razionali non negativi compaiono prima o poi nella lista. Possiamo ora modificarla per trovarne una biiettiva, eliminando i numeri che compaiono già in precedenza:

$$\begin{array}{ccccccccccccccccccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & 17 & \dots \\ \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \downarrow & \\ \frac{0}{1} & \frac{1}{1} & \frac{1}{2} & \frac{2}{1} & \frac{1}{3} & \frac{3}{1} & \frac{1}{4} & \frac{2}{3} & \frac{3}{2} & \frac{4}{1} & \frac{1}{5} & \frac{5}{1} & \frac{1}{6} & \frac{2}{5} & \frac{3}{4} & \frac{4}{3} & \frac{5}{2} & \frac{6}{1} & \dots \end{array}$$

A questo punto, è facile modificarla ancora per ottenere tutti i numeri razionali:

0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	...
↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	↓	...
$\frac{0}{1}$	$\frac{1}{1}$	$-\frac{1}{1}$	$\frac{1}{2}$	$-\frac{1}{2}$	$\frac{2}{1}$	$-\frac{2}{1}$	$\frac{1}{3}$	$-\frac{1}{3}$	$\frac{3}{1}$	$-\frac{3}{1}$	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{2}{3}$	$-\frac{2}{3}$	...

In questo modo, abbiamo costruito una funzione $\varphi : \mathbb{N} \rightarrow \mathbb{Q}$ biiettiva. Diremo quindi che $\mathbb{Q}$ è un insieme “numerabile”.

Si può invece dimostrare che $\mathbb{R} \setminus \mathbb{Q}$ non è numerabile.

Intervalli

Chiamiamo “intervallo” un sottoinsieme non vuoto I di $\mathbb{R}$ con la seguente proprietà: comunque presi due suoi elementi α, β , l'insieme I contiene anche tutti i numeri tra essi compresi.

Si può dimostrare che gli intervalli sono di uno dei seguenti tipi, con relativa notazione:

$[a, b] = \{x : a \leq x \leq b\}$	(chiuso e limitato),
$]a, b[= \{x : a < x < b\}$	(aperto e limitato),
$[a, b[= \{x : a \leq x < b\}$	(né aperto né chiuso, limitato),
$]a, b] = \{x : a < x \leq b\}$	(né aperto né chiuso, limitato),
$[a, +\infty[= \{x : x \geq a\}$	(chiuso, non limitato superiormente),
$]a, +\infty[= \{x : x > a\}$	(aperto, non limitato superiormente),
$] - \infty, b] = \{x : x \leq b\}$	(chiuso, non limitato inferiormente),
$] - \infty, b[= \{x : x < b\}$	(aperto, non limitato inferiormente),
$\mathbb{R}$, talvolta denotato con $] - \infty, +\infty[$.	

Nella lista si possono anche includere gli insiemi costituiti da un unico punto, cioè del tipo $[a, a]$. In tal caso, si tratta di un intervallo degenere.

Teorema (di Cantor). *Data una successione di intervalli chiusi e limitati $I_n = [a_n, b_n]$, con $a_n \leq b_n$, tali che*

$$I_0 \supseteq I_1 \supseteq I_2 \supseteq I_3 \supseteq \dots,$$

esiste un elemento $c \in \mathbb{R}$ che appartiene a tutti gli I_n .

Dimostrazione. Definiamo gli insiemi

$$A = \{a_n : n \in \mathbb{N}\},$$
$$B = \{b_n : n \in \mathbb{N}\}.$$

Preso un elemento a_n di A e un elemento b_m di B (non necessariamente con lo stesso indice), vediamo che $a_n \leq b_m$. Infatti, se $n \leq m$, allora $I_n \supseteq I_m$, per cui $a_n \leq a_m \leq b_m \leq b_n$. Se invece $n \geq m$, si ha $I_m \supseteq I_n$, per cui $a_m \leq a_n \leq b_n \leq b_m$. In ogni caso, $a_n \leq b_m$. Possiamo quindi usare la proprietà di separazione, e troviamo un $c \in \mathbb{R}$ tale che

$$\forall a \in A \quad \forall b \in B \quad a \leq c \leq b.$$

In particolare, $a_n \leq c \leq b_n$, cioè $c \in I_n$, per ogni $n \in \mathbb{N}$. ■

I numeri complessi

Consideriamo l'insieme

$$\mathbb{R} \times \mathbb{R} = \{(a, b) : a \in \mathbb{R}, b \in \mathbb{R}\},$$

che spesso si indica con $\mathbb{R}^2$. Definiamo un'operazione di “addizione”:

$$(a, b) + (a', b') = (a + a', b + b').$$

Si verificano le seguenti proprietà:

- a) (associativa) $(a, b) + ((a', b') + (a'', b'')) = ((a, b) + (a', b')) + (a'', b'');$
- b) esiste un “elemento neutro” $(0, 0)$: si ha $(a, b) + (0, 0) = (a, b)$;
- c) ogni elemento (a, b) ha un “opposto” $-(a, b) = (-a, -b)$: si ha

$$(a, b) + (-a, -b) = (0, 0);$$

- d) (commutativa) $(a, b) + (a', b') = (a', b') + (a, b)$;

Definiamo un'operazione di “moltiplicazione”:

$$(a, b) \cdot (a', b') = (aa' - bb', ab' + ba').$$

Si può verificare che valgono le seguenti proprietà:

- a) (associativa) $(a, b) \cdot ((a', b') \cdot (a'', b'')) = ((a, b) \cdot (a', b')) \cdot (a'', b'');$
- b) esiste un “elemento neutro” $(1, 0)$: si ha $(a, b) \cdot (1, 0) = (a, b)$;
- c) ogni elemento $(a, b) \neq (0, 0)$ ha un “reciproco” $(a, b)^{-1} = (\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2})$: si ha

$$(a, b) \left(\frac{a}{a^2 + b^2}, \frac{-b}{a^2 + b^2} \right) = (1, 0);$$

- d) (commutativa) $(a, b) \cdot (a', b') = (a', b') \cdot (a, b)$;
- e) (distributiva) $(a, b) \cdot ((a', b') + (a'', b'')) = ((a, b) \cdot (a', b')) + ((a, b) \cdot (a'', b''))$.

(Nel seguito, ometteremo spesso di scrivere il “ $\cdot$ ”). In questo modo, $(\mathbb{R}^2, +, \cdot)$ risulta essere un campo, che verrà indicato con $\mathbb{C}$ e si dirà il “campo complesso”. I suoi elementi si chiameranno “numeri complessi”.

Si può pensare $\mathbb{C}$ come un’estensione di $\mathbb{R}$ in questo modo: si identificano tutti gli elementi della forma $(a, 0)$ con il corrispondente numero reale a . Le operazioni di somma e moltiplicazione indotte su $\mathbb{R}$ sono effettivamente quelle preesistenti:

$$\begin{aligned}(a, 0) + (a', 0) &= (a + a', 0), \\ (a, 0) \cdot (a', 0) &= (aa', 0).\end{aligned}$$

Notiamo che vale la seguente uguaglianza:

$$(a, b) = (a, 0) + (0, 1)(b, 0).$$

È allora conveniente introdurre un nuovo simbolo per indicare l’elemento $(0, 1)$. Scriveremo

$$(0, 1) = i.$$

In questo modo, avendo identificato $(a, 0)$ con a e $(b, 0)$ con b , possiamo scrivere

$$(a, b) = a + ib.$$

Posto $z = a + ib$, il numero a si dice “parte reale” di z e si scrive $a = \operatorname{Re}(z)$. Il numero b si dice “parte immaginaria” di z e si scrive $b = \operatorname{Im}(z)$.

Osserviamo ora che si ha

$$i^2 = (0, 1)(0, 1) = (-1, 0) = -1.$$

Usando questa semplice informazione, possiamo verificare che valgono le usuali proprietà simboliche formali: ad esempio,

$$(a + ib) + (a' + ib') = (a + a') + i(b + b').$$

$$(a + ib)(a' + ib') = (aa' - bb') + i(ab' + ba').$$

Sia $z = a + ib$ un numero complesso fissato. Cerchiamo le soluzioni in $\mathbb{C}$ dell’equazione

$$u^2 = z.$$

Queste vengono talvolta dette “radici quadrate” del numero complesso z (attenzione però a non confonderle con la radice quadrata di un numero reale non negativo). Se $b = 0$, ho

$$u = \begin{cases} \pm\sqrt{a} & \text{se } a \geq 0, \\ \pm i\sqrt{-a} & \text{se } a < 0. \end{cases}$$

Se invece $b \neq 0$, scriviamo $u = x + iy$. Allora

$$x^2 - y^2 = a, \quad 2xy = b.$$

Essendo $b \neq 0$, si ha $x \neq 0$ e $y \neq 0$. Posso quindi scrivere $y = \frac{b}{2x}$, e ottengo

$$x^4 - ax^2 - \frac{b^2}{4} = 0,$$

da cui

$$x^2 = \frac{a + \sqrt{a^2 + b^2}}{2}.$$

Determinati così x e y , abbiamo due soluzioni della nostra equazione:

$$u = \pm \left[\sqrt{\frac{a + \sqrt{a^2 + b^2}}{2}} + i \frac{b}{\sqrt{2(a + \sqrt{a^2 + b^2})}} \right].$$

Possiamo ora considerare un'equazione del secondo grado

$$Au^2 + Bu + C = 0,$$

dove A, B, C sono numeri complessi fissati, con $A \neq 0$. Come facilmente si vede, l'equazione è equivalente a

$$\left(u + \frac{B}{2A}\right)^2 = \frac{B^2 - 4AC}{(2A)^2}.$$

Ponendo $v = u + \frac{B}{2A}$ e $z = \frac{B^2 - 4AC}{(2A)^2}$, ci si riconduce al problema delle radici quadrate che abbiamo già risolto.

Per concludere, consideriamo l'equazione polinomiale più generale

$$A_n u^n + A_{n-1} u^{n-1} + \dots + A_1 u + A_0 = 0,$$

dove $A_0, A_1, \dots, A_n$ sono numeri complessi fissati, con $A_n \neq 0$. In altri termini, vogliamo trovare le radici di un polinomio a coefficienti complessi. Il seguente teorema, che enunciamo senza dimostrazione, è noto come **teorema fondamentale dell'algebra**.

Teorema. *Ogni equazione polinomiale ha, nel campo complesso, almeno una soluzione.*

Il problema di trovare una formula generale che fornisca le soluzioni è però tutt'altro che facile. Lo abbiamo affrontato nel caso $n = 2$ e si può risolvere anche se $n = 3$ o 4 . Se $n \geq 5$, però, è stato dimostrato che non esiste alcuna formula algebrica generale che fornisca una radice del polinomio.

Introduciamo ora alcune nozioni associate ai numeri complessi. Se $z = a + ib$, si definisce il “modulo” di z :

$$|z| = \sqrt{a^2 + b^2},$$

Si noti che, se $z = a \in \mathbb{R}$, ritroviamo il “valore assoluto”

$$|a| = \sqrt{a^2} = \begin{cases} a & \text{se } a \geq 0, \\ -a & \text{se } a < 0. \end{cases}$$

Dati due numeri complessi z_1 e z_2 , verifichiamo che

$$|z_1 z_2| = |z_1| |z_2|.$$

Infatti, se $z_1 = a_1 + ib_1$ e $z_2 = a_2 + ib_2$, si ha

$$\begin{aligned} |z_1 z_2|^2 &= (a_1 a_2 - b_1 b_2)^2 + (a_1 b_2 + b_1 a_2)^2 \\ &= a_1^2 a_2^2 - 2a_1 a_2 b_1 b_2 + b_1^2 b_2^2 + a_1^2 b_2^2 + 2a_1 b_2 b_1 a_2 + b_1^2 a_2^2 \\ &= a_1^2 a_2^2 + b_1^2 b_2^2 + a_1^2 b_2^2 + b_1^2 a_2^2 \\ &= (a_1^2 + b_1^2)(a_2^2 + b_2^2) \\ &= |z_1|^2 |z_2|^2. \end{aligned}$$

In particolare, se i due numeri coincidono, si ha

$$|z^2| = |z|^2.$$

Ne segue per induzione che, per $n \in \mathbb{N}$,

$$|z^n| = |z|^n.$$

Inoltre, se $z \neq 0$, essendo $|z^{-1}z| = 1$, si ha

$$|z^{-1}| = |z|^{-1}.$$

Ecco allora che, preso un intero positivo n ,

$$|z^{-n}| = |(z^{-1})^n| = |z^{-1}|^n = (|z|^{-1})^n = |z|^{-n}.$$

Pertanto, l'uguaglianza $|z^n| = |z|^n$ vale per ogni $n \in \mathbb{Z}$.

Dato un numero complesso $z = a + ib$, si introduce il numero $z^* = a - ib$, detto il “complesso coniugato” di z . Valgono le seguenti proprietà:

$$\begin{aligned}(z_1 + z_2)^* &= z_1^* + z_2^*; \\ (z_1 z_2)^* &= z_1^* z_2^*; \\ z^{**} &= z; \\ |z^*| &= |z|; \\ z z^* &= |z|^2; \\ \operatorname{Re}(z) &= \frac{1}{2}(z + z^*), \quad \operatorname{Im}(z) = \frac{1}{2i}(z - z^*); \\ |\operatorname{Re}(z)| &\leq |z|, \quad |\operatorname{Im}(z)| \leq |z|.\end{aligned}$$

Se $z \neq 0$, è

$$z^{-1} = \frac{z^*}{|z|^2}.$$

Dimostriamo ora che vale la seguente “disuguaglianza triangolare”:

$$|z_1 + z_2| \leq |z_1| + |z_2|.$$

Infatti, si ha che

$$\begin{aligned}|z_1 + z_2|^2 &= (z_1 + z_2)(z_1 + z_2)^* \\ &= (z_1 + z_2)(z_1^* + z_2^*) \\ &= z_1 z_1^* + z_1 z_2^* + z_2 z_1^* + z_2 z_2^* \\ &= |z_1|^2 + z_1 z_2^* + (z_1 z_2^*)^* + |z_2|^2 \\ &= |z_1|^2 + 2\operatorname{Re}(z_1 z_2^*) + |z_2|^2 \\ &\leq |z_1|^2 + 2|z_1 z_2^*| + |z_2|^2 \\ &= |z_1|^2 + 2|z_1| |z_2^*| + |z_2|^2 \\ &= |z_1|^2 + 2|z_1| |z_2| + |z_2|^2 \\ &= (|z_1| + |z_2|)^2,\end{aligned}$$

e la disuguaglianza cercata segue dal Lemma di pagina 10.

Lo spazio $\mathbb{R}^N$

Consideriamo l'insieme $\mathbb{R}^N$, costituito dalle N -uple $(x_1, x_2, \dots, x_N)$, dove $x_1, x_2, \dots, x_N$ sono numeri reali. Indicheremo i suoi elementi con i simboli

$$\mathbf{x}, \mathbf{x}', \mathbf{x}'', \dots$$

Cominciamo con l'introdurre un'operazione di addizione in $\mathbb{R}^N$: dati due elementi $\mathbf{x} = (x_1, x_2, \dots, x_N)$ e $\mathbf{x}' = (x'_1, x'_2, \dots, x'_N)$, si definisce $\mathbf{x} + \mathbf{x}'$ in questo modo:

$$\mathbf{x} + \mathbf{x}' = (x_1 + x'_1, x_2 + x'_2, \dots, x_N + x'_N).$$

Valgono le seguenti proprietà:

- a) (associativa) $(\mathbf{x} + \mathbf{x}') + \mathbf{x}'' = \mathbf{x} + (\mathbf{x}' + \mathbf{x}'')$;
- b) esiste un “elemento neutro” $\mathbf{0} = (0, 0, \dots, 0)$: si ha $\mathbf{x} + \mathbf{0} = \mathbf{x} = \mathbf{0} + \mathbf{x}$;
- c) ogni elemento $\mathbf{x} = (x_1, x_2, \dots, x_N)$ ha un “opposto” $(-\mathbf{x}) = (-x_1, -x_2, \dots, -x_N)$: si ha $\mathbf{x} + (-\mathbf{x}) = \mathbf{0} = (-\mathbf{x}) + \mathbf{x}$;
- d) (commutativa) $\mathbf{x} + \mathbf{x}' = \mathbf{x}' + \mathbf{x}$.

Pertanto, $(\mathbb{R}^N, +)$ è un “gruppo abeliano”. Normalmente, si usa scrivere $\mathbf{x} - \mathbf{x}'$ per indicare $\mathbf{x} + (-\mathbf{x}')$.

Definiamo ora la moltiplicazione di un elemento di $\mathbb{R}^N$ per un numero reale: considerati $\mathbf{x} = (x_1, x_2, \dots, x_N) \in \mathbb{R}^N$ e un numero reale $\alpha \in \mathbb{R}$, si definisce $\alpha\mathbf{x}$ in questo modo:

$$\alpha\mathbf{x} = (\alpha x_1, \alpha x_2, \dots, \alpha x_N).$$

Valgono le seguenti proprietà:

- a) $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$;
- b) $(\alpha + \beta)\mathbf{x} = (\alpha\mathbf{x}) + (\beta\mathbf{x})$;
- c) $\alpha(\mathbf{x} + \mathbf{x}') = (\alpha\mathbf{x}) + (\alpha\mathbf{x}')$;
- d) $1\mathbf{x} = \mathbf{x}$.

Pertanto, con le operazioni introdotte, $\mathbb{R}^N$ è uno “spazio vettoriale”. Chiameremo i suoi elementi “vettori”; i numeri reali, in questo ambito, verranno chiamati “scalari”.

È utile introdurre il “prodotto scalare” tra due vettori: dati $\mathbf{x} = (x_1, x_2, \dots, x_N)$ e $\mathbf{x}' = (x'_1, x'_2, \dots, x'_N)$, si definisce il numero reale $\mathbf{x} \cdot \mathbf{x}'$ in questo modo:

$$\mathbf{x} \cdot \mathbf{x}' = \sum_{k=1}^N x_k x'_k.$$

Il prodotto scalare è spesso indicato con simboli diversi, quali ad esempio

$$\langle \mathbf{x} | \mathbf{x}' \rangle, \quad \langle \mathbf{x}, \mathbf{x}' \rangle, \quad (\mathbf{x} | \mathbf{x}'), \quad (\mathbf{x}, \mathbf{x}').$$

Valgono le seguenti proprietà:

- a) $\mathbf{x} \cdot \mathbf{x} \geq 0$;
- b) $\mathbf{x} \cdot \mathbf{x} = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$;
- c) $(\mathbf{x} + \mathbf{x}') \cdot \mathbf{x}'' = (\mathbf{x} \cdot \mathbf{x}'') + (\mathbf{x}' \cdot \mathbf{x}'')$;
- d) $(\alpha\mathbf{x}) \cdot \mathbf{x}' = \alpha(\mathbf{x} \cdot \mathbf{x}')$;
- e) $\mathbf{x} \cdot \mathbf{x}' = \mathbf{x}' \cdot \mathbf{x}$;

A partire dal prodotto scalare, possiamo definire la “norma” di un vettore $\mathbf{x} = (x_1, x_2, \dots, x_N)$:

$$\|\mathbf{x}\| = \sqrt{\mathbf{x} \cdot \mathbf{x}} = \sqrt{\sum_{k=1}^N x_k^2}.$$

Valgono le seguenti proprietà:

- a) $\|\mathbf{x}\| \geq 0$;
- b) $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$;
- c) $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$;
- d) $\|\mathbf{x} + \mathbf{x}'\| \leq \|\mathbf{x}\| + \|\mathbf{x}'\|$.

Per dimostrare la d), abbiamo bisogno della seguente **disuguaglianza di Schwarz**.

Teorema. *Presi due vettori $\mathbf{x}, \mathbf{x}'$, si ha*

$$|\mathbf{x} \cdot \mathbf{x}'| \leq \|\mathbf{x}\| \|\mathbf{x}'\|.$$

Dimostrazione. La disuguaglianza è sicuramente verificata se $\mathbf{x}' = \mathbf{0}$, essendo in tal caso $\mathbf{x} \cdot \mathbf{x}' = 0$ e $\|\mathbf{x}'\| = 0$. Supponiamo quindi $\mathbf{x}' \neq \mathbf{0}$. Per ogni $\alpha \in \mathbb{R}$, si ha

$$0 \leq \|\mathbf{x} - \alpha \mathbf{x}'\|^2 = (\mathbf{x} - \alpha \mathbf{x}') \cdot (\mathbf{x} - \alpha \mathbf{x}') = \|\mathbf{x}\|^2 - 2\alpha \mathbf{x} \cdot \mathbf{x}' + \alpha^2 \|\mathbf{x}'\|^2.$$

Prendendo $\alpha = \frac{1}{\|\mathbf{x}'\|^2} \mathbf{x} \cdot \mathbf{x}'$, si ottiene

$$0 \leq \|\mathbf{x}\|^2 - 2 \frac{1}{\|\mathbf{x}'\|^2} (\mathbf{x} \cdot \mathbf{x}')^2 + \frac{1}{\|\mathbf{x}'\|^4} (\mathbf{x} \cdot \mathbf{x}')^2 \|\mathbf{x}'\|^2 = \|\mathbf{x}\|^2 - \frac{1}{\|\mathbf{x}'\|^2} (\mathbf{x} \cdot \mathbf{x}')^2,$$

da cui la tesi. ■

Nota. La disuguaglianza di Schwarz diventa un’uguaglianza se e solo se i due vettori sono linearmente dipendenti. Infatti, vale sicuramente l’uguaglianza se uno dei due vettori è nullo. Se sono non nulli e $\mathbf{x}$ è del tipo $\alpha \mathbf{x}'$, si ha

$$|(\alpha \mathbf{x}') \cdot \mathbf{x}'| = |\alpha| |\mathbf{x}' \cdot \mathbf{x}'| = |\alpha| \|\mathbf{x}'\|^2 = \|\alpha \mathbf{x}'\| \|\mathbf{x}'\|.$$

Viceversa, se sono non nulli e vale l’uguaglianza $\mathbf{x} \cdot \mathbf{x}' = \|\mathbf{x}\| \|\mathbf{x}'\|$, allora

$$\left\| \frac{\mathbf{x}}{\|\mathbf{x}\|} - \frac{\mathbf{x}'}{\|\mathbf{x}'\|} \right\|^2 = 1 - 2 \frac{\mathbf{x} \cdot \mathbf{x}'}{\|\mathbf{x}\| \|\mathbf{x}'\|} + 1 = 0,$$

per cui $\frac{\mathbf{x}}{\|\mathbf{x}\|} = \frac{\mathbf{x}'}{\|\mathbf{x}'\|}$. I due vettori hanno la stessa direzione, per cui sono linearmente dipendenti. Analogamente, se $\mathbf{x} \cdot \mathbf{x}' = -\|\mathbf{x}\| \|\mathbf{x}'\|$, allora succede che $\frac{\mathbf{x}}{\|\mathbf{x}\|} = -\frac{\mathbf{x}'}{\|\mathbf{x}'\|}$.

Dimostriamo ora la proprietà d) della norma, usando la disuguaglianza di Schwarz:

$$\begin{aligned}\|\mathbf{x} + \mathbf{x}'\|^2 &= (\mathbf{x} + \mathbf{x}') \cdot (\mathbf{x} + \mathbf{x}') \\ &= \|\mathbf{x}\|^2 + 2\mathbf{x} \cdot \mathbf{x}' + \|\mathbf{x}'\|^2 \\ &\leq \|\mathbf{x}\|^2 + 2\|\mathbf{x}\| \|\mathbf{x}'\| + \|\mathbf{x}'\|^2 \\ &= (\|\mathbf{x}\| + \|\mathbf{x}'\|)^2,\end{aligned}$$

da cui la disuguaglianza cercata.

Notiamo ancora la seguente **identità del parallelogramma**, di semplice verifica:

$$\|\mathbf{x} + \mathbf{x}'\|^2 + \|\mathbf{x} - \mathbf{x}'\|^2 = 2(\|\mathbf{x}\|^2 + \|\mathbf{x}'\|^2).$$

Definiamo ora, a partire dalla norma, la “distanza euclidea” tra due vettori $\mathbf{x} = (x_1, x_2, \dots, x_N)$ e $\mathbf{x}' = (x'_1, x'_2, \dots, x'_N)$:

$$d(\mathbf{x}, \mathbf{x}') = \|\mathbf{x} - \mathbf{x}'\| = \sqrt{\sum_{k=1}^N (x_k - x'_k)^2}.$$

Valgono le seguenti proprietà:

- a) $d(\mathbf{x}, \mathbf{x}') \geq 0$;
- b) $d(\mathbf{x}, \mathbf{x}') = 0 \Leftrightarrow \mathbf{x} = \mathbf{x}'$;
- c) $d(\mathbf{x}, \mathbf{x}') = d(\mathbf{x}', \mathbf{x})$;
- d) $d(\mathbf{x}, \mathbf{x}'') \leq d(\mathbf{x}, \mathbf{x}') + d(\mathbf{x}', \mathbf{x}'')$.

Quest'ultima viene spesso chiamata “disuguaglianza triangolare”; la dimostriamo:

$$\begin{aligned}d(\mathbf{x}, \mathbf{x}'') &= \|\mathbf{x} - \mathbf{x}''\| \\ &= \|(\mathbf{x} - \mathbf{x}') + (\mathbf{x}' - \mathbf{x}'')\| \\ &\leq \|\mathbf{x} - \mathbf{x}'\| + \|\mathbf{x}' - \mathbf{x}''\| \\ &= d(\mathbf{x}, \mathbf{x}') + d(\mathbf{x}', \mathbf{x}'').\end{aligned}$$

Spazi metrici

Dato un insieme non vuoto E , una funzione $d : E \times E \rightarrow \mathbb{R}$ si chiama “distanza” (su E) se soddisfa alle seguenti proprietà:

- a) $d(x, x') \geq 0$;
- b) $d(x, x') = 0 \Leftrightarrow x = x'$;
- c) $d(x, x') = d(x', x)$;
- d) $d(x, x'') \leq d(x, x') + d(x', x'')$

(la disuguaglianza triangolare). L'insieme E , dotato della distanza d , si dice “spazio metrico”. I suoi elementi verranno spesso chiamati “punti”.

Abbiamo visto che $\mathbb{R}^N$, dotato della distanza euclidea, è uno spazio metrico (nel seguito, parlando dello spazio metrico $\mathbb{R}^N$, se non altrimenti specificato sottintenderemo che la distanza sia sempre quella euclidea). Nel caso $N = 1$, abbiamo la distanza usuale su $\mathbb{R}$: $d(\alpha, \beta) = |\alpha - \beta|$.

È però possibile considerare diverse distanze su uno stesso insieme. Ad esempio, presi due vettori $\mathbf{x} = (x_1, x_2, \dots, x_N)$ e $\mathbf{x}' = (x'_1, x'_2, \dots, x'_N)$, la funzione

$$d_*(\mathbf{x}, \mathbf{x}') = \sum_{k=1}^N |x_k - x'_k|$$

rappresenta anch'essa una distanza in $\mathbb{R}^N$. Lo stesso dicasi per la funzione

$$d_{**}(\mathbf{x}, \mathbf{x}') = \max\{|x_k - x'_k| : k = 1, 2, \dots, N\}.$$

Oppure, si può definire la seguente:

$$\hat{d}(\mathbf{x}, \mathbf{x}') = \begin{cases} 0 & \text{se } \mathbf{x} = \mathbf{x}', \\ 1 & \text{se } \mathbf{x} \neq \mathbf{x}'. \end{cases}$$

Anche questa è una distanza, per quanto strana possa sembrare.

Dati $x_0 \in E$ e un numero $\rho > 0$, definiamo la palla aperta di centro x_0 e raggio ρ :

$$B(x_0, \rho) = \{x \in E : d(x, x_0) < \rho\};$$

analogamente definiamo la palla chiusa

$$\overline{B}(x_0, \rho) = \{x \in E : d(x, x_0) \leq \rho\},$$

e la sfera

$$S(x_0, \rho) = \{x \in E : d(x, x_0) = \rho\}.$$

In $\mathbb{R}$, ogni intervallo² $]a, b[$ è una palla aperta e ogni intervallo $[a, b]$ è una palla chiusa: si ha

$$]a, b[= B\left(\frac{a+b}{2}, \frac{b-a}{2}\right), \quad [a, b] = \overline{B}\left(\frac{a+b}{2}, \frac{b-a}{2}\right).$$

Una sfera in $\mathbb{R}$ è quindi costituita da due soli punti.

In $\mathbb{R}^2$, con la distanza euclidea, una palla è un cerchio: la palla aperta non comprende i punti della circonferenza esterna, la palla chiusa sì. Una sfera è semplicemente una circonferenza.

²Usiamo qui le notazioni $]a, b[= \{x \in \mathbb{R} : a < x < b\}$ e $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$.

Se in $\mathbb{R}^2$ consideriamo la distanza d_* definita in precedenza, una palla sarà un quadrato, con i lati inclinati di 45 gradi, avente x_0 come punto centrale. Una sfera sarà il perimetro di tale quadrato. Se invece consideriamo la distanza d_{**} , la palla sarà ancora un quadrato, ma con i lati paralleli agli assi cartesiani.

Se invece prendiamo la distanza $\hat{d}$, su un qualsiasi insieme E , allora

$$B(x_0, \rho) = \begin{cases} \{x_0\} & \text{se } \rho \leq 1, \\ E & \text{se } \rho > 1, \end{cases} \quad \overline{B}(x_0, \rho) = \begin{cases} \{x_0\} & \text{se } \rho < 1, \\ E & \text{se } \rho \geq 1, \end{cases}$$

per cui

$$S(x_0, \rho) = \begin{cases} E \setminus \{x_0\} & \text{se } \rho = 1, \\ \emptyset & \text{se } \rho \neq 1. \end{cases}$$

Un insieme $U \subseteq E$ si dice “intorno” di un punto x_0 se esiste un $\rho > 0$ tale che $B(x_0, \rho) \subseteq U$; in tal caso, il punto x_0 si dice “interno” ad U . L'insieme dei punti interni ad U si chiama “l'interno” di U e si denota con $\mathring{U}$. Chiaramente, si ha sempre $\mathring{U} \subseteq U$. Si dice che U è un “insieme aperto” se coincide con il suo interno, ossia se $\mathring{U} = U$.

Teorema. *Una palla aperta è un insieme aperto.*

Dimostrazione. Sia $B(x_0, \rho)$ la palla in questione; prendiamo un $x_1 \in B(x_0, \rho)$. Scelto $r > 0$ tale che $r \leq \rho - d(x_0, x_1)$, si ha che $B(x_1, r) \subseteq B(x_0, \rho)$; infatti, se $x \in B(x_1, r)$, allora

$$d(x, x_0) \leq d(x, x_1) + d(x_1, x_0) < r + d(x_1, x_0) \leq \rho,$$

per cui $x \in B(x_0, \rho)$. Abbiamo quindi dimostrato che ogni punto x_1 di $B(x_0, \rho)$ è interno a $B(x_0, \rho)$. ■

Consideriamo ora tre esempi particolari: nel primo, l'insieme U coincide con E ; nel secondo, U è l'insieme vuoto; nel terzo, esso è costituito da un unico punto.

Ogni punto di E è interno all'insieme E stesso, in quanto ogni palla è per definizione contenuta in E . Quindi, l'interno di E coincide con tutto E , ossia $\mathring{E} = E$. Questo significa che E è un insieme aperto.

L'insieme vuoto non può avere punti interni. Quindi, l'interno di $\emptyset$, non avendo elementi, è vuoto. In altri termini, $\mathring{\emptyset} = \emptyset$, il che significa che $\emptyset$ è anch'esso un insieme aperto.

L'insieme $U = \{x_0\}$, costituito da un unico punto, in generale non è un insieme aperto (ad esempio in $\mathbb{R}^N$ con la distanza euclidea), ma può esserlo in casi particolari, quando x_0 è un “punto isolato” di E . Ad esempio, se si considera la distanza $\hat{d}$, oppure in $\mathbb{N}$, tutti i punti sono isolati.

Teorema. *L'interno di un insieme è un insieme aperto.*

Dimostrazione. Se $\mathring{U}$ è vuoto, la tesi è sicuramente vera. Supponiamo allora che $\mathring{U}$ sia non vuoto. Sia $x_1 \in \mathring{U}$. Allora esiste un $\rho > 0$ tale che $B(x_1, \rho) \subseteq U$. Se proviamo che $B(x_1, \rho) \subseteq \mathring{U}$ la dimostrazione sarà completa, in quanto avremo dimostrato che ogni punto x_1 di $\mathring{U}$ è interno a $\mathring{U}$.

Per dimostrare che $B(x_1, \rho) \subseteq \mathring{U}$, sia x un elemento di $B(x_1, \rho)$. Siccome $B(x_1, \rho)$ è un insieme aperto, esiste un $r > 0$ tale che $B(x, r) \subseteq B(x_1, \rho)$. Allora $B(x, r) \subseteq U$, per cui x appartiene a $\mathring{U}$. La dimostrazione è completa. ■

Si può dimostrare la seguente implicazione:

$$U_1 \subseteq U_2 \quad \Rightarrow \quad \mathring{U}_1 \subseteq \mathring{U}_2 .$$

Da essa segue che $\mathring{U}$ è il più grande insieme aperto contenuto in U : se A è un aperto e $A \subseteq U$, allora $A \subseteq \mathring{U}$.

Diremo che il punto x_0 è “aderente” all'insieme U se per ogni $\rho > 0$ si ha che $B(x_0, \rho) \cap U \neq \emptyset$. L'insieme dei punti aderenti ad U si chiama “la chiusura” di U e si denota con $\overline{U}$. Chiaramente, si ha sempre $U \subseteq \overline{U}$. Si dice che U è un “insieme chiuso” se coincide con la sua chiusura, ossia se $U = \overline{U}$.

Teorema. *Una palla chiusa è un insieme chiuso.*

Dimostrazione. Sia $U = \overline{B}(x_0, \rho)$ la palla in questione; voglio dimostrare che $\overline{U} \subseteq U$. A tal fine vedremo che $\mathcal{C}U \subseteq \mathcal{C}\overline{U}$.³ Prendiamo un $x_1 \in \mathcal{C}U$, ossia $x_1 \notin \overline{B}(x_0, \rho)$. Scelto $r > 0$ tale che $r \leq d(x_0, x_1) - \rho$, si ha che $B(x_1, r) \cap \overline{B}(x_0, \rho) = \emptyset$; infatti, se per assurdo esistesse un $x \in B(x_1, r) \cap \overline{B}(x_0, \rho)$, allora si avrebbe

$$d(x_0, x_1) \leq d(x_0, x) + d(x, x_1) < \rho + r ,$$

in contrasto con la scelta fatta per r . Quindi, $x_1 \notin \overline{U}$, ossia $x_1 \in \mathcal{C}\overline{U}$. ■

Essendo E il l'insieme universo, ogni punto aderente ad E deve comunque appartenere ad E stesso. Quindi, la chiusura di E coincide con E , ossia $\overline{E} = E$. Questo significa che E è un insieme chiuso.

Notiamo che non esiste alcun punto aderente all'insieme $\emptyset$. Infatti, qualsiasi sia il punto x_0 , per ogni $\rho > 0$ si ha che $B(x_0, \rho) \cap \emptyset = \emptyset$. Quindi, la chiusura di $\emptyset$, non avendo elementi, è vuota. In altri termini, $\overline{\emptyset} = \emptyset$, il che significa che $\emptyset$ è un insieme chiuso.

L'insieme $U = \{x_0\}$, costituito da un unico punto, è sempre un insieme chiuso. Infatti, preso un $x_1 \notin U$, scegliendo $\rho > 0$ tale che $\rho < d(x_0, x_1)$ si ha che $B(x_1, \rho) \cap U = \emptyset$, per cui x_1 non è aderente ad U .

³Denotiamo con $\mathcal{C}U$ il complementare di U in E , ossia l'insieme $E \setminus U$.

Teorema. *La chiusura di un insieme è un insieme chiuso.*

Dimostrazione. Poniamo $V = \overline{U}$. Se $V = E$, la tesi è verificata. Supponiamo quindi che sia $V \neq E$. Sia $x_1 \notin V$. Allora esiste un $\rho > 0$ tale che $B(x_1, \rho) \cap U = \emptyset$. Vediamo che anche $B(x_1, \rho) \cap V = \emptyset$. Infatti, se per assurdo ci fosse un $x \in B(x_1, \rho) \cap V$, essendo $B(x_1, \rho)$ un insieme aperto, esisterebbe un $r > 0$ tale che $B(x, r) \subseteq B(x_1, \rho)$. Siccome $x \in V = \overline{U}$, dovrebbe essere $B(x, r) \cap U \neq \emptyset$ e quindi anche $B(x_1, \rho) \cap U \neq \emptyset$, in contraddizione con quanto sopra. Quindi, nessun punto x_1 al di fuori di V può essere aderente a V . In altri termini, V contiene tutti i punti ad esso aderenti, pertanto è chiuso. ■

Si può dimostrare che

$$U_1 \subseteq U_2 \quad \Rightarrow \quad \overline{U}_1 \subseteq \overline{U}_2.$$

Da questa implicazione segue che $\overline{U}$ è il più piccolo insieme chiuso che contiene U : se C è un chiuso e $C \supseteq U$, allora $C \supseteq \overline{U}$.

Si può dimostrare che l'unione e l'intersezione di due insiemi aperti [chiusi] sono insiemi aperti [chiusi]. Lo stesso vale per un numero finito di insiemi aperti [chiusi]: lo si dimostra per induzione. Se invece si considera un numero infinito di insiemi, la cosa cambia. L'unione di un numero infinito di insiemi aperti è un insieme aperto, l'intersezione in generale non lo è. Ad esempio, in $\mathbb{R}$, prendendo gli aperti

$$A_n = \left] -\frac{1}{n+1}, \frac{1}{n+1} \right[,$$

con $n \in \mathbb{N}$, la loro intersezione è $\{0\}$, che non è un aperto. Analogamente, l'intersezione di un numero infinito di insiemi chiusi è un insieme chiuso, mentre l'unione in generale non lo è. Ad esempio, considerando i chiusi

$$C_n = \left[-1 + \frac{1}{n+1}, 1 - \frac{1}{n+1} \right] ,$$

con $n \in \mathbb{N}$, la loro unione è l'intervallo $] -1, 1[$, che non è un chiuso.

Cercheremo ora di capire le analogie incontrate tra le nozioni di interno e chiusura di un insieme, e quelle di insieme aperto e chiuso.

Teorema. *Valgono le seguenti relazioni:*

$$\overline{\mathcal{C}U} = \mathcal{C}\mathring{U}, \quad (\mathcal{C}\mathring{U}) = \mathcal{C}\overline{U}.$$

Dimostrazione. Vediamo la prima uguaglianza. Se $U = E$, allora $\mathcal{C}U = \emptyset$, per cui $\overline{\mathcal{C}U} = \emptyset$; d'altra parte, $\mathring{U} = E$, per cui $\mathcal{C}\mathring{U} = \emptyset$. L'uguaglianza è così verificata in

questo caso. Supponiamo ora che sia $U \neq E$, per cui $\mathcal{C}U \neq \emptyset$. Si ha:

$$\begin{aligned} x \in \overline{\mathcal{C}U} &\Leftrightarrow \forall \rho > 0 \quad B(x, \rho) \cap \mathcal{C}U \neq \emptyset \\ &\Leftrightarrow \forall \rho > 0 \quad B(x, \rho) \not\subseteq U \\ &\Leftrightarrow x \notin \overset{\circ}{U} \\ &\Leftrightarrow x \in \mathcal{C}\overset{\circ}{U}. \end{aligned}$$

Questo dimostra la prima uguaglianza. Possiamo ora usarla per dedurre la seguente:

$$\mathcal{C}(\mathcal{C}\overset{\circ}{U}) = \overline{\mathcal{C}(\mathcal{C}U)} = \overline{U}.$$

Passando ai complementari, si ottiene la seconda uguaglianza. ■

Abbiamo quindi che

$$\overline{U} = \mathcal{C}(\mathcal{C}\overset{\circ}{U}), \quad \overset{\circ}{U} = \mathcal{C}(\overline{\mathcal{C}U}).$$

Come immediato corollario, abbiamo il seguente.

Teorema. *Un insieme è aperto [chiuso] se e solo se il suo complementare è chiuso [aperto].*

Dimostrazione. Se U è aperto, $U = \overset{\circ}{U}$ e quindi

$$\overline{\mathcal{C}U} = \mathcal{C}\overset{\circ}{U} = \mathcal{C}U,$$

per cui $\mathcal{C}U$ è chiuso.

Se U è chiuso, $U = \overline{U}$ e quindi

$$(\mathcal{C}\overset{\circ}{U}) = \mathcal{C}\overline{U} = \mathcal{C}U,$$

per cui $\mathcal{C}U$ è aperto. ■

Si definisce la “frontiera” di un insieme U come differenza tra la sua chiusura e il suo interno:

$$\partial U = \overline{U} \setminus \overset{\circ}{U}.$$

Ad esempio, in $\mathbb{R}$ abbiamo:

$$\overline{\mathbb{Q}} = \mathbb{R}, \quad \overset{\circ}{\mathbb{Q}} = \emptyset, \quad \partial \mathbb{Q} = \mathbb{R}.$$

È bene essere prudenti su alcune conclusioni che possono esserci suggerite dalla nostra intuizione basata sulla distanza euclidea. Ad esempio, le uguaglianze

$$\overline{B(\mathbf{x}_0, \rho)} = \overline{B}(\mathbf{x}_0, \rho), \quad \partial B(\mathbf{x}_0, \rho) = S(\mathbf{x}_0, \rho).$$

valgono sicuramente in $\mathbb{R}^N$ con la distanza euclidea, ma possono non valere in altri casi. Prendiamo ad esempio la distanza $\hat{d}$ considerata sopra. Allora $B(\mathbf{x}_0, 1) = \{\mathbf{x}_0\}$, che è un insieme chiuso, e $\overline{B}(\mathbf{x}_0, 1) = E$ per cui $\overline{B}(\mathbf{x}_0, 1) \neq \overline{B}(\mathbf{x}_0, 1)$. Inoltre, $\partial B(\mathbf{x}_0, 1) = \emptyset$, mentre $S(\mathbf{x}_0, 1) = E \setminus \{\mathbf{x}_0\}$, per cui $\partial B(\mathbf{x}_0, 1) \neq S(\mathbf{x}_0, 1)$.

Nota. Nel seguito, qualora non specificato altrimenti, quando parleremo di $\mathbb{R}^N$ come spazio metrico o normato sarà sempre sottinteso che la distanza e la norma su di esso considerate siano quelle euclidee.

Funzioni continue

Intuitivamente, una funzione f è “continua” se $f(x)$ varia gradualmente al variare di x nel dominio, cioè quando non si verificano variazioni brusche nei valori della funzione. Per rendere rigorosa questa idea intuitiva, sarà conveniente focalizzare la nostra attenzione fissando un x_0 nel dominio e provando a precisare cosa intendiamo per

f è “continua” in x_0 .

Procederemo per gradi.

Primo tentativo. Diremo che f è “continua” in x_0 quando si verifica la cosa seguente:

se x è vicino a x_0 , allora $f(x)$ è vicino a $f(x_0)$.

Osserviamo subito che, sebbene l’idea di continuità vi sia già abbastanza ben formulata, la proposizione precedente non è una definizione accettabile, perché la parola “vicino”, che vi compare due volte, non ha un significato preciso. Innanzitutto, per poter misurare quanto vicino sia x a x_0 e quanto vicino sia $f(x)$ a $f(x_0)$, abbiamo bisogno di introdurre delle distanze. Più precisamente, dovremo supporre che il dominio e il codominio della funzione siano due spazi metrici.

Siano quindi E ed F due spazi metrici, con le loro distanze d_E e d_F , rispettivamente. Sia x_0 un punto di E e $f : E \rightarrow F$ una funzione. Possiamo riformulare il tentativo di definizione precedente come segue.

Secondo tentativo. Diremo che f è “continua” in x_0 quando si verifica la cosa seguente:

se la distanza $d_E(x, x_0)$ è piccola, allora la distanza $d_F(f(x), f(x_0))$ è piccola.

Ci rendiamo subito conto che il problema riscontrato nel primo tentativo non è stato affatto risolto con questo secondo tentativo, in quanto vi compare ora per due volte la parola “piccola”, che non ha un significato preciso. Ci chiediamo allora: *quanto piccola* vogliamo che sia la distanza $d_F(f(x), f(x_0))$? L’idea che abbiamo in mente è che questa distanza possa essere resa piccola quanto si voglia (purché la distanza $d(x, x_0)$ sia sufficientemente piccola, s’intende). Per poterla misurare, introdurremo quindi un numero reale positivo, che chiameremo ε , e chiederemo che sia $d_F(f(x), f(x_0)) < \varepsilon$, qualora $d(x, x_0)$ sia sufficientemente piccola. L’arbitrarietà di tale ε ci permetterà di prenderlo piccolo quanto si voglia.

Terzo tentativo. Diremo che f è “continua” in x_0 quando si verifica la cosa seguente: preso un qualsiasi numero $\varepsilon > 0$,

se la distanza $d_E(x, x_0)$ è piccola, allora $d_F(f(x), f(x_0)) < \varepsilon$.

Adesso la parola “piccola” compare una sola volta, mentre la distanza $d_F(f(x), f(x_0))$ viene semplicemente controllata dal numero ε . Quindi, almeno la seconda parte della proposizione ha ora un significato ben preciso. Potremmo allora cercare di fare altrettanto con la distanza $d(x, x_0)$, introducendo un nuovo numero reale positivo, che chiameremo δ , che la controlli.

Quarto tentativo (quello buono!). Diremo che f è “continua” in x_0 quando si verifica la cosa seguente: preso un qualsiasi numero $\varepsilon > 0$, è possibile trovare un numero $\delta > 0$ per cui,

se $d_E(x, x_0) < \delta$, allora $d_F(f(x), f(x_0)) < \varepsilon$.

Quest’ultima proposizione, a differenza delle precedenti, non presenta alcun termine impreciso. Le distanze $d_E(x, x_0)$ e $d_F(f(x), f(x_0))$ sono semplicemente controllate da due numeri positivi δ e ε , rispettivamente. Riscriviamola quindi in modo formale.

Definizione. Diremo che f è “continua” in x_0 se, comunque preso un numero positivo ε , è possibile trovare un numero positivo δ tale che, se x è un qualsiasi elemento del dominio E che disti da x_0 per meno di δ , allora $f(x)$ dista da $f(x_0)$ per meno di ε . In simboli:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \forall x \in E \quad d_E(x, x_0) < \delta \Rightarrow d_F(f(x), f(x_0)) < \varepsilon.$$

In questa formulazione, spesso la scrittura “ $\forall x \in E$ ” verrà sottintesa.

Si può osservare che una o entrambe le disuguaglianze $d_E(x, x_0) < \delta$ e $d_F(f(x), f(x_0)) < \varepsilon$ possono essere sostituite rispettivamente da $d_E(x, x_0) \leq \delta$ e $d_F(f(x), f(x_0)) \leq \varepsilon$, ottenendo definizioni che sono tutte tra loro equivalenti. Questo è dovuto al fatto, da un lato, che ε è un *qualunque* numero positivo e, dall’altro lato, che se l’implicazione della definizione vale per un certo numero positivo δ , essa vale a maggior ragione prendendo al posto di quel δ un qualsiasi numero positivo più piccolo.

Una rilettura della definizione di continuità ci mostra che f è continua in x_0 se e solo se:

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : f(B(x_0, \delta)) \subseteq B(f(x_0), \varepsilon).$$

Inoltre, è del tutto equivalente considerare una palla chiusa al posto di una palla aperta; risulta inoltre utile la seguente formulazione equivalente, per cui f è continua in x_0 se e solo se:

per ogni intorno V di $f(x_0)$ esiste un intorno U di x_0 tale che $f(U) \subseteq V$.

Nel caso in cui la funzione f sia continua in ogni punto x_0 del dominio E , diremo che “ f è continua su E ”, o semplicemente “ f è continua”.

Vediamo ora alcuni esempi.

1) La funzione costante: per un certo $\bar{c} \in F$, si ha che $f(x) = \bar{c}$, per ogni $x \in E$. Essendo $d_F(f(x), f(x_0)) = d_F(\bar{c}, \bar{c}) = 0$ per ogni $x \in E$, tale funzione è chiaramente continua (ogni scelta di $\delta > 0$ va bene).

2) Supponiamo che x_0 sia un “punto isolato” di E : esiste cioè un $\rho > 0$ per cui non ci sono punti di E che distino da x_0 per meno di ρ , tranne x_0 stesso. Vediamo che, in questo caso, qualsiasi funzione $f : E \rightarrow F$ risulta continua in x_0 . Infatti, dato $\varepsilon > 0$ qualsiasi, prendendo $\delta = \rho$, avremo che $B(x_0, \delta) = \{x_0\}$, per cui $f(B(x_0, \delta)) = \{f(x_0)\} \subseteq B(f(x_0), \varepsilon)$.

3) Siano $E = \mathbb{R}^N$ e $F = \mathbb{R}^N$. Fissato un numero $\alpha \in \mathbb{R}$, consideriamo la funzione $f : \mathbb{R}^N \rightarrow \mathbb{R}^N$ definita da $f(\mathbf{x}) = \alpha \mathbf{x}$. Vediamo che è continua. Infatti, se $\alpha = 0$, si tratta della funzione costante con valore $\mathbf{0}$, e sappiamo che tale funzione è continua. Sia ora $\alpha \neq 0$. Allora, fissato $\varepsilon > 0$, essendo

$$\|f(\mathbf{x}) - f(\mathbf{x}_0)\| = \|\alpha \mathbf{x} - \alpha \mathbf{x}_0\| = \|\alpha(\mathbf{x} - \mathbf{x}_0)\| = |\alpha| \|\mathbf{x} - \mathbf{x}_0\|,$$

basta prendere $\delta = \frac{\varepsilon}{|\alpha|}$ per avere l’implicazione

$$\|\mathbf{x} - \mathbf{x}_0\| < \delta \Rightarrow \|f(\mathbf{x}) - f(\mathbf{x}_0)\| < \varepsilon.$$

4) Siano $E = \mathbb{R}^N$ e $F = \mathbb{R}$. Vediamo che la funzione $f : \mathbb{R}^N \rightarrow \mathbb{R}$ definita da $f(\mathbf{x}) = \|\mathbf{x}\|$ è continua su $\mathbb{R}^N$. Questo seguirà facilmente dalla disuguaglianza

$$\left| \|\mathbf{x}\| - \|\mathbf{x}'\| \right| \leq \|\mathbf{x} - \mathbf{x}'\|,$$

che ora dimostriamo. Si ha:

$$\begin{aligned} \|\mathbf{x}\| &= \|(\mathbf{x} - \mathbf{x}') + \mathbf{x}'\| \leq \|\mathbf{x} - \mathbf{x}'\| + \|\mathbf{x}'\|, \\ \|\mathbf{x}'\| &= \|(\mathbf{x}' - \mathbf{x}) + \mathbf{x}\| \leq \|\mathbf{x}' - \mathbf{x}\| + \|\mathbf{x}\|. \end{aligned}$$

Essendo $\|\mathbf{x} - \mathbf{x}'\| = \|\mathbf{x}' - \mathbf{x}\|$, si ha che

$$\|\mathbf{x}\| - \|\mathbf{x}'\| \leq \|\mathbf{x} - \mathbf{x}'\| \quad \text{e} \quad \|\mathbf{x}'\| - \|\mathbf{x}\| \leq \|\mathbf{x} - \mathbf{x}'\|,$$

da cui la disuguaglianza cercata. A questo punto, considerato un $\mathbf{x}_0 \in \mathbb{R}^N$ e fissato un $\varepsilon > 0$, basta prendere $\delta = \varepsilon$ per avere che

$$\|\mathbf{x} - \mathbf{x}_0\| < \delta \Rightarrow \left| \|\mathbf{x}\| - \|\mathbf{x}_0\| \right| < \varepsilon.$$

5) Siano $E = \mathbb{R}$ e $F = \mathbb{R}$, e consideriamo la “funzione segno” $f : \mathbb{R} \rightarrow \mathbb{R}$ definita da

$$f(x) = \begin{cases} -1 & \text{se } x < 0 \\ 0 & \text{se } x = 0 \\ 1 & \text{se } x > 0. \end{cases}$$

Si può vedere che questa funzione è continua in tutti i punti tranne che in $x_0 = 0$. Infatti, se $x_0 \neq 0$, basterà prendere $\delta < |x_0|$ per avere che f è costante sull'intervallo $]x_0 - \delta, x_0 + \delta[$, quindi continua in x_0 . Per vedere che f non è continua in 0, fissiamo un $\varepsilon \in]0, 1[$; per ogni scelta di $\delta > 0$, è possibile trovare un $x \in]-\delta, \delta[$ tale che $|f(x)| = 1$, per cui $|f(x) - f(0)| > \varepsilon$.

6) La “funzione di Dirichlet” $f : \mathbb{R} \rightarrow \mathbb{R}$ è definita da

$$f(x) = \begin{cases} 1 & \text{se } x \in \mathbb{Q} \\ 0 & \text{se } x \notin \mathbb{Q}. \end{cases}$$

Questa funzione non è continua, in alcun punto x_0 . Infatti, fissato un $\varepsilon \in]0, 1[$, siccome sia $\mathbb{Q}$ che $\mathbb{R} \setminus \mathbb{Q}$ sono densi in $\mathbb{R}$, per ogni x_0 e ogni scelta di $\delta > 0$ ci saranno sicuramente un razionale x' e un irrazionale x'' in $]x_0 - \delta, x_0 + \delta[$; quindi, a seconda che x_0 sia razionale o irrazionale, si avrà che $|f(x'') - f(x_0)| > \varepsilon$ o $|f(x') - f(x_0)| > \varepsilon$.

Enunciamo ora alcune proprietà delle funzioni continue aventi come codominio $F = \mathbb{R}$.

Teorema. Se $f, g : E \rightarrow \mathbb{R}$ sono continue in x_0 , anche $f + g$ lo è.

Dimostrazione. Fissiamo $\varepsilon > 0$. Per la continuità di f e g esistono $\delta_1 > 0$ e $\delta_2 > 0$ tali che

$$\begin{aligned} d(x, x_0) < \delta_1 &\Rightarrow |f(x) - f(x_0)| < \varepsilon, \\ d(x, x_0) < \delta_2 &\Rightarrow |g(x) - g(x_0)| < \varepsilon. \end{aligned}$$

Quindi, se $\delta = \min\{\delta_1, \delta_2\}$, si ha

$$d(x, x_0) < \delta \Rightarrow |(f + g)(x) - (f + g)(x_0)| \leq |f(x) - f(x_0)| + |g(x) - g(x_0)| < 2\varepsilon.$$

Data l'arbitrarietà di ε , ciò dimostra che $f + g$ è continua in x_0 . ■

Teorema. Se $f, g : E \rightarrow \mathbb{R}$ sono continue in x_0 , anche $f \cdot g$ lo è.

Dimostrazione. Fissiamo $\varepsilon > 0$. Non è restrittivo supporre $\varepsilon \leq 1$, in quanto possiamo sempre porre $\varepsilon' = \min\{\varepsilon, 1\}$ e procedere con ε' al posto di ε . Per la continuità di f e g esistono $\delta_1 > 0$ e $\delta_2 > 0$ tali che

$$\begin{aligned} d(x, x_0) < \delta_1 &\Rightarrow |f(x) - f(x_0)| < \varepsilon, \\ d(x, x_0) < \delta_2 &\Rightarrow |g(x) - g(x_0)| < \varepsilon. \end{aligned}$$

Notiamo che, essendo $\varepsilon \leq 1$, da $|f(x) - f(x_0)| < \varepsilon$ segue che $|f(x)| < |f(x_0)| + 1$. Quindi, se $\delta = \min\{\delta_1, \delta_2\}$, si ha

$$\begin{aligned} d(x, x_0) < \delta &\Rightarrow |(f \cdot g)(x) - (f \cdot g)(x_0)| = \\ &= |f(x)g(x) - f(x)g(x_0) + f(x)g(x_0) - f(x_0)g(x_0)| \\ &\leq |f(x)| \cdot |g(x) - g(x_0)| + |g(x_0)| \cdot |f(x) - f(x_0)| \\ &\leq (|f(x_0)| + 1) \cdot |g(x) - g(x_0)| + |g(x_0)| \cdot |f(x) - f(x_0)| \\ &< (|f(x_0)| + |g(x_0)| + 1)\varepsilon. \end{aligned}$$

Data l'arbitrarietà di ε , ciò dimostra che $f \cdot g$ è continua in x_0 . ■

Teorema. Se $f, g : E \rightarrow \mathbb{R}$ sono continue in x_0 , anche $f - g$ lo è.

Dimostrazione. Segue immediatamente dai due teoremi precedenti e dal fatto che ogni funzione costante è continua, in quanto $f - g = f + (-1) \cdot g$. ■

Teorema (della permanenza del segno). Se $g : E \rightarrow \mathbb{R}$ è continua in x_0 e $g(x_0) > 0$, allora esiste un $\delta > 0$ tale che

$$d(x, x_0) < \delta \Rightarrow g(x) > 0.$$

Dimostrazione. Fissiamo $\varepsilon = g(x_0)$. Per la continuità, esiste un $\delta > 0$ tale che

$$d(x, x_0) < \delta \Rightarrow g(x_0) - \varepsilon < g(x) < g(x_0) + \varepsilon \Rightarrow 0 < g(x) < 2g(x_0).$$

■

Naturalmente, se $g(x_0) < 0$, allora esiste un $\delta > 0$ tale che

$$d(x, x_0) < \delta \Rightarrow g(x) < 0.$$

Teorema. Se $f, g : E \rightarrow \mathbb{R}$ sono continue in x_0 e $g(x_0) \neq 0$, anche $\frac{f}{g}$ è continua in x_0 .

Dimostrazione. Si noti che, per la proprietà di permanenza del segno, esiste un $\rho > 0$ tale che il rapporto $\frac{f(x)}{g(x)}$ è definito almeno per tutti gli x di E che distano da x_0 per meno di ρ . Essendo $\frac{f}{g} = f \cdot \frac{1}{g}$, basterà dimostrare che $\frac{1}{g}$ è continua in x_0 . Fissiamo $\varepsilon > 0$; possiamo supporre senza perdita di generalità che $\varepsilon < \frac{|g(x_0)|}{2}$. Per la continuità di g , esiste un $\delta > 0$ tale che

$$d(x, x_0) < \delta \Rightarrow |g(x) - g(x_0)| < \varepsilon.$$

Ma allora, essendo $\varepsilon < \frac{|g(x_0)|}{2}$, anche

$$d(x, x_0) < \delta \Rightarrow |g(x)| > |g(x_0)| - \varepsilon > \frac{|g(x_0)|}{2}.$$

Ne segue che

$$d(x, x_0) < \delta \Rightarrow \left| \frac{1}{g}(x) - \frac{1}{g}(x_0) \right| = \frac{|g(x_0) - g(x)|}{|g(x)g(x_0)|} < \frac{2}{|g(x_0)|^2} \varepsilon.$$

Per l'arbitrarietà de ε , questo dimostra che $\frac{1}{g}$ è continua in x_0 . ■

Sappiamo da quanto sopra che le funzioni costanti sono continue, così come la funzione $f(x) = x$. Usando i teoremi precedenti, abbiamo quindi che tutte le funzioni polinomiali sono continue, così come le funzioni razionali, definite dal rapporto di due polinomi. Più precisamente, esse sono continue sul loro dominio, ossia sull'insieme dei punti in cui il denominatore non si annulla.

Vediamo ora come si comporta una funzione composta di due funzioni continue.

Teorema. Siano $f : E \rightarrow F$ continua in x_0 e $g : F \rightarrow G$ continua in $f(x_0)$; allora $g \circ f$ è continua in x_0 .

Dimostrazione. Fissato un intorno W di $[g \circ f](x_0) = g(f(x_0))$, per la continuità di g in $f(x_0)$ esiste un intorno V di $f(x_0)$ tale che $g(V) \subseteq W$. Allora, per la continuità di f in x_0 , esiste un intorno U di x_0 tale che $f(U) \subseteq V$. Ne segue che $[g \circ f](U) \subseteq W$. ■

Intervalli e continuità

È molto importante la seguente proprietà delle funzioni continue definite su un intervallo.

Teorema (degli zeri, o di Bolzano). Se $f : [a, b] \rightarrow \mathbb{R}$ è una funzione continua tale che

$$f(a) < 0 < f(b) \quad \text{oppure} \quad f(a) > 0 > f(b),$$

allora esiste un $c \in]a, b[$ tale che $f(c) = 0$.

Dimostrazione. Considereremo il caso $f(a) < 0 < f(b)$, essendo l'altro del tutto analogo. Scriviamo $I_0 = [a, b]$ e consideriamo il punto medio $\frac{a+b}{2}$ dell'intervallo I_0 . Se f si annulla in esso, abbiamo trovato il punto c cercato. Altrimenti, $f(\frac{a+b}{2}) < 0$ o $f(\frac{a+b}{2}) > 0$. Se $f(\frac{a+b}{2}) < 0$, chiamiamo I_1 l'intervallo $[\frac{a+b}{2}, b]$; se $f(\frac{a+b}{2}) > 0$, chiamiamo invece I_1 l'intervallo $[a, \frac{a+b}{2}]$. Prendendo ora il punto medio di I_1 e ripetendo il ragionamento, possiamo definire un intervallo I_2 e, per ricorrenza, una successione di intervalli $I_n = [a_n, b_n]$ tali che

$$I_0 \supseteq I_1 \supseteq I_2 \supseteq I_3 \supseteq \dots$$

e, per ogni n , $f(a_n) < 0 < f(b_n)$. Per il teorema di Cantor, esiste un $c \in \mathbb{R}$ appartenente a tutti gli intervalli. Dimostriamo che $f(c) = 0$. Per assurdo, se

$f(c) < 0$, per la permanenza del segno esiste un $\delta > 0$ tale che $f(x) < 0$ per ogni $x \in]c - \delta, c + \delta[$. Ma siccome $b_n - c \leq b_n - a_n$ e, per $n \geq 1$, $b_n - a_n = \frac{b-a}{2^n} < \frac{b-a}{n}$, prendendo $n > \frac{b-a}{\delta}$ si ha che $b_n \in]c - \delta, c + \delta[$. Ma allora dovrebbe essere $f(b_n) < 0$, in contraddizione con quanto sopra. Un ragionamento analogo porta a una contraddizione supponendo $f(c) > 0$. ■

Come conseguenza del teorema degli zeri, abbiamo che una funzione continua “manda intervalli in intervalli”:

Corollario. *Sia E un sottoinsieme di $\mathbb{R}$ e $f : E \rightarrow \mathbb{R}$ una funzione continua. Se $I \subseteq E$ è un intervallo, allora anche $f(I)$ è un intervallo.*

Dimostrazione. Escludendo i casi banali in cui I o $f(I)$ consistono di un unico punto, prendiamo $\alpha, \beta \in f(I)$, con $\alpha < \beta$ e sia γ tale che $\alpha < \gamma < \beta$. Vogliamo vedere che $\gamma \in f(I)$. Consideriamo la funzione $g : E \rightarrow \mathbb{R}$ definita da

$$g(x) = f(x) - \gamma.$$

Siano a, b in I tali che $f(a) = \alpha$ e $f(b) = \beta$. Essendo I un intervallo, la funzione g è definita su $[a, b]$ (o $[b, a]$, nel caso in cui $b < a$) ed è ivi continua. Inoltre, $g(a) < 0 < g(b)$ e quindi, per il teorema degli zeri, esiste un $c \in]a, b[$ tale che $g(c) = 0$, ossia $f(c) = \gamma$. ■

Funzioni monotone

Sia E un sottoinsieme di $\mathbb{R}$. Diremo che una funzione $f : E \rightarrow \mathbb{R}$ è:

“crescente” se $[x_1 < x_2 \Rightarrow f(x_1) \leq f(x_2)]$;

“decrescente” se $[x_1 < x_2 \Rightarrow f(x_1) \geq f(x_2)]$;

“strettamente crescente” se $[x_1 < x_2 \Rightarrow f(x_1) < f(x_2)]$;

“strettamente decrescente” se $[x_1 < x_2 \Rightarrow f(x_1) > f(x_2)]$.

Diremo che è “monotona” se è crescente o decrescente; “strettamente monotona” se è strettamente crescente o strettamente decrescente.

Esempio. La funzione $f : [0, +\infty[\rightarrow \mathbb{R}$ definita da $f(x) = x^n$ è strettamente crescente. Il caso $n = 2$ è stato stabilito nel Lemma a pagina 10. Il caso generale si vede per induzione.

Vediamo ora un teorema sulla continuità delle funzioni invertibili.

Teorema. *Siano I e J due intervalli e $f : I \rightarrow J$ una funzione invertibile. Allora*

$$f \text{ è continua} \quad \Leftrightarrow \quad f \text{ è strettamente monotona}.$$

In tal caso, anche $f^{-1} : J \rightarrow I$ è strettamente monotona e continua.

Dimostrazione. Supponiamo f continua e, per assurdo, non strettamente monotona. Allora esistono $x_1 < x_2 < x_3$ in I tali che

$$f(x_1) < f(x_2) \quad \text{e} \quad f(x_2) > f(x_3),$$

oppure

$$f(x_1) > f(x_2) \quad \text{e} \quad f(x_2) < f(x_3).$$

(Le uguaglianze non possono valere, essendo la funzione f iniettiva.) Consideriamo il primo caso, l'altro essendo analogo. Scegliendo $\gamma \in \mathbb{R}$ tale che $f(x_1) < \gamma < f(x_2)$ e $f(x_2) > \gamma > f(x_3)$, per il corollario al teorema degli zeri si trova che esistono $a \in]x_1, x_2[$ e $b \in]x_2, x_3[$ tali che $f(a) = \gamma = f(b)$, in contraddizione con l'iniettività di f .

Supponiamo ora f strettamente monotona, ad esempio crescente: l'altro caso è del tutto analogo. Preso $x_0 \in I$, vogliamo dimostrare che f è continua in x_0 . Considereremo due casi distinti.

Supponiamo dapprima che x_0 non sia un estremo di I , e pertanto $y_0 = f(x_0)$ non sia un estremo di J . Fissiamo $\varepsilon > 0$; possiamo supporre senza perdita di generalità che $[y_0 - \varepsilon, y_0 + \varepsilon] \subseteq J$. Poniamo $x_1 = f^{-1}(y_0 - \varepsilon)$ e $x_2 = f^{-1}(y_0 + \varepsilon)$, per cui $x_1 < x_0 < x_2$. Essendo $f(x_1) = f(x_0) - \varepsilon$ e $f(x_2) = f(x_0) + \varepsilon$, prendendo $\delta = \min\{x_0 - x_1, x_2 - x_0\}$, si ha

$$d(x, x_0) < \delta \Rightarrow x_1 < x < x_2 \Rightarrow f(x_1) < f(x) < f(x_2) \Rightarrow d(f(x), f(x_0)) < \varepsilon,$$

per cui f è continua in x_0 .

Consideriamo ora l'eventualità che $x_0 = \min I$ e quindi $y_0 = \min J$. Fissiamo $\varepsilon > 0$; possiamo supporre senza perdita di generalità che $[y_0, y_0 + \varepsilon] \subseteq J$. Poniamo come sopra $x_2 = f^{-1}(y_0 + \varepsilon)$. Essendo $f(x_2) = f(x_0) + \varepsilon$, prendendo $\delta = x_2 - x_0$, si ha (per ogni $x \in I$)

$$d(x, x_0) < \delta \Rightarrow x_0 < x < x_2 \Rightarrow f(x_0) < f(x) < f(x_2) \Rightarrow d(f(x), f(x_0)) < \varepsilon,$$

per cui f è continua in x_0 . Il caso eventuale in cui $x_0 = \max I$ si tratta in modo analogo.

Infine, si può vedere che

$$\begin{aligned} f \text{ strettamente crescente} &\Rightarrow f^{-1} \text{ strettamente crescente,} \\ f \text{ strettamente decrescente} &\Rightarrow f^{-1} \text{ strettamente decrescente.} \end{aligned}$$

Quindi, se f è strettamente monotona, anche f^{-1} lo è, e pertanto è anche continua. ■

La funzione esponenziale

Indichiamo con $\mathbb{R}_P$ l'insieme dei numeri reali positivi:

$$\mathbb{R}_P =]0, +\infty[= \{x \in \mathbb{R} : x > 0\}.$$

Enunciamo senza dimostrare il seguente risultato.

Teorema. Dato $a > 0$, esiste un'unica funzione continua $f : \mathbb{R} \rightarrow \mathbb{R}_P$ tale che

$$(i) \quad f(x_1 + x_2) = f(x_1)f(x_2), \text{ per ogni } x_1, x_2 \text{ in } \mathbb{R},$$

$$(ii) \quad f(1) = a.$$

Se inoltre $a \neq 1$, tale funzione è invertibile.

La funzione f si chiama “esponenziale di base a ” e si denota con $\exp_a$. Se $a \neq 1$, la funzione inversa $f^{-1} : \mathbb{R}_P \rightarrow \mathbb{R}$ si chiama “logaritmo di base a ” e si denota con $\log_a$. Per quanto visto sopra, è una funzione continua. Si ha quindi, per $x \in \mathbb{R}$ e $y \in \mathbb{R}_P$,

$$\exp_a(x) = y \quad \Leftrightarrow \quad x = \log_a(y).$$

Dalle proprietà dell'esponenziale

$$(i) \quad \exp_a(x_1 + x_2) = \exp_a(x_1) \exp_a(x_2),$$

$$(ii) \quad \exp_a(1) = a,$$

seguono le corrispondenti proprietà del logaritmo

$$(i') \quad \log_a(y_1 y_2) = \log_a(y_1) + \log_a(y_2),$$

$$(ii') \quad \log_a(a) = 1.$$

Siccome la funzione costante $f(x) = 1$ verifica (i) e (ii) con $a = 1$, si ha che $f = \exp_1$; in altri termini, $\exp_1(x) = 1$ per ogni x .

Vediamo ora alcune proprietà della funzione esponenziale. Osserviamo innanzitutto che, siccome $\exp_a(1) = \exp_a(1 + 0) = \exp_a(1) \exp_a(0)$, si ha che

$$\exp_a(0) = 1.$$

Dimostriamo che, per ogni $x \in \mathbb{R}$ e ogni $n \in \mathbb{N}$, si ha che

$$\exp_a(nx) = (\exp_a(x))^n.$$

Procediamo per induzione. Se $n = 0$, vediamo che

$$\exp_a(0x) = 1, \quad (\exp_a(x))^0 = 1,$$

per cui l'uguaglianza è sicuramente verificata. Supponiamo ora che la formula valga per un certo $n \in \mathbb{N}$. Allora

$$\begin{aligned} \exp_a((n+1)x) &= \exp_a(nx + x) = \exp_a(nx) \exp_a(x) \\ &= (\exp_a(x))^n \exp_a(x) = (\exp_a(x))^{n+1}, \end{aligned}$$

per cui essa vale anche per $n+1$. La dimostrazione è così completa.

Prendendo $x = 1$, vediamo che

$$\exp_a(n) = a^n,$$

per ogni $n \in \mathbb{N}$. Per queste analogie con le potenze, spesso si scrive a^x invece di $\exp_a(x)$, per ogni $x \in \mathbb{R}$.

Se prendiamo $n \in \mathbb{N} \setminus \{0\}$ e scriviamo

$$a = \exp_a(1) = \exp_a\left(n \frac{1}{n}\right) = \left(\exp_a\left(\frac{1}{n}\right)\right)^n,$$

vediamo che $\exp_a\left(\frac{1}{n}\right)$ è quel numero $u \in \mathbb{R}_P$ che risolve l'equazione $u^n = a$. Tale u è la “radice n -esima di a ” e si scrive $u = \sqrt[n]{a}$: si ha quindi

$$\exp_a\left(\frac{1}{n}\right) = \sqrt[n]{a}.$$

Da qui, se $m \in \mathbb{N}$ e $n \in \mathbb{N} \setminus \{0\}$,

$$\exp_a\left(\frac{m}{n}\right) = \exp_a\left(m \frac{1}{n}\right) = \left(\exp_a\left(\frac{1}{n}\right)\right)^m = (\sqrt[n]{a})^m.$$

D'altra parte, scrivendo

$$1 = \exp_a(0) = \exp_a(x - x) = \exp_a(x) \exp_a(-x),$$

vediamo che vale la formula

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

In particolare, se $m \in \mathbb{N}$ e $n \in \mathbb{N} \setminus \{0\}$,

$$\exp_a\left(\frac{-m}{n}\right) = \exp_a\left(-\frac{m}{n}\right) = \frac{1}{\exp_a\left(\frac{m}{n}\right)} = \frac{1}{(\sqrt[n]{a})^m} = (\sqrt[n]{a})^{-m}.$$

(Qui abbiamo usato la ben nota convenzione $x^{-m} = (x^{-1})^m = 1/x^m$, per ogni $x \neq 0$. Supponiamo inoltre note le proprietà delle potenze intere.)

Dimostriamo ora che, per ogni $k \in \mathbb{Z}$, si ha

$$(\sqrt[n]{a})^k = \sqrt[n]{a^k}.$$

Infatti, se $b = \sqrt[n]{a}$, si ha $a^k = (b^n)^k = b^{nk} = (b^k)^n$, da cui $b^k = \sqrt[n]{a^k}$.

Possiamo così concludere che

$$\exp_a\left(\frac{m}{n}\right) = \sqrt[n]{a^m}, \quad \text{per ogni } \frac{m}{n} \in \mathbb{Q}.$$

Se $a \neq 1$, la funzione esponenziale $\exp_a : \mathbb{R} \rightarrow \mathbb{R}_P$, è continua e invertibile, quindi strettamente monotona. Siccome $\exp_a(0) = 1$ e $\exp_a(1) = a$, avremo che

$$\exp_a \text{ è : } \begin{cases} \text{strettamente crescente} & \text{se } a > 1; \\ \text{strettamente decrescente} & \text{se } 0 < a < 1. \end{cases}$$

Lo stesso dicasi per il logaritmo $\log_a$.

Enunciamo infine le seguenti tre proprietà dell'esponenziale:

$$(ab)^x = a^x b^x, \quad \left(\frac{1}{a}\right)^x = \frac{1}{a^x} = a^{-x}, \quad (a^y)^x = a^{yx}.$$

La prima segue dal fatto che la funzione $f(x) = a^x b^x$ verifica la proprietà (i) e $f(1) = ab$, per cui $f = \exp_{ab}$. La seconda è analoga, prendendo $f(x) = \frac{1}{a^x}$; per la terza, si prenda $f(x) = a^{yx}$.

Concludiamo con due utili proprietà del logaritmo:

$$\log_a(x^y) = y \log_a(x), \quad \log_b(x) = \frac{\log_a(x)}{\log_a(b)}.$$

Verifichiamo la prima: poniamo $u = \log_a(x^y)$ e $v = \log_a(x)$. Allora $a^u = x^y$ e $a^v = x$, da cui $a^u = (a^v)^y = a^{vy}$. Ne segue che $u = vy$, che è quanto volevasi dimostrare. Un procedimento analogo permette di verificare anche la seconda.

Le funzioni trigonometriche

Vogliamo ora introdurre le funzioni trigonometriche, in un modo analogo a quanto fatto per la funzione esponenziale.

Dato $T > 0$, una funzione $F : \mathbb{R} \rightarrow \Omega$ (qui Ω è un insieme qualsiasi) si dice “periodica di periodo T ” se

$$F(x + T) = F(x),$$

per ogni $x \in \mathbb{R}$. Chiaramente, se T è un periodo per la funzione F , anche $2T, 3T, \dots$ lo sono. Diremo che T è il “periodo minimo” se non ci sono periodi più piccoli. Introduciamo l'insieme

$$S^1 = \{z \in \mathbb{C} : |z| = 1\}.$$

Si tratta della circonferenza centrata nell'origine, di raggio 1, pensata come sottoinsieme del campo complesso. Enunciamo senza dimostrare il seguente risultato.

Teorema. Dato $T > 0$, esiste un'unica funzione $h_T : \mathbb{R} \rightarrow S^1$, continua e periodica di periodo minimo T , tale che

- (j) $h_T(x_1 + x_2) = h_T(x_1)h_T(x_2)$, per ogni x_1, x_2 in $\mathbb{R}$,
- (jj) $h_T\left(\frac{T}{4}\right) = i$.

La funzione h_T si chiama “funzione circolare di base T ”. Pensando al codominio S^1 come sottoinsieme di $\mathbb{R}^2$, la funzione h_T ha due componenti, che denotiamo con $\cos_T$ e $\sin_T$: sono il “coseno di base T ” e il “seno di base T ”, rispettivamente. Scriveremo quindi

$$h_T(x) = (\cos_T(x), \sin_T(x)), \quad \text{oppure} \quad h_T(x) = \cos_T(x) + i \sin_T(x),$$

per ogni $x \in \mathbb{R}$. Tali funzioni sono continue e periodiche di periodo T , e dalle proprietà della funzione circolare ricaviamo le seguenti:

- (a) $(\cos_T(x))^2 + (\sin_T(x))^2 = 1$,
- (b) $\cos_T(x_1 + x_2) = \cos_T(x_1)\cos_T(x_2) - \sin_T(x_1)\sin_T(x_2)$,
- (c) $\sin_T(x_1 + x_2) = \sin_T(x_1)\cos_T(x_2) + \cos_T(x_1)\sin_T(x_2)$,
- (d) $\cos_T\left(\frac{T}{4}\right) = 0$, $\sin_T\left(\frac{T}{4}\right) = 1$.

Concentriamo ora l'attenzione sull'intervallo $[0, T[$. Scrivendo

$$i = h_T\left(\frac{T}{4}\right) = h_T\left(0 + \frac{T}{4}\right) = h_T(0)h_T\left(\frac{T}{4}\right) = h_T(0)i,$$

ne ricaviamo che $h_T(0) = 1$. Inoltre,

$$h_T\left(\frac{T}{2}\right) = h_T\left(\frac{T}{4} + \frac{T}{4}\right) = h_T\left(\frac{T}{4}\right)h_T\left(\frac{T}{4}\right) = i^2 = -1,$$

mentre

$$h_T\left(\frac{3T}{4}\right) = h_T\left(\frac{T}{2} + \frac{T}{4}\right) = h_T\left(\frac{T}{2}\right)h_T\left(\frac{T}{4}\right) = (-1)i = -i,$$

Riassumendo:

$$\begin{aligned} \cos_T(0) &= 1, & \sin_T(0) &= 0, \\ \cos_T\left(\frac{T}{4}\right) &= 0, & \sin_T\left(\frac{T}{4}\right) &= 1, \\ \cos_T\left(\frac{T}{2}\right) &= -1, & \sin_T\left(\frac{T}{2}\right) &= 0, \\ \cos_T\left(\frac{3T}{4}\right) &= 0, & \sin_T\left(\frac{3T}{4}\right) &= -1. \end{aligned}$$

Le funzioni $\cos_T : \mathbb{R} \rightarrow \mathbb{R}$ e $\sin_T : \mathbb{R} \rightarrow \mathbb{R}$ sono continue. Infatti, fissato $x_0 \in \mathbb{R}$, si ha che

$$h_T(x) - h_T(x_0) = (\cos_T(x) - \cos_T(x_0)) + i(\sin_T(x) - \sin_T(x_0)),$$

da cui

$$\begin{aligned} |\cos_T(x) - \cos_T(x_0)| &\leq |h_T(x) - h_T(x_0)|, \\ |\sin_T(x) - \sin_T(x_0)| &\leq |h_T(x) - h_T(x_0)|. \end{aligned}$$

Osserviamo ora che, dalla

$$1 = h_T(0) = h_T(x - x) = h_T(x)h_T(-x),$$

abbiamo che $h_T(-x) = h_T(x)^{-1} = h_T(x)^*$, essendo $|h_T(x)| = 1$. Quindi,

$$\cos_T(-x) = \cos_T(x), \quad \sin_T(-x) = -\sin_T(x),$$

ossia la funzione $\cos_T$ è pari, mentre $\sin_T$ è dispari.

Dimostreremo ora che $\tilde{h}_T : [0, T[\rightarrow S^1$, la restrizione della funzione circolare h_T all'intervallo $[0, T[$, è biiettiva. Vediamo dapprima l'iniettività. Siano $\alpha < \beta$ in $[0, T[$. Se per assurdo fosse $h_T(\alpha) = h_T(\beta)$, si avrebbe che

$$h_T(\beta - \alpha) = h_T(\beta)h_T(-\alpha) = \frac{h_T(\beta)}{h_T(\alpha)} = 1.$$

Ma allora

$$h_T(x + (\beta - \alpha)) = h_T(x)h_T(\beta - \alpha) = h_T(x),$$

per ogni $x \in \mathbb{R}$, per cui $\beta - \alpha$ sarebbe un periodo di h_T minore di T , mentre sappiamo che T è il periodo minimo.

Vediamo ora che

$$\cos_T(x) \begin{cases} > 0 & \text{se } 0 < x < \frac{T}{4} \\ < 0 & \text{se } \frac{T}{4} < x < \frac{3T}{4} \\ > 0 & \text{se } \frac{3T}{4} < x < T \end{cases}, \quad \sin_T(x) \begin{cases} > 0 & \text{se } 0 < x < \frac{T}{2} \\ < 0 & \text{se } \frac{T}{2} < x < T \end{cases}.$$

Ad esempio, per $x \in]0, \frac{T}{2}[$, non si può certamente avere $\sin_T(x) = 0$, perchè altrimenti i valori in x di $\cos_T$, $\sin_T$ coinciderebbero con i valori in 0 o in $\frac{T}{2}$, mentre abbiamo visto che, se per $\alpha, \beta \in [0, T[$ si ha $\cos_T(\alpha) = \cos_T(\beta)$ e $\sin_T(\alpha) = \sin_T(\beta)$, allora $\alpha = \beta$. Pertanto, per la continuità, $\sin_T$ dovrà essere sempre positiva o sempre negativa in $]0, \frac{T}{2}[$ (teorema degli zeri). Essendo $\sin_T(\frac{T}{4}) = 1$, deve essere sempre positiva.

Per concludere, dimostriamo che $\tilde{h}_T$ è suriettiva (abbiamo già dimostrato prima che è iniettiva). Prendiamo un punto $P = (X_1, X_2) \in S^1$. Si ha che $X_1 \in [-1, 1]$. I due casi in cui $X_1 = -1$ o $X_1 = 1$ si trattano immediatamente, essendo $h_T(\frac{T}{2}) = (-1, 0)$ e $h_T(0) = (1, 0)$. Supponiamo quindi che sia $X_1 \in]-1, 1[$. Sappiamo che $\cos_T(\frac{T}{2}) = -1$, $\cos_T(0) = 1$ e che $\cos_T$ è una funzione continua e T -periodica. Per il corollario al teorema degli zeri, esiste un $\bar{x} \in]0, \frac{T}{2}[$ tale che $\cos_T(\bar{x}) = X_1$. Allora

$$|\sin_T(\bar{x})| = \sqrt{1 - (\cos_T(\bar{x}))^2} = \sqrt{1 - X_1^2} = |X_2|.$$

Abbiamo due possibilità: o $\sin_T(\bar{x}) = X_2$, per cui $h_T(\bar{x}) = P$, oppure $\sin_T(\bar{x}) = -X_2$, nel qual caso

$$h_T(T - \bar{x}) = h_T(-\bar{x}) = h_T(\bar{x})^* = (X_1, X_2) = P.$$

Essendo $T - \bar{x} \in]\frac{T}{2}, T[$, ciò mostra che $\tilde{h}_T$ è suriettiva.

La nozione di limite

Consideriamo due spazi metrici E, F , un punto x_0 di E e una funzione

$$f : E \rightarrow F, \quad \text{oppure} \quad f : E \setminus \{x_0\} \rightarrow F,$$

non necessariamente definita in x_0 .

Definizione. Se esiste un $l \in F$ tale che la funzione $\tilde{f} : E \rightarrow F$, definita da

$$\tilde{f}(x) = \begin{cases} f(x) & \text{se } x \neq x_0, \\ l & \text{se } x = x_0, \end{cases}$$

risulti continua in x_0 , si dice che l è il “limite di f in x_0 ”, o anche “limite di $f(x)$ per x che tende a x_0 ” e si scrive

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

In altri termini, si ha che l è il limite di f in x_0 se

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \forall x \in E \quad 0 < d(x, x_0) < \delta \Rightarrow d(f(x), l) < \varepsilon,$$

o equivalentemente,

$$\forall V, \text{ intorno di } l \quad \exists U, \text{ intorno di } x_0 : f(U \setminus \{x_0\}) \subseteq V.$$

Talvolta si scrive anche $f(x) \rightarrow l$ per $x \rightarrow x_0$.

Sappiamo che, se x_0 è un punto isolato, ogni funzione risulterà continua in x_0 . Il problema non presenta pertanto alcun interesse in questo caso. Supporremo quindi che x_0 non sia un punto isolato, ossia che x_0 sia un “punto di accumulazione” di E : ogni intorno di x_0 contiene punti di E distinti da x_0 stesso.⁴ Nel seguito, supporremo sempre che x_0 sia un punto di accumulazione di E .

Per cominciare, verifichiamo l’unicità del limite.

⁴Un semplice ragionamento mostra che, in questo caso, ogni intorno di x_0 contiene infiniti punti di E . Trovatone uno, x_1 , si prende un intorno di x_0 che non lo contenga, nel quale se ne può trovare un secondo, x_2 , e così via...

Teorema. *Se esiste, il limite di f in x_0 è unico.*

Dimostrazione. Supponiamo per assurdo che ce ne siano due diversi, l e l' . Prendiamo $\varepsilon = \frac{1}{2}d(l, l')$. Allora esiste un $\delta > 0$ tale che

$$0 < d(x, x_0) < \delta \Rightarrow d(f(x), l) < \varepsilon,$$

ed esiste un $\delta' > 0$ tale che

$$0 < d(x, x_0) < \delta' \Rightarrow d(f(x), l') < \varepsilon.$$

Sia $x \neq x_0$ tale che $d(x, x_0) < \delta$ e $d(x, x_0) < \delta'$ (tale x esiste perché x_0 è di accumulazione). Allora

$$d(l', l) \leq d(l, f(x)) + d(f(x), l') < 2\varepsilon = d(l', l),$$

una contraddizione. ■

Il seguente teorema è una riformulazione del legame stretto che intercorre tra i concetti di limite e di continuità.

Teorema. *Considerata la funzione $f : E \rightarrow F$, si ha che*

$$f \text{ è continua in } x_0 \Leftrightarrow \lim_{x \rightarrow x_0} f(x) = f(x_0).$$

Dimostrazione. In questo caso, si ha che la funzione $\tilde{f}$ coincide con f . ■

Un'osservazione generale, che potrebbe essere utile in seguito: si ha

$$\lim_{x \rightarrow x_0} f(x) = l \Leftrightarrow \lim_{x \rightarrow x_0} d(f(x), l) = 0.$$

Esempi. 1. Cominciamo con la funzione $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$, definita da

$$f(x) = \sin_T\left(\frac{1}{x}\right).$$

Se $x_0 = 0$, il limite di f non esiste, perché in ogni intorno di 0 ci sono valori di x per cui $f(x) = 1$ e valori di x per cui $f(x) = -1$.

2. Dimostriamo invece che

$$\lim_{x \rightarrow 0} x \sin_T\left(\frac{1}{x}\right) = 0.$$

Per far questo, è utile osservare che

$$\left| x \sin_T\left(\frac{1}{x}\right) \right| \leq |x|, \quad \text{per ogni } x \in \mathbb{R} \setminus \{0\}.$$

Ecco allora che, fissato $\varepsilon > 0$, basta prendere $\delta = \varepsilon$ per avere che

$$0 < |x - 0| < \delta \Rightarrow |f(x) - 0| < \varepsilon.$$

Funzioni a valori in $\mathbb{R}$

Iniziamo a vedere le proprietà dei limiti che vengono direttamente ereditate dalle funzioni continue. Nei due teoremi seguenti, con relativo corollario, le funzioni f e g sono definite su E o su $E \setminus \{x_0\}$, indifferentemente, e hanno valori in $F = \mathbb{R}$.

Teorema (della permanenza del segno). *Se*

$$\lim_{x \rightarrow x_0} g(x) > 0,$$

allora esiste un $\delta > 0$ tale che

$$0 < d(x, x_0) < \delta \Rightarrow g(x) > 0.$$

Analogamente, se

$$\lim_{x \rightarrow x_0} g(x) < 0,$$

allora esiste un $\delta > 0$ tale che

$$0 < d(x, x_0) < \delta \Rightarrow g(x) < 0.$$

Corollario. *Se $g(x) \leq 0$ per ogni x in un intorno di x_0 , allora, qualora il limite esista, si ha*

$$\lim_{x \rightarrow x_0} g(x) \leq 0.$$

Analogamente, se $g(x) \geq 0$ per ogni x in un intorno di x_0 , allora, qualora il limite esista, si ha

$$\lim_{x \rightarrow x_0} g(x) \geq 0.$$

Naturalmente, si hanno enunciati analoghi qualora g sia di segno opposto.

Teorema. *Se*

$$l_1 = \lim_{x \rightarrow x_0} f(x), \quad l_2 = \lim_{x \rightarrow x_0} g(x),$$

allora

$$\lim_{x \rightarrow x_0} [f(x) + g(x)] = l_1 + l_2,$$

$$\lim_{x \rightarrow x_0} [f(x) - g(x)] = l_1 - l_2,$$

$$\lim_{x \rightarrow x_0} [f(x)g(x)] = l_1 l_2;$$

se $l_2 \neq 0$,

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \frac{l_1}{l_2}.$$

Consideriamo ora il caso in cui E è uno spazio metrico qualunque ed $F = \mathbb{R}$. Risulterà talvolta utile il seguente “teorema dei due carabinieri”.

Teorema. Supponiamo di avere due funzioni f_1, f_2 per cui

$$\lim_{x \rightarrow x_0} f_1(x) = \lim_{x \rightarrow x_0} f_2(x) = l.$$

Se $f : E \setminus \{x_0\} \rightarrow \mathbb{R}$ è tale che, per ogni x ,

$$f_1(x) \leq f(x) \leq f_2(x),$$

allora

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

Dimostrazione. Fissato $\varepsilon > 0$, esistono $\delta_1 > 0$ e $\delta_2 > 0$ tali che

$$0 < d(x, x_0) < \delta_1 \Rightarrow l - \varepsilon < f_1(x) < l + \varepsilon,$$

$$0 < d(x, x_0) < \delta_2 \Rightarrow l - \varepsilon < f_2(x) < l + \varepsilon.$$

Se $\delta = \min\{\delta_1, \delta_2\}$, allora

$$0 < d(x, x_0) < \delta \Rightarrow l - \varepsilon < f_1(x) \leq f(x) \leq f_2(x) < l + \varepsilon,$$

il che dimostra la tesi. ■

Corollario. Se

$$\lim_{x \rightarrow x_0} f(x) = 0,$$

ed esiste un $C > 0$ tale che $|g(x)| \leq C$ per ogni x , allora

$$\lim_{x \rightarrow x_0} f(x)g(x) = 0.$$

Dimostrazione. Si ha

$$-C|f(x)| \leq f(x)g(x) \leq C|f(x)|,$$

e il risultato segue dal teorema precedente, tenuto conto che si ha

$$\lim_{x \rightarrow x_0} f(x) = 0 \Leftrightarrow \lim_{x \rightarrow x_0} |f(x)| = 0,$$

prendendo $f_1(x) = -C|f(x)|$ e $f_2(x) = C|f(x)|$. ■

Cambiamento di variabile nel limite

Consideriamo ora una funzione composta $g \circ f$. Abbiamo due possibili situazioni.

Teorema 1. *Sia $f : E \rightarrow F$, oppure $f : E \setminus \{x_0\} \rightarrow F$, tale che*

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

Se $g : F \rightarrow G$ è continua in l , allora

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

In altri termini,

$$\lim_{x \rightarrow x_0} g(f(x)) = g(\lim_{x \rightarrow x_0} f(x)).$$

Dimostrazione. Riguardando la definizione di limite, si ha che $\tilde{f} : E \rightarrow \mathbb{R}$ ivi definita è continua in x_0 e g è continua in $l = \tilde{f}(x_0)$. Pertanto, $g \circ \tilde{f}$ è continua in x_0 , da cui

$$\lim_{x \rightarrow x_0} g(f(x)) = \lim_{x \rightarrow x_0} g(\tilde{f}(x)) = g(\tilde{f}(x_0)) = g(l).$$

■

Teorema 2. *Sia $f : E \rightarrow F$, oppure $f : E \setminus \{x_0\} \rightarrow F$, tale che*

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

Supponiamo che l sia un punto di accumulazione di F e che la funzione

$$g : F \rightarrow G, \quad \text{oppure} \quad g : F \setminus \{l\} \rightarrow G,$$

non necessariamente definita in l , sia tale che

$$\lim_{y \rightarrow l} g(y) = L.$$

Se $f(x) \neq l$ per ogni $x \in E \setminus \{x_0\}$, allora

$$\lim_{x \rightarrow x_0} g(f(x)) = L.$$

Dimostrazione. Consideriamo nuovamente la funzione $\tilde{f} : E \rightarrow F$, continua in x_0 con $\tilde{f}(x_0) = l$. Analogamente, consideriamo la funzione $\tilde{g} : F \rightarrow G$ così definita:

$$\tilde{g}(y) = \begin{cases} g(y) & \text{se } y \neq l, \\ L & \text{se } y = l. \end{cases}$$

Essa è continua in l con $\tilde{g}(l) = L$. Consideriamo la funzione composta $\tilde{g} \circ \tilde{f}$, che per quanto sopra è continua in x_0 con $\tilde{g}(\tilde{f}(x_0)) = \tilde{g}(l) = L$. Essendo $f(x) \neq l$ per ogni x , si ha che, per $x \in E \setminus \{x_0\}$,

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

e pertanto,

$$\lim_{x \rightarrow x_0} g(f(x)) = \lim_{x \rightarrow x_0} \tilde{g}(\tilde{f}(x)) = \tilde{g}(\tilde{f}(x_0)) = L.$$

■

Alcune considerazioni sull'ultimo teorema dimostrato. Si noti che la sua conclusione si riassume con la formula

$$\lim_{x \rightarrow x_0} g(f(x)) = \lim_{y \rightarrow \lim_{x \rightarrow x_0} f(x)} g(y).$$

Spesso si dice che si è operato il “cambio di variabile $y = f(x)$ ”. Riguardando inoltre le ipotesi dello stesso teorema, si vede subito che è sufficiente richiedere che sia $f(x) \neq l$ per gli x tali che $0 < d(x, x_0) < \delta$. Ciò è dovuto al fatto che la nozione di limite è, in un certo senso, di tipo “locale”. Questa osservazione vale in generale e verrà spesso usata in seguito.

Esempi. 1. Dimostriamo che

$$\lim_{x \rightarrow 0} \cos\left(x \sin\left(\frac{1}{x}\right)\right) = 1.$$

In effetti, se $f : \mathbb{R} \setminus \{0\} \rightarrow \mathbb{R}$ è definita da $f(x) = x \sin(1/x)$ e $g : \mathbb{R} \rightarrow \mathbb{R}$ è definita da $g(y) = \cos(y)$, sappiamo che $\lim_{x \rightarrow 0} f(x) = 0$, e che g è continua. Per il Teorema 1,

$$\lim_{x \rightarrow 0} g(f(x)) = g\left(\lim_{x \rightarrow 0} f(x)\right) = g(0) = 1.$$

2. Sia ora f come nell'esempio precedente, e sia $g : \mathbb{R} \rightarrow \mathbb{R}$ definita da

$$g(y) = \begin{cases} 1 & \text{se } y \neq 0, \\ 2 & \text{se } y = 0. \end{cases}$$

Si può vedere che, in ogni intorno di $x_0 = 0$, la funzione $g(f(x))$ assume infinite volte il valore 1 e infinite volte il valore 2. Pertanto, in questo caso,

$$\text{il limite } \lim_{x \rightarrow 0} g(f(x)) \text{ non esiste.}$$

Limite delle restrizioni

Finora abbiamo considerato due spazi metrici E, F , un punto x_0 di accumulazione per E e una funzione $f : E \rightarrow F$, oppure $f : E \setminus \{x_0\} \rightarrow F$. Siccome l'eventuale valore di f in x_0 è ininfluente ai fini dell'esistenza o meno del limite, nonché del suo effettivo valore, da ora in poi per semplicità considereremo solo il caso $f : E \setminus \{x_0\} \rightarrow F$.

Si può verificare che tutte le considerazioni fatte continuano a valere per una funzione $f : \hat{E} \setminus \{x_0\} \rightarrow F$, con $\hat{E} \subseteq E$, purché x_0 sia di accumulazione per $\hat{E}$: ogni intorno di x_0 deve contenere infiniti punti di $\hat{E}$.

Sia ora $f : E \setminus \{x_0\} \rightarrow F$, e sia $\hat{E} \subseteq E$. Possiamo considerare la restrizione di f a $\hat{E} \setminus \{x_0\}$: è la funzione $\hat{f} : \hat{E} \setminus \{x_0\} \rightarrow F$ i cui valori coincidono con quelli di f : si ha $\hat{f}(x) = f(x)$ per ogni $x \in \hat{E} \setminus \{x_0\}$. Talvolta si scrive $\hat{f} = f|_{\hat{E}}$.

Teorema. *Se esiste il limite di f in x_0 e x_0 è di accumulazione anche per $\hat{E}$, allora esiste anche il limite di $\hat{f}$ in x_0 e ha lo stesso valore:*

$$\lim_{x \rightarrow x_0} \hat{f}(x) = \lim_{x \rightarrow x_0} f(x).$$

Dimostrazione. Segue immediatamente dalla definizione di $\hat{f}$. ■

Il teorema precedente viene spesso usato per stabilire la non esistenza del limite per la funzione f : a tal scopo, è sufficiente trovare due diverse restrizioni lungo le quali i valori del limite differiscono.

Esempi. 1. La funzione $f : \mathbb{R}^2 \setminus \{(0,0)\} \rightarrow \mathbb{R}$, definita da

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

non ha limite per $(x, y) \rightarrow (0, 0)$, come si vede considerando le restrizioni alle due rette $\{(x, y) : x = 0\}$ e $\{(x, y) : x = y\}$.

2. Più sorprendente è la funzione definita da

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

per la quale le restrizioni a tutte le rette passanti per $(0, 0)$ hanno limite 0, ma la restrizione alla parabola $\{(x, y) : y = x^2\}$ vale costantemente $\frac{1}{2}$.

3. Dimostriamo invece che

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

Fissiamo un $\varepsilon > 0$. Dopo aver verificato che

$$\frac{x^2 y^2}{x^2 + y^2} \leq \frac{1}{2} (x^2 + y^2),$$

risulta naturale prendere $\delta = \sqrt{2\varepsilon}$, per avere che

$$d((x, y), (0, 0)) < \delta \Rightarrow \left| \frac{x^2 y^2}{x^2 + y^2} - 0 \right| < \varepsilon.$$

Sia ora $E \subseteq \mathbb{R}$. Possiamo considerare le due restrizioni $\hat{f}_1$ e $\hat{f}_2$ agli insiemi $\hat{E}_1 = E \cap]-\infty, x_0]$ e $\hat{E}_2 = E \cap [x_0, +\infty[$. Se x_0 è di accumulazione per $\hat{E}_1$, chiameremo “limite sinistro” di f , quando esiste, il limite di $\hat{f}_1(x)$ per x che tende a x_0 ; lo denoteremo con

$$\lim_{x \rightarrow x_0^-} f(x).$$

Analogamente, se x_0 è di accumulazione per $\hat{E}_2$, chiameremo “limite destro” di f , quando esiste, il limite di $\hat{f}_2(x)$ per x che tende a x_0 ; lo denoteremo con

$$\lim_{x \rightarrow x_0^+} f(x).$$

Si può dimostrare il fatto seguente.

Teorema. *Se x_0 è di accumulazione per $\hat{E}_1$ e per $\hat{E}_2$, il limite di $f(x)$ per x che tende a x_0 esiste se e solo se esistono sia il limite sinistro che il limite destro e hanno lo stesso valore.*

Dimostrazione. Sappiamo già che, se esiste il limite, tutte le restrizioni devono avere lo stesso limite. Viceversa, supponiamo che esistano e coincidano i limiti sinistro e destro, e sia ℓ il loro valore. Fissiamo un $\varepsilon > 0$. Allora esistono $\delta_1 > 0$ e $\delta_2 > 0$ tali che, se $x \in E$,

$$x_0 - \delta_1 < x < x_0 \Rightarrow d(f(x), \ell) < \varepsilon,$$

$$x_0 < x < x_0 + \delta_2 \Rightarrow d(f(x), \ell) < \varepsilon.$$

Preso $\delta = \min\{\delta_1, \delta_2\}$, abbiamo quindi che, se $x \neq x_0$,

$$x_0 - \delta < x < x_0 + \delta \Rightarrow d(f(x), \ell) < \varepsilon,$$

per cui il limite di f in x_0 esiste ed è uguale a ℓ . ■

Esempio. La funzione “segno”, ossia $f : \mathbb{R} \rightarrow \mathbb{R}$ definita da

$$f(x) = \begin{cases} 1 & \text{se } x > 0 \\ 0 & \text{se } x = 0 \\ -1 & \text{se } x < 0 \end{cases}$$

non ha limite in $x_0 = 0$, essendo che $\lim_{x \rightarrow 0^-} f(x) = -1$ e $\lim_{x \rightarrow 0^+} f(x) = 1$.

La retta ampliata

Consideriamo la funzione $\varphi : \mathbb{R} \rightarrow]-1, 1[$, definita da

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

Si tratta di una funzione invertibile, con inversa $\varphi^{-1} :]-1, 1[\rightarrow \mathbb{R}$, definita da

$$\varphi^{-1}(y) = \frac{y}{1 - |y|}.$$

Possiamo allora definire una nuova distanza su $\mathbb{R}$:

$$\tilde{d}(x, x') = |\varphi(x) - \varphi(x')|.$$

Per la nuova distanza, la palla aperta di centro $x_0 \in \mathbb{R}$ e raggio ρ è data da

$$\tilde{B}(x_0, \rho) = \{x : |\varphi(x) - \varphi(x_0)| < \rho\}.$$

È importante notare che gli intorni di un punto $x_0 \in \mathbb{R}$ rimangono gli stessi di quelli definiti dalla distanza usuale in $\mathbb{R}$. Infatti, essendo φ continua in x_0 , per ogni $\rho_1 > 0$ esiste un $\rho_2 > 0$ tale che

$$|x - x_0| < \rho_2 \quad \Rightarrow \quad |\varphi(x) - \varphi(x_0)| < \rho_1,$$

ossia

$$]x_0 - \rho_2, x_0 + \rho_2[\subseteq \tilde{B}(x_0, \rho_1).$$

Viceversa, essendo φ^{-1} continua in $y_0 = \varphi(x_0) \in]-1, 1[$, per ogni $\rho_1 > 0$ esiste un $\rho_2 > 0$ tale che

$$|y - y_0| < \rho_2 \quad \Rightarrow \quad y \in]-1, 1[\quad \text{e} \quad |\varphi^{-1}(y) - \varphi^{-1}(y_0)| < \rho_1;$$

In particolare, prendendo $y = \varphi(x)$,

$$|\varphi(x) - \varphi(x_0)| < \rho_2 \quad \Rightarrow \quad \varphi(x) \in]-1, 1[\quad \text{e} \quad |x - x_0| < \rho_1,$$

ossia

$$\tilde{B}(x_0, \rho_2) \subseteq]x_0 - \rho_1, x_0 + \rho_1[.$$

Da quanto visto, si deduce che ogni intorno per la nuova distanza è anche intorno per la vecchia distanza, e viceversa.

Introduciamo ora il nuovo insieme $\tilde{\mathbb{R}}$, definito come unione di $\mathbb{R}$ e di due nuovi elementi, che indicheremo con $-\infty$ e $+\infty$:

$$\tilde{\mathbb{R}} = \mathbb{R} \cup \{-\infty, +\infty\}.$$

L'insieme $\tilde{\mathbb{R}}$ risulta totalmente ordinato se si mantiene l'ordine esistente tra coppie di numeri reali e si pone inoltre, per ogni $x \in \mathbb{R}$,

$$-\infty < x < +\infty.$$

Consideriamo la funzione $\tilde{\varphi} : \tilde{\mathbb{R}} \rightarrow [-1, 1]$, definita da

$$\tilde{\varphi}(x) = \begin{cases} -1 & \text{se } x = -\infty, \\ \varphi(x) & \text{se } x \in \mathbb{R}, \\ 1 & \text{se } x = +\infty. \end{cases}$$

Essa è invertibile, con inversa $\tilde{\varphi}^{-1} : [-1, 1] \rightarrow \tilde{\mathbb{R}}$ definita da

$$\tilde{\varphi}^{-1}(y) = \begin{cases} -\infty & \text{se } y = -1, \\ \varphi^{-1}(y) & \text{se } y \in]-1, 1[, \\ +\infty & \text{se } y = 1. \end{cases}$$

Definiamo, per $x, x' \in \tilde{\mathbb{R}}$,

$$\tilde{d}(x, x') = |\tilde{\varphi}(x) - \tilde{\varphi}(x')|;$$

si verifica facilmente che $\tilde{d}$ è una distanza su $\tilde{\mathbb{R}}$. In questo modo, $\tilde{\mathbb{R}}$ risulta uno spazio metrico. Vediamo ad esempio cos'è una palla aperta centrata in $+\infty$:

$$B(+\infty, \rho) = \{x \in \tilde{\mathbb{R}} : |\tilde{\varphi}(x) - 1| < \rho\} = \{x \in \tilde{\mathbb{R}} : \tilde{\varphi}(x) > 1 - \rho\},$$

e quindi

$$B(+\infty, \rho) = \begin{cases} \tilde{\mathbb{R}} & \text{se } \rho > 2, \\]-\infty, +\infty] & \text{se } \rho = 2, \\]\varphi^{-1}(1 - \rho), +\infty] & \text{se } \rho < 2, \end{cases}$$

dove abbiamo usato le notazioni

$$]a, +\infty] = \{x \in \tilde{\mathbb{R}} : x > a\} =]a, +\infty[\cup \{+\infty\}.$$

Possiamo quindi affermare che un intorno di $+\infty$ è un insieme che contiene, oltre al punto $+\infty$, un intervallo del tipo $] \alpha, +\infty[$, per un certo $\alpha \in \mathbb{R}$.

Analogamente, un intorno di $-\infty$ è un insieme che contiene, oltre a $-\infty$, un intervallo del tipo $] -\infty, \beta[$, per un certo $\beta \in \mathbb{R}$.

Vediamo ora come si traduce la definizione di limite in alcuni casi in cui compaiono gli elementi $+\infty$ o $-\infty$. Ad esempio, sia $E \subseteq \mathbb{R}$, F uno spazio metrico e $f : E \rightarrow F$ una funzione. Considerando E come sottoinsieme di $\widetilde{\mathbb{R}}$, si ha che $+\infty$ è punto di accumulazione per E se e solo se E non è limitato superiormente. In tal caso, si ha:

$$\begin{aligned} \lim_{x \rightarrow +\infty} f(x) = l \in F &\Leftrightarrow \forall V \text{ intorno di } l \ \exists U \text{ intorno di } +\infty : \\ &f(U \cap E) \subseteq V \\ &\Leftrightarrow \forall \varepsilon > 0 \ \exists \alpha \in \mathbb{R} : \ x > \alpha \Rightarrow d(f(x), l) < \varepsilon. \end{aligned}$$

Analogamente, se E non è limitato inferiormente, si ha:

$$\lim_{x \rightarrow -\infty} f(x) = l \in F \Leftrightarrow \forall \varepsilon > 0 \ \exists \beta \in \mathbb{R} : \ x < \beta \Rightarrow d(f(x), l) < \varepsilon.$$

Si noti che

$$\lim_{x \rightarrow +\infty} f(x) = l \Leftrightarrow \lim_{x \rightarrow -\infty} f(-x) = l.$$

Vediamo ora il caso in cui E sia uno spazio metrico ed $F = \mathbb{R}$, considerato come sottoinsieme di $\widetilde{\mathbb{R}}$. Supponiamo che x_0 sia di accumulazione per E e consideriamo una funzione $f : E \rightarrow \mathbb{R}$, o $f : E \setminus \{x_0\} \rightarrow \mathbb{R}$. Si ha:

$$\begin{aligned} \lim_{x \rightarrow x_0} f(x) = +\infty &\Leftrightarrow \forall V \text{ intorno di } +\infty \ \exists U \text{ intorno di } x_0 : \\ &f(U \setminus \{x_0\}) \subseteq V \\ &\Leftrightarrow \forall \alpha \in \mathbb{R} \ \exists \delta > 0 : \ 0 < d(x, x_0) < \delta \Rightarrow f(x) > \alpha; \end{aligned}$$

analogamente,

$$\lim_{x \rightarrow x_0} f(x) = -\infty \Leftrightarrow \forall \beta \in \mathbb{R} \ \exists \delta > 0 : \ 0 < d(x, x_0) < \delta \Rightarrow f(x) < \beta.$$

Si noti che

$$\lim_{x \rightarrow x_0} f(x) = +\infty \Leftrightarrow \lim_{x \rightarrow x_0} (-f(x)) = -\infty.$$

Le situazioni considerate in precedenza possono talvolta presentarsi assieme. Ad esempio, se $E \subseteq \mathbb{R}$ non è limitato superiormente ed $F = \mathbb{R}$, si avrà

$$\begin{aligned} \lim_{x \rightarrow +\infty} f(x) = +\infty &\Leftrightarrow \forall V \text{ intorno di } +\infty \ \exists U \text{ intorno di } +\infty : \\ &f(U \cap E) \subseteq V \\ &\Leftrightarrow \forall \alpha \in \mathbb{R} \ \exists \alpha' \in \mathbb{R} : \ x > \alpha' \Rightarrow f(x) > \alpha; \end{aligned}$$

analogamente,

$$\lim_{x \rightarrow +\infty} f(x) = -\infty \quad \Leftrightarrow \quad \forall \beta \in \mathbb{R} \quad \exists \alpha \in \mathbb{R} : \quad x > \alpha \Rightarrow f(x) < \beta.$$

Se invece $E \subseteq \mathbb{R}$ non è limitato inferiormente ed $F = \mathbb{R}$, si avrà

$$\lim_{x \rightarrow -\infty} f(x) = +\infty \quad \Leftrightarrow \quad \forall \alpha \in \mathbb{R} \quad \exists \beta \in \mathbb{R} : \quad x < \beta \Rightarrow f(x) > \alpha;$$

analogamente,

$$\lim_{x \rightarrow -\infty} f(x) = -\infty \quad \Leftrightarrow \quad \forall \beta \in \mathbb{R} \quad \exists \beta' \in \mathbb{R} : \quad x < \beta' \Rightarrow f(x) < \beta.$$

Operazioni con i limiti $+\infty$ e $-\infty$

Qualora i limiti siano $+\infty$ o $-\infty$, non si possono usare i teoremi sulle operazioni con i limiti. A titolo illustrativo, enunciamo alcuni teoremi validi in questi casi. Nel seguito, tutte le funzioni saranno definite in uno spazio metrico E , oppure in $E \setminus \{x_0\}$, con x_0 di accumulazione. Iniziamo con l'addizione:

Teorema. Se

$$\lim_{x \rightarrow x_0} f(x) = +\infty$$

ed esiste un $\gamma \in \mathbb{R}$ tale che, per ogni x in un intorno di x_0 ,

$$g(x) \geq \gamma,$$

allora

$$\lim_{x \rightarrow x_0} [f(x) + g(x)] = +\infty.$$

Dimostrazione. Fissiamo $\alpha \in \mathbb{R}$. Considerato $\alpha' = \alpha - \gamma$, esiste un $\delta > 0$ tale che

$$0 < d(x, x_0) < \delta \Rightarrow f(x) > \alpha'.$$

Quindi,

$$0 < d(x, x_0) < \delta \Rightarrow f(x) + g(x) > \alpha' + \gamma = \alpha.$$

■

Corollario. Se

$$\lim_{x \rightarrow x_0} f(x) = +\infty \quad \text{e} \quad \lim_{x \rightarrow x_0} g(x) = l \in \mathbb{R} \quad (\text{o } +\infty),$$

allora

$$\lim_{x \rightarrow x_0} [f(x) + g(x)] = +\infty.$$

Dimostrazione. Se il limite di g è $l \in \mathbb{R}$, esiste un $\delta > 0$ tale che

$$0 < d(x, x_0) < \delta \Rightarrow g(x) > l - 1.$$

Se invece il limite è $+\infty$, esiste un $\delta > 0$ tale che

$$0 < d(x, x_0) < \delta \Rightarrow g(x) > 0.$$

In ogni caso, si può applicare il teorema precedente per concludere. ■

Come regola mnemonica, scriveremo brevemente

$$\begin{aligned} (+\infty) + l &= +\infty, \text{ se } l \text{ è un numero reale;} \\ (+\infty) + (+\infty) &= +\infty. \end{aligned}$$

In modo del tutto analogo, si possono enunciare un teorema e il relativo corollario nel caso in cui il limite di f sia $-\infty$. Come regola mnemonica, scriveremo allora

$$\begin{aligned} (-\infty) + l &= -\infty, \text{ se } l \text{ è un numero reale;} \\ (-\infty) + (-\infty) &= -\infty. \end{aligned}$$

Similmente per quanto riguarda il prodotto:

Teorema. Se

$$\lim_{x \rightarrow x_0} f(x) = +\infty$$

ed esiste un $\gamma > 0$ tale che, per ogni x in un intorno di x_0 ,

$$g(x) \geq \gamma,$$

allora

$$\lim_{x \rightarrow x_0} [f(x)g(x)] = +\infty.$$

Dimostrazione. Fissiamo $\alpha \in \mathbb{R}$. Possiamo supporre che sia $\alpha > 0$. Posto $\alpha' = \frac{\alpha}{\gamma}$, esiste un $\delta > 0$ tale che

$$0 < d(x, x_0) < \delta \Rightarrow f(x) > \alpha'.$$

Quindi,

$$0 < d(x, x_0) < \delta \Rightarrow f(x)g(x) > \alpha'\gamma = \alpha.$$

■

Corollario. Se

$$\lim_{x \rightarrow x_0} f(x) = +\infty \quad e \quad \lim_{x \rightarrow x_0} g(x) = l > 0 \quad (o \quad +\infty),$$

allora

$$\lim_{x \rightarrow x_0} [f(x)g(x)] = +\infty.$$

Dimostrazione. Se il limite di g è un numero reale $l > 0$, esiste un $\delta > 0$ tale che

$$0 < d(x, x_0) < \delta \Rightarrow g(x) > \frac{l}{2}.$$

Se invece il limite è $+\infty$, esiste un $\delta > 0$ tale che

$$0 < d(x, x_0) < \delta \Rightarrow g(x) > 1.$$

In ogni caso, si può applicare il teorema precedente per concludere. ■

Come sopra, scriveremo brevemente

$$\begin{aligned} (+\infty) \cdot l &= +\infty, \text{ se } l > 0 \text{ è un numero reale;} \\ (+\infty) \cdot (+\infty) &= +\infty, \end{aligned}$$

con tutte le varianti del caso:

$$\begin{aligned} (+\infty) \cdot l &= -\infty, \text{ se } l < 0 \text{ è un numero reale;} \\ (-\infty) \cdot l &= -\infty, \text{ se } l > 0 \text{ è un numero reale;} \\ (-\infty) \cdot l &= +\infty, \text{ se } l < 0 \text{ è un numero reale;} \\ (+\infty) \cdot (-\infty) &= -\infty; \\ (-\infty) \cdot (-\infty) &= +\infty. \end{aligned}$$

Passiamo ora a un altro tipo di risultati.

Teorema. Se

$$\lim_{x \rightarrow x_0} |f(x)| = +\infty,$$

allora

$$\lim_{x \rightarrow x_0} \frac{1}{f(x)} = 0.$$

Dimostrazione. Fissiamo un $\varepsilon > 0$. Posto $\alpha = \frac{1}{\varepsilon}$, esiste un $\delta > 0$ tale che

$$0 < d(x, x_0) < \delta \Rightarrow |f(x)| > \alpha.$$

Quindi,

$$0 < d(x, x_0) < \delta \Rightarrow \left| \frac{1}{f(x)} - 0 \right| = \frac{1}{|f(x)|} < \frac{1}{\alpha} = \varepsilon. \quad \blacksquare$$

Teorema. Se

$$\lim_{x \rightarrow x_0} f(x) = 0$$

e $f(x) > 0$ per ogni x in un intorno di x_0 , allora

$$\lim_{x \rightarrow x_0} \frac{1}{f(x)} = +\infty.$$

Se invece $f(x) < 0$ per ogni x in un intorno di x_0 , allora

$$\lim_{x \rightarrow x_0} \frac{1}{f(x)} = -\infty.$$

Dimostrazione. Vediamo solo il primo caso, essendo il secondo analogo. Fissiamo $\alpha \in \mathbb{R}$; possiamo supporre $\alpha > 0$. Posto $\varepsilon = \frac{1}{\alpha}$, esiste un $\delta > 0$ tale che

$$0 < d(x, x_0) < \delta \Rightarrow 0 < f(x) < \varepsilon.$$

Allora,

$$0 < d(x, x_0) < \delta \Rightarrow \frac{1}{f(x)} > \frac{1}{\varepsilon} = \alpha.$$

■

Presentiamo due varianti del teorema dei due carabinieri: nel caso in cui il limite vale $+\infty$, si ha il seguente

Teorema. Sia f_1 tale che

$$\lim_{x \rightarrow x_0} f_1(x) = +\infty.$$

Se f è tale che, per ogni x in un intorno di x_0 ,

$$f_1(x) \leq f(x),$$

allora

$$\lim_{x \rightarrow x_0} f(x) = +\infty.$$

Dimostrazione. Ponendo $g(x) = f(x) - f_1(x)$, si ha che $g(x) \geq 0$ per ogni x in un intorno di x_0 e $f(x) = f_1(x) + g(x)$. Il risultato segue quindi direttamente dal primo teorema visto a lezione. ■

Nel caso in cui il limite sia $-\infty$, si ha l'analogo

Teorema. Sia f_2 tale che

$$\lim_{x \rightarrow x_0} f_2(x) = -\infty.$$

Se f è tale che, per ogni x in un intorno di x_0 ,

$$f(x) \leq f_2(x),$$

allora

$$\lim_{x \rightarrow x_0} f(x) = -\infty.$$

Calcoleremo ora alcuni limiti elementari per x che tende a $+\infty$ o $-\infty$. Consideriamo la funzione

$$f(x) = x^n,$$

dove n è un numero intero. Si può verificare per induzione che, se $n \geq 1$,

$$x \geq 1 \quad \Rightarrow \quad x^n \geq x.$$

Siccome chiaramente $\lim_{x \rightarrow +\infty} x = +\infty$, ne segue che

$$\lim_{x \rightarrow +\infty} x^n = \begin{cases} +\infty & \text{se } n \geq 1, \\ 1 & \text{se } n = 0, \\ 0 & \text{se } n \leq -1. \end{cases}$$

Tenendo poi conto che

$$(-x)^n = x^n \text{ se } n \text{ è pari}, \quad (-x)^n = -x^n \text{ se } n \text{ è dispari},$$

si vede che

$$\lim_{x \rightarrow -\infty} x^n = \begin{cases} +\infty & \text{se } n \geq 1 \text{ è pari}, \\ -\infty & \text{se } n \geq 1 \text{ è dispari}, \\ 1 & \text{se } n = 0, \\ 0 & \text{se } n \leq -1. \end{cases}$$

Consideriamo ora la funzione polinomiale

$$f(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0,$$

dove $n \geq 1$ e $a_n \neq 0$. Scrivendo

$$f(x) = x^n \left(a_n + \frac{a_{n-1}}{x} + \cdots + \frac{a_2}{x^{n-2}} + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n} \right)$$

e usando il fatto che

$$\lim_{x \rightarrow +\infty} \left(a_n + \frac{a_{n-1}}{x} + \cdots + \frac{a_2}{x^{n-2}} + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n} \right) = a_n,$$

si vede che

$$\lim_{x \rightarrow +\infty} f(x) = \begin{cases} +\infty & \text{se } a_n > 0, \\ -\infty & \text{se } a_n < 0, \end{cases}$$

mentre

$$\lim_{x \rightarrow -\infty} f(x) = \begin{cases} +\infty & \text{se } [n \text{ è pari e } a_n > 0], \text{ oppure } [n \text{ è dispari e } a_n < 0], \\ -\infty & \text{se } [n \text{ è pari e } a_n < 0], \text{ oppure } [n \text{ è dispari e } a_n > 0]. \end{cases}$$

Consideriamo ora una funzione razionale

$$f(x) = \frac{a_n x^n + a_{n-1} x^{n-1} + \dots + a_2 x^2 + a_1 x + a_0}{b_m x^m + b_{m-1} x^{m-1} + \dots + b_2 x^2 + b_1 x + b_0},$$

dove $n, m \geq 1$ e $a_n, b_m \neq 0$. Similmente a quanto sopra, scrivendo

$$f(x) = x^{n-m} \frac{a_n + \frac{a_{n-1}}{x} + \dots + \frac{a_2}{x^{n-2}} + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n}}{b_m + \frac{b_{m-1}}{x} + \dots + \frac{b_2}{x^{m-2}} + \frac{b_1}{x^{m-1}} + \frac{b_0}{x^m}},$$

possiamo concludere che

$$\lim_{x \rightarrow +\infty} f(x) = \begin{cases} +\infty & \text{se } n > m \text{ e } a_n, b_m \text{ hanno lo stesso segno}, \\ -\infty & \text{se } n > m \text{ e } a_n, b_m \text{ hanno segno opposto}, \\ \frac{a_n}{b_m} & \text{se } n = m, \\ 0 & \text{se } n < m. \end{cases}$$

Può risultare utile osservare che

$$\lim_{x \rightarrow +\infty} f(x) = \lim_{x \rightarrow +\infty} \frac{a_n}{b_m} x^{n-m}.$$

In modo analogo si vede che

$$\lim_{x \rightarrow -\infty} f(x) = \lim_{x \rightarrow -\infty} \frac{a_n}{b_m} x^{n-m},$$

con tutta la casistica che ne consegue.

Limiti di funzioni monotone

Vedremo ora che la monotonia di una funzione f permette di stabilire l'esistenza del limite sinistro e del limite destro.

Siano E un sottoinsieme di $\mathbb{R}$, x_0 un punto di accumulazione per $E \cap]-\infty, x_0[$, e consideriamo una funzione monotona $f : E \cap]-\infty, x_0[\rightarrow \mathbb{R}$.

Teorema. Se f è crescente, allora

$$\lim_{x \rightarrow x_0^-} f(x) = \sup f(E \cap] - \infty, x_0[).$$

Se invece f è decrescente, allora

$$\lim_{x \rightarrow x_0^-} f(x) = \inf f(E \cap] - \infty, x_0[).$$

Dimostrazione. Dimostriamo la prima affermazione (la dimostrazione della seconda è analoga). Sia $s = \sup f(E \cap] - \infty, x_0[)$. Se $s \in \mathbb{R}$, fissiamo $\varepsilon > 0$. Per le proprietà dell'estremo superiore, esiste un $\bar{y} \in f(E \cap] - \infty, x_0[)$ tale che $\bar{y} > s - \varepsilon$. Quindi, preso $\bar{x} \in E \cap] - \infty, x_0[$ tale che $f(\bar{x}) = \bar{y}$, per la crescenza di f abbiamo

$$\bar{x} < x < x_0 \Rightarrow s - \varepsilon < f(x) \leq s,$$

il che completa la dimostrazione in questo caso.

Se invece $s = +\infty$, fissiamo un $\alpha \in \mathbb{R}$. Allora esiste un $\bar{x} \in E \cap] - \infty, x_0[$ tale che $f(\bar{x}) > \alpha$. Per la crescenza di f ,

$$\bar{x} < x < x_0 \Rightarrow f(x) > \alpha,$$

per cui $\lim_{x \rightarrow x_0^-} f(x) = +\infty$. ■

Si osservi che il teorema precedente include anche il caso in cui $x_0 = +\infty$. Si può enunciare un teorema analogo nel caso in cui x_0 è un punto di accumulazione per $E \cap]x_0, +\infty[$ e $f : E \cap]x_0, +\infty[\rightarrow \mathbb{R}$ è monotona (qui potrebbe anche essere $x_0 = -\infty$).

Teorema. Se f è crescente, allora

$$\lim_{x \rightarrow x_0^+} f(x) = \inf f(E \cap]x_0, +\infty[).$$

Se invece f è decrescente, allora

$$\lim_{x \rightarrow x_0^+} f(x) = \sup f(E \cap]x_0, +\infty[).$$

Per quanto riguarda le successioni di numeri reali, possiamo enunciare il seguente

Corollario. Ogni successione monotona di numeri reali ha limite.

Dimostrazione. Se $(a_n)_n$ è crescente, allora

$$\lim_n a_n = \sup \{a_n : n \in \mathbb{N}\},$$

e questo limite può essere un numero reale o $+\infty$. Similmente, se $(a_n)_n$ è decrescente, il limite sarà un numero reale o $-\infty$. ■

Consideriamo ad esempio la successione $(a_n)_n$, così definita per $n \geq 1$:

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

Vediamo che $(a_n)_n$ è crescente:

$$\begin{aligned} \frac{a_{n+1}}{a_n} &= \frac{\left(1 + \frac{1}{n+1}\right)^{n+1}}{\left(1 + \frac{1}{n}\right)^n} \\ &= \left(\frac{n+2}{n+1}\right)^{n+1} \left(\frac{n}{n+1}\right)^{n+1} \frac{n+1}{n} \\ &= \left(\frac{n^2 + 2n}{(n+1)^2}\right)^{n+1} \frac{n+1}{n} \\ &= \left(1 + \frac{-1}{(n+1)^2}\right)^{n+1} \frac{n+1}{n}, \end{aligned}$$

quindi, per la disuguaglianza di Bernoulli,

$$\frac{a_{n+1}}{a_n} \geq \left(1 + (n+1) \frac{-1}{(n+1)^2}\right) \frac{n+1}{n} = 1.$$

Analogamente, consideriamo la successione

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

Si ha che $a_n < b_n$, per ogni $n \geq 1$. Vediamo che $(b_n)_n$ è decrescente:

$$\begin{aligned} \frac{b_n}{b_{n+1}} &= \frac{\left(1 + \frac{1}{n}\right)^{n+1}}{\left(1 + \frac{1}{n+1}\right)^{n+2}} \\ &= \frac{n}{n+1} \left(\frac{n+1}{n}\right)^{n+2} \left(\frac{n+1}{n+2}\right)^{n+2} \\ &= \frac{n}{n+1} \left(\frac{(n+1)^2}{n^2 + 2n}\right)^{n+2} \\ &= \frac{n}{n+1} \left(1 + \frac{1}{n^2 + 2n}\right)^{n+2} \\ &\geq \frac{n}{n+1} \left(1 + (n+2) \frac{1}{n^2 + 2n}\right) \\ &= 1. \end{aligned}$$

Pertanto, le successioni $(a_n)_n$ e $(b_n)_n$ hanno entrambe limite finito. Essendo

$$\lim_n \frac{b_n}{a_n} = \lim_n \left(1 + \frac{1}{n}\right) = 1,$$

possiamo concludere che le due successioni hanno lo stesso limite, un numero reale. Esso si chiama “numero di Nepero” e si denota con e . Scriveremo

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

Si può dimostrare che è un numero irrazionale:

$$e = 2.7182818284590452353602874713526624977572470936999595 \dots$$

Limiti notevoli per l'esponenziale e il logaritmo

Dimostriamo ora che, al variare di x in $\mathbb{R}$,

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

Consideriamo, per $x \geq 0$, il numero naturale $n(x)$ tale che

$$n(x) \leq x < n(x) + 1$$

(detto “parte intera di x ”). Allora, per $x \geq 1$,

$$\begin{aligned} \left(1 + \frac{1}{n(x) + 1}\right)^{n(x)} &< \left(1 + \frac{1}{x}\right)^{n(x)} \leq \\ &\leq \left(1 + \frac{1}{x}\right)^x < \\ &< \left(1 + \frac{1}{x}\right)^{n(x)+1} \leq \left(1 + \frac{1}{n(x)}\right)^{n(x)+1}. \end{aligned}$$

Notiamo che $\lim_{x \rightarrow +\infty} n(x) = +\infty$, quindi

$$\begin{aligned} \lim_{x \rightarrow +\infty} \left(1 + \frac{1}{n(x)}\right)^{n(x)+1} &= \lim_n \left(1 + \frac{1}{n}\right)^{n+1} \\ &= \lim_n \left(1 + \frac{1}{n}\right)^n \left(1 + \frac{1}{n}\right) \\ &= e \cdot 1 = e. \end{aligned}$$

e analogamente

$$\begin{aligned}\lim_{x \rightarrow +\infty} \left(1 + \frac{1}{n(x) + 1}\right)^{n(x)} &= \lim_n \left(1 + \frac{1}{n}\right)^{n-1} \\ &= \lim_n \left(1 + \frac{1}{n}\right)^n \left(1 + \frac{1}{n}\right)^{-1} \\ &= e \cdot 1 = e\end{aligned}$$

Per il “teorema dei due carabinieri”, si ha che anche il limite cercato vale e .

Dimostriamo ora che si ha anche

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

Infatti, usando la formula $\lim_{x \rightarrow -\infty} f(x) = \lim_{x \rightarrow +\infty} f(-x)$, abbiamo che

$$\begin{aligned}\lim_{x \rightarrow -\infty} \left(1 + \frac{1}{x}\right)^x &= \lim_{x \rightarrow +\infty} \left(1 - \frac{1}{x}\right)^{-x} = \lim_{x \rightarrow +\infty} \left(1 + \frac{1}{x-1}\right)^x \\ &= \lim_{y \rightarrow +\infty} \left(1 + \frac{1}{y}\right)^{y+1} = \lim_{y \rightarrow +\infty} \left(1 + \frac{1}{y}\right)^y \left(1 + \frac{1}{y}\right) = e \cdot 1 = e.\end{aligned}$$

Possiamo ora enunciare il seguente importante

Teorema. *Si ha*

$$\lim_{x \rightarrow 0} \frac{\log_a(1+x)}{x} = \log_a(e), \quad \lim_{x \rightarrow 0} \frac{a^x - 1}{x} = \frac{1}{\log_a(e)}.$$

Dimostrazione. Abbiamo che

$$\lim_{x \rightarrow 0^+} \frac{\log_a(1+x)}{x} = \lim_{y \rightarrow +\infty} y \log_a \left(1 + \frac{1}{y}\right) = \lim_{y \rightarrow +\infty} \log_a \left(1 + \frac{1}{y}\right)^y = \log_a(e),$$

e lo stesso vale per il limite sinistro. Inoltre,

$$\lim_{x \rightarrow 0} \frac{a^x - 1}{x} = \lim_{y \rightarrow 0} \frac{y}{\log_a(1+y)} = \frac{1}{\log_a(e)}.$$

■

Si noti che la scelta della base $a = e$ semplifica le espressioni: si ha

$$\lim_{x \rightarrow 0} \frac{\log_e(1+x)}{x} = 1, \quad \lim_{x \rightarrow 0} \frac{e^x - 1}{x} = 1.$$

È per questo motivo che, da ora in poi, sceglieremo come base dell'esponenziale e del logaritmo il numero di Nepero e , che viene anche chiamato la “base naturale”. Scriveremo $\exp(x)$ (o anche $\exp x$) invece di $\exp_e(x)$ e $\ln(x)$ (o anche $\ln x$) invece di $\log_e(x)$. Potrebbero essere utili le formule seguenti:

$$a^x = e^{x \ln(a)}, \quad \log_a(x) = \frac{\ln(x)}{\ln(a)}.$$

Limiti notevoli per le funzioni trigonometriche

Definiamo la successione $(\ell_n)_n$ in questo modo:

$$\ell_1 = 2, \quad \ell_{n+1} = \sqrt{2 - \sqrt{4 - \ell_n^2}}.$$

(Geometricamente, si può vedere che ℓ_n corrisponde alla lunghezza del lato di un poligono regolare di 2^n lati inscritto ad una circonferenza di lato 1.) Si ha:

$$\begin{aligned} \ell_2 &= \sqrt{2} \\ \ell_3 &= \sqrt{2 - \sqrt{2}} \\ \ell_4 &= \sqrt{2 - \sqrt{2 + \sqrt{2}}} \\ \ell_5 &= \sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{2}}}} \\ &\dots \end{aligned}$$

Poniamo

$$a_n = 2^{n-1} \ell_n.$$

(Geometricamente, a_n corrisponde al semiperimetro di tale poligono.) In modo analogo, definiamo, per $n \geq 2$,

$$b_n = 2^n \frac{\ell_n}{\sqrt{4 - \ell_n^2}}.$$

(Geometricamente, si può vedere che b_n corrisponde al semiperimetro di un poligono regolare di 2^n lati circoscritto alla circonferenza di raggio 1.) Si ha che $a_n < b_n$ per

ogni $n \geq 2$. Ecco come si sviluppano le due successioni:

$$\begin{array}{ll}
 a_2 = 2\sqrt{2} & b_2 = 4 \\
 a_3 = 4\sqrt{2 - \sqrt{2}} & b_3 = 8 \frac{\sqrt{2 - \sqrt{2}}}{\sqrt{2 + \sqrt{2}}} \\
 a_4 = 8\sqrt{2 - \sqrt{2 + \sqrt{2}}} & b_4 = 16 \frac{\sqrt{2 - \sqrt{2 + \sqrt{2}}}}{\sqrt{2 + \sqrt{2 + \sqrt{2}}}} \\
 a_5 = 16\sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{2}}}} & b_5 = 32 \frac{\sqrt{2 - \sqrt{2 + \sqrt{2 + \sqrt{2}}}}}{\sqrt{2 + \sqrt{2 + \sqrt{2 + \sqrt{2}}}}} \\
 \dots & \dots
 \end{array}$$

Vediamo che la successione $(a_n)_n$ è strettamente crescente:

$$\frac{a_{n+1}}{a_n} = 2 \frac{\ell_{n+1}}{\ell_n} = 2 \frac{\sqrt{2 - \sqrt{4 - \ell_n^2}}}{\ell_n} = \frac{2}{\sqrt{2 + \sqrt{4 - \ell_n^2}}} > \frac{2}{\sqrt{2 + 2}} = 1.$$

Inoltre, la successione $(b_n)_n$ è strettamente decrescente:

$$\begin{aligned}
 \frac{b_n}{b_{n+1}} &= \frac{1}{2} \frac{\ell_n}{\sqrt{4 - \ell_n^2}} \frac{\sqrt{4 - \ell_{n+1}^2}}{\ell_{n+1}} \\
 &= \frac{1}{2} \frac{\ell_n}{\sqrt{4 - \ell_n^2}} \frac{\sqrt{2 + \sqrt{4 - \ell_n^2}}}{\sqrt{2 - \sqrt{4 - \ell_n^2}}} \\
 &= \frac{1}{2} \frac{2 + \sqrt{4 - \ell_n^2}}{\sqrt{4 - \ell_n^2}} \\
 &= \frac{1}{2} \left(\frac{2}{\sqrt{4 - \ell_n^2}} + 1 \right) \\
 &> \frac{1}{2}(1 + 1) = 1.
 \end{aligned}$$

Pertanto, le successioni $(a_n)_n$ e $(b_n)_n$ hanno entrambe limite finito. Essendo quindi

$$\lim_n \ell_n = \lim_n \frac{a_n}{2^{n-1}} = 0,$$

si ha

$$\lim_n \frac{b_n}{a_n} = \lim_n \frac{2}{\sqrt{4 - \ell_n^2}} = 1,$$

per cui possiamo concludere che le due successioni hanno lo stesso limite, un numero reale, che chiameremo “pi greco” e denoteremo con π . Si può dimostrare che è un numero irrazionale:

$$\pi = 3.1415926535897932384626433832795028841971693993751 \dots$$

Si può dimostrare il seguente

Teorema. *Si ha*

$$\lim_{x \rightarrow 0} \frac{\sin_T(x)}{x} = \frac{2\pi}{T}.$$

La sua dimostrazione risulta piuttosto complicata a questo livello e viene pertanto omessa. Si noti che la scelta della base $T = 2\pi$ semplifica le espressioni del limite. È per questo motivo che, da ora in poi, sceglieremo come base delle funzioni trigonometriche il numero 2π : scriveremo $\cos(x)$ (o anche $\cos x$) invece di $\cos_{2\pi}(x)$ e $\sin(x)$ (o anche $\sin x$) invece di $\sin_{2\pi}(x)$. Potranno essere utili le seguenti formule:

$$\cos_T(x) = \cos\left(\frac{2\pi}{T}x\right), \quad \sin_T(x) = \sin\left(\frac{2\pi}{T}x\right).$$

Si ha quindi

$$\lim_{x \rightarrow 0} \frac{\sin x}{x} = 1.$$

Spesso si forniscono delle dimostrazioni intuitive di questa formula, in cui si utilizza l’interpretazione geometrica della funzione “seno”. La “misura” dell’“angolo” x (in “radianti”) corrisponde alla “lunghezza” dell’arco di circonferenza unitaria che parte da $(1, 0)$ e arriva in $(\cos x, \sin x)$. Rimane a livello intuitivo il fatto che si possa “misurare” un arco di curva di questo tipo. Una definizione precisa di tale misura utilizza, ad esempio, la teoria dell’integrale.

Dimostriamo che vale la seguente uguaglianza:

$$\lim_{x \rightarrow 0} \frac{\cos x - 1}{x^2} = -\frac{1}{2}.$$

Infatti, ricordando che la funzione coseno è continua in 0, abbiamo

$$\begin{aligned} \lim_{x \rightarrow 0} \frac{\cos x - 1}{x^2} &= \lim_{x \rightarrow 0} \frac{\cos x - 1}{x^2} \frac{\cos x + 1}{\cos x + 1} \\ &= \lim_{x \rightarrow 0} \frac{-1}{\cos x + 1} \left(\frac{\sin x}{x}\right)^2 = \frac{-1}{\cos(0) + 1} \cdot 1^2 = -\frac{1}{2}, \end{aligned}$$

Definiamo ora la funzione “tangente di base T ”:

$$\tan_T(x) = \frac{\sin_T(x)}{\cos_T(x)}.$$

Il suo dominio naturale è l'insieme $\{x \in \mathbb{R} : x \neq \frac{T}{4} + k\frac{T}{2}, k \in \mathbb{Z}\}$. Essendo seno e coseno funzioni continue, anche la tangente lo è (sul suo dominio). Inoltre, essa è periodica: il suo periodo minimo è $\frac{T}{2}$. Come per le funzioni trigonometriche, se $T = 2\pi$ scriveremo semplicemente $\tan(x)$ (o anche $\tan x$) invece di $\tan_{2\pi}(x)$. Dimostriamo che

$$\lim_{x \rightarrow 0} \frac{\tan x}{x} = 1.$$

Infatti, si ha:

$$\lim_{x \rightarrow 0} \frac{\tan x}{x} = \lim_{x \rightarrow 0} \frac{1}{\cos x} \lim_{x \rightarrow 0} \frac{\sin x}{x} = \frac{1}{\cos(0)} \cdot 1 = 1.$$

Ancora sui limiti

Vediamo ora come si comportano l'esponenziale e il logaritmo a $+\infty$. Usando le proprietà di monotonia e suriettività di $\exp_a : \mathbb{R} \rightarrow]0, +\infty[$, vediamo che

$$\lim_{x \rightarrow +\infty} a^x = \begin{cases} +\infty & \text{se } a > 1, \\ 1 & \text{se } a = 1, \\ 0 & \text{se } a < 1, \end{cases}$$

mentre

$$\lim_{x \rightarrow +\infty} \log_a(x) = \begin{cases} +\infty & \text{se } a > 1, \\ -\infty & \text{se } a < 1. \end{cases}$$

Scrivendo $x^\alpha = \exp(\ln x^\alpha) = \exp(\alpha \ln x)$, si vede che

$$\lim_{x \rightarrow +\infty} x^\alpha = \begin{cases} +\infty & \text{se } \alpha > 0, \\ 1 & \text{se } \alpha = 0, \\ 0 & \text{se } \alpha < 0. \end{cases}$$

Teorema. Per ogni $\alpha > 0$, si ha

$$\lim_{x \rightarrow +\infty} \frac{e^x}{x^\alpha} = +\infty, \quad \lim_{x \rightarrow +\infty} \frac{\ln x}{x^\alpha} = 0.$$

Dimostrazione. Cominciamo con il dimostrare che, se $a > 1$,

$$\lim_n \frac{a^n}{n} = +\infty.$$

Infatti, scrivendo $a = 1 + b$, con $b > 0$, si ha che

$$a^n = (1 + b)^n = 1 + nb + \frac{n(n-1)}{2} b^2 + \dots + b^n > \frac{n(n-1)}{2} b^2.$$

Quindi,

$$\frac{a^n}{n} > \frac{n-1}{2} b^2,$$

da cui segue il risultato. Vediamo ora che, per ogni numero intero $k \geq 1$, si ha che

$$\lim_n \frac{a^n}{n^k} = +\infty.$$

Infatti, scrivendo

$$\frac{a^n}{n^k} = \left(\frac{a^{n/k}}{n} \right)^k = \left(\frac{(\sqrt[k]{a})^n}{n} \right)^k,$$

si può usare il fatto che $\lim_n \frac{(\sqrt[k]{a})^n}{n} = +\infty$ e concludere.

Siccome siamo interessati a calcolare un limite per $x \rightarrow +\infty$, supporremo ora $x \geq 1$. Siano $n(x)$ e $n(\alpha)$ i numeri naturali tali che

$$n(x) \leq x < n(x) + 1, \quad n(\alpha) \leq \alpha < n(\alpha) + 1.$$

Ponendo $k = n(\alpha) + 1$, per $x \geq 1$ si ha

$$\frac{e^x}{x^\alpha} \geq \frac{e^x}{x^{n(\alpha)+1}} = \frac{e^x}{x^k} \geq \frac{e^{n(x)}}{(n(x)+1)^k}.$$

Inoltre,

$$\lim_{x \rightarrow +\infty} \frac{e^{n(x)}}{(n(x)+1)^k} = \lim_n \frac{e^n}{(n+1)^k} = \frac{1}{e} \lim_n \frac{e^{n+1}}{(n+1)^k} = \frac{1}{e} \lim_m \frac{e^m}{m^k} = +\infty.$$

Ne segue la tesi per quanto riguarda il primo limite.

Passiamo ora al secondo limite. Con il cambio di variabile “ $y = \ln x$ ”, si ha

$$\lim_{x \rightarrow +\infty} \frac{\ln x}{x^\alpha} = \lim_{y \rightarrow +\infty} \frac{y}{(e^y)^\alpha} = \lim_{y \rightarrow +\infty} \left(\frac{y^{1/\alpha}}{e^y} \right)^\alpha = \lim_{y \rightarrow +\infty} \left(\frac{e^y}{y^{1/\alpha}} \right)^{-\alpha} = 0,$$

che è quanto volevasi dimostrare. ■

Concludiamo questa parte con due esempi.

1. Vogliamo far vedere che

$$\lim_n \frac{e^n}{n!} = 0.$$

A tal fine, dimostriamo per induzione che, per ogni $n \geq 3$, si ha

$$(P_n) \quad 0 < \frac{e^n}{n!} \leq \frac{e^3}{2n},$$

dopodichè il risultato segue dal “teorema dei due carabinieri”. Se $n = 3$, abbiamo un’uguaglianza, per cui la proposizione è vera. Supponiamo ora la proposizione (P_n) vera per un certo $n \geq 3$. Allora, essendo $e \in]2, 3[$,

$$\frac{e^{n+1}}{(n+1)!} = \frac{e}{n+1} \frac{e^n}{n!} \leq \frac{e}{n+1} \frac{e^3}{2n} = \frac{e}{n} \frac{e^3}{2(n+1)} \leq \frac{e^3}{2(n+1)},$$

per cui anche la proposizione (P_{n+1}) è vera.

2. Vediamo ora che

$$\lim_n \frac{n!}{n^n} = 0.$$

A tal scopo dimostriamo per induzione che, per ogni $n \geq 1$, si ha

$$(Q_n) \quad 0 < \frac{n!}{n^n} \leq \frac{1}{n},$$

e il risultato seguirà di nuovo usando il “teorema dei due carabinieri”. Se $n = 1$, la proposizione è sicuramente vera. Supponiamo ora (Q_n) vera per un certo $n \geq 1$. Allora

$$\begin{aligned} 0 < \frac{(n+1)!}{(n+1)^{n+1}} &= \frac{n!}{(n+1)^n} = \frac{n!}{n^n} \left(\frac{n}{n+1} \right)^n \\ &\leq \frac{1}{n} \left(\frac{n}{n+1} \right)^n = \frac{1}{n} \frac{n}{n+1} \left(\frac{n}{n+1} \right)^{n-1} = \frac{1}{n+1} \left(\frac{n}{n+1} \right)^{n-1} \leq \frac{1}{n+1}, \end{aligned}$$

per cui anche la proposizione (Q_{n+1}) è vera.

La derivata

Introdurremo ora il concetto di “derivata” di una funzione definita su un sottoinsieme di $\mathbb{R}$, a valori in $\mathbb{R}$.

Sia E un sottoinsieme di $\mathbb{R}$, dominio di una funzione $f : E \rightarrow \mathbb{R}$, e $x_0 \in E$ un punto di accumulazione per E . Se x è un punto di E diverso da x_0 , possiamo considerare il “rapporto incrementale”

$$\frac{f(x) - f(x_0)}{x - x_0};$$

si tratta del coefficiente angolare della retta nel piano passante per i punti $(x_0, f(x_0))$ e $(x, f(x))$.

Definizione. Qualora esso esista, chiameremo il limite

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

“derivata” di f nel punto x_0 , e lo denoteremo con uno dei seguenti simboli:

$$f'(x_0), \quad Df(x_0), \quad \frac{df}{dx}(x_0).$$

Si dice invece che f è “derivabile” in x_0 qualora la derivata sia un numero reale (e non $+\infty$ o $-\infty$). In tal caso, la retta nel piano passante per il punto $(x_0, f(x_0))$ con coefficiente angolare $f'(x_0)$, di equazione

$$y = f(x_0) + f'(x_0)(x - x_0),$$

si chiama “retta tangente” al grafico di f nel punto $(x_0, f(x_0))$.

Si noti che, in alcuni casi, la derivata di f in x_0 potrebbe essere solo un limite destro o un limite sinistro. Questo si verifica tipicamente quando E è un intervallo e x_0 coincide con uno degli estremi.

Osserviamo inoltre che si ha

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

Esempi. 1) Sia $f : \mathbb{R} \rightarrow \mathbb{R}$ definita da $f(x) = mx + q$. Allora

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{(mx + q) - (mx_0 + q)}{x - x_0} = m.$$

La retta tangente, in questo caso, coincide con il grafico della funzione. Il caso particolare in cui $m = 0$ ci mostra che la derivata di una funzione costante è sempre nulla.

2) Sia $f(x) = x^n$. Allora

$$f'(x_0) = \lim_{x \rightarrow x_0} \frac{x^n - x_0^n}{x - x_0} = \lim_{x \rightarrow x_0} \left(\sum_{k=0}^{n-1} x^k x_0^{n-1-k} \right) = nx_0^{n-1}.$$

Lo vediamo anche in un altro modo:

$$\begin{aligned} f'(x_0) &= \lim_{h \rightarrow 0} \frac{(x_0 + h)^n - x_0^n}{h} \\ &= \lim_{h \rightarrow 0} \frac{1}{h} \left(\sum_{k=0}^n \binom{n}{k} x_0^{n-k} h^k - x_0^n \right) \\ &= \lim_{h \rightarrow 0} \left(\sum_{k=1}^n \binom{n}{k} x_0^{n-k} h^{k-1} \right) \\ &= nx_0^{n-1}. \end{aligned}$$

3) Sia $f(x) = e^x$. Allora

$$f'(x_0) = \lim_{h \rightarrow 0} \frac{e^{x_0+h} - e^{x_0}}{h} = \lim_{h \rightarrow 0} e^{x_0} \frac{e^h - 1}{h} = e^{x_0}.$$

4) Sia $f(x) = \cos x$. Allora

$$\begin{aligned} f'(x_0) &= \lim_{h \rightarrow 0} \frac{\cos(x_0 + h) - \cos(x_0)}{h} \\ &= \lim_{h \rightarrow 0} \frac{\cos(x_0) \cos(h) - \sin(x_0) \sin(h) - \cos(x_0)}{h} \\ &= -\cos(x_0) \lim_{h \rightarrow 0} \frac{1 - \cos(h)}{h} - \sin(x_0) \lim_{h \rightarrow 0} \frac{\sin(h)}{h} \\ &= -\sin(x_0). \end{aligned}$$

5) Sia $g(x) = \sin x$. Allora

$$\begin{aligned} g'(x_0) &= \lim_{h \rightarrow 0} \frac{\sin(x_0 + h) - \sin(x_0)}{h} \\ &= \lim_{h \rightarrow 0} \frac{\sin(x_0) \cos(h) + \cos(x_0) \sin(h) - \sin(x_0)}{h} \\ &= -\sin(x_0) \lim_{h \rightarrow 0} \frac{1 - \cos(h)}{h} + \cos(x_0) \lim_{h \rightarrow 0} \frac{\sin(h)}{h} \\ &= \cos(x_0). \end{aligned}$$

Il seguente teorema ci fornisce una caratterizzazione della derivabilità.

Teorema. La funzione f è derivabile in x_0 se e solo se esiste un numero reale ℓ per cui si possa scrivere

$$f(x) = f(x_0) + \ell(x - x_0) + r(x),$$

dove r è una funzione tale che

$$\lim_{x \rightarrow x_0} \frac{r(x)}{x - x_0} = 0.$$

In tal caso, si ha $\ell = f'(x_0)$.

Dimostrazione. Supponiamo che f sia derivabile in x_0 . Allora

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

Quindi, ponendo $r(x) = f(x) - f(x_0) - f'(x_0)(x - x_0)$, essa verifica le proprietà richieste, con $\ell = f'(x_0)$.

Viceversa, supponiamo che $f(x) = f(x_0) + \ell(x - x_0) + r(x)$, con

$$\lim_{x \rightarrow x_0} \frac{r(x)}{x - x_0} = 0.$$

Allora

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0) - \ell(x - x_0)}{x - x_0} = 0,$$

da cui

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \ell.$$

■

Vediamo ora che la derivabilità implica la continuità.

Teorema. *Se f è derivabile in x_0 , allora f è continua in x_0 .*

Dimostrazione. Si ha

$$\begin{aligned} \lim_{x \rightarrow x_0} f(x) &= \lim_{x \rightarrow x_0} \left[f(x_0) + \frac{f(x) - f(x_0)}{x - x_0} (x - x_0) \right] \\ &= f(x_0) + f'(x_0) \cdot 0 = f(x_0), \end{aligned}$$

il che è equivalente a dire che f è continua in x_0 .

■

Alcune formule di derivazione

Vediamo ora alcune regole che si usano abitualmente.

Teorema. *Se $f, g : E \rightarrow \mathbb{R}$ sono derivabili in x_0 , anche $f + g$ lo è, e si ha*

$$(f + g)'(x_0) = f'(x_0) + g'(x_0).$$

Dimostrazione. Si ha:

$$\begin{aligned} \lim_{x \rightarrow x_0} \frac{(f + g)(x) - (f + g)(x_0)}{x - x_0} &= \lim_{x \rightarrow x_0} \left[\frac{f(x) - f(x_0)}{x - x_0} + \frac{g(x) - g(x_0)}{x - x_0} \right] \\ &= \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} + \lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0} \\ &= f'(x_0) + g'(x_0), \end{aligned}$$

per cui la formula è dimostrata.

■

Teorema. Se $f, g : E \rightarrow \mathbb{R}$ sono derivabili in x_0 , anche $f \cdot g$ lo è, e si ha

$$(f \cdot g)'(x_0) = f'(x_0)g(x_0) + f(x_0)g'(x_0).$$

Dimostrazione. Si ha

$$\begin{aligned} \lim_{x \rightarrow x_0} \frac{(f \cdot g)(x) - (f \cdot g)(x_0)}{x - x_0} &= \lim_{x \rightarrow x_0} \left[\frac{f(x) - f(x_0)}{x - x_0} g(x_0) + f(x) \frac{g(x) - g(x_0)}{x - x_0} \right] \\ &= \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} g(x_0) + \lim_{x \rightarrow x_0} f(x) \lim_{x \rightarrow x_0} \frac{g(x) - g(x_0)}{x - x_0}, \end{aligned}$$

e si conclude, ricordando che $\lim_{x \rightarrow x_0} f(x) = f(x_0)$, essendo f continua in x_0 . ■

Il caso particolare in cui g è costante con valore $\alpha \in \mathbb{R}$ ci fornisce la formula seguente:

$$(\alpha f)'(x_0) = \alpha f'(x_0).$$

Inoltre, scrivendo $f - g = f + (-1)g$, si ha:

$$(f - g)'(x_0) = f'(x_0) - g'(x_0).$$

Teorema. Se $f, g : E \rightarrow \mathbb{R}$ sono derivabili in x_0 e $g(x_0) \neq 0$, anche $\frac{f}{g}$ lo è, e si ha

$$\left(\frac{f}{g}\right)'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{[g(x_0)]^2}.$$

Dimostrazione. Si ha che $\frac{f}{g} = f \cdot \frac{1}{g}$, per cui dimostreremo dapprima che $\frac{1}{g}$ è derivabile in x_0 :

$$\lim_{x \rightarrow x_0} \frac{\frac{1}{g}(x) - \frac{1}{g}(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{g(x_0) - g(x)}{(x - x_0)g(x)g(x_0)} = -\frac{g'(x_0)}{[g(x_0)]^2}.$$

Quindi,

$$\left(\frac{f}{g}\right)'(x_0) = f'(x_0)\frac{1}{g}(x_0) + f(x_0)\left(\frac{1}{g}\right)'(x_0) = \frac{f'(x_0)}{g(x_0)} - f(x_0)\frac{g'(x_0)}{[g(x_0)]^2},$$

da cui la tesi. ■

Esempi. 1) Consideriamo la funzione “tangente”:

$$F(x) = \tan x = \frac{\sin x}{\cos x}.$$

Prendendo $f(x) = \sin x$ e $g(x) = \cos x$, si ha ⁵

$$F'(x_0) = \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{[g(x_0)]^2} = \frac{\cos^2(x_0) + \sin^2(x_0)}{\cos^2(x_0)} = \frac{1}{\cos^2(x_0)}.$$

2) Sono derivabili tutte le funzioni polinomiali

$$F(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_2 x^2 + a_1 x + a_0,$$

con derivata

$$F'(x_0) = n a_n x_0^{n-1} + (n-1) a_{n-1} x_0^{n-2} + \cdots + 2 a_2 x_0 + a_1.$$

Ne segue che sono derivabili anche tutte le funzioni razionali, del tipo

$$F(x) = \frac{p(x)}{q(x)},$$

dove $p(x)$ e $q(x)$ sono polinomi, con l'accortezza di scegliere un punto x_0 in cui $q(x_0) \neq 0$.

Vediamo ora come si calcola la derivata di una funzione composta.

Teorema. Se $f : E \rightarrow \mathbb{R}$ è derivabile in x_0 , e $g : E' \rightarrow \mathbb{R}$ è derivabile in $f(x_0)$, dove E' è un sottoinsieme di $\mathbb{R}$, contenente $f(E)$, per cui $f(x_0)$ è di accumulazione, allora $g \circ f$ è derivabile in x_0 , e si ha

$$(g \circ f)'(x_0) = g'(f(x_0)) f'(x_0).$$

Dimostrazione. Ponendo $y_0 = f(x_0)$, sia $R : E' \rightarrow \mathbb{R}$ la funzione ausiliaria così definita:

$$R(y) = \begin{cases} \frac{g(y) - g(y_0)}{y - y_0} & \text{se } y \neq y_0, \\ g'(y_0) & \text{se } y = y_0. \end{cases}$$

Osserviamo che la funzione R è continua in y_0 e

$$g(y) - g(y_0) = R(y)(y - y_0), \quad \text{per ogni } y \in E'.$$

Quindi, se $x \neq x_0$,

$$\frac{g(f(x)) - g(f(x_0))}{x - x_0} = R(f(x)) \frac{f(x) - f(x_0)}{x - x_0}.$$

⁵Qui e nel seguito scriveremo $\cos^2(x)$ e $\sin^2(x)$ per indicare $(\cos(x))^2$ e $(\sin(x))^2$, rispettivamente. Anche in questo caso, si può scrivere $\cos^2 x$ e $\sin^2 x$.

Siccome f è continua in x_0 e R è continua in $y_0 = f(x_0)$, la funzione composta $R \circ f$ è continua in x_0 e pertanto

$$\begin{aligned}\lim_{x \rightarrow x_0} \frac{g(f(x)) - g(f(x_0))}{x - x_0} &= \lim_{x \rightarrow x_0} R(f(x)) \lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} \\ &= R(f(x_0))f'(x_0) = g'(y_0)f'(x_0),\end{aligned}$$

che è quanto volevasi dimostrare. ■

Esempi. 1) Sia $h : \mathbb{R} \rightarrow \mathbb{R}$ definita da $h(x) = \cos(e^x)$. Si ha che $h = g \circ f$, con $f(x) = e^x$ e $g(y) = \cos y$. Fissato $x_0 \in \mathbb{R}$, si ha che $f'(x_0) = e^{x_0}$. Se $y_0 = f(x_0)$, abbiamo che $g'(y_0) = -\sin y_0$. Pertanto, la derivata di h in x_0 è

$$h'(x_0) = g'(f(x_0))f'(x_0) = -\sin(e^{x_0})e^{x_0}.$$

2) Sia ora $h : \mathbb{R} \rightarrow \mathbb{R}$ definita da $h(x) = e^{\cos x}$. Allora $h = g \circ f$, con $f(x) = \cos x$ e $g(y) = e^y$. Fissato $x_0 \in \mathbb{R}$, si ha che $f'(x_0) = -\sin x_0$. Se $y_0 = f(x_0)$, abbiamo che $g'(y_0) = e^{y_0}$. Pertanto, la derivata di h in x_0 è

$$h'(x_0) = g'(f(x_0))f'(x_0) = e^{\cos x_0}(-\sin x_0).$$

Vedremo ora come calcolare la derivata dell'inversa di una funzione invertibile. Nel seguito considereremo solo intervalli non degeneri, ossia non ridotti ad un solo punto.

Teorema. Siano I, J due intervalli e $f : I \rightarrow J$ una funzione continua invertibile. Se f è derivabile in x_0 e $f'(x_0) \neq 0$, allora f^{-1} è derivabile in $y_0 = f(x_0)$ e si ha che

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)}.$$

Dimostrazione. Osserviamo che, nelle ipotesi del teorema, la funzione $f : I \rightarrow J$ è sicuramente strettamente monotona, e la sua inversa $f^{-1} : J \rightarrow I$ è continua. Applicando il teorema del limite di una funzione composta, abbiamo

$$\lim_{y \rightarrow y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \lim_{\substack{x \rightarrow \lim_{y \rightarrow y_0} f^{-1}(y) \\ y \rightarrow y_0}} \frac{x - x_0}{f(x) - f(x_0)}.$$

Essendo f^{-1} continua, si ha che $\lim_{y \rightarrow y_0} f^{-1}(y) = f^{-1}(y_0) = x_0$, da cui la tesi. ■

Esempio. Se $f(x) = e^x$, si ha che $f^{-1}(y) = \ln y$, per cui, essendo $y_0 = e^{x_0}$,

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{e^{x_0}} = \frac{1}{y_0}.$$

Sia ora α un numero reale e $h :]0, +\infty[\rightarrow \mathbb{R}$ la funzione definita da $h(x) = x^\alpha$. Essendo

$$x^\alpha = e^{\alpha \ln x},$$

si ha che $h = g \circ f$, con $f(x) = \alpha \ln x$ e $g(y) = e^y$. Allora

$$h'(x_0) = g'(f(x_0))f'(x_0) = e^{\alpha \ln x_0} \alpha \frac{1}{x_0} = x_0^\alpha \alpha \frac{1}{x_0} = \alpha x_0^{\alpha-1}.$$

Quindi, la stessa formula trovata per un esponente n naturale continua a valere anche per un esponente α non intero.

La funzione derivata

Consideriamo una funzione $f : I \rightarrow \mathbb{R}$, dove $I \subseteq \mathbb{R}$ è un intervallo. Diremo che “ f è derivabile” se lo è in ogni punto di I . In tal caso, ad ogni $x \in I$ resta associato il numero reale $f'(x)$, per cui è ben definita una funzione $f' : I \rightarrow \mathbb{R}$, detta “funzione derivata”. Abbiamo la seguente tabella:

$f(x)$	$f'(x)$
x^α	$\alpha x^{\alpha-1}$
e^x	e^x
$\ln x$	$\frac{1}{x}$
$\cos x$	$-\sin x$
$\sin x$	$\cos x$
$\tan x$	$\frac{1}{\cos^2 x}$
$\dots$	$\dots$

Ci si può ora chiedere se la funzione derivata sia a sua volta derivabile in qualche punto di I . Se f' è derivabile in un punto x_0 , chiameremo la sua derivata $(f')'(x_0)$ “derivata seconda” di f in x_0 e la denoteremo con uno dei seguenti simboli:

$$f''(x_0), \quad D^2 f(x_0), \quad \frac{d^2 f}{dx^2}(x_0).$$

Si può procedere per induzione e definire, in generale, la derivata n -esima di f in x_0 , che denoteremo con uno dei seguenti simboli:

$$f^{(n)}(x_0), \quad D^n f(x_0), \quad \frac{d^n f}{dx^n}(x_0);$$

si ha $f^{(n)}(x_0) = (f^{(n-1)})'(x_0)$.

Se una funzione f possiede derivata n -esima in un punto x_0 per ogni $n \geq 1$, si dice che essa è “derivabile infinite volte” in x_0 . Ad esempio, la funzione esponenziale $f(x) = e^x$ lo è, in ogni punto $x_0 \in \mathbb{R}$. In questo caso, si ha

$$D^n e^x = e^x, \quad \text{per ogni } n \geq 1.$$

Successioni e sottosuccessioni

Utilizzeremo ora le successioni e i loro limiti per caratterizzare alcuni concetti introdotti in precedenza. A tal fine, riscriviamo la definizione di limite per una successione in uno spazio metrico E :

$$\lim_n a_n = \ell \quad \Leftrightarrow \quad \forall \varepsilon > 0 \quad \exists \bar{n} \in \mathbb{N} : \quad n \geq \bar{n} \Rightarrow d(a_n, \ell) < \varepsilon.$$

Sia ora U un sottoinsieme dello spazio metrico E . Possiamo caratterizzare la nozione di punto aderente a U facendo uso delle successioni.

Teorema. *Un punto $x \in E$ è aderente a U se e solo se esiste una successione $(a_n)_n$ in U tale che $\lim_n a_n = x$.*

Dimostrazione. Se x è aderente a U , allora, per ogni $n \in \mathbb{N}$, l'intersezione $B\left(x, \frac{1}{n+1}\right) \cap U$ è non vuota, per cui posso sceglierne un elemento, che chiamo a_n . In questo modo, ho costruito una successione $(a_n)_n$ in U , ed è facile vedere che essa ha limite x . Una delle due implicazioni è così dimostrata.

Supponiamo ora che esista una successione $(a_n)_n$ in U tale che $\lim_n a_n = x$. Allora, fissato $\rho > 0$, esiste un $\bar{n} \in \mathbb{N}$ tale che

$$n \geq \bar{n} \Rightarrow d(a_n, x) < \rho,$$

ossia $a_n \in B(x, \rho)$. Quindi, $B(x, \rho) \cap U$ è non vuoto, e questo dimostra che x è aderente a U . ■

Consideriamo ora due spazi metrici E, F e una funzione $f : E \rightarrow F$. Vogliamo caratterizzare la continuità di f in un punto $x_0 \in E$, facendo uso delle successioni.

Teorema. *La funzione f è continua in x_0 se e solo se, presa una successione $(a_n)_n$ in E , si ha*

$$\lim_n a_n = x_0 \quad \Rightarrow \quad \lim_n f(a_n) = f(x_0).$$

Dimostrazione. Supponiamo che f sia continua in x_0 , e sia $(a_n)_n$ una successione in E tale che $\lim_n a_n = x_0$. Per il Teorema 1 sul limite di una funzione composta,

$$\lim_n f(a_n) = f(\lim_n a_n) = f(x_0),$$

cosicchè una delle due implicazioni è dimostrata.

Ragioniamo ora per contrapposizione, e supponiamo che f non sia continua in x_0 . Questo significa che esiste un $\varepsilon > 0$ tale che, per ogni $\delta > 0$, esiste almeno un $x \in E$ per cui $d(x, x_0) < \delta$ e $d(f(x), f(x_0)) \geq \varepsilon$. Prendendo $\delta = \frac{1}{n+1}$, per ogni $n \in \mathbb{N}$ esiste pertanto un a_n in E tale che $d(a_n, x_0) < \frac{1}{n+1}$ e $d(f(a_n), f(x_0)) \geq \varepsilon$. Ne segue che $\lim_n a_n = x_0$, ma sicuramente non può essere che $\lim_n f(a_n) = f(x_0)$. ■

Data che sia una successione $(a_n)_n$, una sua “sottosuccessione” si ottiene selezionando una successione strettamente crescente di indici $(n_k)_k$ e considerando la funzione composta

$$k \mapsto n_k \mapsto a_{n_k}.$$

Teorema. *Se una successione ha limite, allora tutte le sue sottosuccessioni hanno lo stesso limite.*

Dimostrazione. Essendo gli indici n_k in $\mathbb{N}$, dalla $n_{k+1} > n_k$ si deduce che $n_{k+1} \geq n_k + 1$ e, per induzione, che $n_k \geq k$, per ogni k . Ne segue che $\lim_k n_k = +\infty$. Pertanto,

$$\lim_{k \rightarrow +\infty} a_{n_k} = \lim_{n \rightarrow \lim_k n_k} a_n = \lim_{n \rightarrow +\infty} a_n. \quad \blacksquare$$

Teorema. *Ogni successione di numeri reali ha una sottosuccessione monotona.*

Dimostrazione. Sia $(a_n)_n$ una successione in $\mathbb{R}$. Diremo che $\bar{n}$ è un “punto di vedetta” per la successione se $a_{\bar{n}} > a_n$ per ogni $n > n$. Distinguiamo ora tre casi.

Caso 1: ci sono infiniti punti di vedetta. Allora li ordiniamo in una successione strettamente crescente di indici $(n_k)_k$. In tal modo, la sottosuccessione $(a_{n_k})_k$ è strettamente decrescente.

Caso 2: ci sono solamente un numero finito di punti di vedetta. Sia N il più grande tra essi, e scegliamo $n_0 > N$. Allora, siccome n_0 non è un punto di vedetta, esiste un $n_1 > n_0$ tale che $a_{n_1} \geq a_{n_0}$. Per induzione, costruiamo una successione strettamente crescente di indici $(n_k)_k$ in questo modo: una volta definito n_k , siccome esso non è un punto di vedetta, esiste un $n_{k+1} > n_k$ tale che $a_{n_{k+1}} \geq a_{n_k}$. La sottosuccessione $(a_{n_k})_k$ così costruita è crescente.

Caso 3: non ci sono punti di vedetta. In questo caso, si può scegliere n_0 arbitrariamente e procedere come nel Caso 2. ■

La nozione di compattezza

Enunciamo ora la seguente proprietà degli intervalli chiusi e limitati di $\mathbb{R}$, ossia il “Teorema di Bolzano–Weierstrass”.

Teorema. *Ogni successione $(a_n)_n$ in $[a, b]$ possiede una sottosuccessione $(a_{n_k})_k$ che ha limite in $[a, b]$.*

Dimostrazione. Abbiamo dimostrato che ogni successione $(a_n)_n$ in $\mathbb{R}$ ha una sottosuccessione $(a_{n_k})_k$ monotona, che pertanto possiede un limite $\lim_k a_{n_k} = l$. Siccome $a \leq a_{n_k} \leq b$ per ogni k , deve essere $a \leq l \leq b$. ■

In uno spazio metrico E , diremo che un sottoinsieme U è “compatto” se ogni successione $(a_n)_n$ in U possiede una sottosuccessione $(a_{n_k})_k$ che ha limite in U . Il teorema di Bolzano–Weierstrass afferma quindi che, se $E = \mathbb{R}$, gli intervalli del tipo $U = [a, b]$ sono compatti.

Nel seguito, diremo che una funzione $f : U \rightarrow \mathbb{R}$ è “limitata superiormente” (o “limitata inferiormente”) se lo è la sua immagine $f(U)$. Diremo che f è “limitata” se è sia limitata superiormente che inferiormente. Diremo che “ f ha massimo” (o “ f ha minimo”) se $f(U)$ ce l’ha. Nel caso in cui f abbia massimo, chiameremo “punto di massimo” ogni $\bar{x}$ per cui $f(\bar{x}) = \max f(U)$; analoga definizione per “punto di minimo”.

Teorema (di Weierstrass). *Se U è un insieme compatto e $f : U \rightarrow \mathbb{R}$ è una funzione continua, allora f ha massimo e minimo.*

Dimostrazione. Sia $s = \sup f(U)$. Dimostreremo che esiste un punto di massimo, ossia un $\bar{x} \in U$ tale che $f(\bar{x}) = s$.

Notiamo che è possibile trovare una successione $(y_n)_n$ in $f(U)$ tale che $\lim_n y_n = s$: se $s \in \mathbb{R}$, per ogni $n \geq 1$ possiamo trovare un $y_n \in f(U)$ per cui $s - \frac{1}{n} < y_n \leq s$; se invece $s = +\infty$, per ogni n esiste un $y_n \in f(U)$ tale che $y_n > n$.

In corrispondenza, possiamo trovare una successione $(x_n)_n$ in U tale che $f(x_n) = y_n$. Essendo U compatto, esiste una sottosuccessione $(x_{n_k})_k$ che ha un limite $\bar{x} \in U$. Siccome $\lim_n y_n = s$ e $y_{n_k} = f(x_{n_k})$, la sottosuccessione $(y_{n_k})_k$ ha anch’essa limite s . Allora, per la continuità di f ,

$$f(\bar{x}) = f(\lim_k x_{n_k}) = \lim_k f(x_{n_k}) = \lim_k y_{n_k} = s.$$

Il teorema è così dimostrato, per quanto riguarda l’esistenza del massimo. Per il minimo, si procede in modo analogo (oppure, si considera la funzione continua $g = -f$ e si usa il fatto che g ha massimo). ■

Ricordo ora che una funzione $f : E \rightarrow F$ si dice “continua” se è continua in ogni punto $x_0 \in E$. In altri termini, se

$$\forall x_0 \in E \quad \forall \varepsilon > 0 \quad \exists \delta > 0 : \forall x \in E \quad d(x, x_0) < \delta \Rightarrow d(f(x), f(x_0)) < \varepsilon.$$

Si noti che, in generale, la scelta di δ dipende sia da ε che da x_0 . Nel caso in cui tale δ non dipenda da x_0 , diremo che la funzione è “uniformemente continua”: In tal caso, si ha che

$$\forall \varepsilon > 0 \quad \exists \delta > 0 : \forall x_0 \in E \quad \forall x \in E \quad d(x, x_0) < \delta \Rightarrow d(f(x), f(x_0)) < \varepsilon.$$

Proprietà notevoli della funzione derivata

Diremo che $x_0 \in I$ è un “punto di massimo locale” per la funzione $f : E \rightarrow \mathbb{R}$ se esiste un intorno U di x_0 per cui

$$x \in U \Rightarrow f(x) \leq f(x_0).$$

Se invece vale la disuguaglianza opposta, avremo un “punto di minimo locale”.

Calcoliamo ora la derivata nei punti di massimo o di minimo locale, che siano *interni* ad I .

Teorema (di Fermat). *Siano I un intervallo, x_0 un punto interno ad I , e $f : I \rightarrow \mathbb{R}$ derivabile in x_0 . Se x_0 è un punto di massimo o di minimo locale per f , allora $f'(x_0) = 0$.*

Dimostrazione. Se x_0 è punto di massimo locale, per x in un intorno di x_0 contenuto in I si ha che

$$\frac{f(x) - f(x_0)}{x - x_0} \begin{cases} \geq 0 & \text{se } x < x_0, \\ \leq 0 & \text{se } x > x_0. \end{cases}$$

Siccome f è derivabile in x_0 , abbiamo che esiste il limite del rapporto incrementale e coincide con i limiti destro e sinistro:

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

Da quanto sopra, per il corollario al teorema della permanenza del segno,

$$\lim_{x \rightarrow x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \leq 0 \leq \lim_{x \rightarrow x_0^-} \frac{f(x) - f(x_0)}{x - x_0},$$

e quindi deve essere $f'(x_0) = 0$. Nel caso in cui x_0 sia un punto di minimo locale, si procede in modo analogo. ■

Normalmente la derivata, essendo un limite, ci dà un’informazione di tipo locale sul comportamento della funzione. Il seguente teorema, invece, con la generalizzazione che ne seguirà, ci porterà all’uso della derivata per avere informazioni generali sull’andamento del grafico di una funzione.

Teorema (di Rolle). *Se $f : [a, b] \rightarrow \mathbb{R}$ è una funzione continua, derivabile su $]a, b[$ e tale che*

$$f(a) = f(b),$$

allora esiste un punto $\xi \in]a, b[$ tale che $f'(\xi) = 0$.

Dimostrazione. Se la funzione è costante, allora la sua derivata si annulla in tutti i punti, e la conclusione è banalmente vera. Supponiamo ora che f non sia costante. Esiste quindi un $\bar{x} \in]a, b[$ tale che

$$f(\bar{x}) < f(a) = f(b), \quad \text{oppure} \quad f(\bar{x}) > f(a) = f(b).$$

Supponiamo valga il primo caso. Per il teorema di Weierstrass, f ha minimo in $[a, b]$, e nel caso considerato un punto di minimo deve necessariamente essere in $]a, b[$. Sia $\xi \in]a, b[$ un tale punto. Per il teorema di Fermat, avremo che $f'(\xi) = 0$.

La situazione è analoga nel secondo caso. Per il teorema di Weierstrass, f ha massimo in $[a, b]$, e in questo caso un punto di massimo deve necessariamente essere in $]a, b[$. Se $\xi \in]a, b[$ è un tale punto, per il teorema di Fermat avremo che $f'(\xi) = 0$. ■

Enunciamo ora una generalizzazione del teorema di Rolle.

Teorema (di Lagrange). *Se $f : [a, b] \rightarrow \mathbb{R}$ è una funzione continua, derivabile su $]a, b[$, allora esiste un punto $\xi \in]a, b[$ tale che*

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}.$$

Dimostrazione. Definiamo la funzione

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a) - f(a).$$

Si ha che $g : [a, b] \rightarrow \mathbb{R}$ è una funzione continua, derivabile su $]a, b[$ e tale che

$$g(a) = 0 = g(b).$$

Per il teorema di Rolle, esiste un punto $\xi \in]a, b[$ tale che

$$g'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{b - a} = 0,$$

da cui la tesi. ■

Corollario. *Sia I un intervallo e $f : I \rightarrow \mathbb{R}$ una funzione continua, derivabile su $\overset{\circ}{I}$. Si ha che:*

- a) se $f'(x) \geq 0$ per ogni $x \in \overset{\circ}{I}$, allora f è crescente;
- b) se $f'(x) > 0$ per ogni $x \in \overset{\circ}{I}$, allora f è strettamente crescente;
- c) se $f'(x) \leq 0$ per ogni $x \in \overset{\circ}{I}$, allora f è decrescente;
- d) se $f'(x) < 0$ per ogni $x \in \overset{\circ}{I}$, allora f è strettamente decrescente;
- e) se $f'(x) = 0$ per ogni $x \in \overset{\circ}{I}$, allora f è costante.

Dimostrazione. Dimostriamo a): siano $x_1 < x_2$ in I . Per il teorema di Lagrange, esiste un $\xi \in]x_1, x_2[$ tale che

$$f'(\xi) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Quindi, essendo $f'(\xi) \geq 0$, si deve avere che $f(x_1) \leq f(x_2)$. Questo dimostra che f è crescente.

Le altre si dimostrano in modo analogo. ■

Si noti che, se f è crescente, allora ogni rapporto incrementale di f è sempre maggiore o uguale a zero e quindi $f'(x) \geq 0$ per ogni $x \in \overset{\circ}{I}$. Quindi in a), e così anche in c) ed e), vale anche l'implicazione opposta. Ma così non è per b) e d): se f è strettamente crescente, in generale non è vero che $f'(x) > 0$ per ogni $x \in \overset{\circ}{I}$: la derivata potrebbe annullarsi in qualche punto (vedi ad esempio $f(x) = x^3$).

Funzioni trigonometriche inverse

Tenuto conto della formula per la derivata e delle proprietà di segno delle funzioni trigonometriche, abbiamo che

$$\begin{aligned} \cos x \text{ è } & \begin{cases} \text{strettamente decrescente su } [0, \pi], \\ \text{strettamente crescente su } [\pi, 2\pi], \end{cases} \\ \sin x \text{ è } & \begin{cases} \text{strettamente crescente su } \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \\ \text{strettamente decrescente su } \left[\frac{\pi}{2}, \frac{3\pi}{2}\right]. \end{cases} \end{aligned}$$

Consideriamo le funzioni $F : [0, \pi] \rightarrow [-1, 1]$ e $G : [-\frac{\pi}{2}, \frac{\pi}{2}] \rightarrow [-1, 1]$ definite da $F(x) = \cos x$ e $G(x) = \sin x$. Sono strettamente monotone, quindi iniettive. Inoltre, essendo continue, la loro immagine è un intervallo e, siccome $F(\pi) = -1 = G(-\frac{\pi}{2})$ e $F(0) = 1 = G(\frac{\pi}{2})$, deve coincidere con $[-1, 1]$. Esse sono pertanto biettive. Chiameremo le due funzioni $F^{-1} : [-1, 1] \rightarrow [0, \pi]$ e $G^{-1} : [-1, 1] \rightarrow [-\frac{\pi}{2}, \frac{\pi}{2}]$ rispettivamente “arco coseno” e “arco seno” e scriveremo

$$F^{-1}(y) = \arccos y, \quad G^{-1}(y) = \arcsin y.$$

La prima è strettamente decrescente, la seconda strettamente crescente. Calcoliamone le derivate: ponendo $y = F(x)$, per $x \in]0, \pi[$ si ha

$$(F^{-1})'(y) = \frac{1}{F'(x)} = -\frac{1}{\sin x} = -\frac{1}{\sqrt{1 - \cos^2 x}} = -\frac{1}{\sqrt{1 - y^2}},$$

mentre ponendo $y = G(x)$, per $x \in] - \frac{\pi}{2}, \frac{\pi}{2}[$ si ha

$$(G^{-1})'(y) = \frac{1}{G'(x)} = \frac{1}{\cos x} = \frac{1}{\sqrt{1 - \sin^2 x}} = \frac{1}{\sqrt{1 - y^2}}.$$

Si può notare che la funzione $\arccos + \arcsin$ ha derivata nulla e pertanto è costante. Calcolandola in 0, si trova quindi che

$$\arccos y + \arcsin y = \frac{\pi}{2}, \quad \text{per ogni } y \in [-1, 1].$$

Consideriamo ora la funzione $H :] - \frac{\pi}{2}, \frac{\pi}{2}[\rightarrow \mathbb{R}$ definita da $H(x) = \tan x$. Per lo stesso tipo di considerazioni, essa risulta invertibile. Chiameremo la funzione $H^{-1} : \mathbb{R} \rightarrow] - \frac{\pi}{2}, \frac{\pi}{2}[$ “arco tangente” e scriveremo

$$H^{-1}(y) = \arctan y.$$

Essa è strettamente crescente e si ha:

$$\lim_{y \rightarrow -\infty} \arctan y = -\frac{\pi}{2}, \quad \lim_{y \rightarrow +\infty} \arctan y = \frac{\pi}{2}.$$

Calcoliamone la derivata: ponendo $y = H(x)$, per $x \in] - \frac{\pi}{2}, \frac{\pi}{2}[$ si ha

$$(H^{-1})'(y) = \frac{1}{H'(x)} = \cos^2 x = \frac{1}{1 + \tan^2 x} = \frac{1}{1 + y^2}.$$

Funzioni iperboliche e loro inverse

Sono interessanti le “funzioni iperboliche”:

$$\cosh(x) = \frac{e^x + e^{-x}}{2}, \quad \sinh(x) = \frac{e^x - e^{-x}}{2},$$

Esse soddisfano le seguenti proprietà, di facile verifica:

- (a) $(\cosh(x))^2 - (\sinh(x))^2 = 1$,
- (b) $\cosh(x_1 + x_2) = \cosh(x_1) \cosh(x_2) + \sinh(x_1) \sinh(x_2)$,
- (c) $\sinh(x_1 + x_2) = \sinh(x_1) \cosh(x_2) + \cosh(x_1) \sinh(x_2)$.

Si definisce inoltre la funzione “tangente iperbolica”:

$$\tanh(x) = \frac{\sinh(x)}{\cosh(x)}.$$

Essa è definita su tutto $\mathbb{R}$.

Si possono agevolmente calcolare le derivate

$$D \cosh(x) = \sinh(x), \quad D \sinh(x) = \cosh(x),$$

e inoltre

$$D \tanh(x) = \frac{1}{\cosh^2(x)}.$$

La funzione $\sinh : \mathbb{R} \rightarrow \mathbb{R}$ è strettamente crescente e invertibile. Si vede infatti che

$$\sinh^{-1}(y) = \ln(y + \sqrt{y^2 + 1}).$$

La derivata si può calcolare direttamente, oppure usando la formula della funzione inversa: se $y = \sinh(x)$, si ha

$$D \sinh^{-1}(y) = \frac{1}{D \sinh(x)} = \frac{1}{\cosh(x)} = \frac{1}{\sqrt{1 + \sinh^2(x)}} = \frac{1}{\sqrt{1 + y^2}}.$$

La funzione $\cosh : \mathbb{R} \rightarrow \mathbb{R}$ non è né iniettiva (è una funzione pari) né suriettiva: si ha $\cosh x \geq 1$ per ogni $x \in \mathbb{R}$. D'altra parte, la funzione $F : [0, +\infty[\rightarrow [1, +\infty[$, definita da $F(x) = \cosh x$, è strettamente crescente, invertibile e la sua inversa $F^{-1} : [1, +\infty[\rightarrow [0, +\infty[$ è data da

$$F^{-1}(y) = \ln(y + \sqrt{y^2 - 1}).$$

Essa si denota spesso, impropriamente, con $\cosh^{-1}$. Calcoliamone la derivata: ponendo $y = \cosh(x)$, con $x \geq 0$, si ha

$$D \cosh^{-1}(y) = \frac{1}{D \cosh(x)} = \frac{1}{\sinh(x)} = \frac{1}{\sqrt{\cosh^2(x) - 1}} = \frac{1}{\sqrt{y^2 - 1}}.$$

La funzione $\tanh : \mathbb{R} \rightarrow \mathbb{R}$ non è suriettiva: si ha $-1 < \tanh x < 1$ per ogni $x \in \mathbb{R}$. D'altra parte, la funzione $H : \mathbb{R} \rightarrow]-1, 1[$, definita da $H(x) = \tanh x$, è strettamente crescente, invertibile e la sua inversa $H^{-1} :]-1, 1[\rightarrow \mathbb{R}$ è data da

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

Essa si denota spesso, impropriamente, con $\tanh^{-1}$. Ne calcoliamo la derivata: ponendo $y = \tanh(x)$, si ha

$$D \tanh^{-1}(y) = \frac{1}{D \tanh(x)} = \cosh^2(x) = \frac{1}{1 - \tanh^2(x)} = \frac{1}{1 - y^2}.$$

Riassumiamo nella tabella sottostante le derivate delle funzioni elementari fin qui trovate.

$f(x)$	$f'(x)$	$f(x)$	$f'(x)$
x^α	$\alpha x^{\alpha-1}$	$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$
e^x	e^x	$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$
$\ln x$	$\frac{1}{x}$	$\arctan x$	$\frac{1}{1+x^2}$
$\cos x$	$-\sin x$	$\cosh^{-1} x$	$\frac{1}{\sqrt{x^2-1}}$
$\sin x$	$\cos x$	$\sinh^{-1} x$	$\frac{1}{\sqrt{x^2+1}}$
$\tan x$	$\frac{1}{\cos^2 x}$	$\tanh^{-1} x$	$\frac{1}{1-x^2}$
$\cosh x$	$\sinh x$	$\dots$	$\dots$
$\sinh x$	$\cosh x$		
$\tanh x$	$\frac{1}{\cosh^2 x}$		

Convessità e concavità

Sia $I \subseteq \mathbb{R}$ un intervallo.

Definizione. Diremo che una funzione $f : I \rightarrow \mathbb{R}$ è “convessa” se, comunque presi tre punti $x_1 < x_2 < x_3$ in I , si ha che

$$(a) \quad \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

Vediamo che sono equivalenti ad (a) le seguenti:

$$(b) \quad \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1},$$

$$(c) \quad \frac{f(x_3) - f(x_1)}{x_3 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

Infatti,

$$\begin{aligned}
& \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_2)}{x_3 - x_2} \Leftrightarrow \\
& \Leftrightarrow (f(x_2) - f(x_1))(x_3 - x_2) \leq (f(x_3) - f(x_2))(x_2 - x_1) \\
& \Leftrightarrow (f(x_2) - f(x_1))(x_3 - x_1 + x_1 - x_2) \leq (f(x_3) - f(x_1) + f(x_1) - f(x_2))(x_2 - x_1) \\
& \Leftrightarrow (f(x_2) - f(x_1))(x_3 - x_1) \leq (f(x_3) - f(x_1))(x_2 - x_1) \\
& \Leftrightarrow \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leq \frac{f(x_3) - f(x_1)}{x_3 - x_1},
\end{aligned}$$

per cui $(a) \Leftrightarrow (b)$; analogamente si vede che $(a) \Leftrightarrow (c)$.

Osserviamo che $f : I \rightarrow \mathbb{R}$ è convessa se e solo se, per ogni x_0 in I , la funzione “rapporto incrementale” $F : I \setminus \{x_0\} \rightarrow \mathbb{R}$, definita da

$$F(x) = \frac{f(x) - f(x_0)}{x - x_0},$$

è crescente. Infatti, presi x, x' in $I \setminus \{x_0\}$ tali che $x < x'$, si ha $F(x) \leq F(x')$, e questo accade in tutti e tre i casi possibili: $x < x' < x_0$, oppure $x < x_0 < x'$, oppure $x_0 < x < x'$. A questo punto, diventa naturale la seguente caratterizzazione della convessità.

Teorema. Se $f : I \rightarrow \mathbb{R}$ è continua, derivabile su $\overset{\circ}{I}$, allora f è convessa se e solo se f' è crescente su $\overset{\circ}{I}$.

Dimostrazione. Supponiamo che f sia convessa. Siano $\alpha < \beta$ due punti in $\overset{\circ}{I}$. Se $\alpha < x < \beta$, per (b) si ha

$$\frac{f(x) - f(\alpha)}{x - \alpha} \leq \frac{f(\beta) - f(\alpha)}{\beta - \alpha},$$

da cui, essendo f derivabile in α ,

$$f'(\alpha) = \lim_{x \rightarrow \alpha^+} \frac{f(x) - f(\alpha)}{x - \alpha} \leq \frac{f(\beta) - f(\alpha)}{\beta - \alpha}.$$

Analogamente, per (c) si ha

$$\frac{f(\beta) - f(\alpha)}{\beta - \alpha} \leq \frac{f(\beta) - f(x)}{\beta - x}.$$

da cui, essendo f derivabile in β ,

$$f'(\beta) = \lim_{x \rightarrow \beta^-} \frac{f(\beta) - f(x)}{\beta - x} \geq \frac{f(\beta) - f(\alpha)}{\beta - \alpha}.$$

Ne segue che $f'(\alpha) \leq f'(\beta)$, il che dimostra che f' è crescente.

Viceversa, supponiamo f' crescente. Presi $x_1 < x_2 < x_3$, per il teorema di Lagrange abbiamo che

$$\begin{aligned}\exists \xi_1 \in]x_1, x_2[: f'(\xi_1) &= \frac{f(x_2) - f(x_1)}{x_2 - x_1} \\ \exists \xi_2 \in]x_2, x_3[: f'(\xi_2) &= \frac{f(x_3) - f(x_2)}{x_3 - x_2}.\end{aligned}$$

Essendo f' crescente, si ha che $f'(\xi_1) \leq f'(\xi_2)$; ne segue (a). ■

Diremo che f è “strettamente convessa” se, comunque presi tre punti $x_1 < x_2 < x_3$ in I , si ha

$$(a') \quad \frac{f(x_2) - f(x_1)}{x_2 - x_1} < \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

Equivalentemente, possiamo scrivere le analoghe

$$(b') \quad \frac{f(x_2) - f(x_1)}{x_2 - x_1} < \frac{f(x_3) - f(x_1)}{x_3 - x_1},$$

$$(c') \quad \frac{f(x_3) - f(x_1)}{x_3 - x_1} < \frac{f(x_3) - f(x_2)}{x_3 - x_2}.$$

Vale la seguente caratterizzazione.

Teorema. Se $f : I \rightarrow \mathbb{R}$ è continua, derivabile su $\overset{\circ}{I}$, allora f è strettamente convessa se e solo se f' è strettamente crescente su $\overset{\circ}{I}$.

Dimostrazione. Dovremo modificare un pochino la dimostrazione del teorema precedente. Supponiamo che f sia strettamente convessa e siano $\alpha < \beta$ due punti in $\overset{\circ}{I}$. Se $\alpha < x < \frac{1}{2}(\alpha + \beta)$, per (b') si ha

$$\frac{f(x) - f(\alpha)}{x - \alpha} < \frac{f(\frac{\alpha+\beta}{2}) - f(\alpha)}{\frac{\alpha+\beta}{2} - \alpha} < \frac{f(\beta) - f(\alpha)}{\beta - \alpha},$$

da cui

$$f'(\alpha) = \lim_{x \rightarrow \alpha^+} \frac{f(x) - f(\alpha)}{x - \alpha} \leq \frac{f(\frac{\alpha+\beta}{2}) - f(\alpha)}{\frac{\alpha+\beta}{2} - \alpha} < \frac{f(\beta) - f(\alpha)}{\beta - \alpha}.$$

Analogamente, se $\frac{1}{2}(\alpha + \beta) < x < \beta$, per (c') si ha

$$\frac{f(\beta) - f(\alpha)}{\beta - \alpha} < \frac{f(\beta) - f(\frac{\alpha+\beta}{2})}{\beta - \frac{\alpha+\beta}{2}} < \frac{f(\beta) - f(x)}{\beta - x}.$$

da cui

$$f'(\beta) = \lim_{x \rightarrow \beta^-} \frac{f(\beta) - f(x)}{\beta - x} \geq \frac{f(\beta) - f(\frac{\alpha+\beta}{2})}{\beta - \frac{\alpha+\beta}{2}} > \frac{f(\beta) - f(\alpha)}{\beta - \alpha}.$$

Ne segue che $f'(\alpha) < f'(\beta)$, il che dimostra che f' è strettamente crescente.

Viceversa, supponiamo f' crescente. Presi $x_1 < x_2 < x_3$, usando il teorema di Lagrange, esattamente come per il teorema precedente si dimostra che vale (a') . ■

Diremo che f è “concava” se la funzione $(-f)$ è convessa o, equivalentemente, se vale (a) ma con il segno di disuguaglianza invertito. Diremo che f è “strettamente concava” se la funzione $(-f)$ è strettamente convessa o, equivalentemente, se vale (a') ma con il segno di disuguaglianza invertito. Si possono scrivere, naturalmente, gli analoghi teoremi che caratterizzano la concavità (o la stretta concavità) di f con la decrescenza (o la stretta decrescenza) di f' .

Arriviamo quindi al seguente corollario, che trova spesso applicazione in situazioni pratiche.

Corollario. Sia I un intervallo e $f : I \rightarrow \mathbb{R}$ una funzione continua, derivabile due volte su $\overset{\circ}{I}$. Si ha che:

- a) se $f''(x) \geq 0$ per ogni $x \in \overset{\circ}{I}$, allora f è convessa;
- b) se $f''(x) > 0$ per ogni $x \in \overset{\circ}{I}$, allora f è strettamente convessa;
- c) se $f''(x) \leq 0$ per ogni $x \in \overset{\circ}{I}$, allora f è concava;
- d) se $f''(x) < 0$ per ogni $x \in \overset{\circ}{I}$, allora f è strettamente concava.

Analogamente a quanto già osservato per le funzioni monotone, anche qui in a) e c) valgono anche le implicazioni opposte: se f è convessa, allora $f''(x) \geq 0$ per ogni $x \in \overset{\circ}{I}$, e similmente se f è concava. Ma così non è per b) e d) (vedi ad esempio $f(x) = x^4$).

Esempi. 1) La funzione esponenziale $f(x) = e^x$ è strettamente convessa: si ha

$$f''(x) = e^x > 0,$$

per ogni $x \in \mathbb{R}$. La sua inversa $\ln(x)$, il logaritmo naturale, è una funzione strettamente concava.

2) Tenuto conto delle derivate delle funzioni trigonometriche, si ha che:

$$\begin{aligned} \cos x \text{ è } & \begin{cases} \text{strettamente concava su } \left[-\frac{\pi}{2}, \frac{\pi}{2}\right], \\ \text{strettamente convessa su } \left[\frac{\pi}{2}, \frac{3\pi}{2}\right], \end{cases} \\ \sin x \text{ è } & \begin{cases} \text{strettamente concava su } [0, \pi], \\ \text{strettamente convessa su } [\pi, 2\pi]. \end{cases} \end{aligned}$$

I punti che separano un intervallo in cui si ha convessità da un altro in cui si ha concavità si chiamano “punti di flesso”.

Analoghe considerazioni si possono fare per le altre funzioni elementari fin qui studiate.

Le regole di de l'Hôpital

Iniziamo con l'introdurre la seguente generalizzazione del teorema di Lagrange.

Teorema (di Cauchy). *Se $f, g : [a, b] \rightarrow \mathbb{R}$ sono due funzioni continue, derivabili su $]a, b[$, con $g'(x) \neq 0$ per ogni $x \in]a, b[$, allora esiste un punto $\xi \in]a, b[$ tale che*

$$\frac{f'(\xi)}{g'(\xi)} = \frac{f(b) - f(a)}{g(b) - g(a)}.$$

Dimostrazione. Consideriamo la funzione $h : [a, b] \rightarrow \mathbb{R}$ definita da

$$h(x) = (g(b) - g(a))f(x) - (f(b) - f(a))g(x).$$

Si vede che essa è continua, derivabile su $]a, b[$, e $h(a) = h(b)$. Per il teorema di Rolle, esiste un punto $\xi \in]a, b[$ tale che $h'(\xi) = 0$. Ne segue la tesi. ■

Il seguente risultato è noto come “regola di de l'Hôpital” nel caso indeterminato del tipo $\frac{0}{0}$.

Teorema. *Sia I un intervallo e $f, g : I \setminus \{x_0\} \rightarrow \mathbb{R}$ due funzioni derivabili, con $g'(x) \neq 0$ per ogni $x \in I \setminus \{x_0\}$, tali che*

$$\lim_{x \rightarrow x_0} f(x) = \lim_{x \rightarrow x_0} g(x) = 0.$$

Se esiste il limite

$$\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)},$$

allora esiste anche il limite

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)},$$

e i due coincidono.

Dimostrazione. Sia $l = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$ (possibilmente $l = +\infty$ o $-\infty$); estendiamo le due funzioni anche al punto x_0 ponendo $f(x_0) = g(x_0) = 0$. In questo modo f e g saranno continue su tutto I . Per il teorema di Cauchy, per ogni $x \neq x_0$ esiste un punto $\xi_x \in]x_0, x[$ (che dipende da x) ⁶ tale che

$$\frac{f'(\xi_x)}{g'(\xi_x)} = \frac{f(x) - f(x_0)}{g(x) - g(x_0)} = \frac{f(x)}{g(x)}.$$

⁶Qui e nel seguito, nel caso in cui x sia minore di x_0 , con il simbolo $]x_0, x[$ si intende indicare l'intervallo $]x, x_0[$. Si noti che la formula del teorema di Cauchy rimane valida anche in questo caso.

Se $x \rightarrow x_0$, si ha che anche $\xi_x \rightarrow x_0$, per cui, usando il teorema sul limite di una funzione composta,

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow x_0} \frac{f'(\xi_x)}{g'(\xi_x)} = \lim_{y \rightarrow x_0} \frac{f'(y)}{g'(y)} = l.$$

■

Esempio. Siano $x_0 = 0$, $f(x) = \sin x - x$ e $g(x) = x^3$. Allora

$$\lim_{x \rightarrow 0} \frac{f'(x)}{g'(x)} = \lim_{x \rightarrow 0} \frac{\cos x - 1}{3x^2} = -\frac{1}{6},$$

per cui anche

$$\lim_{x \rightarrow 0} \frac{\sin x - x}{x^3} = -\frac{1}{6}.$$

Risulta talvolta utile il seguente

Corollario. Sia I un intervallo contenente x_0 e $f : I \rightarrow \mathbb{R}$ una funzione, continua in x_0 , e derivabile in ogni $x \neq x_0$. Se esiste il limite

$$l = \lim_{x \rightarrow x_0} f'(x),$$

allora esiste anche la derivata di f in x_0 e si ha $f'(x_0) = l$.

Dimostrazione. Siano $F(x) = f(x) - f(x_0)$ e $G(x) = x - x_0$. Abbiamo che

$$\lim_{x \rightarrow x_0} F(x) = \lim_{x \rightarrow x_0} G(x) = 0,$$

e

$$\lim_{x \rightarrow x_0} \frac{F'(x)}{G'(x)} = \lim_{x \rightarrow x_0} f'(x) = l.$$

La regola di de l'Hôpital ci dice quindi che

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{x - x_0} = \lim_{x \rightarrow x_0} \frac{F(x)}{G(x)} = l,$$

ossia $f'(x_0) = l$.

■

Il teorema precedente non esclude la possibilità che x_0 sia un estremo dell'intervallo I , nel qual caso si parlerà di limite destro o limite sinistro. Esso si estende anche ai casi in cui $x_0 = +\infty$ o $-\infty$. Vediamo qui il primo caso.

Teorema. Sia I un intervallo non limitato superiormente e $f, g : I \rightarrow \mathbb{R}$ due funzioni derivabili, con $g'(x) \neq 0$ per ogni $x \in I$, tali che

$$\lim_{x \rightarrow +\infty} f(x) = \lim_{x \rightarrow +\infty} g(x) = 0.$$

Se esiste il limite

$$\lim_{x \rightarrow +\infty} \frac{f'(x)}{g'(x)},$$

allora esiste anche il limite

$$\lim_{x \rightarrow +\infty} \frac{f(x)}{g(x)},$$

e i due coincidono.

Dimostrazione. Sia $l = \lim_{x \rightarrow +\infty} \frac{f'(x)}{g'(x)}$; definendo le due funzioni $F(x) = f(x^{-1})$ e $G(x) = g(x^{-1})$, si ha che $G'(x) \neq 0$ per ogni x e

$$\lim_{x \rightarrow 0^+} F(x) = \lim_{x \rightarrow 0^+} G(x) = 0.$$

Inoltre,

$$\lim_{x \rightarrow 0^+} \frac{F'(x)}{G'(x)} = \lim_{x \rightarrow 0^+} \frac{f'(x^{-1})(-x^{-2})}{g'(x^{-1})(-x^{-2})} = \lim_{x \rightarrow 0^+} \frac{f'(x^{-1})}{g'(x^{-1})} = \lim_{y \rightarrow +\infty} \frac{f'(y)}{g'(y)} = l.$$

Per il teorema precedente, si ha che anche $\lim_{x \rightarrow 0^+} \frac{F(x)}{G(x)} = l$; pertanto,

$$\lim_{x \rightarrow +\infty} \frac{f(x)}{g(x)} = \lim_{u \rightarrow 0^+} \frac{f(u^{-1})}{g(u^{-1})} = \lim_{u \rightarrow 0^+} \frac{F(u)}{G(u)} = l.$$

■

Osserviamo che la conclusione del teorema è scritta in forma di implicazione: se esiste il limite del rapporto delle due derivate, allora esiste il limite del rapporto delle due funzioni. L'implicazione opposta non è vera, come si vede dal seguente esempio: siano $x_0 = 0$,

$$f(x) = x^2 \sin\left(\frac{1}{x}\right), \quad g(x) = x.$$

Allora $\lim_{x \rightarrow 0} f(x) = \lim_{x \rightarrow 0} g(x) = 0$,

$$\lim_{x \rightarrow 0} \frac{f(x)}{g(x)} = \lim_{x \rightarrow 0} x \sin\left(\frac{1}{x}\right) = 0,$$

mentre

$$\frac{f'(x)}{g'(x)} = 2x \sin\left(\frac{1}{x}\right) - \cos\left(\frac{1}{x}\right),$$

per cui non esiste il $\lim_{x \rightarrow 0} \frac{f'(x)}{g'(x)}$.

Vediamo ora che la regola di de l'Hôpital continua a valere anche nei casi indeterminati del tipo $\frac{\infty}{\infty}$, dove ∞ può essere $+\infty$ o $-\infty$. Ad esempio, nel caso in cui x_0 sia un numero reale, si ha il seguente

Teorema. Sia I un intervallo contenente x_0 e $f, g : I \setminus \{x_0\} \rightarrow \mathbb{R}$ due funzioni derivabili, con $g'(x) \neq 0$ per ogni $x \in I \setminus \{x_0\}$, tali che

$$\lim_{x \rightarrow x_0} f(x) = \infty, \quad \lim_{x \rightarrow x_0} g(x) = \infty.$$

Se esiste il limite

$$\lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)},$$

allora esiste anche il limite

$$\lim_{x \rightarrow x_0} \frac{f(x)}{g(x)},$$

e i due coincidono.

Dimostrazione. Sia $l = \lim_{x \rightarrow x_0} \frac{f'(x)}{g'(x)}$. Supponiamo dapprima $l \in \mathbb{R}$, e che x_0 non sia l'estremo destro dell'intervallo I . Fissiamo $\varepsilon > 0$. Allora esiste un $\delta_1 > 0$ tale che

$$x_0 < x < x_0 + \delta_1 \Rightarrow \left| \frac{f'(x)}{g'(x)} - l \right| \leq \frac{\varepsilon}{2}.$$

Per il teorema di Cauchy, per ogni $x \in]x_0, x_0 + \delta_1[$, esiste un $\xi_x \in]x, x_0 + \delta_1[$ tale che

$$\frac{f'(\xi_x)}{g'(\xi_x)} = \frac{f(x_0 + \delta_1) - f(x)}{g(x_0 + \delta_1) - g(x)},$$

per cui

$$x_0 < x < x_0 + \delta_1 \Rightarrow \left| \frac{f(x_0 + \delta_1) - f(x)}{g(x_0 + \delta_1) - g(x)} - l \right| \leq \frac{\varepsilon}{2}.$$

Possiamo inoltre supporre che δ_1 sia tale che

$$x_0 < x < x_0 + \delta_1 \Rightarrow f(x) \neq 0 \text{ e } g(x) \neq 0.$$

Scriviamo

$$\frac{f(x_0 + \delta_1) - f(x)}{g(x_0 + \delta_1) - g(x)} = \psi(x) \frac{f(x)}{g(x)},$$

e osserviamo che

$$\lim_{x \rightarrow x_0} \psi(x) = \lim_{x \rightarrow x_0} \frac{1 - f(x_0 + \delta_1)/f(x)}{1 - g(x_0 + \delta_1)/g(x)} = 1.$$

In particolare,

$$\lim_{x \rightarrow x_0} \frac{1}{\psi(x)} \left(l - \frac{\varepsilon}{2} \right) = l - \frac{\varepsilon}{2}, \quad \lim_{x \rightarrow x_0} \frac{1}{\psi(x)} \left(l + \frac{\varepsilon}{2} \right) = l + \frac{\varepsilon}{2}.$$

Pertanto, esiste un $\delta \in]0, \delta_1[$ tale che, se $x_0 < x < x_0 + \delta$, allora

$$\frac{1}{\psi(x)} \left(l - \frac{\varepsilon}{2} \right) \geq l - \varepsilon, \quad \frac{1}{\psi(x)} \left(l + \frac{\varepsilon}{2} \right) \leq l + \varepsilon.$$

Quindi, se $x_0 < x < x_0 + \delta$, si ha

$$l - \varepsilon \leq \frac{1}{\psi(x)} \left(l - \frac{\varepsilon}{2} \right) \leq \frac{1}{\psi(x)} \frac{f(x_0 + \delta_1) - f(x)}{g(x_0 + \delta_1) - g(x)} \leq \frac{1}{\psi(x)} \left(l + \frac{\varepsilon}{2} \right) \leq l + \varepsilon,$$

da cui

$$\left| \frac{f(x)}{g(x)} - l \right| \leq \varepsilon.$$

Abbiamo così dimostrato che

$$\lim_{x \rightarrow x_0^+} \frac{f(x)}{g(x)} = l.$$

In modo del tutto analogo si dimostra che, se x_0 non è l'estremo sinistro dell'intervallo I , allora

$$\lim_{x \rightarrow x_0^-} \frac{f(x)}{g(x)} = l,$$

per cui il teorema è dimostrato, nel caso in cui $l \in \mathbb{R}$.

Supponiamo ora $l = +\infty$ e che x_0 non sia l'estremo destro dell'intervallo I . Fissiamo $\alpha > 0$. Allora esiste un $\delta_1 > 0$ tale che

$$x_0 < x < x_0 + \delta_1 \Rightarrow \frac{f'(x)}{g'(x)} \geq 2\alpha.$$

Procedendo come sopra, possiamo dedurre che

$$x_0 < x < x_0 + \delta_1 \Rightarrow \frac{f(x_0 + \delta_1) - f(x)}{g(x_0 + \delta_1) - g(x)} \geq 2\alpha.$$

Possiamo inoltre supporre che δ_1 sia tale che

$$x_0 < x < x_0 + \delta_1 \Rightarrow f(x) \neq 0 \text{ e } g(x) \neq 0.$$

Definiamo $\psi(x)$ come sopra. Esiste un $\delta \in]0, \delta_1[$ tale che

$$x_0 < x < x_0 + \delta \Rightarrow \psi(x) \leq 2.$$

Quindi, se $x_0 < x < x_0 + \delta$, si ha

$$\frac{1}{\psi(x)} \frac{f(x_0 + \delta_1) - f(x)}{g(x_0 + \delta_1) - g(x)} \geq \frac{1}{\psi(x)} 2\alpha \geq \alpha,$$

da cui

$$\frac{f(x)}{g(x)} \geq \alpha.$$

Abbiamo così dimostrato che

$$\lim_{x \rightarrow x_0^+} \frac{f(x)}{g(x)} = +\infty.$$

In modo del tutto analogo si dimostra che, se x_0 non è l'estremo sinistro dell'intervallo I , allora

$$\lim_{x \rightarrow x_0^-} \frac{f(x)}{g(x)} = +\infty,$$

per cui il teorema è dimostrato, nel caso in cui $l = +\infty$. Il caso $l = -\infty$ è del tutto analogo al precedente. ■

Esempio. Vogliamo calcolare, se esiste, il

$$\lim_{x \rightarrow 0^+} x \ln x.$$

Poniamo $f(x) = \ln x$ e $g(x) = 1/x$. Notiamo che $\lim_{x \rightarrow 0^+} f(x) = -\infty$ e $\lim_{x \rightarrow 0^+} g(x) = +\infty$. Inoltre,

$$\lim_{x \rightarrow 0^+} \frac{f'(x)}{g'(x)} = \lim_{x \rightarrow 0^+} \frac{1/x}{-1/x^2} = \lim_{x \rightarrow 0^+} (-x) = 0.$$

Pertanto, anche

$$\lim_{x \rightarrow 0^+} x \ln x = \lim_{x \rightarrow 0^+} \frac{f(x)}{g(x)} = 0.$$

Anche nel caso indeterminato del tipo $\frac{\infty}{\infty}$ si possono scrivere gli analoghi teoremi se $x_0 = +\infty$ o $-\infty$. Vediamo il primo caso.

Teorema. Sia I un intervallo non limitato superiormente e $f, g : I \rightarrow \mathbb{R}$ due funzioni derivabili, con $g'(x) \neq 0$ per ogni $x \in I$, tali che

$$\lim_{x \rightarrow +\infty} f(x) = \infty, \quad \lim_{x \rightarrow +\infty} g(x) = \infty.$$

Se esiste il limite

$$\lim_{x \rightarrow +\infty} \frac{f'(x)}{g'(x)},$$

allora esiste anche il limite

$$\lim_{x \rightarrow +\infty} \frac{f(x)}{g(x)},$$

e i due coincidono.

La dimostrazione è analoga a quella del caso $\frac{0}{0}$.

La formula di Taylor

Il seguente teorema ci fornisce la cosiddetta “formula di Taylor con resto di Lagrange”.

Teorema. Siano $x \neq x_0$ due punti di un intervallo I e $f : I \rightarrow \mathbb{R}$ una funzione derivabile $n + 1$ volte su I . Allora esiste un $\xi \in]x_0, x[$ tale che

$$f(x) = p_n(x) + r_n(x),$$

dove

$$p_n(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2!}f''(x_0)(x - x_0)^2 + \cdots + \frac{1}{n!}f^{(n)}(x_0)(x - x_0)^n$$

è il “polinomio di Taylor di grado n associato alla funzione f nel punto x_0 ” e

$$r_n(x) = \frac{1}{(n+1)!}f^{(n+1)}(\xi)(x - x_0)^{n+1}$$

è il “resto di Lagrange”.

Dimostrazione. Osserviamo che il polinomio p_n soddisfa alle seguenti proprietà:

$$\begin{cases} p_n(x_0) = f(x_0), \\ p'_n(x_0) = f'(x_0), \\ p''_n(x_0) = f''(x_0), \\ \vdots \\ p^{(n)}_n(x_0) = f^{(n)}(x_0). \end{cases}$$

Applicando il teorema di Cauchy, troviamo un $\xi_1 \in]x_0, x[$ tale che

$$\frac{f(x) - p_n(x)}{(x - x_0)^{n+1}} = \frac{(f(x) - p_n(x)) - (f(x_0) - p_n(x_0))}{(x - x_0)^{n+1} - (x_0 - x_0)^{n+1}} = \frac{f'(\xi_1) - p'_n(\xi_1)}{(n+1)(\xi_1 - x_0)^n}.$$

Applicando di nuovo il teorema di Cauchy, troviamo un $\xi_2 \in]x_0, \xi_1[$ tale che

$$\frac{f'(\xi_1) - p'_n(\xi_1)}{(n+1)(\xi_1 - x_0)^n} = \frac{(f'(\xi_1) - p'_n(\xi_1)) - (f'(x_0) - p'_n(x_0))}{(n+1)(\xi_1 - x_0)^n - (n+1)(x_0 - x_0)^n} = \frac{f''(\xi_2) - p''_n(\xi_2)}{(n+1)n(\xi_2 - x_0)^{n-1}}.$$

Procedendo per induzione, troviamo $n + 1$ elementi $\xi_1, \xi_2, \dots, \xi_{n+1}$ tali che

$$\begin{aligned} \frac{f(x) - p_n(x)}{(x - x_0)^{n+1}} &= \frac{f'(\xi_1) - p'_n(\xi_1)}{(n+1)(\xi_1 - x_0)^n} \\ &= \frac{f''(\xi_2) - p''_n(\xi_2)}{(n+1)n(\xi_2 - x_0)^{n-1}} \\ &\vdots \\ &= \frac{f^{(n+1)}(\xi_{n+1}) - p_n^{(n+1)}(\xi_{n+1})}{(n+1)!(\xi_{n+1} - x_0)^0}. \end{aligned}$$

Se $x > x_0$, si ha

$$x_0 < \xi_{n+1} < \xi_n < \dots < \xi_2 < \xi_1 < x,$$

mentre se $x < x_0$ si ha l'ordine opposto. Essendo la derivata $(n+1)$ -esima di un polinomio di grado n sempre nulla, si ha che $p_n^{(n+1)}(\xi_{n+1}) = 0$ e ponendo $\xi = \xi_{n+1}$ si ottiene la tesi. ■

Osserviamo che, se $n = 0$, si ha l'equivalente del teorema di Lagrange:

$$f(x) = f(x_0) + f'(\xi)(x - x_0), \quad \text{per un certo } \xi \in]x_0, x[.$$

Si noti che il polinomio di Taylor

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

potrebbe in realtà avere un grado inferiore a n (ad esempio, per una funzione costante, $p_n(x)$ ha sempre grado 0).

Esempi. Determiniamo il polinomio di Taylor di alcune funzioni considerando per semplicità il caso $x_0 = 0$.

1) Sia $f(x) = e^x$. Si ha:

$$p_n(x) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \dots + \frac{x^n}{n!} = \sum_{k=0}^n \frac{x^k}{k!}.$$

2) Sia $f(x) = \cos x$. Allora, se $n = 2m$ o $n = 2m + 1$,

$$p_n(x) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \dots + (-1)^m \frac{x^{2m}}{(2m)!} = \sum_{k=0}^m (-1)^k \frac{x^{2k}}{(2k)!}.$$

3) Sia $f(x) = \sin x$. Allora, se $n = 2m + 1$ o $n = 2m + 2$,

$$p_n(x) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \dots + (-1)^m \frac{x^{2m+1}}{(2m+1)!} = \sum_{k=0}^m (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

Teorema. Per ogni $x \in \mathbb{R}$, si ha che

$$e^x = \lim_n \left(1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} \right).$$

Dimostrazione. La formula è chiaramente vera se $x = 0$. Se $x \neq 0$, per la formula di Taylor con resto di Lagrange, esiste un $\xi \in]0, x[$ tale che $f(x) = p_n(x) + r_n(x)$, con

$$r_n(x) = e^\xi \frac{x^{n+1}}{(n+1)!}.$$

Vogliamo dimostrare che $\lim_n r_n(x) = 0$. Osserviamo che

$$|r_n(x)| \leq e^{|x|} \frac{|x|^{n+1}}{(n+1)!},$$

e sappiamo che, per ogni $a > 0$, si ha $\lim_n \frac{a^n}{n!} = 0$. Ne segue la tesi. ■

Scriveremo brevemente

$$e^x = \sum_{k=0}^{\infty} \frac{x^k}{k!},$$

la “serie di Taylor” associata alla funzione esponenziale nel punto $x_0 = 0$.

Con analoga dimostrazione, si ha pure il seguente

Teorema. Per ogni $x \in \mathbb{R}$, si ha che

$$\begin{aligned} \cos x &= \lim_m \left(1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots + (-1)^m \frac{x^{2m}}{(2m)!} \right), \\ \sin x &= \lim_m \left(x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^m \frac{x^{2m+1}}{(2m+1)!} \right). \end{aligned}$$

Scriveremo brevemente:

$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!}, \quad \sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!}.$$

Queste funzioni, per cui si ha che $f(x) = \lim_n p_n(x)$, si chiamano *analitiche*. Non tutte le funzioni lo sono. Per esempio, la funzione

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

è derivabile infinite volte, e $f^{(k)}(0) = 0$ per ogni $k \in \mathbb{N}$, per cui $p_n(x)$ è identicamente nullo.

Riportiamo nella seguente tabella i polinomi di Taylor associati ad alcune funzioni elementari, nel punto $x_0 = 0$.

$f(x)$	$p_n(x)$ nel punto $x_0 = 0$
e^x	$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!}$
$\ln(1+x)$	$x - \frac{x^2}{2} + \frac{x^3}{3} - \frac{x^4}{4} + \cdots + \frac{(-1)^{n-1}x^n}{n}$
$\cos x$	$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \cdots + (-1)^m \frac{x^{2m}}{(2m)!}$
$\sin x$	$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \frac{x^7}{7!} + \cdots + (-1)^m \frac{x^{2m+1}}{(2m+1)!}$
$\arctan x$	$x - \frac{x^3}{3} + \frac{x^5}{5} - \frac{x^7}{7} + \cdots + (-1)^m \frac{x^{2m+1}}{2m+1}$
$\cosh x$	$1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \frac{x^6}{6!} + \cdots + \frac{x^{2m}}{(2m)!}$
$\sinh x$	$x + \frac{x^3}{3!} + \frac{x^5}{5!} + \frac{x^7}{7!} + \cdots + \frac{x^{2m+1}}{(2m+1)!}$
$\tanh^{-1} x$	$x + \frac{x^3}{3} + \frac{x^5}{5} + \frac{x^7}{7} + \cdots + \frac{x^{2m+1}}{2m+1}$

Non risulta invece elementare la formula del polinomio di Taylor per le funzioni $\tan x$ e $\tanh x$, di cui riportiamo solo i primi termini.

$$\begin{array}{l|l} \tan x & x + \frac{x^3}{3} + \frac{2x^5}{15} + \frac{17x^7}{315} + \cdots \\ \tanh x & x - \frac{x^3}{3} + \frac{2x^5}{15} - \frac{17x^7}{315} + \cdots \end{array}$$

Tenendo fisso x_0 , vogliamo ora vedere cosa succede se $x \rightarrow x_0$. Per ogni $x \neq x_0$, per evidenziare il fatto che il nostro $\xi \in]x_0, x[$ dipende da x , scriviamo $\xi = \xi_x$. Allora, se $f^{(n+1)}$ è *limitata in un intorno di x_0* , avremo che

$$\lim_{x \rightarrow x_0} \frac{r_n(x)}{(x - x_0)^n} = \lim_{x \rightarrow x_0} \frac{1}{(n+1)!} f^{(n+1)}(\xi_x)(x - x_0) = 0.$$

Talvolta per questa relazione di limite si usa la seguente notazione:

$$r_n(x) = o(|x - x_0|^n) \quad \text{se } x \rightarrow x_0.$$

Notiamo ancora che, se $f^{(n+1)}$ è *continua in* x_0 , allora

$$\lim_{x \rightarrow x_0} \frac{r_n(x)}{(x - x_0)^{n+1}} = \lim_{x \rightarrow x_0} \frac{1}{(n+1)!} f^{(n+1)}(\xi_x) = \frac{1}{(n+1)!} f^{(n+1)}(x_0).$$

Questo ci permette, in alcuni casi, di determinare gli eventuali punti di massimo o di minimo locale di una funzione. Ad esempio, se x_0 è un punto, interno al dominio I di f , in cui $f'(x_0) = 0$ e $f''(x_0) > 0$, allora

$$\lim_{x \rightarrow x_0} \frac{f(x) - f(x_0)}{(x - x_0)^2} = \lim_{x \rightarrow x_0} \frac{f(x) - [f(x_0) + f'(x_0)(x - x_0)]}{(x - x_0)^2} = \frac{1}{2!} f''(x_0) > 0,$$

da cui si deduce, usando il teorema sulla permanenza del segno, che esiste un intorno U di x_0 tale che $f(x) > f(x_0)$, per ogni $x \in U \setminus \{x_0\}$. Pertanto,

se $f'(x_0) = 0$ e $f''(x_0) > 0$, allora x_0 è un punto di minimo locale.

Analogamente si vede che

se $f'(x_0) = 0$ e $f''(x_0) < 0$, allora x_0 è un punto di massimo locale.

Qualora $f'(x_0) = 0$ e anche $f''(x_0) = 0$, dovremo guardare alla derivata terza. Se $f'''(x_0) \neq 0$, allora x_0 non è nè di minimo nè di massimo locale. Se anche $f'''(x_0) = 0$, allora si vede che

*se $f'(x_0) = f''(x_0) = f'''(x_0) = 0$ e $f^{(4)}(x_0) > 0$,
allora x_0 è un punto di minimo locale.*

mentre

*se $f'(x_0) = f''(x_0) = f'''(x_0) = 0$ e $f^{(4)}(x_0) < 0$,
allora x_0 è un punto di massimo locale.*

Il procedimento può essere continuato. Tralasciamo i dettagli, per brevità.

L'integrale - alcune premesse

In questa parte del corso denoteremo con I l'intervallo *compatto* $[a, b]$. Innanzitutto scegliamo in I alcuni punti

$$a = a_0 < a_1 < \cdots < a_{m-1} < a_m = b,$$

ottenendo così una “partizione” di I costituita dagli intervalli $[a_{j-1}, a_j]$, con $j = 1, \dots, m$. Ora, per ogni j , scegliamo un punto

$$x_j \in [a_{j-1}, a_j].$$

Otteniamo così una “partizione puntata” o “P-partizione” di I :

$$\mathring{\mathcal{P}} = \left\{ (x_1, [a_0, a_1]), \dots, (x_m, [a_{m-1}, a_m]) \right\}.$$

Esempi. Sia $I = [0, 1]$. Queste sono alcune P-partizioni di I :

$$\mathring{\mathcal{P}} = \left\{ \left(\frac{1}{6}, [0, 1] \right) \right\}$$

$$\mathring{\mathcal{P}} = \left\{ \left(0, \left[0, \frac{1}{3} \right] \right), \left(\frac{1}{2}, \left[\frac{1}{3}, 1 \right] \right) \right\}$$

$$\mathring{\mathcal{P}} = \left\{ \left(\frac{1}{3}, \left[0, \frac{1}{3} \right] \right), \left(\frac{1}{3}, \left[\frac{1}{3}, \frac{2}{3} \right] \right), \left(\frac{2}{3}, \left[\frac{2}{3}, 1 \right] \right) \right\}$$

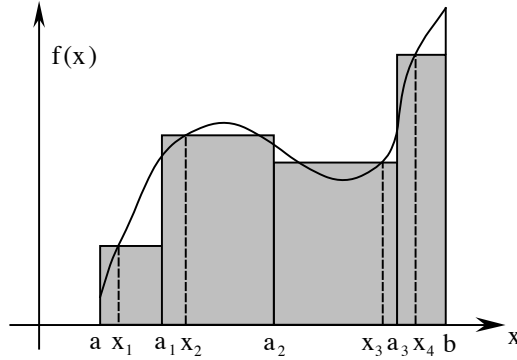
$$\mathring{\mathcal{P}} = \left\{ \left(\frac{1}{8}, \left[0, \frac{1}{4} \right] \right), \left(\frac{3}{8}, \left[\frac{1}{4}, \frac{1}{2} \right] \right), \left(\frac{5}{8}, \left[\frac{1}{2}, \frac{3}{4} \right] \right), \left(\frac{7}{8}, \left[\frac{3}{4}, 1 \right] \right) \right\}.$$

Consideriamo ora una funzione $f : I \rightarrow \mathbb{R}$. Per ogni P-partizione $\mathring{\mathcal{P}}$ come sopra definiamo il numero

$$S(f, \mathring{\mathcal{P}}) = \sum_{j=1}^m f(x_j)(a_j - a_{j-1}),$$

che chiameremo “somma di Riemann” associata a f e $\mathring{\mathcal{P}}$.

Per comprendere meglio questa definizione, assumiamo per semplicità che la funzione f sia positiva. In tal caso, $S(f, \mathring{\mathcal{P}})$ rappresenta la somma delle aree dei rettangoli aventi base $[a_{j-1}, a_j]$ e altezza $[0, f(x_j)]$.



Se f non è positiva su I , le aree verranno considerate con segno positivo o negativo a seconda che $f(x_j)$ sia rispettivamente positivo o negativo. Se $f(x_j) = 0$, il j -esimo termine della somma sarà chiaramente uguale a zero.

Esempio. Sia $f : [0, 1] \rightarrow \mathbb{R}$ definita da $f(x) = 4x^2 - 1$, e sia

$$\mathring{\mathcal{P}} = \left\{ \left(\frac{1}{8}, \left[0, \frac{1}{4} \right] \right), \left(\frac{1}{2}, \left[\frac{1}{4}, \frac{3}{4} \right] \right), \left(\frac{7}{8}, \left[\frac{3}{4}, 1 \right] \right) \right\}.$$

Allora,

$$S(f, \mathring{\mathcal{P}}) = -\frac{15}{16} \cdot \frac{1}{4} + 0 \cdot \frac{1}{2} + \frac{33}{16} \cdot \frac{1}{4} = \frac{9}{32}.$$

Con l'obiettivo di misurare quanto “fine” sia una P-partizione, dovremo far uso di un “calibro”, cioè di una funzione *positiva* $\delta : I \rightarrow \mathbb{R}$.

Se δ è un calibre su I , diciamo che la P-partizione $\mathring{P}$ è “ δ -fine” se, per ogni $j = 1, \dots, m$,

$$x_j - a_{j-1} \leq \delta(x_j) \quad \text{e} \quad a_j - x_j \leq \delta(x_j);$$

equivalentemente, potremmo scrivere

$$[a_{j-1}, a_j] \subseteq [x_j - \delta(x_j), x_j + \delta(x_j)].$$

Mostriamo ora che è sempre possibile trovare una P-partizione δ -fine dell'intervallo compatto I , qualunque sia il calibre δ .

Teorema. *Per ogni calibre δ su $I = [a, b]$ esiste una P-partizione δ -fine di I .*

Dimostrazione. Poniamo $I_0 = I$ e assumiamo per assurdo che ci sia un calibre $\delta : I_0 \rightarrow \mathbb{R}$ per il quale non esistano P-partizioni δ -fini. Prendendo il punto medio di I_0 , lo dividiamo in due sottointervalli *chiusi*. Almeno uno di questi due sottointervalli non ha alcuna P-partizione δ -fine (altrimenti potremmo incollare insieme le due P-partizioni δ -fini per ottenere una P-partizione δ -fine dell'intervallo originale I_0). Scegliamolo e denotiamolo con I_1 . Ora iteriamo la stessa procedura, costruendo così una successione

$$I = I_0 \supseteq I_1 \supseteq I_2 \supseteq I_3 \supseteq \dots$$

di sottointervalli chiusi, ciascuno dei quali non ha P-partizioni δ -fini. Per il Teorema di Cantor esiste un punto c che appartiene a tutti questi intervalli. Per n sufficientemente grande, I_n sarà contenuto in $[c - \delta(c), c + \delta(c)]$. Ma allora l'insieme $\mathring{P} = \{(c, I_n)\}$, il cui unico elemento è la coppia (c, I_n) , è una P-partizione δ -fine di I_n , una contraddizione. ■

Esempi. Forniamo alcuni esempi di P-partizioni δ -fini dell'intervallo $I = [0, 1]$.

Iniziamo con un calibre costante: $\delta(x) = \frac{1}{5}$. Poiché il teorema precedente non fornisce alcuna informazione su come trovare una P-partizione δ -fine, procederemo per tentativi. Come prima prova scegliamo gli a_j equidistanti e gli x_j come punti medi degli intervalli $[a_{j-1}, a_j]$, ossia

$$a_j = \frac{j}{m}, \quad x_j = \frac{a_{j-1} + a_j}{2} = \frac{2j-1}{2m}.$$

Affinché la P-partizione corrispondente sia δ -fine, deve essere

$$x_j - a_{j-1} = \frac{1}{2m} \leq \frac{1}{5} \quad \text{e} \quad a_j - x_j = \frac{1}{2m} \leq \frac{1}{5}.$$

Queste disuguaglianze vengono soddisfatte scegliendo $m \geq 3$. Se $m = 3$, abbiamo la P-partizione δ -fine

$$\mathring{\mathcal{P}} = \left\{ \left(\frac{1}{6}, \left[0, \frac{1}{3} \right] \right), \left(\frac{1}{2}, \left[\frac{1}{3}, \frac{2}{3} \right] \right), \left(\frac{5}{6}, \left[\frac{2}{3}, 1 \right] \right) \right\} .$$

Se invece di prendere i punti x_j al centro dei rispettivi intervalli vorremmo sceglierli, ad esempio, all'estremità sinistra, cioè $x_j = \frac{j-1}{m}$, per avere una P-partizione δ -fine bisognerebbe che sia

$$x_j - a_{j-1} = 0 \leq \frac{1}{5} \quad \text{e} \quad a_j - x_j = \frac{1}{m} \leq \frac{1}{5} .$$

Queste disuguaglianze sono verificate se $m \geq 5$. Ad esempio, se $m = 5$, abbiamo la P-partizione δ -fine

$$\mathring{\mathcal{P}} = \left\{ \left(0, \left[0, \frac{1}{5} \right] \right), \left(\frac{1}{5}, \left[\frac{1}{5}, \frac{2}{5} \right] \right), \left(\frac{2}{5}, \left[\frac{2}{5}, \frac{3}{5} \right] \right), \left(\frac{3}{5}, \left[\frac{3}{5}, \frac{4}{5} \right] \right), \left(\frac{4}{5}, \left[\frac{4}{5}, 1 \right] \right) \right\} .$$

Si noti che, con tale scelta degli a_j , se $m \geq 5$, i punti x_j possono effettivamente essere presi arbitrariamente nei rispettivi intervalli $[a_{j-1}, a_j]$, ottenendo ancora P-partizioni δ -fini.

L'esempio precedente mostra come sia possibile costruire P-partizioni δ -fini nel caso di un calibro δ che è costante con valore $\frac{1}{5}$. È chiaro che una procedura simile può essere utilizzata per un calibro costante con valore positivo arbitrario. Consideriamo ora il caso in cui δ sia una funzione continua. In tal caso, il Teorema di Weierstrass dice che $\delta(x)$ ha un valore minimo positivo: sia esso $\bar{\delta}$. Consideriamo quindi il calibro costante con valore $\bar{\delta}$, e costruiamo una P-partizione $\bar{\delta}$ -fine con la procedura che abbiamo visto sopra. Chiaramente, una tale P-partizione deve essere anche δ -fine. Questo argomento mostra come il caso di un calibro continuo possa essere ricondotto a quello di un calibro costante.

Consideriamo ora il calibro non continuo

$$\delta(x) = \begin{cases} \frac{1}{2} & \text{se } x = 0, \\ \frac{x}{2} & \text{se } x \in]0, 1]. \end{cases}$$

Come prima, procediamo per tentativi. Proviamo come sopra, prendendo gli a_j equidistanti e gli x_j come punti medi degli intervalli $[a_{j-1}, a_j]$. Questa volta, però, il tentativo è destinato a fallire; infatti, dovremmo avere che

$$x_1 = x_1 - a_0 \leq \delta(x_1) = \frac{x_1}{2},$$

il che è chiaramente impossibile se $x_1 > 0$. L'unico modo per risolvere questo problema è scegliere $x_1 = 0$. Decidiamo allora, ad esempio, di prendere gli x_j coincidenti con a_{j-1} , come fatto sopra. Troviamo così la P-partizione δ -fine

$$\mathring{\mathcal{P}} = \left\{ \left(0, \left[0, \frac{1}{2} \right] \right), \left(\frac{1}{2}, \left[\frac{1}{2}, \frac{3}{4} \right] \right), \left(\frac{3}{4}, \left[\frac{3}{4}, 1 \right] \right) \right\} .$$

Si noti che una scelta più economica avrebbe potuto essere

$$\mathring{\mathcal{P}} = \left\{ \left(0, \left[0, \frac{1}{2}\right]\right), \left(1, \left[\frac{1}{2}, 1\right]\right) \right\}.$$

La scelta $x_1 = 0$ è comunque inevitabile.

Infine, una volta fissato un punto $c \in]0, 1[$, definiamo il calibro $\delta : [0, 1] \rightarrow \mathbb{R}$ come

$$\delta(x) = \begin{cases} \frac{c-x}{2} & \text{se } x \in [0, c[, \\ \frac{1}{5} & \text{se } x = c, \\ \frac{x-c}{2} & \text{se } x \in]c, 1]. \end{cases}$$

Considerazioni analoghe a quelle fatte nel caso precedente portano alla conclusione che, per avere una P-partizione δ -fine, è necessario che uno degli x_j sia uguale a c . Ad esempio, se $c = \frac{1}{2}$, una scelta possibile è

$$\mathring{\mathcal{P}} = \left\{ \left(0, \left[0, \frac{1}{4}\right]\right), \left(\frac{1}{4}, \left[\frac{1}{4}, \frac{3}{8}\right]\right), \left(\frac{1}{2}, \left[\frac{3}{8}, \frac{5}{8}\right]\right), \left(\frac{3}{4}, \left[\frac{5}{8}, \frac{3}{4}\right]\right), \left(1, \left[\frac{3}{4}, 1\right]\right) \right\}.$$

L'integrale - definizione

Vogliamo ora definire una sorta di convergenza delle somme di Riemann, quando le P-partizioni diventano “sempre più fini”. La seguente definizione è dovuta a Jaroslav Kurzweil e Ralph Henstock.

Definizione. Si dice che una funzione $f : I \rightarrow \mathbb{R}$ è “integrabile” se esiste un numero reale $\mathcal{J}$ con la seguente proprietà: dato $\varepsilon > 0$, è possibile trovare un calibro $\delta : I \rightarrow \mathbb{R}$ tale che, per ogni P-partizione δ -fine $\mathring{\mathcal{P}}$ di I ,

$$|S(f, \mathring{\mathcal{P}}) - \mathcal{J}| \leq \varepsilon.$$

Diremo anche che f è “integrabile su I ”.

Dimostriamo che esiste al massimo un $\mathcal{J} \in \mathbb{R}$ che verifica le condizioni della definizione. Se ce ne fosse un secondo, diciamo $\mathcal{J}'$, allora per ogni $\varepsilon > 0$ ci sarebbero due calibri δ e δ' su I associati rispettivamente a $\mathcal{J}$ e $\mathcal{J}'$, soddisfacenti la condizione della definizione. Definiamo il calibro

$$\delta''(x) = \min\{\delta(x), \delta'(x)\}.$$

Una volta scelta una P-partizione δ'' -fine $\mathring{\mathcal{P}}$ di I , abbiamo che $\mathring{\mathcal{P}}$ è sia δ -fine che δ' -fine, quindi

$$|\mathcal{J} - \mathcal{J}'| \leq |\mathcal{J} - S(f, \mathring{\mathcal{P}})| + |S(f, \mathring{\mathcal{P}}) - \mathcal{J}'| \leq 2\varepsilon.$$

Poiché questo vale per ogni $\varepsilon > 0$, deve necessariamente essere $\mathcal{J} = \mathcal{J}'$.

Se $f : I \rightarrow \mathbb{R}$ è una funzione integrabile, l'unico elemento $\mathcal{J} \in \mathbb{R}$ che verifica le condizioni della definizione è detto “integrale” di f su I ed è indicato con uno dei seguenti simboli:

$$\int_I f, \quad \int_a^b f, \quad \int_I f(x) dx, \quad \int_a^b f(x) dx.$$

La presenza della lettera x nella notazione precedente non ha un'importanza particolare. Potrebbe essere sostituita da qualsiasi altra lettera $t, u, \alpha, \dots$, o da qualsiasi altro simbolo, a meno che non sia già utilizzato con altro significato. Per ragioni che verranno spiegate in seguito, poniamo inoltre

$$\int_b^a f = - \int_a^b f \quad \text{e} \quad \int_a^a f = 0.$$

Esempi. 1. Come primo esempio consideriamo una funzione costante $f(x) = c$. In questo caso, per qualsiasi P-partizione $\mathring{\mathcal{P}}$ di $[a, b]$,

$$S(f, \mathring{\mathcal{P}}) = \sum_{j=1}^m c(a_j - a_{j-1}) = c \sum_{j=1}^m (a_j - a_{j-1}) = c(b - a),$$

quindi anche

$$\int_a^b c dx = c(b - a).$$

Infatti, una volta fissato $\varepsilon > 0$, in questo caso semplice si vede facilmente che qualsiasi calibro $\delta : [a, b] \rightarrow \mathbb{R}$ soddisfa la condizione della definizione, con $\mathcal{J} = c(b - a)$, poiché per ogni P-partizione δ -fine $\mathring{\mathcal{P}}$ di I si ha $|S(f, \mathring{\mathcal{P}}) - \mathcal{J}| = 0 < \varepsilon$.

2. Come secondo esempio, consideriamo la funzione $f(x) = x$, per cui

$$S(f, \mathring{\mathcal{P}}) = \sum_{j=1}^m x_j(a_j - a_{j-1}).$$

Per trovare un candidato per l'integrale, consideriamo la P-partizione in cui gli x_j sono i punti medi degli intervalli $[a_{j-1}, a_j]$. In questo caso particolare, abbiamo

$$\sum_{j=1}^m x_j(a_j - a_{j-1}) = \sum_{j=1}^m \frac{a_{j-1} + a_j}{2}(a_j - a_{j-1}) = \frac{1}{2} \sum_{j=1}^m (a_j^2 - a_{j-1}^2) = \frac{1}{2}(b^2 - a^2).$$

Vogliamo ora dimostrare che la funzione $f(x) = x$ è integrabile su $[a, b]$ e che il suo

integrale è proprio $\frac{1}{2}(b^2 - a^2)$. Fissiamo $\varepsilon > 0$. Per qualsiasi P-partizione $\mathring{\mathcal{P}}$, abbiamo:

$$\begin{aligned} \left| S(f, \mathring{\mathcal{P}}) - \frac{1}{2}(b^2 - a^2) \right| &= \left| \sum_{j=1}^m x_j(a_j - a_{j-1}) - \sum_{j=1}^m \frac{a_{j-1} + a_j}{2}(a_j - a_{j-1}) \right| \\ &\leq \sum_{j=1}^m \left| x_j - \frac{a_{j-1} + a_j}{2} \right| (a_j - a_{j-1}) \\ &\leq \sum_{j=1}^m \frac{a_j - a_{j-1}}{2} (a_j - a_{j-1}). \end{aligned}$$

Se scegliamo che il calibro δ sia costante con valore $\frac{\varepsilon}{b-a}$, per ogni P-partizione δ -fine $\mathring{\mathcal{P}}$ abbiamo che

$$\left| S(f, \mathring{\mathcal{P}}) - \frac{1}{2}(b^2 - a^2) \right| \leq \sum_{j=1}^m \frac{2\delta}{2} (a_j - a_{j-1}) = \frac{\varepsilon}{b-a} \sum_{j=1}^m (a_j - a_{j-1}) = \varepsilon.$$

Con questa scelta del calibro si verifica quindi la condizione della definizione, e abbiamo dimostrato che

$$\int_a^b x \, dx = \frac{1}{2}(b^2 - a^2).$$

Proprietà elementari dell'integrale

Siano $f : I \rightarrow \mathbb{R}$ e $g : I \rightarrow \mathbb{R}$ due funzioni reali, e $\alpha \in \mathbb{R}$ sia una costante. È facile verificare che, per ogni P-partizione $\mathring{\mathcal{P}}$ di I ,

$$S(f + g, \mathring{\mathcal{P}}) = S(f, \mathring{\mathcal{P}}) + S(g, \mathring{\mathcal{P}}),$$

e

$$S(\alpha f, \mathring{\mathcal{P}}) = \alpha S(f, \mathring{\mathcal{P}}).$$

Queste proprietà di linearità sono ereditate dall'integrale, come sarà dimostrato nelle due proposizioni seguenti.

Proposizione. *Se f e g sono integrabili su I , allora $f + g$ è integrabile su I e*

$$\int_I (f + g) = \int_I f + \int_I g.$$

Dimostrazione. Poniamo $\mathcal{J}_1 = \int_I f$ e $\mathcal{J}_2 = \int_I g$. Una volta fissato $\varepsilon > 0$, ci sono due calibri δ_1 e δ_2 su I tali che, per ogni P-partizione $\mathring{\mathcal{P}}$ di I , se $\mathring{\mathcal{P}}$ è δ_1 -fine, allora

$$|S(f, \mathring{\mathcal{P}}) - \mathcal{J}_1| \leq \frac{\varepsilon}{2},$$

mentre se $\mathring{\mathcal{P}}$ è δ_2 -fine, allora

$$|S(g, \mathring{\mathcal{P}}) - \mathcal{J}_2| \leq \frac{\varepsilon}{2}.$$

Definiamo il calibro $\delta : I \rightarrow \mathbb{R}$ come $\delta(x) = \min\{\delta_1(x), \delta_2(x)\}$. Sia $\mathring{\mathcal{P}}$ una P-partizione δ -fine di I . Essa è sia δ_1 -fine che δ_2 -fine, per cui

$$\begin{aligned} |S(f + g, \mathring{\mathcal{P}}) - (\mathcal{J}_1 + \mathcal{J}_2)| &= |S(f, \mathring{\mathcal{P}}) - \mathcal{J}_1 + S(g, \mathring{\mathcal{P}}) - \mathcal{J}_2| \\ &\leq |S(f, \mathring{\mathcal{P}}) - \mathcal{J}_1| + |S(g, \mathring{\mathcal{P}}) - \mathcal{J}_2| \\ &\leq \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon. \end{aligned}$$

Questo completa la dimostrazione. ■

Proposizione. Se f è integrabile su I e $\alpha \in \mathbb{R}$, allora αf è integrabile su I e

$$\int_I (\alpha f) = \alpha \int_I f.$$

Dimostrazione. Se $\alpha = 0$, l'identità è sicuramente vera. Se $\alpha \neq 0$, poniamo $\mathcal{J} = \int_I f$ e fissiamo $\varepsilon > 0$. Esiste un calibro δ su I tale che

$$|S(f, \mathring{\mathcal{P}}) - \mathcal{J}| \leq \frac{\varepsilon}{|\alpha|},$$

per ogni P-partizione δ -fine $\mathring{\mathcal{P}}$ di I . Quindi, per ogni P-partizione δ -fine $\mathring{\mathcal{P}}$ di I , abbiamo che

$$|S(\alpha f, \mathring{\mathcal{P}}) - \alpha \mathcal{J}| = |\alpha S(f, \mathring{\mathcal{P}}) - \alpha \mathcal{J}| = |\alpha| |S(f, \mathring{\mathcal{P}}) - \mathcal{J}| \leq |\alpha| \frac{\varepsilon}{|\alpha|} = \varepsilon,$$

e la dimostrazione è così completata. ■

Abbiamo appena dimostrato che l'insieme delle funzioni integrabili è uno spazio vettoriale e che l'integrale è una funzione lineare su di esso.

Studiamo ora il comportamento dell'integrale rispetto alla relazione d'ordine in $\mathbb{R}$.

Proposizione. Se f è integrabile su I e $f(x) \geq 0$ per ogni $x \in I$, allora

$$\int_I f \geq 0.$$

Dimostrazione. Fissiamo $\varepsilon > 0$. C'è un calibro δ su I tale che

$$\left| S(f, \mathring{\mathcal{P}}) - \int_I f \right| \leq \varepsilon,$$

per ogni P-partizione δ -fine $\mathring{\mathcal{P}}$ di I . Quindi, essendo chiaramente $S(f, \mathring{\mathcal{P}}) \geq 0$, si ha che

$$\int_I f \geq S(f, \mathring{\mathcal{P}}) - \varepsilon \geq -\varepsilon.$$

Poiché questo è vero per ogni $\varepsilon > 0$, deve essere $\int_I f \geq 0$, dimostrando così il risultato. ■

Corollario. Se f e g sono integrabili su I e $f(x) \leq g(x)$ per ogni $x \in I$, allora

$$\int_I f \leq \int_I g.$$

Dimostrazione. È sufficiente applicare la proposizione precedente alla funzione $g - f$ e usare l'uguaglianza $\int_I (g - f) = \int_I g - \int_I f$. ■

Corollario. Se f e $|f|$ sono integrabili su I , allora

$$\left| \int_I f \right| \leq \int_I |f|.$$

Dimostrazione. Applicando il corollario precedente alle disuguaglianze

$$-|f| \leq f \leq |f|,$$

abbiamo che

$$-\int_I |f| \leq \int_I f \leq \int_I |f|,$$

da cui la conclusione. ■

Definizione. Diremo che la funzione integrabile $f : I \rightarrow \mathbb{R}$ è “*L-integrabile*” (o “*integrabile secondo Lebesgue*”) se anche $|f|$ è integrabile.

Il Teorema Fondamentale

Il seguente teorema stabilisce un legame inaspettato tra il calcolo differenziale e la teoria dell'integrale. Si tratta del **Teorema Fondamentale** del calcolo differenziale e integrale.

Teorema Fondamentale - I Sia $F : [a, b] \rightarrow \mathbb{R}$ una funzione derivabile e sia f la sua derivata: $F'(x) = f(x)$ per ogni $x \in [a, b]$. Allora f è integrabile su $[a, b]$ e

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

Dimostrazione. Sia $\varepsilon > 0$ fissato. Sappiamo che, per ogni $x \in [a, b]$,

$$f(x) = F'(x) = \lim_{u \rightarrow x} \frac{F(u) - F(x)}{u - x}.$$

Allora, per ogni $x \in [a, b]$ esiste un $\delta(x) > 0$ tale che, per ogni $u \in [a, b]$,

$$0 < |u - x| \leq \delta(x) \Rightarrow \left| \frac{F(u) - F(x)}{u - x} - f(x) \right| \leq \frac{\varepsilon}{b - a},$$

cioè

$$|u - x| \leq \delta(x) \Rightarrow |F(u) - F(x) - f(x)(u - x)| \leq \frac{\varepsilon}{b - a} |u - x|.$$

Abbiamo così definito un calibro $\delta : [a, b] \rightarrow \mathbb{R}$.

Consideriamo ora una P-partizione δ -fine di I ,

$$\mathring{\mathcal{P}} = \{(x_1, [a_0, a_1]), \dots, (x_m, [a_{m-1}, a_m])\}.$$

Poiché, per ogni $j = 1, \dots, m$,

$$|a_{j-1} - x_j| \leq \delta(x_j) \quad \text{e} \quad |a_j - x_j| \leq \delta(x_j),$$

abbiamo che

$$\begin{aligned} & \left| F(a_j) - F(a_{j-1}) - f(x_j)(a_j - a_{j-1}) \right| = \\ & = \left| [F(a_j) - F(x_j) - f(x_j)(a_j - x_j)] + [F(x_j) - F(a_{j-1}) + f(x_j)(a_{j-1} - x_j)] \right| \\ & \leq \left| F(a_j) - F(x_j) - f(x_j)(a_j - x_j) \right| + \left| F(a_{j-1}) - F(x_j) - f(x_j)(a_{j-1} - x_j) \right| \\ & \leq \frac{\varepsilon}{b - a} |a_j - x_j| + \frac{\varepsilon}{b - a} |a_{j-1} - x_j| \\ & = \frac{\varepsilon}{b - a} (a_j - x_j + x_j - a_{j-1}) = \frac{\varepsilon}{b - a} (a_j - a_{j-1}). \end{aligned}$$

Quindi, usando l'uguaglianza

$$F(b) - F(a) = \sum_{j=1}^m [F(a_j) - F(a_{j-1})],$$

si ha che

$$\begin{aligned}
|F(b) - F(a) - S(f, \mathring{\mathcal{P}})| &= \left| \sum_{j=1}^m [F(a_j) - F(a_{j-1})] - \sum_{j=1}^m f(x_j)(a_j - a_{j-1}) \right| \\
&= \left| \sum_{j=1}^m [F(a_j) - F(a_{j-1}) - f(x_j)(a_j - a_{j-1})] \right| \\
&\leq \sum_{j=1}^m |F(a_j) - F(a_{j-1}) - f(x_j)(a_j - a_{j-1})| \\
&\leq \sum_{j=1}^m \frac{\varepsilon}{b-a} (a_j - a_{j-1}) = \varepsilon,
\end{aligned}$$

e il teorema è dimostrato. ■

Funzioni primitivabili

In questa sezione e nella successiva indichiamo con $\mathcal{I}$ qualsiasi intervallo di $\mathbb{R}$ (non necessariamente un intervallo compatto).

Una funzione $f : \mathcal{I} \rightarrow \mathbb{R}$ si dice “primitivabile” (o “primitivabile su $\mathcal{I}$ ”) se esiste una funzione derivabile $F : \mathcal{I} \rightarrow \mathbb{R}$ tale che $F'(x) = f(x)$ per ogni $x \in \mathcal{I}$. Tale funzione F si chiama “primitiva” di f .

Il Teorema Fondamentale stabilisce che tutte le funzioni primitivabili definite su un intervallo compatto $I = [a, b]$ sono integrabili, e che il loro integrale è facilmente calcolabile, una volta nota una sua primitiva. Può essere riformulato come segue.

Teorema Fondamentale - II Sia $f : [a, b] \rightarrow \mathbb{R}$ una funzione primitivabile e sia F una delle sue primitive. Allora f è integrabile su $[a, b]$ e

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

Talvolta è utile denotare la differenza $F(b) - F(a)$ con uno dei seguenti simboli:

$$[F]_a^b, \quad [F(x)]_{x=a}^{x=b},$$

o varianti di questi, ad esempio $[F(x)]_a^b$, qualora non ci siano ambiguità.

Esempio. Consideriamo la funzione $f(x) = x^n$. È facile vedere che $F(x) = \frac{1}{n+1}x^{n+1}$ ne è una primitiva. Allora, per il Teorema Fondamentale,

$$\int_a^b x^n dx = \left[\frac{1}{n+1} x^{n+1} \right]_a^b = \frac{1}{n+1} (b^{n+1} - a^{n+1}).$$

Il fatto che la differenza $F(b) - F(a)$ non dipenda dalla primitiva scelta è spiegato dalla seguente proposizione.

Proposizione. *Sia $f : \mathcal{I} \rightarrow \mathbb{R}$ una funzione primitivabile e sia F sia una delle sue primitive. Allora, una funzione $G : \mathcal{I} \rightarrow \mathbb{R}$ è una primitiva di f se e solo se $F - G$ è una funzione costante su $\mathcal{I}$.*

Dimostrazione. Se $F - G$ è costante, allora

$$G'(x) = (F - (F - G))'(x) = F'(x) - (F - G)'(x) = F'(x) = f(x),$$

per ogni $x \in \mathcal{I}$, e quindi G è una primitiva di f . D'altra parte, se G è una primitiva di f , abbiamo

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

per ogni $x \in \mathcal{I}$. Di conseguenza, $F - G$ è costante su $\mathcal{I}$. ■

Si noti che, se $f : \mathcal{I} \rightarrow \mathbb{R}$ è una funzione primitivabile, lo è anche su ogni sottointervallo di $\mathcal{I}$. In particolare è integrabile su ogni intervallo $[a, x] \subseteq \mathcal{I}$, e quindi è possibile definire una funzione

$$x \mapsto \int_a^x f,$$

che chiamiamo “funzione integrale” di f , e denotiamo con uno dei seguenti simboli:

$$\int_a^\cdot f, \quad \int_a^\cdot f(t) dt.$$

In quest'ultima notazione conviene usare una lettera diversa da x per la variabile f ; ad esempio, qui abbiamo scelto la lettera t . Il Teorema Fondamentale ci dice che, se F è una primitiva di f , allora

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

Vediamo quindi che $\int_a^x f$ differisce da $F(x)$ per una costante, da cui quanto segue.

Corollario. *Sia $f : [a, b] \rightarrow \mathbb{R}$ una funzione primitivabile. Allora la funzione integrale $\int_a^\cdot f$ è una delle sue primitive; è derivabile su $[a, b]$ e*

$$\left(\int_a^\cdot f \right)'(x) = f(x), \quad \text{per ogni } x \in [a, b].$$

Si noti che la scelta del punto a nella definizione di $\int_a^\cdot f$ non è affatto obbligatoria. Se $f : \mathcal{I} \rightarrow \mathbb{R}$ è primitivabile, si potrebbe prendere qualsiasi punto $\omega \in \mathcal{I}$ e considerare la funzione $\int_\omega^\cdot f$. Le convenzioni fatte sull'integrale con gli estremi scambiati

sono tali che il precedente corollario continua a valere con questa nuova funzione integrale. Infatti, se F è una primitiva di f , anche se $x < \omega$ abbiamo

$$\int_{\omega}^x f = - \int_x^{\omega} f = -(F(\omega) - F(x)) = F(x) - F(\omega),$$

per cui $\int_{\omega} f$ è ancora una primitiva di f . Possiamo quindi scrivere

$$\frac{d}{dx} \int_{\omega}^x f = f(x), \quad \text{o, equivalentemente,} \quad \frac{d}{dx} \int_{\omega}^x f(t) dt = f(x).$$

Questa formula può essere generalizzata; se $\alpha, \beta : [a, b] \rightarrow \mathbb{R}$ sono due funzioni derivabili, allora

$$\frac{d}{dx} \int_{\alpha(x)}^{\beta(x)} f(t) dt = f(\beta(x))\beta'(x) - f(\alpha(x))\alpha'(x).$$

Infatti, se F è una primitiva di f , la formula precedente si ottiene facilmente scrivendo $\int_{\alpha(x)}^{\beta(x)} f(t) dt = F(\beta(x)) - F(\alpha(x))$ e derivando.

Indicheremo l'insieme di tutte le primitive di f con uno dei seguenti simboli:

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

Bisogna fare attenzione alla notazione $\int$ introdotta per le primitive, che assomiglia a quella per gli integrali, anche se i due concetti sono completamente diversi. Per quanto riguarda l'uso della x , un'osservazione analoga a quella fatta per l'integrale si può fare anche qui: la si può sostituire con qualsiasi altra lettera o simbolo, con le dovute precauzioni. Quando si applica la teoria a problemi pratici, tuttavia, se F denota una primitiva di f , invece di scrivere correttamente

$$\int f = \{F + c : c \in \mathbb{R}\},$$

è frequente utilizzare notazioni improprie del tipo

$$\int f(x) dx = F(x) + c,$$

dove $c \in \mathbb{R}$ rappresenta una costante arbitraria; ci adatteremo anche noi a questa

consuetudine. Facciamo un elenco delle primitive di alcune funzioni elementari:

$$\begin{aligned}\int e^x dx &= e^x + c, \\ \int \sin x dx &= -\cos x + c, \\ \int \cos x dx &= \sin x + c, \\ \int x^\alpha dx &= \frac{x^{\alpha+1}}{\alpha+1} + c, \quad \text{con } \alpha \neq -1, \\ \int \frac{1}{x} dx &= \ln |x| + c, \\ \int \frac{1}{1+x^2} dx &= \arctan x + c, \\ \int \frac{1}{\sqrt{1-x^2}} dx &= \arcsin x + c.\end{aligned}$$

Si noti che la definizione di funzione primitivabile ha senso anche in alcuni casi in cui f non è necessariamente definita su un intervallo, e infatti le formule sopra riportate vanno interpretate sui domini naturali delle funzioni considerate. Per esempio,

$$\int \frac{1}{x} dx = \begin{cases} \ln x + c & \text{se } x \in]0, +\infty[, \\ \ln(-x) + c & \text{se } x \in]-\infty, 0[. \end{cases}$$

Esempio. Usando il Teorema Fondamentale troviamo:

$$\int_0^\pi \sin x dx = [-\cos x]_0^\pi = -\cos \pi + \cos 0 = 2.$$

Si noti che la presenza della costante arbitraria c a volte può portare a risultati apparentemente diversi. Sappiamo, ad esempio, che $\int \frac{1}{\sqrt{1-x^2}} dx = \arcsin x + c$, ma è facilmente verificabile che si ha anche

$$\int \frac{1}{\sqrt{1-x^2}} dx = -\arccos x + c.$$

Ciò è spiegato dal fatto che $\arcsin x = \frac{\pi}{2} - \arccos x$ per ogni $x \in [-1, 1]$, per cui la differenza tra $\arcsin$ e $-\arccos$ è costante. La stessa notazione c per la costante arbitraria nelle due formule a volte potrebbe essere fuorviante!

Dalle proprietà note delle derivate si possono facilmente dimostrare le due seguenti proposizioni.

Proposizione. Siano f e g primitivabili su $\mathcal{I}$, e siano F e G due rispettive primitive. Allora, $f + g$ è primitivabile su $\mathcal{I}$, e $F + G$ è una delle sue primitive; scriveremo brevemente⁷

$$\int (f + g) = \int f + \int g.$$

Proposizione. Sia f primitivabile su $\mathcal{I}$, e sia F una delle sue primitive. Se $\alpha \in \mathbb{R}$ è una qualsiasi costante, allora αf è primitivabile su $\mathcal{I}$, e αF è una delle sue primitive; scriveremo brevemente

$$\int (\alpha f) = \alpha \int f.$$

Come conseguenza di queste proposizioni abbiamo che l'insieme delle funzioni primitivabili su $\mathcal{I}$ è uno spazio vettoriale.

Concludiamo questa sezione evidenziando un'interessante classe di funzioni integrabili ma non primitivabili. Consideriamo una funzione $f : [a, b] \rightarrow \mathbb{R}$ tale che l'insieme

$$E = \{x \in [a, b] : f(x) \neq 0\}$$

sia finito o numerabile (ad esempio, una funzione che è zero ovunque tranne che in un punto, oppure la funzione di Dirichlet $\mathcal{D} : [a, b] \rightarrow \mathbb{R}$, definita da $\mathcal{D}(x) = 1$ se x è razionale e $\mathcal{D}(x) = 0$ se x è irrazionale).

Proviamo che tale funzione è integrabile, con $\int_a^b f = 0$. Assumiamo ad esempio che E sia infinito (il caso in cui E sia finito può essere trattato in modo analogo). Essendo numerabile, possiamo scrivere $E = \{e_n : n \in \mathbb{N}\}$. Fissato $\varepsilon > 0$, costruiamo un calibro δ su $[a, b]$ in questo modo: se $x \notin E$, poniamo $\delta(x) = 1$; se invece per un certo n è $x = e_n$, poniamo

$$\delta(e_n) = \frac{\varepsilon}{2^{n+3}|f(e_n)|}.$$

Sia ora $\mathring{\mathcal{P}} = \{(x_1, [a_0, a_1]), \dots, (x_m, [a_{m-1}, a_m])\}$ una P-partizione δ -fine di $[a, b]$. Allora la somma di Riemann associata può essere scritta in questo modo:

$$S(f, \mathring{\mathcal{P}}) = \sum_{\{1 \leq j \leq m : x_j \in E\}} f(x_j)(a_j - a_{j-1}).$$

Poiché $[a_{j-1}, a_j] \subseteq [x_j - \delta(x_j), x_j + \delta(x_j)]$, abbiamo che $a_j - a_{j-1} \leq 2\delta(x_j)$. Se x_j appartiene ad E , deve essere che $x_j = e_n$ per qualche $n \in \mathbb{N}$. Si noti però che a ciascuno di tali e_n possono corrispondere anche due degli x_j . Se denotiamo con N

⁷Qui e nel seguito utilizziamo in modo intuitivo le operazioni algebriche che coinvolgono gli insiemi. Per essere precisi, la somma di due insiemi A e B è definita come

$$A + B = \{a + b : a \in A, b \in B\}.$$

il più grande degli n per cui $x_j = e_n$ per un qualche $j \in \{1, \dots, m\}$, avremo quindi che

$$\begin{aligned} \left| \sum_{\{1 \leq j \leq m : x_j \in E\}} f(x_j)(a_j - a_{j-1}) \right| &\leq 2 \sum_{n=0}^N |f(e_n)| 2\delta(e_n) \\ &= 4 \sum_{n=0}^N \frac{\varepsilon}{2^{n+3}} = \frac{\varepsilon}{2} \sum_{n=0}^N \left(\frac{1}{2}\right)^n = \frac{\varepsilon}{2} \frac{1 - \left(\frac{1}{2}\right)^{N+1}}{1 - \frac{1}{2}} < \varepsilon. \end{aligned}$$

Ciò dimostra che f è integrabile su $[a, b]$ con $\int_a^b f = 0$.

Vediamo ora che, se E non è vuoto, allora f non è primitivabile su $[a, b]$. Infatti, se così fosse, la sua funzione integrale $\int_a^x f$ dovrebbe essere una delle sue primitive. Ma la procedura precedente mostra che $\int_a^x f = 0$, per ogni $x \in [a, b]$. Allora f dovrebbe essere identicamente uguale a zero, essendo la derivata di una funzione costante. Una contraddizione.

Primitivazione per parti e per sostituzione

Presentiamo ora due metodi frequentemente utilizzati per la ricerca delle primitive di alcune funzioni. Il primo è noto come metodo di “primitivazione per parti”.

Proposizione. Siano $F, G : \mathcal{I} \rightarrow \mathbb{R}$ due funzioni derivabili, e siano f, g le rispettive derivate. Si ha che fG è primitivabile su $\mathcal{I}$ se e solo se tale è Fg , nel qual caso si ottiene una primitiva di fG sottraendo da FG una primitiva di Fg ; scriveremo brevemente

$$\int fG = FG - \int Fg.$$

Dimostrazione. Poiché F e G sono derivabili, tale è anche FG , e abbiamo

$$(FG)' = fG + Fg.$$

Quindi $fG + Fg$ è primitivabile su $\mathcal{I}$ con primitiva FG , ed essendo $\int (fG + Fg) = \int fG + \int Fg$ si conclude facilmente. ■

Esempio. Vorremmo trovare una primitiva della funzione $h(x) = xe^x$. Definiamo le seguenti funzioni: $f(x) = e^x$, $G(x) = x$ e di conseguenza $F(x) = e^x$, $g(x) = 1$. Applicando la formula data dalla proposizione precedente, abbiamo

$$\int e^x x dx = e^x x - \int e^x dx = xe^x - e^x + c,$$

dove c rappresenta, come al solito, una costante arbitraria.

Come conseguenza immediata della formula di primitivazione per parti, abbiamo la corrispondente formula di “integrazione per parti”:

$$\int_a^b fG = F(b)G(b) - F(a)G(a) - \int_a^b Fg.$$

Esempi. Torniamo alla funzione $h(x) = xe^x$ dell'esempio precedente. Usando la formula di integrazione per parti, abbiamo che

$$\int_0^1 e^x x dx = e^1 \cdot 1 - e^0 \cdot 0 - \int_0^1 e^x dx = e - [e^x]_0^1 = e - (e^1 - e^0) = 1.$$

Notiamo che avremmo potuto ottenere lo stesso risultato utilizzando il Teorema Fondamentale, avendo già scoperto in precedenza che una primitiva di $h(x)$ è data da $H(x) = xe^x - e^x$. Infatti,

$$\int_0^1 e^x x dx = H(1) - H(0) = (e - e) - (0 - 1) = 1.$$

Vediamo qualche altro esempio. Sia $h(x) = \sin^2 x$. Con l'ovvia scelta delle funzioni f e G , troviamo che

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

da cui otteniamo

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

Consideriamo ora il caso della funzione $h(x) = \ln x$, con $x > 0$. per applicare la formula di primitivazione per parti, scegliamo le funzioni $f(x) = 1$, $G(x) = \ln x$. In questo modo troviamo

$$\int \ln x dx = x \ln x - \int x \frac{1}{x} dx = x \ln x - \int 1 dx = x \ln x - x + c.$$

Il secondo metodo che vogliamo studiare è quello della “primitivazione per sostituzione”.

Proposizione. Siano $\varphi : \mathcal{I} \rightarrow \mathbb{R}$ una funzione derivabile e $f : J \rightarrow \mathbb{R}$ una funzione primitivabile, dove J è un intervallo contenente $\varphi(\mathcal{I})$, con primitiva F . Allora la funzione $(f \circ \varphi)\varphi'$ è primitivabile su $\mathcal{I}$, e una sua primitiva è data da $F \circ \varphi$. Scriveremo brevemente

$$\int (f \circ \varphi)\varphi' = \left(\int f \right) \circ \varphi.$$

Dimostrazione. La funzione $F \circ \varphi$ è derivabile su $\mathcal{I}$ e

$$(F \circ \varphi)' = (F' \circ \varphi)\varphi' = (f \circ \varphi)\varphi'.$$

Ne consegue che $(f \circ \varphi)\varphi'$ è primitivabile su $\mathcal{I}$, con primitiva $F \circ \varphi$. ■

Come primo esempio, cerchiamo una primitiva della funzione $h(x) = xe^{x^2}$. Definendo $\varphi(x) = x^2$, $f(t) = \frac{1}{2}e^t$ (si consiglia di utilizzare lettere diverse per indicare le variabili di φ e f), abbiamo che $h = (f \circ \varphi)\varphi'$. Poiché si vede che una primitiva di f è $F(t) = \frac{1}{2}e^t$, una primitiva di h sarà $F \circ \varphi$, cioè

$$\int xe^{x^2} dx = F(\varphi(x)) + c = \frac{1}{2}e^{x^2} + c.$$

La formula di primitivazione per sostituzione viene spesso scritta nella forma

$$\int f(\varphi(x))\varphi'(x) dx = \int f(t) dt \Big|_{t=\varphi(x)},$$

dove, se F è una primitiva di f , il termine di destra deve essere letto come

$$\int f(t) dt \Big|_{t=\varphi(x)} = F(t) + c \Big|_{t=\varphi(x)} = F(\varphi(x)) + c.$$

Formalmente si ha un “cambio di variabile” $t = \varphi(x)$, e il simbolo dt va a sostituire $\varphi'(x) dx$ (la notazione Leibniz $\frac{dt}{dx} = \varphi'(x)$ risulta in questo caso un’utile regola mnemonica).

Esempio. Per trovare una primitiva della funzione $h(x) = \frac{\ln x}{x}$, possiamo scegliere $\varphi(x) = \ln x$ e applicare la formula

$$\int \frac{\ln x}{x} dx = \int t dt \Big|_{t=\ln x} = \frac{1}{2} t^2 + c \Big|_{t=\ln x} = \frac{1}{2} (\ln x)^2 + c.$$

In questo caso, scrivendo $t = \ln x$, il simbolo dt sostituisce $\frac{1}{x} dx$.

Come conseguenza delle formule precedenti, abbiamo la regola di “integrazione per sostituzione”:

$$\int_a^b f(\varphi(x))\varphi'(x) dx = \int_{\varphi(a)}^{\varphi(b)} f(t) dt .$$

Infatti, se F è una primitiva di f su $\varphi(\mathcal{I})$, applicando il Teorema Fondamentale, abbiamo

$$\int_a^b (f \circ \varphi)\varphi' = (F \circ \varphi)(b) - (F \circ \varphi)(a) = F(\varphi(b)) - F(\varphi(a)) = \int_{\varphi(a)}^{\varphi(b)} f .$$

Esempio. Tornando alla funzione $h(x) = xe^{x^2}$ considerata in precedenza, abbiamo

$$\int_0^2 xe^{x^2} dx = \int_0^4 \frac{1}{2}e^t dt = \frac{1}{2}[e^t]_0^4 = \frac{e^4 - 1}{2} .$$

Chiaramente lo stesso risultato è ottenibile direttamente dal Teorema Fondamentale, una volta noto che una primitiva di $h(x)$ è data da $H(x) = \frac{1}{2}e^{x^2}$. Infatti,

$$\int_0^2 xe^{x^2} dx = H(2) - H(0) = \frac{1}{2}e^4 - \frac{1}{2}e^0 = \frac{e^4 - 1}{2} .$$

Quando la funzione $\varphi : \mathcal{I} \rightarrow \varphi(\mathcal{I})$ è invertibile, si può anche scrivere

$$\int f(t) dt = \int f(\varphi(x))\varphi'(x) dx \Big|_{x=\varphi^{-1}(t)} ,$$

con la corrispondente formula dell'integrale:

$$\int_{\alpha}^{\beta} f(t) dt = \int_{\varphi^{-1}(\alpha)}^{\varphi^{-1}(\beta)} f(\varphi(x))\varphi'(x) dx .$$

Esempio. Se cerchiamo una primitiva di $f(t) = \sqrt{1-t^2}$, con $t \in]-1, 1[$, possiamo usare la funzione $\varphi :]0, \pi[\rightarrow]-1, 1[$ definita da $\varphi(x) = \cos x$, in modo da avere

$$f(\varphi(x))\varphi'(x) = \sqrt{1 - \cos^2 x} (-\sin x) = -\sin^2 x ,$$

essendo $\sin x > 0$ quando $x \in]0, \pi[$. Come abbiamo già dimostrato, quest'ultima funzione è primitivabile, quindi possiamo scrivere

$$\begin{aligned} \int \sqrt{1-t^2} dt &= - \int \sin^2 x dx \Big|_{x=\arccos t} \\ &= - \frac{1}{2}(x - \cos x \sin x) + c \Big|_{x=\arccos t} \\ &= - \frac{1}{2} \left(\arccos t - t\sqrt{1-t^2} \right) + c . \end{aligned}$$

Integrabilità su sottointervalli

In questa sezione vedremo che se una funzione è integrabile su un intervallo $I = [a, b]$, allora è anche integrabile su uno qualunque dei suoi sottointervalli. In particolare, è possibile considerare la sua funzione integrale. Vedremo inoltre che se una funzione è integrabile su due intervalli contigui, allora essa è integrabile anche sulla loro unione.

In seguito, sarà utile considerare le cosiddette “sotto-P-partizioni” dell’intervallo I . Una sotto-P-partizione è un insieme del tipo

$$\Xi = \{(\xi_j, [\alpha_j, \beta_j]) : j = 1, \dots, m\},$$

dove gli intervalli $[\alpha_j, \beta_j]$ non si sovrappongono ma non sono necessariamente contigui, e $\xi_j \in [\alpha_j, \beta_j]$ per ogni $j = 1, \dots, m$. Per una sotto-P-partizione Ξ , possiamo ancora definire la somma di Riemann ad essa associata:

$$S(f, \Xi) = \sum_{j=1}^m f(\xi_j)(\beta_j - \alpha_j).$$

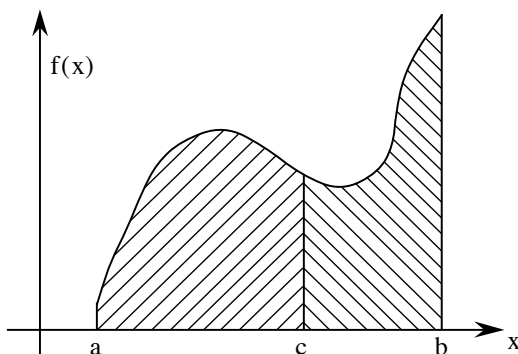
Inoltre, dato un calibro δ su I , diremo che la sotto-P-partizione Ξ è δ -fine se, per ogni j , si ha

$$\xi_j - \alpha_j \leq \delta(\xi_j) \quad \text{e} \quad \beta_j - \xi_j \leq \delta(\xi_j).$$

Enunciamo la proprietà di “additività su sottointervalli”.

Teorema. *Dati tre punti $a < c < b$, la funzione $f : [a, b] \rightarrow \mathbb{R}$ è integrabile su $[a, b]$ se e solo se è integrabile sia su $[a, c]$ che su $[c, b]$. In questo caso,*

$$\int_a^b f = \int_a^c f + \int_c^b f.$$



Esempio. Consideriamo la funzione $f : [0, 2] \rightarrow \mathbb{R}$ definita da

$$f(x) = \begin{cases} 3 & \text{se } x \in [0, 1], \\ 5 & \text{se } x \in]1, 2]. \end{cases}$$

Essendo f costante su $[0, 1]$ con valore 3, è ivi integrabile, e $\int_0^1 f = 3$. Inoltre, sull'intervallo $[1, 2]$ la funzione f differisce dalla costante 5 solo in un punto: abbiamo che la funzione $g(x) = f(x) - 5$ è uguale a zero eccetto per $x = 1$. Per quanto già dimostrato in precedenza, g è integrabile su $[1, 2]$ con integrale nullo e quindi, essendo $f(x) = g(x) + 5$, anche f è integrabile e

$$\int_1^2 f(x) dx = \int_1^2 g(x) dx + \int_1^2 5 dx = 0 + 5 = 5.$$

Pertanto,

$$\int_0^2 f(x) dx = \int_0^1 f(x) dx + \int_1^2 f(x) dx = 3 + 5 = 8.$$

È facile dedurre dal teorema sopra dimostrato che se una funzione è integrabile su un intervallo I , lo è pure su un qualsiasi sottointervallo di I . Inoltre, abbiamo quanto segue.

Corollario. Se $f : I \rightarrow \mathbb{R}$ è integrabile, scelti arbitrariamente tre punti u, v, w in I si ha che

$$\int_u^w f = \int_u^v f + \int_v^w f.$$

Dimostrazione. Il caso $u < v < w$ segue immediatamente dal teorema precedente. Gli altri casi possibili si dimostrano facilmente utilizzando le convenzioni sugli integrali a estremi scambiati o uguali. ■

Funzioni R -integrabili e funzioni continue

Introduciamo un'importante classe di funzioni integrabili. Come al solito, I è l'intervallo $[a, b]$.

Definizione. Diciamo che è una funzione integrabile $f : I \rightarrow \mathbb{R}$ è “ R -integrabile” (o “integrabile secondo Riemann”) se tra tutti i possibili calibri $\delta : I \rightarrow \mathbb{R}$ che verificano la definizione di integrabilità è sempre possibile sceglierne uno che sia costante su I .

Vogliamo ora stabilire l'integrabilità delle funzioni continue. In effetti, come vedremo nei due teoremi seguenti, esse sono sia R -integrabili che primitivabili.

Teorema. Ogni funzione continua $f : I \rightarrow \mathbb{R}$ è R -integrabile.

Teorema. Ogni funzione continua $f : I \rightarrow \mathbb{R}$ è primitivabile.

Si noti che non è sempre possibile trovare un'espressione elementare per la primitiva di una funzione continua. Ad esempio, la funzione $f(x) = \sin(x^2)$ è primitivabile, essendo continua, ma non esiste una formula elementare che definisca alcuna delle sue primitive. Per “formula elementare” si intende una formula analitica in cui compaiono solo polinomi, esponenziali, logaritmi, funzioni trigonometriche e le loro inverse.

Integrazione su intervalli non compatti

Iniziamo considerando una funzione $f : [a, b[\rightarrow \mathbb{R}$, dove $b \leq +\infty$. Supponiamo che f sia integrabile su ogni intervallo compatto del tipo $[a, c]$, con $c \in]a, b[$. Questo avviene, per esempio, quando f è continua in $[a, b[$.

Definizione. Diciamo che una funzione $f : [a, b[\rightarrow \mathbb{R}$ è “integrabile” se f è integrabile su $[a, c]$ per ogni $c \in]a, b[$, e il limite

$$\lim_{c \rightarrow b^-} \int_a^c f$$

esiste ed è finito. In tal caso, questo limite si chiama “integrale” di f su $[a, b[$ e si indica con $\int_a^b f$, o con $\int_a^b f(x) dx$.

In particolare, se $b = +\infty$, scriveremo: $\int_a^{+\infty} f$, o $\int_a^{+\infty} f(x) dx$.

Esempi. Sia $a > 0$; si vede facilmente che la funzione $f : [a, +\infty[\rightarrow \mathbb{R}$ definita da $f(x) = x^{-\alpha}$ è integrabile se e solo se $\alpha > 1$ (Figura 1 a sinistra), nel qual caso abbiamo

$$\int_a^{+\infty} \frac{dx}{x^\alpha} = \frac{a^{1-\alpha}}{\alpha-1}.$$

Consideriamo ora il caso $a < b < +\infty$. Si può verificare che la funzione $f : [a, b[\rightarrow \mathbb{R}$ definita da $f(x) = (b-x)^{-\beta}$ è integrabile se e solo se $\beta < 1$ (Figura 1 a destra), nel qual caso abbiamo

$$\int_a^b \frac{dx}{(b-x)^\beta} = \frac{(b-a)^{1-\beta}}{1-\beta}.$$

Si dice anche che *l'integrale converge* se la funzione f è integrabile su $[a, b[$, ossia quando il limite $\lim_{c \rightarrow b^-} \int_a^c f$ esiste ed è finito. Se il limite non esiste, si dice che *l'integrale è indeterminato*. Se esiste e vale $+\infty$ o $-\infty$, si dice che *l'integrale diverge* rispettivamente a $+\infty$ o a $-\infty$.

È chiaro che la convergenza dell'integrale dipende esclusivamente dal comportamento della funzione “vicino” al punto b . In altre parole, modificando la funzione al di fuori di un intorno di b , la convergenza dell'integrale non viene compromessa.

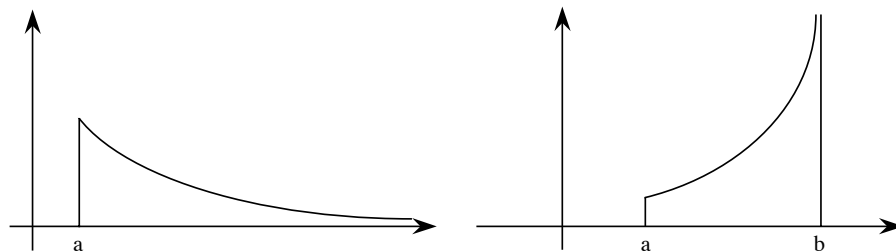


Figure 1: Due possibili situazioni per l'integrale generalizzato.

Vale il seguente **Criterio del Confronto**.

Teorema. Sia $f : [a, b[\rightarrow \mathbb{R}$ una funzione integrabile su $[a, c]$, per ogni $c \in]a, b[$. Se esiste una funzione integrabile $g : [a, b[\rightarrow \mathbb{R}$ tale che, per ogni $x \in [a, b[$,

$$|f(x)| \leq g(x),$$

allora anche f è integrabile su $[a, b[$.

Si noti che sarebbe stato sufficiente supporre la disuguaglianza $|f(x)| \leq g(x)$ su $[\bar{c}, b[$, per un certo $\bar{c} \in]a, b[$. Come conseguenza immediata, abbiamo quanto segue.

Corollario. Sia $f : [a, b[\rightarrow \mathbb{R}$ una funzione integrabile su $[a, c]$, per ogni $c \in]a, b[$. Se $|f|$ è integrabile su $[a, b[$, allora anche f lo è, e si ha che

$$\left| \int_a^b f \right| \leq \int_a^b |f|.$$

Enunciamo ora un corollario del Criterio del Confronto che è spesso utilizzato nella pratica.

Corollario. Siano $f, g : [a, b[\rightarrow \mathbb{R}$ due funzioni a valori positivi, integrabili su $[a, c]$ per ogni $c \in]a, b[$. Supponiamo che esista il limite

$$L = \lim_{x \rightarrow b^-} \frac{f(x)}{g(x)}.$$

Allora valgono le seguenti conclusioni:

- a) se $L \in]0, +\infty[$, allora f è integrabile su $[a, b[$ se e solo se lo è g ;
- b) se $L = 0$ e g è integrabile su $[a, b[$, allora lo è anche f ;
- c) se $L = +\infty$ e g non è integrabile su $[a, b[$, allora neppure f lo è.

Dimostrazione. Caso a). Se $L \in]0, +\infty[$, esiste un $\bar{c} \in]a, b[$ tale che

$$x \in [\bar{c}, b[\Rightarrow \frac{L}{2} \leq \frac{f(x)}{g(x)} \leq \frac{3L}{2},$$

cioè,

$$x \in [\bar{c}, b[\Rightarrow f(x) \leq \frac{3L}{2}g(x) \quad \text{e} \quad g(x) \leq \frac{2}{L}f(x).$$

La conclusione segue allora dal Criterio del Confronto.

Caso b). Se $L = 0$, esiste un $\bar{c} \in]a, b[$ tale che, se $x \in [\bar{c}, b[$, allora $f(x) \leq g(x)$, e si applica il Criterio del Confronto.

Caso c). Se $L = +\infty$ ci riconduciamo al caso b) scambiando i ruoli di f e g . ■

Esempio. Consideriamo la funzione $f : [0, +\infty[\rightarrow \mathbb{R}$ definita da

$$f(x) = e^{1/(x^2+1)} - 1.$$

Come funzione di confronto, prendiamo

$$g(x) = \frac{1}{x^2 + 1}.$$

Essa è integrabile su $[0, +\infty[$, con

$$\int_0^{+\infty} \frac{1}{x^2 + 1} dx = \lim_{c \rightarrow +\infty} [\arctan x]_0^c = \frac{\pi}{2}.$$

Dato che

$$\lim_{x \rightarrow +\infty} \frac{f(x)}{g(x)} = \lim_{t \rightarrow 0^+} \frac{e^t - 1}{t} = 1,$$

concludiamo che anche f è integrabile su $[0, +\infty[$.

Consideriamo ora il caso di una funzione $f :]a, b] \rightarrow \mathbb{R}$, con $a \geq -\infty$. Ecco una definizione analoga del suo integrale.

Definizione. Diciamo che una funzione $f :]a, b] \rightarrow \mathbb{R}$ è “integrabile” se f è integrabile su $[c, b]$ per ogni $c \in]a, b[$, e il limite

$$\lim_{c \rightarrow a^+} \int_c^b f$$

esiste ed è finito. In tal caso, questo limite si chiama “integrale” di f su $]a, b]$ e si indica con $\int_a^b f$, o con $\int_a^b f(x) dx$.

Data la funzione $f :]a, b] \rightarrow \mathbb{R}$, è possibile considerare la funzione $g : [a', b'[\rightarrow \mathbb{R}$, con $a' = -b$ e $b' = -a$, definita da $g(x) = f(-x)$. È facile vedere che f è integrabile su $]a, b]$ se e solo se g è integrabile su $[a', b'[$. Ci riconduciamo in questo modo alla situazione precedente.

Definiamo anche l'integrale di una funzione $f :]a, b[\rightarrow \mathbb{R}$, con $-\infty \leq a < b \leq +\infty$, nel modo seguente.

Definizione. Diciamo che $f :]a, b[\rightarrow \mathbb{R}$ è “integrabile” se, una volta fissato un punto $p \in]a, b[$, la funzione f è integrabile su $[p, b[$ e su $]a, p]$. In tal caso, si pone

$$\int_a^b f = \int_a^p f + \int_p^b f.$$

È facile verificare che la definizione data non dipende dalla scelta di $p \in]a, b[$.

Esempi. Se $a, b \in \mathbb{R}$, si può verificare che la funzione

$$f(x) = \frac{1}{[(x-a)(b-x)]^\beta}$$

è integrabile su $]a, b[$ se e solo se $\beta < 1$. In questo caso è possibile scegliere, ad esempio, $p = (a+b)/2$.

Un altro caso si presenta quando $a = -\infty$ e $b = +\infty$. Per esempio, si verifica facilmente che la funzione $f(x) = (x^2+1)^{-1}$ è integrabile su $]-\infty, +\infty[$. Prendendo ad esempio $p = 0$, abbiamo:

$$\int_{-\infty}^{+\infty} \frac{1}{x^2+1} dx = \int_{-\infty}^0 \frac{1}{x^2+1} dx + \int_0^{+\infty} \frac{1}{x^2+1} dx = \frac{\pi}{2} + \frac{\pi}{2} = \pi.$$

Un'altra situazione che si potrebbe presentare nelle applicazioni è il caso di una funzione il cui dominio sia un intervallo privato di un punto interno.

Definizione. Dati $a < q < b$, diciamo che una funzione $f : [a, b] \setminus \{q\} \rightarrow \mathbb{R}$ è “integrabile” se f è integrabile sia su $[a, q[$ che su $]q, b]$. In tal caso, poniamo

$$\int_a^b f = \int_a^q f + \int_q^b f.$$

Ad esempio, se $a < 0 < b$, la funzione $f(x) = \sqrt{|x|}/x$ è integrabile su $[a, b] \setminus \{0\}$ e

$$\int_a^b \frac{\sqrt{|x|}}{x} dx = \int_a^0 \frac{-1}{\sqrt{-x}} dx + \int_0^b \frac{1}{\sqrt{x}} dx = 2\sqrt{b} - 2\sqrt{-a}.$$

D'altro canto, la funzione $f(x) = 1/x$ non è integrabile su $[-1, 1] \setminus \{0\}$, anche se il fatto che f sia dispari potrebbe portarci a supporre che il suo integrale sia pari a zero. Tuttavia, così facendo si perderebbero alcune proprietà importanti dell'integrale, come ad esempio l'additività sui sotto-intervalli.

La funzione esponenziale complessa

Abbiamo visto che, per ogni $x \in \mathbb{R}$, ponendo

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

si ha che

$$\lim_{n \rightarrow +\infty} p_n(x) = e^x.$$

Definiamo ora le funzioni $q_n : \mathbb{R} \rightarrow \mathbb{C}$ in questo modo:

$$q_n(y) = p_n(iy) = 1 + iy - \frac{y^2}{2!} - i \frac{y^3}{3!} + \frac{y^4}{4!} + i \frac{y^5}{5!} - \frac{y^6}{6!} - i \frac{y^7}{7!} + \cdots + i^n \frac{y^n}{n!}.$$

Abbiamo quindi che

$$q_n(y) = q_n^{(1)}(y) + i q_n^{(2)}(y),$$

dove

$$q_n^{(1)}(y) = 1 - \frac{y^2}{2!} + \frac{y^4}{4!} - \frac{y^6}{6!} + \cdots + (-1)^m \frac{y^{2m}}{(2m)!}$$

se $n = 2m$ o $n = 2m + 1$, mentre

$$q_n^{(2)}(y) = y - \frac{y^3}{3!} + \frac{y^5}{5!} - \frac{y^7}{7!} + \cdots + (-1)^m \frac{y^{2m+1}}{(2m+1)!}$$

se $n = 2m + 1$ o $n = 2m + 2$. Essendo

$$\lim_{n \rightarrow +\infty} q_n^{(1)}(y) = \cos y, \quad \lim_{n \rightarrow +\infty} q_n^{(2)}(y) = \sin y,$$

concludiamo quindi che

$$\lim_{n \rightarrow \infty} q_n(y) = \cos y + i \sin y,$$

ossia

$$\lim_{n \rightarrow \infty} p_n(iy) = \cos y + i \sin y.$$

Per analogia con la formula trovata in precedenza, si decide di denotare questa espressione con il simbolo e^{iy} . Ecco allora che, per ogni numero complesso $z = x + iy$, risulterà naturale porre

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

Resta così definita la funzione *esponenziale complessa* $\exp : \mathbb{C} \rightarrow \mathbb{C}$, ponendo

$$\exp(x + iy) = e^{x+iy}.$$

Si noti che

$$|e^{x+iy}| = |e^x| |e^{iy}| = e^x,$$

per cui $\exp(z) \neq 0$, per ogni $z \in \mathbb{C}$.

Ricordando la funzione circolare $h_T : \mathbb{R} \rightarrow S^1$, abbiamo che

$$h_{2\pi}(t) = \cos t + i \sin t = e^{it}.$$

Sappiamo che $\tilde{h}_{2\pi} : [0, 2\pi[\rightarrow S^1$, la restrizione di $h_{2\pi}$ all'intervallo $[0, 2\pi[$, è biiettiva. Ora, siccome ogni numero complesso $z \neq 0$ si può scrivere come

$$z = |z| \frac{z}{|z|}, \quad \text{con } |z| > 0 \text{ e } \frac{z}{|z|} \in S^1,$$

vediamo che esso determina univocamente due numeri $\rho = |z|$ e $\theta \in [0, 2\pi[$ tale che

$$\frac{z}{|z|} = h_{2\pi}(\theta) = e^{i\theta}.$$

Ecco quindi che ogni numero complesso z si potrà scrivere come

$$z = \rho e^{i\theta},$$

dove $\rho \geq 0$ è il *modulo* e $\theta \in [0, 2\pi[$ è l'*argomento* di z (se però $z = 0$, l'argomento non risulta univocamente definito).

Con le note proprietà di omomorfismo viste durante il corso, si può verificare che, per ogni z_1, z_2 in $\mathbb{C}$, si ha che

$$\exp(z_1 + z_2) = \exp(z_1) \exp(z_2).$$

Infatti, se $z_1 = x_1 + iy_1$ e $z_2 = x_2 + iy_2$, si ha

$$\begin{aligned} e^{z_1+z_2} &= e^{(x_1+x_2)+i(y_1+y_2)} = e^{x_1+x_2} e^{i(y_1+y_2)} \\ &= e^{x_1+x_2} h_{2\pi}(y_1 + y_2) = e^{x_1} e^{x_2} h_{2\pi}(y_1) h_{2\pi}(y_2) \\ &= e^{x_1} h_{2\pi}(y_1) e^{x_2} h_{2\pi}(y_2) = e^{x_1} e^{iy_1} e^{x_2} e^{iy_2} = e^{x_1+iy_1} e^{x_2+iy_2} = e^{z_1} e^{z_2}. \end{aligned}$$

Possiamo pertanto affermare che la funzione esponenziale complessa è un omomorfismo tra $(\mathbb{C}, +)$ e $(\mathbb{C} \setminus \{0\}, \cdot)$.

Osserviamo che, scrivendo

$$z_1 = \rho_1 e^{i\theta_1}, \quad z_2 = \rho_2 e^{i\theta_2},$$

si ha che

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

Si ha quindi che il modulo del prodotto è uguale al prodotto dei due moduli (cosa a noi già nota), mentre l'argomento del prodotto è uguale alla somma dei due argomenti, se questa somma non supera 2π ; altrimenti, a questa somma bisognerà sottrarre 2π . Per evitare di dover considerare casi diversi, come in questo caso, talvolta si considera l'argomento di un numero complesso definito a meno di un multiplo intero di 2π .

Si noti che, scrivendo per $t \in \mathbb{R}$

$$e^{it} = \cos t + i \sin t, \quad e^{-it} = \cos t - i \sin t,$$

si trova che

$$\cos t = \frac{e^{it} + e^{-it}}{2}, \quad \sin t = \frac{e^{it} - e^{-it}}{2i}.$$

Queste formule possono essere usate per estendere le funzioni $\cos$ e $\sin$ al campo complesso, assumendole valide anche per $t \in \mathbb{C}$. Anche le funzioni iperboliche possono essere estese a $\mathbb{C}$ con le loro formule

$$\cosh z = \frac{e^z + e^{-z}}{2}, \quad \sinh z = \frac{e^z - e^{-z}}{2}.$$

Si avrà quindi che

$$\cos t = \cosh(it), \quad \sin t = -i \sinh(it).$$

In questo contesto risulta ben chiaro il legame di parentela che c'è tra queste funzioni.