Analisi Matematica I

Fabio Fagnani, Gabriele Grillo
Dipartimento di Matematica
Politecnico di Torino

Queste dispense contengono il materiale delle lezioni del corso di Analisi Matematica I rivolto agli studenti del primo anno di Ingegneria della III facoltà del Politecnico di Torino.

Sono pensate, come il corso del resto, per studenti in possesso di una cultura matematica quale è quella fornita mediamente dalle scuole medie superiori. Si presuppone in particolare la conoscenza dei seguenti argomenti: polinomi, esponenziali, logaritmi, elementi di trigonometria, geometria analitica (rette, coniche), equazioni e disequazioni algebriche e trascendenti. Tali elementi non saranno trattati in queste note. Si presuppone anche che lo studente conosca le basi del linguaggio insiemistico: alcuni richiami sono comunque fatti nell'Appendice A1 che consigliamo di leggere come prima cosa.

Queste dispense contengono anche un certo numero di esercizi di diverse difficoltà. Gli esercizi senza particolari contrassegni sono da considerarsi di difficoltà normale. Saperli risolvere è condizione necessaria per poter continuare a seguire il corso con profitto; rappresentano la difficoltà richiesta per superare l'esame del corso. Gli esercizi contrassegnati con * sono di difficoltà superiore. La capacità dello studente di risolverli significa un ottimo livello di padronanza dei concetti del corso. Degli esercizi con risposta numerica o logica è fornito, alla fine di ciascuno, il risultato.

Capitolo 1

Insiemi di numeri

1.1 Naturali, interi, razionali

I numeri sono così pervasivi del nostro mondo da far si che se ne cominci a fare conoscenza nei primi anni di vita e se ne continui a fare un uso via via più approfondito nella vita quotidiana e nel percorso scolastico fino alle scuole superiori. Resta tuttavia indispensabile per gli studenti che hanno deciso di intraprendere studi universitari di carattere scientifico o tecnico ritornarci di nuovo sopra nei corsi matematici di base. L'esigenza nasce dalla necessità di fare alcuni chiarimenti su alcuni aspetti delicati e profondi dei numeri che giocano poi un ruolo fondamentale in tutta quanta la matematica. Non si tratta tanto di questioni fondazionali sul concetto di numero, che in questa sede non verranno affrontate, quanto di questioni concrete da tenere ben presenti da chiunque voglia utilizzare lo strumento matematico con perizia e sicurezza.

I primi numeri che si incontrano sono gli interi positivi, detti anche *numeri naturali*: $1, 2, 3, \ldots$. L'insieme dei numeri naturali si indica con il simbolo $\mathbb{N}$. Sono i numeri che servono a contare e che hanno fatto la prima comparsa nelle società umane svariate migliaia di anni fa. Per fare misure di quantità fisiche come lunghezze, aree, tempi, temperature, ecc., è tuttavia necessario poter disporre di sottoparti dell'unità e considerare quindi numeri frazionari m/n dove $m, n \in \mathbb{N}$ con $n \neq 0$. E' poi conveniente anche introdurre i numeri con il segno per poter trattare grandezze negative come possono essere la temperatura, la velocità e molte altre grandezze fisiche. Si ottengono così i seguenti insiemi numerici:

$$\mathbb{Z} = \{0, \pm 1, \pm 2, \ldots\} \qquad \text{numeri interi relativi}\,,$$

$$\mathbb{Q} = \{\tfrac{m}{n} \mid m, n \in \mathbb{Z},\ n \neq 0\} \qquad \text{numeri razionali}\,.$$

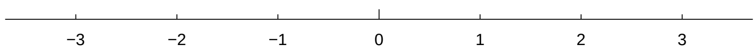
Si hanno le evidenti inclusioni $\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q}$. Talvolta è anche utile considerare gli insiemi

$$\mathbb{Z}^+ = \{m \in \mathbb{Z} \mid m \geq 0\} \qquad \text{numeri interi non negativi}\,,$$

$$\mathbb{Q}^+ = \{q \in \mathbb{Q} \mid q \geq 0\} \qquad \text{numeri razionali non negativi}\,.$$

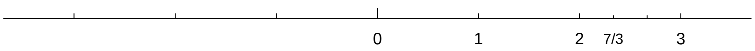
1.2 Perchè servono altri numeri?

I numeri fin qui introdotti sono suscettibili di una semplice interpretazione geometrica. Su di una retta r, fissiamo un punto che indicheremo con 0 ed un altro punto, a destra di 0, denominato 1. Usando come unità di lunghezza quella del segmento da 0 a 1 ed i due versi possibili (a destra e a sinistra di 0), si possono così facilmente rappresentare, sulla retta r, i numeri interi relativi come mostrato nella seguente figura.

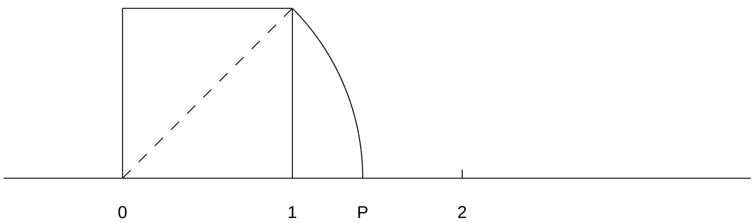


Dato invece un razionale m/n, esso può sempre venire espresso, tramite una semplice divisione, come $m/n = q + m'/n$ dove $q \in \mathbb{Z}$ e $0 \leq m' < n$. Il modo di rappresentare m/n sulla retta r diviene ora operativamente chiaro: scegliamo il segmento da q a $q+1$, dividiamolo in n parti uguali e consideriamo il punto ottenuto partendo da q, spostandosi di m' segmentini di lunghezza $1/n$ nella direzione di $q+1$. Otterremo ovviamente in questo modo un numero compreso tra q e $q+1$. Ad esempio $7/3 = 2+1/3$ è rappresentato nella figura sotto:

Ogni numero razionale è così univocamente rappresentato da un punto sulla retta. Sarà vera la cosa contraria? In realtà il problema non è ben posto in quanto non abbiamo dato una definizione esatta di retta. Affidiamoci tuttavia alla nostra intuizione di retta come un continuo di punti allineati



intendendo per continuo il fatto che non ci siano 'buchi' nella retta. Costruiamo ora sul segmento da 0 a 1 un quadrato; poi, con un compasso, facciamo centro in 0, apriamo con raggio determinato dal vertice del quadrato opposto a 0 e tracciamo un arco di circonferenza fino ad incontrare la retta r in un certo punto P. Che numero rappresenta P?



Considerato che il numero associato ad un punto della retta può essere pensato come la lunghezza, con eventuale segno, del segmento dal punto all'origine 0, risulta chiaro che P deve rappresentare il numero $\sqrt{2}$. Ma chi è $\sqrt{2}$? Può essere rappresentato come frazione? La risposta è nota da almeno due millenni, ma vale la pena ricordarla nella proposizione sotto dove ne presentiamo anche la classica, elegante dimostrazione.

Proposizione 1.1 *$\sqrt{2}$ non è un numero razionale.*

Dimostrazione Supponiamo, per assurdo, che lo sia, cioè che esistano naturali m, n con $n \neq 0$ tali che $\sqrt{2} = m/n$. Ovviamente si può ipotizzare che m e n siano primi tra loro. Elevando al quadrato si ottiene $2 = m^2/n^2$ o anche

$$2n^2 = m^2. \tag{1.1}$$

Questo significa che m^2 deve essere divisibile per 2, cioè deve essere un numero pari; questo implica (pensate perchè) che m è un numero pari. Quindi si può scrivere $m = 2q$ per qualche naturale q. Sostituendo in (1.1) si ottiene così

$$2n^2 = 4q^2 \ \Rightarrow \ n^2 = 2q^2. \tag{1.2}$$

Quest'ultima formula implica però che n^2, e di conseguenza n, è un numero pari. Quindi sia m che n sono numeri pari e questo è assurdo per l'ipotesi fatta che fossero primi tra loro. Ne consegue che $\sqrt{2}$ non può essere razionale. ■

Esercizio 1.1 Si mostri che $\sqrt{3}$ non è un numero razionale.

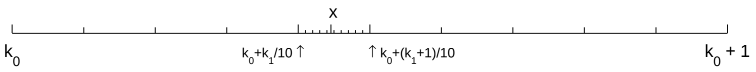
Esercizio 1.2 * Si mostri che $\sqrt{x}$ non è un numero razionale se x non è un quadrato perfetto (cioè se x non è del tipo $x = n^2$ per qualche $n \in \mathbb{N}$).

Esercizio 1.3 * Si mostri che $\sqrt[3]{2}$ non è un numero razionale.

1.3 I numeri reali

I punti della retta r sono quindi 'di più' dei numeri razionali. Che tipo di numeri servono per poter rappresentare tutti i punti della retta? Sono i numeri reali che introdurremo attraverso le rappresentazioni decimali. Fissiamo prima alcune notazioni. Sulla retta r vi è un ordinamento naturale: se a e b sono due punti di r, scriveremo che $a < b$ se a sta a sinistra di b e, $a \leq b$ se $a < b$ o se $a = b$. Dati a e b di r con $a < b$, indicheremo con $[a, b]$ il segmento dei punti tra a e b estremi inclusi, mentre con il simbolo $]a, b[$ indicheremo lo stesso segmento senza estremi. Il segmento con uno soltanto dei due estremi verrà indicato, rispettivamente, con $[a, b[$ se contiene a, con $]a, b]$ se contiene b. I sottoinsiemi della retta del tipo $[a, b]$, $]a, b[$, $[a, b[$, $]a, b]$ verranno anche detti *intervalli*.

Consideriamo ora un punto $x > 0$ di r. Chiaramente ci sarà un intero $k_0 \geq 0$ tale che
$$k_0 \leq x < k_0 + 1.$$
Dividiamo ora l'intervallo $[k_0, k_0 + 1[$ in dieci parti eguali
$$\left[k_0, k_0 + \frac{1}{10}\right[, \quad \left[k_0 + \frac{1}{10}, k_0 + \frac{2}{10}\right[, \quad \dots, \quad \left[k_0 + \frac{9}{10}, k_0 + 1\right[.$$
x dovrà stare in uno di questi. Supponiamo che
$$k_0 + \frac{k_1}{10} \leq x < k_0 + \frac{k_1 + 1}{10}$$
per un qualche $k_1 = 0, \dots, 9$.



Andiamo avanti così dividendo a sua volta l'intervallino $[k_0 + \frac{k_1}{10}, k_0 + \frac{k_1+1}{10}[$ in dieci parti (che misurano quindi un centesimo di quello iniziale $[k_0, k_0+1[$) individuando quello in cui sta x:

$$k_0 + \frac{k_1}{10} + \frac{k_2}{100} \leq x < k_0 + \frac{k_1}{10} + \frac{k_2+1}{100}$$

per un qualche $k_2 = 0, \dots, 9$. Continuando così si determina una sequenza infinita di numeri naturali $k_0, k_1, k_2, \dots$ con k_0 qualunque e tutti gli altri compresi tra 0 e 9 che risultano collegati al punto x nel modo seguente

$$k_0 + \frac{k_1}{10} + \frac{k_2}{100} + \cdots \frac{k_n}{10^n} \leq x < k_0 + \frac{k_1}{10} + \frac{k_2}{100} + \cdots \frac{k_n+1}{10^n}\,.$$

Introduciamo la notazione compatta

$$x_n = k_0 + \frac{k_1}{10} + \frac{k_2}{100} + \cdots + \frac{k_n}{10^n} = \sum_{i=0}^{n} \frac{k_i}{10^i}\,.$$

Allora la diseguaglianza precedente può essere riscritta come

$$x_n \leq x < x_n + \frac{1}{10^n}\,.$$

Si noti che x_n e $x_n + 1/10^n$ sono entrambi numeri razionali che distano tra di loro $1/10^n$. Poichè x sta in mezzo, vuol dire che entrambi hanno distanza da x non superiore ad $1/10^n$: più precisamente x_n *approssima per difetto* x a meno di $1/10^n$ mentre $x_n + 1/10^n$ *approssima per eccesso* sempre a meno di $1/10^n$. All'aumentare di n essi 'si avvicinano' quanto vogliamo al vero punto x. Tale punto x non coincide, in generale, con nessuno dei punti x_n (a meno che la sequenza dei k_n sia fatta da un certo punto in poi da tutti zeri); diremo invece che x è rappresentato dalla sequenza infinita, detta *allineamento decimale*,

$$k_0, k_1 k_2 k_3 \cdots k_n \cdots$$

Diremo anche che quella sopra è la *rappresentazione decimale* di x. Nel caso invece in cui $x < 0$, si considera il suo simmetrico $-x$ rispetto al punto 0. $-x$

ha una rappresentazione decimale $k_0, k_1k_2k_3\cdots k_n\cdots$. La rappresentazione decimale di x è allora convenzionalmente indicata come $-k_0, k_1k_2k_3\cdots k_n\cdots$. Si noti che valgono le seguenti diseguaglianze

$$-\left[k_0+\frac{k_1}{10}+\frac{k_2}{100}+\cdots\frac{k_n+1}{10^n}\right]\leq x<-\left[k_0+\frac{k_1}{10}+\frac{k_2}{100}+\cdots\frac{k_n}{10^n}\right].$$

Useremo, analogamente a prima, la notazione

$$x_n=-\sum_{i=0}^{n}\frac{k_i}{10^i}.$$

Per indicare la ripetizione infinita di una cifra o di un gruppo di cifre in una rappresentazione decimale useremo una barretta sopra: Ad esempio $0,\overline{8}=0,88888\cdots$, $0,1\overline{23}=0,1232323\cdots$. Nel caso in cui a ripetersi sia la cifra 0, essa verrà in generale omessa; scriveremo, ad esempio, $0,2317\overline{0}=0,2317$: questi ultimi verranno detti *allineamenti decimali finiti.*

Ad ogni punto della retta abbiamo così univocamente associato un allineamento decimale. Resta ora da vedere se questo ragionamento si può invertire, cioè se ad ogni allineamento decimale $\pm k_0, k_1k_2k_3\cdots k_n\cdots$ corrisponda un punto determinato della retta: questo è in un certo senso il punto più delicato di tutta la storia. Per simmetria basta farlo per gli allineamenti decimali positivi. Se un tale x esiste, deve stare in tutti gli intervalli

$$[k_0,k_0+1]\,,\quad\left[k_0+\frac{k_1}{10},k_0+\frac{k_1+1}{10}\right],\quad\left[k_0+\frac{k_1}{10}+\frac{k_2}{100},k_0+\frac{k_1}{10}+\frac{k_2+1}{100}\right],\ldots$$

Poichè questi intervalli sono l'uno incapsulato nell'altro e la loro larghezza diventa piccola quanto vogliamo sembra piuttosto evidente che esista un punto della retta, ed uno soltanto, che sta in tutti essi. Questa evidenza non è dimostrabile, ma va ipotizzata come proprietà di continuità della retta a cui facevamo cenno prima. Quindi supponiamo che esista un punto x della retta r che sta esattamente in tutti gli intervallini sopra. L'unica cosa che resta da verificare è se effettivamente la rappresentazione decimale di x, come definita sopra, sia proprio data da $k_0, k_1k_2k_3\cdots k_n\cdots$. Per come si è introdotta la rappresentazione decimale, è chiaro che questo risulta vero se accade che x sta dentro tutti gli intervallini semiaperti

$$[k_0,k_0+1[\,,\quad\left[k_0+\frac{k_1}{10},k_0+\frac{k_1+1}{10}\right[,\quad\left[k_0+\frac{k_1}{10}+\frac{k_2}{100},k_0+\frac{k_1}{10}+\frac{k_2+1}{100}\right[,\ldots$$

Sfortunatamente questo potrebbe non essere vero come mostriamo ora con un semplice esempio.

Esempio 1 Consideriamo l'allineamento decimale $0,\overline{9}$. Il punto x ad esso associato deve stare in tutti gli intervalli

$$[0,1], \quad \left[\frac{9}{10},1\right], \quad \left[\frac{9}{10}+\frac{9}{100},1\right], \ldots$$

ed è chiaro che l'unico punto con queste proprietà è 1. Tuttavia, la rappresentazione decimale di 1, come definita inizialmente, non è data da $0,\overline{9}$, ma da $1 = 1,\overline{0}$.

L'esempio sopra mostra come ci siano 'più' allineamenti decimali che punti della retta. $1 = 1,\overline{0}$ e $0,\overline{9}$ rappresentano lo stesso punto della retta, 1 appunto. Questa ambiguità capita solo e soltanto per gli allineamenti decimali che terminano con un numero infinito di 0 o di 9. Non è difficile rendersi conto che i due allineamenti decimali

$$k_0, k_1k_2k_3\cdots k_n\overline{9} \qquad k_0, k_1k_2k_3\cdots(k_n+1)$$

rappresentano sempre lo stesso punto della retta. E si può dimostrare che questi sono gli unici casi in cui si possono avere delle ambiguità. Se ci restringiamo agli allineamenti decimali che non terminano con una coda infinita di 9, allora vi è una perfetta corrispondenza biunivoca tra punti della retta e allineamenti decimali. Questo sottoinsieme di allineamenti decimali sarà detto *l'insieme dei numeri reali* e sarà indicato con il simbolo $\mathbb{R}$. Terremo sempre presente la corrispondenza con i punti della retta e parleremo infatti spesso di *retta reale*. Con $\mathbb{R}^+$ indicheremo invece la semiretta destra dei numeri reali non negativi.

Osservazione: Per come è stato costruito, $\mathbb{R}$ contiene i numeri razionali e dunque anche gli altri insiemi numerici fin qui introdotti. A che tipo di allineamenti decimali corrispondono i numeri razionali? Si noti intanto che gli allineamenti decimali finiti sono sicuramente razionali: in effetti se $x = \pm k_0, k_1k_2\cdots k_m$, si ha che

$$x = \pm\sum_{i=0}^{m}\frac{k_i}{10^i} = \pm\frac{\sum_{i=0}^{m}10^{m-i}k_i}{10^m}$$

e quindi x è un numero razionale esprimibile per mezzo di una frazione con denominatore una potenza di 10. E' facile vedere che tutti i numeri razionali di questo tipo hanno effettivamente un allineamento decimale finito. Che si può dire degli altri razionali? Se anzichè allineamenti finiti, aventi cioè una coda infinita di 0, consideriamo allineamenti decimali che hanno una coda costituita dalla ripetizione infinita di un gruppo di cifre, otteniamo ancora numeri razionali. In effetti si ha, ad esempio,

$$0,\overline{3} = \frac{3}{9} \qquad 1,\overline{41} = \frac{140}{99}\,.$$

Questo fatto non dovrebbe esservi nuovo e dovreste sapere come operativamente passare, in generale, dall'allineamento decimale con coda periodica alla corrispondente frazione; ci torneremo comunque più avanti. Fatto interessante (che noi non dimostreremo) è che la cosa si inverte: la rappresentazione decimale di un qualunque numero razionale ha sempre una coda periodica costituita cioè dalla ripetizione infinita di un certo gruppo di cifre.

1.4 Proprietà algebriche e di ordinamento

Ci aspettiamo di poter definire la somma ed il prodotto di numeri reali in modo che valgano le usuali proprietà algebriche come per i numeri razionali. Questo si può in effetti fare anche se c'è qualche dettaglio tecnico da superare. Come si fa ad esempio a sommare due allineamenti decimali $x = k_0, k_1k_2\cdots$ e $y = h_0, h_1h_2\cdots$? Lo sappiamo sicuramente fare se sono entrambi finiti (ce lo hanno insegnato alle scuole elementari), in tal caso in effetti possono anche essere entrambi pensati come numeri razionali. Con qualche accorgimento l'algoritmo delle elementari si adatta anche al caso in cui uno dei due sia un allineamento decimale infinito ed uno invece finito. Le cose si fanno un po' più complicate quando invece sono entrambi infiniti. Un'idea potrebbe essere considerare gli approssimanti finiti $x_n = k_0, k_1k_2\cdots k_n$ e $y_n = h_0, h_1h_2\cdots h_n$. Essi sono per costruzione degli allineamenti decimali finiti e possiamo quindi sommarli e considerare $x_n + y_n$. Possiamo pensare questi come gli approssimanti decimali di $x + y$?. Si noti che se $x = 0,\overline{1}$ e $y = 0,\overline{8}$, allora

$$x_n = 0,\ \underbrace{1\cdots1}_{n} \qquad y_n = 0,\ \underbrace{8\cdots8}_{n} \qquad x_n + y_n = 0,\ \underbrace{9\ldots9}_{n}\,.$$

D'altra parte non è difficile intuire che $x + y = 0,\overline{9} = 1$. Quindi l'approssimante decimale di ordine n di $x + y$ non è dato da $x_n + y_n$, ma dal numero 1 stesso. Si noti tuttavia che, man mano che n cresce, $x_n + y_n$, pur non essendo l'approssimante decimale per difetto $(x+y)_n$ di $x+y$, 'si avvicina' al numero x+y=1. Questo concetto di convergenza, che sarà studiato più avanti, permette di definire rigorosamente il concetto di somma e anche quello di prodotto di due qualunque numeri reali. Non insisteremo oltre su queste questioni, ma è importante che lo studente si renda conto che per definire correttamente le operazioni di somma e prodotto sui reali vi sono queste difficoltà concettuali.

Le operazioni di somma e prodotto per i numeri reali soddisfano alle stesse regole che per i razionali e che qui sotto brevemente richiamiamo. Nel seguito x, y, e z sono tre qualunque numeri reali.

(P1)	$(x+y)+z = x+(y+z)$	associatività +
(P2)	$x+y = y+x$	commutatività +
(P3)	$x+0 = 0+x = x$	elemento neutro +
(P4)	$\forall\, x,\ \exists! -x:\ \ x+(-x) = 0$	elemento opposto
(P5)	$(x \cdot y) \cdot z = x \cdot (y \cdot z)$	associatività $\cdot$
(P6)	$x \cdot y = y \cdot x$	commutatività $\cdot$
(P7)	$x \cdot 1 = 1 \cdot x = x$	elemento neutro $\cdot$
(P8)	$\forall\, x \neq 0,\ \exists!\, x^{-1}:\ \ x \cdot x^{-1} = 1$	elemento reciproco
(P9)	$x \cdot (y+z) = x \cdot y + x \cdot z$	distributività

Un qualunque insieme dotato di due operazioni + e $\cdot$ soddisfacenti tutte le proprietà precedenti viene detto *campo*. Parleremo così del *campo dei numeri reali*.

Il segno di moltiplicazione $\cdot$ verrà spesso omesso, qualora questo non crei ambiguità. Inoltre, se $x \in \mathbb{R}$ e $n \in \mathbb{N}$, il simbolo x^n indicherà, come al solito, il prodotto di n fattori uguali a x. Se $x \neq 0$, si pone inoltre per convenzione $x^0 = 1$.

Le nove proprietà sopra esposte, ne implicano in realtà molte altre che, come queste, vi sarà già capitato di vedere e di usare più o meno consciamente. Ne riportiamo un paio abbastanza importanti.

(Q1)	$x \cdot y = 0$ se e solo se uno dei due fattori x o y è 0
(Q2)	$(-x) \cdot y = x \cdot (-y) = -(xy)$

Facciamo vedere come la (Q1), nota anche come *legge di annullamento del prodotto*, si deduce dalle proprietà di campo, lasciando la verifica dell'altra allo studente.

Dimostrazione di (Q1): Dimostriamo prima il 'se', cioè che se uno dei due fattori è zero, ad esempio y, allora il prodotto è 0. Si ha

$$x \cdot 0 = x \cdot (0 + 0) = x \cdot 0 + x \cdot 0 \tag{1.3}$$

(la prima eguaglianza segue da (P3), la seconda da (P9)). Sommando ad ambo i membri il numero $-(x \cdot 0)$ si ottiene così

$$0 = -(x \cdot 0) + x \cdot 0 = (-(x \cdot 0) + x \cdot 0) + x \cdot 0 = 0 + x \cdot 0 = x \cdot 0.$$

(la prima eguaglianza segue da (P4), la seconda da (1.3) e da (P1), la terza da (P4), la quarta infine da (P3)).

Dimostriamo ora il 'solo se', cioè facciamo vedere che se $x \cdot y = 0$, allora, necessariamente, uno dei due fattori deve essere 0. Se $x = 0$ siamo a posto, altrimenti supponiamo $x \neq 0$ e moltiplichiamo a sinistra per x^{-1}. Otteniamo

$$0 = x^{-1} \cdot 0 = x^{-1} \cdot (x \cdot y) = (x^{-1} \cdot x) \cdot y = 1 \cdot y = y$$

(la prima eguaglianza segue dal 'se' appena dimostrato, la seconda dall'ipotesi fatta, la terza da (P5), la quarta da (P8)). La dimostrazione è così completata. ■

Come avevamo notato prima, la corrispondenza con la retta determina una struttura di ordinamento naturale sui numeri reali (indicato ancora con i simboli $<$, $>$, $\leq$, $\geq$). In termini delle rappresentazioni decimali, la relazione di ordine può essere vista nel modo seguente: consideriamo ad esempio due numeri positivi $x = k_0, k_1 \cdots$ e $y = h_0, h_1 \cdots$. Allora,

$$x < y \iff \exists r \in \mathbb{N} \text{ tale che } k_i = h_i \text{ per } i = 0, \ldots, r-1 \text{ e } k_r < h_r\,.$$

L'ordinamento sui reali, come quello sui razionali, gode di alcune proprietà che, seppure evidenti, giocano un ruolo fondamentale in moltissime utilizzazioni dei numeri reali

(P10)	$x \leq x$	riflessività
(P11)	$x \leq y,\ \ y \leq x \ \Rightarrow\ x = y$	antisimmetria
(P12)	$x \leq y,\ \ y \leq z \ \Rightarrow\ x \leq z$	transitività

Inoltre, la relazione di ordine che vi è su $\mathbb{R}$ ha delle proprietà importanti di connessione con la struttura algebrica di campo che riportiamo qui sotto.

(P13)	$x \leq y,\ z \leq w,\ \Rightarrow\ x + z \leq y + w$
(P14)	$x \leq y,\ z \geq 0,\ \Rightarrow\ x \cdot z \leq y \cdot z$

Un campo dotato di una relazione di ordine $\leq$ che soddisfi le proprietà (P10)-(P14) è detto un *campo ordinato*. I numeri reali formano un campo ordinato.

Dalle proprietà miste (P13) e (P14) se ne deducono altre come ad esempio

$$
\begin{array}{lll}
\text{(Q3)} & x \leq y,\ z \leq 0, & \Rightarrow\ x \cdot z \geq y \cdot z \\
\text{(Q4)} & x + z \leq y + z, & \Rightarrow\ x \leq y \\
\text{(Q5)} & x \cdot z \leq y \cdot z,\ z > 0 & \Rightarrow\ x \leq y \\
\text{(Q6)} & 0 < x \leq y & \Rightarrow\ 0 < y^{-1} \leq x^{-1}
\end{array}
$$

Lo studente certamente noterà come le proprietà (P13), (P14) e (Q3)-(Q6) sono continuamente utilizzate nella risoluzione di disequazioni.

Introduciamo ora un concetto molto utile, quello di valore assoluto di un numero reale. Il *valore assoluto* familiarmente (ma impropriamente) conosciuto da molti studenti come 'il numero senza segno' ha la seguente precisa definizione:

$$
x \in \mathbb{R} \quad |x| = \begin{cases} x & \text{se } x \geq 0, \\ -x & \text{se } x < 0 \end{cases} .
$$

Nonostante la semplicità della definizione, il valore assoluto è foriero di molti errori. Spesso ci troveremo a dover considerare diseguaglianze del tipo $|x| \leq a$ dove $a \in \mathbb{R}$. Poichè per definizione $|x| \geq 0$ si ha che la suddetta diseguaglianza non ha soluzioni se $a < 0$. Nel caso invece in cui $a \geq 0$ si ha che

$$
|x| \leq a \iff -a \leq x \leq a\,. \tag{1.4}
$$

Se invece consideriamo $|x| \geq a$, essa è sempre soddisfatta se $a \leq 0$, mentre, se $a > 0$ si ha

$$
|x| \geq a \iff x \leq -a \text{ oppure } x \geq a\,. \tag{1.5}
$$

Il valore assoluto gode di alcune importanti proprietà:

$$
\begin{array}{ll}
\text{(Q7)} & |x + y| \leq |x| + |y| \\
\text{(Q8)} & |x \cdot y| = |x| \cdot |y|
\end{array}
$$

Dimostreremo l'importante (Q7) lasciando la dimostrazione (più semplice) di (Q8) allo studente.

Dimostrazione di (Q7): Poichè sicuramente $|x| \leq |x|$ e $|y| \leq |y|$, segue da (1.4) che

$$
\begin{aligned}
-|x| \leq x \leq |x|\,, \\
-|y| \leq y \leq |y|\,.
\end{aligned}
$$

Sommando membro a membro, si ottiene

$$-(|x|+|y|) \leq x+y \leq (|x|+|y|)$$

che, per la (1.5), è equivalente a

$$|x+y| \leq |x|+|y|\,.$$

■

Si noti che se x e y hanno lo stesso segno, allora la (Q7) è addirittura un'eguaglianza: $|x+y| = |x|+|y|$. La diseguaglianza si ha nei casi in cui i segni sono discordi.

Esercizio 1.4 Si risolva la disequazione $|3x-|x|| < |x|+1$
R: $-1/3 < x < 1$

Esercizio 1.5 Si descriva sul piano xy l'insieme delle soluzioni della diseguaglianza $|x-y| \leq 1$.

Esercizio 1.6 * Si dimostri che se x e y sono numeri reali, si ha

$$|x-y| \geq ||x|-|y||\ .$$

1.5 Proprietà di continuità

Le proprietà algebriche e di ordinamento illustrate nel paragrafo precedente non sono esclusive dei numeri reali. In effetti anche i razionali godono delle stesse proprietà, in altri termini anche $\mathbb{Q}$ è un campo ordinato. Ciò che in effetti differenzia i due insiemi numerici riguarda, come è stato discusso prima, la 'continuità' dell'insieme dei numeri reali, la sua struttura di 'retta senza buchi'. Vogliamo qui formalizzare meglio questo concetto rendendo rigoroso il concetto di continuità in termini della definizione che abbiamo assunto di numeri reali come allineamenti decimali.

Cominciamo con alcune definizioni.

Definizione 1.2 Sia $A \subseteq \mathbb{R}$ un sottoinsieme.

- Un elemento $M \in A$ è detto *massimo di A* se $x \leq M$ per ogni $x \in A$.
- Un elemento $m \in A$ è detto *minimo di A* se $x \geq m$ per ogni $x \in A$.

Si usano le notazioni $M = \max A$, $m = \min A$. E' facile verificare (lo studente è invitato a riflettere sul perchè) che se A ammette un elemento massimo, esso è unico; similmente per il minimo.

Mostriamo alcuni esempi.

Esempio 2 $A = [a, b]$. Allora $\max A = b$ e $\min A = a$.

Esempio 3 $A = \{(-1)^n \frac{1}{n} \mid n = 2, 3, \ldots\}$. A consiste di numeri sia positivi che negativi. Prendendo n pari si ottengono i numeri

$$\frac{1}{2}, \frac{1}{4}, \frac{1}{6}, \ldots$$

mentre prendendo n dispari si ottengono i numeri

$$-\frac{1}{3}, -\frac{1}{5}, -\frac{1}{7}, \ldots$$

Si ha quindi $\max A = 1/2$ e $\min A = -1/3$.

E' facile tuttavia costruire esempi di insiemi che non ammettono massimo e/o minimo:

Esempio 4 $A = \mathbb{N}$. Allora non esiste il massimo mentre il minimo è 1.

Esempio 5 $A = \mathbb{Z}$. Allora non esiste nè il massimo, nè il minimo

Negli esempi precedenti la mancanza di minimo o massimo è collegata ad una 'illimitatezza' dell'insieme stesso. Introduciamo il seguente concetto:

Definizione 1.3 Sia $A \subseteq \mathbb{R}$ un sottoinsieme.

- A è detto *superiormente limitato* se esiste $L \in \mathbb{R}$ tale che $x \leq L$ per ogni $x \in A$.
- A è detto *inferiormente limitato* se esiste $l \in \mathbb{R}$ tale che $x \geq l$ per ogni $x \in A$.
- A è detto *limitato* se è sia inferiormente che superiormente limitato.

Chiaramente se A ammette massimo, esso è superiormente limitato e se ammette minimo è inferiormente limitato. Sarà vero che un insieme superiormente limitato necessariamente ammette massimo e che un insieme inferiormente limitato necessariamente ammette minimo? La risposta è in entrambi casi sul negativo come mostrano i seguenti:

Esempio 6 Sia $A =]0,1[$. A è limitato; facciamo vedere che non ammette massimo. Per assurdo supponiamo che il massimo ci sia e chiamiamolo $M \in]0,1[$. Tale numero sarà del tipo $M = 0, k_1 k_2 k_3 \cdots$ con non tutti i k_i eguali a 9 (infatti ce ne saranno infiniti non eguali a 9). Supponiamo che $k_s < 9$ e consideriamo il numero $\tilde{M} = 0, k_1 k_2 \cdots k_{s-1}(k_s + 1)k_{s+1} \cdots$. Chiaramente $\tilde{M} \in]0,1[$ e $\tilde{M} > M$ e questo significa che M non poteva essere il massimo di A. Similmente si fa vedere che A non ammette minimo.

Esempio 7 Sia $A = \{1/n \mid n = 1, 2, 3, \ldots\}$. A è limitato; chiaramente esiste il massimo di A che è 1. Non esiste invece il minimo: in effetti se per assurdo $a \in A$ fosse il minimo si avrebbe $a = 1/n$ per qualche n (essendo questi gli elementi di A). Ma $1/(n+1) \in A$ e $1/(n+1) < 1/n$ il che significa che $a = 1/n$ non può essere il minimo.

Gli esempi sopra suggeriscono un modo di generalizzare i concetti di massimo e di minimo di un insieme. Nel caso $A =]0,1[$ pur non esistendo nè massimo nè minimo, vi sono due numeri in un certo senso speciali per A: 0 e 1. 1 non è il massimo perchè non sta in A, però ha una notevole proprietà: se L è un qualunque numero che sta alla destra di A cioè tale che $L \geq x$ per ogni $x \in A$, allora $1 \leq L$; in altri termini 1 è il più piccolo dei numeri che stanno alla destra di A. Similmente 0 può essere caratterizzato come il numero più grande che sta alla sinistra di A. Quanto è generale questo nostro ragionamento? Può essere fatto per ogni insieme limitato? La risposta è affermativa e conduce al cuore del problema di continuità. Per formalizzare i ragionamenti che faremo è conveniente prima introdurre qualche altra notazione.

Definizione 1.4 Sia dato un insieme $A \subseteq \mathbb{R}$. Un numero reale L è detto *maggiorante di* A se $L \geq x$ per ogni $x \in A$. Un numero reale l è detto *minorante di* A se $l \leq x$ per ogni $x \in A$. L'insieme dei maggioranti di A lo indicheremo con il simbolo A^+, mentre quello dei minoranti con il simbolo A^-.

E' chiaro che A è superiormente limitato se e soltanto se esiste almeno un maggiorante, cioè se A^+ è non vuoto. Similmente, A è inferiormente limitato se e soltanto se A^- è non vuoto. Inoltre il massimo, se esiste, è un maggiorante, mentre il minimo, se esiste, è un minorante. Vale il seguente fondamentale risultato:

Teorema 1.5 *Sia $A \subseteq \mathbb{R}$. Allora:*

(i) Se A è superiormente limitato, A^+ ammette minimo che viene detto l'estremo superiore di A e indicato con $\sup A = \min A^+$.

(ii) Se A è inferiormente limitato, A^- ammette massimo che viene detto l'estremo inferiore di A e indicato con $\inf A = \max A^-$.

Dimostrazione Diamo solo un'idea della dimostrazione che contiene delle idee piuttosto interessanti. Dimostriamo (i) nel caso particolare in cui $A \cap \mathbb{R}^+ \neq \emptyset$, così che $A^+ \subseteq \mathbb{R}^+$. Gli elementi di A^+ saranno quindi del tipo $x = k_0, k_1 k_2 k_3 \cdots$. Consideriamo

$$\tilde{k}_0 = \min\{k_0 \mid x \in A^+\}\,, \qquad A_0^+ = \{x \in A^+ \mid k_0 = \tilde{k}_0\}\,,$$

$$\tilde{k}_1 = \min\{k_1 \mid x \in A_0^+\}\,, \qquad A_1^+ = \{x \in A_0^+ \mid k_1 = \tilde{k}_1\}\,,$$

e così via, iterando,

$$\tilde{k}_n = \min\{k_n \mid x \in A_{n-1}^+\}\,, \qquad A_n^+ = \{x \in A_{n-1}^+ \mid k_n = \tilde{k}_n\}\,.$$

Si ha chiaramente $A^+ \supseteq A_0^+ \supseteq A_1^+ \supseteq \cdots$ e tutti gli A_r^+ sono, per costruzione, non vuoti. Consideriamo $L = \tilde{k}_0, \tilde{k}_1 \tilde{k}_2 \tilde{k}_3 \cdots$ e dimostriamo che questo è il minimo di A^+. Per come è stato costruito è facile rendersi conto che $L \leq x$ per ogni $x \in A^+$. Rimane da dimostrare che L sta in A^+. Se per assurdo $L \notin A^+$, vuol dire che non è un maggiorante di A, quindi esiste $y \in A$ tale che $y > L$. Avremo $y = h_0, h_1 h_2 \cdots$ ed esisterà un indice $r \geq 0$ tale che

$$\tilde{k}_i = h_i \ \text{ per } i = 0, 1, \ldots, r-1 \quad \text{e} \ \ \tilde{k}_r < h_r\,.$$

Scegliamo un qualunque $z \in A_r^+$. z è un maggiorante e la sua rappresentazione decimale è del tipo $z = \tilde{k}_0, \tilde{k}_1 \tilde{k}_2 \cdots \tilde{k}_r k_{r+1} \cdots$. Quindi $z < y$ e questo è assurdo perchè $y \in A$ e $z \in A^+$. Quindi L deve stare in A^+ e quindi è il minimo di A^+. Lo studente pensi a come estendere la dimostrazione di (i) al caso generale. La dimostrazione di (ii) si fa in modo analogo: vale la pena notare che sfruttando la simmetria dell'insieme dei numeri reali rispetto allo 0, si può far discendere (ii) da (i); lasciamo allo studente il compito di formalizzare il procedimento. ■

Nel caso in cui un insieme A non sia, rispettivamente, superiormente o inferiormente limitato, si pone, per convenzione $\sup A = +\infty$ o $\inf A = -\infty$.

Esempio 8 Riprendiamo l'Esempio 6: $A =]0,1[$. Allora $A^+ = [1,+\infty[$: '$\supseteq$' è evidente, mentre '$\subseteq$' segue dal fatto che per le considerazioni svolte nell'Esempio 6, non ci sono maggioranti più piccoli di 1. Similmente, $A^- =]-\infty,0]$. Quindi $\sup A = 1$ e $\inf A = 0$.

Esempio 9 Riprendiamo l'Esempio 7: $A = \{1/n \,|\, n = 1,2,3,\ldots\}$. Allora chiaramente $A^+ = [1,+\infty[$, mentre $A^- =]-\infty,0]$: '$\supseteq$' è evidente, mentre '$\subseteq$' segue dal fatto che non ci sono minoranti più grandi di 0 (si rifletta sul perchè). Quindi $\sup A = 1$ e $\inf A = 0$.

Esempio 10 $A = \{x \in \mathbb{R} \mid x^3 < x^2\}$. La diseguaglianza $x^3 < x^2$ è risolta da $x < 0$ e da $0 < x < 1$. Quindi, $A =]-\infty,0[\,\cup\,]0,1[$. Si ha, in questo caso, $A^+ = [1,+\infty[$ e $A^- = \emptyset$. Quindi, $\sup A = 1$ e $\inf A = -\infty$.

C'e' un'utile caratterizzazione per gli estremi superiore ed inferiore di un insieme:

Proposizione 1.6 *Sia A un insieme superiormente limitato e sia $L \in \mathbb{R}$. Sono equivalenti:*

1) $L = \sup A$.

2) L gode delle seguenti due proprietà:

(A) $L \geq x$ per ogni $x \in A$.

(B) Per ogni numero $\epsilon > 0$, esiste $x \in A$ tale che $x > L - \epsilon$.

Dimostrazione 1) $\Rightarrow$ 2): supponiamo che $L = \sup A$ e dimostriamo (A) e (B). Per quanto riguarda (A) si noti che essa dice semplicemente che L è un maggiorante di A: essa è quindi verificata essendo l'estremo superiore un maggiorante. Veniamo a (B): se, per assurdo essa fosse falsa vorrebbe dire che esiste $\epsilon > 0$ tale che $x \leq L - \epsilon$ per ogni $x \in A$. Quindi $L - \epsilon$ è anch'esso un maggiorante di A e $L - \epsilon < L$: ma questo è assurdo poichè L è, per ipotesi, il più piccolo dei maggioranti. Quindi anche (B) deve valere.

2) $\Rightarrow$ 1): supponiamo ora che L soddisfi le proprietà (A) e (B) e dimostriamo che $L = \sup A$. Dobbiamo far vedere che L è il minimo dei maggioranti. Il fatto che sia un maggiorante lo dice (A). Se non fosse il minimo, vorrebbe dire che esisterebbe un altro maggiorante $L' < L$. Sicuramente si può scrivere $L' = L - \epsilon$ per qualche numero $\epsilon > 0$. Poichè $L - \epsilon$ è un maggiorante, si ha che $x \leq L - \epsilon$ per ogni $x \in A$. Ma questo contraddice l'ipotesi (B). Quindi L è necessariamente il minimo dei maggioranti, cioè l'estremo superiore di A. ■

Proposizione 1.7 *Sia A un insieme inferiormente limitato e sia $l \in \mathbb{R}$. Sono equivalenti:*

1) $l = \inf A$.

2) l gode delle seguenti due proprietà:

(A) $l \leq x$ *per ogni* $x \in A$.

(B) Per ogni numero $\epsilon > 0$, esiste $x \in A$ tale che $x < l + \epsilon$.

Dimostrazione Completamente analoga alla precedente. E' lasciata come utile esercizio per lo studente. ■

Esercizio 1.7 Determinare gli estremi superiore ed inferiore per i seguenti insiemi, specificando se si tratta di massimi e/o minimi:

$$A_1 = \{x \in \mathbb{R} \mid |x| \leq 2x\}\,,$$

$$A_2 = \{n \in \mathbb{N} \mid n^2 \leq 4n\} \cup]-2, 2[\,,$$

$$A_3 = \{(-1)^n \tfrac{2n}{n+1} \mid n \in \mathbb{N}\}\,.$$

R: $\sup A_1 = +\infty$, $\min A_1 = 0$; $\inf A_2 = -2$, $\max A_2 = 4$; $\sup A_3 = 2$, $\inf A_3 = -2$.

Esercizio 1.8 * Siano A e B due sottoinsiemi di $\mathbb{R}$. Si dimostri che $A \cup B$ è superiormente limitato se e soltanto se lo sono entrambi A e B, e che inoltre vale la relazione:

$$\sup(A \cup B) = \max\{\sup A, \sup B\}\,.$$

Si enunci e si dimostri poi l'analoga proprietà per l'estremo inferiore.

L'esistenza dell'estremo superiore per insiemi superiormente limitati (o di quello inferiore per insiemi inferiormente limitati) è uno dei modi più eleganti per esprimere la proprietà di continuità dell'insieme dei numeri reali. E' una proprietà che l'insieme dei numeri razionali invece non ha: dentro $\mathbb{Q}$ vi sono sottoinsiemi superiormente limitati che non ammettono tuttavia estremo superiore come mostreremo tra poco. Intanto presentiamo l'importante risultato dell'esistenza delle radici che segue dall'esistenza dell'estremo superiore.

Teorema 1.8 *Sia $n \in \mathbb{N}$ e sia $b \geq 0$. Allora esiste uno ed un solo numero reale $a \geq 0$ tale che $a^n = b$. a viene detto la radice n-esima positiva di b ed indicato con i simbolo $\sqrt[n]{b}$ o $b^{1/n}$.*

Dimostrazione (Idea) Non forniamo una dimostrazione di questo risultato per il momento. Mostriamo soltanto la strada che utilizza direttamente il concetto di estremo superiore. Si consideri

$$A = \{x \in \mathbb{R}^+ \mid x^n \leq b\}\,.$$

Lunghi, ma non concettualmente difficili, passaggi mostrano che $a = \sup A$ ha le proprietà richieste, cioè $a^n = b$.

Per quanto concerne l'unicità della radice n-esima positiva invece il ragionamento è molto più semplice. Siano $a_1, a_2 \in \mathbb{R}^+$ tali che $a_1^n = b$ e $a_2^n = b$. Quindi $a_1^n = a_2^n$ e da questo segue (ci vuole l'induzione per rendere rigoroso questo punto assai intuitivo) che $a_1 = a_2$. ■

Osservazione: Se ogni sottoinsieme superiormente limitato di $\mathbb{Q}$ possedesse l'estremo superiore, il Teorema 1.8 varrebbe anche per $\mathbb{Q}$. In particolare esisterebbe in $\mathbb{Q}$ la radice quadrata di 2 cosa che sappiamo essere falsa per la Proposizione 1.1. Quindi vi devono necessariamente essere sottoinsiemi di $\mathbb{Q}$, superiormente limitati, che non ammettono estremo superiore razionale. Uno di questi è proprio

$$A = \{x \in \mathbb{Q}^+ \,|\, x^2 \leq 2\}$$

che ha come estremo superiore $\sqrt{2}$.

A partire dalle radici n-esime si possono definire gli elevamenti ad una qualunque potenza razionale come segue. Sia $a \geq 0$ e sia $q = m/n$ con $n, m \in \mathbb{N}$. Si pone

$$a^q = (a^m)^{1/n} = (a^{1/n})^m\,,$$

dove la prima eguaglianza è da intendersi come definizione, mentre la seconda segue dall'osservazione che $(a^{1/n})^m$ è la radice n-esima positiva di a^m in quanto $((a^{1/n})^m)^n = ((a^{1/n})^n)^m = a^m$. Se invece $q \in \mathbb{Q}$ è negativo, si definisce

$$a^q = \frac{1}{a^{-q}}\,.$$

Come sicuramente sapete, l'elevazione a potenza si può estendere a potenze reali qualsiasi in modo tale che le proprietà fondamentali delle potenze continuino a valere. Un modo rigoroso per definire questi elevamenti a potenza reale è il seguente: sia ancora $a \geq 0$ e sia $x \geq 0$. Indichiamo con x_n

l'approssimante decimale n-esimo per difetto di x. Allora si pone

$$\begin{aligned} a^x &= \sup\{a^{x_n} \mid n \in \mathbb{N}\} \qquad \text{se } a > 1\,, \\ a^x &= \inf\{a^{x_n} \mid n \in \mathbb{N}\} \qquad \text{se } a < 1\,. \end{aligned} \tag{1.6}$$

Si può dimostrare, anche se è estremamente noioso farlo, che con tale definizione l'elevamento a potenza gode delle ben note proprietà:

$$a^{x_1+x_2} = a^{x_1}a^{x_2}\,, \qquad a^{x_1x_2} = (a^{x_1})^{x_2}\,, \qquad (a_1a_2)^x = a_1^x a_2^x\,, \tag{1.7}$$

dove $a, a_1, a_2 \geq 0$ e x, x_1, x_2 numeri reali qualsiasi. Inoltre si ha che

$$\begin{array}{lll} a > 1, \;\; x_1 < x_2 & \Rightarrow & a^{x_1} < a^{x_2}\,, \\ a < 1, \;\; x_1 < x_2 & \Rightarrow & a^{x_1} > a^{x_2}\,. \end{array} \tag{1.8}$$

Fissiamo ora $a > 0$ e $b > 0$ e consideriamo l'equazione

$$a^x = b\,. \tag{1.9}$$

Segue dalla (1.8) che vi è al più una soluzione della (1.9). Si può dimostrare (lo faremo in seguito) che (1.9) ammette una ed una sola soluzione che viene detta il *logaritmo* in base a di b ed indicata con il simbolo $\log_a b$. Dalle (1.7) si possono ricavare le ben note proprietà algebriche dei logaritmi.

1.6 Questioni di approssimazione.

Consideriamo un numero reale $x = \pm k_0, k_1k_2\cdots$. I suoi approssimanti decimali finiti $x_n = \pm k_0, k_1k_2\cdots k_n$ sono dei numeri razionali che approssimano il numero x: tra x e x_n c'è una distanza di al più $1/10^n$. In un senso che sarà precisato nel prossimo capitolo si ha che i numeri x_n convergono verso x; quello che a noi importa, per il momento, sottolineare è che vi sono razionali vicini quanto vogliamo al numero x: questa proprietà si esprime dicendo che i razionali sono *densi* all'interno della retta reale.

Non sempre tuttavia conosciamo un numero reale direttamente attraverso il suo allineamento decimale, talvolta invece attraverso qualche proprietà caratterizzante. Ad esempio $\sqrt{2}$ è quel numero reale positivo (che sappiamo esistere) tale che $\sqrt{2}\cdot\sqrt{2} = 2$. Come si fa però a determinare la sua rappresentazione decimale, o comunque ad approssimarlo esplicitamente con numeri razionali? Questo è tutt'altro che un problema teorico quando dobbiamo fare dei conti ed il calcolatore che vi da il numero decimale $1,414\cdots$ quando digitate $\sqrt{2}$ avrà pure un metodo operativo per ottenerlo. Nel seguito presentiamo un modo per approssimare $\sqrt{2}$ rimandando al prossimo capitolo gli aspetti formali.

Sia $a_0 \in \mathbb{Q}$ tale che $a_0 \geq 2$, e si consideri

$$a_1 = \frac{1}{2}\left(a_0 + \frac{2}{a_0}\right) . \tag{1.10}$$

Si ha che

$$\begin{aligned} a_1 - \sqrt{2} &= \tfrac{1}{2}\left(a_0 + \tfrac{2}{a_0}\right) - \sqrt{2} \\ &= \frac{1}{2}\frac{\left(a_0^2 + 2 - 2\sqrt{2}a_0\right)}{a_0} \\ &= \frac{1}{2}\frac{(a_0 - \sqrt{2})}{a_0}(a_0 - \sqrt{2})\,. \end{aligned} \tag{1.11}$$

Quindi $a_1 \geq \sqrt{2}$. Inoltre, poichè sicuramente, per come è stato scelto a_0,

$$\frac{|a_0 - \sqrt{2}|}{a_0} \leq 1\,,$$

segue dalla (1.11) che

$$a_1 - \sqrt{2} \leq \frac{1}{2}\left|a_0 - \sqrt{2}\right| . \tag{1.12}$$

Questo significa che, comunque fosse stato scelto a_0, a_1 approssima (per eccesso) $\sqrt{2}$ meglio di a_0 di un fattore almeno $1/2$. Tutto questo suggerisce uno schema iterativo per trovare delle approssimazioni sempre migliori di $\sqrt{2}$:

$$\begin{aligned} a_1 &= \tfrac{1}{2}\left(a_0 + \tfrac{2}{a_0}\right) \\ a_2 &= \tfrac{1}{2}\left(a_1 + \tfrac{2}{a_1}\right) \\ a_3 &= \tfrac{1}{2}\left(a_2 + \tfrac{2}{a_2}\right) \\ &\vdots \\ a_{n+1} &= \tfrac{1}{2}\left(a_n + \tfrac{2}{a_n}\right) . \end{aligned} \tag{1.13}$$

Iterando la stima (1.12) si ottiene,

$$\begin{aligned} a_n - \sqrt{2} &\leq \frac{1}{2}|a_{n-1} - \sqrt{2}| \\ &\leq \frac{1}{2^2}|a_{n-2} - \sqrt{2}| \\ &\vdots \\ &\leq \frac{1}{2^n}|a_0 - \sqrt{2}|\,. \end{aligned}$$

Ad ogni passo si guadagna un fattore 1/2 nell'approssimare $\sqrt{2}$. Scegliendo, ad esempio, inizialmente $a_0 = 2$ e sfruttando il fatto che $1 < \sqrt{2} < 2$, si ha che $|a_0 - \sqrt{2}| < 1$ e quindi

$$a_n - \sqrt{2} \leq \frac{1}{2^n}\,.$$

Si ha quindi una stima assoluta di quanto a_n disti da $\sqrt{2}$ e prendendo n opportunamente grande ci si può avvicinare a $\sqrt{2}$ tanto quanto vogliamo. Riportiamo qui sotto le prime cifre decimali dei valori dei primi a_n:

$$a_0 = 2, \;\; a_1 = 1,5, \;\; a_2 = 1,4166\cdots, \;\; a_3 = 1,4142\cdots$$

Capitolo 2

Vettori e numeri complessi

2.1 Perchè altri numeri

Per molte applicazioni il campo dei numeri reali risulta ancora troppo piccolo. Ad esempio su di esso non si possono estrarre radici di numeri negativi: questo di per sè non sembra essere un problema pratico molto importante, tuttavia porta ad una serie di problemi teorici alcuni dei quali saranno toccati nel seguito del corso. Vorremmo un campo che contiene $\mathbb{R}$ nel quale si possano fare radici quadrate di qualunque elemento. Fortunatamente un tale campo esiste e ci sono vari modi per costruirlo: noi optiamo per la strada più geometrica che ci permetterà tra le altre cose di discutere alcune questioni sui vettori di interesse indipendente.

2.2 Vettori del piano

Consideriamo un riferimento cartesiano ortogonale $0XY$. Fissate unità di riferimento sugli assi X e Y si determina una corrispondenza biunivoca tra punti del piano e coppie ordinate di numeri reali: dato un punto P, si considerano le due proiezioni, rispettivamente a e b sui due assi, dette le *coordinate* di P. E' chiaro che il punto P risulta in questo modo univocamente determinato dalla coppia di numeri (a, b); useremo anche la notazione $P = (a, b)$. Il piano si può quindi pensare come il prodotto cartesiano della retta reale per se stessa $\mathbb{R} \times \mathbb{R}$.

Gli elementi di $\mathbb{R} \times \mathbb{R}$ vengono anche detti *vettori*. Il motivo di questa terminologia è dovuta al fatto che ad ogni punto del piano (e quindi ad ogni elemento di $\mathbb{R} \times \mathbb{R}$) si può associare un vettore geometrico inteso come segmento orientato che spicca dall'origine 0 del sistema di riferimento e finisce nel punto in questione.

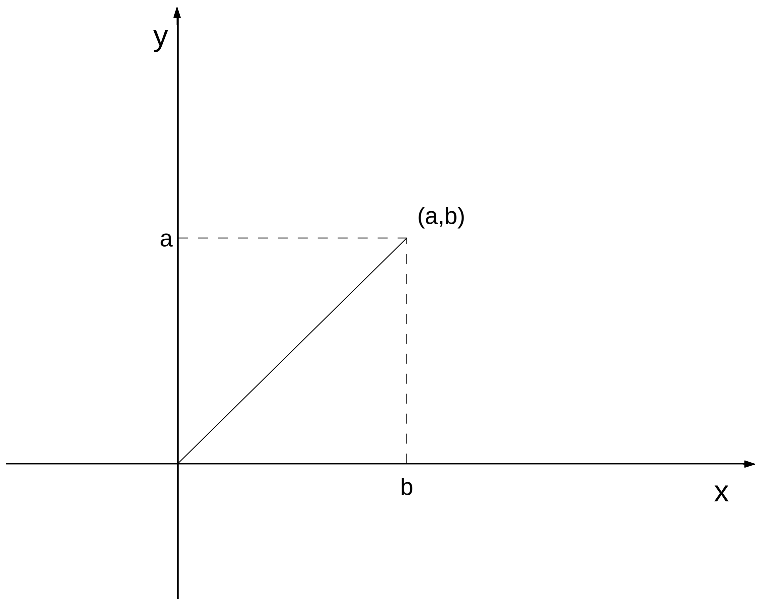


Figura 2.1:

Tale corrispondenza ci è utile per introdurre alcune operazioni. Consideriamo due punti (a_1, b_1) e (a_2, b_2) e i vettori geometrici corrispondenti. I vettori geometrici possono essere sommati con la nota regola del parallelogramma: il vettore somma è individuato in Figura 2.2 dal punto di coordinate (a, b).

Come si determinano a e b a partire dalle coordinate dei due punti di partenza? La risposta è sorprendentemente semplice: basta sommare le coordinate rispettive, cioè

$$a = a_1 + a_2 \,, \qquad b = b_1 + b_2 \,.$$

Abbiamo così introdotto su $\mathbb{R} \times \mathbb{R}$ un'operazione di somma che formalmente gode delle stesse proprietà (P1-P4) dei numeri reali.

Vi è anche un'altra importante operazione sui vettori geometrici: il prodotto per uno scalare. Consideriamo il vettore geometrico relativo a (a, b) e sia $\lambda \in \mathbb{R}$: la moltiplicazione per lo scalare λ consiste nel considerare il

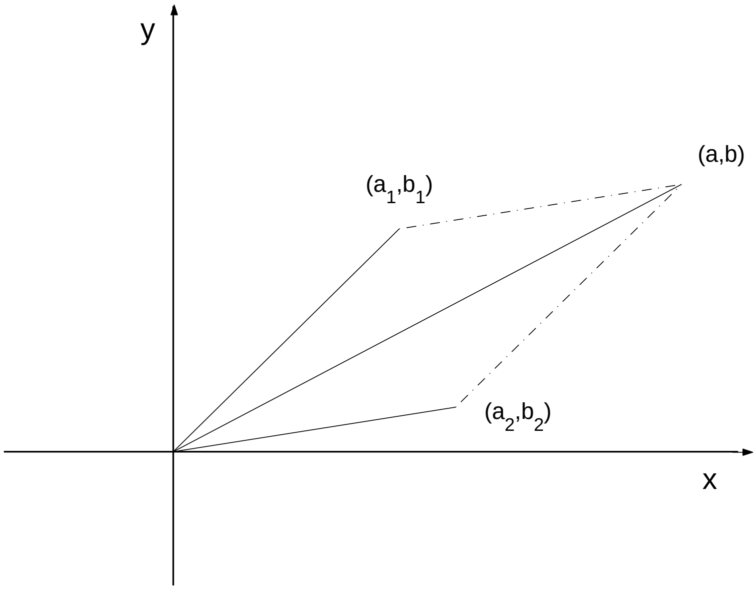


Figura 2.2:

vettore geometrico che ha la stessa direzione di quello iniziale, lunghezza modificata di un fattore $|\lambda|$, stesso verso se $\lambda \geq 0$, verso opposto se $\lambda < 0$. E' facile rendersi conto che il punto finale del vettore così ottenuto ha coordinate $(\lambda a, \lambda b)$. Abbiamo così ottenuto una moltiplicazione di elementi di $\mathbb{R} \times \mathbb{R}$ per numeri reali. Si noti che dato un qualunque (a, b) si può sempre scrivere:

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

Ogni elemento di $\mathbb{R} \times \mathbb{R}$ si scrive così come *combinazione lineare* dei due elementi base $(1, 0)$ e $(0, 1)$.

2.3 La costruzione dei numeri complessi

Vorremmo introdurre su $\mathbb{R} \times \mathbb{R}$ anche un'operazione di moltiplicazione interna, cioè tra coppie di numeri reali in modo tale da renderlo un campo come $\mathbb{R}$. La prima cosa che può venire in mente è di considerare la moltiplicazione coordinata per coordinata mimando così la definizione di somma. Questa non è tuttavia una buona idea in quanto non soddisferebbe ad esempio alla proprietà (P8) (si pensi al perchè). Dobbiamo prendere un'altra strada.

Ricordiamoci che il campo che vogliamo costruire deve contenere $\mathbb{R}$ al suo interno. Ora ci sono ovviamente tante rette possibili dentro il piano $\mathbb{R} \times \mathbb{R}$, scegliamone una, l'asse X, questa vogliamo che sia il vecchio campo $\mathbb{R}$: su di essa somma e moltiplicazione devono essere come sui reali. Quindi si deve avere:

$$(a_1, 0) \cdot (a_2, 0) = (a_1 \cdot a_2, 0)\,.$$

Si noti in particolare che l'elemento $(1, 0)$ è l'elemento neutro della moltiplicazione per i punti della retta X, e quindi deve essere anche l'elemento neutro per la moltiplicazione come operazione su tutto $\mathbb{R} \times \mathbb{R}$. Prendiamo ora due qualunque elementi e moltiplichiamoli. Se devono valere le varie proprietà (P1-P8) si deve necessariamente avere:

$$\begin{aligned} (a_1, b_1) \cdot (a_2, b_2) &= [a_1(1, 0) + b_1(0, 1)] \cdot [a_2(1, 0) + b_2(0, 1)] \\ &= a_1a_2(1, 0) + b_1b_2(0, 1) \cdot (0, 1) + [a_1b_2 + b_1a_2](0, 1)\,. \end{aligned}$$

L'unica cosa aperta che rimane da decidere è, a questo punto, quanto fa $(0, 1) \cdot (0, 1)$. Ci sono molte possibilità, la più semplice e che ci permette di costruire il campo che volevamo è scegliere

$$(0, 1) \cdot (0, 1) = -(1, 0)\,. \tag{2.1}$$

Si ottiene così

$$(a_1, b_1) \cdot (a_2, b_2) = [a_1a_2 - b_1b_2](1, 0) + [a_1b_2 + b_1a_2](0, 1)$$

o anche,

$$(a_1, b_1) \cdot (a_2, b_2) = (a_1a_2 - b_1b_2, a_1b_2 + b_1a_2)\,. \tag{2.2}$$

Non è difficile, anche se piuttosto noioso, verificare direttamente a posteriori che questa definizione di moltiplicazione, insieme alla somma definita prima, rende $\mathbb{R} \times \mathbb{R}$ un campo. Esso viene indicato con il simbolo $\mathbb{C}$ e chiamato il *campo dei numeri complessi*.

L'elemento $(0, 1)$ viene detto *unità immaginaria* e indicato con il simbolo i. Si usa anche la convenzione di indicare i numeri complessi che stanno sulla retta reale del tipo $(a, 0)$, semplicemente come a: sono i numeri reali dentro il piano complesso. Con queste notazioni, l'espressione (2.1) diventa

$$i^2 = -1\,. \tag{2.3}$$

Ogni numero complesso (a, b) può quindi essere espresso come

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

Questa notazione è molto pratica e permettte di ricordare facilmente le operazioni di somma e prodotto sui complessi. In effetti, usando le regole di campo ed il fatto che $i^2 = -1$, si ha

$$(a_1 + ib_1) + (a_2 + ib_2) = a_1 + a_2 + i(b_1 + b_2)\,,$$

$$(a_1+ib_1)\cdot(a_2+ib_2) = a_1a_2+ia_1b_2+ib_1a_2+i^2b_1b_2 = (a_1a_2-b_1b_2)+i(a_1b_2+b_1a_2)\,.$$

Spesso si usa la lettera z per indicare un generico numero complesso; se $z = a + ib$, diremo anche che a è la *parte reale* e b è la *parte immaginaria* del numero complesso z e si scrive anche $a = \Re e\, z$, $b = \Im m\, z$.

Esercizio 2.1 Trovare parte reale e immaginaria di $z = (a + ib)^3$, dove a e b sono numeri reali.
R: $\Re e\, z = a^3 - 3ab^2$, $\Im m\, z = 3a^2b - b^3$.

2.4 Modulo e coniugio

Sia $z = a + ib$ un numero complesso. Si definisce il *coniugio* di z come il numero complesso $\overline{z} = a - ib$. Sul piano, $\overline{z}$ rappresenta il simmetrico di z rispetto alla retta reale (vedi Figura 2.3).

L'operazione di coniugio gode di alcune importanti proprietà di immediata dimostrazione che riportiamo sotto:

$$\begin{array}{lll}
\text{(C1)} & \overline{z_1 + z_2} = \overline{z_1} + \overline{z_2} & \forall z_1,\, z_2 \in \mathbb{C}\,, \\
\text{(C2)} & \overline{z_1 \cdot z_2} = \overline{z_1} \cdot \overline{z_2} & \forall z_1,\, z_2 \in \mathbb{C}\,, \\
\text{(C3)} & \overline{z} = z \;\Leftrightarrow\; z \in \mathbb{R}\,. &
\end{array}$$

Dato un numero complesso $z = a+ib$ si può poi considerare il suo *modulo* definito nel modo dei vettori geometrici:

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

Osserviamo che se z_1 e z_2 sono due numeri complessi, allora $|z_1 - z_2|$ rappresenta la distanza euclidea nel piano tra i punti z_1 e z_2. Se z_0 è un numero complesso, la diseguaglianza $|z - z_0| \leq r$ rappresenta il cerchio, nel piano complesso di centro z_0 e raggio r.

Il modulo dei complessi gode di proprietà molto simili al valore assoluto di numeri reali. Si ha in particolare

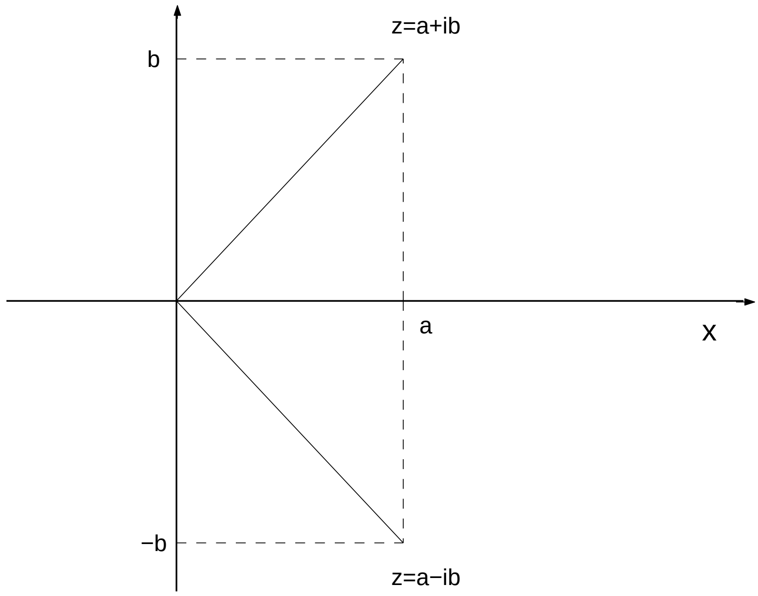


Figura 2.3:

$$
\begin{array}{lll}
\text{(M1)} & |z_1 + z_2| \leq |z_1| + |z_2| & \forall z_1,\, z_2 \in \mathbb{C}\,, \\
\text{(M2)} & |z_1 \cdot z_2| = |z_1| \cdot |z_2| & \forall z_1,\, z_2 \in \mathbb{C}\,, \\
\text{(M3)} & |\overline{z}| = |z| & \forall z \in \mathbb{C}\,, \\
\text{(M4)} & \overline{z} \cdot z = |z|^2 & \forall z \in \mathbb{C}\,, \\
\text{(M5)} & |\Re e\, z| \leq |z|, \quad |\Im m\, z| \leq |z| & \forall z \in \mathbb{C}\,.
\end{array}
$$

La diseguaglianza (M1) è generalmente nota come *diseguaglianza triangolare*. In effetti se consideriamo l'interpretazione geometrica della somma, si ha che $|z_1|$, $|z_2|$ e $|z_1 + z_2|$ rappresentano le lunghezze dei tre lati di un triangolo, da cui la diseguaglianza. Ovviamente la (M1) può anche essere dimostrata per via puramente analitica, cosa che noi non faremo. La (M2) si dimostra direttamente utilizzando le definizioni di modulo e di moltiplicazione (provare a farla per esercizio). La (M3) è immediata dalla definizione. La (M4) pure semplice si vede nel modo seguente: se $z = a + ib$ si ha che

$$\overline{z} \cdot z = (a - ib)(a + ib) = a^2 - (ib)^2 = a^2 + b^2 = |z|^2\,.$$

Infine, la prima delle (M5) segue dal fatto che

$$|z| = \sqrt{(\Re e\, z)^2 + (\Im m\, z)^2} \geq \sqrt{(\Re e\, z)^2} = |\Re e\, z|$$

e similmente l'altra.

La proprietà (M4) offre un modo piuttosto semplice per determinare il reciproco di un numero complesso. Sia $z \in \mathbb{C}$ con $z \neq 0$: segue dalla (M4) che

$$z \cdot \frac{\overline{z}}{|z|^2} = 1 \ \Rightarrow\ z^{-1} = \frac{\overline{z}}{|z|^2}\,.$$

Vediamo un semplice esempio:

Esempio 11 Sia $z = 1 + i$. Allora, $\overline{z} = 1 - i$ e $|z|^2 = 2$. Quindi,

$$z^{-1} = \frac{\overline{z}}{z^2} = \frac{1}{2}(1 - i) = \frac{1}{2} - i\frac{1}{2}\,.$$

Esercizio 2.2 Calcolare

$$4i + \frac{2+3i}{2+i}, \qquad \frac{1}{2i}(i^5 - i^{-5})\,.$$

R: $7/5 + 24i/5$, 1.

Esercizio 2.3 Dimostrare che se z un numero complesso tale che $\Im m\, z > 0$, allora $\Im m\, \frac{z-1}{z+1} > 0$.

Esercizio 2.4 * Provare l'identità del parallelogramma:

$$|z - w|^2 + |z + w|^2 = 2|z|^2 + 2|w|^2 \qquad \forall z, w \in \mathbb{C}\,.$$

2.5 La forma trigonometrica dei numeri complessi

Dato un numero complesso $z \in \mathbb{C} \setminus 0$, consideriamo il suo modulo $\rho = |z|$ e l'angolo θ, detto *argomento* (o *fase*), che esso forma, pensato come vettore geometrico, con la semiretta positiva dell'asse X:

Questi due numeri (ρ, θ) completamente determinano il numero z che da essi può essere ricostruito semplicemente considerando:

$$z = (\rho\cos\theta, \rho\sin\theta) = \rho\cos\theta + i\rho\sin\theta\,. \tag{2.4}$$

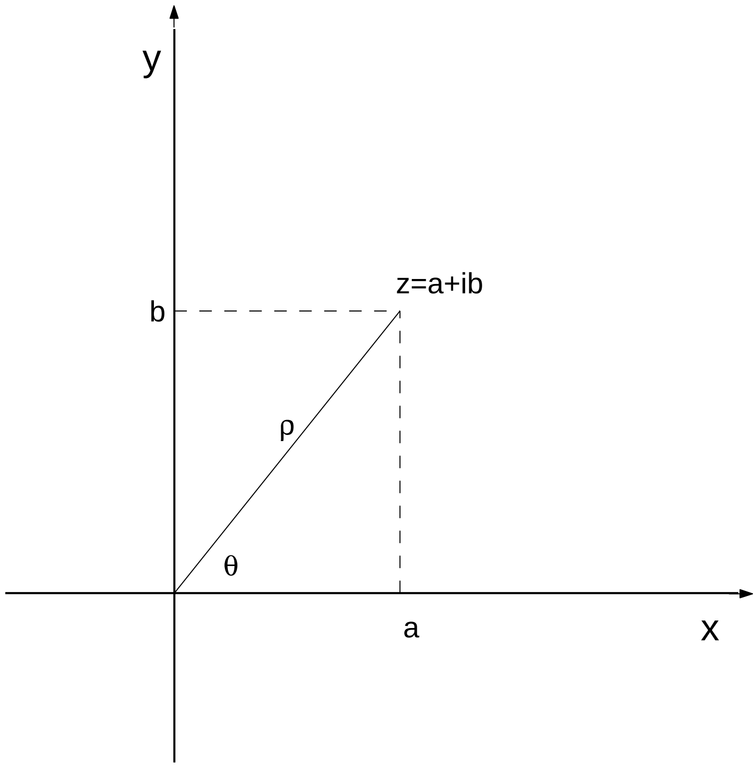


Figura 2.4:

La rappresentazione (2.4) è detta *forma trigonometrica* dei numeri complessi. Essa risulta molto utile sopratutto per il modo in cui permette di esprimere la moltiplicazione. Consideriamo due numeri complessi in forma trigonometrica

$$z_1 = \rho_1 \cos\theta_1 + i\rho_1 \sin\theta_1, \qquad z_2 = \rho_2 \cos\theta_2 + i\rho_2 \sin\theta_2 \,.$$

Si ha:

$$\begin{aligned} z_1 z_2 &= (\rho_1 \cos\theta_1 + i\rho_1 \sin\theta_1)(\rho_2 \cos\theta_2 + i\rho_2 \sin\theta_2) \\ &= \rho_1\rho_2(\cos\theta_1 \cos\theta_2 - \sin\theta_1 \sin\theta_2) + i\rho_1\rho_2(\cos\theta_1 \sin\theta_2 + \sin\theta_1 \cos\theta_2) \\ &= \rho_1\rho_2 \cos(\theta_1 + \theta_2) + i\rho_1\rho_2 \sin(\theta_1 + \theta_2) \,. \end{aligned} \tag{2.5}$$

Questo mostra che quando si moltiplicano due numeri complessi, i loro moduli si moltiplicano, mentre i loro argomenti si sommano.

Osservazione: Segue dalle considerazioni precedenti che, moltiplicando un numero complesso z per l'unità immaginaria i, si ottiene il numero iz che ha lo stesso modulo di z ma fase variata di un addendo $\pi/2$: cioè iz è il ruotato di z d un angolo $\pi/2$ in senso antiorario.

Segue infine da (2.5) che

$$z = \rho\cos\theta + i\rho\sin\theta \;\Rightarrow\; z^n = \rho^n\cos(n\theta) + i\rho^n\sin(n\theta)\,. \tag{2.6}$$

Esercizio 2.5 Calcolare, per ogni valore di θ nell'intervallo $[0, 2\pi)$, modulo e argomento di

$$\frac{1+\cos\theta - i\sin\theta}{1+\cos\theta + i\sin\theta}\,.$$

R: 1, $-\theta$.

2.6 Equazioni sui complessi, radici dell'unità

Un polinomio complesso è un'espressione del tipo

$$p(z) = a_0 + a_1 z + a_2 z^2 + \cdots + a_n z^n$$

dove $a_0, a_1, \ldots, a_n \in \mathbb{C}$ e z è una indeterminata che può prendere valori complessi. I polinomi complessi godono di proprietà algebriche analoghe al caso reale: si possono sommare, moltiplicare, dividere nello stesso modo. L'equazione

$$p(z) = 0\,, \tag{2.7}$$

come accadeva nel caso reale, non può avere più di n soluzioni (in effetti per ogni soluzione z_0 si ha che $z - z_0$ divide $p(z)$ ed essendo $p(z)$ di grado n non può avere più di n divisori di grado 1). Un teorema di importanza basilare, noto come il *teorema fondamentale dell'algebra*, assicura che (se $a_n \neq 0$) l'equazione sopra ha sempre esattamente n soluzioni (se contate correttamente con le loro molteplicità). Questo non accadeva sui reali: ad esempio $x^2 + 1 = 0$ non ha soluzioni su $\mathbb{R}$ pur essendo di grado 2. Quello che ora faremo sarà trovare le soluzioni di particolare equazioni polinomiali del tipo (2.7).

Cominciamo con l'osservare che l'espressione (2.3) dice che i è una radice quadrata di -1, essa cioè risolve, insieme a $-i$ l'equazione $z^2 + 1 = 0$. La stessa equazione che sui reali non aveva soluzioni ne ammette due sui complessi come previsto dal teorema fondamentale dell'algebra. L'esistenza di radici quadrate di numeri negativi è in effetti una fondamentale differenza rispetto al campo dei numeri reali dove tali radici non esistono. Si noti che se $\alpha > 0$ si ha che

$$z^2 + \alpha = (z + i\sqrt{\alpha})(z - i\sqrt{\alpha})\,.$$

Quindi le radici quadrate di $-\alpha$ cioè le soluzioni di $z^2 + \alpha = 0$ sono date da $\pm i\sqrt{\alpha}$.

Ci occupiamo ora di studiare l'equazione

$$z^n - 1 = 0\,. \tag{2.8}$$

Le sue soluzioni sono le radici n-esime di 1. Tra esse vi deve ovviamente essere sempre $z = 1$ e, se n è pari, anche $z = -1$. Queste sono le sole soluzioni reali. Per trovare le altre utilizziamo la forma trigonometrica e, in particolare, l'espressione (2.6). Sia $z = \rho\cos\theta + i\rho\sin\theta$, così che $z^n = \rho^n\cos(n\theta) + i\rho^n\sin(n\theta)$. Quindi,

$$z^n - 1 = 0 \;\Leftrightarrow\; \begin{cases} \rho^n = 1 \\ n\theta \in \{2k\pi \mid k = 0, \pm 1, \pm 2, \ldots\}\,. \end{cases}$$

Si ha quindi $\rho = 1$, mentre i possibili argomenti θ sono dati da

$$\theta_1 = 0,\ \theta_2 = \frac{2\pi}{n},\ \theta_3 = \frac{4\pi}{n}, \ldots, \theta_n = \frac{2(n-1)\pi}{n}\,.$$

Si noti che è inutile andare ulteriormente avanti con k: se $k = n$ si ottiene in effetti $2n\pi/n = 2\pi$ che rappresenta la stessa fase di $\theta_1 = 0$ e i successivi valori di k non farebbero che riottenere le fasi θ_i già trovate. Possiamo quindi scrivere esplicitamente le soluzioni di (2.8) che sono:

$$z_k = \cos\frac{2(k-1)\pi}{n} + i\sin\frac{2(k-1)\pi}{n}, \qquad k = 1, 2, \ldots, n\,.$$

Se $w \in \mathbb{C}$, possiamo ora anche trovare tutte le sue radici n-esime, cioè risolvere l'equazione

$$z^n - w = 0\,. \tag{2.9}$$

Scriviamo

$$w = \overline{\rho}\cos\overline{\theta} + i\overline{\rho}\sin\overline{\theta}, \qquad z = \rho\cos\theta + i\rho\sin\theta\,.$$

Usando (2.5) è facile trovare una soluzione particolare. Basta scegliere ρ e θ nel modo seguente

$$\begin{cases} \rho^n = \overline{\rho} \\ n\theta = \overline{\theta} \end{cases} \;\Rightarrow\; \begin{cases} \rho = \sqrt[n]{\overline{\rho}} \\ \theta = \frac{\overline{\theta}}{n}\,. \end{cases}$$

Dunque

$$z = \sqrt[n]{\rho}\cos\frac{\overline{\theta}}{n} + \sqrt[n]{\rho}\sin\frac{\overline{\theta}}{n}$$

è una soluzione di (2.9). Per trovare tutte le altre basta ora considerare le n radici n-esime dell'unità: $z_1, z_2, \ldots, z_n$. E' facile rendersi conto che $zz_1, zz_2, \ldots, zz_n$ sono le n soluzioni distinte di (2.9).

Esercizio 2.6 Si risolvano le equazioni

$$z^3 - 1 = 0, \quad z^4 + 1 = 0, \quad z^4 + i = 0, \quad z^3 + 8 = 0, \quad z^4 + 1 - i = 0$$

e se ne disegnino le rispettive soluzioni sul piano complesso.
R: 1, $-1/2 \pm i\sqrt{3}/2$; $\pm\sqrt{2}/2 \pm \sqrt{2}i/2$; $\cos(k\pi/8) + i\sin(k\pi/8)$, $k = 3, 7, 11, 15$; -2, $1 \pm \sqrt{3}i$; $\sqrt[8]{2}[\cos(k\pi/16) + i\sin(k\pi/16)]$, $k = 3, 11, 19, 27$.

Esercizio 2.7 Trovare tutte le soluzioni complesse dell'equazione ciclotomica:

$$z^n + z^{n-1} + \cdots + z + 1 = 0$$

(sugg.: si moltiplichi l'equazione per $z - 1$).
R: $\cos 2k\pi/(n+1) + i\sin 2k\pi/(n+1)$, $k = 1, \ldots, n$.

Esercizio 2.8 * Si risolva l'equazione

$$z^3 = i\overline{z}\,.$$

R: 0, $\cos k\pi/8 + i\sin k\pi/8$, $k = 1, 5, 9, 13$.

Capitolo 3

Successioni e serie

3.1 Successioni e definizione di limite

Nell'introdurre i numeri reali abbiamo spesso sfiorato il concetto di convergenza. In effetti la rappresentazione decimale stessa suggerisce questo concetto: se consideriamo il numero $x = \pm x_0, x_1 x_2 \cdots$, l'idea è che gli approssimanti decimali $x_n = \pm x_0, x_1 \cdots x_n$ si avvicinano al numero x quando n diventa sempre più grande. Avevamo anche notato come questo concetto di 'avvicinamento' fosse quello che ci voleva per introdurre correttamente la somma ed il prodotto di reali. Infine il concetto di convergenza è ritornato sulle questioni di approssimazione ed in particolare sul metodo iterativo introdotto per approssimare $\sqrt{2}$. In questo capitolo preciseremo in modo rigoroso il concetto di convergenza e ne studieremo molte sue proprietà.

Cominciamo col definire un concetto preliminare.

Definizione 3.1 Una *successione* (di numeri reali) è un' applicazione da $\mathbb{N}$ a $\mathbb{R}$, cioè una legge che associa ad ogni naturale n un ben determinato reale a_n. Per indicare la successione useremo la notazione $n \mapsto a_n$, o, la più compatta, (a_n).

Osservazione: Talvolta una successione (a_n) è definita per n appartenente ad un dominio diverso da $\mathbb{N}$; talvolta più grande come $\mathbb{N} \cup \{0\}$, talvolta più piccolo come ad esempio $\mathbb{N} \setminus \{1\}$.

Nel primo capitolo abbiamo già incontrato vari esempi di successioni:

Esempio 12 Sia $x = \pm k_0, k_1 k_2 \cdots$ un numero reale espresso in forma di allineamento decimale. Gli approssimanti decimali finiti formano una successione (x_n) dove

$$x_n = \pm k_0, k_1 k_2 \cdots k_n = \pm \sum_{i=0}^{n} \frac{k_i}{10^i} \qquad n = 0, 1, 2, \ldots$$

Esempio 13 Si consideri la successione degli approssimanti di $\sqrt{2}$ illustrati nel capitolo precedente. Tale successione (a_n) è determinata da

$$\begin{cases} a_{n+1} & = \dfrac{1}{2}\left(a_n + \dfrac{2}{a_n}\right) \\ a_0 & = 2 \end{cases}$$

A differenza dell'esempio precedente non si ha in questo caso una formula esplicita che permetta, dato n, di calcolare a_n; si ha invece una formula che permette di calcolare a_n se è noto a_{n-1}. Poichè a_0 è noto, da esso possiamo calcolare a_1. Da a_1 applicando di nuovo la formula possiamo calcolare a_2 e così via. Tali successioni vengono dette *successioni per ricorrenza.*

Si noti che una successione (a_n) è qualcosa di più dell'insieme dei valori che essa assume, cioè della sua immagine $\{a_n \mid n \in \mathbb{N}\}$: come in ogni applicazione, nel concetto di successione è codificata l'informazione di come tali valori vengono assunti. Successioni diverse possono benissimo assumere gli stessi valori come nel seguente esempio:

Esempio 14 Consideriamo le tre successioni (a_n), (b_n) e (c_n) date da

$$\begin{aligned} a_n &= (-1)^n, & n \in \mathbb{N}\,, \\ b_n &= (-1)^{n+1}, & n \in \mathbb{N}\,, \\ c_n &= \begin{cases} -1 & \text{se} \quad n \leq 7 \\ 1 & \text{se} \quad n \geq 8\,. \end{cases} \end{aligned}$$

Allora,

$$\{a_n \mid n \in \mathbb{N}\} = \{b_n \mid n \in \mathbb{N}\} = \{c_n \mid n \in \mathbb{N}\} = \{1, -1\}\,,$$

ma le tre successioni sono evidentemente diverse.

Vogliamo ora formalizzare il concetto intuitivo di una successione (a_n) che si avvicina ad un certo numero reale l quando n diventa sempre più grande.

Poichè $|a_n - l|$ è la distanza di a_n da l, l'idea è che questo dovrebbe diventare piccolo quanto vogliamo al crescere di n: cioè fissato un margine di errore qualunque $\epsilon > 0$ si dovrà avere che $|a_n - l| < \epsilon$ se n è abbastanza grande. Ecco dunque formalizzata la definizione fondamentale:

Definizione 3.2 Si dice che la successione (a_n) *converge* al numero $l \in \mathbb{R}$ se

$$\forall \epsilon > 0 \;\; \exists n_0 \in \mathbb{N} \; : \; \forall n \geq n_0 \text{ si ha } |a_n - l| < \epsilon \,. \tag{3.1}$$

l viene detto *limite* della successione (a_n) e si usano le notazioni equivalenti

$$\lim_{n \to +\infty} a_n = l, \qquad a_n \to l \,.$$

Osservazioni sulla definizione di limite: il concetto appena introdotto, di limite, è fondamentale e vale la pena di fare alcune considerazioni per chiarire a fondo la definizione data.

(A) La diseguaglianza $|a_n - l| < \varepsilon$ è equivalente a $l - \epsilon < a_n < l + \epsilon$ cioè al fatto che a_n si trova all'interno dell'intervallo $]l - \epsilon, l + \epsilon[$. Tali intervalli si dicono anche *intorni (centrati)* di l. La successione (a_n) converge ad l quindi se per ogni intorno fissato di l, esiste $n_0 \in \mathbb{N}$ tale che a_n sta in questo intorno se $n \geq n_0$.

(B) Si noti l'ordine dei vari quantificatori all'interno della definizione: 'per ogni $\epsilon > 0$ esiste $n_0 \in \mathbb{N}$ tale che...'. Ciò significa che n_0 è scelto sulla base della scelta fatta per ϵ ed in generale sarà tanto più grande quanto più piccolo è stato scelto ϵ.

(C) Dire che la successione (a_n) converge ad l è equivalente a dire che la successione $(a_n - l)$ converge a 0, o anche che la successione dei valori assoluti $(|a_n - l|)$ converge a 0. Questo dovrebbe risultare evidente dalla definizione, tuttavia è bene che lo studente ci rifletta per convincersene del tutto.

(D) Come vedremo, una successione può non ammettere limite. Dovrebbe risultare tuttavia chiaro che se tale limite esiste esso è unico: una sucessione non può 'avvicinarsi' contemporaneamente a due numeri distinti. Omettiamo una dimostrazione rigorosa di questo fatto intuitivo.

Prima di procedere oltre con esempi conviene introdurre un linguaggio utile per trattare proprietà delle successioni come la convergenza.

Definizione 3.3 Data una succesione (a_n) diremo che una certa proprietà P è valida *definitivamente* per (a_n) se esiste $n_0 \in \mathbb{N}$ tale che la successione gode della proprietà P per $n \geq n_0$.

Chiariamo la definizione con qualche esempio concreto:

Esempio 15

- La successione (a_n) data da $a_n = n - 5$ è definitivamente positiva in quanto $a_n > 0$ per $n \geq 6$.
- La successione (a_n) data da $a_n = 1/n$ è definitivamente minore di $0,1$ in quanto effettivamente $1/n < 0,1$ se $n > 10$.

Con tale linguaggio la definizione di convergenza può essere equivalentemente formulata come segue:

- Una successione (a_n) converge ad l se per ogni $\epsilon > 0$ si ha che $|a_n - l| < \epsilon$ definitivamente.

oppure, considerando anche l'osservazione (A) precedente

- Una successione (a_n) converge ad l se per ogni intorno centrato di l si ha che a_n definitivamente sta in questo intorno.

Queste riformulazioni mettono anche in luce un altro importante aspetto della convergenza: se (a_n) e (b_n) sono due successioni definitivamente uguali, allora una converge se e soltanto se converge l'altra ed il limite è lo stesso.

Presentiamo ora alcuni semplici esempi.

Esempio 16 Sia $c \in \mathbb{R}$ e sia (a_n) data da $a_n = c$ per ogni n. Tale successioni sono dette *costanti*. Chiaramente si ha che $a_n \to c$. Lo stesso accade se la successione soltanto definitivamente eguale a c.

Esempio 17 Sia (a_n) data da $a_n = 1/n$. Facciamo vedere che $a_n \to 0$. Fissiamo $\epsilon > 0$. Dobbiamo far vedere che $|a_n - 0| = 1/n < \epsilon$ definitivamente. In effetti,

$$\frac{1}{n} < \epsilon \iff n > \frac{1}{\epsilon}\,.$$

Questo dimostra la nostra tesi

Esempio 18 Sia (a_n) data da $a_n = \frac{n+2}{n+3}$. Facciamo vedere che $a_n \to 1$. Fissiamo $\epsilon > 0$. Dobbiamo far vedere che $|a_n - 1| = |\frac{n+2}{n+3} - 1| < \epsilon$ definitivamente. Si ha

$$\left|\frac{n+2}{n+3} - 1\right| = \left|-\frac{1}{n+3}\right| = \frac{1}{n+3}$$

e

$$\frac{1}{n+3} < \epsilon \iff n+3 > \frac{1}{\epsilon} \iff n > \frac{1}{\epsilon} - 3\,.$$

Questo dimostra la nostra tesi.

Intuitivamente il limite di una successione è quel numero, se esiste, al quale i valori della successione si avvicinano. Questo implica un rapporto tra il segno del limite e quello dei valori della successione come mostra il seguente.

Teorema 3.4 (della permanenza del segno) *Sia (a_n) una successione convergente ad un limite $l \neq 0$. Allora, definitivamente, il segno di a_n e di l sono identici.*

Dimostrazione Supponiamo $l > 0$ (l'altro caso si dimostra in modo analogo). Fissiamo $\epsilon = l/2$. Per la definizione di limite definitivamente si ha che $l - l/2 < a_n < l + l/2$. In particolare si ha che $a_n > l/2$ definitivamente. ∎

Corollario 3.5 *Sia (a_n) una successione che converge ad l e tale che, definitivamente, $a_n \geq 0$ (rispettivamente $a_n \leq 0$). Allora, $l \geq 0$ (rispettivamente $l \leq 0$).*

Dimostrazione Supponiamo che $a_n \geq 0$ definitivamente. Se per assurdo si avesse $l < 0$, per il Teorema della permanenza del segno, si avrebbe $a_n < 0$ definitivamente. Questo non può essere, e dunque $l \geq 0$. L'altro caso si dimostra analogamente. ∎

3.2 Limiti: prime proprietà

Per poter determinare i limiti delle successioni con una certa disinvoltura, senza ogni volta dover ricorrere alla definizione, è necessario sviluppare un po' di teoria. Iniziamo con un risultato chiave:

Teorema 3.6 (del confronto) *Siano (a_n), (b_n) e (c_n) tre successioni tali che*

$$a_n \leq b_n \leq c_n \quad \forall n \in \mathbb{N}\,. \tag{3.2}$$

Supponiamo inoltre che (a_n) e (c_n) convergano con

$$\lim_{n\to+\infty} a_n = l = \lim_{n\to+\infty} c_n\,. \tag{3.3}$$

Allora anche (b_n) risulta convergente e

$$\lim_{n\to+\infty} b_n = l\,.$$

Dimostrazione Fissiamo un numero $\epsilon > 0$. Per l'ipotesi (3.3) si ha che

$$\begin{array}{ll} l-\epsilon < a_n < l+\epsilon & \text{definitivamente} \\ l-\epsilon < c_n < l+\epsilon & \text{definitivamente}\,; \end{array}$$

il che significa (si pensi perchè) che, definitivamente, sono simultaneamente vere

$$l-\epsilon < a_n < l+\epsilon, \qquad l-\epsilon < c_n < l+\epsilon\,.$$

Usando ora (3.2), e la prima e la quarta delle diseguaglianze sopra si ha che, definitivamente,

$$l-\epsilon < a_n \leq b_n \leq c_n < l+\epsilon$$

il che significa che

$$l-\epsilon < b_n < l+\epsilon \quad \text{definitivamente}\,.$$

■

Osservazione: L'ipotesi (3.2) del Teorema 3.6 può essere indebolita richiedendo che essa valga soltanto definitivamente.

Esempio 19 Sia $x \in \mathbb{R}$ e si consideri la successione (x^n). Essa è detta *progressione geometrica*. Verifichiamo qui che se $|x| < 1$, allora $x^n \to 0$. Si noti che se $x = 0$ questo è ovvio in quanto allora $x^n = 0$ per ogni n. Supponiamo dunque $x \neq 0$. Poichè $|x| < 1$ si ha che $1/|x| > 1$. Si può quindi scrivere $1/|x| = 1 + \delta$ per un qualche numero $\delta > 0$. Allora

$$\frac{1}{|x|^n} = (1+\delta)^n \geq 1 + n\delta$$

(l'ultima diseguaglianza è stata dimostrata nella Proposizione A.6). Quindi,

$$0 \leq |x|^n \leq \frac{1}{1+n\delta}\,.$$

Poichè $1/(1 + n\delta) \to 0$ (lo si verifichi direttamente) si ha che, per il Teorema del confronto, $|x|^n \to 0$ e anche $x^n \to 0$. La verifica del fatto che $x^n \to 0$ poteva anche essere fatta direttamente utilizzando i logaritmi e le loro proprietà; si provi a farlo per esercizio. Più avanti studieremo la progressione geometrica anche per gli altri valori di x.

Esempio 20 Consideriamo un numero reale $x = \pm k_0, k_1 k_2 \cdots$ e la successione delle sue approssimazioni decimali finite (x_n) date da $x_n = \pm k_0, k_1 k_2 \cdots k_n$. Allora $x_n \to x$. In effetti si ha che

$$0 \leq |x - x_n| \leq 10^{-n}\,.$$

La successione costantemente uguale a 0 converge a 0 così come la successione (10^{-n}) (vedi Esempio 19). Quindi in virtù del Teorema del confronto $|x - x_n| \to 0$ o, equivalentemente (vedi Osservazione (C) dopo la definizione di limite) $x_n \to x$.

Esempio 21 Consideriamo la successione per ricorrenza (a_n) dell'Esempio 13. Nella Sezione 1.6 del capitolo precedente è stato mostrato che

$$0 \leq a_n - \sqrt{2} \leq \frac{1}{2^n} \qquad \forall n \in \mathbb{N}\,.$$

Segue quindi dal Teorema del confronto che $a_n - \sqrt{2} \to 0$ e quindi che $a_n \to \sqrt{2}$.

Introduciamo un altro concetto importante:

Definizione 3.7 Una successione (a_n) si dice *limitata* se lo è la sua immagine, cioè se esiste $L \in \mathbb{R}$ tale che $|a_n| \leq L$ per ogni $n \in \mathbb{N}$.

La limitatezza è, a differenza del concetto di limite, una proprietà che riguarda soltanto l'immagine della successione che è un sottoinsieme di numeri reali. Vale però il seguente fatto:

Proposizione 3.8 *Sia (a_n) una successione che ammette limite, allora essa è limitata.*

Dimostrazione Supponiamo che $a_n \to l$. Allora per la definizione di limite si ha che esiste $n_0 \in \mathbb{N}$ tale che $l - 1 < a_n < l + 1$ per ogni $n \geq n_0$. In particolare si ha che $|a_n| < |l| + 1$ per ogni $n \geq n_0$. Si ha quindi sicuramente

$$|a_n| \leq \max\{|a_1|, |a_2|, \ldots, |a_{n_0-1}|, |l| + 1\}\,.$$

■

Esempio 22 Si consideri nuovamente la progressione geometrica (x^n) questa volta supponendo $|x| > 1$. Essa non è limitata; in effetti, fissato un qualunque $M > 0$, la diseguaglianza $|x^n| > M$ è equivalente, per le proprietà dei logaritmi, a $n > \frac{\log M}{\log |x|}$ e quindi ci sono sempre soluzioni in $\mathbb{N}$. In particolare, la progressione geometrica x^n con $|x| > 1$ non ammette limite. Gli unici valori di x per i quali non abbiamo ancora studiato la convergenza sono rimasti ± 1. Per quanto riguarda il caso $x = 1$, esso è banale in quanto in tal caso $x^n = 1$ per ogni n. Il caso $x = -1$ sarà considerato nell'Esempio 24.

Esempio 23 Sia (a_n) data da $a_n = n^2 - n$. (a_n) non è limitata: in effetti se fissiamo un qualunque $M > 0$ e consideriamo la diseguaglianza $x^2 - x > M$, essa è in particolare risolta da $x > (1 + \sqrt{4M})/2$. Quindi se $n \in \mathbb{N}$ è tale che $n > (1 + \sqrt{4M})/2$, si ha che $a_n > M$. Quindi (a_n) non può neppure ammettere limite.

Il risultato inverso della Proposizione 3.8 non vale: ci sono successioni limitate che non ammettono limite come mostra il seguente:

Esempio 24 Sia (a_n) data da $a_n = (-1)^n$. Allora (a_n) è chiaramente limitata. Tuttavia essa non ammette limite: con l'aumentare di n, a_n continua ad oscillare tra -1 e 1. Questa è la spiegazione intuitiva della mancanza di limite, il ragionamento può comunque essere reso rigoroso nel modo seguente. Supponiamo per assurdo che esista un numero $l \in \mathbb{R}$ tale che $a_n \to l$. Fissiamo $\epsilon = 1/2$. Allora deve esistere $n_0 \in \mathbb{N}$ tale che

$$|a_n - l| < \frac{1}{2} \quad \forall n \geq n_0 \,. \tag{3.4}$$

Siano ora $n_1, n_2 \geq n_0$. Si ha che

$$|a_{n_1} - a_{n_2}| = |a_{n_1} - l + l - a_{n_2}| \leq |a_{n_1} - l| + |a_{n_2} - l| < \frac{1}{2} + \frac{1}{2} = 1$$

(la prima diseguaglianza segue dalla proprietà (Q7) del Capitolo 1, la seconda dalla (3.4)). Questo però è assurdo poichè se $n_2 = n_1 + 1$, si ha che $|a_{n_1} - a_{n_2}| = 2$. Questo ragionamento mostra che (a_n) non può possedere limite.

Si noti come la successione (c_n) dell'Esempio 14, che ha la stessa immagine della successione (a_n) dell'Esempio 24 sopra, sia definitivamente uguale ad 1 e che quindi ammetta limite uguale ad 1. Questo mostra come la proprietà di limite, a differenza della limitatezza, dipenda in modo essenziale dalla successione e non meramente dalla sua immagine.

C'è un altro concetto di limite che riguarda successioni non limitate, e che è importante introdurre:

Definizione 3.9 Sia (a_n) una successione.

- Si dice che (a_n) tende a $+\infty$ se fissato un qualunque $M \in \mathbb{R}$ si ha che $a_n > M$ definitivamente. Si usano le notazioni:

$$\lim_{n\to+\infty} a_n = +\infty, \qquad a_n \to +\infty\,.$$

- Si dice che (a_n) tende a $-\infty$ se fissato un qualunque $M \in \mathbb{R}$ si ha che $a_n < M$ definitivamente. Si usano le notazioni:

$$\lim_{n\to+\infty} a_n = -\infty, \qquad a_n \to -\infty\,.$$

Successioni che ammettono questo tipo di limite sono già apparse negli esempi precedenti. Ad esempio la progressione geometrica (x^n) con $x > 1$ tende a $+\infty$ come mostra il ragionamento fatto nell'Esempio 22. Similmente tende a $+\infty$ la successione trattata nell'Esempio 23. D'altra parte non è vero che ogni successione illimitata necessariamente tenda a $+\infty$ o a $-\infty$. Un esempio è il seguente

Esempio 25 Si consideri ancora la progressione geometrica (x^n) con $x < -1$. Si noti che se tendesse a $+\infty$ dovrebbe essere definitivamente positiva, se invece tendesse a $-\infty$ dovrebbe ssere definitivamente negativa; poichè essa assume valori positivi per n pari, negativi per n dispari nessuna delle due possibilità si può verificare.

Qui sotto, per comodità, riassumiamo il comportamento al limite della progressione geometrica quale si deriva dagli Esempi 19, 22 e 25:

$$\lim_{n\to+\infty} x^n = \begin{cases} 0 & \text{se} \quad |x| < 1 \\ 1 & \text{se} \quad x = 1 \\ +\infty & \text{se} \quad x > 1 \\ \nexists & \text{se} \quad x \le -1\,. \end{cases} \tag{3.5}$$

Anche per i limiti infiniti si ha una versione del teorema del confronto la cui dimostrazione è lasciata per esercizio.

Teorema 3.10 (del confronto) *Siano (a_n) e (b_n) due successioni tali che definitivamente*

$$a_n \le b_n\,. \tag{3.6}$$

Allora,

$$\lim_{n\to+\infty} a_n = +\infty \quad \Rightarrow \quad \lim_{n\to+\infty} b_n = +\infty\,,$$
$$\lim_{n\to+\infty} b_n = -\infty \quad \Rightarrow \quad \lim_{n\to+\infty} a_n = -\infty\,.$$

Esempio 26 Consideriamo la successione (a_n) data da $a_n = n!$. Si ha che $n! \geq n$ per ogni $n \in \mathbb{N}$. Poichè $n \to +\infty$ (lo si verifichi), per il Teorema del confronto anche $n! \to +\infty$.

3.3 Limiti: proprietà algebriche

Siano (a_n) e (b_n) due successioni. Si può considerare la successione somma (a_n+b_n), la successione prodotto $(a_n \cdot b_n)$ e, se $b_n \neq 0$ per ogni n, la successione quoziente (a_n/b_n). Il comportamento al limite di tali successioni segue le regole dell'algebra:

Proposizione 3.11 *Siano (a_n) e (b_n) due successioni convergenti: $a_n \to l_1$ e $b_n \to l_2$. Allora*

$$a_n + b_n \to l_1 + l_2, \qquad a_n \cdot b_n \to l_1 \cdot l_2\,.$$

Inoltre se $b_n \neq 0$ per ogni n, e $l_2 \neq 0$, si ha anche

$$a_n/b_n \to l_1/l_2\,.$$

Dimostrazione La dimostrazione è piuttosto noiosa e priva di idee interessanti. Per dare un'idea dimostreremo che $a_n \cdot b_n \to l_1 \cdot l_2$ che è la più delicata. Le altre sono lasciate allo studente. Si ha

$$\begin{aligned} |a_n b_n - l_1 l_2| &= |a_n b_n - a_n l_2 + a_n l_2 - l_1 l_2| \\ &\leq |a_n b_n - a_n l_2| + |a_n l_2 - l_1 l_2| \\ &= |a_n||b_n - l_2| + |a_n - l_1||l_2|\,. \end{aligned} \tag{3.7}$$

Fissiamo ora $\epsilon > 0$. Poichè (a_n) per ipotesi converge essa è limitata:

$$|a_n| \leq L \qquad \forall n \in \mathbb{N}\,. \tag{3.8}$$

D'altra parte poichè $a_n \to l_1$ e $b_n \to l_2$ si ha che

$$\begin{aligned} |a_n - l_1| &< \frac{\epsilon}{2(|l_2|+1)} \quad \text{definitivamente}\,, \\ |b_n - l_2| &< \frac{\epsilon}{2L} \qquad\quad \text{definitivamente}\,. \end{aligned} \tag{3.9}$$

Di nuovo per il significato stesso di definitivamente si ha che le tre diseguaglianze precedenti (3.8) e (3.9) sono simultaneamente vere, definitivamente. Usandole insieme a (3.7) si ottiene quindi che, definitivamente,

$$|a_n b_n - l_1 l_2| < L\frac{\epsilon}{2L} + \frac{\epsilon}{2(|l_2|+1)}|l_2| \leq \frac{\epsilon}{2} + \frac{\epsilon}{2} = \epsilon\,.$$

La dimostrazione è così completata. ■

Le successioni che convergono a zero vengono anche dette *infinitesime*. Somma e prodotto di funzioni infinitesime sono infinitesime in virtù della Proposizione 3.11. In realtà per quanto riguarda il prodotto si può dire qualcosa di più:

Proposizione 3.12 *Sia (a_n) una successione infinitesima e sia (b_n) una successione limitata. Allora la successione prodotto $(a_n b_n)$ è infinitesima.*

Dimostrazione Per l'ipotesi di limitatezza, esiste $L \geq 0$ tale che $|a_n| \leq L$ per ogni $n \in \mathbb{N}$. Si ha quindi

$$|a_n b_n| = |a_n||b_n| \leq L|a_n|\,.$$

D'altra parte, poichè $a_n \to 0$, fissato $\epsilon > 0$ si ha che, definitivamente,

$$|a_n| < \epsilon/L\,.$$

Le due diseguaglianze implicano che $|a_n b_n| < L(\epsilon/L) = \epsilon$ definitivamente. ■

Esempio 27 Si consideri la successione (a_n) data da $a_n = \sin n/n$. Poichè $(\sin n)$ è limitata ($|\sin n| \leq 1$), mentre $(1/n)$ è infinitesima, ne segue che, in virtù della Proposizione 3.12, anche la successione (a_n) è infinitesima.

Si noti che poichè non ogni successione limitata ammette limite, la Proposizione 3.12 non può essere direttamente derivata dalla Proposizione 3.11. Si noti inoltre che la limitatezza della (b_n) è fondamentale come mostra il seguente

Esempio 28 Siano (a_n) e (b_n) due successioni date da $a_n = 1/n$ e $b_n = n$. (a_n) è infinitesima, ma (b_n) è evidentemente non limitata. Il prodotto è dato da $a_n b_n = 1$ per ogni n. In particolare, $(a_n b_n)$ non è infinitesima.

Anche per le successioni che tendono a $\pm\infty$ ci sono una serie di risultati di tipo algebrico la dimostrazione dei quali è lasciata per esercizio.

Proposizione 3.13 *Siano (a_n) e (b_n) due successioni. Allora,*

$$\begin{array}{lllll}
a_n \to +\infty, & b_n \to l \text{ (finito o } +\infty) & \Longrightarrow & a_n + b_n \to +\infty\,, \\
a_n \to -\infty, & b_n \to l \text{ (finito o } -\infty) & \Longrightarrow & a_n + b_n \to -\infty\,, \\
a_n \to +\infty, & b_n \to l\ (> 0 \text{ o } +\infty) & \Longrightarrow & a_n.b_n \to +\infty\,, \\
a_n \to +\infty, & b_n \to l\ (< 0 \text{ o } -\infty) & \Longrightarrow & a_n.b_n \to -\infty\,, \\
a_n \to \pm\infty, & & \Longrightarrow & 1/a_n \to 0\,, \\
a_n \to 0, & a_n > 0 & \Longrightarrow & 1/a_n \to +\infty\,.
\end{array}$$

Come per le successioni infinitesime, anche per quelle che tendono a $\pm\infty$ si possono avere risultati algebrici più forti sul tipo di quello espresso nella Proposizione 3.12. Alcuni di questi sono presentati nella proposizione seguente: anche di questa le dimostrazioni sono lasciate per esercizio.

Proposizione 3.14 *Siano (a_n) e (b_n) due successioni. Allora,*

$$\begin{array}{llll}
a_n \to +\infty, & b_n \text{ (infer. limitata)} & \Longrightarrow & a_n + b_n \to +\infty\,, \\
a_n \to -\infty, & b_n \text{ (super. limitata)} & \Longrightarrow & a_n + b_n \to -\infty\,, \\
a_n \to +\infty, & b_n \geq \delta > 0 & \Longrightarrow & a_n.b_n \to +\infty\,, \\
a_n \to +\infty, & b_n \leq \delta < 0 & \Longrightarrow & a_n.b_n \to -\infty\,.
\end{array}$$

Si noti come nè la Proposizione 3.11, nè la Proposizione 3.13 diano alcuna informazione sulle situazioni elencate sotto:

$$\begin{array}{llll}
a_n \to 0, & b_n \to 0 & \Longrightarrow & a_n/b_n \to ? \\
a_n \to \pm\infty, & b_n \to \pm\infty & \Longrightarrow & a_n/b_n \to ? \\
a_n \to \pm\infty, & b_n \to 0 & \Longrightarrow & a_n.b_n \to ? \\
a_n \to +\infty, & b_n \to -\infty & \Longrightarrow & a_n + b_n \to ?
\end{array}$$

Le situazioni sopra vengono solitamente dette *forme di indeterminazione* o *di indecisione*. In questi casi può effettivamente accadere qualunque cosa: il limite a destra non esistere od esistere ed essere un qualunque numero. Non c'è una tecnica generale per 'risolvere' queste indeterminazioni; molte indeterminazioni notevoli saranno studiate più avanti. Presentiamo qui alcuni esempi elementari:

Esempio 29 Consideriamo due polinomi

$$\begin{aligned}
p(x) &= p_0 + p_1 x + \cdots + p_r x^r\,, \\
q(x) &= q_0 + q_1 x + \cdots + q_s x^s
\end{aligned}$$

con p_r e q_s diversi da 0. Supponiamo inoltre che $q(n) \neq 0$ per ogni $n \in \mathbb{N}$. Si consideri la successione (a_n) data da $a_n = p(n)/q(n)$. Vogliamo calcolarne il limite:

$$\begin{aligned}\frac{p(n)}{q(n)} = \frac{p_0 + p_1 n + \cdots + p_r n^r}{q_0 + q_1 n + \cdots + q_s n^s} &= \frac{n^r \left(\frac{p_0}{n^r} + \frac{p_1}{n^{r-1}} + \cdots + p_r\right)}{n^s \left(\frac{q_0}{n^s} + \frac{q_1}{n^{s-1}} + \cdots + q_s\right)} \\ &= n^{r-s} \frac{\frac{p_0}{n^r} + \frac{p_1}{n^{r-1}} + \cdots + p_r}{\frac{q_0}{n^s} + \frac{q_1}{n^{s-1}} + \cdots + q_s} .\end{aligned}$$

Si noti ora che tutti i termini del tipo a/n^i con $i > 0$ convergono a 0 in virtù del fatto che sono il prodotto di i successioni $1/n$ che convergono a 0 (vedi Esempio 17) e della successione costante (a) che converge ad a (vedi Esempio 16). Quindi,

$$\frac{p_0}{n^r} + \frac{p_1}{n^{r-1}} + \cdots + p_r \to p_r, \qquad \frac{q_0}{n^s} + \frac{q_1}{n^{s-1}} + \cdots + q_s \to q_s .$$

D'altra parte:

$$n^{r-s} \to \begin{cases} 0 & \text{se} \quad r < s \\ 1 & \text{se} \quad r = s \\ +\infty & \text{se} \quad r > s . \end{cases}$$

Utilizzando ancora i risultati sul comportamento algebrico dei limiti si può concludere che

$$\lim_{n\to+\infty} \frac{p_0 + p_1 n + \cdots + p_r n^r}{q_0 + q_1 n + \cdots + q_s n^s} = \begin{cases} 0 & \text{se} \quad r < s \\ p_r/q_s & \text{se} \quad r = s \\ +\infty & \text{se} \quad r > s \text{ e } p_r/q_s > 0 \\ -\infty & \text{se} \quad r > s \text{ e } p_r/q_s < 0 . \end{cases}$$

Esempio 30 Consideriamo la successione (a_n) definita da

$$a_n = \frac{n!}{n^n} .$$

Abbiamo visto nell'Esempio 26 che $n! \to +\infty$. Poichè chiaramente vale

$$n^n = n \cdot n \cdots n \geq n \cdot (n-1) \cdots 1 = n! ,$$

ne segue che per confronto anche $n^n \to +\infty$. La nostra successione presenta dunque una forma di indeterminazione del tipo ∞/∞. Si noti che si può stimare:

$$\frac{n!}{n^n} = \frac{n \cdot (n-1) \cdots 2 \cdot 1}{n \cdot n \cdots n \cdot n} = \frac{n}{n} \frac{n-1}{n} \cdots \frac{2}{n} \frac{1}{n} \leq 1 \cdot 1 \cdots 1 \cdot \frac{1}{n} = \frac{1}{n} .$$

Si ha dunque
$$0 \leq a_n \leq \frac{1}{n}\,.$$
Poichè $1/n \to 0$ segue dal Teorema 3.6 del confronto che anche $a_n \to 0$.

Esercizio 3.1 Per ciascuna delle successioni (a_n) definite sotto, dire se essa ammette limite finito o infinito, ed in caso affermativo calcolarlo:

$$\begin{aligned}
a_n &= \frac{1}{n!}\,, & a_n &= \frac{n!}{(2n)!}\,, & a_n &= \frac{n^7-1}{n^6+1}\,,\\
a_n &= \frac{4^n-2^n}{4^n+2^n}\,, & a_n &= \frac{n^2\sin n+3n^3}{1+n^3}\,, & a_n &= (1+(-1)^n)n^2\,,\\
a_n &= n^3+(-1)^n n^2\,, & a_n &= \frac{n^3+1}{n!}\,, & a_n &= n2^n+\cos(n^2)\,.
\end{aligned}$$

R: (da sinistra a destra e dall'alto in basso) 0, 0, $+\infty$, 1, 3, $\nexists$, $+\infty$, 0, $+\infty$.

Esercizio 3.2 Sia (a_n) una successione tale che $a_n \geq 0$ per ogni n e tale che $a_n \to l$. Si mostri che allora $\sqrt{a_n} = \sqrt{l}$.

Esercizio 3.3 * Si dimostri che, se $\epsilon > 0$ si ha
$$(1+\varepsilon)^n \geq \frac{n(n-1)}{2}\epsilon^2 \quad \forall n \in \mathbb{N}\,.$$
Si utilizzi poi la diseguaglianza precedente per dimostrare che
$$\lim_{n\to+\infty} \sqrt[n]{n} = 1\,.$$

3.4 Successioni monotone

Introduciamo ora una classe molto importante di successioni:

Definizione 3.15 Una successione (a_n) si dice

- *crescente* se $a_{n+1} \geq a_n$ per ogni n;
- *strettamente crescente* se $a_{n+1} > a_n$ per ogni n;
- *decrescente* se $a_{n+1} \leq a_n$ per ogni n;

- *strettamente decrescente* se $a_{n+1} \leq a_n$ per ogni n.

Le successioni crescenti o decrescenti vengono anche dette *monotone*.

Esempi di successioni monotone sono già apparsi precedentemente. In effetti si può vedere che, ad esempio, la successione nell'Esempio 17 è decrescente, mentre quella nell'Esempio 18 è crescente. La progressione geometrica (x^n) risulta crescente se $x \geq 1$, decrescente se $0 \leq x \leq 1$; non è invece monotona se $x < 0$.

Il seguente teorema esprime la proprietà fondamentale delle successioni monotone.

Teorema 3.16 *Sia (a_n) una successione monotona. Allora, (a_n) ammette sempre limite: finito se è limitata, $\pm\infty$ se è illimitata. Si ha, più precisamente,*

$$\begin{aligned} \lim_{n\to+\infty} a_n &= \sup\{a_n \mid n \in \mathbb{N}\} \qquad \textit{se } (a_n) \textit{ è crescente}, \\ \lim_{n\to+\infty} a_n &= \inf\{a_n \mid n \in \mathbb{N}\} \qquad \textit{se } (a_n) \textit{ è decrescente}. \end{aligned} \tag{3.10}$$

Dimostrazione Dimostreremo il teorema nel caso in cui (a_n) è crescente, essendo l'altro caso del tutto analogo. Supponiamo prima che (a_n) sia limitata e denotiamo $L = \sup\{a_n \mid n \in \mathbb{N}\}$. Fissiamo $\epsilon > 0$. Dobbiamo far vedere che, definitivamente

$$L - \epsilon < a_n < L + \epsilon\,. \tag{3.11}$$

Si noti innazitutto che $a_n \leq L$ per ogni n e quindi, a maggior ragione, $a_n < L + \epsilon$ per ogni n. D'altra parte, per la caratterizzazione dell'estremo superiore data dalla Proposizione 1.6, esiste sicuramente n_0 tale che $L - \epsilon < a_{n_0}$. Poichè (a_n) è crescente, ne segue che $L - \epsilon < a_n$ per ogni $n \geq n_0$. Si ha quindi che (3.11) è vera per ogni $n \geq n_0$.

Se invece (a_n) non è limitata, ne segue che, essendo crescente, deve necessariamente essere superiormente non limitata. Ne segue che, fissato un qualunque $M \in \mathbb{R}$, esiste $n_0 \in \mathbb{N}$ tale che $a_{n_0} > M$. D'altra parte, poichè è crescente si ha che $a_n \geq a_{n_0} > M$ per ogni $n \geq n_0$. Questo mostra che, in questo caso, $a_n \to +\infty$. Questo conclude la dimostrazione. ∎

Osservazione: Se indeboliamo le ipotesi del Teorema 3.16 supponendo che la (a_n) sia definitivamente monotona, oltre che limitata, si ha che ancora essa ammette limite. Tuttavia le relazioni (3.10) in generale non varranno più.

Esempio 31 Consideriamo la successione per ricorrenza (a_n) dell'Esempio 13. Avevamo notato che $a_n \geq \sqrt{2}$ per ogni n. Faremo ora vedere che (a_n) è monotona decrescente, dando alla fine una dimostrazione alternativa a quella nell'Esempio 21 del fatto che essa converge a $\sqrt{2}$. Mostriamo innanzitutto che

$$x \geq \sqrt{2} \Rightarrow x \geq \frac{1}{2}\left(x + \frac{2}{x}\right). \tag{3.12}$$

In effetti:

$$x \geq \frac{1}{2}\left(x + \frac{2}{x}\right) \iff 2x \geq x + \frac{2}{x} \iff x \geq \frac{2}{x} \iff x^2 \geq 2\,.$$

Da (3.12), ricordando il fatto, mostrato nel capitolo precedente, che $a_n \geq \sqrt{2}$ per ogni n, segue che

$$a_n \geq \frac{1}{2}\left(a_n + \frac{2}{a_n}\right) = a_{n+1} \qquad \forall n \in \mathbb{N}\,.$$

Poichè (a_n) è monotona decrescente e inferiormente limitata ($a_n \geq \sqrt{2}$), si ha che essa ammette limite finito $l \geq \sqrt{2}$. Segue dalle proprietà algebriche dei limiti che

$$\frac{1}{2}\left(a_n + \frac{2}{a_n}\right) \to \frac{1}{2}\left(l + \frac{2}{l}\right),$$

mentre

$$a_{n+1} \to l\,.$$

Deve quindi essere:

$$l = \frac{1}{2}\left(l + \frac{2}{l}\right).$$

E' facile vedere che l'equazione sopra ha due soluzioni : $\pm\sqrt{2}$. Poichè $l \geq \sqrt{2}$, si ha che necessariamente $l = \sqrt{2}$.

Esercizio 3.4 Per ciascuna delle successioni definite sotto si dica se è monotona, definitivamente monotona, limitata:

$$a_n = \frac{n}{2^n}\,, \qquad a_n = \frac{4^n}{n!}\,, \qquad a_n = (-1)^n\,,$$

$$a_n = \frac{n^2}{n^3+1}\,, \qquad a_n = 3^n - 2^n\,, \qquad a_n = \sqrt[n]{n}\,.$$

R: (da sinistra a destra e dall'alto al basso) mon. lim.; def.mon., lim.; lim.; mon., lim.; mon.; def.mon., lim..

Esercizio 3.5 * Si consideri la successione definita per ricorrenza:

$$\begin{cases} a_{n+1} &= \dfrac{1}{2}\left(a_n + \dfrac{b}{a_n}\right) \\ a_0 &= b \end{cases}$$

dove $b > 1$. Si dimostri che $a_n \to \sqrt{b}$. (Si suggerisce di dimostrare preventivamente che $a_n \geq \sqrt{b}$ per ogni n, poi di far vedere che (a_n) è decrescente, infine di determinare il limite).

3.5 Limiti notevoli

In questa parte, presentiamo alcuni importanti limiti di successioni. Per ottenerli, avremo bisogno di sviluppare qualche tecnica più sofisticata per il calcolo dei limiti che presenta notevole interesse di per sè. Queste tecniche risultano molto utili come vedremo quando si è in presenza di forme di indeterminazione che i teoremi di algebrici non permettono di analizzare.

Sia (a_n) una successione tale che $a_n > 0$ per ogni n e consideriamo la successione dei rapporti (q_n) definita da:

$$q_n = \frac{a_{n+1}}{a_n}, \quad n \in \mathbb{N}.$$

L'analisi della successione (q_n) permette di ottenere utili informazioni sulla successione di partenza (a_n). Si noti intanto che

$$a_{n+1} \geq a_n \iff q_n \geq 1.$$

Ne segue che (a_n) è monotona crescente se e soltanto se $q_n \geq 1$ per ogni n. Similmente, (a_n) è monotona decrescente se e solo se $q_n \leq 1$ per ogni n. Il seguente risultato mette in relazione i comportamenti limite delle due successioni (a_n) e (q_n).

Proposizione 3.17 *Supponiamo che esista il limite*

$$R = \lim_{n\to+\infty} q_n.$$

Valgono le affermazioni seguenti:

(i) Se $R < 1$, allora $a_n \to 0$.

(ii) Se $R > 1$ (incluso anche $R = +\infty$), allora $a_n \to +\infty$.

Dimostrazione (i): Si noti innanzitutto che $q_n - 1 \to R - 1 < 0$. Per il Teorema 3.4 della permanenza del segno, ne segue che, definitivamente $q_n - 1 < 0$, cioè $q_n < 1$. In virtù delle considerazioni precedenti ne segue che (a_n) è definitivamente monotona decrescente; quindi, per il Teorema 3.16, ammette limite $l \geq 0$ (essendo $a_n > 0$ per ogni n). Rimane da far vedere che $l = 0$. Se per assurdo si avesse $l > 0$, si avrebbe, applicando le regole di calcolo algebrico dei limiti (Proposizione 3.11) che

$$q_n = \frac{a_{n+1}}{a_n} \to \frac{l}{l} = 1\,,$$

che contraddice le ipotesi fatte essendo $R < 1$. Ne segue che deve essere quindi $l = 0$.

(ii): si consideri la successione reciproco (b_n): $b_n = 1/a_n$. Si ha che

$$\frac{b_{n+1}}{b_n} = \frac{a_n}{a_{n+1}} = \frac{1}{q_n} \to \frac{1}{R} < 1\,.$$

Per il punto (i) ne segue che $b_n \to 0$. In virtù della Proposizione 3.13 ne segue che $b_n \to +\infty$. ∎

Osservazione: Nel caso in cui $R = 1$ non si può concludere alcunchè sul comportamento asintotico della successione (a_n). Si noti infatti che le successioni (a_n) e (b_n) definite da $a_n = n$ e $b_n = 1/n$ hanno comportamento asintotici del tutto diversi. Per entrambe tuttavia la successione dei quozienti tende a 1.

Esempio 32 Sia $x > 1$ e $r \in \mathbb{N}$. Si ha

$$\lim_{n\to+\infty} \frac{n^r}{x^n} = 0\,. \tag{3.13}$$

Consideriamo in effetti la successione dei quozienti:

$$q_n = \frac{(n+1)^r}{x^{n+1}}\frac{x^n}{n^r} = \frac{1}{x}\left(1 + \frac{1}{n}\right)^r \to \frac{1}{x}\,(1+0)^r = \frac{1}{x}\,.$$

(Abbiamo utilizzato il fatto che $1/n \to 0$ e le proprietà algebriche dei limiti). Poichè $1/x < 1$, la tesi segue subito dalla Proposizione 3.17.

Esempio 33 Sia $x \in \mathbb{R}$. Si ha

$$\lim_{n\to+\infty} \frac{x^n}{n!} = 0\,. \tag{3.14}$$

Consideriamo, anche in questo caso, la successione dei quozienti:

$$q_n = \frac{x^{n+1}}{(n+1)!}\frac{n!}{x^n} = x\frac{1}{n+1} \to x \cdot 0 = 0\,.$$

La tesi segue quindi ancora dalla Proposizione 3.17.

Da questi limiti notevoli possiamo ora ottenerne altri.

Esempio 34 Si ha

$$\lim_{n\to+\infty} \sqrt[n]{n} = 1\,. \tag{3.15}$$

Si noti in effetti che, fissato $\epsilon > 0$, si ha che

$$\sqrt[n]{n} < 1+\epsilon \iff n > (1+\epsilon)^n \iff \frac{n}{(1+\epsilon)^n} < 1\,.$$

Poichè, per la (3.13), $n/(1+\epsilon)^n \to 0$ si ha che $n/(1+\epsilon)^n < 1$ definitivamente. Quindi, $\sqrt[n]{n} < 1+\epsilon$ definitivamente. Per concludere basta ora notare che, poichè $n \geq 1$, si ha $\sqrt[n]{n} \geq 1$ per ogni $n \in \mathbb{N}$. Quindi si ha che, definitivamente

$$1-\epsilon < 1 \leq \sqrt[n]{n} < 1+\epsilon\,.$$

Esempio 35 Si ha

$$\lim_{n\to+\infty} \sqrt[n]{n!} = +\infty\,. \tag{3.16}$$

Si noti in effetti che, fissato $M \in \mathbb{R}$ si ha che $\sqrt[n]{n!} > M$ è sempre verificato se $M < 0$; mentre, se $M \geq 0$ si ha

$$\sqrt[n]{n!} > M \iff n! > M^n \iff \frac{M^n}{n!} < 1\,.$$

Poichè, per l'Esempio 3.14, $M^n/n! \to 0$ si ha che $M^n/n! < 1$ definitivamente. Quindi, $\sqrt[n]{n!} > M$ definitivamente.

Arriviamo ora ad uno dei limiti più importanti. Alcune considerazioni su questo limite sono rimandate alla sezione successiva sulle serie.

Esempio 36 Consideriamo la successione $((1+1/n)^n)$. Facciamo vedere che essa è monotona crescente. Per farlo abbiamo bisogno di utilizzare la formula del binomio di Newton (Proposizione A.5 presentata nell'appendice). In base ad essa si può scrivere:

$$\begin{aligned}
\left(1+\frac{1}{n}\right)^n &= \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} \\
&= \sum_{k=0}^{n} \frac{n(n-1)\cdots(n-k+1}{n^k}\frac{1}{k!} \qquad (3.17) \\
&= \sum_{k=0}^{n} \frac{n}{n}\frac{n-1}{n}\cdots\frac{n-k+1}{n}\frac{1}{k!}\,.
\end{aligned}$$

Si noti ora che le varie frazioni

$$\frac{n}{n}\,,\ \frac{n-1}{n}\,,\quad \frac{n-k+1}{n}$$

sono tutte crescenti in n (verificarlo per esercizio). Quindi, trattandosi di tutti termini positivi, gli addendi

$$\frac{n(n-1)\cdots(n-k+1}{n^k}\frac{1}{k!}$$

risultano essere crescenti in n. Possiamo ora concludere che effettivamente la nostra successione è crescente. In effetti, come si vede nella rappresentazione sopra, essa è costituita da $n+1$ addendi positivi: quando n aumenta, il numero degli addendi aumenta e, al contempo, i vari addendi crescono con n. Più avanti mostreremo che questa successione è superiormente limitata, più precisamente mostreremo che $((1+1/n)^n) \leq 3$ per ogni n. Ne consegue, in virtù del Teorema 3.16, che esiste finito

$$\lim_{n\to+\infty}\left(1+\frac{1}{n}\right)^n\,.$$

IL valore di tale limite si indica con la lettera e ed è detta *costante di Nepero*. Per le considerazioni precedenti si ha che $e \leq 3$. Si noti, d'altra parte, che $(1+1/1)^1 = 2$. Essendo la successione crescente si ha che $(1+1/n)^n \geq 2$ per ogni n. Quindi $2 \leq e \leq 3$. Le prime cifre della rappresentazione decimale di e sono: $e = 2,71\cdots$.

3.6 Serie: prime proprietà

Un tipo molto importante di successioni sono le serie definite qui sotto:

Definizione 3.18 Sia (a_n) una successione. Si definisce *serie* associata ad (a_n), la successione delle somme parziali (s_n) data da

$$s_n = \sum_{k=1}^{n} a_k\,.$$

Nel caso in cui $a_n \geq 0$ per ogni n, (s_n) si dice *serie a termini non negativi*

Il limite di una serie ha una terminologia e simbologia particolari:

Definizione 3.19 Supponiamo che la serie associata alla successione (a_n) abbia limite. Esso è detto *somma della serie* ed indicato con il simbolo

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

Una serie si dice *convergente* o *sommabile* se tale limite è finito.

Osservazione: Come accadeva per le successioni, anche le somme parziali possono partire da un valore di k diverso da 1.

Osservazione: Se due successioni (a_n) e (b_n) sono definitivamente uguali, si ha che la serie associata ad una converge se e soltanto se la serie associata all'altra converge. In effetti, se abbiamo che $a_n = b_n$ per $n \geq n_0$ si ha che, se $n \geq n_0$,

$$\begin{aligned}
\sum_{k=1}^{n} a_k &= \sum_{k=1}^{n_0} a_k + \sum_{k=n_0+1}^{n} a_k \\
&= \sum_{k=1}^{n_0} a_k + \sum_{k=n_0+1}^{n} b_k \\
&= \sum_{k=1}^{n_0} a_k - \sum_{k=1}^{n_0} b_k + \sum_{k=1}^{n_0} b_k + \sum_{k=n_0+1}^{n} b_k \\
&= \left[\sum_{k=1}^{n_0} a_k - \sum_{k=1}^{n_0} b_k\right] + \sum_{k=1}^{n} b_k \,.
\end{aligned}$$

il che significa che le due serie, definitivamente, differiscono per una costante. Questo, per la Proposizione 3.11 implica la tesi. Si noti tuttavia che in generale le due serie non avranno la stessa somma.

Generalmente, il simbolo $\sum_{k=1}^{+\infty} a_k$ viene anche utilizzato per indicare la serie, cioè la successione delle some parziali $(\sum_{k=1}^{n} a_k)$, oltre che la sua eventuale somma. Lo studente deve comunque mantenere ben distinti i due concetti: una serie è una successione di somme parziali, la somma della serie è il limite di detta successione.

Ovviamente la convergenza delle serie può essere, in via di principio, studiata con le stesse tecniche utilizzate per le successioni. La difficoltà nasce di solito nel sapere trattare queste somme che aumentano sempre di addendi. Vedremo più avanti che si possono formulare tecniche più specifiche per la loro analisi, per intanto mostriamo un paio di esempi che si riescono a trattare direttamente.

Esempio 37 Serie di Mengoli-Cauchy: E' la serie

$$\sum_{k=1}^{+\infty} \frac{1}{k(k+1)} .$$

Si osservi che

$$\frac{1}{k(k+1)} = \frac{1}{k} - \frac{1}{k+1} .$$

Questo permette di scrivere

$$\begin{aligned} \sum_{k=1}^{n} \tfrac{1}{k(k+1)} &= \left(1 - \tfrac{1}{2}\right) + \left(\tfrac{1}{2} - \tfrac{1}{3}\right) + \cdots + \left(\tfrac{1}{n-1} - \tfrac{1}{n}\right) + \left(\tfrac{1}{n} - \tfrac{1}{n+1}\right) \\ &= 1 - \tfrac{1}{n+1} . \end{aligned}$$

Poichè $1/(n+1) \to 0$ si ha che la serie di Mengoli-Cauchy è convergente e si ha

$$\sum_{k=1}^{+\infty} \frac{1}{k(k+1)} = 1 .$$

Esempio 38 Serie geometrica: sia $x \in \mathbb{R}$ e si consideri la serie (detta serie geometrica) associata alla progressione geometrica (x^n) introdotta nell'Esempio 19:

$$\sum_{k=0}^{+\infty} x^k .$$

Se $|x| < 1$ si può scrivere (vedi Esercizio A.2)

$$\sum_{k=0}^{n} x^k = \frac{1 - x^{n+1}}{1 - x} .$$

Poichè in tal caso $x^n \to 0$ (vedi Esempio 19), applicando la Proposizione 3.11, si ottiene che

$$\sum_{k=0}^{+\infty} x^k = \frac{1}{1 - x} . \tag{3.18}$$

Di diretta derivazione dalla Proposizione 3.11 si ha la seguente:

Proposizione 3.20 *Siano (a_n) e (b_n) due successioni e sia $c \in \mathbb{R}$. Allora,*

(i) Se $\sum_{k=1}^{+\infty} a_k$ e $\sum_{k=1}^{+\infty} b_k$ convergono, anche $\sum_{k=1}^{+\infty}(a_k + b_k)$ converge, e si ha

$$\sum_{k=1}^{+\infty}(a_k + b_k) = \sum_{k=1}^{+\infty} a_k + \sum_{k=1}^{+\infty} b_k .$$

(ii) Se $\sum_{k=1}^{+\infty} a_k$ *converge, anche* $\sum_{k=1}^{+\infty} ca_k$ *converge, e si ha*

$$\sum_{k=1}^{+\infty} ca_k = c\sum_{k=1}^{+\infty} a_k\,.$$

Dimostrazione Poichè,

$$\sum_{k=1}^{n}(a_k+b_k)=\sum_{k=1}^{n} a_k+\sum_{k=1}^{n} b_k, \qquad \sum_{k=1}^{n} ca_k = c\sum_{k=1}^{n} a_k$$

tutto segue dalla Proposizione 3.11. ■

Il seguente importante criterio da un'utile condizione necessaria per la convergenza di una serie:

Proposizione 3.21 *Sia* (a_n) *una successione tale che* $\sum_{k=1}^{+\infty} a_k$ *converge. Allora,* (a_n) *è infinitesima.*

Dimostrazione Possiamo scrivere, per ogni $n \in \mathbb{N}$,

$$a_n = \sum_{k=1}^{n} a_k - \sum_{k=1}^{n-1} a_k\,.$$

Poichè $\left(\sum_{k=1}^{n} a_k\right)$ e $\left(\sum_{k=1}^{n-1} a_k\right)$ entrambe convergono e allo stesso numero, si ha che la loro differenza necessariamente converge a 0. ■

Esempio 39 Serie geometrica (completamento): sia $x \in \mathbb{R}$ tale che $|x| \geq 1$ e si consideri la serie geometrica

$$\sum_{k=0}^{+\infty} x^k\,.$$

Segue dagli Esempi 22 e 24 che la successione (x^n) per $|x| \geq 1$ non è mai infinitesima. Quindi la serie associata non converge.

3.7 Serie a termini non negativi

Nel resto di questa sezione concentreremo lo studio sulle serie associate a successioni con tutti i termini non negativi; le altre verranno studiate in seguito. Serie di questo tipo sono già apparse in questo capitolo. In effetti la successione degli approssimanti decimali di un numero reale non negativo illustrata nell'Esempio 12 è proprio una serie a termini non negativi.

L'osservazione fondamentale da fare sulle serie a termini non negativi è che esse sono evidentemente successioni crescenti così che esse ammettono sempre limite per il Teorema 3.16. Si ha, più precisamente,

Corollario 3.22 *Sia (a_n) una successione a termini non negativi. Si ha che*

$$\sum_{k=0}^{+\infty} a_k = \sup\left\{\sum_{k=0}^{n} a_k \mid n \in \mathbb{N}\right\}. \tag{3.19}$$

In particolare, $\sum_{k=0}^{+\infty} a_k$ è sommabile se e soltanto se essa è limitata.

Nel caso in cui la serie (a termini non negativi) non sia limitata, si dice anche che la serie è *divergente*.

Possiamo ora esporre quello che è il risultato base per la convergenza delle serie a termini positivi.

Teorema 3.23 (del confronto tra serie) *Siano (a_n) e (b_n) due successioni non negative tali che*

$$0 \le a_n \le b_n \qquad \forall n \in \mathbb{N}. \tag{3.20}$$

Allora,

(i) Se $\sum_{k=0}^{+\infty} b_k$ converge anche $\sum_{k=0}^{+\infty} a_k$ converge e si ha

$$\sum_{k=0}^{+\infty} a_k \le \sum_{k=0}^{+\infty} b_k < +\infty. \tag{3.21}$$

(ii) Se $\sum_{k=0}^{+\infty} a_k$ diverge anche $\sum_{k=0}^{+\infty} b_k$ diverge.

Dimostrazione (i): Segue dalla (3.20) che

$$\sum_{k=0}^{n} a_k \le \sum_{k=0}^{n} b_k \qquad \forall n \in \mathbb{N}. \tag{3.22}$$

Si ha quindi che, usando il Corollario 3.22,

$$\sum_{k=0}^{+\infty} a_k = \sup\left\{\sum_{k=0}^{n} a_k \mid n \in \mathbb{N}\right\} \le \sup\left\{\sum_{k=0}^{n} b_k \mid n \in \mathbb{N}\right\} = \sum_{k=0}^{+\infty} b_k.$$

(ii) conseguenza immediata della (i). ■

Osservazione: Se l'ipotesi (3.20) nel Teorema 3.23 è verificata soltanto definitivamente la tesi del teorema continua a valere anche se la diseguaglianza (3.21) non è più necessariamente vera.

La precedente osservazione è alla base di un altro criterio di confronto di derivazione dal Teorema 3.23, molto utile in certe applicazioni, e che presentiamo qui sotto.

Teorema 3.24 (del confronto asintotico tra serie) *Siano (a_n) e (b_n) due successioni positive tali che*

$$\lim_{n\to+\infty}\frac{a_n}{b_n}=l\neq 0\,.$$

Allora, $\sum\limits_{k=0}^{+\infty} b_k$ *converge se e soltanto se* $\sum\limits_{k=0}^{+\infty} a_k$ *converge.*

Dimostrazione Segue dalla definizione di limite che, definitivamente

$$\frac{a_n}{b_n}\le l+1$$

o, equivalentemente, che, definitivamente

$$a_n\le(l+1)b_n\,.$$

Se la serie $\sum_{k=0}^{+\infty} b_k$ converge, anche la serie $\sum_{k=0}^{+\infty}(l+1)b_k$ converge per la Proposizione 3.20. Quindi anche la serie $\sum_{k=0}^{+\infty} a_k$ converge per il Teorema 3.23 e l'osservazione sopra. Per dimostrare l'altra parte del teorema basta notare che il ruolo di (a_n) e di (b_n) è interscambiabile in quanto si ha anche, per la Proposizione 3.11,

$$\lim_{n\to+\infty}\frac{a_n}{b_n}=\frac{1}{l}\,.$$

■

Presentiamo ora alcuni importanti esempi.

Esempio 40 Serie armonica: E' la serie

$$\sum_{k=1}^{+\infty}\frac{1}{k}\,.$$

Si consideri la successione (a_n) costruita nel modo seguente:

$$a_1=1,\ a_2=\frac{1}{2},\ a_3=a_4=\frac{1}{4},\ a_5=a_6=a_7=a_8=\frac{1}{8},\ldots$$

La formula generale è

$$a_n = \frac{1}{2^m}, \quad \text{se} \quad 2^{m-1} + 1 \leq n \leq 2^m .$$

Chiaramente,

$$\frac{1}{n} \leq a_n \quad \forall n = 1, 2, \ldots$$

D'altra parte,

$$\begin{aligned}
\sum_{k=1}^{2^m} a_k &= 1 + \tfrac{1}{2} + \underbrace{\frac{1}{4} + \frac{1}{4}} + \underbrace{\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}} + \cdots + \underbrace{\frac{1}{2^m} + \cdots + \frac{1}{2^m}}_{2^{m-1}} \\
&= 1 + \tfrac{1}{2} + \tfrac{1}{2} + \tfrac{1}{2} + \cdots + \tfrac{1}{2} \\
&= 1 + \tfrac{m}{2} .
\end{aligned}$$

Questo mostra che la serie associata ad (a_n) è divergente. Per confronto lo è anche la serie armonica:

$$\sum_{k=1}^{+\infty} \frac{1}{k} = +\infty .$$

Esempio 41 Serie armonica generalizzata: E' la serie

$$\sum_{k=1}^{+\infty} \frac{1}{k^\alpha} .$$

dove α è un parametro reale non negativo. Per $\alpha = 1$ abbiamo visto, nell'esempio precedente, che la serie diverge. Vediamo ora che cosa accade se $\alpha = 2$. Si noti che

$$\frac{\frac{1}{n^2}}{\frac{1}{n(n+1)}} = \frac{n(n+1)}{n^2} \to 1 .$$

Poichè la serie associata alla successione $(\frac{1}{n(n+1)})$ converge (vedi Esempio 37), per confronto asintotico si ha che

$$\sum_{k=1}^{+\infty} \frac{1}{k^2} < +\infty .$$

Si noti ora che se $\alpha \geq 2$ si ha che

$$\frac{1}{n^\alpha} \leq \frac{1}{n^2} \qquad n = 1, 2, 3, \ldots$$

Quindi, per confronto, si ha anche

$$\sum_{k=1}^{+\infty} \frac{1}{k^\alpha} < +\infty \qquad \forall \alpha \geq 2\,.$$

D'altra parte, se $\alpha \leq 1$ si ha che

$$\frac{1}{n^\alpha} \geq \frac{1}{n} \qquad n = 1, 2, 3, \ldots$$

Quindi, per confronto, si ha anche

$$\sum_{k=1}^{+\infty} \frac{1}{k^\alpha} = +\infty \qquad \forall \alpha \leq 1\,.$$

Rimane da stabilire che cosa succeda se $1 < \alpha < 2$: si può dimostrare, ma non lo faremo in questa sede, che per tali valori di α la serie armonica generalizzata continua a convergere. Riassumendo si ha che:

$$\sum_{k=1}^{+\infty} \frac{1}{k^\alpha} \begin{cases} < +\infty & \text{se } \alpha > 1\,, \\ = +\infty & \text{se } 0 \leq \alpha \leq 1\,. \end{cases}$$

Esempio 42 Consideriamo la serie

$$\sum_{k=1}^{+\infty} \frac{k!}{k^k}\,. \tag{3.23}$$

Si noti che, se $n \geq 2$,

$$\frac{n!}{n^n} = \frac{1 \cdot 2 \cdot 3 \cdots n}{n \cdot n \cdot n \cdots n} = \frac{1}{n} \cdot \frac{2}{n} \cdot \frac{3}{n} \cdots \frac{n}{n} \leq \frac{1}{n} \cdot \frac{2}{n} = \frac{2}{n^2}\,.$$

Poichè (vedi l'Esempio 41) $\sum_{k=1}^{+\infty} 1/k^2$ converge, per confronto, anche la serie (3.23) converge.

Dal Teorema 3.23 si ricavano utili criteri per testare la convergenza delle serie a termini positivi.

Teorema 3.25 (Criterio del rapporto) *Sia (a_n) una successione tale che $a_n > 0$ per ogni n. Consideriamo il rapporto a_{n+1}/a_n di due successivi elementi della successione. Allora,* *(i) Se esiste* $0 \le r < 1$ *tale che* $a_{n+1}/a_n \le r$ *per ogni* $n \in \mathbb{N}$*, la serie* $\sum_{k=1}^{+\infty} a_k$ *converge e si ha*

$$\sum_{k=1}^{+\infty} a_k \le \frac{a_1}{1-r}\,. \tag{3.24}$$

(ii) Se $a_{n+1}/a_n \ge 1$ *per ogni* $n \in \mathbb{N}$*, la serie* $\sum_{k=1}^{+\infty} a_k$ *diverge.*

Dimostrazione (i): Applicando iterativamente la diseguaglianza $\frac{a_{n+1}}{a_n} \le r$ si ottiene:

$$a_{n+1} \le r a_n \le r^2 a_{n-1} \le \cdots \le r^n a_1$$

(per verificare la relazione sopra rigorosamente si deve utilizzare l'induzione, si provi a farlo). La serie associata alla successione $(r^n a_1)$ converge per l'Esempio 38 e la Proposizione 3.20. Dunque $\sum_{k=1}^{+\infty} a_k$ converge per confronto.

(ii): Applicando iterativamente la diseguaglianza $\frac{a_{n+1}}{a_n} \ge 1$ si ottiene:

$$a_{n+1} \ge a_n \ge a_{n-1} \ge \cdots \ge a_1 > 0\,.$$

Poichè la serie associata alla successione costante (a_1) diverge, anche $\sum_{k=1}^{+\infty} a_k$ diverge per confronto. ■

Teorema 3.26 (Criterio della radice) *Sia* (a_n) *una successione tale che* $a_n \ge 0$ *per ogni* n*. Consideriamo l'espressione* $\sqrt[n]{a_n}$*. Allora,*

(i) Se esiste $0 \le r < 1$ *tale che* $\sqrt[n]{a_n} \le r$ *per ogni* $n \in \mathbb{N}$*, la serie* $\sum_{k=1}^{+\infty} a_k$ *converge e si ha*

$$\sum_{k=1}^{+\infty} a_k \le \frac{1}{1-r}\,. \tag{3.25}$$

(ii) Se $\sqrt[n]{a_n} \ge 1$ *per ogni* $n \in \mathbb{N}$*, la serie* $\sum_{k=1}^{+\infty} a_k$ *diverge.*

Dimostrazione (i): Dalla diseguaglianza $\sqrt[n]{a_n} \le r$ segue che

$$a_n \le r^n \qquad \forall n \in \mathbb{N}\,.$$

Per l'Esempio 38 la serie geometrica $\sum_{k=1}^{+\infty} r^k$ converge. Dunque $\sum_{k=1}^{+\infty} a_k$ converge per confronto.

(ii): La diseguaglianza $\sqrt[n]{a_n} \ge 1$ implica

$$a_n \ge 1 \qquad \forall n \in \mathbb{N}\,.$$

Poichè la serie associata alla successione costante (1) diverge, anche $\sum_{k=1}^{+\infty} a_k$ diverge per confronto. ■

Osservazione Si noti che se sappiamo soltanto che $a_{n+1}/a_n < 1$ non possiamo concludere alcunchè sulla convergenza della serie. Si noti in effetti che se $a_n = 1/n$ si ha che $a_{n+1}/a_n = n/n+1 < 1$ per ogni n. D'altra parte sappiamo (vedi Esempio 40) che la serie associata diverge. Una simile osservazione si può fare per il criterio della radice: se sappiamo soltanto che $\sqrt[n]{a_n} < 1$ non possiamo concludere niente sulla convergenza della serie. L'esempio precedente si può utilizzare anche in questo caso.

Osservazione Si noti come i Teoremi 3.25 e 3.26 possano essere applicati anche nel caso in cui le ipotesi ($a_{n+1}/a_n \leq r < 1$ o $\sqrt[n]{a_n} \leq r < 1$) siano verificate soltanto definitivamente. La convergenza è ancora assicurata anche se le specifiche stime (3.24) e (3.25) non saranno più necessariamente vere.

L'osservazione precedente permette, come nel caso del Teorema del confronto, di fornire delle versioni di tipo asintotico dei due criteri precedenti:

Teorema 3.27 (Criterio del rapporto asintotico) *Sia (a_n) una successione tale che $a_n > 0$ per ogni n. Supponiamo che esista*

$$\lim_{n\to+\infty} \frac{a_{n+1}}{a_n} = R\,.$$

Allora,

(i) Se $R < 1$, la serie $\sum_{k=1}^{+\infty} a_k$ converge.

(ii) Se $R > 1$ (incluso la possibilità $R = +\infty$), la serie $\sum_{k=1}^{+\infty} a_k$ diverge.

Dimostrazione (i): se $R < 1$, esiste sicuramente $\epsilon > 0$ tale che $R + \epsilon < 1$. Per la definizione di limite si ha che, definitivamente

$$\frac{a_{n+1}}{a_n} \leq R + \epsilon < 1\,.$$

La tesi quindi segue dal Teorema 3.25 e dall'Osservazione seguente.

(ii) si dimostra in modo analogo. ■

Osservazione: Si noti come il Teorema 3.27 insieme alla Proposizione 3.21 implichi la Proposizione 3.17.

Teorema 3.28 (Criterio della radice asintotico) *Sia (a_n) una successione tale che $a_n \geq 0$ per ogni n. Supponiamo che esista*

$$\lim_{n\to+\infty} \sqrt[n]{a_n} = R\,.$$

Allora,

(i) Se $R < 1$, la serie $\sum_{k=1}^{+\infty} a_k$ converge.

(ii) Se $R > 1$ (incluso la possibilità $R = +\infty$), la serie $\sum_{k=1}^{+\infty} a_k$ diverge.

Dimostrazione Si lascia per esercizio.

Esempio 43 (Serie esponenziale) E' una delle serie più importanti della matematica:

$$\sum_{k=0}^{+\infty} \frac{x^k}{k!}\,.$$

Supponiamo qui che $x > 0$. Verifichiamo la sua convergenza applicando il criterio del rapporto asintotico:

$$\lim_{n\to+\infty} \frac{\dfrac{x^{n+1}}{(n+1)!}}{\dfrac{x^n}{n!}} = \lim_{n\to+\infty} \frac{n!}{(n+1)!}x = \lim_{n\to+\infty} \frac{1}{n+1}x = 0\,.$$

Questo dimostra che la serie esponenziale converge per ogni $x > 0$. C'è un profondo legame con la costante di Nepero introdotta precedentemente; in effetti si ha che:

$$\sum_{k=0}^{+\infty} \frac{1}{k!} = \lim_{n\to+\infty}\left(1+\frac{1}{n}\right)^n\,.$$

Per dimostrare questa eguaglianza, partiamo dalla relazione (3.17):

$$\left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{n} \frac{n}{n}\frac{n-1}{n}\cdots\frac{n-k+1}{n}\frac{1}{k!} \leq \sum_{k=0}^{n}\frac{1}{k!} \leq \sum_{k=0}^{+\infty}\frac{1}{k!}\,.$$

Questo mostra che la successione $(1+1/n)^n)$ è superiormente limitata (questo fatto non era stato dimostrato nella sezione precedente). Essendo monotona crescente, essa ammette limite finito e vale

$$\lim_{n\to+\infty}\left(1+\frac{1}{n}\right)^n \leq \sum_{k=0}^{+\infty}\frac{1}{k!}\,.$$

Per dimostrare che la relazione sopra è in realtà un'eguaglianza, si osservi che se $n > m$ si ha

$$\left(1+\frac{1}{n}\right)^n = \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} > \sum_{k=0}^{m} \frac{1}{k!} \frac{n(n-1)\cdots(n-k+1)}{n^k}.$$

Passando ora al limite per $n \to +\infty$ e tenuto conto che risulta

$$\frac{n(n-1)\cdots(n-k+1)}{n^k} \to 1\,,$$

otteniamo, per ogni $m \in \mathbb{N}$,

$$e \geq \sum_{k=0}^{m} \frac{1}{k!}\,.$$

Passando ora al limite per $m \to +\infty$ si ottiene

$$e \geq \sum_{k=0}^{+\infty} \frac{1}{k!}\,.$$

Quindi si ha proprio

$$e = \sum_{k=0}^{+\infty} \frac{1}{k!}\,.$$

Mostriamo infine che $e \leq 3$. Vale la seguente diseguaglianza

$$k! \geq 2^{k-1}\,, \quad \forall k \geq 1\,,$$

(questo fatto si dimostra per induzione su n). Ne segue che

$$\frac{1}{k!} \leq \frac{1}{2^{k-1}}\,, \quad \forall k \geq 1\,.$$

Quindi per confronto e utilizzando la (3.18) si ottiene

$$\sum_{k=0}^{+\infty} \frac{1}{k!} = 1 + \sum_{k=1}^{+\infty} \frac{1}{k!} \leq 1 + \sum_{k=1}^{+\infty} \frac{1}{2^{k-1}} = 1 + \frac{1}{1-\frac{1}{2}} = 3\,.$$

Esempio 44 Consideriamo la serie

$$\sum_{k=0}^{+\infty} k^r x^k\,,$$

dove $r \in \mathbb{N}$ e $x > 0$. Si ha che

$$\lim_{k \to +\infty} \frac{(k+1)^r x^{k+1}}{k^r x^k} = \lim_{k \to +\infty} \left(1 + \frac{1}{k}\right) x = x\,.$$

Quindi, per il Criterio del rapporto asintotico, si ha che la serie converge se e solo se $x < 1$. Si noti che, in realtà, il criterio del rapporto non permette di concludere niente se $x = 1$; per questo valore di x tuttavia la successione non è neppure limitata e quindi la serie diverge.

Esempio 45 Consideriamo la serie

$$\sum_{k=0}^{+\infty} \left(\frac{1+k}{1+2k}\right)^k .$$

Si ha che

$$\lim_{n \to +\infty} \sqrt[n]{\left(\frac{1+n}{1+2n}\right)^n} = \lim_{n \to +\infty} \frac{1+n}{1+2n} = \frac{1}{2}\,.$$

Quindi la serie converge in virtù del Criterio della radice asintotico.

Non sempre si possono applicare i criteri asintotici come mostra il seguente:

Esempio 46 Consideriamo la serie

$$\sum_{k=0}^{+\infty} \frac{7 + (-1)^k}{2^k}\,.$$

Si noti che

$$\frac{\frac{7+(-1)^{k+1}}{2^{k+1}}}{\frac{7+(-1)^k}{2^k}} = \frac{7+(-1)^{k+1}}{7+(-1)^k}\frac{1}{2}\,.$$

Non è difficile rendersi conto che l'espressione sopra non ammette limite; quindi il criterio del rapporto asintotico non può essere usato. Si noti tuttavia che

$$\frac{7+(-1)^{k+1}}{7+(-1)^k}\frac{1}{2} \le \frac{8}{6} \cdot \frac{1}{2} = \frac{2}{3} < 1\,.$$

Quindi per il criterio del rapporto la serie converge.

Osservazione: Si noti che nel caso in cui il limite R dei Criteri del rapporto asintotico o della radice asintotico, è uguale ad 1, nulla si può dedurre sul comportamento della serie. In effetti si ha, ad esempio, che se (a_n) e (b_n) sono due successioni date, rispettivamente, da $a_n = 1/n$ e $b_n = 1/n^2$, si ha che per entrambe $a_{n+1}/a_n \to 1$ e $b_{n+1}/b_n \to 1$. D'altra parte sappiamo che $\sum_{k=1}^{+\infty} a_k$ diverge mentre $\sum_{k=1}^{+\infty} b_k$ converge.

Esercizio 3.6 Studiare la convergenza delle serie seguenti:

$$\sum_{n=0}^{+\infty} \frac{n!}{(2n)!}\,, \qquad \sum_{n=0}^{+\infty} \frac{n^2+n}{n^4+1}\,, \qquad \sum_{n=0}^{+\infty} (\sqrt{n+1}-\sqrt{n})\,,$$

$$\sum_{n=0}^{+\infty} \frac{(n!)^2}{n(2n)!}\,, \qquad \sum_{n=0}^{+\infty} \frac{1+2+\cdots+n}{n^4}\,, \qquad \sum_{n=0}^{+\infty} (1+|\sin n|)\,,$$

$$\sum_{n=0}^{+\infty} \frac{n}{n+1}\,, \qquad \sum_{n=0}^{+\infty} [(2n)!-n!]\,, \qquad \sum_{n=0}^{+\infty} \frac{1+\sin^6 n}{n^2}\,.$$

R: (da sinistra a destra e dall'alto al basso) convergono 1, 2, 4, 5, 9.

Esercizio 3.7 * Sia (a_n) una successione tale che $a_n \geq 0$ per ogni $n \in \mathbb{N}$. Dimostrare che

$$\sum_{k=1}^{+\infty} a_k < +\infty \iff \sum_{k=1}^{+\infty} \frac{a_k}{1+a_k} < +\infty\,.$$

Esercizio 3.8 Dimostrare che

$$0,\overline{1} = \frac{1}{9}\,.$$

Esercizio 3.9 Dimostrare che

$$k_0,\overline{k_1} = \frac{10\cdot k_0,k_1 - k_0}{9}\,.$$

Esercizio 3.10 * Dimostrare che

$$k_0,k_1\cdots k_{n_0}\overline{k_{n_0+1}\cdots k_{n_0+n_1}} = \frac{10^{n_0+n_1}(k_0,k_1\cdots k_{n_0+n_1}) - 10^{n_0}(k_0,k_1\cdots k_{n_0})}{(10^{n_1}-1)10^{n_0}}\,.$$

3.8 Serie a segno qualunque

Uno dei concetti più importanti per studiare la convergenza di serie che non siamo a termini non negativi è quello di convergenza assoluta che proponiamo qui sotto:

Definizione 3.29 Sia (a_n) una successione qualsiasi. Si dice che la serie $\sum_{k=0}^{+\infty} a_k$ è *assolutamente convergente* se converge la serie $\sum_{k=0}^{+\infty} |a_k|$.

Il termine stesso 'convergenza assoluta' sembra indicare un concetto più forte della convergenza; ci aspettiamo dunque che la convergenza assoluta implichi la convergenza. Le cose stanno proprio così anche se la dimostrazione non è del tutto immediata.

Cominciamo con l'introdurre un concetto che sarà di utilità anche in altre parti del corso. Dato $x \in \mathbb{R}$, definiamo

$$x^+ = \begin{cases} x & \text{se} \quad x \geq 0 \\ 0 & \text{se} \quad x < 0\,, \end{cases}$$

$$x^- = \begin{cases} 0 & \text{se} \quad x \geq 0 \\ -x & \text{se} \quad x < 0\,. \end{cases}$$

x^+ è detta la *parte positiva* di x, mentre x^- la *parte negativa*. Si hanno le seguenti relazioni di facile dimostrazione:

$$x = x^+ - x^-, \qquad |x| = x^+ + x^-\,.$$

Siamo ora pronti per enunciare e dimostrare il risultato cercato:

Teorema 3.30 *Se una serie* $\sum_{k=0}^{+\infty} a_k$ *converge assolutamente, allora converge, e si ha*

$$\left|\sum_{k=0}^{+\infty} a_k\right| \leq \sum_{k=0}^{+\infty} |a_k|\,. \tag{3.26}$$

Dimostrazione Consideriamo le successioni ottenute prendendo, per ogni n, la parte positiva e la parte negativa di a_n: (a_n^+) e (a_n^-). Per il modo in cui sono state definite queste due successioni sono non negative e valgono inoltre le relazioni:

$$a_n = a_n^+ - a_n^-, \qquad |a_n| = a_n^+ + a_n^-\,. \tag{3.27}$$

Dalla seconda in (3.27) si ha, in particolare, che

$$a_n^+ \leq |a_n| \qquad a_n^- \leq |a_n|\,.$$

Poichè, per ipotesi, $\sum_{k=0}^{+\infty} |a_k|$ converge, per il Teorema del confronto tra serie (sono non-negative), si ha che $\sum_{k=0}^{+\infty} a_k^+$ e $\sum_{k=0}^{+\infty} a_k^-$ entrambe convergono. Usando ora la prima eguaglianza in (3.27) e la Proposizione 3.20 si ha che $\sum_{k=0}^{+\infty} a_k$ converge e si ha

$$\sum_{k=0}^{+\infty} a_k = \sum_{k=0}^{+\infty} a_k^+ - \sum_{k=0}^{+\infty} a_k^- \leq \sum_{k=0}^{+\infty} a_k^+ + \sum_{k=0}^{+\infty} a_k^- = \sum_{k=0}^{+\infty} |a_k|\,.$$

D'altra parte, si ha anche,

$$\sum_{k=0}^{+\infty} a_k = \sum_{k=0}^{+\infty} a_k^+ - \sum_{k=0}^{+\infty} a_k^- \geq -\sum_{k=0}^{+\infty} a_k^+ - \sum_{k=0}^{+\infty} a_k^- = -\sum_{k=0}^{+\infty} |a_k|\,.$$

Le due diseguaglianze sopra insieme sono equivalenti a (3.26). ■

Esempio 47 Consideriamo la serie

$$\sum_{k=0}^{+\infty} \frac{\sin k}{k^2}\,.$$

Studiamo la convergenza assoluta. Si ha

$$\left|\frac{\sin k}{k^2}\right| = \frac{|\sin k|}{k^2} \leq \frac{1}{k^2}\,.$$

Poichè la serie associata alla succesione $(1/k^2)$ converge, per confronto converge anche la serie associata a

$$\left|\frac{\sin k}{k^2}\right|\,.$$

Questo vuol dire che la serie di partenza converge assolutamente.

Il concetto di convergenza assoluta è molto utile perchè riconduce ad una serie a termini non negativi per le quali abbiamo a disposizione molti risultati per studiarne la convergenza. Cosa succede però se una serie non è assolutamente convergente? Può accadere che tuttavia essa sia convergente? Non ci addentreremo molto in questi aspetti molto delicati della teoria delle serie; ci limitiamo a presentare un risultato senza dimostrazione ed un esempio che da una risposta alla domanda precedente.

Teorema 3.31 (Criterio di Leibnitz) *Sia (a_n) una successione decrescente infinitesima (e quindi a termini non negativi). Allora la serie*

$$\sum_{k=1}^{+\infty}(-1)^k a_k$$

è convergente.

Esempio 48 Consideriamo la serie $\sum_{k=1}^{+\infty}(-1)^k/k$. Essa è convergente per il Teorema 3.31. Si noti tuttavia che essa non è assolutamente convergente in quanto $|(-1)^k/k| = 1/k$ e come sappiamo (vedi Esempio 40) la serie armonica non è convergente. Quindi esistono serie convergenti che non sono assolutamente convergenti.

Esercizio 3.11 Studiare la convergenza delle serie seguenti, specificando i casi in cui si ha convergenza assoluta:

$$\sum_{n=0}^{+\infty}\frac{1-n}{1+n}, \qquad \sum_{n=1}^{+\infty}\frac{\cos \pi n}{\sqrt{n}}, \qquad \sum_{n=0}^{+\infty}\left(\frac{\sin n}{2}\right)^n,$$

$$\sum_{n=0}^{+\infty}\frac{1+(-1)^n n}{n^2+1}, \qquad \sum_{n=0}^{+\infty}\frac{2^n-7n^2}{3^n+1}, \qquad \sum_{n=0}^{+\infty}(1-7\sin n)2^{-n}.$$

R: (da sinistra a destra e dall'alto al basso) convergono assolutamente 3, 5, 6; convergono (non ass.) 2, 4.

Capitolo 4

Funzioni di variabile reale

4.1 Introduzione

Il concetto di funzione è stato introdotto nella sua forma più astratta nell'Appendice. In questa sede ci concentreremo su funzioni che agiscono tra sottoinsiemi di numeri reali. Esse servono a formalizzare l'idea secondo la quale il valore di una certa grandezza numerica che ci interessa *dipende* dal valore assunto da un altra grandezza numerica tramite una legge nota. Vediamo alcuni esempi.

1. Lo spazio percorso da un corpo di massa m in caduta libera e trascurando effetti di attrito, dopo un tempo t, è dato da
$$s(t) = \frac{1}{2}gt^2 + v_0 t$$
dove g è l'accelerazione di gravità e v_0 è la velocità iniziale del corpo.

2. Un corpo di massa unitaria, lanciato a velocità iniziale v_0 ad un angolo φ rispetto al terreno, atterra a una distanza
$$d(\varphi) = \frac{v_0^2}{g}\sin(2\varphi)$$
dal punto di lancio.

3. L'area A di un triangolo equilatero di lato l vale
$$A(l) = \frac{\sqrt{3}}{4}l^2\,.$$

4. La probabilità p che un generatore di numeri casuali tra zero e uno fornisca un numero maggiore di $x \in (0, 1)$ vale
$$p(x) = 1 - x\,.$$

In tutti questi casi l'idea fondamentale è la seguente: assegnato il valore di una *variabile indipendente* (ingresso), è univocamente assegnato il valore di una *variabile dipendente* (uscita). Naturalmente i valori che può assumere la variabile indipendente dipendono dall'esempio specifico considerato. In particolare negli esempi precedenti abbiamo che il tempo t può essere un qualunque numero positivo, il lato l può essere anch'esso un qualunque numero positivo, il numero x può variare tra 0 e 1, e infine l'angolo φ può variare tra 0 e $\pi/2$.

Osserviamo inoltre che i valori assegnati in uscita tramite le funzioni considerate sono sempre numeri reali: in questo corso non faremo che poche eccezioni a questa regola, sempre sottintesa da ora in avanti. Osserviamo però che in realtà per ognuno dei casi sopra riportati si può dire qualcosa di più: per esempio l'angolo $\vartheta(t)$ certamente risulterà compreso tra $-\pi$ e π.

Dunque, ricapitolando in tutti questi casi, si tratta di funzioni
$$f : A \subseteq \mathbb{R} \to \mathbb{R}$$
aventi come dominio un sottoinsieme di numeri reali e come codominio l'insieme dei numeri reali. Tali funzioni sono spesso dette *funzioni reali di variabile reale*. In tutto questo capitolo il termine funzione si riferirà implicitamente a questo tipo particolare di funzioni.

Osservazione: Talvolta può risultare utile utilizzare codomini di tipo diverso, considerando funzioni del tipo
$$f : A \subset \mathbb{R} \to B \subset \mathbb{R}\,,$$
cioè funzioni per le quali si sa a priori che l'insieme dei possibili valori assunti dalla f è un insieme *contenuto* in $\mathbb{R}$ ma non necessariamente coincidente con esso; questo specialmente quando si discute l'invertibilità di una funzione (vd. i punti seguenti). Il codominio è un insieme a cui si sa a priori (senza cioè un'indagine accurata sulla funzione) che i valori di f appartengono. Come tale, esso di per sé non è univocamente determinato. L'immagine invece è esattamente l'insieme dei valori assunti da f, cioè il più piccolo codominio possibile, dove il termine "più piccolo si riferisce all'inclusione insiemistica.

Assegnata una funzione, possiamo innanzitutto definirne il *grafico*. Esso è un sottoinsieme del piano cartesiano, ed é definito come l'insieme delle coppie $(x, f(x))$ con x appartenente al dominio A della funzione. Alcuni ben noti esempi sono illustrati nelle Figure 4.1 e 4.2.

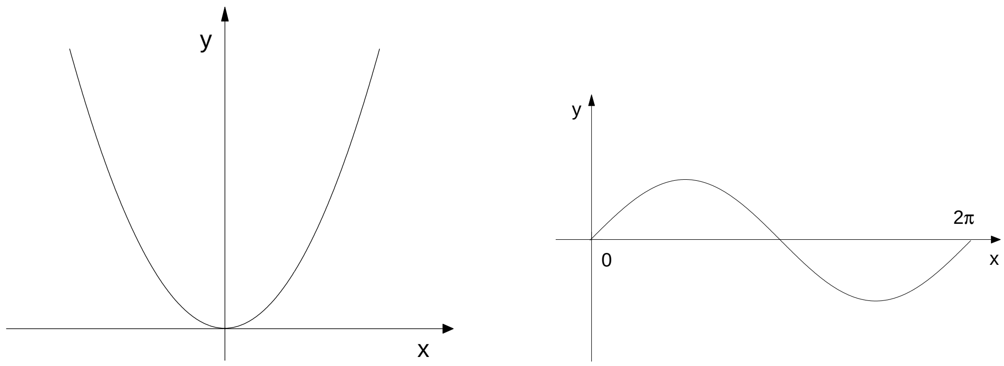


Figura 4.1:

4.2 Invertibilità, funzioni composte

Nell'appendice sono state introdotte le definizioni di iniettività, suriettività ed invertibilità in generale. La verifica di tali proprietà per funzioni reali di variabile reale, può essere fatta per via analitica: data $f : A \to B$, si tratta di studiare l'equazione $f(x) = y$ per ogni fissato $y \in B$. Se tale equazione ha sempre (per ogni $y \in B$) almeno una soluzione $x \in A$, la funzione è surgettiva; se non ha mai più di una soluzione, essa è iniettiva. E' talvolta utile utilizzare un metodo grafico per verificare queste proprietà. Si consideri il grafico di f e, preso un punto $y \in B$, si tracci la retta orizzontale passante per $(0, y)$. Se, qualunque sia $y \in B$, tale retta incontra sempre il grafico di f, vuol dire che la funzione è surgettiva; se tale retta non incontra mai più di una volta il grafico, vuol dire che la funzione è iniettiva.

Presentiamo ora alcuni esempi concreti.

Esempio 49 Prendiamo il primo degli esempi elementari prima considerati: $f : \mathbb{R} \to \mathbb{R}$ definita come

$$f(x) = x^2 .$$

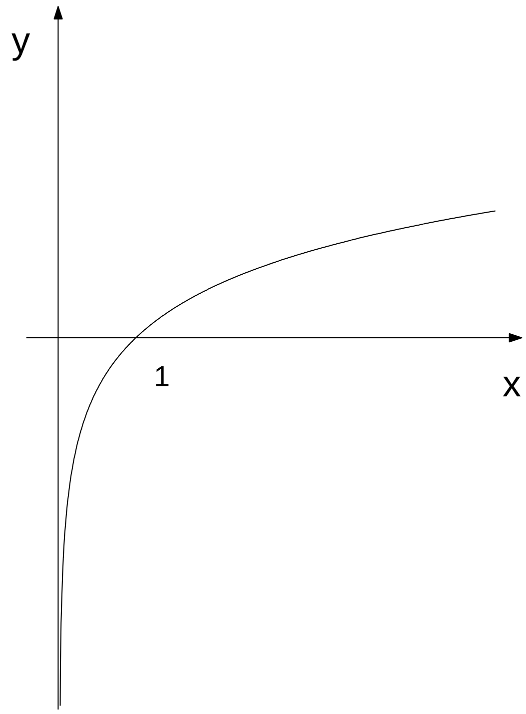


Figura 4.2:

La funzione che stiamo considerando non è né iniettiva né surgettiva. In effetti, è chiaro che l'immagine di f è la semiretta $[0,+\infty[$: quindi se $y < 0$ non esiste alcun $x \in R$ tale che $x^2 = y$, e dunque f non è surgettiva. Ma la funzione data non è neanche iniettiva, poiché $f(x) = f(-x)$ per ogni x. Non si può quindi certo invertire f: eppure in un qualche senso $\sqrt{x}$ ha almeno qualcuna delle proprietà che abbiamo richiesto per una funzione inversa. In effetti

$$(\sqrt{x})^2 = x,$$

Però

$$\sqrt{x^2} = |x|$$

e non x come si sarebbe forse tentati di scrivere. Potremmo anche tentare di usare la funzione $-\sqrt{x}$, ma avremmo analogamente $(-\sqrt{x})^2 = x$ e $-\sqrt{x^2} = -|x|$. Cerchiamo allora di restringere il dominio in modo che la funzione sia almeno iniettiva. Se consideriamo f sul dominio $A = [0,+\infty[$, l'equazione $y = x^2$ ha una e una sola soluzione $x \in A$ quale che sia $y \geq 0$. Quindi f, che si dice essere la *restrizione* della funzione originaria all'insieme A, è iniettiva. Potete verificare per esercizio che, se $A = [a, b]$, allora la restrizione di f ad A è iniettiva se e soltanto se o il punto $x = 0$ non appartiene ad A oppure $a = 0$ oppure $b = 0$.

E circa la suriettività? Siccome x^2 è sempre un numero non negativo, non possiamo sperare che f sia surgettiva a meno di modificare il codominio: in effetti se definiamo il codominio come l'insieme $B = [0, +\infty[$ la funzione f risulta surgettiva sia nel caso in cui il dominio sia $\mathbb{R}$, sia nel caso in cui esso sia $[0, +\infty[$. è di nuovo lasciata per esercizio la verifica del fatto che, se il codominio B è scelto essere $[0. + \infty[$, la funzione f è surgettiva se e soltanto se il dominio A contiene $[0, +\infty[$ oppure $] - \infty, 0]$.

In conclusione, la funzione $f : [0, +\infty[\to [0, +\infty[$ definita da $f(x) = x^2$ è invertibile, in quanto iniettiva e surgettiva. Quale ne è l'inversa? Quella che ci aspettiamo: $f^{-1}(x) = \sqrt{x}$. In effetti abbiamo visto sopra che $(\sqrt{x})^2 = x$ e che $\sqrt{x^2} = |x|$, ma poiché x è per ipotesi non negativo, allora $\sqrt{x^2} = x$.

Possiamo però verificare altrettanto facilmente che la funzione $f :] - \infty, 0] \to [0, +\infty[$ definita da $f(x) = x^2$ è anch'essa invertibile e ha come inversa la funzione $f^{-1}(x) = -\sqrt{x}$. In conclusione, anche se la forma analitica di f non cambia, la modifica di dominio e codominio ha conseguenze sia sull'invertibilità che, a maggior ragione, sull'espressione esplicita dell'inversa.

Esercizio 4.1 Mostrare che la funzione $f : \mathbb{R} \to \mathbb{R}$ definita da $f(x) = 2x + 3$ è invertibile, e determinarne esplicitamente l'inversa. Determinare poi, dato un intervallo $A = [a, b]$, se esiste un insieme $B \subset \mathbb{R}$ tale che $f : A \to B$ definita come sopra sia invertibile.

R: $f^{-1}(x) = x/2 - 3/2$

Esercizio 4.2 Stabilire se la funzione $f : \mathbb{R} \setminus 0 \to \mathbb{R}$ definita da $f(x) = 1/x$ è surgettiva, iniettiva, invertibile. Nel caso in cui essa non risulti invertibile, stabilire se esiste $B \subset \mathbb{R}$ tale che $f : \mathbb{R} \setminus 0 \to B$ definita come sopra sia invertibile, determinando poi esplicitamente la funzione inversa.

È immediato dalla definizione, e importante in pratica, che il grafico di una funzione inversa f^{-1} si ricava da quello di f *per simmetria rispetto alla bisettrice del primo e terzo quadrante.* In effetti per definizione di inversa un punto (x, y) appartiene al dominio di f se e soltanto se (y, x) appartiene al grafico di f^{-1}. E potrete facilmente rendervi conto che i punti di coordinate

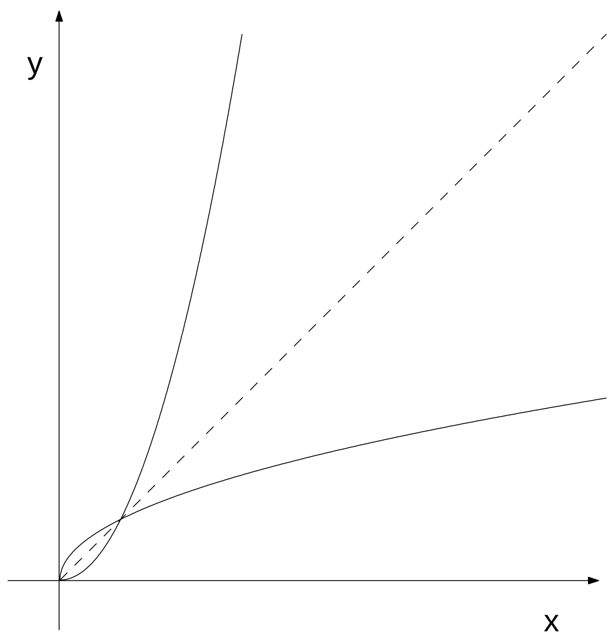


Figura 4.3:

(x, y) e (y, x) sono simmetrici rispetto alla retta $y = x$, sono cioè punti che stanno su una retta perpendicolare alla retta $y = x$ e sono da tale retta equidistanti. Si veda la Figura 4.3

Affrontiamo ora il problema di comporre due funzioni reali. Siano $f : A_1 \subset \mathbb{R} \to \mathbb{R}$, $g : A_2 \subset \mathbb{R} \to \mathbb{R}$. Ha sempre senso scrivere $f(g(x))$? La risposta è no: in effetti se $f(x) = \log x$ e $g(x) = -x^2$ allora $f(g(x))$ non ha senso dato che non si può prendere il logaritmo di un numero negativo. Se però l'immagine di g, $g(A_2)$, è contenuta nel dominio di f (cosa che non accade nell'esempio appena fatto!), allora ha perfettamente senso, come sappiamo, considerare $f(g(x))$ per ogni $x \in A_2$: per esempio se $f(x) = \log x$ e $g(x) = x^2 + 1$ allora $\log(x^2 + 1)$ è un oggetto ben definito per ogni $x \in \mathbb{R}$. Il prodotto di composizione non è commutativo:

Osservazione: È essenziale precisare l'ordine in cui agiscono le funzioni g ed f. In effetti se, come sopra, $f(x) = \log x$ e $g(x) = x^2 + 1$, allora

$$\begin{aligned} &[f \circ g](x) = \log(x^2 + 1) \qquad \forall x \in \mathbb{R}\,, \\ &[g \circ f](x) = (\log x)^2 + 1 \qquad \forall x > 0\,. \end{aligned}$$

Quindi $f \circ g \neq g \circ f$ sia dal punto di vista della loro espressione esplicita che dal punto di vista del dominio. Addirittura potrebbe darsi che che $f \circ g$ abbia un dominio ben definito e che $g \circ f$ non sia mai definita.

4.3 Monotonia, limitatezza, simmetrie, periodicità

In questa sezione introduciamo alcune definizioni che ci saranno utili in seguito. La prima di queste riguarda la proprietà di *monotonia* che una funzione può o meno soddisfare. Già conosciamo questo concetto nell'ambito delle successioni, e si tratta solo di fare le ovvie modifiche come segue.

Definizione 4.1 Una funzione $f : A \subset \mathbb{R} \to B \subset \mathbb{R}$ si dice

- *monotona crescente* se
$$x < y \in A \Longrightarrow f(x) \leq f(y)\,,$$
- *monotona strettamente crescente* se
$$x < y \in A \Longrightarrow f(x) < f(y)\,,$$
- *monotona decrescente* se
$$x < y \in A \Longrightarrow f(x) \geq f(y)\,,$$
- *monotona strettamente decrescente* se
$$x < y \in A \Longrightarrow f(x) > f(y)\,.$$

Bisogna fare attenzione nel verificare la condizione di monotonia, come si vede dai seguenti esempi. Essi mostrano che la proprietà di monotonia dipende in modo essenziale dal dominio della funzione.

Esempio 50 Consideriamo la funzione
$$f : \mathbb{R} \setminus \{0\} \to \mathbb{R}, \quad f(x) = \frac{1}{x}\,.$$

Si vede immediatamente che tale funzione non è monotona né crescente né decrescente: ad esempio $f(-1) = -1 \leq 1 = f(1)$ mentre $f(1) = 1 > 1/2 = f(2)$. Tuttavia le restrizioni di f a $\mathbb{R}^-$ e, rispettivamente, a $\mathbb{R}^+$ sono entrambe monotone strettamente decrescenti: ad esempio la disuguaglianza

$$\frac{1}{x_1} \geq \frac{1}{x_2}$$

è soddisfatta, nel caso $x_1, x_2 > 0$, se e soltanto se $x_1 < x_2$.

Esempio 51 Consideriamo la funzione $f : \mathbb{R} \to \mathbb{R}$ definita da $f(x) = \sin x$. Allora certamente f non è monotona dato che, ad esempio, $f(2k\pi + \pi/2) = 1 \ \forall k \in \mathbb{Z}$ e $f(2k\pi - \pi/2) = -1 \ \forall k \in \mathbb{Z}$. Tuttavia le restrizioni di f agli intervalli del tipo

$$\left[-\frac{\pi}{2} + 2k\pi, \frac{\pi}{2} + 2k\pi\right]$$

sono, per ogni fissato $k \in \mathbb{Z}$, monotone strettamente crescenti, mentre le restrizioni di f agli intervalli del tipo

$$\left[\frac{\pi}{2} + 2k\pi, \frac{3\pi}{2} + 2k\pi\right]$$

sono monotone strettamente decrescenti. Ciò può essere dedotto dalla definizione geometrica di seno di un angolo ma potrebbe anche essere dimostrato analiticamente.

Un importante proprietà delle funzioni monotone è la seguente:

Proposizione 4.2 *Sia $f : A \subset \mathbb{R} \to f(A)$ una funzione monotona strettamente crescente o strettamente decrescente. Allora f è invertibile.*

Prima di dare la semplice dimostrazione osserviamo che è essenziale che il codominio di f coincida con la sua immagine. Pensate ad esempio alla funzione $f(x) = \operatorname{artg} x$, che è monotona crescente e ha valori compresi tra -1 e 1, e quindi non è certo invertibile come funzione da $\mathbb{R}$ in $\mathbb{R}$.

Dimostrazione Dobbiamo mostrare che, per ogni elemento y del codominio, esiste uno e un solo elemento x del dominio tale che $f(x) = y$. Ma in questo caso il codominio coincide con l'immagine di f, e quindi per definizione di immagine si ha che, dato $y \in f(A)$, esiste almeno un $\hat{x} \in A$ tale che $f(\hat{x}) = y$. Resta da dimostrare che tale y è unico. Infatti, supponendo ad esempio f strettamente crescente, si ha che per ogni $x_1, x_2 \in A$ con $x_1 < \hat{x} < x_2$, vale

$$f(x_1) < f(\hat{x}) = y < f(x_2)$$

cosicché in particolare $f(x) \neq y$ se $x \neq y$. Il caso di funzioni strettamente decrescenti si tratta in modo analogo. ■

Esercizio 4.3 Stabilire per quali $n \in \mathbb{N}$ la funzione $f : \mathbb{R} \to \mathbb{R}$ definita da $f(x) = x^n$ è monotona.

Conosciamo già il concetto di limitatezza di un insieme numerico. Sappiamo anche cosa intendiamo dicendo che una successione è limitata; in particolare abbiamo definito tale concetto chiedendo che sia limitato il codominio della successione. Procediamo analogamente anche per le funzioni.

Definizione 4.3 Una funzione $f : A \subset \mathbb{R} \to B \subset \mathbb{R}$ si dice *inferiormente limitata* se la sua immagine è inferiormente limitata, cioè se esiste $L \in \mathbb{R}$ tale che $f(x) \geq L$ per ogni $x \in A$. Si dice *superiormente limitata* se la sua immagine è superiormente limitata, cioè se esiste $M \in \mathbb{R}$ tale che $f(x) \geq M$ per ogni $x \in A$. Si dice *limitata* se è sia inferiormente che superiormente limitata.

Esercizio 4.4 Stabilire quali tra le seguenti funzioni sono inferiormente limitate, quali sono superiormente limitate e quali limitate:

$$
\begin{aligned}
&f(x) = \log(x+5) && \text{se } x > 0\,,\\
&f(x) = \sin(x^2) && \text{se } x \in \mathbb{R}\,,\\
&f(x) = x^2 - x + 2 && \text{se } x \in \mathbb{R}\,,\\
&f(x) = -2^{-x} && \text{se } x \in \mathbb{R}\,.
\end{aligned}
$$

R: infer. lim; lim.; infer. lim.; super. lim..

Alcune funzioni possiedono *proprietà di simmetria*. Quelle a cui ci interesseremo sono particolarmente comprensibili pensando al grafico delle funzioni in esame: in particolare vogliamo dare una precisa definizione che corrisponda all'idea che un grafico sia simmetrico rispetto all'origine o rispetto all'asse delle ordinate. Vediamo come.

Definizione 4.4 Una funzione $f : \mathbb{R} \to \mathbb{R}$ si dice *pari* se

$$f(x) = f(-x) \;\; \forall x \in \mathbb{R}\,.$$

Si dice *dispari* se

$$f(x) = -f(-x) \;\; \forall x \in \mathbb{R}\,.$$

Cosa significano i concetti appena introdotti? Riflettiamo sul fatto che il grafico di una funzione è definito come l'insieme dei punti del piano cartesiano di coordinate $(x, f(x))$. Se allora f è pari, ciò significa che il punto di coordinate $(-x, f(x))$ appartiene al grafico. Ma tale punto altro non è che quello ottenuto per simmetria rispetto all'asse y partendo dal punto $(x, f(x))$ (l'ascissa cambia segno, l'ordinata rimane invariata). Quindi dire che una funzione è pari significa dire che il suo grafico è simmetrico rispetto all'asse y. Analogamente una funzione è dispari, il punto $(-x, -f(x))$ appartiene al grafico. Ma tale punto è ottenuto dal punto $(x, f(x))$ per simmetria rispetto all'origine (sia l'ascissa che l'ordinata cambiano segno). Quindi dire che una funzione è dispari significa dire che il suo grafico è simmetrico rispetto all'origine. La figura (4.4) mostra i grafici di una funzione pari e di una funzione dispari.

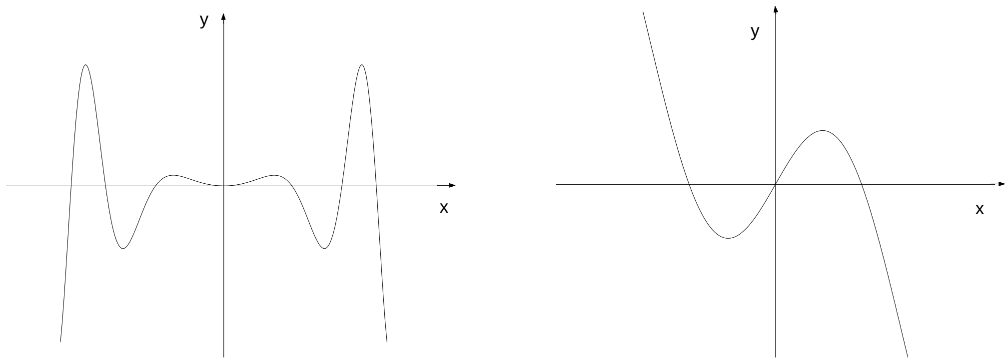


Figura 4.4:

Esercizio 4.5 Stabilire quali tra le funzioni $f : \mathbb{R} \to \mathbb{R}$ definite da $f(x) = x^n$ per $n \in \mathbb{N}$ sono pari e quali dispari.

Osservazione: Le definizioni di funzione pari e di funzione dispari possono essere generalizzate a funzioni $f : A \subset \mathbb{R} \to \mathbb{R}$ non necessariamente definite su tutto $\mathbb{R}$. È però necessario che il dominio A sia simmetrico rispetto all'origine, cioè che se $x \in A$ anche $-x \in A$. Se A ha questa proprietà possiamo dare una definizione di funzione pari e di funzione dispari identica alla precedente. Un esempio tipico è la funzione $f(x) = 1/x$ definita sull'insieme $\mathbb{R} \setminus 0$, che è effettivamente simmetrico rispetto all'origine. Tale funzione è chiaramente dispari.

Concludiamo questa sezione dando la definizione di funzione *periodica*.

Definizione 4.5 Sia $f : A \subset \mathbb{R} \to \mathbb{R}$ una funzione. Essa si dice periodica se esiste un numero $T > 0$ tale che

$$f(x+T) = f(x) \quad \forall x \in A \,. \tag{4.1}$$

Inoltre, se l'insieme di tali numeri ammette minimo, tale minimo si dice *periodo* della funzione.

Il lettore attento avrà notato che la definizione appena data richiede un fatto piuttosto delicato: l'esistenza del minimo per l'insieme dei numeri T che soddisfano (4.1). Ciò in effetti è *falso* in generale! Si pensi al fatto che la funzione (detta *funzione di Dirichlet*)

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

soddisfa la proprietà $f(x+T) = f(x)$ $\forall x \in \mathbb{R}$ *quale che sia* $T \in \mathbb{Q}$. Ma la proprietà sopra menzionata vale se si fa qualche ipotesi ulteriore sulla funzione: tali ipotesi hanno a che fare col concetto di *continuità* di una funzione, che sarà discusso nella prossima sezione. Ci accontenteremo quindi al momento di *supporre* che esista il minimo dell'insieme dei T che soddisfano (4.1).

Si deve anche osservare che il dominio A non può essere qualsiasi se la funzione è periodica, perché esso deve soddisfare la proprietà

$$x \in A \Longrightarrow x + T \in A$$

per ogni $x \in A$.

Gli esempi più consueti di funzioni periodiche sono le funzioni trigonometriche: In effetti $\sin x$, $\cos x$, $\operatorname{tg} x$, $\operatorname{cotg} x$... sono funzioni periodiche (di quale periodo?). Inoltre funzioni composte del tipo

$$[f \circ \sin]x = f(\sin x) \,,$$

con f definita su un insieme che includa l'intervallo $[-1,1]$, sono periodiche: ad esempio $f(x) = 1/(2+\sin x)$ è una funzione periodica. Analogamente accade per funzioni che dipendano solo da $\sin x$, $\cos x$ o da altre funzioni periodiche.

Osserviamo infine che, ovviamente, vi sono funzioni periodiche che non si costruiscono a partire da funzioni trigonometriche. Ad esempio la *funzione a gradino*

$$f(x) = \begin{cases} 0 & \text{se } [2k, 2k+1),\ k \in \mathbb{Z} \\ 1 & \text{se } [2k-1, 2k),\ k \in \mathbb{Z} \end{cases}$$

è periodica di periodo $T = 2$.

4.4 Funzioni continue

Una delle proprietà essenziali che è opportuno studiare quando viene assegnata una funzione è quella di *continuità*. Essa sorge in maniera naturale quando ci si chieda se è o meno vero il seguente fatto (enunciato in modo qualitativo): se la variabile indipendente x varia di poco, accade lo stesso per la variabile dipendente $y = f(x)$? Che questo non sia sempre vero lo mostra la funzione

$$f(x) = \begin{cases} 0 & \text{se } x \leq 0 \\ 1 & \text{se } x > 0 \end{cases} \tag{4.2}$$

dato che se $x = 0$ $f(x) = 0$ mentre $f(x) = 1$ quale che sia $x > 0$ anche molto vicino all'origine.

Più formalmente daremo la seguente

Definizione 4.6 Una funzione $f :]a, b[\to \mathbb{R}$ si dice *continua nel punto* $x_0 \in]a, b[$ se per ogni $\varepsilon > 0$ $\exists \delta > 0$, in generale dipendente da ε, tale che

$$|f(x) - f(x_0)| \leq \varepsilon$$

per ogni $x \in]a, b[$ tale che $|x - x_0| \leq \delta$. La funzione si dice *continua su* $]a, b[$ se è continua in ogni punto di $]a, b[$.

Osservazioni sul concetto di continuità: Possiamo descrivere colloquialmente quanto richiesto nella definizione precedente in almeno due modi (nessuno dei quali deve però sostituire la corretta definizione!).

- In effetti, ciò che si chiede è che, fissata l'ampiezza ε di una striscia *orizzontale* del tipo

$$A_\varepsilon = \{(x, y) : f(x_0) - \varepsilon \leq y \leq f(x_0) + \varepsilon\},$$

accada che il grafico di $f(x)$ stia nella striscia purché x sia abbastanza vicino a x_0 nel senso che $|x - x_0| \leq \delta$, e che ciò sia vero quale che sia l'ampiezza ε della striscia: è chiaro che δ deve in generale dipendere da ε.

- Ancora più informalmente possiamo dire che una funzione continua ammette un grafico che può essere disegnato senza staccare la penna dal foglio. Ciò naturalmente non succede per la funzione a gradino sopra scritta.

Esempio 52 Dimostriamo che la funzione

$$f(x) = mx + n, \quad x \in \mathbb{R}$$

è continua in ogni punto di $\mathbb{R}$ qualunque siano $m, n \in \mathbb{R}$. Fissiamo $x_0 \in \mathbb{R}$ e notiamo che

$$|f(x) - f(x_0)| = |mx + n - (mx_0 + n)| = |m(x - x_0)| = |m||x - x_0|\,.$$

Se $m = 0$ si ha che $|f(x) - f(x_0)| = 0$ sempre e quindi la continuità è chiaramente verificata. Supponiamo dunque che $m \neq 0$. Fissato un qualunque $\epsilon > 0$, se scegliamo $\delta = \epsilon/|m|$, allora si ha che

$$|x - x_0| < \delta \;\Rightarrow\; |f(x) - f(x_0)| = |m||x - x_0| < |m|\delta = \epsilon\,.$$

Questo dice esattamente che f è continua nel punto x_0.

Esempio 53 Dimostriamo che la funzione

$$f(x) = x^2, \quad x \in \mathbb{R}$$

è continua su $\mathbb{R}$, cioè che è continua in ogni punto. In effetti, sia $x_0 \in \mathbb{R}$ e sia $\varepsilon > 0$ fissato. Allora, se $|x - x_0| \leq \delta$:

$$\begin{aligned} |x^2 - x_0^2| &= |x - x_0|\,|x + x_0| \\ &\leq \delta|x + x_0| \\ &\leq \delta(|x - x_0| + 2|x_0|) \\ &\leq \delta(\delta + 2|x_0|)\,. \end{aligned}$$

Si noti ora che l'equazione (in δ):

$$\delta(\delta + 2|x_0|) = \varepsilon$$

è risolta in particolare da $\delta = \sqrt{\varepsilon + x_0^2} - |x_0|$. Per tale valore di δ si ha dunque che, se $|x - x_0| \leq \delta$, allora $|f(x) - f(x_0)| \leq \varepsilon$. La continuità in x_0 è dunque provata. Si noti infine che $\delta = \sqrt{\varepsilon + x_0^2} - |x_0|$ una quantità positiva e che si avvicina a zero quando ε si avvicina a zero. Il senso del verbo "avvicinarsi in questo contesto sarà formalizzato a breve con una naturale estensione del concetto di limite a voi già noto dallo studio delle successioni.

Esempio 54 Mostriamo che la funzione (4.2) non è continua nell'origine. Occorre mostrare che esiste $\varepsilon > 0$ tale che, per ogni $\delta > 0$, esista x con $|x| \leq \delta$ e $|f(x)| > \varepsilon$. In effetti, basta prendere $\varepsilon = 1/2$ e osservare che per ogni $x > 0$

$$|f(x)| = f(x) = 1 > \frac{1}{2}\,.$$

Esempio 55 Facciamo un altro esempio di funzione discontinua nell'origine. L'esempio è differente dal precedente perché, mentre nel primo caso $f(x)$ si "avvicina a valori differenti a seconda che x si "avvicini all'origine da destra o da sinistra, nel presente esempio non c'è alcun valore cui si possa dire che $f(x)$ si "avvicini se x è vicino all'origine. La funzione è definita come

$$f(x) = \begin{cases} \sin\left(\dfrac{1}{x}\right) & \text{se } x \neq 0 \\ 0 & \text{se } x = 0\,. \end{cases} \tag{4.3}$$

La funzione oscilla con sempre maggiore rapidità man mano che ci si avvicina all'origine. Per formalizzare la manzanza di continuità ragioniamo in questo modo. Notiamo prima di tutto che per ogni $\delta > 0$, al variare di $x \in]-\delta, \delta[$, la funzione $f(x)$ copre tutto l'intervallo $[-1, 1]$, e questo comunque δ sia stato scelto piccolo. Se ora fissiamo $0 < \epsilon < 1$, è chiaro, dalla considerazione precedente, che la diseguaglianza $|f(x)| < \epsilon$ non potrà mai essere soddisfatta su alcun intorno di 0 del tipo $]-\delta, \delta[$.

4.5 Proprietà delle funzioni continue

Vediamo ora alcune proprietà elementari ma fondamentali delle funzioni continue. La prima di esse ci dice che se f è una funzione continua in x_0 allora essa è limitata se x varia in un intervallo centrato in x_0 e di ampiezza non troppo grande. Questo è naturale visto che la definizione di continuità ci dice che $f(x)$ si avvicina a $f(x_0)$ quando x si avvicina a x_0.

Proposizione 4.7 *Se $f :]a, b[\to \mathbb{R}$ è continua in $x_0 \in]a, b[$ allora esiste $\delta > 0$ tale che la restrizione di f a $]x_0 - \delta, x_0 + \delta[$ sia una funzione limitata.*

Dimostrazione Sappiamo per ipotesi che, scelto ad esempio $\varepsilon = 1$, $\exists \delta > 0$ tale che

$$|f(x) - f(x_0)| \leq 1$$

per ogni $x \in]a, b[$ tale che $|x - x_0| \leq \delta$. Questo implica per la disuguaglianza triangolare che, per tali x, si abbia

$$\begin{aligned} |f(x)| &= |f(x) - f(x_0) + f(x_0)| \\ &\leq |f(x) - f(x_0)| + |f(x_0)| \\ &\leq 1 + |f(x_0)| \,. \end{aligned}$$

■

Il prossimo risultato è altrettanto intuitivo. Se assumiamo la continuità, e quindi di nuovo il fatto che, quando x sia avvicina a x_0 allora $f(x)$ si avvicina a $f(x_0)$, allora $f(x)$ avrà lo stesso segno di $f(x_0)$ se x è abbastanza vicino a x_0.

Proposizione 4.8 (della permanenza del segno) *Se $f :]a, b[\to \mathbb{R}$ è continua in $x_0 \in]a, b[$ e $f(x_0) > 0$ (rispettivamente $f(x_0) < 0$) allora esiste $\delta > 0$ tale che $f(x) > 0$ (rispettivamente $f(x) < 0$) per ogni $x \in]x_0 - \delta, x_0 + \delta[$. Più precisamente esiste un $\delta > 0$ tale che $f(x) \geq f(x_0)/2$ (risp. $f(x) < f(x_0)/2$) per ogni $x \in]x_0 - \delta, x_0 + \delta[$.*

Dimostrazione Basta dimostrare il caso $f(x_0) > 0$. In tale ipotesi avremo che, scelto $\varepsilon = f(x_0)/2$, esiste $\delta > 0$ tale che

$$|f(x) - f(x_0)| \leq \frac{f(x_0)}{2} \tag{4.4}$$

per tutti gli $x \in]a, b[$ tali che $|x - x_0| \leq \delta$. Ciò implica in particolare che

$$f(x) - f(x_0) \geq -\frac{f(x_0)}{2} \,,$$

cioè che

$$f(x) \geq \frac{f(x_0)}{2} > 0$$

per tutti i tali x. ■

Il prossimo risultato ci dice che la proprietà di continuità è stabile rispetto a somme (quindi anche a differenze), a prodotti e a rapporti (purché, nell'ultimo caso, il rapporto abbia senso, cioè il denominatore non sia zero). La dimostrazione è istruttiva solo per l'uso che viene fatto della proprietà di permanenza del segno.

Proposizione 4.9 *Siano* $f, g\ :]a, b[\to \mathbb{R}$ *funzioni continue in* $x_0 \in]a, b[$. *Allora sono continue in* x_0 *anche le seguenti funzioni:*

- $f + g$;
- fg;
- *se* $g(x_0) \neq 0$, $\dfrac{f}{g}$.

Dimostrazione Dimostriamo il caso del quoziente, e assumiamo per fissare le idee che $g(x_0) > 0$: la dimostrazione nel caso $g(x_0) < 0$ è analoga. Allora la Proposizione 4.8 applicata a g ci garantisce che esiste $\delta_0 > 0$ tale che $g(x) > g(x_0)/2$ per ogni x tale che $|x - x_0| \leq \delta_0$. Le ipotesi ci dicono inoltre che per ogni $\varepsilon > 0$ $\exists \delta_1, \delta_2 > 0$ tali che

$$|f(x) - f(x_0)| \leq \varepsilon$$

per ogni $x \in]a, b[$ tale che $|x - x_0| \leq \delta_1$ e

$$|g(x) - g(x_0)| \leq \varepsilon$$

per ogni $x \in]a, b[$ tale che $|x - x_0| \leq \delta_2$. Se ora $x \in]a, b[$ è tale che $|x - x_0| \leq \delta = \min(\delta_0, \delta_1, \delta_2)$, si ha che:

$$\begin{aligned}
\left|\frac{f(x)}{g(x)} - \frac{f(x_0)}{g(x_0)}\right| &= \left|\frac{f(x)g(x_0) - f(x_0)g(x)}{g(x)g(x_0)}\right| \\
&= \left|\frac{(f(x) - f(x_0))g(x_0) - f(x_0)(g(x) - g(x_0))}{g(x)g(x_0)}\right| \\
&\leq \frac{|(f(x) - f(x_0))|g(x_0) + |f(x_0)|g(x) - g(x_0)|}{g(x)g(x_0)} \\
&\leq 2\frac{|(f(x) - f(x_0))|g(x_0) + |f(x_0)||g(x) - g(x_0)|}{g(x_0)^2} \\
&\leq 2\varepsilon\frac{g(x_0) + |f(x_0)|}{g(x_0)^2}\ .
\end{aligned}$$

Si noti che l'ultima espressione pur non essendo esattamente ε, è un'espressione che può essere assegnata arbitrariamente piccola scegliendo opportunamente ε. Questo mostra la continuità del quoziente. ■

I prossimi risultati ci mostrano come la proprietà di continuità sia stabile rispetto a due operazioni che conosciamo: la composizione di funzioni e l'operazione di inversione di una funzione (quando tali operazioni sono possibili, e sotto opportune condizioni nel secondo caso).

Proposizione 4.10 *Sia* $f\ :]a, b[\to]c, d[$ *una funzione continua in* $x_0 \in]a, b[$ *e* $f\ :]c, d[\to \mathbb{R}$ *una funzione continua in* $y_0 := f(x_0) \in]c, d[$. *Allora la funzione* $g \circ f\ :]a, b[\to \mathbb{R}$ *è continua in* x_0.

In altre parole la composizione di funzioni continue è ancora una funzione continua. Omettiamo la dimostrazione, che è però molto semplice. Provate allora a svolgerla:

Esercizio 4.6 Dimostrare la proposizione precedente.

Il secondo problema è più complicato: quando si può dire che l'inversa di una funzione continua è continua? Che la questione non sia banale lo mostra il seguente

Esempio 56 Consideriamo la funzione $f : [-1,0]\cup]1,2] : [-1,1]$ definita da

$$f(x) = \begin{cases} x & \text{se } x \in [-1,0] \\ x-1 & \text{se } x \in \,]1,2]\,. \end{cases}$$

Finora avevamo sempre considerato funzioni definite su un intervallo, ma non ci sono problemi a considerare anche funzioni definite sull'unione di più intervalli come in questo caso. Inoltre finora abbiamo parlato di continuità solo in punti *interni* a un intervallo di definizione della funzione, ma è possibile generalizzare questa definizione al caso in cui x_0 è un estremo dell'intervallo: basta richiedere, informalmente parlando, che x si avvicini a x_0 "da destra, se x_0 è l'estremo sinistro dell'intervallo considerato, o "da sinistra in caso contrario. Si vedano per maggiori dettagli la successiva formula (4.5) e la Definizione 4.15.

È facile convincersi che la funzione è continua in ogni punto nel quale è definita (in effetti il suo grafico mostra che essa è monotona, quindi invertibile dal dominio all'immagine). La sua funzione inversa si può scrivere ricordando il fatto che il grafico di una funzione f e quello dell'inversa f^{-1} sono simmetrici rispetto alla bisettrice del primo e terzo quadrante. Allora:

$$f^{-1}(y) = \begin{cases} y & \text{se } y \in [-1,0] \\ y+1 & \text{se } y \in]0,1]\,. \end{cases}$$

Si tratta quindi di una funzione discontinua nell'origine $y = 0$.

Qual'è il problema nell'esempio precedente? In parole povere si tratta del fatto che il dominio di f è fatto di due intervalli disgiunti. In effetti la funzione inversa risulta discontinua proprio nel punto $y_0 = 0$, e si deve osservare che $f(0) = 0$. Quindi il problema si verifica in un punto che è l'immagine tramite f di un estremo del primo dei due intervalli su cui è definita f. Non potremo trattare la questione in piena generalità, ma daremo almeno il seguente risultato

Proposizione 4.11 *Se $f : I \to A \subset \mathbb{R}$ una funzione continua e invertibile su un intervallo $I \subset \mathbb{R}$. Allora la funzione inversa $f^{-1} : A \to I$ è continua su A.*

Non dimostreremo questo fatto, ma osserviamo solo di nuovo che l'ipotesi che I sia un intervallo (comprendente o meno uno o entrambi gli estremi) è essenziale per la validità della Proposizione.

4.6 Limiti

Molti di voi avranno riconosciuto nella definizione di continuità un concetto imparentato con la definizione di limite data per le successioni. In quest'ultima si fissava un ε e si chiedeva che, per n abbastanza grande, lo scarto tra il limite l e la successione fosse al massimo ε. Nella definizione di continuità abbiamo qualcosa di simile ma alla variabile *discreta* $n \in \mathbb{N}$ si sostituisce la variabile *continua* $x \in \mathbb{R}$, alla successione (a_n) (essa altro non è che una *funzione* della variabile discreta $n \in \mathbb{N}$) si sostituisce la funzione $f(x)$, al limite l della successione si sostituisce il valore $f(x_0)$.

Potremmo quindi pensare a dare una definizione di limite di funzione (quando "$x \to x_0$) che permetta di riformulare la definizione di continuità in questi termini: *una funzione $f :]a, b[\to \mathbb{R}$ è continua in $x_0 \in \mathbb{R}$ se e soltanto se il suo limite per $x \to x_0$ esiste finito e vale $f(x_0)$.*

Questo può essere effettivamente fatto, e la definizione risultante è la seguente:

Definizione 4.12 Sia $f :]a, b[\to \mathbb{R}$ una funzione, e $x_0 \in]a, b[$. Si dice che il *limite* per x che tende a x_0 di $f(x)$ vale $l \in \mathbb{R}$, e si scrive

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

oppure

$$f(x) \to l \quad \text{se } x \to x_0$$

qualora sia verificata la seguente condizione: per ogni $\varepsilon > 0$ esiste $\delta > 0$, in generale funzione di ε, tale che

$$|f(x) - l| \leq \varepsilon$$

per ogni $x \in]a, b[$ tale che $|x - x_0| < \delta$, eccettuato al più il punto $x = x_0$.

È a questo punto evidente la seguente implicazione, da tenere ben presente:

$$f \text{ continua in } x_0 \Longleftrightarrow \lim_{x\to x_0} f(x) = f(x_0). \tag{4.5}$$

Esercizio 4.7 Dimostrare che il limite, se esiste, è unico, cioè che se $l_1, l_2 \in \mathbb{R}$ sono tali che $\lim_{x\to x_0} f(x) = l_1$, $\lim_{x\to x_0} f(x) = l_2$, allora $l_1 = l_2$.

Va notata la richiesta che conclude la definizione di limite: $f(x)$ deve distare poco da l se x è vicino a x_0, ma sul valore di $f(x_0)$ non ci sono richieste. Questo perchè di una funzione f che differisca da una funzione continua g soltanto in un punto x_0 vogliamo poter dire egualmente che il limite quando $x \to x_0$ esiste, anche se esso non coincide con $f(x_0)$.

Oltre alle analogie formali tra i concetti di limite di funzioni e di limite di successioni, c'è un importante teorema (la cui dimostrazione è omessa) che collega più profondamente i due concetti.

Teorema 4.13 *Sia $f :]a,b[\to \mathbb{R}$ una funzione. Allora $f(x) \to l$ per $x \to \overline{x}$ se e soltanto se $f(a_n) \to l$ per ogni successione (a_n) tale che $a_n \to \overline{x}$ e tale che $a_n \neq \overline{x}$ definitivamente. In particolare f è continua in $\overline{x} \in]a,b[$ se e soltanto se è continua per successioni, cioè se e soltanto se, per ogni successione (a_n) tale che $a_n \to \overline{x}$, si ha $f(a_n) \to f(\overline{x})$ per $n \to \infty$.*

Quindi la continuità di una funzione in $\overline{x}$ può essere studiata esaminando la convergenza delle *successioni* $f(a_n)$, dove $a_n \to \overline{x}$ per $n \to +\infty$. Bisogna però ovviamente lavorare su *tutte* le possibili successioni (a_n) che convergono a $\overline{x}$. Il Teorema 4.13 è spesso anche utile per far vedere che un limite non esiste: esibendo una successione (a_n) tale che $a_n \to \overline{x}$ e tale che $f(a_n)$ non converge; oppure due successioni (a_n) e (b_n) che entrambi convergono a $\overline{x}$ e tali che le successioni $f(a_n)$ e $f(b_n)$ convergono a limiti distinti.

Esempio 57 (Continuazione Esempio 55). Un altro modo per mostrare la mancanza di continuità in 0 della funzione scalino (4.2) è proprio attraverso il Teorema 4.13. Definiamo la successione

$$x_n = \frac{1}{\dfrac{\pi}{2} + 2n\pi} \quad \forall n \in \mathbb{N}\,.$$

Si ha che (x_n) tende a zero e $f(x_n) = 1$ per ogni $n \in \mathbb{N}$. Quindi, dato che $f(0) = 0$, f non è continua nell'origine. Possiamo osservare più in generale (farlo per esercizio!) che per ogni $l \in [-1,1]$ fissato si può trovare una successione (x_n) (dipendente da l) tale che $x_n \to 0$ per $n \to +\infty$ e $f(x_n) = l$ per ogni $n \in \mathbb{N}$.

Passiamo ora a considerare altre forme utili del concetto di limite. Come accade per le successioni sappiamo che è necessaria una definizione apposita quando il limite vale $+\infty$ o $-\infty$: eccola qui di seguito.

Definizione 4.14 Sia $f :]a, b[\to \mathbb{R}$ una funzione, e $x_0 \in]a, b[$. Si dice che il limite per x che tende a x_0 di $f(x)$ vale $+\infty$ (risp. $-\infty$), e si scrive

$$\lim_{x \to x_0} f(x) = +\infty \quad (\text{risp. } -\infty)$$

oppure

$$f(x) \to +\infty \text{ (risp. } -\infty) \quad \text{se } x \to x_0$$

se per ogni $M \in \mathbb{R}$ esiste $\delta > 0$, in generale dipendente da M, tale che

$$f(x) > M \quad (\text{risp. } f(x) < M)$$

per ogni $x \in]a, b[$ tale che $|x - x_0| < \delta$.

Altre forme di limite nascono poi modificando il tipo di convergenza della variabile indipendente x. Una possibilità è imporre che x 'si avvicini' a x_0 soltanto da destra o da sinistra come è formalmente proposto nella seguente definizione:

Definizione 4.15 Si dice che $f(x) \to l \in \mathbb{R}$ quando x tende a x_0 da destra (risp. da sinistra), e si scrive

$$\lim_{x \to x_0^+} f(x) = l \quad (\text{risp. } \lim_{x \to x_0^-} f(x) = l)\,,$$

se per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che $|f(x) - l| \leq \varepsilon$ quando $0 < x - x_0 \leq \delta$ (risp. quando $\delta \leq x - x_0 < 0$).

In buona sostanza richiediamo *esattamente la stessa condizione* richiesta nella definizione di limite, ma restringiamo l'insieme degli x per cui essa è richiesta valere. Naturalmente una definizione analoga si può dare anche per funzioni che tendono a $+\infty$ o a $-\infty$. E' chiaro che se una funzione ha, per $x \to x_0$ da destra e da sinistra, valori limite diversi, non può sicuramente essere continua.

Esempio 58 Ritorniamo alla funzione (4.2) che avevamo notato non essere continua nell'origine. Si noti che i limiti destro e sinistro esistono in 0 e sono tra loro diversi:
$$\lim_{x\to 0+} f(x) = 1\,, \qquad \lim_{x\to 0-} f(x) = -1$$

Concludiamo infine con un'ultima definizione di limite che riguarda il caso in cui la variabile x tende non a un valore finito, ma all'infinito proprio come l'indice n nelle successioni.

Definizione 4.16

- Si dice che il limite per x che tende a $+\infty$ di $f(x)$ vale l, e si scrive
$$\lim_{x\to+\infty} f(x) = l$$
oppure
$$f(x) \to l \quad \text{se } x \to +\infty$$
se per ogni $\varepsilon > 0$ esiste K, tale che
$$|f(x) - l| < \varepsilon$$
per ogni x tale che $x > K$.

- Si dice che il limite per x che tende a $+\infty$ di $f(x)$ vale $+\infty$ (risp. $-\infty$), e si scrive
$$\lim_{x\to+\infty} f(x) = +\infty \quad (\text{risp. } -\infty)$$
oppure
$$f(x) \to +\infty \ (\text{risp. } -\infty) \quad \text{se } x \to +\infty$$
se per ogni $M \in \mathbb{R}$ esiste $K \in \mathbb{R}$, tale che
$$f(x) > M \quad (\text{risp. } f(x) < M)$$
per ogni x tale che $x > K$.

- Si dice che il limite per x che tende a $-\infty$ di $f(x)$ vale l, e si scrive
$$\lim_{x\to-\infty} f(x) = l$$

oppure

$$f(x) \to l \quad \text{se } x \to -\infty$$

se per ogni $\varepsilon > 0$ esiste K, tale che

$$|f(x) - l| < \varepsilon$$

per ogni x tale che $x < K$.

- Si dice che il limite per x che tende a $-\infty$ di $f(x)$ vale $+\infty$ (risp. $-\infty$), e si scrive

$$\lim_{x \to -\infty} f(x) = +\infty \quad (\text{risp. } -\infty)$$

oppure

$$f(x) \to +\infty \ (\text{risp. } -\infty) \quad \text{se } x \to -\infty$$

se per ogni $M \in \mathbb{R}$ esiste $K \in \mathbb{R}$, tale che

$$f(x) > M \ \ (\text{risp. } f(x) < M)$$

per ogni x tale che $x < K$.

Il Teorema 4.13 ammette delle ovvie generalizzazioni per queste altre forme del concetto di limite che lasciamo da formulare ai lettori.

Non entreremo poi nel dettaglio dei Teoremi elementari sui limiti di funzioni, che sono del tutto analoghi a quelli relativi ai limiti di successioni e che possono essere anche dimostrati utilizzando il Teorema 4.13. Si tratta di risultati del tipo seguente

Proposizione 4.17 *Siano $f :]a, b[\to \mathbb{R}$ e $g :]a, b[\to \mathbb{R}$ due funzioni tali che*

$$\lim_{x \to x_0} f(x) = l_1 \in \mathbb{R}, \ \lim_{x \to x_0} g(x) = l_2 \in \mathbb{R},$$

dove $x_0 \in \mathbb{R}$. Sia inoltre c una costante. Allora, si ha che

$$\lim_{x \to x_0} [cf(x)] = cl_1 \quad \forall c \in \mathbb{R};$$

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

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

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \frac{l_1}{l_2} \ (\text{se } l_2 \neq 0).$$

Inoltre, se f e g sono definite su semirette, i risultati precedenti valgono anche per $x \to \pm\infty$.

Enunciamo ora qualche altro risultato molto utile nelle applicazioni, la cui dimostrazione (lasciata per esercizio) può sempre essere fatta utilizzando il Teorema 4.13.

Teorema 4.18 (del confronto tra funzioni) *Siano $f :]a,b[\to \mathbb{R}$, $g :]a,b[\to \mathbb{R}$ e $h :]a,b[\to \mathbb{R}$ tre funzioni tali che*

$$f(x) \leq g(x) \leq h(x) \quad \forall x \in]a,b[\,.$$

Supponiamo inoltre che $f(x)$ e $g(x)$ ammettano limite per $x \to x_0 \in]a,b[$ con

$$\lim_{x\to x_0} f(x) = l = \lim_{x\to x_0} h(x)\,.$$

Allora anche $g(x)$ ammette limite per $x \to x_0$, e si ha

$$\lim_{x\to x_0} g(x) = l\,.$$

Osservazione: Il precedente teorema del confronto ammette varie possibili estensioni: x_0 può essere rimpiazzato da $\pm\infty$ e si possono anche considerare varianti sul tipo del Teorema 3.10 per le successioni in cui invece il limite l è rimpiazzato da $\pm\infty$.

Esempio 59 Segue dalla definizione stessa di seno di un angolo, che vale sempre la diseguaglianza

$$-x \leq \sin x \leq x$$

(in realtà geometricamente si vede se $|x| <\leq \pi/2$, ma per gli altri valori di x è facile estendere). Poichè x e $-x$ tendono entrambe a 0 per $x \to 0$, segue dal Teorema del confronto che

$$\lim_{x\to 0} \sin x = 0 \tag{4.6}$$

Proposizione 4.19 (della composizione dei limiti) *Sia $f :]a,b[\to \mathbb{R}$ tale che*

$$\lim_{x\to x_0} f(x) = y_0$$

con $x_0 \in]a,b[$ e con y_0 appartenente a un intervallo del tipo $]c,d[$. Sia inoltre $g :]c,d[\to \mathbb{R}$ continua in y_0. Allora,

$$\lim_{x\to x_0} g(f(x)) = g(y_0)\,.$$

Osservazione: La proposizione precedente continua a valere anche nel caso in cui x_0 sia $+\infty$. Similmente si può avere y_0 eguale a $+\infty$: in tal caso però la continuità di g va rimpiazzata con l'esistenza del limite a $+\infty$.

Proposizione 4.20 (del limite di funzioni monotone) *Sia $f :]a,b[\to \mathbb{R}$ una funzione monotona. Allora esistono sempre, finiti o infiniti, i limiti della $f(x)$ per $x \to b-$ e $x \to a+$ e precisamente si ha*

$$f(x) \text{ crescente} \qquad \lim_{x \to b-} f(x) = \sup_{x \in]a,b[} f(x)\,, \qquad \lim_{x \to a+} f(x) = \inf_{x \in]a,b[} f(x)$$

$$f(x) \text{ decrescente} \qquad \lim_{x \to b-} f(x) = \inf_{x \in]a,b[} f(x)\,, \qquad \lim_{x \to a+} f(x) = \sup_{x \in]a,b[} f(x)$$

Si possono anche riformulare le Proposizioni 3.12, 3.13 e 3.14 nell'ambito delle funzioni. Si provi a farlo per esercizio. Come nel caso delle successioni ci sono però dei casi in cui non è possibile dire nulla di generale: ad esempio se $f(x) \to 0$ e $g(x) \to +\infty$ per $x \to x_0$, non si può dire a priori nulla del limite del prodotto $f(x)g(x)$. Si suole parlare in casi come questi di *forme di indeterminazione* o *di indecisione*, e nel caso appena descritto si indica *simbolicamente* il problema scrivendo che si sta discutendo una forma di indecisione del tipo "$0 \cdot \infty$. Le più comuni forme di indecisione, anch'esse indicate in analogia con il caso precedente, sono le seguenti:

$$0 \cdot \infty\,,\ \frac{0}{0}\,,\ \frac{\infty}{\infty}\,,\ +\infty - \infty\,,\ 0^0\,,\ 1^\infty\,,\ \infty^0\,.$$

Se vi trovate in uno di questi casi non potete dir nulla di generale ma dovrete lavorare caso per caso, tipicamente usando i *limiti notevoli* e gli *sviluppi in serie di Taylor* di alcune funzioni elementari come discusso in seguito.

4.7 Continuità di alcune classi di funzioni

La domanda più naturale che possiamo porre a questo punto è la seguente: esiste una classe di funzioni "abbastanza ampia della cui continuità possiamo essere certi a priori, senza cioè necessità di una verifica caso per caso? Fortunatamente sì : una larga classe di *funzioni elementari* sono continue ovunque esse siano definite. In effetti si ha:

Teorema 4.21 *Tutte le funzioni sotto elencate sono continue sul proprio naturale insieme di definizione:*

1. *Tutte le funzioni razionali*

$$f(x) = \frac{a_0 + a_1 x + \dots a_{n-1}x^{n-1} + a_n x^n}{b_0 + b_1 x + \dots b_{m-1}x^{m-1} + b_m x^m}\,, \qquad x \in \mathbb{R}\,;$$

2. *Le funzioni potenza*

$$f(x) = x^\alpha\,, \qquad x > 0\,,\ \alpha \in \mathbb{R}\,;$$

3. *Le funzioni esponenziali*

$$f(x) = a^x\,, \qquad x \in \mathbb{R}\,,\ a > 0\,;$$

4. *I logaritmi*

$$f(x) = \log_a x\,, \qquad x > 0\,,\ a > 0\,,\ a \neq 1\,;$$

5. *Le funzioni trigonometriche elementari*

$$\sin x\,, \cos x\,, \tan x \qquad x \in \mathbb{R}\,;$$

e le loro inverse

$$\arcsin x\,, \arccos x\,, \arctan x \qquad x \in \mathbb{R}\,;$$

6. *La funzione valore assoluto*

$$f(x) = |x|\,.$$

Dimostrazione 1.: Le funzioni razionali si costruiscono da due tipi di funzioni, le costanti $f(x) = c$, e la funzione identità $f(x) = x$, attraverso le quattro operazioni elementari. Poichè queste funzioni base come ben sappiamo sono ovunque continue, il risultato segue dalla Proposizione 4.9.

3. Lo dimostriamo per $a > 1$ lasciandolo per esercizio nel caso $a < 1$. Sia $x_0 \in \mathbb{R}$. Abbiamo che

$$a^x - a^{x_0} = a^{x_0}[a^{x - x_0} - 1]\,.$$

Questo mostra che è sufficiente dimostrare che

$$\lim_{x \to x_0} a^{x - x_0} = 1$$

cioè che

$$\lim_{x \to 0} a^x = 1$$

che è poi la continuità dell'esponenziale nel punto 0. Si ricordi che la funzione a^x risulta crescente (essendo $a > 1$). Ne segue quindi che esistono finiti i due limiti:

$$\lim_{x \to 0+} a^x = l^+ \,, \quad \lim_{x \to 0-} a^x = l^-$$

D'altra parte, sappiamo che la successione $a^{1/n}$ tende a 1 per $n \to +\infty$. In virtù del Teorema 4.13 ne segue che necessariamente $l^+ = 1$. Analogamente $a^{-1/n}$ tende a 1 per $n \to +\infty$ e quindi anche $l^- = 1$. Questo dimostra il limite sopra.

4.: segue dalla Proposizione 4.11 considerando il fatto che il logaritmo è la funzione inversa dell'esponenziale.

2.: Vale la seguente identità (che per inciso è molto utile in vari contesti)

$$x^\alpha = a^{\log_a x^\alpha} = a^{\alpha \log_a x}$$

Poichè $\log_a x$ è continua (per $x > 0$) per il punto 4. e α è una costante, il prodotto $\alpha \log_a x$ è continuo per la Proposizione 4.9. Essendo l'esponenziale a^x continuo per il punto 2., segue quindi dalla Proposizione 4.10 che x^α è continua su tutto $\mathbb{R}$.

5.: Osserviamo che

$$\sin(x) - \sin(x_0) = 2 \sin\left(\frac{x - x_0}{2}\right) \cos\left(\frac{x + x_0}{2}\right) .$$

Si conclude quindi notando che il coseno è una funzione limitata, mentre $\sin((x - x_0)/2) \to 0$ per $x \to x_0 \to 0$ in base alla (4.6) e nuovamente alla Proposizione 4.10.

La continuità del coseno, ora segue dal fatto che $\cos x = \sin(x - \pi/2)$ e dalla Proposizione 4.10. Infine la continuità della tangente (per $x \neq \pi/2 + k\pi$) segue dalll'espressione $\tan x = \sin x / \cos x$ e dalla Proposizione 4.9.

6.: Osserviamo che

$$||x| - |x_0|| \leq |x - x_0|$$

Quindi tutto segue nuovamente dal Teorema del confronto. ■

Notiamo infine che si possono costruire molte altre funzioni continue partendo dal precedente elenco. In effetti somme, prodotti, rapporti (ove definiti) e composizioni (ove definite) di funzioni continue danno luogo, come sappiamo, a funzioni che risultano continue in ogni punto nel quale sono definite.

Ad esempio sono anche continue su $\mathbb{R}$ le *funzioni iperboliche*

$$\sinh x = \frac{e^x - e^{-x}}{2} \,,$$
$$\cosh x = \frac{e^x + e^{-x}}{2} \,.$$

Ma, senza bisogno di altri calcoli, possiamo anche concludere ad esempio che la funzione

$$f(x) = \frac{\log\left(1+\sqrt{|\sin^2 x - \cos^3(x^2)|}\right)}{e^{\frac{x^2-3x+1}{x}}+5}$$

è continua ovunque è definita, cioè per $x \neq 0$. In effetti tale funzione si ottiene con somme, prodotti, rapporti e composizioni di alcune tra le funzioni dell'elenco precedente.

La conoscenza di queste classi di funzioni continue ci permette anche di calcolare varie relazioni di limite come mostriamo nel seguente:

Esempio 60 Consideriamo la funzione

$$f(x) = \sin\frac{1}{x}$$

defnita su $\mathbb{R} \setminus \{0\}$. Vogliamo calcolarne il limite per $x \to +\infty$. Sappiamo che $1/x \to 0$ per $x \to +\infty$. Segue allora dalla continuità della funzione seno e dalla Proposizione 4.19 sulla composizione di limiti che

$$\lim_{x\to+\infty} \sin\frac{1}{x} = \lim_{y\to 0} \sin y = \sin 0 = 0$$

Esercizio 4.8 Calcolare i seguenti limiti:

$$\lim_{x\to 0} e^{-\frac{1}{x^2}}, \qquad \lim_{x\to+\infty} \cos\left(\frac{x+1}{x^2+1}\right), \qquad \lim_{x\to-\infty} \ln\cos\frac{\sin x+6}{x^2+3}, \qquad \lim_{x\to 0+} e^{-e^{\frac{1}{x}}}.$$

R: 0, 1, 0, 0.

4.8 Discontinuità

Quando una funzione è discontinua, può essere utile avere informazioni ulteriori riguardo al *tipo* di discontinuità della funzione che stiamo studiando. I due esempi visti in precedenza (la funzione a gradino e la funzione $f(x) = \sin(1/x)$) hanno comportamenti del tutto diversi tra loro. Nel primo caso per $x \to 0$ da destra e da sinistra la funzione tende a valori diversi. Nel secondo caso non esistono proprio neppure i limiti destro e sinistro.

Diamo qui di seguito una classificazione molto rozza dei tipi di discontinuità. Ne troverete varie altre in altri testi, ma non ci è necessario scendere in maggior dettaglio.

Definizione 4.22 Sia $f :]a, b[\to \mathbb{R}$ una funzione e $x_0 \in]a, b[$. Si supponga che f non sia continua in x_0 (si dirà allora che f è discontinua in x_0). Diremo allora che f ha in x_0 una discontinuità:

- *eliminabile*, se $\lim_{x\to x_0} f(x) = l \in \mathbb{R}$ ma $l \neq f(x_0)$, cioè se il limite in x_0 esiste finito ma è diverso dal valore della funzione in tale punto;
- *di prima specie*, se esistono finiti
$$\lim_{x\to x_0^+} f(x) = l_1 \in \mathbb{R}\,,$$
$$\lim_{x\to x_0^+} f(x) = l_2 \in \mathbb{R}\,,$$
ma tali valori sono diversi tra loro;
- *di seconda specie* se la discontinuità non è né eliminabile né di prima specie.

La Figura 4.5 mostra discontinuità di prima e seconda specie.

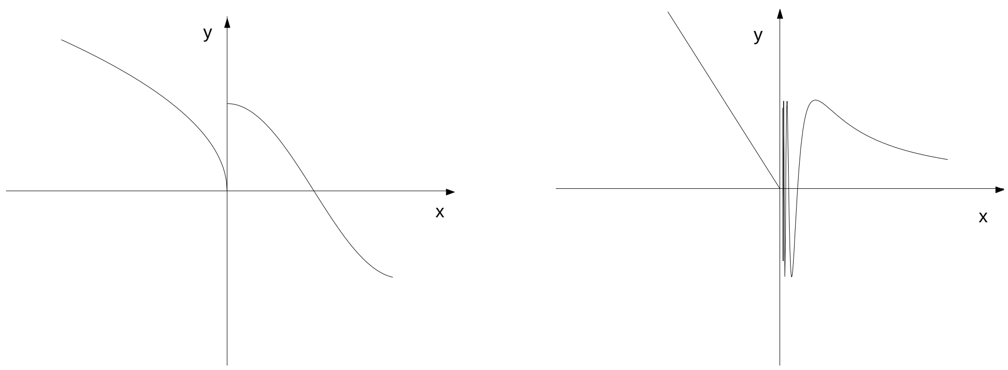


Figura 4.5:

Esempio 61 La funzione
$$f(x) = \begin{cases} x & \text{se } x \neq 0 \\ 1 & \text{se } x = 0 \end{cases}$$
ammette una discontinuità eliminabile in $x = 0$. In effetti ciò significa che basta modificare la definizione di f in un solo punto (in questo caso $x = 0$) per ottenere una funzione continua. La discontinuità non è una caratteristica per così dire costitutiva della funzione, ma è stata imposta in un certo senso “artificialmente.

Esempio 62 La funzione a gradino (4.2) ha una discontinuità di prima specie nell'origine: il limite da destra vale uno mentre quello da sinistra vale zero.

Esempio 63 La funzione (4.3) ha una discontinuità di seconda specie nell'origine, poichè il limite per $x \to x_0$ di tale funzione non esiste. In effetti ciò segue dai commenti che seguono la definizione (4.3) utilizzando il Teorema 4.13.

Peraltro vi sono esempi di funzioni con discontinuità di seconda specie dalle caratteristiche del tutto diverse:

Esempio 64 La funzione

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

ha una discontinuità di seconda specie nell'origine: i limiti destro e sinistro esistono, ma valgono, rispettivamente, $+\infty$ e $-\infty$.

Esercizio 4.9 Stabilire la continuità o il tipo di discontinuità delle seguenti funzioni per $x = 0$:

$$f(x) = \begin{cases} e^{-\frac{1}{|x|}} & x \neq 0 \\ 0 & x = 0 \end{cases} \qquad f(x) = \begin{cases} \sin x & x \leq 0 \\ \cos x & x > 0 \end{cases}$$

$$f(x) = \begin{cases} \ln|x| & x \neq 0 \\ 1 & x = 0 \end{cases} \qquad f(x) = \begin{cases} x \sin\frac{1}{x} & x \neq 0 \\ 0 & x = 0 \end{cases}$$

R: (da sinistra a destra e dall'alto al basso) cont, I, II, cont.

4.9 Limiti notevoli

In questa sezione, presentiamo alcuni importanti limiti notevoli, alcuni collegati a limiti notevoli di successioni visti in precedenza.

Proposizione 4.23

$$\lim_{x \to +\infty} \left(1 + \frac{1}{x}\right)^x = e\,. \tag{4.7}$$

Dimostrazione (Idea). Ricordiamo che vale

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

Per esercizio (*) si faccia poi vedere che se $a_n \to +\infty$ per $n \to +\infty$, allora

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

Usando allora il Teorema 4.13 nella sua variante per limiti all'infinito, si ottiene la tesi per $x \to +\infty$. ■

Osservazione: Vale anche

$$\lim_{n\to-\infty}\left(1+\frac{1}{n}\right)^n = e\,. \tag{4.8}$$

e il risultato si dimostra in modo analogo.

Dal risultato precedente e con l'utilizzo delle varie proprietà dei limiti, si possono ricavare molti altri limiti notevoli illustrati negli esempi che seguono.

Esempio 65 Vale la relazione:

$$\lim_{x\to 0}\frac{\log_a(1+x)}{x} = \log_a e\,.$$

In effetti, per le proprietà dei logaritmi si ha che

$$\frac{\log_a(1+x)}{x} = \log_a\left[(1+x)^{1/x}\right]\,.$$

Noi sappiamo, in base al risultato precedente e alla Proposizione 4.19 di composizione dei limiti, che

$$\lim_{x\to 0+}(1+x)^{1/x} = \lim_{x\to+\infty}\left(1+\frac{1}{x}\right)^x = e\,.$$

Usando nuovamente la composizione dei limiti e la continutà della funzione logaritmo si ha dunque

$$\lim_{x\to 0+}\frac{\log_a(1+x)}{x} = \log_a e\,.$$

Similmente si mostra la stessa relazione per il limite sinistro e si ha dunque la tesi.

Esempio 66 Vale la relazione:

$$\lim_{x\to 0}\frac{a^x-1}{x}=\ln a$$

per ogni $a>0$, $a\neq 1$. Facciamo la sostituzione $a^x-1=t$ equivalente a $x=\log_a(t+1)$. Si noti che quando $t\to 0$ si ha che $x\to 0$. Per la composizione dei limiti e per l'esempio precedente si ha dunque,

$$\lim_{x\to 0}\frac{a^x-1}{x}=\lim_{t\to 0}\frac{t}{\log_a(t+1)}=\frac{1}{\log_a e}=\ln a\,.$$

Esempio 67 Vale la relazione:

$$\lim_{x\to 0}\frac{(1+x)^a-1}{x}=a\,.$$

per ogni $a\in\mathbb{R}$. Scriviamo,

$$\frac{(1+x)^a-1}{x}=\frac{(1+x)^a-1}{\ln(1+x)}\frac{\ln(1+x)}{x}\,.$$

Già sappiamo, dall'Esempio 65, che

$$\lim_{x\to 0}\frac{\ln(1+x)}{x}=1\,.$$

Dobbiamo calcolare il limite dell'altro pezzo. Poniamo $t=(1+x)^a-1$ che equivale a $\ln(1+x)=\ln(1+t)/a$. Si noti inoltre che per $x\to 0$, si ha anche $t\to 0$ (in base alla continuità della funzione esponenziale). Dunque componendo i limiti e sfruttando di nuovo l'Esempio 65 otteniamo

$$\lim_{x\to 0}\frac{(1+x)^a-1}{\ln(1+x)}=\lim_{t\to 0}\frac{ta}{\ln(1+t)}=a\,.$$

Presentiamo ora un esempio che mostra come questi limiti notevoli possano essere utilizzati per il calcolo di limiti.

Esempio 68 Supponiamo di voler calcolare

$$\lim_{x\to+\infty}\left(\frac{x^2-1}{x^2+1}\right)^x\,.$$

Si tratta di una indeterminazione del tipo 1^∞. Possiamo scrivere

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

Si noti ora che per la (4.8) e la solita composizione di limiti, si ha che

$$\lim_{x\to+\infty}\left(1+\frac{1}{\frac{x^2+1}{-2}}\right)^{\frac{x^2+1}{-2}}=e$$

D'altra parte,

$$\lim_{x\to+\infty}\frac{-2}{x^2+1}x=0\,.$$

Si ha dunque,

$$\lim_{x\to+\infty}\left(\frac{x^2-1}{x^2+1}\right)^x=e^0=1\,.$$

Esempio 69 Consideriamo ora una situazione un po' più complessa:

$$\lim_{x\to e}(\ln x)^{\frac{1}{x-e}}\,.$$

Si tratta ancora di una indeterminazione del tipo 1^∞. Rappresentiamo

$$(\ln x)^{\frac{1}{x-e}}=e^{\ln\left[(\ln x)^{\frac{1}{x-e}}\right]}\,.$$

Si noti che questo è un passaggio che si può sempre fare per qualsiasi funzione; è la cosidetta rappresentazione esponenziale. Studiamo

$$\lim_{x\to e}\ln\left[(\ln x)^{\frac{1}{x-e}}\right]\,.$$

Possiamo scrivere

$$\ln\left[(\ln x)^{\frac{1}{x-e}}\right]=\frac{\ln(\ln x)}{x-e}\,.$$

Si noti che ci siamo riportati ad una indeterminazione del tipo 0/0. D'altra parte,

$$\frac{\ln(\ln x)}{x-e}=\frac{\ln\left(\ln e\frac{x}{e}\right)}{x-e}=\frac{\ln\left(1+\ln\frac{x}{e}\right)}{x-e}=\frac{\ln\left(1+\ln\frac{x}{e}\right)}{\ln\frac{x}{e}}\,\frac{\ln\frac{x}{e}}{x-e}$$

Si noti ora che poichè, per la continuità del logaritmo, $\ln(x/e)\to 0$ per $x\to e$ segue dall'Esempio 65 che

$$\lim_{x\to e}\frac{\ln\left(1+\ln\frac{x}{e}\right)}{\ln\frac{x}{e}}=1$$

D'altra parte,

$$\frac{\ln\frac{x}{e}}{x-e}=\frac{\ln\left(1+\left(\frac{x}{e}-1\right)\right)}{e\left(\frac{x}{e}-1\right)}$$

converge a 1 per $x\to e$ sempre in virtù dell'Esempio 65. Si ha dunque

$$\lim_{x\to e}(\ln x)^{\frac{1}{x-e}}=\frac{1}{e}\,.$$

Vi sono poi vari altri limiti notevoli di importanza essenziale, la cui dimostrazione può essere trovata su qualsiasi testo di Analisi Matematica ed è probabilmente ben nota a molti di voi. Ci limitiamo a riportare qui di seguito senza dimostrazione tali risultati. Una dimostrazione alternativa verrà comunque data più avanti con l'utilizzo degli sviluppi di Taylor.

$$\lim_{x\to 0} \frac{\sin x}{x} = 1\,, \tag{4.9}$$

$$\lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2}\,, \tag{4.10}$$

Esercizio 4.10 Mostrare che

$$\lim_{x\to 0} \frac{\operatorname{tg} x}{x} = 1\,. \tag{4.11}$$

Esercizio 4.11 Mostrare che

$$\lim_{x\to 0} \frac{\sinh x}{x} = 1\,.$$

Esercizio 4.12 Calcolare, se esistono,

$$\lim_{x\to 0} \frac{\sin(x\cos x)}{x}\,, \qquad \lim_{x\to 0} \frac{\ln(\cos x)}{x^2}$$

Conludiamo enunciando, senza dimostrazione, altri tre limiti notevoli che possono essere di qualche utilità e che coinvolgono le funzioni trigonometriche inverse e che possono essere facilmente dimostrati con l'ausilio della Proposizione 4.19 della composizione di limiti:

$$\lim_{x\to 0} \frac{\operatorname{artg} x}{x} = 1\,,$$

$$\lim_{x\to 0} \frac{\operatorname{arsin} x}{x} = 1\,,$$

$$\lim_{x\to 0} \frac{1-\operatorname{arcos} x}{x} = 1\,.$$

4.10 Equivalenza asintotica, parte principale

Abbiamo visto che un modo per capire come si comporta, ad esempio, la funzione $\cos x$ vicino all'origine $x = 0$ è quello di prendere i primi termini del suo sviluppo in serie. Se $|x|$ è piccolo abbiamo visto che

$$\frac{\cos x - 1}{-x^2/2}$$

è vicino a uno, nel senso che il limite dell'espressione precedente per $x \to 0$ vale proprio uno. Possiamo scrivere *formalmente* che

$$\cos x - 1 \underset{x\to 0}{\sim} -\frac{x^2}{2}$$

dove però il simbolo "$\sim$ non significa ancora niente di preciso, ma è usato per descrivere il fatto che il membro di sinistra "è ben approssimato, per $x \to 0$, da quello di destra. Vediamo ora alcune definizioni che ci permetteranno di avere utili notazioni con cui operare.

Definizione 4.24 Siano f e g due funzioni definite su un intervallo $]a, b[$, e sia $x_0 \in]a, b[$. Supponiamo inoltre che esista $\delta > 0$ tale che $g(x) \neq 0$ se $x \neq x_0$ è tale che $|x-x_0| \leq \delta$. Allora si dirà che f e g sono asintoticamente equivalenti (o, più brevemente, asintotiche) quando $x \to x_0$, se vale la condizione

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = 1\,. \tag{4.12}$$

Si scriverà in tal caso

$$f(x) \underset{x\to x_0}{\sim} g(x)\,.$$

La definizione data è ragionevole: diciamo che f e g sono asintotici per $x \to x_0$ se il loro rapporto tende a uno in tale limite.

Esempio 70 Riformuliamo alcuni dei limiti notevoli visti prima in termini di equivalenza asintotica. Ad esempio

$$\begin{aligned} e^x - 1 &\underset{x\to 0}{\sim} x\,, \\ \ln(1+x) &\underset{x\to 0}{\sim} x\,, \\ \sin x &\underset{x\to 0}{\sim} x\,, \\ \cos x - 1 &\underset{x\to 0}{\sim} -\frac{x^2}{2}\,. \end{aligned}$$

Osservazione: Siano f_1, f_2, g_1 e g_2 quattro funzioni definite su un intervallo $]a,b[$ sempre diverse da 0, e sia $x_0 \in]a,b[$. Supponiamo che

$$f_1(x) \underset{x\to x_0}{\sim} f_2(x)\,, \qquad g_1(x) \underset{x\to x_0}{\sim} g_2(x)\,.$$

Allora, segue dalla definizione stessa di equivalenza asintotica che i due rapporti,

$$\frac{f_1(x)}{g_1(x)}\,, \qquad \frac{f_2(x)}{g_2(x)}$$

hanno lo stesso comportamento asintotico per $x{\to}x_0$ nel senso che se una delle due funzioni rapporto ammette limite per $x \to x_0$, anche l'altra ammette limite e i limiti sono uguali. Tutto segue in effetti dall'identit`:

$$\frac{f_1(x)}{g_1(x)} = \frac{f_1(x)}{f_2(x)}\frac{f_2(x)}{g_2(x)}\frac{g_2(x)}{g_1(x)}\,,$$

sfruttando la definizione di equivalenza asintotica. Questo fatto è molto importante: dice che si può sostituire, ai fini di un calcolo di un limite, il numeratore e il denominatore di una frazione, con espressioni ad esse asintoticamente equivalenti.

Un altro concetto importante è il seguente:

Definizione 4.25 Siano f e g due funzioni definite su un intervallo $]a,b[$, e sia $x_0 \in]a,b[$. Supponiamo inoltre che esista $\delta > 0$ tale che $g(x) \neq 0$ se $x \neq x_0$ è tale che $|x - x_0| \leq \delta$. Allora si dirà che f è trascurabile rispetto a g quando $x \to x_0$ se vale la condizione

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = 0\,. \tag{4.13}$$

Si scriverà in tal caso che

$$f(x) = o(g(x)) \quad \text{per } x \to 0\,.$$

Si suole anche dire in tal caso che f è un “o piccolo di g.

Anche questa definizione è ragionevole: diciamo che f è trascurabile rispetto a g per $x \to x_0$ se il loro rapporto tende a zero in tale limite.

Esercizio 4.13 Dimostrare che, se $h > k > 0$ allora $x^h = o(x^k)$ quando $x \to 0^+$, cioè quando x si avvicina a zero da destra (vale a dire mantenendosi positivo).

Vi è una semplice relazione tra il concetto di “o piccolo e quello di equivalenza asintotica.

Lemma 4.26 *Siano f e g due funzioni definite su un intervallo $]a,b[$, e sia $x_0 \in]a,b[$. Allora*

$$f(x) \underset{x\to x_0}{\sim} g(x) \Longleftrightarrow f(x) = g(x) + o(g(x))\,.$$

Dimostrazione $\Rightarrow$: La condizione $f \sim g$ implica che

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

con $h(x) \to 0$ per $x \to x_0$. Ma allora, poiché $h(x)g(x) = o(g(x))$ si avrà che

$$f(x) = g(x) + g(x)h(x) = g(x) + o(g(x)) \quad \text{per } x \to 0\,.$$

$\Leftarrow$: segue invertendo i ragionamenti precedenti. ■

Esempio 71 Riformuliamo alcuni dei limiti notevoli visti prima in termini di ‘o piccolo’, per $x \to 0$:

$$\begin{aligned} e^x &= 1 + x + o(x)\,, \\ \ln(1+x) &= x + o(x)\,, \\ \sin x &= x + o(x)\,, \\ \cos x &= 1 - \tfrac{x^2}{2} + o(x^2)\,. \end{aligned}$$

Enunciamo ora, senza dimostrazione, alcune ulteriori proprietà elementari dei simboli appena introdotti.

Proposizione 4.27 *Siano f, g, h, k funzioni definite su un intervallo $]a, b[$, e siano $x_0 \in]a, b[$ e $l \in \mathbb{R}$. Si ha che*

- $f(x) = o(1)$ *per* $x \to x_0$ *se e soltanto se* $f(x) \to 0$ *per* $x \to x_0$*;*
- $f(x) \sim l$ *per* $x \to x_0$ *se e soltanto se* $f(x) \to l$ *per* $x \to x_0$*;*
- $f(x) \sim l$ *per* $x \to x_0$ *se e soltanto se* $f(x) = l + o(1)$ *per* $x \to x_0$*;*
- $f(x)o(g(x)) = o(f(x)g(x))$ *per* $x \to x_0$*;*
- *se* $f(x) = o(g(x))$ *per* $x \to x_0$ *allora* $o(f(x)) + o(g(x)) = o(g(x))$ *per* $x \to x_0$*;*
- *se* $f(x) = o(g(x))$ *e* $g(x) \sim h(x)$ *per* $x \to x_0$ *allora* $f(x) = o(h(x))$ *per* $x \to x_0$*;*
- *se* $f(x) \sim g(x)$ *e* $h(x) \sim k(x)$ *per* $x \to x_0$ *allora*

  $$f(x)h(x) \sim g(x)k(x), \qquad \frac{f(x)}{h(x)} \sim \frac{g(x)}{k(x)}$$

  per $x \to x_0$ *purché i denominatori siano diversi da zero vicino a* x_0 *salvo al più nel punto* x_0*;*
- *se* $f(x) \sim g(x)$ *e* $h(x) \sim k(x)$ *per* $x \to x_0$ *allora*

  $$f(x) + h(x) \sim g(x) + k(x)$$

  per $x \to x_0$ *purché le funzioni coinvolte abbiano tutte lo stesso segno vicino a* x_0*.*

In tutto quanto abbiamo fatto precedentemente, x_0 era un numero reale. Non ci sono però difficoltà a ripetere parola per parola quanto detto nel caso in cui x_0 sia $\pm\infty$. Usiamo questo fatto nel seguente esempio.

Esempio 72 Discutiamo qui alcuni limiti notevoli, l'analogo di parte dei quali è stato visto in precedenza per le successioni (si veda l'esempio 40). Ad esempio si dimostra che

$$\lim_{x\to+\infty} \frac{x^\alpha}{a^x} = 0 \tag{4.14}$$

quali che siano i parametri $\alpha > 0$, $\beta > 1$. Questo si può scrivere, con le nostre notazioni, così :

$$x^\alpha = o\left(a^x\right) \quad \text{per } x \to +\infty$$

se $\alpha > 0$, $\beta > 1$.

Come si dimostra la validità del limite precedente? Diamone solo un cenno. Sappiamo già che

$$\lim_{n\to+\infty} \frac{n^\alpha}{a^n} = 0 \tag{4.15}$$

ed è possibile dimostrare che, se al posto della successione $c_n = n$ in (4.15) si prende una *qualsiasi* successione c_n che diverge a $+\infty$, si ottiene egualmente

$$\lim_{n\to+\infty} \frac{{c_n}^\alpha}{a^{c_n}} = 0\,. \tag{4.16}$$

Se (4.16) è vera allora il Lemma 4.13 ci garantisce che (4.14) è vera. Si mostra con la stessa linea di ragionamento (prima per le successioni che tendono a $+\infty$, poi per la funzione) che:

$$\lim_{x\to+\infty} \frac{(\log x)^a}{x^\alpha} = 0\,,$$

per ogni $a, \alpha > 0$. Questo si scrive anche come

$$(\log x)^a = o(x^\alpha) \quad \text{per } x \to +\infty\,.$$

Ci restano da introdurre alcuni altri concetti importanti: quelli di *ordine di infinitesimo*, di *ordine di infinito* e di *parte principale*. Per introdurre l'argomento pensate al fatto che, nella discussione precedente, abbiamo visto che, spesso, è possibile approssimare funzioni assegnate, la cui espressione può essere molto complicata, con funzioni assai più semplici, i *polinomi*: vedremo in seguito che è anche possibile dare stime numeriche dell'errore che si compie.

Cominciamo con un esempio.

Esempio 73 Abbiamo dimostrato in precedenza che vale

$$f(x) := \cos x - 1 = -\frac{x^2}{2} + o(x^2) \quad \text{per } x \to 0\,. \tag{4.17}$$

Si noti che non c'era dubbio che la funzione a sinistra avesse limite zero per $x \to 0$: noi abbiamo concluso che la funzione si avvicina a zero "come $-x^2/2$, a meno ovviamente di termini trascurabili rispetto a x^2 stesso. Diremo in questo caso quanto segue:

- $f(x)$ ha ordine di infinitesimo pari a due rispetto all'infinitesimo campione $g(x) = x$;
- la parte principale di $f(x)$ per $x \to 0$ è $-x^2/2$.

Il significato di queste affermazioni dovrebbe essere chiaro: se cerco di confrontare f con le potenze di x, (4.17) mostra che l'esponente corretto è due. Inoltre, una volta stabilito questo fatto, tra i monomi della forma cx^2 con $c \in \mathbb{R}$, quello che approssima correttamente la funzione vicino a zero (a meno di termini trascurabili rispetto a x^2 stesso) è $-x^2/2$.

Diamo quindi la seguente definizione.

Definizione 4.28 Sia $f :]a,b[\to \mathbb{R}$ una funzione, e $x_0 \in]a,b[$. Sia $g :]a,b[\to \mathbb{R}$ un'altra funzione fissata, e si supponga che

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

(rispettivamente che

$$\lim_{x\to x_0} f(x) = \lim_{x\to x_0} g(x) = +\infty)\,.$$

Diremo che f ha ordine di infinitesimo (risp. di infinito) $k \in \mathbb{N}$ rispetto all'infinitesimo campione (risp. all'infinito campione) g se

$$f(x) = cg(x)^k + o(g(x)^k) \quad \text{per } x \to x_0$$

per un opportuno $c \neq 0$. In tal caso si dice che $cg(x)^k$ è la parte principale di f (fissato il campione g) quando $x \to x_0$.

Osservazione: Si noti che nella definizione sopra $f(x)$ è equivalente alla sua parte principale, per $x \to x_0$. Quando si deve calcolare un limite di una frazione in forma indeterminata $0/0$ si può dunque calcolare la parte principale del numeratore e del denominatore e sostituirlio ad essi nel calcolo del limite. Vedremo più avanti moltissime applicazioni di questo principio.

Osservazione: Si può anche considerare ordini di infinitesimo e di infinito non interi nella definizione sopra, purchè si considerino funzioni campione positive in un intorno del punto x_0.

Una definizione identica può essere data anche per $x_0 = \pm\infty$, e con modifiche minori anche nel caso in cui f e g tendano a $-\infty$. Evitiamo di ripetere parola per parola quanto scritto sopra.

Osserviamo che nella maggior parte dei casi si sceglie $g(x) = x - x_0$ (se f tende a zero per $x \to x_0$) o $g(x) = x$ (se f ha limite $\pm\infty$ per $x \to x_0$). A meno che non venga esplicitamente detto il contrario, supporremo sempre di aver fatto tale scelta. Osservate anche che *non è affatto detto* che, fissato il campione g, l'ordine di infinitesimo e la parte principale di f rispetto a g siano ben definiti: pensate ad esempio al caso

$$f(x) = e^{-1/x^2}, \quad g(x) = x\,.$$

x

Esercizio 4.14 Calcolare la parte principale delle seguenti funzioni:

$$f(x) = e^{[\ln(1-x^2)]^2} - \cos x \quad \text{per } x \to 0\,,$$

$$f(x) = \frac{(1-2\cos x)^2}{\sin(x-\pi/3)} \qquad \text{per } x \to \pi/3\,,$$

$$f(x) = \frac{(x^2-5x+6)^2}{\ln(x-2)} \qquad \text{per } x \to 3\,.$$

R: $x^2/2,\ 3(x-\pi/3),\ (x-3)$.

4.11 Altre proprietà delle funzioni continue

Studieremo in questa sezione due proprietà fondamentali delle funzioni continue; l'esistenza dei valori intermedi e l'esistenza di massimo e minimo per funzioni definite su un intervallo chiuso e limitato $[a, b]$.

Entrambi questi risultati sono di notevole importanza. Il primo di essi riguarda l'esistenza di zeri per funzioni continue che passino da valori positivi a valori negativi e, più in generale, il fatto che tutti i valori intermedi tra due valori assunti da una funzione continua sono anch'essi valori effettivamente assunti da tale funzione. Intuitivamente non è difficile convincersi di questo fatto: se, ad esempio, una funzione continua f vale uno in $x = 0$ e tre in $x = 1$, il fatto che il suo grafico possa essere disegnato "senza staccare la penna dal foglio dovrà ragionevolmente implicare che, dato un qualunque valore l tra zero e tre, ci sia un punto $\overline{x} \in]0, 1[$ in cui la funzione vale l, cioè che il grafico della funzione intersechi la retta orizzontale $y = l$.

Enunciamo e dimostriamo il primo di tali risultati.

Teorema 4.29 (degli zeri) *Sia $f : [a, b] \to \mathbb{R}$ una funzione continua su $[a, b]$ tale che $f(a)f(b) < 0$, cioè tale che $f(a)$ e $f(b)$ abbiano segno diverso. Allora esiste $\overline{x} \in]a, b[$ tale che $f(\overline{x}) = 0$.*

Dimostrazione Supponiamo per fissare le idee che $f(a) < 0$ e $f(b) > 0$ e dimostriamo il Teorema in tale ipotesi. Se tale condizione non fosse vera per f allora lo sarebbe per $-f$, e quindi la dimostrazione fornirebbe uno zero per $-f$, quindi per f stessa.

Definiamo allora $x_0 = a$, $y_0 = b$, e sia $z_0 = (x_0 + y_0)/2$ il punto medio dell'intervallo $]a, b[$. Se $f(z_0) = 0$ non c'è altro da dimostrare. Se invece $f(z_0) \neq 0$ definiamo la coppia (x_1, y_1) come segue:

$$(x_1, y_1) = \begin{cases} (x_0, z_0) & \text{se } f(z_0) > 0 \\ (z_0, y_0) & \text{se } f(z_0) < 0. \end{cases}$$

Si noti che $|x_1 - y_1| = |x_0 - y_0|/2$. In ogni caso poi si ha $f(x_1) < 0$, $f(y_1) > 0$ e inoltre $x_1 \geq x_0$, $y_1 \leq y_0$. Ripetiamo la costruzione sull'intervallo $[x_1, y_1]$, ponendo $z_1 = (x_1 + x_2)/2$: z_1 è il punto medio di $[x_1, x_2]$. Di nuovo se $f(z_0) = 0$ non c'è altro da dimostrare, altrimenti si definisca

$$(x_2, y_2) = \begin{cases} (x_1, z_1) & \text{se } f(z_1) > 0 \\ (z_1, y_1) & \text{se } f(z_1) < 0. \end{cases}$$

e si noti che $|x_2 - y_2| = |x_0 - y_0|/4$, che $f(x_2) < 0$, $f(y_2) > 0$, e inoltre che $x_2 \geq x_1$, $y_2 \leq y_1$. Dovremmo ormai aver capito l'idea: si procede per ricorrenza costruendo la coppia (x_n, y_n) in termini di quella (x_{n-1}, y_{n-1}), per ogni $n \in \mathbb{N}$. Per far ciò si definisce, per ogni $n \in \mathbb{N}$, prima $z_{n-1} = (x_{n-1} + y_{n-1})/2$, punto medio di (x_{n-1}, y_{n-1}), e poi

$$(x_n, y_n) = \begin{cases} (x_{n-1}, z_{n-1}) & \text{se } f(z_{n-1}) > 0 \\ (z_{n-1}, y_{n-1}) & \text{se } f(z_{n-1}) < 0. \end{cases}$$

Si ha allora $|x_n - y_n| = 2^{-n}|x_0 - y_0|$, che $f(x_n) < 0$, $f(y_n) > 0$ e inoltre che $x_n \geq x_{n-1}$, $y_n \leq y_{n-1}$ per ogni $n \in \mathbb{N}$, a meno che naturalmente per qualche $\hat{n}$ il procedimento non si arresti nel senso che $f(z_{\hat{n}}) = 0$, caso nel quale non c'è però altro da dimostrare.

Abbiamo quindi costruito, se il procedimento non si arresta, due successioni (x_n), (y_n) con le seguenti proprietà:

1. (x_n) è monotona crescente, (y_n) è monotona decrescente;
2. $x_n < y_n \forall n \in \mathbb{N}$;
3. $|x_n - y_n| = 2^{-n}|x_0 - y_0| \forall n \in \mathbb{N}$;
4. $f(x_n) < 0$, $f(y_n) > 0 \forall n \in \mathbb{N}$.

Sappiamo allora per il punto 1) che esiste il limite di (x_n) quando $n \to +\infty$. Tale limite coincide sempre con $\sup\{x_n : n \in \mathbb{N}\}$ e quindi nel presente caso è finito, in quanto per costruzione sappiamo che $x_n \leq b$ per ogni $n \in \mathbb{N}$, cosicché (x_n) è superiormente limitata.

Chiamiamo l_1 tale limite. Analogamente si mostra che il limite di (y_n) quando $n \to +\infty$ esiste anch'esso finito: chiamiamo l_2 tale limite, e osserviamo che $l_1 \leq l_2$ perché, per il punto 2),

$$\sup\{x_n : n \in \mathbb{N}\} \leq \inf\{y_n : n \in \mathbb{N}\}. \tag{4.18}$$

Mostriamo ora che $l_1 = l_2$. In effetti

$$0 \leq |l_1 - l_2| \leq |l_1 - x_n| + |x_n - y_n| + |y_n - l_2|\,. \tag{4.19}$$

Per ipotesi, $(|l_1 - x_n|)$ e $(|y_n - l_2|)$ sono infinitesime, così come $(|x_n - y_n|)$ per il punto 3.. Segue dunque che il secondo membro di (4.19) è infinitesimo. Per il Teorema del confronto di successioni, segue che anche la successione costante $(|l_1 - l_2|)$ deve essere infinitesima. Questo implica $l_1 = l_2$. Chiamiamo $\overline{x}$ questo valore comune.

Chiaramente $\overline{x} \in]a, b[$. Utilizziamo ora finalmente la continuità di f in ogni punto $x \in [a, b]$ (si osservi che fino a ora questo non era stato fatto), e in particolare nel punto $x = \overline{x}$. Poiché sia (x_n) che (y_n) convergono a $\overline{x}$ e f è continua per successioni per il Lemma 4.13, si ha che

$$f(\overline{x}) = \lim_{n\to+\infty} f(x_n) = \lim_{n\to+\infty} f(y_n)\,.$$

Ma d'altronde il Teorema 3.4 della permamenza del segno per le successioni ci assicura che

$$\lim_{n\to+\infty} f(x_n) \leq 0,\ \lim_{n\to+\infty} f(y_n) \geq 0\,,$$

e quindi che $f(\overline{x}) = 0$. ■

Esercizio 4.15 * Dimostrare che il polinomio $P(x) = x^3 - 4x^2 + x + 1$ ha tre radici reali, una delle quali negativa, una seconda delle quali compresa tra zero e uno, la terza delle quali maggiore di uno.

Vi sarete chiesti se la scelta del valore $y = 0$ abbia qualche particolare importanza. La risposta è no, come mostra il seguente Corollario.

Corollario 4.30 *Sia $f : [a,b] \to \mathbb{R}$ una funzione continua su $[a,b]$. Allora f assume tutti i valori compresi nell'intervallo*

$$[\min\{f(a), f(b)\}, \max\{f(a), f(b)\}]$$

Dimostrazione È sufficiente, detto c un qualunque valore appartenente all'intervallo sopra indicato, applicare il Teorema precedente a $g(x) := f(x) - c$, osservando che, per come c è stato scelto, le due quantità $f(a) - c$ e $f(b) - c$, se entrambe non nulle, sono sicuramente di segno opposto. ■

Il prossimo risultato è forse il più importante nello studio delle funzioni continue da noi intrapreso.

Diamo prima una definizione.

Definizione 4.31 Sia $f : A \subset \mathbb{R} \to \mathbb{R}$. Un punto $x_0 \in A$ si dice

- *punto di massimo assoluto* per f su A se

$$f(x_0) \geq f(x) \quad \forall x \in A\,;$$

- *punto di minimo assoluto* per f su A se

$$f(x_0) \leq f(x) \quad \forall x \in A\,.$$

I corrispondenti valori assunti dalla f si dicono, rispettivamente *il massimo ed il minimo assoluto della funzione* f su A.

Si noti come mentre i punti di massimo e di minimo possono essere svariati, il massimo assoluto ed il minimo assoluto di una funzione, se esistono, sono unici. Possiamo caratterizzarli come il massimo ed il minimo dell'insieme immagine $f(A)$: $\max f(A)$ e $\min f(A)$.

Il Teorema che andremo ad enunciare e dimostrare ci dice che una funzione continua su un intervallo della forma $[a,b]$ ammette necessariamente massimo e minimo assoluti, cioè che esistono $x_1, x_2 \in [a,b]$ tali che

$$\begin{aligned} f(x_1) &\geq f(x) \quad \forall x \in [a,b]\,, \\ f(x_2) &\leq f(x) \quad \forall x \in [a,b]\,. \end{aligned}$$

Cominciamo con qualche esempio, per vedere se c'è qualche speranza che il risultato sopra descritto valga anche senza qualcuna delle ipotesi appena introdotte.

Esempio 74 Se f non è continua su $[a,b]$ allora il massimo e il minimo assoluto possono non esistere, come mostra la funzione $f:[-1,1]\to\mathbb{R}$ definita da

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

In effetti potete facilmente convincervi che l'unico candidato a essere massimo assoluto per la funzione è il valore uno: ma non esiste nessun $x\in[-1,1]$ per cui $f(x)=1$. Analogamente per il minimo.

Esempio 75 È anche essenziale che l'intervallo di definizione sia un intervallo chiuso, cioè che contenga i suoi estremi. Se si lavorasse su un intervallo aperto $]a,b[$ la conclusione sarebbe falsa anche per funzioni continue. Si pensi alla funzione dell'esempio precedente, definita sull'insieme $]-1,1[$, che analogamente a quanto detto prima non ammette massimo e minimo assoluti pur essendo limitata e *continua* sull'intervallo considerato, o alla funzione $f:]-\pi/2,\pi/2[\to\mathbb{R}$ definita da

$$f(x)=\operatorname{tg} x$$

che è invece addirittura illimitata (sia dal basso che dall'alto) nell'intervallo considerato.

Esempio 76 È infine necessario lavorare su un intervallo limitato. Basti pensare alla funzione $f:\mathbb{R}\to\mathbb{R}$ definita da

$$f(x)=\operatorname{artg} x$$

che è continua, monotona crescente e limitata su $\mathbb{R}$, ma non ammette massimo né minimo assoluti su $\mathbb{R}$, in quanto in tal caso $f(\mathbb{R})=]-\pi/2,\pi/2[$. $\pi/2$ è dunque l'estremo superiore dei valori assunti dalla funzione, ma non è il massimo in quanto non è mai assunto, non vi è infatti alcun $x\in\mathbb{R}$ tale che $\operatorname{artg} x=\pi/2$.

Enunciamo e dimostriamo dunque il fondamentale Teorema di Weierstrass.

Teorema 4.32 (di Weierstrass) *Una funzione $f:[a,b]\to\mathbb{R}$ continua su $[a,b]$ è ivi limitata e ammette massimo e minimo assoluti.*

Dimostrazione La dimostrazione sfrutta alcune delle idee contenute nella dimostrazione del Teorema degli zeri. In effetti, poniamo

$$M=\sup\{f(x):x\in[a,b]\}$$

che in linea di principio può essere un numero reale oppure $+\infty$. Dividiamo ora l'intervallo $I = [a, b]$ in due sottointervalli di ugual lunghezza, diciamo $J_{1,1} = [a, (a + b)/2]$, $J_{1,2} = [(a + b)/2, b]$ e poniamo

$$M_1 = \sup\{f(x) : x \in [a, (a + b)/2]\},$$
$$M_2 = \sup\{f(x) : x \in [(a + b)/2, b]\}.$$

Se $M = +\infty$ allora significa che non ci sono maggioranti dell'immagine di f quando la si veda come funzione su $[a, b]$: ma allora non ci sono maggioranti neppure dell'immagine di almeno una delle due restrizioni di f a $J_{1,1}$, $J_{1,2}$. In tal caso si ha quindi che M_1 o M_2 è eguale a $+\infty$. Se invece $M \in \mathbb{R}$ abbiamo che $M_1 \leq M$, e che $M_2 \leq M$ visto che gli estremi superiori che compaiono nei membri di sinistra di tali disuguaglianze sono presi su insiemi contenuti in quello che compare nel membro di destra. Quindi $\max\{M_1, M_2\} \leq M$. Inoltre $\max\{M_1, M_2\}$ è un maggiorante dell'immagine di f definita sull'intero $[a, b]$, e quindi $M \leq \max\{M_1, M_2\}$. In conclusione

$$M = \max\{M_1, M_2\}.$$

Quindi l'estremo superiore dell'immagine della funzione definita in uno dei due sottointervalli costruiti deve coincidere con M, estremo superiore dell'immagine della funzione definita su tutto $[a, b]$. Sia I_1 tale sottointervallo (se entrambi avessero la proprietà richiesta, se ne scelga uno qualunque). Procedendo ricorsivamente otteniamo una successione di intervalli (I_n) della forma $I_n = [x_n, y_n]$ tali che:

1. $I_{n+1} \subset I_n \quad \forall n \in \mathbb{N}$;
2. la lunghezza di I_n vale $2^{-n}|b - a|$ per ogni $n \in \mathbb{N}$;
3. $$\sup\{f(x) : x \in I_n\} = \sup\{f(x) : x \in [a, b]\} \quad \forall n \in \mathbb{N}.$$

Come nella dimostrazione del Teorema degli zeri possiamo concludere che le successioni (x_n) e (y_n) convergono entrame a un medesimo valore $l \in [a, b]$.

Siccome f è continua in l, per ogni $\varepsilon > 0$ fissato si ha che

$$f(x) \leq f(l) + \varepsilon$$

se $|x - l| \leq \delta$ opportuno. Questo vuol dire che per n abbastanza grande il numero $f(l) + \varepsilon$ è un maggiorante di f su I_n. Questo significa, per la proprietà 3), che

$$M \leq f(l) + \varepsilon \quad \forall \varepsilon > 0.$$

Quindi $M \leq f(l)$, ma la disuguaglianza opposta $M \geq f(l)$ è evidente dalla definizione dal fatto che M è l'estremo superiore dei valori assunti dalla f in $[a, b]$, quindi è certamente non più piccolo del valore assunto da f in un qualunque punto fissato. Infine

$$M = f(l)$$

cioè

$$f(l) = \sup\{f(x) : x \in [a, b]\},$$

cosicché M non vale $+\infty$ e inoltre x_0 è un punto di massimo assoluto per f. La dimostrazione dell'esistenza di un minimo assoluto per f è analoga. ■

Capitolo 5

Derivate e calcolo differenziale

5.1 Il concetto di derivata

Il concetto di continuità di per sè non quantifica il legame tra la variazione della variabile indipendente x e la corrispondente variazione della variabile dipendente $y = f(x)$. Tale legame è insito nella dipendenza di δ da ϵ nella definizione di continuità ma non viene esplicitamente quantificato.

Studiare più in profondità tale legame (tra variazioni della x e della y) è ovviamente importante da un punto di vista computazionale: per avere il risultato $y = f(x)$ con una determinata precisione, con quale precisione dobbiamo conoscere x? Ma lo è anche per approfondire lo studio qualitativo dei grafici delle funzioni come vedremo tra breve.

Le rette sono sicuramente le funzioni per le quali questo rapporto tra variazione della x e corrispondente variazione della y è il più semplice possibile. In effetti se abbiamo

$$f(x) = mx + n$$

si ha che

$$f(x_1) - f(x_2) = m(x_1 - x_2), \quad \forall x_1, x_2 \in \mathbb{R}\,.$$

Abbiamo dunque, in questo caso, una proporzionalità diretta, con costante di proporzionalità data dal coefficiente angolare m, tra la variazione $x_1 - x_2$ e la corrispondente variazione $f(x_1) - f(x_2)$.

Vogliamo ora introdurre un concetto che permetta di studiare, anche per funzioni non lineari, il rapporto tra la variazione della x e la corrispondente variazione della y. E' chiaro che tale concetto sarà collegato alla 'pendenza' del grafico della funzione stessa.

Partiamo da una funzione $f :]a,b[\to \mathbb{R}$ e sia $x_0 \in]a,b[$. Vogliamo definire il concetto di 'pendenza' del grafico della f nel punto $(x_0, f(x_0))$. Intuitivamente, questa pendenza dovrebbe essere quella della retta 'tangente' al grafico suddetto nel punto $(x_0, f(x_0))$, cioè il suo coefficiente angolare. Il problema è che non abbiamo ancora definito il concetto di retta tangente ad una curva. Questo si può introdurre come segue. Consideriamo un altro punto x sull'asse X e la retta secante il grafico nei punti $(x_0, f(x_0))$ e $(x, f(x))$. Al tendere di x a x_0, tale retta tenderà ad assumere la posizione della tangente, ammesso che tale limite si possa in qualche senso fare.

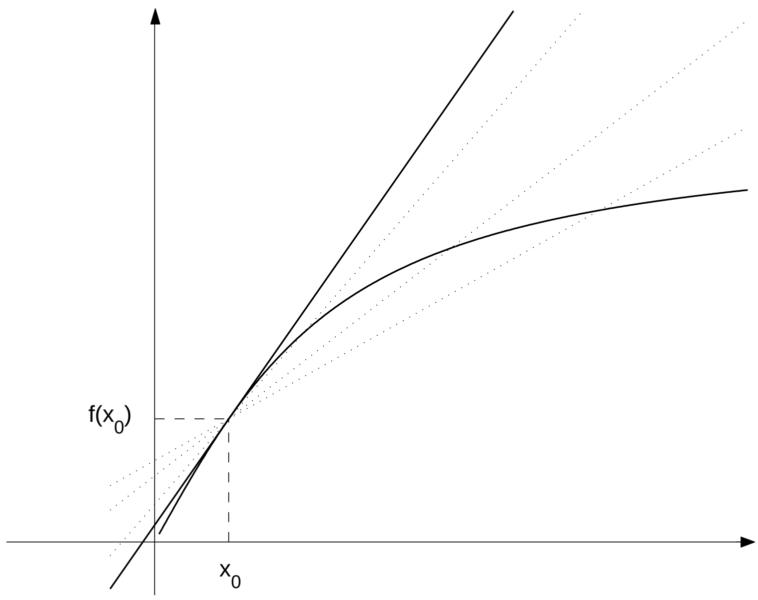


Figura 5.1:

Il coefficiente angolare della secante è dato da

$$\frac{f(x) - f(x_0)}{x - x_0}$$

che è detto il *rapporto incrementale* della funzione f tra x_0 e x. Il coefficiente angolare della tangente si dovrebbe quindi ottenere facendo il limite, ammesso che esista, del rapporto incrementale per $x \to x_0$. Possiamo ora formalizzare la definizione fondamentale:

Definizione 5.1 Sia $f :]a, b[\to \mathbb{R}$ una funzione e sia $x_0 \in]a, b[$. Si dice che la funzione f è *derivabile* in x_0 se esiste finito il limite

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} . \tag{5.1}$$

Tale limite, quando esso esiste, è detto la *derivata* della funzione f nel punto x_0 e denotato con uno dei simboli equivalenti

$$f'(x_0), \quad Df(x_0), \quad \frac{df}{dx}(x_0) .$$

Se la funzione f è derivabile in ogni punto di $]a, b[$ si può definire la *funzione derivata* della f:

$$f' : x \mapsto f'(x) .$$

Osservazione: Sovente, il limite (5.1) è equivalentemente espresso come

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} . \tag{5.2}$$

Mostriamo ora come partendo dal limite (5.1) si possa riformulare il concetto di derivabilità; queste considerazioni giocheranno un ruolo fondamentale nel seguito. Se f è derivabile in x_0 abbiamo che

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) . \tag{5.3}$$

Equivalentemente,

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

o anche,

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} = 0 . \tag{5.4}$$

Chiamiamo,

$$\omega(x) = f(x) - f(x_0) - f'(x_0)(x - x_0) .$$

Allora abbiamo che

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \omega(x) \tag{5.5}$$

dove, da (5.4), sappiamo che

$$\omega(x) = o(x - x_0)\,.$$

La scrittura (5.4) è una condizione equivalente alla derivabilità nel senso precisato dalla seguente osservazione.

Osservazione: Supponiamo che

$$f(x) = f(x_0) + m(x - x_0) + \omega(x) \tag{5.6}$$

dove $\omega(x) = o(x - x_0)$. Allora si può facilmente dimostrare (farlo per esercizio) che f è derivabile in x_0 e si ha $f'(x_0) = m$.

Come la lipshitzianità, anche la derivabilità è una condizione più forte della continuità:

Proposizione 5.2 *Sia* $f\ :]a,b[\to \mathbb{R}$ *derivabile in* $x_0 \in]a,b[$. *Allora,* f *è continua in* x_0.

Dimostrazione Se f è derivabile in x_0 si ha che

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \omega(x) \tag{5.7}$$

con $\omega(x) = o(x - x_0)$. Quando $x \to x_0$ si ha che $f'(x_0)(x - x_0) \to 0$ e $\omega(x) \to 0$. Quindi, da (5.7), si ottiene

$$\lim_{x\to x_0} f(x) = f(x_0)\,.$$

■

Presentiamo ora alcuni primi esempi di funzioni derivabili e di calcoli delle relative derivate:

Esempio 77 *(Derivata di una retta).* Sia $f(x) = mx + n$. E' allora immediato verificare che f è derivabile in ogni punto e si ha $f'(x) = m$ per ogni $x \in \mathbb{R}$. In particolare, le costanti hanno derivata nulla in ogni punto.

Esempio 78 *(Derivata della funzione esponenziale).* Sia $f(x) = e^x$ e sia $x_0 \in \mathbb{R}$. Allora

$$\lim_{x\to x_0} \frac{e^x - e^{x_0}}{x - x_0} = \lim_{x\to x_0} e^{x_0}\frac{e^{x-x_0} - 1}{x - x_0} = e^{x_0}\,.$$

Quindi f è derivabile in ogni punto e si ha $f'(x) = e^x$.

Vale la pena di notare come la continuità da sola non implichi la derivabilità come mostrato dal seguente:

Esempio 79 Sia $f(x) = |x|$. Sappiamo che essa è continua in ogni punto di $\mathbb{R}$. Essa è chiaramente derivabile per ogni $x \neq 0$: in tali punti è infatti come una retta e si ha

$$f'(x) = \begin{cases} 1 & \forall x > 0 \\ -1 & \forall x < 0 \,. \end{cases}$$

Invece non è derivabile in 0. In effetti

$$\frac{f(x) - f(0)}{x - 0} = \frac{|x|}{x} = \begin{cases} 1 & \forall x > 0 \\ -1 & \forall x < 0 \,. \end{cases}$$

Si ha quindi che il rapporto incrementale ammette limiti destro e sinistro per $x \to 0$ eguali, rispettivamente, a $+1$ e a -1. Essendo tra loro diversi, ne segue che il limite non esiste e che quindi la funzione non è derivabile in 0. La mancanza di derivabilità si vede geometricamente in questo caso come la presenza di un angolo nel grafico.

L'esempio precedete suggerisce la definizione dei concetti di *derivata destra e sinistra* come limiti destro e sinistro, rispettivamente, del rapporto incrementale. Nel caso in cui esistano le derivate destra e sinistra in un punto, ma siano tra loro diverse, il punto in questione viene detto *punto angoloso.*

5.2 Regole di derivazione

L'operazione di derivazione soddisfa ad una serie di proprietà molto importanti esposte nei risultati di questa sezione.

Proposizione 5.3 *Siano $f :]a, b[\to \mathbb{R}$ e $g :]a, b[\to \mathbb{R}$ due funzioni entrambe derivabili in un punto $x_0 \in]a, b[$ e sia $c \in \mathbb{R}$. Si ha allora che anche le funzioni $f + g$, cf, fg, e (se $g(x_0) \neq 0$) f/g sono derivabili nel punto x_0 e si ha:*

$$\begin{aligned} (f+g)'(x_0) &= f'(x_0) + g'(x_0) \,, \\ (cf)'(x_0) &= cf'(x_0) \,, \\ (fg)'(x_0) &= f'(x_0)g(x_0) + f(x_0)g'(x_0) \,, \\ \left(\frac{f}{g}\right)'(x_0) &= \frac{f'(x_0)g(x_0) - f(x_0)g'(x_0)}{g(x_0)^2} \,. \end{aligned} \tag{5.8}$$

Dimostrazione Dimostriamo l'ultima, quella relativa al quoziente; lasciando le altre, più semplici, per esercizio:

$$
\begin{aligned}
\frac{\frac{f(x)}{g(x)}-\frac{f(x_0)}{g(x_0)}}{x-x_0} &= \frac{f(x)g(x_0)-f(x_0)g(x)}{x-x_0}\frac{1}{g(x)g(x_0)} \\
&= \frac{f(x)g(x_0)-f(x_0)g(x_0)+f(x_0)g(x_0)-f(x_0)g(x)}{x-x_0}\frac{1}{g(x)g(x_0)} \\
&= \frac{[f(x)-f(x_0)]g(x_0)-f(x_0)[g(x)-g(x_0)]}{x-x_0}\frac{1}{g(x)g(x_0)} \\
&= \left[\frac{f(x)-f(x_0)}{x-x_0}g(x_0)-f(x_0)\frac{g(x)-g(x_0)}{x-x_0}\right]\frac{1}{g(x)g(x_0)}\,.
\end{aligned}
$$

Passando ora al limite per $x \to x_0$ e sfruttando le regole dei limiti ed il fatto che g è una funzione continua in x_0, si ottiene la tesi. ■

La regola di derivazione del prodotto (terza formula di (5.8)) è generalmente nota come *regola di Leibnitz*.

Mostriamo ora l'applicazione delle proposizioni precedenti attraverso alcuni esempi.

Esempio 80 Sia $f(x) = x^n$. Essa è il prodotto di n volte la funzione $f(x) = x$. Quindi, poichè quest'ultima funzione è derivabile (vedi Esempio 77), segue da un'applicazione iterata della Proposizione 5.3 che anche la funzione di partenza x^n deve esserlo. Calcoliamone la derivata per i primi valori di n sfruttando la formula nella Proposizione 5.3:

$$
\begin{array}{lllllll}
n = 1 & f(x) & = & x & f'(x) & = & 1\,, \\
n = 2 & f(x) & = & x^2 = x\cdot x & f'(x) & = & 1\cdot x + x\cdot 1 = 2x\,, \\
n = 3 & f(x) & = & x^3 = x^2\cdot x & f'(x) & = & 2x\cdot x + x^2\cdot 1 = 3x^2\,, \\
n = 4 & f(x) & = & x^4 = x^3\cdot x & f'(x) & = & 3x^2\cdot x + x^3\cdot 1 = 4x^3\,.
\end{array}
$$

I calcoli precedenti sembrano suggerire la seguente formula per il calcolo della derivata della funzione x^n, qualunque sia $n \in \mathbb{N}$:

$$f(x) = x^n\,, \qquad f'(x) = nx^{n-1}\,.$$

Questa è in effetti la formula esatta ed una verifica formale (lasciata per esercizio) può essere fatta per induzione.

Esempio 81 *Funzioni polinomiali.* Sia

$$f(x) = a_0 + a_1x + a_2x^2 + \cdots a_nx^n\,.$$

Segue allora dalla Proposizione 5.3 e dal precedente esempio che la f è derivabile e la sua derivata è data da

$$f'(x) = a_1 + 2a_2 x + \cdots + n a_n x^{n-1}\,.$$

Esempio 82 *(Derivata delle funzioni trigonometriche).* Sia

$$f(x) = \sin x\,.$$

e sia $x_0 \in \mathbb{R}$. Usando le formule di prostaferesi possiamo scrivere:

$$\frac{\sin x - \sin x_0}{x - x_0} = \frac{2\sin\frac{x-x_0}{2}\cos\frac{x+x_0}{2}}{x - x_0} = \frac{\sin\frac{x-x_0}{2}}{\frac{x-x_0}{2}}\cos\frac{x+x_0}{2}\,.$$

Sappiamo che

$$\lim_{x\to x_0}\frac{\sin\frac{x-x_0}{2}}{\frac{x-x_0}{2}} = \lim_{t\to 0}\frac{\sin t}{t} = 1\,.$$

D'altra parte, in base alla continuità della funzione coseno si ha anche

$$\lim_{x\to x_0}\cos\frac{x+x_0}{2} = \cos x_0\,.$$

Segue dunque che

$$\lim_{x\to x_0}\frac{\sin x - \sin x_0}{x - x_0} = \cos x_0\,.$$

Abbiamo dunque dimostrato che la derivata della funzione $f(x) = \sin x$ è data da $f'(x) = \cos x$.

Per quanto riguarda la funzione $\cos x$, si possono ripetere considerazioni analoghe utilizzando ancora le formule di prostaferesi. Si dimostri per esercizio che anche la funzione $f(x) = \cos x$ è ovunque derivabile e vale $f'(x) = -\sin x$.

Consideriamo infine la tangente $f(x) = \tan x$. La f è definita per $x \in D = \mathbb{R} \setminus \{\pi/2 + k\pi \mid k \in \mathbb{Z}\}$. Essa può essere pensata come il rapporto tra il seno ed il coseno. Utilizzando la Proposizione 5.3 e le derivate delle funzioni seno e coseno precedentemente ricavate, si ottiene che essa è derivabile per ogni $x \in D$ e si ha

$$D\tan(x) = \frac{\cos x\cos x - \sin x(-\sin x)}{\cos^2 x} = \frac{1}{\cos^2 x}\,.$$

Altre importanti regole di derivazione sono quelle per le funzioni composte ed inverse.

Proposizione 5.4 (Derivata della funzione composta). *Siano*

$$f :]a, b[\to]c, d[, \qquad g :]c, d[\to \mathbb{R}$$

due funzioni e sia $x_0 \in]a, b[$. Supponiamo che f sia derivabile nel punto x_0 e che g sia derivabile nel punto $y_0 = f(x_0)$. Si ha allora che la funzione composta $g \circ f$ è derivabile nel punto x_0 e si ha:

$$(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)\,. \tag{5.9}$$

Dimostrazione Per ipotesi abbiamo che

$$\begin{aligned} f(x) &= f(x_0) + f'(x_0)(x - x_0) + o(x - x_0)\,, \\ g(y) &= g(y_0) + g'(y_0)(y - y_0) + o(y - y_0)\,. \end{aligned}$$

Ricordandoci che $y_0 = f(x_0)$, prendendo $y = f(x)$, sostituendo nella seconda espressione la prima e sfruttando le proprietà degli infinitesimi, si ottiene

$$\begin{aligned} g(f(x)) &= g(f(x_0)) + g'(f(x_0))[f'(x_0)(x - x_0) \\ &+ o(x - x_0)] + o(f'(x_0)(x - x_0) + o(x - x_0)) \\ &= g(f(x_0)) + g'(f(x_0))f'(x_0)(x - x_0) + o(x - x_0)\,. \end{aligned}$$

Per le osservazioni fatte alla definizione di derivata, questo significa proprio che $g \circ f$ è derivabile in x_0 e si ha che $(g \circ f)'(x_0) = g'(f(x_0))f'(x_0)$. ■

Esempio 83 Sia

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

La f è ottenuta componendo la funzione $x \mapsto -x^2$ con la funzione esponenziale. La sua derivata nel punto x è dunque, in virtù della Proposizione 5.4 e degli Esempi 78 e 81,

$$f'(x) = e^{-x^2}(-2x)\,.$$

Esempio 84 La funzione $f(x) = \cos x$ può essere rappresentata come

$$f(x) = \sin\left(x - \frac{\pi}{2}\right)\,.$$

Ne segue che la sua derivata può essere calcolata, in virtù sempre della Proposizione 5.4, come:

$$f'(x) = \cos\left(x - \frac{\pi}{2}\right) = -\sin x$$

(dove l'ultima eguaglianza segue ad esempio dalle formule di addizione). Abbiamo dunque ritrovato la derivata del coseno in altro modo rispetto a quanto visto nell'Esempio 82.

Proposizione 5.5 (Derivata della funzione inversa). *Sia $f :]a,b[\to]c,d[$ una funzione surgettiva strettamente monotona (e dunque invertibile per la Proposizione 4.2). Supponiamo inoltre che f sia derivabile nel punto $x_0 \in]a,b[$ e che sia $f'(x_0) \neq 0$. La funzione inversa $f^{-1} :]c,d[\to]a,b[$ è allora derivabile nel punto $y_0 = f(x_0)$ e si ha:*

$$(f^{-1})'(y_0) = \frac{1}{f'(x_0)} = \frac{1}{f'(f^{-1}(y_0))} \,. \tag{5.10}$$

Omettiamo una dimostrazione formale di questo risultato. La Figura 5.2 ne fornisce comunque una visione geometrica: le rette tangenti al grafico della f nel punto (x_0, y_0) e al grafico della f^{-1} nel punto (y_0, x_0) sono l'una la simmetrica dell'altra rispetto alla bisettrice. Questo spiega perchè le due derivate nella formula (5.10) siano l'una il reciproco dell'altra.

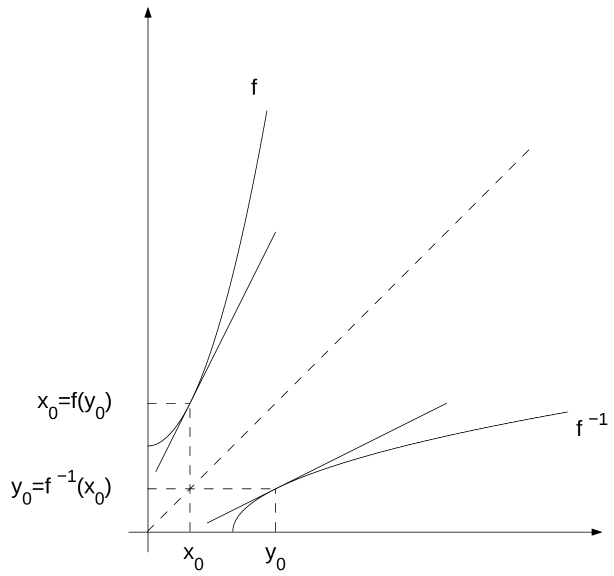


Figura 5.2:

Esempio 85 *(Derivata della funzione logaritmo naturale).* Consideriamo la funzione

$$f(x) = \ln x, \quad x \in (0, +\infty)\,.$$

Essendo l'inversa della funzione esponenziale, ed essendo la funzione esponenziale strettamente crescente sempre derivabile con derivata non nulla, utilizzando la Proposizione 5.5, si ha che f è derivabile in ogni punto e si ha

$$D\ln(x) = \frac{1}{e^{\ln x}} = \frac{1}{x}, \quad x > 0\,.$$

Esempio 86 *(Derivata delle funzioni trigonometriche inverse).* Consideriamo la funzione

$$\begin{aligned} &f : (-1,1) \to (-\pi/2, \pi/2) \\ &f(x) = \arcsin x\,. \end{aligned}$$

Poichè la funzione $\sin x$ è strettamente crescente , derivabile e con derivata non nulla in $(-\pi/2, \pi/2)$, utilizzando nuovamente la Proposizione 5.5 si ottiene che f risulta derivabile in ogni punto del suo dominio e si ha

$$D\arcsin(x) = \frac{1}{\cos(\arcsin x)} = \frac{1}{\sqrt{1-(\sin(\arcsin x))^2}} = \frac{1}{\sqrt{1-x^2}}\,.$$

Si noti che la seconda eguaglianza segue dall'identità trigonometrica $\sin^2 x + \cos^2 x = 1$ e dal fatto che, essendo $\arcsin x \in (-\pi/2, \pi/2)$ per definizione, sicuramente $\cos(\arcsin x) > 0$. In modo analogo si può far vedere (lasciamo i dettagli per esercizio) che

$$\begin{aligned} D\arccos(x) &= -\frac{1}{\sqrt{1-x^2}}\,, \\ D\arctan(x) &= \frac{1}{1+x^2}\,. \end{aligned}$$

Presentiamo qualche esempio che usa tutte le tecniche sinora introdotte.

Esempio 87 Sia $f(x) = a^x$ dove $a > 0$. Allora si può scrivere $f(x) = e^{x\ln a}$. Usando la regola di derivazione delle funzioni composte si ottiene quindi $f'(x) = e^{x\ln a}\ln a$. Dunque:

$$Da^x = a^x \ln a, \quad x \in \mathbb{R}\,.$$

Esempio 88 Sia $f(x) = x^\alpha$ dove $\alpha \in \mathbb{R}$ e $x > 0$. Allora si può scrivere $f(x) = e^{\alpha\ln x}$. Usando la regola di derivazione delle funzioni composte si ottiene quindi $f'(x) = e^{\alpha\ln x}\alpha\frac{1}{x}$. Dunque:

$$Dx^\alpha = \alpha x^{\alpha-1}, \quad x \in \mathbb{R}\,.$$

Esercizio 5.1 Calcolare la derivata delle seguenti funzioni:

$$\log_a x, \quad \arcsin(\sin x - x), \quad \frac{\ln\sin x}{\cos x}, \quad (\ln x)^{\ln x}, \quad (\arctan x)^{x^2+1}\,.$$

R: $\frac{1}{x\ln a}$, $\frac{\cos x - 1}{\sqrt{1-(\sin x - x)^2}}$, $\frac{(\cos x)^2 + (\sin x)^2 \ln\sin x}{\sin x(\cos x)^2}$, $\frac{(\ln x)^{\ln x}}{x}[\ln\ln x + 1]$, $(\arctan x)^{x^2+1}(2x\ln\arctan x + \frac{1}{\arctan x})$.

Esercizio 5.2 Sia $f : \mathbb{R} \to \mathbb{R}$ una funzione pari derivabile in ogni punto. Si dimostri allora che f' è dispari.

Esercizio 5.3 Sia $f : \mathbb{R} \to \mathbb{R}$ una funzione dispari derivabile in ogni punto. Si dimostri allora che f' è pari.

5.3 Il calcolo differenziale

La derivata di una funzione, come vedremo tra breve, contiene importanti informazioni sulla funzione stessa utilissime per farne uno studio qualitativo. Per il momento mostriamo come alcune proprietà delle funzioni si riflettono sulle loro derivate. Cominciamo con il seguente risultato intuitivo.

Proposizione 5.6 *Sia $f :]a, b[\to \mathbb{R}$ derivabile in ogni punto e crescente (risp., decrescente). Allora, $f'(x) \geq 0$ (risp. $f'(x) \leq 0$) per ogni $x \in]a, b[$.*

Dimostrazione Dimostriamolo nel caso crescente, l'altro si può fare in maniera analoga. Fissiamo $x_0 \in]a, b[$ e consideriamo il rapporto incrementale

$$R(x) = \frac{f(x) - f(x_0)}{x - x_0}.$$

Poichè f è crescente, è facile rendersi conto che $R(x) \geq 0$ per ogni $x \neq x_0$ sia destra che a sinistra di x_0. Quindi, il limite di $R(x)$ per $x \to x_0$ (che esiste per ipotesi), per il teorema di permanenza del segno non può essere certo un numero negativo. Quindi $f'(x_0) \geq 0$. ■

Oltre agli intervalli di monotonia ci sono altri elementi interessanti in una funzione: i punti di massimo e minimo locali. Sono punti in un intorno dei quali la funzione assume valori, a seconda, non superiori o non inferiori al valore che assume in detti punti. Ecco la definizione formale:

Definizione 5.7 Sia $f :]a, b[\to \mathbb{R}$ e sia $x_0 \in]a, b[$. x_0 è detto punto di *massimo locale* (risp. di *minimo locale*) se esiste $\delta > 0$ tale che $f(x) \leq f(x_0)$ (risp. $f(x) \geq f(x_0)$) per ogni $x \in]a, b[$ tale che $|x - x_0| \leq \delta$.

Proposizione 5.8 *Sia $f :]a, b[\to \mathbb{R}$ e sia $x_0 \in]a, b[$ un punto di massimo o di minimo locale per la f. Allora, se f è derivabile in x_0, si ha che $f'(x_0) = 0$.*

Dimostrazione Supponiamo che x_0 sia un punto di massimo locale per la f e consideriamo il rapporto incrementale

$$R(x) = \frac{f(x) - f(x_0)}{x - x_0}.$$

Poichè x_0 è punto di massimo locale non è difficile verificare che se x è abbastanza vicino ad x_0 si ha che $R(x) \geq 0$ se x sta a sinistra di x_0, mentre $R(x) \leq 0$ se x sta alla destra. Poichè esiste per ipotesi il limite di $R(x)$ per $x \to x_0$ ne segue che, ancora per il teorema di permanenza del segno, che tale limite non può essere nè un numero negativo, nè un numero positivo. Quindi può soltanto essere eguale a 0. ■

Quello che a noi piacerebbe sarebbe avere dei risultati che in qualche modo invertano le Proposizioni 5.6 e 5.8. Dal segno della derivata vorremmo ricavare informazioni sulla monotonia della funzione e dagli zeri l'eventuale presenza di punti di massimo e minimo locali. Se intuitivamente questo deve essere possibile, da un punto di vista teorico non è un'operazione così facile poichè mentre le proprietà di monotonia sono proprietà globali riguardanti il comportamento della funzione su tutto quanto un intervallo, la derivata è di per sè un concetto locale essendo espresso tramite un limite. Da informazioni di tipo locale (il segno della derivata punto per punto) vorremmo ottenere informazioni di tipo globale (la monotonia su di un intervallo). La chiave per ottenere questi risultati è un teorema fondamentale che sta alla base di tutto il calcolo differenziale, noto come il Teorema di Lagrange.

Teorema 5.9 (di Lagrange) *Sia $f[a,b] \to \mathbb{R}$ una funzione continua in ogni punto di $[a,b]$ e derivabile in ogni punto dell'intervallo $]a,b[$. Allora, esiste $\xi \in]a,b[$ tale che*

$$\frac{f(b)-f(a)}{b-a} = f'(\xi)\,. \tag{5.11}$$

Osservazione: L'interpretazione geometrica del risultato espresso dal teorema precedente è molto semplice. $\frac{f(b)-f(a)}{b-a}$ rappresenta il coefficiente angolare della retta secante il grafico della curva per i punti $(a, f(a))$ e $(b, f(b))$. Il teorema afferma che esiste un punto ξ interno ad $]a,b[$ dove la retta tangente al grafico ha lo stesso coefficiente angolare, è cioè parallela alla suddetta secante. Si veda la Figura 5.3

Si noti un caso particolare del Teorema di Lagrange, noto come Teorema di Rolle:

Teorema 5.10 *Sia $f[a,b] \to \mathbb{R}$ una funzione continua in ogni punto di $[a,b]$ e derivabile in ogni punto dell'intervallo $]a,b[$e tale che $f(a) = f(b)$. Allora, esiste $\xi \in]a,b[$ tale che*

$$f'(\xi) = 0\,.$$

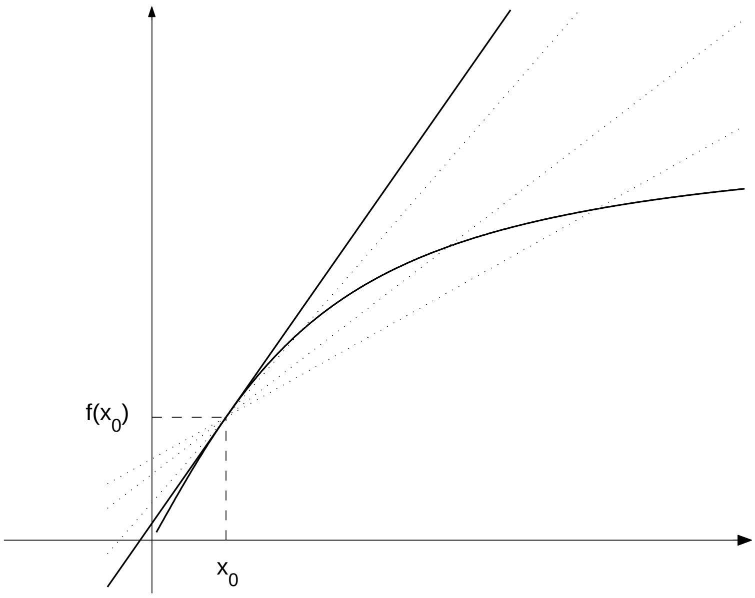


Figura 5.3:

Il Teorema di Rolle si ricava da quello di Lagrange semplicemente sostituendo nella formula (5.11) $f(b) = f(a)$. In realtà è anche vero il viceversa: da Rolle si può ricavare Lagrange come mostrano le considerazioni seguenti.

Dimostrazione (**Rolle** $\Rightarrow$ **Lagrange**). Sia $f : [a, b] \to \mathbb{R}$ una funzione che soddisfa le ipotesi del Teorema di Lagrange. Consideriamo la funzione ausiliaria:

$$g(x) = f(x) - \frac{f(b) - f(a)}{b - a}(x - a)\,.$$

g è ancora continua e derivabile in ogni punto di $]a, b[$. La sua derivata è data da

$$g'(x) = f'(x) - \frac{f(b) - f(a)}{b - a}\,. \tag{5.12}$$

Inoltre è facile vedere che $g(b) = g(a)$. Quindi g soddisfa le ipotesi del Teorema di Rolle e quindi esiste un punto $\xi \in]a, b[$ tale che $g'(\xi) = 0$. Segue allora dall'espressione (5.12), che

$$f'(\xi) = \frac{f(b) - f(a)}{b - a}\,.$$

■

Quello che rimane da dimostrare è quindi il Teorema di Rolle.

Dimostrazione (del Teorema di Rolle). Sia $f : [a,b] \to \mathbb{R}$ una funzione che soddisfa le ipotesi del Teorema di Rolle. Poichè essa è, in particolare, continua, segue dal Teorema di Weierstrass che essa ammette massimo e minimo assoluti, cioè esistono $x_1, x_2 \in [a,b]$ tali che

$$f(x_1) \leq f(x) \leq f(x_2), \quad \forall x \in [a,b]\,. \tag{5.13}$$

Se entrambi i punti x_1 e x_2 stanno sui bordi dell'intervallo $[a,b]$ sono cioè entrambi uguali ad a o b, ne segue che, poichè $f(a) = f(b)$ per ipotesi, sicuramente $f(x_1) = f(x_2)$ e segue allora dalla (5.13) che f è una funzione costante. Le funzioni costanti hanno derivata nulla in ogni punto e quindi, in questo caso, il teorema di Rolle è dimostrato. Se invece almeno uno dei due punti x_1 o x_2 sta in $]a,b[$, diciamo x_1 per fissare le idee (con x_2 si ragiona analogamente), si ha che $f'(x_1) = 0$ per la Proposizione 5.8. ■

E' ora il momento di mostrare importanti corollari del Teorema di Lagrange, che in particolare forniscono degli inversi alle Proposizioni 5.6 e 5.8.

Corollario 5.11 *Sia $f : [a,b] \to \mathbb{R}$ continua, derivabile su $]a,b[$ e tale che $f'(x) = 0$ per ogni $x \in]a,b[$. Allora, f è costante.*

Dimostrazione Per dimostrae che una funzione è costante basta far vedere che, presa una qualunque coppia di punti, in essi assume lo stesso valore. Siano dunque $c, d \in [a,b]$ tali che $c < d$ e consideriamo f ristretta all'intervallo $[c,d]$. Essa soddisfa le ipotesi del teorema di Lagrange e quindi esiste $\xi \in]c,d[$ tale che

$$\frac{f(d) - f(c)}{d - c} = f'(\xi)\,.$$

D'altra parte, per ipotesi, $f'(\xi) = 0$ e quindi si deve necessariamente avere $f(d) = f(c)$.■

Corollario 5.12 *Sia $f : [a,b] \to \mathbb{R}$ una funzione continua, derivabile su $]a,b[$. Allora,*

$$\begin{array}{llllll} f'(x) & \geq & 0, & \forall x \in]a,b[& \Rightarrow & f \text{ è crescente su } [a,b]\,, \\ f'(x) & > & 0, & \forall x \in]a,b[& \Rightarrow & f \text{ è str. crescente su } [a,b]\,, \\ f'(x) & \leq & 0, & \forall x \in]a,b[& \Rightarrow & f \text{ è decrescente su } [a,b]\,, \\ f'(x) & < & 0, & \forall x \in]a,b[& \Rightarrow & f \text{ è str. decrescente su } [a,b]\,. \end{array}$$

Dimostrazione Dimostriamone uno, gli altri si vedono analogamente. Supponiamo che $f'(x) \leq 0$ per ogni $x \in]a,b[$. Siano $c, d \in [a,b]$ tali che $c < d$ e consideriamo f ristretta all'intervallo $[c,d]$. Essa soddisfa le ipotesi del teorema di Lagrange e quindi esiste $\xi \in]c,d[$ tale che

$$\frac{f(d) - f(c)}{d - c} = f'(\xi)\,.$$

D'altra parte, per ipotesi, $f'(\xi) \leq 0$ e quindi si deve necessariamente avere $f(d) \leq f(c)$. ∎

Osservazione: Si noti come il Corollario 5.12 contenga informazioni più precise rispetto alla Proposizione 5.6 dove non si hanno risultati particolari nel caso di monotonia stretta. In effetti, non si possono proprio avere: può accadere che una funzione sia, ad esempio, strettamente crescente, ma non avere derivata strettamente positiva. Si pensi , ad esempio, alla funzione $f(x) = x^3$ definita e derivabile su tutto $\mathbb{R}$. Essa è strettamente crescente

$$x_1 < x_2 \ \Rightarrow \ x_1^3 < x_2^3 \, .$$

Tuttavia la sua derivata è data da $f'(x) = 3x^2$. Abbiamo così che $f'(x) \geq 0$ per ogni $x \in \mathbb{R}$ come previsto dalla Proposizione 5.6. Tuttavia $f'(0) = 0$.

Corollario 5.13 *Sia $f\ :]a,b[\to \mathbb{R}$ una funzione derivabile e sia $x_0 \in]a,b[$ tale che $f'(x_0) = 0$. Se esiste $\delta > 0$ tale che*

$$\begin{array}{ll} f'(x) \leq 0\,, & \forall x < x_0\,,\ \text{t.c. } |x - x_0| \leq \delta\,, \\ f'(x) \geq 0\,, & \forall x > x_0\,,\ \text{t.c. } |x - x_0| \leq \delta\,, \end{array} \tag{5.14}$$

allora x_0 è un punto di minimo locale per f. Invece, se esiste $\delta > 0$ tale che

$$\begin{array}{ll} f'(x) \geq 0\,, & \forall x < x_0\,,\ \text{t.c. } |x - x_0| \leq \delta\,, \\ f'(x) \leq 0\,, & \forall x > x_0\,,\ \text{t.c. } |x - x_0| \leq \delta\,, \end{array} \tag{5.15}$$

allora x_0 è un punto di massimo locale per f.

Dimostrazione Consideriamo il primo caso (5.14). Sia $x < x_0$ e tale che $|x - x_0| < \delta$. Applicando il Teorema di Lagrange ad f sull'intervallo $[x, x_0]$, si ha che esiste $c \in (x, x_0)$ tale che

$$\frac{f(x_0) - f(x)}{x_0 - x} = f'(c) \leq 0$$

che implica $f(x) \geq f(x_0)$. Se invece partiamo da un $x > x_0$ sempre tale che $|x - x_0| < \delta$ e applichiamo il Teorema di Lagrange stavolta sull'intervallo $[x_0, x]$, si ha che

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(c) \geq 0$$

per un qualche $c \in (x_0, x)$. Quindi, anche in questo caso, si ha $f(x) \geq f(x_0)$. ∎

Il Teorema di Lagrange permette anche di ottenere stime della variazione di una funzione $f(x)$ in termini della variazione corrispondente della variabile x:

Corollario 5.14 *Sia* $f :]a,b[\to \mathbb{R}$ *derivabile in ogni punto e tale che esista* $M > 0$ *per cui* $|f'(x)| \leq M$ *per ogni* $x \in]a,b[$. *Allora*

$$|f(x_1) - f(x_2)| \leq M|x_1 - x_2|, \quad \forall x_1, x_2 \in]a,b[$$

(una funzione che soddisfa alla stima sopra è detta Lipschitziana di costante M*).*

Dimostrazione Siano $x_1, x_2 \in]a,b[$. Allora, per il Teorema di Lagrange

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2} = f'(c)$$

per un opportuno punto c tra x_1 e x_2. Equivalentemente,

$$f(x_1) - f(x_2) = f'(c)(x_1 - x_2).$$

Si può quindi stimare

$$|f(x_1) - f(x_2)| = |f'(c)||x_1 - x_2| \leq M|x_1 - x_2|.$$

■

La limitatezza della derivata di f risulta automatica se cambiamo leggermente le altre ipotesi. Supponiamo che $f : [a,b] \to \mathbb{R}$ sia derivabile anche in a e b dove consideriamo la derivata sinistra e destra, rispettivamente. Supponiamo inoltre che la funzione derivata $f' : [a,b] \to \mathbb{R}$ sia continua. Allora essa è automaticamente limitata per il Teorema di Weierstrass (vedi Capitolo 4). L'insieme delle funzioni dotate di derivata prima continua su $[a,b]$ è indicato con il simbolo $C^1([a,b])$. Dunque, per le considerazioni appena fatte, se $f \in C^1([a,b])$, allora f è lipschitziana su $[a,b]$.

Esercizio 5.4 Dimostrare che la funzione $f(x) = x^2$ è lipschitziana su ogni intervallo chiuso e limitato $[a,b]$. Dimostrare poi (*) che $f(x)$ non è lipschitziana su tutto $\mathbb{R}$.

Esercizio 5.5 Dimostrare che la funzione $f(x) = \arctan x$ è lipschitziana su $\mathbb{R}$.

5.3.1 Studio di funzioni

Una delle applicazioni più importanti del calcolo differenziale è lo studio qualitativo di una funzione. L'espressione studiare una funzione deve essere intesa come analisi delle seguenti cose: campo naturale di esistenza, segno e zeri, limiti ai bordi del dominio, intervalli di monotonia, massimi e minimi locali, asintoti, convessità e flessi (quest'ultimi concetti saranno introdotti prossimamente) . Non necessariamente sempre si riesce a fare uno studio completo di tutti gli elementi sopra esposti ed è buona abitudine cercare di tracciare il grafico anche con una conoscenza parziale di tali elementi.

Presentiamo ora un esempio dettagliato di come si svolge uno studio di funzione.

Esempio 89 Si consideri la funzione definita dall'equazione

$$f(x) = \left|\ln\left|e^{2x} - e^x\right|\right| .$$

Studiare la funzione assegnata significa percorrere i passi che verranno delineati qui di seguito, fino a essere in condizione di disegnarne un grafico qualitativo il più possibile preciso. Suggeriamo di seguire il percorso delineato in questo esempio e di indicare per quanto possibile *di passo in passo* su un grafico cartesiano le informazioni ottenute.

La prima cosa da fare è determinare *l'insieme di definizione* della funzione. Questo significa determinare qual'è l'insieme di numeri reali per i quali ha senso considerare l'espressione scritta. In questo caso occorre (e basta) controllare quando la quantità di cui si deve calcolare il logaritmo è *strettamente positiva.* Poiché l'argomento del logaritmo è un valore assoluto, vi sono problemi soltanto per quei valori della x per i quali si ha

$$e^{2x} - e^x = 0 .$$

Semplificando si ottiene $e^x = 1$, cioè $x = 0$. Dunque il *campo di esistenza* D della funzione data è

$$D = \mathbb{R} \setminus \{0\} .$$

Prima di procedere con i passi successivi, osserviamo che la funzione f è definita come valore assoluto di una certa quantità, più precisamente della funzione

$$g(x) = \ln\left|e^{2x} - e^x\right| .$$

La via più breve per studiare la f è quindi quella di studiare la g, ottenendo poi il grafico di f da quello di g semplicemente prendendo i tratti di quest'ultimo con ordinata negativa, e “ribaltandoli rispetto all'asse delle ascisse.

Studiamo quindi la g, che ha naturalmente lo stesso insieme di definizione della f. La seconda cosa da fare è studiare i *limiti alla frontiera* dell'insieme di definizione, in questo caso per $x \to 0^{\pm}$ e per $x \to \pm\infty$. Vediamo: intanto

$$\lim_{x\to 0^{\pm}} g(x) = -\infty$$

perché, per costruzione, l'argomento del logaritmo tende a zero quando $x \to 0$ sia da sinistra che da destra. Inoltre

$$\lim_{x\to+\infty} g(x) = +\infty$$

perché l'argomento del logaritmo tende, in tale limite, a $+\infty$, dato che e^{2x} è un'infinito di ordine superiore a e^x. Infine

$$\lim_{x\to-\infty} g(x) = -\infty$$

perché l'argomento del logaritmo tende a zero in tale limite.

Prima di proseguire osserviamo subito che

$$e^{2x} - e^x = e^x(e^x - 1) > 0$$

se e solo se $x > 0$. Quindi si avrà

$$g(x) = \begin{cases} \ln(e^{2x} - e^x) & \text{se } x > 0\,, \\ \ln(e^x - e^{2x}) & \text{se } x < 0\,. \end{cases}$$

Le ultime osservazioni sui limiti all'infinito rendono *possibile* l'esistenza di asintoti obliqui sia per $x \to +\infty$ che per $x \to -\infty$. Dire che una retta, diciamo di equazione

$$y = mx + q$$

è asintoto obliquo per g per $x \to +\infty$ significa per definizione che

$$g(x) - (mx + q) \to 0 \text{ se } x \to +\infty$$

cioè che la distanza tra la curva g e la retta tende a zero a $+\infty$. Notiamo allora che, se $x > 0$

$$\begin{aligned} g(x) &= \ln(e^{2x} - e^x) \\ &= \ln(e^{2x}) + \ln(1 - e^{-x}) \\ &= 2x + o(1) \text{ se } x \to +\infty \end{aligned}$$

ovvero

$$g(x) - 2x \to 0 \text{ se } x \to +\infty\,.$$

Analogamente, se $x < 0$:

$$\begin{aligned} g(x) &= \ln(e^x - e^{2x}) \\ &= \ln(e^x) + \ln(1 - e^x) \\ &= x + o(1) \text{ se } x \to -\infty \end{aligned}$$

ovvero

$$g(x) - x \to 0 \text{ se } x \to -\infty\,.$$

In entrambi i casi la struttura della funzione ci ha permesso facilmente di identificare la *parte principale* della funzione data, che è risultata essere lineare in x. Ne possiamo concludere che la retta $y = 2x$ è asintoto obliquo per g per $x \to +\infty$, mentre la retta $y = x$ è asintoto obliquo per g per $x \to -\infty$.

Osserviamo qui che per determinare se esiste o meno un asintoto obliquo (per esempio per $x \to +\infty$) si può anche procedere così : per prima cosa si calcola

$$\lim_{x \to +\infty} \frac{g(x)}{x} .$$

Se tale limite esiste *finito*, diciamo uguale a $m \in \mathbb{R}$, allora *è possibile* che *una* retta del tipo $y = mx + q$ sia asintoto obliquo per g a $+\infty$. Per stabilire se l'asintoto esiste effettivamente e per calcolare in tal caso q si verifica se

$$\lim_{x \to +\infty} [g(x) - mx]$$

esiste *finito*. Se ciò accade e il limite vale, diciamo, $q \in \mathbb{R}$, allora la retta $y = mx + q$ è effettivamente asintoto obliquo per g a $+\infty$. È immediato verificare che questo procedimento porta, per la funzione che stiamo studiando, alla medesima conclusione ottenuta prima.

Studiamo ora il *segno* della funzione. Cominciamo con l'osservare che $\ln a$ è positivo se e solo se $a > 1$. Allora, se $x > 0$, g sarà positiva se e solo se $e^{2x} - e^x > 1$. Posto $z = e^x$ ciò equivale a $z^2 - z - 1 > 0$ che è verificata, ricordando che z per costruzione deve essere positivo, se e solo se $z > (1 + \sqrt{5})/2$.. Quindi g è negativa per $x \in]0, \ln((1 + \sqrt{5})/2)[$ e positiva per $x > \ln((1 + \sqrt{5})/2)$. L'unico zero di g per $x > 0$ è il punto $x = \ln((1 + \sqrt{5})/2)$. Se invece x è negativo allora g è positivo se e solo se $e^x - e^{2x} > 1$. Procedendo come prima si vede subito che questa disequazione non è mai soddisfatta, e quindi g è negativa per ogni $x < 0$.

Studiamo ora la *derivata prima* della funzione e, con essa, gli eventuali massimi e minimi di g. Calcoliamo allora, dapprima per $x > 0$:

$$g'(x) = \frac{2e^{2x} - e^x}{e^{2x} - e^x} = \frac{2e^x - 1}{e^x - 1} .$$

È immediato verificare che $g'(x) = 0$ se e solo se $x = -\ln 2$, che però non appartiene all'insieme degli x per i quali abbiamo calcolato la derivata. Quindi non vi sono punti stazionari per g nell'intervallo $]0, +\infty[$ e un rapido calcolo mostra anche che $g'(x) > 0 \ \forall x > 0$. Quindi g è ivi monotona crescente. Osserviamo però subito (fate esplicitamente il calcolo!) che l'espressione prima scritta per g' è valida *anche per* $x < 0$. Quindi il punto $x = -\ln 2$ è stazionario per g. Studiamo ora il segno della derivata prima per $x < 0$. Il denominatore è ovviamente negativo se $x < 0$, mentre il numeratore è positivo se $x > -\ln 2$ e negativo altrimenti. In conclusione

$g'(x) > 0$ se $x < -\ln 2$ e $g'(x) < 0$ per $x \in]-\ln 2, 0[$. Quindi il punto $x = -\ln 2$ è un punto di massimo relativo. Non ve ne sono altri, e peraltro il punto trovato non è un estremo assoluto: si ha in effetti, per le considerazioni precedenti sui limiti della funzione agli estremi dell'insieme di definizione, che $\inf g = -\infty$, e che $\sup g = +\infty$.

Sebbene ancora non si sia introdotto l'argomento, concludiamo per completezza lo studio di funzione con lo studio della derivata seconda. Essa vale, per ogni $x \neq 0$:

$$g''(x) = -\frac{e^x}{(e^x - 1)^2}\,.$$

Essa è chiaramente sempre negativa: quindi la concavità di g è sempre rivolta verso il basso. Per le precedenti considerazioni sulle relazioni tra il grafico di f e quello di g possiamo concludere che il grafico di f è quello mostrato nella Figura 5.4.

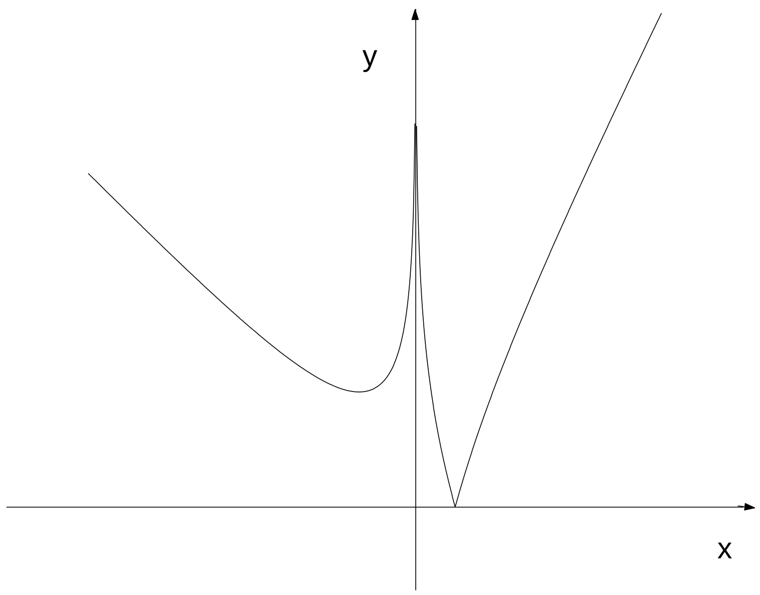


Figura 5.4:

In particolare il punto $x_0 = (1+\sqrt{5})/2$ è un *punto angoloso* per la funzione f, poiché f non è ivi derivabile ma le derivate destra e sinistra di f esistono finite, con

$$f'_+((1+\sqrt{5})/2) \neq f'_-((1+\sqrt{5})/2)\,.$$

Abbiamo anche che tale punto è un punto di minimo *assoluto* per f (si noti che f non è derivabile in x_0 e che quindi non si sarebbe potuto stabilire questo fatto annullando la derivata prima di f...), cosicché $\inf f = \min f = 0$. Non vi sono massimi relativi, e $\sup f = +\infty$. La retta $y = 2x$ è asintoto obliquo per f a $+\infty$, mentre la retta $y = -x$ (si noti il cambio di segno) è asintoto obliquo per f a $-\infty$. La retta $x = 0$ è asintoto verticale per f, e si ha $\lim_{x\to 0^\pm} = +\infty$.

Esercizio 5.6 Studiare le seguenti funzioni e tracciarne un grafico qualitativo:

$$\sinh x, \qquad \frac{1+\ln x}{x}, \qquad \frac{\sin x}{x}, \qquad \frac{|e^x - 1|}{1+|x|}, \qquad \sqrt[5]{x(x^2-1)^2}\,.$$

5.3.2 Problemi di massimo e minimo

In molte applicazioni si è interessati al calcolare il valore massimo e minimo assoluti (quando essi esistono) di una funzione. Naturalmente questo tipo di informazioni si dovrebbero poter dedurre da uno studio qualitativo della funzione stessa. Tuttavia si tratta questo di un problema più specifico che spesso può essere risolto senza dover arrivare a disegnare la funzione stessa.

In questo contesto ci limiteremo a fare alcune osservazioni nel caso in cui si abbia una funzione continua $f : [a,b] \to \mathbb{R}$ definita su di un intervallo chiuso e limitato. In questo caso il massimo ed il minimo assoluti sicuramente esistono in virtù del Teorema di Weierstrass. Come si fa a determinarli? Essi possono essere assunti alle estremità a e b dell'intervallo di definizione oppure all'interno in $]a,b[$. Se la funzione è derivabile in $]a,b[$, gli eventuali punti interni dove la funzione assume valore massimo o minimo assoluti, saranno certamente punti dove la derivata si annulla. Se dunque noi siamo in grado di risolvere l'equazione

$$f'(x) = 0\,, \quad x \in]a,b[$$

trovando un certo numero di soluzioni $x_1, x_2, \ldots, x_n \in]a,b[$ (questo non è il caso più generale, l'equazione sopra potrebbe anche avere infinite soluzioni, però sicuramente comprende molti esempi concreti), allora sicuramente il massimo ed il minimo di f dovranno essere, per il ragionamento appena fatto, tra i valori:

$$f(x_1), f(x_2), \ldots, f(x_n), f(a), f(b)\,.$$

Dunque,

$$\begin{aligned}
\min\{f(x) \mid x \in [a,b]\} &= \min\{f(x_1), f(x_2), \ldots, f(x_n), f(a), f(b)\} \\
\max\{f(x) \mid x \in [a,b]\} &= \max\{f(x_1), f(x_2), \ldots, f(x_n), f(a), f(b)\}\,.
\end{aligned}$$

Il problema si riduce quindi a calcolare il massimo ed il minimo di un insieme costituito da un insieme finito di punti. Si noti che per risolvere il nostro problema non serve neppure andare ad indagare se i singoli punti x_k siano di massimo locale, di minimo locale o di flesso.

Esempio 90 Consideriamo la funzione $f : [0,5] \to \mathbb{R}$ data da

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

Essa è definita su di un intervallo chiuso e limitato ed è ivi continua. Dunque ammette massimo e minimo assoluti. Si ha che

$$f'(x) = \frac{x^2+2x-1}{(x+1)^2}\,.$$

L'equazione $x^2+2x-1=0$ ha un'unica soluzione in $]0,5[$ data da $x_1 = \sqrt{2}-1$. Si osservi ora che $f(0)=1$ mentre $f(5)=26/6$. D'altra parte, $f(x_1) = 2\sqrt{2}-2 < 1$ Quindi,

$$\begin{aligned} \min\{f(x) \mid x \in [0,5]\} &= \min\{2\sqrt{2}-2, 1, 26/6\} = 2\sqrt{2}-2\,, \\ \max\{f(x) \mid x \in [0,5]\} &= \max\{2\sqrt{2}-2, 1, 26/6\} = 26/6\,. \end{aligned}$$

Esercizio 5.7 Calcolare massimo e minimo assoluto delle funzioni seguenti nel dominio specificato:

$$f(x) = 2\sin x - x\,,\ x \in [0,\pi]\,; \qquad g(x) = xe^{\frac{1}{x}}\,,\ x \in [1/2, 5]\,.$$

R: $\max f = \sqrt{3}-\pi/3$, $\min f = -\pi$; $\max g = 5e^{1/5}$, $\min g = e$.

5.4 Formula di Taylor e serie di Taylor

Abbiamo visto nei capitoli precedenti vari esempi di funzioni rappresentabili attraverso polinomi opportuni a meno di termini di ordine superiore. Ad esempio abbiamo che

$$e^x = 1 + o(1)\,,$$

o meglio ancora

$$e^x = 1 + x + o(x)\,.$$

Ci chiediamo se è possibile fare di meglio, ad esempio trovare un'approssimazione di ordine 2, cioè qualcosa del tipo

$$e^x = a_0 + a_1 x + a_2 x^2 + o(x^2)\,,$$

per opportuni coefficienti a_0, a_1 e a_2.

Analizziamo le cose da un punto di vista più generale. Se la funzione $f(x)$ è derivabile su un intervallo $]a, b[$ e $x_0 \in]a, b[$, sappiamo dalla definizione stessa di derivata che vale la relazione:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0)\,.$$

Come ben sappiamo $f(x_0) + f'(x_0)(x - x_0)$ è la funzione lineare che corrisponde alla retta tangente al grafico della $f(x)$ nel punto $(x_0, f(x_0))$. Questa è l'unica funzione lineare che approssima la $f(x)$ a meno di termini trascurabili rispetto a $(x - x_0)$ (si rifletta sul perchè). Vorremmo ottenere una formula analoga al secondo ordine, cioè trovare a_0, a_1 , $a_2 \in \mathbb{R}$ tali che

$$f(x) = a_0 + a_1(x - x_0) + a_2(x - x_0)^2 + o(x - x_0)^2$$

Come fare per determinare questi speciali coefficienti? Si noti intanto che la formula sopra implica che

$$f(x) = a_0 + a_1(x - x_0) + o(x - x_0)$$

dal quale segue che necessariamente $a_0 = f(x_0)$ e $a_1 = f'(x_0)$. Rimane dunque da calcolare soltanto a_2: esso deve essere collegato al modo in cui il grafico di f sta 'curvando' vicino ad x_0, cioè come sta cambiando la pendenza misurata da $f'(x)$ vicino a x_0. Per leggere le variazioni di $f'(x)$ entra naturalmente in gioco la derivata della funzione derivata $f'(x)$ nel punto x_0 (se essa esiste), che prende il nome di derivata seconda della funzione $f(x)$ e si indica con il simbolo $f''(x_0)$. Quale sarà la relazione tra a_2 e $f''(x_0)$? Facciamo qualche considerazione preliminare supponendo che questa derivata seconda esista effettivamente. Calcoliamo le derivate del polinomio di secondo grado

$$g(x) = f(x_0) + f'(x_0)(x - x_0) + a_2(x - x_0)^2\,.$$

Abbiamo,

$$\begin{aligned} g'(x) &= f'(x_0) + 2a_2(x - x_0) \\ g''(x) &= 2a_2 \end{aligned}$$

Dunque,

$$g(x_0) = f(x_0)\,, \quad g'(x_0) = f'(x_0)\,, \quad g''(x_0) = 2a_2\,.$$

E' abbastanza intuitivo che il polinomio di secondo grado che meglio approssimi la funzione $f(x)$ vicino a x_0 sia quello che, oltre ad avere lo stesso valore in x_0 e ad avere la stessa derivata della funzione $f(x)$, abbia coincidente anche la derivata seconda. Questo si ottiene chiaramente scegliendo

$$a_2 = \frac{f''(x_0)}{2}\,.$$

Si tratta dunque di dimostrare a questo punto che vale la formula

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 + o(x - x_0)^2\,.$$

Dimostreremo questa eguaglianza attraverso un risultato intermedio, di grande interesse applicativo, che può pensarsi come una generalizzazione del Teorema di Lagrange.

Teorema 5.15 *Sia $f : [a, b] \to \mathbb{R}$ una funzione derivabile due volte su tutto $[a, b]$ con $f''(x)$ funzione continua. La continuità e la derivabilità sono intese da destra (risp. da sinistra) nell'estremo a (risp. nell'estremo b). Esiste allora un punto ξ compreso tra a e b tale che valga la formula*

$$f(b) = f(a) + f'(a)(b - a) + \frac{f''(\xi)}{2}(b - a)^2\,. \tag{5.16}$$

Dimostrazione Definiamo il numero K tramite la relazione

$$K = \frac{f(b) - f(a) - f'(a)(b - a)}{(b - a)^2}$$

e consideriamo la funzione ausiliaria

$$g(x) = f(x) - [f(a) + f'(a)(x - a) + K(x - a)^2]\,.$$

Si noti che

$$g(a) = g(b) = 0\,.$$

Per il Teorema di Rolle esiste un punto $\eta \in]a, b[$ tale che $g'(\eta)$=0. D'altra parte, poichè

$$g'(x) = f'(x) - f'(a) - 2K(x - a)$$

per ogni x, segue anche che $g'(a) = 0$. Applicando di nuovo il Teorema di Rolle questa volta a g' sull'intervallo $[a, \nu]$, si ottiene che esiste $\xi \in [a, \nu] \subset [a, b]$ tale che

$$g''(\xi) = 0\,.$$

Poiché
$$g''(x) = f''(x) - 2K\,,$$
si ha quindi che
$$f''(\xi) = 2K\,.$$
Dalla definizione di K segue ora subito la formula (5.16). ■

Possiamo a questo punto ottenere il risultato che volevamo

Corollario 5.16 *Sia $f : [a,b] \to \mathbb{R}$ una funzione derivabile due volte su tutto $[a,b]$ con $f''(x)$ funzione continua. La continuità e la derivabilità sono intese da destra (risp. da sinistra) nell'estremo a (risp. nell'estremo b). Sia $x_0 \in]a,b[$. Vale allora la formula*
$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2} f''(x_0)(x - x_0)^2 + o(x - x_0)^2\,. \qquad (5.17)$$

Dimostrazione È sufficiente considerare il caso $x > x_0$, e usare la formula di Taylor con resto di Lagrange nell'intervallo $[x_0, x]$ per ottenere che
$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(\xi)}{2}(x - x_0)^2\,, \qquad (5.18)$$
dove $\xi \in]x_0, x[$.

La continuità di f'' implica poi che, per $x \to x_0$:
$$f''(\xi) = f''(x_0) + o(1)$$
in quanto ξ t necessariamente tende a x_0 quando $x \to x_0$. Allora,
$$\begin{aligned} f(x) &= f(x) = f(x_0) + f'(x_0)(x - x_0) + \tfrac{1}{2}[f''(x_0) + o(1)](x - x_0)^2 \\ &= f(x) = f(x_0) + f'(x_0)(x - x_0) + \tfrac{1}{2} f''(x_0)(x - x_0)^2 + o(x - x_0)^2\,. \end{aligned}$$
■

Tutto questo può essere generalizzato ad ordini superiori al secondo. Per fare questo dobbiamo innanzitutto introdurre le derivate di ordine superiore al secondo. In effetti come abbiamo definito la derivata seconda, possiamo a sua volta definire la derivata terza, quarta e così via, se esse esistono. Si indicano $f'''(x)$, $f^{(4)}(x)$, $f^{(5)}(x)$,e più generalmente per ordine k, $f^{(k)}(x)$ o anche $D_k(f)(x)$. Diamo ora una definizione

Definizione 5.17 Una funzione $f : [a,b] \to \mathbb{R}$ si dice di classe C^k su $[a,b]$, $k \in \mathbb{N}$, se essa è continua e derivabile con derivate continue fino all'ordine k su $[a,b]$. La continuità e la derivabilità sono intese da destra (risp. da sinistra) nell'estremo a (risp. nell'estremo b). f si dice di classe C^∞ su $[a,b]$ se è di classe C^k su $[a,b]$ per ogni $k \in \mathbb{N}$.

Un'analoga definizione si dà per funzioni di classe C^k sull'intervallo *aperto* $]a,b[$.

Il seguente risultato è una generalizzazione del Teorema 5.15 di dimostrazione concettualmente analoga, anche se tecnicamente un po' più complicata, e che quindi verrà omessa.

Teorema 5.18 *Sia $f:[a,b]\to\mathbb{R}$ una funzione di classe $C^n([a,b])$. Esiste allora un punto ξ compreso tra a e b tale che valga la formula*

$$f(b)=\sum_{k=0}^{n-1}\frac{f^{(k)}(a)}{k!}(b-a)^k+\frac{f^{(n)}(\xi)}{n!}(b-a)^n\,. \tag{5.19}$$

La (5.19) prende il nome di *formula di Taylor, con resto di Lagrange, all'ordine n.* Il prossimo risultato invece generalizza il Corollario 5.16 e prende il nome di *formula di Taylor, con resto di Peano, all'ordine n.* La dimostrazione è del tutto simile a quella del Corollario 5.16 ed è quindi omessa.

Corollario 5.19 *Sia $f:[a,b]\to\mathbb{R}$ una funzione di classe $C^n([a,b])$, e $x_0\in\,]a,b[$. Vale allora la formula*

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+o((x-x_0)^n)\,. \tag{5.20}$$

Osservazione. Quale che sia la formula data per il resto, la formula di Taylor ha lo scopo di trovare, fissato n, un polinomio di grado n che *approssimi*, in un senso da precisarsi, la funzione data vicino a x_0. Questo è un'idea fondamentale in matematica: approssimare funzioni complicate con funzioni molto più semplici, in questo caso i polinomi. Avrete certamente notato che i polinomi approssimanti che intervengono sia nella formula di Taylor con il resto di Lagrange che in quella con il resto di Peano, sono sempre gli stessi:

$$P_n(x)=\sum_{k=0}^{n}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k\,. \tag{5.21}$$

Tale polinomio è detto il *polinomio di Taylor di grado n*, della funzione f nel punto x_0. Che cosa ha di speciale questo polinomio? Si vede subito che

$$P_n(x_0)=f(x_0)\,,$$

come certamente deve essere se in un qualunque senso P_n approssima f vicino a x_0. Calcoli semplici mostrano però anche che

$$P_n^{(k)}(x_0) = f^{(k)}(x_0) \quad \forall k = 0, 1, \ldots, n\,.$$

Quindi P_n è quel polinomio di grado n i cui coefficienti sono scelti in modo che le derivate fino alla n–esima di P_n in x_0 coincidano con le corrispondenti derivate di f in x_0 (si vede facilmente che questo determina *univocamente* i coefficienti del polinomio).

Ciò è ragionevole: se vogliamo approssimare una funzione f con una retta vicino al punto x_0, la retta "migliore è intuitivamente la tangente al grafico in x_0, e tale retta è proprio quell'unica retta il cui valore in x_0 coincide con quello di f, e che ha coefficiente angolare pari alla derivata di f in x_0. Il polinomio di Taylor di grado (n) ha proprietà simili per le derivate successive, e compare naturalmente quando si voglia approssimare f con un *polinomio* di grado maggiore di uno invece che con una retta. Si osservi infine che $P_n(x)$ è l'unico polinomio di grado al più n che approssima $f()$ a meno di termini trascurabili rispetto a $(x-x_0)^n$.In effetti, se $q(x)$ fosse un altro polinomio di grado al più n tale che

$$f(x) = q(x) + o(x - x_0)^n\,,$$

confrontando con la (5.20) otterremo che

$$q(x) - P_n(x) = o(x - x_0)^n$$

Poichè $q(x) - P_n(x)$ si può sempre pensare come un polinomio di grado al più n nella variabile $(x-x_0)$, se ne conclude che l'unico modo perchè l'eguaglianza sopra possa valere (si rifletta sul perchè) è che $q(x) - P_n(x) = 0$ identicamente.

Se una funzione $f(x)$ è di classe C^∞ su un intervallo $]a, b[$ e $x_0 \in]a, b[$ possiamo considerare i polinomi di Taylor $P_n(x)$ relativi ad f nel punto x_0, qualunque sia n. La successione dei polinomi di Taylor (5.21) è di fatto una serie, detta *serie di Taylor*, che dipende dalla variabile x. Ci aspettiamo che questa serie approssimi sempre meglio la nostra funzione $f(x)$ all'aumentare di n: ciè che sappiamo di quanto è buona questa approssimazione ci viene dal Teorema 5.18. E' lecito chiedersi se tale serie sia sommabile, e tale somma valga proprio $f(x)$, cioè se valga

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

per lo meno per valori di x sufficientemente vicini a x_0 (si noti che ovviamente per $x = x_0$ l'eguaglianza vale). Per molte funzioni importanti questo è vero.

Facciamo ora qualche esempio calcolando lo sviluppo di Taylor di alcune funzioni notevoli.

Esempio 91 Consideriamo la funzione esponenziale

$$f(x) = e^x \, ,$$

che è di classe C^∞ su tutto $\mathbb{R}$ e calcoliamo il suo sviluppo di Taylor in 0. Poichè $f^{(k)}(x) = e^x$ per ogni k, si ha che $f^{(k)}(0) = 1$ per ogni k. I polinomi di Taylor sono dunque dati da

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

Si tratta proprio della serie esponenziale che avevamo visto a suo tempo e che avevamo dimostrato convergere per valori di x non negativi. La analizziamo ora da un altro punto di vista. Scriviamo la formula di Taylor col resto di Lagrange, di grado $n+1$:

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + e^{\xi} \frac{x^{n+1}}{(n+1)!} \, .$$

dove ξ è un punto compreso tra 0 e x che in generale dipenderà da n. Questa eguaglianza può anche essere equivalentemente scritta come

$$\sum_{k=0}^{n} \frac{x^k}{k!} = e^x - e^{\xi} \frac{x^{n+1}}{(n+1)!} \, . \tag{5.22}$$

Si noti ora che

$$\left| -e^{\xi} \frac{x^{n+1}}{(n+1)!} \right| \le e^{|x|} \frac{|x|^{n+1}}{(n+1)!} \, .$$

Poichè, qualunque sia x, si ha che

$$\lim_{n \to +\infty} \frac{|x|^{n+1}}{(n+1)!} = 0$$

mentre $e^{|x|}$ è costante, segue dal Teorema del confronto per successioni che

$$\lim_{n \to +\infty} \left(-e^{\xi} \frac{x^{n+1}}{(n+1)!} \right) = 0 \, ,$$

il che vuol dire che, per la (5.22), la serie di Taylor di e^x è sommabile qualunque sia $x \in \mathbb{R}$ e si ha

$$\sum_{k=0}^{+\infty} \frac{x^k}{k!} = e^x \, . \tag{5.23}$$

Si noti che questa serie, quando $x < 0$ non è neppure una serie a termini positivi. Tuttavia, con il ragionamento fatto, abbiamo ottenuto che esssa converge anche per $x < 0$ e che e^x è in ogni caso la sua somma.

Esempio 92 Consideriamo ora le funzioni trigonometriche $\sin x$ e $\cos x$ anch'esse di classe C^∞ su tutto $\mathbb{R}$ e calcoliamo i loro sviluppi di Taylor in 0, come fatto prima per l'esponenziale. Cominciamo con $f(x) = \sin x$. Si noti che

$$f'(x) = \cos x\,,\ f''(x) = -\sin x\,,\ f'''(x) = -\cos x\,,\ f^{(4)}(x) = \sin x\,.$$

Abbiamo quindi che le derivate si ripetono con un ciclo di 4. Possiamo scrivere

$$\begin{array}{ll} k = 2m & f^{(2m)}(x) = (-1)^m \sin x \\ \\ k = 2m+1 & f^{(2m+1)}(x) = (-1)^m \cos x \end{array}$$

Segue quindi che

$$\begin{array}{ll} k = 2m & f^{(2m)}(0) = 0 \\ \\ k = 2m+1 & f^{(2m+1)}(x) = (-1)^m \end{array}$$

Il polinomio di Taylor di grado $2m+1$ è dunque dato da

$$P_{2m+1}(x) = \sum_{k=0}^{m} (-1)^k \frac{x^{2k+1}}{(2k+1)!}\,.$$

Si noti che, essendo nulle tutte le derivate pari, si ha $P_{2m}(x) = P_{2m-1}(x)$. La formula di Taylor col resto di Lagrange, di grado $2m$, è:

$$\sin x = \sum_{k=0}^{m-1} (-1)^k \frac{x^{2k+1}}{(2k+1)!} + (-1)^m \sin\xi \frac{x^{2m}}{(2m)!}\,.$$

dove ξ è un punto compreso tra 0 e x che in generale dipenderà da n. Si noti ora che,

$$\left|(-1)^m \sin\xi \frac{x^{2m}}{(2m)!}\right| \leq \frac{|x|^{2m}}{(2m)!}\,.$$

Poichè, qualunque sia x, si ha che

$$\lim_{n\to+\infty} \frac{|x|^{2m}}{(2m)!} = 0\,,$$

come prima abbiamo che

$$\lim_{n\to+\infty} (-1)^m \sin\xi \frac{x^{2m}}{(2m)!} = 0\,.$$

Quindi, come nel caso dell'esponenziale, la serie di Taylor di $\sin x$ è sommabile qualunque sia $x \in \mathbb{R}$ e si ha

$$\sum_{k=0}^{+\infty} (-1)^{k+1} \frac{x^{2k+1}}{(2k+1)!} = \sin x\,. \tag{5.24}$$

Considerazioni molto simili si possono fare per la funzione $\cos x$ sempre con $x_0 = 0$. Con ragionamenti simili ai precedenti, si può in effetti mostrare che i polinomi di Taylor sono in tal caso dati da

$$P_{2m}(x) = \sum_{k=0}^{m} (-1)^k \frac{x^{2k}}{(2k)!}\,, \quad P_{2m+1}(x) = P_{2m}(x).$$

Analogamente al caso del seno, si può far vedere che questa serie è convergente qualunque sia $x \in\ R$ e che la somma è proprio data dalla funzione coseno calcolata nel punto x. Si ha dunque:

$$\sum_{k=0}^{+\infty} (-1)^k \frac{x^{2k}}{(2k)!} = \cos x\,. \tag{5.25}$$

Ci sono alcuni casi in cui si possono scrivere i polinomi di Taylor di una funzione senza fare neppure una derivata. Ecco di seguito alcuni semplici, ma importanti esempi.

Esempio 93 Sappiamo dalla teoria delle serie che

$$\frac{1}{1-x} = \sum_{k=0}^{+\infty} x^k\,, \quad \forall x\,:\, |x| < 1\,. \tag{5.26}$$

Si noti ora che

$$\sum_{k=n+1}^{+\infty} x^k = x^{n+1} \sum_{k=n+1}^{+\infty} x^{k-(n+1)} = x^{n+1} \sum_{h=0}^{+\infty} x^h = \frac{x^{n+1}}{1-x} = o(x^n)$$

per $x \to 0$. Si può dunque scrivere

$$\frac{1}{1-x} = \sum_{k=0}^{n} x^k + \sum_{k=n+1}^{+\infty} x^k = \sum_{k=0}^{n} x^k + o(x^n)\,.$$

Poichè $(1-x)^{-1} \in C^\infty(]-1,1[)$, si ha che necessariamente il polinomio di Taylor di detta funzione in 0 è dato proprio da

$$P_n(x) = \sum_{k=0}^{n} x^k\,.$$

La serie di Taylor della funzione $(1-x)^{-1}$ è dunque coincidente con la serie geometrica. Si noti che questa, a differenza della serie esponenziale converge solamente su $]-1,1[$. Poichè sappiamo che il polinomio di Taylor di una funzione $f(x)$ ha necessariamente la forma (5.21), ne segue che necessariamente

$$D^k\left[(1-x)^{-1}\right](0) = k!\,.$$

Si noti che quest'ultima relazione puà anche essere verificata direttamente e da essa si possono poi costruire i polinomi di Taylor nel modo tradizionale.

Da (5.26) si possono ottenere anche altri sviluppi. Ad esempio si ha, sempre per $|x|<1$,

$$\frac{1}{1+x} = \sum_{k=0}^{+\infty}(-1)^k x^k\,, \quad \forall x\,:\,|x|<1\,, \tag{5.27}$$

dalla quale si ottiene, con gli stessi ragionamenti di sopra, che i polinomi di Taylor della funzione $(1+x)^{-1}$ nel punto 0, sono dati da

$$P_n(x) = \sum_{k=0}^{n}(-1)^k x^k\,.$$

Ancora, sempre da (5.26), possiamo scrivere , sempre per $|x|<1$,

$$\frac{1}{1+x^2} = \sum_{k=0}^{+\infty}(-1)^k x^{2k}\,, \quad \forall x\,:\,|x|<1\,, \tag{5.28}$$

dalla quale si ottiene che i polinomi di Taylor della funzione $(1+x^2)^{-1}$ nel punto 0, sono dati da

$$P_{2n}(x) = \sum_{k=0}^{n}(-1)^k x^{2k}\,, \quad P_{2n+1}(x) = P_{2n}(x)\,.$$

Vediamo un altro importante sviluppo:

Esempio 94 Sia

$$f(x) = \ln(1+x)\,.$$

Vogliamo calcolare lo sviluppo di Taylor in 0. Si noti che anche in questo caso ci troviamo in presenza di una funzione C^∞ quindi essa ammetterà sviluppi di ogni ordine. Il problema è calcolare le derivate in 0. Si noti che

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

della quale abbiamo già calcolato i polinomi di Taylor (5.27). Dalla (5.27) si possono determinare immediatamente le derivate della funzione $f'(x)$ in 0, ottenendo,

$$f^{(k+1)}(0) = (-1)^k k!\,, k = 0, 1, \dots\,,$$

(per esercizio verificare direttamente che queste sono effettivamente le derivate della funzione $f'(x)$ in 0). Ricordando poi che $f(0) = 0$, possiamo scrivere i polinomi di Taylor della $f(x)$ in 0:

$$P_n(x) = \sum_{k=1}^{n} \frac{(-1)^{k-1}(k-1)!}{k!} x^k = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k} x^k\,.$$

Si può dimostrare che, in questo caso, la serie di Taylor converge per $x \in]-1, 1[$. Questo verrà visto nei corsi successivi.

Esercizio 5.8 Determinare i polinomi di Taylor della funzione $\arctan x$ in $x = 0$.

Esercizio 5.9 * Si definisca, per ogni $\alpha \in \mathbb{R}$ e ogni $n \in \mathbb{N}$,

$$\binom{\alpha}{0} = 1\,,$$

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)\cdots(\alpha-n+1)}{n!} \quad \text{se } n \in \mathbb{N}\,.$$

Questo oggetto è una naturale generalizzazione degli usuali coefficienti binomiali, che erano però definiti solo per α intero e maggiore o uguale di n.

Si scrivano i polinomi di Taylor di centro $x_0 = 0$ per la funzione $f(x) = (1+x)^\alpha$, mostrando che

$$P_n(x) = \sum_{k=0}^{n} \binom{\alpha}{k} x^k\,.$$

5.5 Applicazioni della formula di Taylor

5.5.1 Calcolo di limiti

Vediamo ora come calcolare limiti assai più complicati dei precedenti facendo uso degli sviluppi in serie delle funzioni elementari e della simbologia prima introdotta.

Esempio 95 Calcoliamo

$$\lim_{x\to+\infty 0} \frac{e^{\sin x} - 1 - \sin x - \frac{1}{2}(\sin x)^2}{(1-\cos x)\sin x}\,.$$

Chiamiamo $f(x)$ la funzione di cui dobbiamo calcolare il limite. Siamo in presenza di una forma di indecisione del tipo $0/0$, che indagheremo usando gli sviluppi di Taylor delle funzioni e^x, $\sin x$, $\cos x$. Cominciamo dal numeratore. Dobbiamo prima di tutto esaminare il termine $e^{\sin x}$; quando si devono sviluppare funzioni composte come questa, ricordate sempre che è opportuno cominciare a sviluppare il termine *più interno*, cioè in questo caso la funzione seno. Possiamo scrivere che

$$\sin x = x - \frac{x^3}{6} + o(x^3)\,.$$

Chi ci dice che fermarsi al terzo ordine (cioè scegliere di approssinare il seno con un polinomio di grado tre) sia la scelta giusta? A priori ciò non è detto: il grado $n = 3$ potrebbe essere troppo alto, cioè potrebbe darsi che si finiscano per scrivere termini trascurabili inutilmente, così come troppo basso, cioè potrebbe darsi che l'approssimazione scelta non sia abbastanza precisa da permettere di calcolare il limite. Non c'è purtroppo modo di saperlo a priori; occorre fare dei tentativi. Vedremo tra poco cosa sarebbe successo se avessimo usato l'approssimazione $\sin x \sim x$, cioè se ci fossimo fermati al primo ordine nello sviluppo del seno.

Sviluppiamo ora l'esponenziale:

$$e^y = 1 + y + \frac{y^2}{2} + \frac{y^3}{6} + o(y^3)\,,$$

valido per ogni $y \in \mathbb{R}$ implica che, posto

$$y = x - \frac{x^3}{6} + o(x^3)$$

(si osservi che $o(x^3)$ è una *funzione* di x, la cui espressione non è data esplicitamente, ma che si sa essere trascurabile rispetto a x^3 quando $x \to 0$), si abbia

$$\begin{aligned} e^{\sin x} = {} & 1 + \left[x - \frac{x^3}{6} + o(x^3)\right] + \frac{\left[x - \frac{x^3}{6} + o(x^3)\right]^2}{2} + \frac{\left[x - \frac{x^3}{6} + o(x^3)\right]^3}{6} \\ & + o\left(\left[x - \frac{x^3}{6} + o(x^3)\right]^3\right) \\ & = 1 + x + \frac{x^2}{2} + o(x^3)\,. \end{aligned}$$

L'ultimo passaggio vi potrà sembrare un po' misterioso. Così non è: semplicemente si è notato per prima cosa che

$$\left[x - \frac{x^3}{6} + o(x^3)\right]^2 = x^2 + o(x^3)\,, \tag{5.29}$$

visto che sviluppando il quadrato si ha

$$\begin{aligned}&\left[x - \frac{x^3}{6} + o(x^3)\right]^2 \\ &= x^2 + \frac{x^6}{36} + o(x^3)o(x^3) - \frac{x^4}{3} + 2xo(x^3) - \frac{x^3}{3}o(x^3)\,.\end{aligned}$$

È chiaro che i termini del tipo x^4 e x^6 sono trascurabili rispetto a x^2. Inoltre le proprietà del simbolo "o piccolo ci dicono anche che

$$2xo(x^3) = o(x^4), \quad \frac{x^3}{3}o(x^3) = o(x^6)$$

(le costanti numeriche sono irrilevanti: se f è trascurabile rispetto a g nel senso che il loro rapporto tende a zero, lo stesso accade se a f si sostituisce cf con $c \in \mathbb{R}$), e che

$$o(x^3)o(x^3) = o(x^6)\,;$$

quindi (5.29) è vera. Simili considerazioni mostrano che

$$\left[x - \frac{x^3}{6} + o(x^3)\right]^3 = x^3 + o(x^3)\,, \tag{5.30}$$

Quindi, poiché

$$\left[x - \frac{x^3}{6} + o(x^3)\right]^3 \sim x^3 \quad \text{per } x \to 0\,,$$

si ha anche che

$$o\left(\left[x - \frac{x^3}{6} + o(x^3)\right]^3\right) = o(x^3)\,.$$

Procedendo in maniera analoga si mostra che

$$\begin{aligned}(\sin x)^2 &= \left(x - \frac{x^3}{6} + o(x^3)\right)^2 \\ &= x^2 + o(x^3)\,.\end{aligned}$$

Quindi il numeratore di f soddisfa l'identità

$$e^{\sin x} - 1 - \sin x + \frac{1}{2}(\sin x)^2 = 1 + x + \frac{x^2}{2} + o(x^3) - 1 - x + \frac{x^3}{6} - \frac{x^2}{2} + o(x^3)$$
$$= \frac{x^3}{6} + o(x^3)\,.$$

Potete facilmente verificare, usando gli sviluppi del seno e del coseno, che il denominatore di f soddisfa invece l'identità

$$(1 - \cos x)\sin x = \frac{x^3}{2} + o(x^3)\,.$$

Quindi

$$f(x) = \frac{x^3/6 + o(x^3)}{x^3/2 + o(x^3)}\,.$$

A questo punto possiamo concludere che

$$f(x) \sim \frac{x^3/6}{x^3/2} = \frac{1}{3}$$

per $x \to 0$ e quindi che

$$\lim_{x\to 0} f(x) = \frac{1}{3}\,.$$

Se avessimo utilizzato degli sviluppi arrestati a gradi più bassi, non saremmo stati in grado di concludere alcunchè. In effetti, ad esempio, se ci fossimo fermati al primo ordine nello sviluppo del $\sin x$:

$$\sin x = x + o(x)\,,$$

l'unica conclusione che si sarebbe potuta trarre sul numeratore sarebbe stata la seguente

$$e^{\sin x} - 1 - x - 1/2(\sin x)^2 = o(x)\,.$$

Questa conclusione è *corretta*, ma non permette di dir nulla sul limite di f. In effetti ciò mostrerebbe solo, per le proprietà del simbolo o piccolo, che

$$f(x) \sim_{x\to 0} \frac{o(x)}{\frac{x^3}{2}} = o\left(\frac{1}{x^2}\right)\,.$$

Di nuovo ciò è corretto, ma non permette ovviamente di dire alcunché sul valore del limite. Fermarsi quindi a un ordine troppo basso nello sviluppo non è sbagliato, ma non permette di calcolare il limite.

Esercizio 5.10 Calcolare i seguenti limiti:

$$\lim_{x\to 0} \frac{\sin x - x}{x^2}\,,$$

$$\lim_{x\to 0} \frac{x\sin x + 1 - \cos(2x)}{x(e^x - 1)}\,,$$

$$\lim_{x\to 0} \frac{\log(1+x^2) - x^2}{x^2 - (\sin x)^2}\,,$$

$$\lim_{x\to +\infty} x\left[\log\left(1+\frac{1}{x^2}\right) - \frac{1}{x^2}\right]\,,$$

$$\lim_{x\to +\infty} x\left[\left(1+\frac{1}{x}\right)^x - e\right]\,.$$

R: $0,\ 3,\ -3/2,\ 0,\ -e/2$.

5.5.2 Concavità e convessità

Scopo di questa sezione è discutere brevemente due applicazioni della formula di Taylor. La prima di esse consisterà nello studio del concetto di *concavità* e di *convessità* di una funzione. Diamo per cominciare una definizione.

Definizione 5.20 Sia f una funzione di classe $C^1(]a,b[)$. Si dice che f è *convessa* in un punto $x_0 \in]a,b[$ se esiste $\delta > 0$ tale che il grafico di f si trovi, per $x \in]x_0-\delta, x_0+\delta[$, sopra il grafico della retta tangente al grafico di f in x_0, cioè se

$$f(x) \geq f(x_0) + f'(x_0)(x-x_0) \quad \forall x \in]x_0-\delta, x_0+\delta[\,.$$

La funzione f si dice *concava* in x_0 se esiste $\delta > 0$ tale che il grafico di f si trovi, per $x \in]x_0-\delta, x_0+\delta[$, sotto il grafico della retta tangente al grafico di f in x_0, cioè se

$$f(x) \leq f(x_0) + f'(x_0)(x-x_0) \quad \forall x \in]x_0-\delta, x_0+\delta[\,.$$

La funzione f si dice convessa in $]a,b[$ se è convessa in ogni punto di tale intervallo, convessa in $]a,b[$ se è concava in ogni punto del suo intervallo.

Una definizione identica si dà per la convessità e la concavità di funzioni definite su intervalli chiusi $[a, b]$, con le usuali convenzioni sul significato delle condizioni richieste ai bordi dell'intervallo, e anche su intervalli del tipo $[a, b[$ e $]a, b]$ (ciò sarà usato più avanti).

Esempio 96 La funzione $f(x) = x^2$ è convessa su $\mathbb{R}$. In effetti, fissato $x_0 \in \mathbb{R}$, si noti che $f'(x_0) = 2x_0$ e che quindi la disuguaglianza da verificare è la seguente:

$$x^2 - x_0^2 \geq 2x_0(x - x_0)$$

o equivalentemente

$$(x - x_0)(x + x_0) \geq 2x_0(x - x_0)\,, \tag{5.31}$$

almeno per x vicino a x_0. Dividiamo in due casi il problema: se $x > x_0$ la disuguaglianza scritta equivale a chiedere che

$$x + x_0 \geq 2x_0$$

che è ovviamente vera per tali x. Se invece $x < x_0$ la disuguaglianza da verificare è la seguente:

$$x + x_0 \leq 2x_0$$

che di nuovo è vera per gli x considerati.

Osservazione. Alcuni di voi conosceranno una definizione apparentemente del tutto diversa di convessità e di concavità. In effetti la nostra definizione è un po' restrittiva, perché si applica solo a funzioni di classe C^1. Una funzione del tipo

$$f(x) = |x|$$

non è derivabile in $x = 0$ e quindi non ci si può porre il problema della sua convessità nei termini prima scritti. Si dà in effetti di solito una definizione diversa, che può essere applicata a una *generica* funzione: si dice infatti che f è convessa su un intervallo $[a, b]$ se accade che, presi due punti arbitrari $x_1, x_2 \in [a, b]$ e un arbitrario numero $\lambda \in]0, 1[$ si ha:

$$f(\lambda x_1 + (1 - \lambda)x_2) \leq \lambda f(x_1) + (1 - \lambda)f(x_2)\,. \tag{5.32}$$

Si dice che f è concava su $[a, b]$ se per tali x_1, x_2, λ si ha

$$f(\lambda x_1 + (1 - \lambda)x_2) \geq \lambda f(x_1) + (1 - \lambda)f(x_2)\,. \tag{5.33}$$

Graficamente la condizione (5.32) si interpreta (pensate perché!) dicendo che il segmento che congiunge i punti $(x_1, f(x_1))$, $(x_2, f(x_2))$ sta sempre sopra il grafico della funzione in $[x_1, x_2]$. Si tratta quindi di di una definizione abbastanza simile a quella vista prima per funzioni di classe C^1, ma ora non è necessario nemmeno assumere (a priori...) che f sia continua.

In effetti, si può dimostrare che le due definizioni *coincidono* se f è una funzione di classe C^1.

La formula di Taylor ci permette di dare una semplice condizione che implica la proprietà di concavità o di convessità per funzioni di classe C^2. In effetti si ha il seguente risultato.

Teorema 5.21 *Sia f di classe $C^2(]a,b[)$ e sia $x_0 \in]a,b[$. Si supponga che $f''(x_0) > 0$ (risp. $f''(x_0) < 0$). Allora la funzione f è convessa in x_0 (risp. concava in x_0).*

Dimostrazione È sufficiente scrivere la formula di Taylor al secondo ordine con resto di Lagrange per f tra x_0 e x:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(\xi)}{2}(x - x_0)^2$$

e osservare che, se ad esempio $f''(x_0) > 0$, e x è sufficientemente vicino a x_0, allora anche $f''(\xi) > 0$. Ne segue che, per tali x,

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0)\,.$$

Questo implica che f è convessa in x_0. La condizione per la concavità si mostra in modo identico. ■

Esempio 97 La funzione $f(x) = e^x$ è convessa su $\mathbb{R}$. In effetti la sua derivata seconda è ancora e^x, che è sempre positiva. Questo fatto, unito al fatto che la derivata di f in x_0 vale e^{x_0}, implica la validità delle seguenti disuguaglianze, a priori non ovvie:

$$\begin{array}{rcl} e^x & = & f(x) \\ & \geq & f(x_0) + f'(x_0)(x - x_0) \\ & = & e^{x_0} + e^{x_0}(x - x_0) \end{array}$$

per ogni x, x_0 reali. Dividendo per e^{x_0} si ottiene equivalentemente

$$e^{x - x_0} \geq 1 + (x - x_0)$$

o anche, visto che x e x_0 sono arbitrari, la disuguaglianza

$$e^y \geq 1 + y \quad \forall y \in \mathbb{R}\,.$$

Analogamente la funzione $g(x) = \log x$ è concava su $]0, +\infty[$, poichè la sua derivata seconda vale $-1/x^2$ che è sempre negativa. La concavità di g in x_0 si può riscrivere come segue, ricordando che $g'(x) = 1/x$:

$$\log x \leq \log x_0 + \frac{1}{x_0}(x - x_0)\,,$$

per ogni $x, x_0 > 0$, che può anche essere scritta come

$$\log\left(\frac{x}{x_0}\right) \leq \frac{x}{x_0} - 1$$

o anche, per l'arbitrarietà di x e x_0, come

$$\log y \leq y - 1 \quad \forall y > 0\,.$$

Entrambe le disuguaglianze ora dimostrate non sono probabilmente così ovvie, anche se graficamente non sono sorprendenti.

Non è detto che, in un fissato punto x_0, una funzione anche di classe C^2 sia necessariamente concava o convessa. Un esempio tipico è la funzione

$$f(x) = x^3\,,$$

la cui tangente nell'origine è l'asse delle ascisse, che *attraversa il grafico della curva stessa*. Punti di questo tipo si dicono punti di *flesso*, o inflessionali, per f. Più formalmente:

Definizione 5.22 Sia f una funzione di classe $C^1(]a, b[)$. Si dice che f ha un punto di flesso in x_0 se esiste $\delta > 0$ tale che f è convessa su $[x_0 - \delta, x_0[$ e concava su $]x_0, x_0 + \delta]$, o viceversa.

Non è troppo difficile dimostrare (ma ciò non sarà fatto per brevità) che una condizione *necessaria* per l'esistenza di un punto di flesso per una funzione di classe C^2 è la seguente:

Proposizione 5.23 *Sia f una funzione di classe $C^2(]a, b[)$. Se $x_0 \in]a, b[$ è un punto di flesso, allora $f''(x_0) = 0$.*

Esempio 98 La condizione data sopra è solo necessaria ma non, in generale, sufficiente. Un esempio banale è dato dalla funzione

$$f(x) = x^4$$

che ha derivata seconda nulla nell'origine, ma è ovunque convessa.

Diamo una condizione sufficiente per l'esistenza di un punto di flesso.

Teorema 5.24 *Sia f una funzione di classe $C^2(]a,b[)$. Si supponga inoltre che, assegnato $x_0 \in]a,b[$ si abbia, per un opportuno $\delta > 0$,*

$$\begin{array}{ll} f''(x_0) = 0\,, & \\ f''(x) > 0 & \text{se } x \in]x_0, x_0+\delta]\,, \\ f''(x) < 0 & \text{se } x \in [x_0-\delta, x_0[\,, \end{array}$$

oppure che si abbia alternativamente

$$\begin{array}{ll} f''(x_0) = 0\,, & \\ f''(x) < 0 & \text{se } x \in]x_0, x_0+\delta]\,, \\ f''(x) > 0 & \text{se } x \in [x_0-\delta, x_0[\,. \end{array}$$

Allora f ha un punto di flesso in x_0.

Dimostrazione La dimostrazione è immediata. In effetti se vale il primo tipo di ipotesi si ha che, per il Teorema 5.21, la funzione f è convessa in $]x_0, x_0+\delta]$ e concava in $[x_0-\delta, x_0]]$, mentre vale il viceversa sotto il secondo tipo di ipotesi. In entrambi i casi la presenza di un flesso in x_0 segue dalla definizione. ■

5.5.3 Calcolo approssimato

Scopo di questa sezione è mostrare che la formula di Taylor con resto di Lagrange si presta all'approssimazione numerica dei valori di funzioni irrazionali. Vediamo come in un esempio.

Esempio 99 Supponiamo che si voglia avere qualche informazione sul valore numerico della quantità $\sin(1/3)$. Dare informazioni su tale valore vuole dire *approssimarlo* con un numero (possibilmente razionale) e *dare una stima* dell'errore commesso. L'idea è di usare la formula di Taylor con resto di Lagrange, stimando per l'appunto il resto. Sappiamo infatti che, definita la funzione

$$f(x) = \sin x\,,$$

la formula di Taylor con resto di Lagrange per f centrata in $x_0 = 0$ afferma che, per ogni $n \in \mathbb{N}$ e ogni $x \in \mathbb{R}$:

$$f(x) = \sum_{k=0}^{n-1} (-1)^k \frac{x^{2k+1}}{(2k+1)!} + \frac{f^{(2n+1)}(\xi)}{(2n+1)!} x^{2n+1}$$

dove ξ è un opportuno punto compreso tra 0 e x (o tra x e 0, a seconda del segno di x). In particolare, posto $x = 1/3$ se ne ricava

$$\sin\left(\frac{1}{3}\right) = \sum_{k=0}^{n-1}(-1)^k \frac{1}{(2k+1)!\,3^{2k+1}} + \frac{f^{(2n+1)}(\xi)}{(2n+1)!}\frac{1}{3^{2n+1}}\,.$$

Osserviamo ora che il resto nella formula appena scritta può essere stimato facilmente: siccome le derivate di qualunque ordine della funzione $\sin x$ sono sempre seni e coseni (eventualmente con un segno meno), si ha che, posto

$$R_{2n+1} := \frac{f^{(2n+1)}(\xi)}{(2n+1)!}\frac{1}{3^{2n+1}}$$

vale la stima

$$|R_{2n+1}(x)| \leq \frac{1}{3^{2n+1}(2n+1)!}\,.$$

Ma allora, fissato n posso stimare quanto è grande l'errore commesso approssimando la quantità $\sin(1/3)$ con la quantità, facilmente calcolabile,

$$\sum_{k=0}^{n-1}(-1)^k \frac{1}{(2k+1)!3^{2k+1}} \ :$$

tale errore è infatti non superiore a

$$\frac{1}{3^{2n+1}(2n+1)!}\,.$$

Quindi l'errore commesso stimando come sopra $\sin(1/3)$ per, ad esempio, $n = 4$, potete rispondere che

$$\begin{aligned} \sin(1/3) &= \frac{1}{3} - \frac{1}{3!\,3^3} + \frac{1}{5!\,3^5} - \frac{1}{7!\,3^7} + \varepsilon \\ &= 0.327194696 + \varepsilon\,, \end{aligned}$$

dove l'errore ε non supera la quantità

$$\frac{1}{9!\,3^9} \approx 1.4 \times 10^{-10}\,.$$

L'approssimazione per $\sin(1/3)$ ottenuta usando lo sviluppo di Taylor al quarto ordine (quindi a un ordine relativamente basso) è quindi corretta fino alla nona cifra decimale!

Alternativamente, è spesso necessario *fissare* l'errore massimo che si può commettere, e questo consente di determinare l'ordine a cui si deve scrivere lo sviluppo. Se per esempio volessimo approssimare la funzione $\sin x$ nell'intervallo $[0, 1/3]$ con un *polinomio*, in modo che l'errore commesso non superi 10^{-5}, è sufficiente chiedere, per i calcoli precedenti, che n sia tale da far sì che il resto (in valore assoluto) non superi tale errore. Si vede facilmente che basta scegliere $n = 3$, e che quindi il polinomio (di quinto grado, vista la struttura dello sviluppo per la funzione seno)

$$P(x) = x - \frac{x^3}{6} + \frac{x^5}{120}$$

approssima la funzione $\sin x$ nell'intervallo $[0, 1/3]$ con un errore non superiore a 10^{-5}.

Esercizio 5.11 Calcolare il valore di $\sqrt[3]{e}$ con un errore inferiore a 10^{-6}.
R: $1,395612$

Esercizio 5.12 Determinare un polinomio che approssimi la funzione $f(x) = \sinh x$ nell'intervallo $[-1/2, 1/3]$ con un errore inferiore a 10^{-4}.
R: $\sum_{k=0}^{2} \frac{x^{2k+1}}{(2k+1)!}$.

Esercizio 5.13 Dare una stima dell'errore massimo che si commette approssimando la funzione $f(x) = e^x$ col suo polinomio di Taylor di grado 5, centrato in $x_0 = 0$, nell'intervallo $[0, 2]$.
R: 0.65

5.5.4 Irrazionalità del numero e

Per quanto strano può sembrare, la formula di Taylor consente di dimostrare in maniera molto semplice che e è un numero irrazionale. Facciamo vedere come si fa in modo che possiate rendervi conto della versatilità di tale strumento.

In effetti, usiamo la formula di Taylor per la funzione

$$f(x) = e^x$$

con centro in $x_0 = 0$, all'ordine n, per ottenere

$$e^x = \sum_{k=1}^{n-1} \frac{x^k}{k!} + \frac{e^\xi}{n!} x^n \,,$$

per un opportuno ξ compreso tra 0 e x (che assumiamo essere positivo), dove abbiamo usato il fatto che le derivate di qualsiasi ordine di e^x coincidono con e^x stesso. Ponendo $x = 1$ si trova

$$e = \sum_{k=1}^{n-1} \frac{1}{k!} + \frac{e^\xi}{n!}$$

per un opportuno ξ compreso tra 0 e 1. Quindi

$$0 < e - \sum_{k=1}^{n-1} \frac{1}{k!} = \frac{e^\xi}{n!} \leq \frac{e}{n!} < \frac{3}{n!}\,.$$

Moltiplicando entrambi i membri dell'ultima formula per $(n-1)!$ otteniamo quindi

$$0 < e(n-1)! - (n-1)! \sum_{k=1}^{n-1} \frac{1}{k!} < \frac{3}{n}$$

per ogni intero positivo n. Se per assurdo e fosse un numero razionale, $e = p/q$ con p e q interi positivi primi tra loro, allora si potrebbe scegliere un $n \in \mathbb{N}$ tale che $e(n-1)!$ sia intero (basta prendere n tale che $n-1$ sia un multiplo di q) e tale che inoltre $n > 3$. Poiché inoltre la quantità

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

è un numero intero, si sarebbe costruito un numero intero

$$m = e(n-1)! - (n-1)! \sum_{k=1}^{n-1} \frac{1}{k!}$$

che soddisfa la condizione

$$0 < m < 1\,.$$

Ciò è evidentemente impossibile.

Capitolo 6

Il calcolo integrale

6.1 Invertire il procedimento di derivazione

Supponiamo che $f :]a, b[\to \mathbb{R}$ sia una funzione derivabile a noi ignota della quale conosciamo la funzione derivata $g = f' :]a, b[\to \mathbb{R}$. Possiamo da $g(x)$ risalire alla funzione $f(x)$? In altri termini, più geometrici: se di una funzione conosciamo, punto per punto, la pendenza (cioè il coefficiente angolare della retta tangente), possiamo ricostruire la funzione stessa? Così come abbiamo posto la domanda, essa ha risposta negativa. In effetti, se $f(x)$ è una funzione la cui derivata è uguale a $g(x)$ per ogni $x \in]a, b[$, ogni altra funzione del tipo $f(x) + k$ dove k è una costante, ha la proprietà che la sua derivata è ancora $g(x)$ poichè la derivata di una costante è identicamente nulla. Tuttavia, come mostra il seguente risultato, la costante additiva è la sola indeterminazione possibile nel determinare $f(x)$:

Teorema 6.1 *Siano f_1, f_2 e g tre funzioni definite sull'intervallo $]a, b[$ a valori reali, con f_1 e f_2 derivabili su $]a, b[$ tali che*

$$\begin{aligned} f_1'(x) &= g(x) \qquad \forall x \in]a, b[\,, \\ f_2'(x) &= g(x) \qquad \forall x \in]a, b[\,. \end{aligned}$$

Allora esiste una costante k tale che

$$f_2(x) = f_1(x) + k \quad \forall x \in]a, b[\,.$$

Dimostrazione Segue da un'applicazione immediata del Corollario 5.11. In effetti, se consideriamo $f(x) = f_2(x) - f_1(x)$, si ha che

$$f'(x) = f_2'(x) - f_1'(x) = g(x) - g(x) = 0 \quad \forall x \in]a, b[\,.$$

Quindi f è costante, cioè esiste $k \in \mathbb{R}$ tale che $f(x) = k$ per ogni $x \in]a, b[$. Per come $f(x)$ è stata definita, ne segue immediatamente la tesi. ■

Le funzioni derivabile che hanno come derivata la funzione $g(x)$ vengono dette *primitive* della funzione $g(x)$. Il risultato precedente mostra come due primitive di una stessa funzione su di un intervallo differiscano tra di loro per una costante additiva. In altre parole, se abbiamo una primitiva di $g(x)$, tutte le altre si ottengono aggiungendo costanti additive. Il problema che rimane da risolvere è come sapere se una data funzione ammette primitive e come fare in pratica a trovare una qualunque di queste primitive.

Si noti che per alcune funzioni è immediato trovare una primitiva. Ad esempio si ha che se $f(x) = x^n$ con $n \in \mathbb{N}$ una primitiva su tutto quanto $\mathbb{R}$ è data da $F(x) = x^{n+1}/(n+1)$. Basta derivare F e verificare che si ottiene $f(x)$. Più in generale, se $f(x) = x^\alpha$ con $\alpha \neq -1$ una primitiva su $]0, +\infty[$ è data da $F(x) = x^{\alpha+1}/(\alpha+1)$. Invece se $f(x) = x^{-1}$ si ottiene che una primitiva in $]0, +\infty[$ è data da $F(x) = \ln x$.

Tutte le considerazioni fatte in questa sezione si estendono al caso di intervalli chiusi e limitati $[a, b]$ (o anche contenenti uno solo dei due estremi). L'unica cosa da notare è che la derivabilità nei punti a e b è da intendersi come l'esistenza, rispettivamente, della derivata destra e sinistra. Questa estensione ci tornerà utile in seguito.

Osservazione. Si noti come ci siano sicuramente funzioni che non ammettono una primitiva. Ad esempio consideriamo $f :]-1, 1[\to \mathbb{R}$ definita da

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

Se ammettesse una primitiva $F(x)$ tale funzione sarebbe la primitiva di -1 su $]-1, 0[$ e di 1 su $[0, 1[$. Quindi sarebbe del tipo

$$F(x) = \begin{cases} -x + k_1 & \text{se } x \in]-1, 0[\\ x + k_2 & \text{se } x \in [0, 1[\,. \end{cases}$$

per qualche $k_1, k_2 \in \mathbb{R}$. Si noti tuttavia che per nessun valore di k_1 e k_2 tale funzione $F(x)$ è derivabile su $]-1, 1[$ e dunque non può essere una primitiva su $]-1, 1[$ di $f(x)$. Più in generale si può provare, con tecniche simili, che nessuna funzione che presenti discontinuità di prima specie ammette primitive.

Osservazione. Si noti infine che il Teorema 6.1 non si estende al caso di domini che non siano intervalli. Ad esempio le funzioni

$$f_1(x) = 1, \;\; x \in \mathbb{R} \setminus \{0\}\,,$$

$$f_2(x) = \begin{cases} 1, & x > 0 \\ 0, & x < 0 \end{cases}$$

sono entrambe primitive della funzione costantemente uguale a 0 su $\mathbb{R} \setminus \{0\}$. Tuttavia, esse non differiscono per una costante.

6.2 Il calcolo delle aree

Introduciamo ora un altro problema che ha apparentemente niente a che vedere con il precedente: il calcolo dell'area di figure planari. La più semplice figura per il calcolo dell'area è certamente il rettangolo; da questa si passa ai triangoli e successivamente a tutte le figure poligonali in quanto decomponibili in un numero finito di triangoli. Che si può dire delle figure non poligonali? Il cerchio, ad esempio, non rientra in questa categoria e tuttavia la formula dell'area è ben nota da un paio di millenni. Come la si è ottenuta? Con un procedimento di approssimazione, considerando poligoni iscritti (o circoscritti) al cerchio con un numero di lati sempre più grande. Dietro la formula dell'area del cerchio c'è dunque un procedimento limite. In questo capitolo noi presenteremo un procedimento limite del tutto generale per il calcolo delle aree di regioni del piano del tipo che ora descriviamo. Siano $f_1 : [a,b] \to \mathbb{R}$ e $f_2 : [a,b] \to \mathbb{R}$ tali che $f_1(x) \leq f_2(x)$ per ogni $x \in [a,b]$ e consideriamo

$$A = \big\{(x,y) \in \mathbb{R} \times \mathbb{R} \,|\, a \leq x \leq b, \;\; f_1(x) \leq y \leq f_2(x)\big\}\,.$$

A viene detto un *dominio normale*; si veda la Figura 6.1

Esempio 100 Il cerchio è un dominio normale. In effetti se consideriamo il cerchio di raggio r centrato nell'origine del sistema di assi cartesiani

$$B(0,r) = \{(x,y) \in \mathbb{R} \times \mathbb{R} \,|\, x^2 + y^2 \leq r^2\}$$

si ha che esso può essere anche descritto come

$$B(0,r) = \Big\{(x,y) \in \mathbb{R} \times \mathbb{R} \,|\, -r \leq x \leq r, \;\; -\sqrt{r^2 - x^2} \leq y \leq \sqrt{r^2 - x^2}\Big\}\,.$$

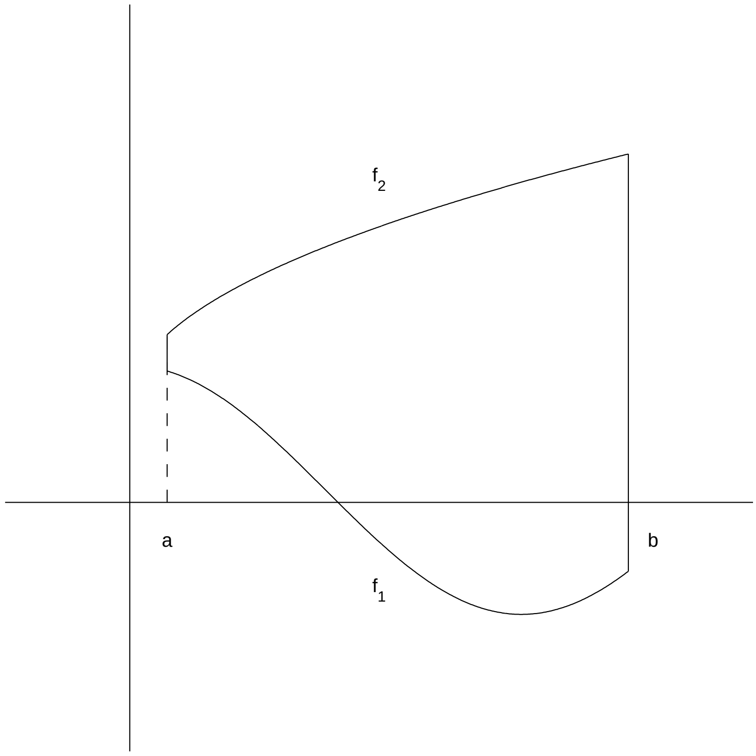


Figura 6.1:

Domini normali particolarmente semplici sono quelli del tipo

$$A = \{(x, y) \in \mathbb{R} \times \mathbb{R} \,|\, a \leq x \leq b,\ 0 \leq y \leq f(x)\}$$

dove $f : [a, b] \to \mathbb{R}$ è una funzione a valori non negativi.

L'idea del procedimento per il calcolo delle aree, che illustreremo in modo formale nella prossima sezione è la seguente: dividiamo l'intervallo $[a, b]$ in N sottointervalli (di uguale ampiezza, anche se ciò non è essenziale) e consideriamo il *plurirettangolo iscritto* come in Figura 6.2 (verrà definito rigorosamente più tardi). L'area del plurirettangolo, che dipende ovviamente da N, s_N cresce all'aumentare di N e 'dovrebbe' convergere all'area del dominio sotteso da f. Similmente potremo considerare il *plurirettangolo circoscritto* come in Figura 6.3 la cui area S_N decresce all'aumentare di N e 'dovrebbe' anch'essa convergere all'area del dominio sotteso da f.

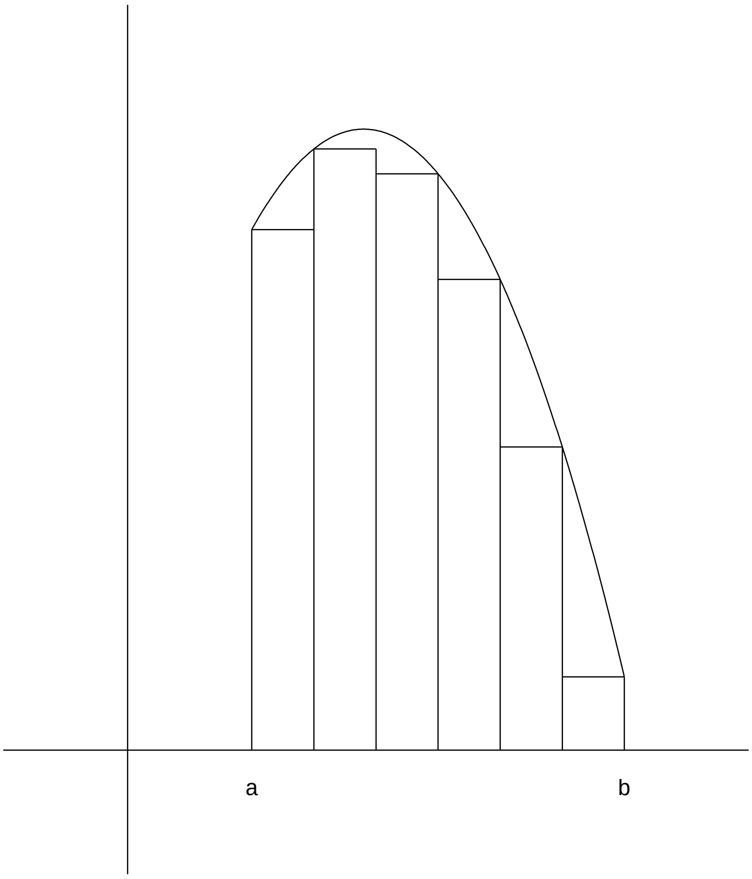


Figura 6.2:

Questi processi di convergenza vanno studiati con un po' di attenzione e come vedremo sarà necessario fare delle opportune ipotesi su f affinchè tutto funzioni.

6.3 L'integrazione definita

Il procedimento di approssimazione delle aree attraverso i plurirettangoli, informalmente illustrato alla fine della sezione precedente prende il nome di *integrazione definita* e verrà qui presentato in generale anche per funzioni

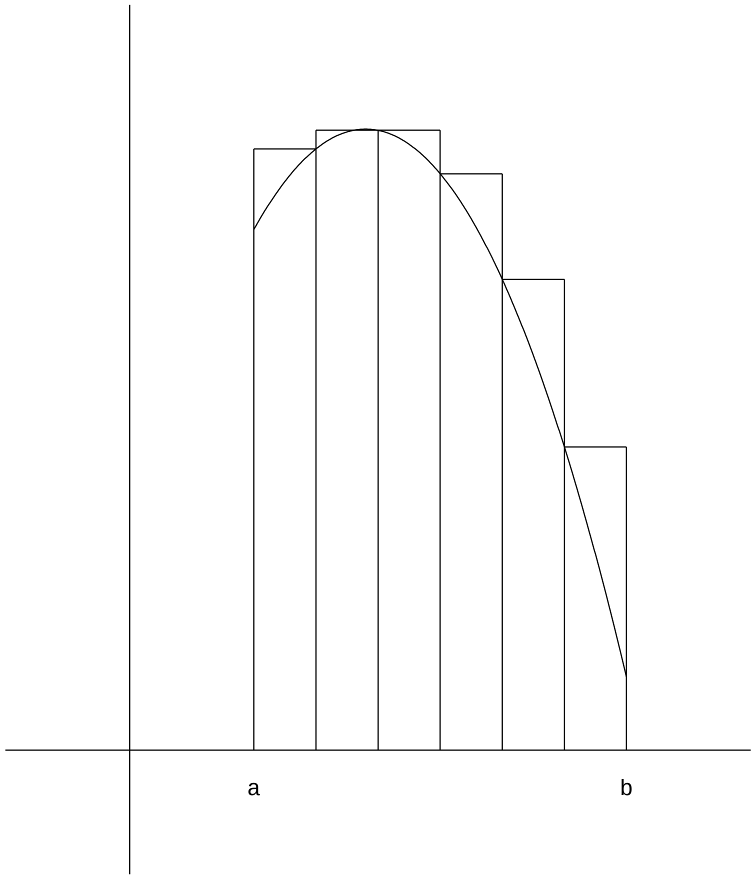


Figura 6.3:

non necessariamente a valori positivi. Ne discuteremo poi l'interpretazione geometrica.

Partiamo dunque da una funzione $f : [a, b] \to \mathbb{R}$ e facciamo la sola ipotesi che f sia limitata. Una *partizione* di $[a, b]$ è un sottoinsieme finito di punti di $[a, b]$ del tipo

$$\delta = \{x_1 = a < x_2 < x_3 < \cdots < x_{n+1} = b\} .$$

Denotiamo con $\Delta_{[a,b]}$ l'insieme di tutte le possibili partizioni dell'intervallo $[a, b]$. Ogni partizione $\delta \in \Delta$ come sopra determina una suddivisione dell'intervallo $[a, b]$ in n sottointervalli

$$I_1 = [x_1, x_2],\ I_2 = [x_2, x_3],\ \ldots,\ I_n = [x_n, x_{n+1}]\,.$$

Sia

$$L_k = \sup_{x \in I_k} f(x)\,, \qquad l_k = \inf_{x \in I_k} f(x)\,.$$

Definiamo la *somma superiore* e la *somma inferiore* di f relativa alla partizione δ come, rispettivamente:

$$S_\delta(f) = \sum_{k=1}^{n} L_k(x_{k+1} - x_k)\,, \qquad s_\delta(f) = \sum_{k=1}^{n} l_k(x_{k+1} - x_k)\,.$$

Definiamo ora l'*integrale definito inferiore* e l'*integrale definito superiore* di f su $[a, b]$ come, rispettivamente:

$$\overline{\int_a^b} f(x)\,\mathrm{d}x = \inf_{\delta \in \Delta_{[a,b]}} S_\delta(f)\,, \qquad \underline{\int_a^b} f(x)\,\mathrm{d}x = \sup_{\delta \in \Delta_{[a,b]}} s_\delta(f) v\,. \tag{6.1}$$

Finalmente diamo la seguente:

Definizione 6.2 Una funzione $f : [a, b] \to \mathbb{R}$ si dice *integrabile (secondo Riemann)* se

$$\overline{\int_a^b} f(x)\,\mathrm{d}x = \underline{\int_a^b} f(x)\,\mathrm{d}x\,.$$

Il loro valore comune viene detto l'*integrale definito* di f su $[a, b]$ e indicato con il simbolo

$$\int_a^b f(x)\,\mathrm{d}x\,.$$

Osservazione 1. Diamo ora un'interpretazione geometrica dei vari concetti introdotti nel caso in cui $f(x) \geq 0$ per ogni $x \in [a, b]$. L'insieme dei rettangoli aventi base I_k e altezze l_k formano quello che precedentemente avevamo chiamato il plurirettangolo iscritto, mentre l'insieme dei rettangoli aventi base I_k e altezze L_k formano il plurirettangolo circoscritto. Le somme $s_\delta(f)$ e $S_\delta(f)$ sono proprio le aree di questi due plurirettangoli. Dunque $\underline{\int_a^b} f(x)\,\mathrm{d}x$ rappresenta l'estremo superiore delle aree di tutti i possibili plurirettangoli inscritti, mentre $\overline{\int_a^b} f(x)\,\mathrm{d}x$ rappresenta l'estremo inferiore delle aree di tutti i possibili plurirettangoli circoscritti. Intuitivamente l'area sottesa dal grafico di f deve essere un numero compreso tra questi due valori. L'integrabilità di f, cioè il fatto che questi due valori siano eguali ci dice che in tal caso l'area può essere senza ambiguità definita proprio come questo valore comune; inoltre essa può venire approssimata, con arbitraria precisione, sia dall'interno con l'area dei plurirettangoli inscritti, che dall'esterno con l'area dei plurirettangoli circoscritti.

Osservazione 2. Nel caso in cui invece la funzione $f : [a, b] \to \mathbb{R}$ sia tale che $f(x) \leq 0$ per ogni $x \in [a, b]$, un ragionamento simile al precedente mostra che se f è integrabile, il valore dell'integrale $\int_a^b f(x)\,\mathrm{d}x$ può essere interpretato come l'area tra il grafico della f e l'asse x con il segno $-$ davanti. Nel caso più generale in cui f abbia cambiamenti di segno all'interno dell'intervallo di definizione $[a, b]$ vedremo più avanti (Osservazione 2 dopo la Proposizione 6.11) che l'integrale può essere visto come una sorta di somma algebrica di aree, positive quelle situate sopra l'asse delle x, negative quelle situate sotto.

Osservazione 3. Vedremo più avanti le motivazioni della notazione adottata per descrivere l'integrale. Si noti per ora che la variabile x che compare nella formula dell'integrale non giuoca nessun ruolo specifico. L'integrale definito

$$\int_a^b f(x)\,\mathrm{d}x$$

è un numero, la x entra solo all'interno e potrebbe essere sostituita da qualunque altra variabile, ad esempio

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b f(t)\,\mathrm{d}t = \int_a^b f(y)\,\mathrm{d}y\,.$$

Si ha il seguente semplice risultato:

Proposizione 6.3 *Sia $f : [a, b] \to \mathbb{R}$ una funzione integrabile. Allora si ha*

$$(b-a)\inf_{x\in[a,b]} f(x) \leq \int_a^b f(x)\,\mathrm{d}x \leq (b-a)\sup_{x\in[a,b]} f(x)\,.$$

Dimostrazione Basta osservare che

$$(b-a)\inf_{x\in[a,b]} f(x)\,, \qquad (b-a)\sup_{x\in[a,b]} f(x)$$

sono, rispettivamente, le somme inferiore e superiori di f relative alla partizione banale $\delta = \{a, b\}$ e poi sfruttare la definizione (6.1). ■

6.4 Criteri di integrabilità

Vogliamo ora cercare condizioni su f che garantiscano la sua integrabilità. Cominciamo con alcuni risultati preliminari.

Fissiamo un intervallo $[a, b]$ e consideriamo due partizioni $\delta_1, \delta_2 \in \Delta_{[a,b]}$. Si dice che δ_2 è *più fine* di δ_1 se $\delta_1 \subseteq \delta_2$ cioè se tutti i punti di δ_1 sono anche punti di δ_2. Ci servirà ancora un concetto. Sia

$$\delta = \big\{x_1 = a < x_2 < x_3 < \cdots < x_{n+1} = b\big\} \in \Delta_{[a,b]}\,.$$

Si definisce *parametro di finezza* della partizione δ, il numero

$$r(\delta) = \max\{x_2 - x_1, x_3 - x_2, \ldots, x_{n+1} - x_n\}\,.$$

Il parametro $r(\delta)$ è quindi il massimo delle ampiezze degli intervallini che definiscono la partizione.

Vale il seguente risultato intuitivo che dice che le somme superiori diminuiscono e quelle inferiori aumentano all'aumentare della finezza della partizione.

Proposizione 6.4 *Sia $f : [a, b] \to \mathbb{R}$ una funzione limitata e siano $\delta_1, \delta_2 \in \Delta_{[a,b]}$ con δ_2 più fine di δ_1. Allora si ha,*

$$S_{\delta_2}(f) \le S_{\delta_1}(f)\,, \qquad s_{\delta_2}(f) \ge s_{\delta_1}(f)\,. \tag{6.2}$$

Dimostrazione La partizione δ_2 si ottiene dalla δ_1 aggiungendo un numero finito di punti dell'intervallo $[a, b]$. Si può quindi pensare di passare dalla δ_1 alla δ_2 in un numero finito di passi aggiungendo un solo punto alla volta. Se facciamo vedere che le diseguaglianze (6.2) valgono quando si aggiunge un punto, per transitività esse varranno anche nel passare da δ_1 a δ_2. Supponiamo dunque di essere nel caso particolare in cui $\delta_2 = \delta_1 \cup \{\overline{x}\}$. Sia

$$\delta_1 = \{x_1 = a < x_2 < x_3 < \cdots < x_{n+1} = b\}$$

e definiamo

$$L_k = \sup_{x\in I_k} f(x)\,, \qquad l_k = \inf_{x\in I_k} f(x)\,.$$

dove, come prima, $I_k = [x_k, x_{k+1}]$. Supponiamo ora che $\overline{x} \in I_{\overline{k}} = [x_{\overline{k}}, x_{\overline{k}+1}]$. Definiamo

$$\begin{aligned} I'_{\overline{k}} &= [x_{\overline{k}}, \overline{x}]\,, & I''_{\overline{k}} &= [\overline{x}, x_{\overline{k}+1}]\,, \\ L'_{\overline{k}} &= \sup_{x \in I'_{\overline{k}}} f(x)\,, & L''_{\overline{k}} &= \sup_{x \in I''_{\overline{k}}} f(x)\,, \\ l'_{\overline{k}} &= \inf_{x \in I'_{\overline{k}}} f(x)\,, & l''_{\overline{k}} &= \inf_{x \in I''_{\overline{k}}} f(x)\,. \end{aligned}$$

Chiaramente

$$L'_{\overline{k}}, L''_{\overline{k}} \leq L_{\overline{k}}\,, \qquad l'_{\overline{k}}, l''_{\overline{k}} \geq l_{\overline{k}}\,.$$

Si ha dunque

$$\begin{aligned} S_{\delta_2}(f) &= \sum_{k=1}^{\overline{k}-1} L_k(x_{k+1} - x_k) + L'_{\overline{k}}(\overline{x} - x_{\overline{k}}) + L''_{\overline{k}}(x_{\overline{k}+1} - \overline{x}) + \sum_{k=\overline{k}+1}^{n} L_k(x_{k+1} - x_k) \\ &\leq \sum_{k=1}^{\overline{k}-1} L_k(x_{k+1} - x_k) + L_{\overline{k}}(\overline{x} - x_{\overline{k}}) + L_{\overline{k}}(x_{\overline{k}+1} - \overline{x}) + \sum_{k=\overline{k}+1}^{n} L_k(x_{k+1} - x_k) \\ &= \sum_{k=1}^{\overline{k}-1} L_k(x_{k+1} - x_k) + L_{\overline{k}}(x_{\overline{k}+1} - x_{\overline{k}}) + \sum_{k=\overline{k}+1}^{n} L_k(x_{k+1} - x_k) \\ &= \sum_{k=1}^{n} L_k(x_{k+1} - x_k) = S_{\delta_1}(f)\,. \end{aligned}$$

Un analogo calcolo mostra invece che $s_{\delta_2}(f) \geq s_{\delta_1}(f)$. ■

A questo punto possiamo provare la seguente:

Proposizione 6.5 *Sia $f : [a, b] \to \mathbb{R}$ una funzione limitata. Si ha sempre*

$$\overline{\int_a^b} f(x)\,\mathrm{d}x \geq \underline{\int_a^b} f(x)\,\mathrm{d}x\,.$$

Dimostrazione Faremo vedere che ogni somma superiore è non inferiore ad ogni somma inferiore. Questo, per come sono definiti gli integrali superiore e inferiore, implicherà la tesi. Fissiamo dunque due qualsiasi partizioni $\delta_1, \delta_2 \in \Delta_{[a,b]}$ e facciamo vedere che $S_{\delta_1} \geq s_{\delta_2}$. Si noti che nel caso in cui $\delta_1 = \delta_2$ questo è ovvio per come sono state definite le somme superiori e inferiori. Per vederlo nel caso generale si considera la partizione $\delta_1 \cup \delta_2$ ottenuta considerando tutti i punti sia di δ_1 che di δ_2. Si noti che per come è costruita $\delta_1 \cup \delta_2$ è più fine sia di δ_1 che di δ_2. Applicando dunque la Proposizione 6.4 si ottiene

$$S_{\delta_1} \geq S_{\delta_1 \cup \delta_2} \geq s_{\delta_1 \cup \delta_2} \geq s_{\delta_2}\,.$$

Il risultato è così dimostrato. ■

Siamo ora pronti per enunciare e dimostrare il criterio fondamentale di integrabilità.

Teorema 6.6 *Sia $f : [a,b] \to \mathbb{R}$ una funzione limitata. Allora f è integrabile su $[a,b]$ se e soltanto se per ogni $\epsilon > 0$ esiste una partizione $\delta \in \Delta_{[a,b]}$ tale che $S_\delta - s_\delta \leq \epsilon$.*

Dimostrazione Supponiamo prima che f sia integrabile. Allora,

$$\inf_{\delta \in \Delta_{[a,b]}} S_\delta(f) = \sup_{\delta \in \Delta_{[a,b]}} s_\delta(f) = \theta\,.$$

D'altra parte, per le proprietà degli estremi superiore ed inferiore (vedi Proposizioni 1.6 e 1.7), si ha che esistono partizioni $\delta_1, \delta_2 \in \Delta_{[a,b]}$ tali che

$$S_{\delta_1} \leq \theta + \frac{\epsilon}{2} \qquad s_{\delta_2} \geq \theta - \frac{\epsilon}{2}\,. \tag{6.3}$$

Consideriamo ora $\delta = \delta_1 \cup \delta_2$. In virtù della Proposizione 6.4 si ha che $S_\delta \leq S_{\delta_1}$ e $s_\delta \geq s_{\delta_2}$. Usando queste diseguaglianze e (6.3) si ha che

$$S_\delta - s_\delta \leq S_{\delta_1} - s_{\delta_2} \leq \theta + \frac{\epsilon}{2} - \left(\theta - \frac{\epsilon}{2}\right) = \epsilon$$

che è quello che volevamo provare.

Supponiamo ora che invece valga la condizione: per ogni $\epsilon > 0$ esiste una partizione $\delta \in \Delta_{[a,b]}$ tale che $S_\delta - s_\delta \leq \epsilon$ e facciamo vedere che allora f è integrabile. Supponiamo per assurdo che non lo sia, cioè che (per la Proposizione 6.5)

$$\overline{\int_a^b} f(x)\,\mathrm{d}x > \underline{\int_a^b} f(x)\,\mathrm{d}x\,.$$

Fissiamo ora un qualunque $\epsilon > 0$ tale che

$$\epsilon < \overline{\int_a^b} f(x)\,\mathrm{d}x - \underline{\int_a^b} f(x)\,\mathrm{d}x\,.$$

Allora qualunque sia $\delta \in \Delta_{[a,b]}$ si ha che

$$S_\delta - s_\delta \geq \overline{\int_a^b} f(x)\,\mathrm{d}x - \underline{\int_a^b} f(x)\,\mathrm{d}x > \epsilon$$

il che contraddice la condizione ipotizzata. Il teorema è così dimostrato. ■

Il criterio precedente ammette un'utile variante che è la seguente

Teorema 6.7 *Sia $f : [a,b] \to \mathbb{R}$ una funzione limitata. Allora f è integrabile su $[a,b]$ se e soltanto se per ogni $\epsilon > 0$ esiste $r > 0$ tale che per ogni partizione $\delta \in \Delta_{[a,b]}$ tale che $r(\delta) \leq r$ si ha che $S_\delta - s_\delta \leq \epsilon$.*

Dimostrazione La condizione espressa in questo criterio è, da un punto di vista logico, più forte di quella espressa nell'altro criterio di integrabilità, quindi essa implica a maggior ragione l'integrabilità. Far vedere invece che l'integrabilità implica questa condizione più forte è delicato e più difficile che per l'altro criterio. La dimostrazione viene così omessa.

■

I risultati precedenti si applicano spesso nella forma seguente. Sia δ_n la partizione ottenuta dividendo $[a,b]$ in n parti eguali (queste partizioni si dicono *uniformi*). Si noti che $r(\delta_n) = (b-a)/n$. Vale il seguente

Corollario 6.8 *Sia $f : [a,b] \to \mathbb{R}$ una funzione limitata. Allora f è integrabile se e soltanto se*

$$\lim_{n\to+\infty} [S_{\delta_n} - s_{\delta_n}] = 0\,. \tag{6.4}$$

Inoltre, in tal caso si ha:

$$\lim_{n\to+\infty} S_{\delta_n} = \int_a^b f(x)\,\mathrm{d}x = \lim_{n\to+\infty} s_{\delta_n}\,. \tag{6.5}$$

Dimostrazione Supponiamo f integrabile e facciamo vedere che vale (6.4). Fissiamo $\epsilon > 0$. Per il Teorema 6.7, esiste $r > 0$ tale che $S_\delta - s_\delta < \epsilon$ per ogni partizione δ tale che $r(\delta) \le r$. Poichè $r(\delta_n) \to 0$ segue che $r(\delta_n) \le r$ definitivamente. Quindi

$$|S_{\delta_n} - s_{\delta_n}| = S_{\delta_n} - s_{\delta_n} < \epsilon$$

definitivamente. Viceversa, se vale (6.4) allora dal Teorema 6.6 segue che f è integrabile. Infine le relazioni (6.5) semplicemente seguono da

$$s_{\delta_n} \le \int_a^b f(x)\,\mathrm{d}x \le S_{\delta_n}$$

e da (6.4). ■

Osservazione. Il risultato limite (6.5) offre una spiegazione informale del simbolismo adottato per denotare l'integrale. Al tendere di n verso $+\infty$, le somme finite delle aree di rettangoli S_{δ_n} e s_{δ_n} tendono a diventare delle somme infinite "$\int$ di rettangolini di base infinitesima $\mathrm{d}x$ e altezza $f(x)$. In realtà c'è un'altra motivazione per l'utilizzo di questo formalismo che apparirà più avanti.

Esempio 101 Consideriamo la funzione $f : [0,1] \to \mathbb{R}$ data da $f(x) = x^2$. Poichè f è crescente su $[0,1]$ si ha che

$$S_{\delta_n}(f) = \sum_{k=1}^{n} \left(\frac{k}{n}\right)^2 \frac{1}{n} = \frac{\sum\limits_{k=1}^{n} k^2}{n^3} = \frac{n(n+1)(2n+1)}{6n^3},$$

mentre

$$s_{\delta_n}(f) = \sum_{k=0}^{n-1} \left(\frac{k}{n}\right)^2 \frac{1}{n} = \frac{\sum\limits_{k=0}^{n-1} k^2}{n^3} = \frac{(n-1)n(2n-1)}{6n^3}.$$

Entrambe, per $n \to +\infty$ convergono a $1/3$. Questo significa, per il Corollario 6.8 che x^2 è integrabile su $[0,1]$ e si ha

$$\int_0^1 x^2 \, \mathrm{d}x = \frac{1}{3}.$$

Vorremmo ora mostrare con un altro esempio che esistono effettivamente funzioni non integrabili.

Esempio 102 Consideriamo la funzione $f : [0,1] \to \mathbb{R}$ data da

$$f(x) = \begin{cases} 0 & x \in [0,1] \cap \mathbb{Q} \\ 1 & x \in [0,1] \setminus \mathbb{Q}. \end{cases}$$

Questa funzione è detta *funzione di Dirichlet.* Consideriamo la successione delle partizioni uniformi δ_n dell'intervallo $[0,1]$. Poichè in ogni intervallino di lunghezza maggiore di 0 cadono sempre necessariamente sia punti razionali che non, ne segue che su ogni intervallino della partizione uniforme l'estremo superiore della f è 1, mentre l'estremo inferiore è 0. Si ha quindi, per ogni $n \in \mathbb{N}$,

$$S_{\delta_n} = \sum_{k=1}^{n} 1 \cdot \frac{1}{n} = 1, \qquad s_{\delta_n} = \sum_{k=1}^{n} 0 \cdot \frac{1}{n} = 0.$$

Segue quindi dal Corollario 6.8 che la funzione di Dirichlet non è integrabile.

6.5 Proprietà delle funzioni integrabili

In questa sezione mostriamo come l'integrale goda di molte importanti proprietà. Cominciamo con il seguente risultato che mostra come l'integrazione sia un'operazione lineare.

Proposizione 6.9 *Siano $f : [a,b] \to \mathbb{R}$ e $g : [a,b] \to \mathbb{R}$ due funzioni integrabili su $[a,b]$ e sia $\lambda \in \mathbb{R}$. Allora:*

(i) $f+g$ è integrabile su $[a,b]$ e si ha

$$\int_a^b (f+g)(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x\,.$$

(ii) λf è integrabile su $[a,b]$ e si ha

$$\int_a^b \lambda f(x)\,\mathrm{d}x = \lambda \int_a^b f(x)\,\mathrm{d}x\,.$$

Dimostrazione Dimostreremo (i) lasciando (ii) allo studente per esercizio. Faremo uso dei precedenti criteri di integrabilità. Notiamo innanzitutto che fissata una partizione

$$\delta = \{x_1 = a < x_2 < x_3 < \cdots < x_{n+1} = b\}$$

sfruttando il fatto che vale sempre (si pensi al perchè)

$$\begin{array}{l}\sup\limits_{x\in[x_k,x_{k+1}]}[f(x)+g(x)] \le \sup\limits_{x\in[x_k,x_{k+1}]} f(x) + \sup\limits_{x\in[x_k,x_{k+1}]} f(x)\,,\\ \inf\limits_{x\in[x_k,x_{k+1}]}[f(x)+g(x)] \ge \inf\limits_{x\in[x_k,x_{k+1}]} f(x) + \inf\limits_{x\in[x_k,x_{k+1}]} f(x)\end{array}$$

si ottiene

$$S_\delta(f+g) \le S_\delta(f) + S_\delta(g)\,, \qquad s_\delta(f+g) \ge s_\delta(f) + s_\delta(g)\,.$$

Consideriamo ora la successione di partizioni uniformi δ_n. Si ha che

$$s_{\delta_n}(f) + s_{\delta_n}(g) \le s_{\delta_n}(f+g) \le S_{\delta_n}(f+g) \le S_{\delta_n}(f) + S_{\delta_n}(g)\,. \tag{6.6}$$

Passando ora al limite per $n \to +\infty$ si ha che, per il Teorema 6.7, le due successioni estreme di (6.6) convergono entrambe a

$$\int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x\,. \tag{6.7}$$

Quindi, per il Teorema del confronto tra successioni, anche $S_{\delta_n}(f+g)$ e $s_{\delta_n}(f+g)$ convergono a (6.7). Questo, in virtù del Corollario 6.8, mostra che $f+g$ è integrabile. Inoltre, sempre per il Corollario 6.8 segue che

$$\int_a^b (f+g)(x)\,\mathrm{d}x = \lim_{n\to+\infty} S_{\delta_n}(f+g) = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x\,.$$

La dimostrazione è quindi completa. ■

La seguente illustra invece le proprietà di monotonia dell'integrale.

Proposizione 6.10 *Siano $f : [a,b] \to \mathbb{R}$ e $g : [a,b] \to \mathbb{R}$ due funzioni integrabili su $[a,b]$ e tali che $f(x) \leq g(x)$ per ogni $x \in [a,b]$. Allora:*

(i)
$$\int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x\,.$$

(ii) $|f|$ è integrabile su $[a,b]$ e si ha
$$\left|\int_a^b f(x)\,\mathrm{d}x\right| \leq \int_a^b |f(x)|\,\mathrm{d}x\,.$$

Dimostrazione Per dimostrare (i) si noti intanto che, per la definizione stessa di integrale, se una funzione integrabile f è tale che $f(x) \geq 0$ per ogni x allora necessariamente $\int_a^b f(x)\,\mathrm{d}x \geq 0$ (e tale valore rappresenta proprio l'area sottesa dal grafico). Consideriamo allora la differenza $g - f$. Per ipotesi essa assume valori non negativi e utilizzando la linearità espressa dalla Proposizione 6.9 si ottiene che

$$0 \leq \int_a^b (f-g)(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x - \int_a^b g(x)\,\mathrm{d}x$$

che implica la tesi.

Per quanto concerne (ii) la questione più delicata è far vedere che $|f(x)|$ è ancora integrabile. Questo si fa utilizzando il criterio di integrabilità illustrato nel Teorema 6.6, sfruttando il fatto che se I è un qualunque intervallo si ha che

$$\sup_{x\in I} |f(x)| - \inf_{x\in I} |f(x)| \leq \sup_{x\in I} f(x) - \inf_{x\in I} f(x)\,.$$

Per dimostrare poi la diseguaglianza, basta osservare che essa è equivalente a

$$-\int_a^b |f(x)|\,\mathrm{d}x \leq \int_a^b f(x)\,\mathrm{d}x \leq \int_a^b |f(x)|\,\mathrm{d}x$$

e questo segue da risultato già dimostrato (i) poichè si ha

$$-|f(x)| \leq f(x) \leq |f(x)|$$

per ogni $x \in [a,b]$. ■

Presentiamo infine un'ulteriore proprietà dell'integrale: l'additività rispetto al dominio di integrazione.

Proposizione 6.11 *Sia $f : [a,b] \to \mathbb{R}$ una funzione. Sia $c \in]a,b[$ tale che f sia integrabile su $[a,c]$ e, rispettivamente, su $[c,b]$. Allora f è anche integrabile su tutto quanto $[a,b]$ e si ha*

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x \tag{6.8}$$

La dimostrazione è lasciata per esrcizio. Anche in questo caso si tratta di un'applicazione abbastanza immediata dei criterio di integrabilità: il trucco è lavorare con partizioni che contengano il punto c .... Si noti inoltre come, nel caso in cui f prenda valori non negativi, questo risultato esprima la proprietà di additività dell'area.

■

Per motivi che saranno chiari in seguito, risulta utile definire l'integrale anche su intervalli invertiti. Precisamente, si pone convenzionalmente, se $a > b$ e se f è integrabile su $[b,a]$:

$$\int_a^b f(x)\,\mathrm{d}x = -\int_b^a f(x)\,\mathrm{d}x\,.$$

Si pone anche

$$\int_a^a f(x)\,\mathrm{d}x = 0$$

qualunque sia il punto a. Con queste estensioni è possibile vedere che la Proposizione 6.11 continua a valere anche nel caso in cui il punto c stia fuori dell'intervallo $[a,b]$.

Osservazione 1. La Proposizione 6.11 ammette una sorta di formulazione alternativa: si può in effetti dimostrare che se $f : [a,b] \to \mathbb{R}$ è una funzione integrabile e $c \in [a,b]$, allora f è anche integrabile sui sottointervalli $[a,c]$ e $[c,d]$ e continua naturalmente a valere la formula (6.8).

Osservazione 2. Attraverso la Proposizione 6.11 possiamo fornire una interpretazione geometrica dell'integrale definito per funzioni che cambiano di segno. Supponiamo ad esempio di avere una funzione $f : [a,b] \to \mathbb{R}$ integrabile. Supponiamo che esistano due punti $c < d \in]a,b[$ con $f(x) \geq 0$ su $[a,c]$ e su $[d,b]$ e $f(x) \leq 0$ su $[c,d]$ come in Figura 6.4.

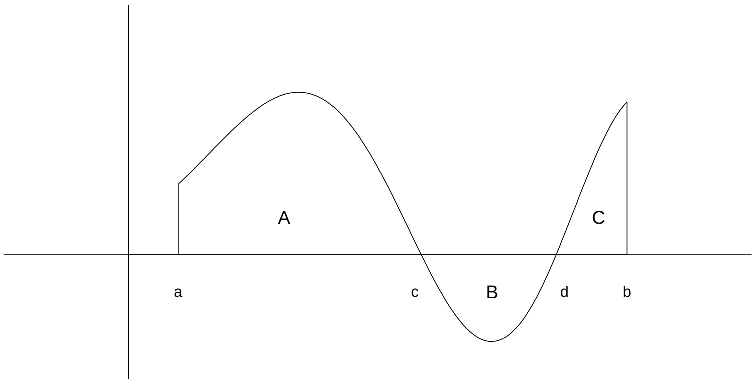


Figura 6.4:

Possiamo scrivere

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^d f(x)\,\mathrm{d}x + \int_d^b f(x)\,\mathrm{d}x\,.$$

$\int_a^c f(x)\,\mathrm{d}x$ e $\int_d^b f(x)\,\mathrm{d}x$ rappresentano le aree delle due parti sopra l'asse x, A e C rispettivamente, mentre $\int_c^d f(x)\,\mathrm{d}x$ rappresenta l'area della parte B con il segno $-$ davanti. Quindi $\int_a^b f(x)\,\mathrm{d}x$ rappresenta una sorta di somma algebrica di aree, considerando positive quelle sopra l'asse x, negative quelle sotto.

6.6 L'integrale delle funzioni continue

Possiamo ora presentare un'ampia classe di funzioni integrabili.

Teorema 6.12 *Sia $f : [a,b] \to \mathbb{R}$ una funzione continua. Allora f è integrabile su $[a,b]$.*

Dimostrazione Ci limiteremo a dimostrare il teorema nell'ipotesi più forte che f sia lipschitziana (concetto introdotto nel Corollario 5.14). Il caso generale richiede la conoscenza di proprietà delle funzioni continue che in questo corso non sono state studiate. Supponiamo dunque che esista $M > 0$ tale che

$$|f(x_1) - f(x_2)| \leq M|x_1 - x_2| \quad \forall x_1, x_2 \in [a,b]\,.$$

Faremo uso del Teorema 6.6. Fissiamo dunque $\epsilon > 0$ e consideriamo una qualunque partizione $\delta \in \Delta_{[a,b]}$ avente parametro di finezza $r(\delta) < \epsilon/[M(b-a)]$. Consideriamo la differenza tra la somma superiore e quella inferiore relative alla partizione δ:

$$S_\delta - s_\delta = \sum_{k=1}^{n} (L_k - l_k)(x_{k+1} - x_k)\,, \tag{6.9}$$

dove

$$L_k = \sup_{x \in I_k} f(x)\,, \qquad l_k = \inf_{x \in I_k} f(x)$$

e dove, come prima, $I_k = [x_k, x_{k+1}]$. Poichè f è continua su $[a,b]$, segue dal Teorema 4.32 di Weierstrass che f ammette massimo e minimo assoluti su ogni intervallo I_k. In altri termini esistono punti $\alpha_k, \beta_k \in I_k$ tali che $L_k = f(\alpha_k)$ e $l_k = f(\beta_k)$ per ogni k. Sostituendo nella (6.9), utilizzando la lipschitzianità e il fatto che, per la scelta della partizione, $|\alpha_k - \beta_k| \leq \epsilon/[M(b-a)]$ si ha che

$$\begin{aligned} S_\delta - s_\delta &= \sum_{k=1}^{n} (f(\alpha_k) - f(\beta_k))(x_{k+1} - x_k) \\ &\leq \sum_{k=1}^{n} M|\alpha_k - \beta_k|(x_{k+1} - x_k) \\ &\leq \sum_{k=1}^{n} M \tfrac{\epsilon}{M(b-a)} (x_{k+1} - x_k) \\ &= \tfrac{\epsilon}{b-a} \sum_{k=1}^{n} (x_{k+1} - x_k) \\ &= \tfrac{\epsilon}{b-a} (b-a) = \epsilon\,. \end{aligned} \tag{6.10}$$

Questo, per il Teorema 6.6, dimostra la tesi. ■

Teorema 6.13 (della media integrale). *Sia $f : [a,b] \to \mathbb{R}$ una funzione continua. Allora esiste $c \in [a,b]$ tale che*

$$\int_a^b f(x)\,\mathrm{d}x = (b-a) f(c)\,.$$

Dimostrazione Segue dalla Proposizione 6.3 che

$$\inf_{x \in [a,b]} f(x) \leq \frac{1}{b-a} \int_a^b f(x)\,\mathrm{d}x \leq \sup_{x \in [a,b]} f(x)\,.$$

Poichè f è una funzione continua su $[a,b]$ segue dal Corollario 4.30 che essa assume ogni valore compreso tra il suo minimo ed il suo massimo. Dunque, esiste $c \in [a,b]$ tale che

$$\frac{1}{b-a} \int_a^b f(x)\,\mathrm{d}x = f(c)\,.$$

Questo implica la tesi. ■

Osservazione 1. L'interpretazione geometrica del Teorema 6.13 della media integrale è la seguente. Supponiamo per semplicità $f(x) \geq 0$. Esso dice che c'è un punto $c \in [a, b]$ tale che l'area del rettangolo di base $[a, b]$ e altezza $f(c)$ è esattamente la stessa dell'area sottesa al grafico di f sull'intervallo $[a, b]$. Tale valore $f(c)$ è dunque da interpretarsi come l'altezza media del grafico di f rispetto all'asse x.

Osservazione 2. Si noti che il Teorema della media integrale vale anche nel caso in cui $a > b$.

Le funzioni continue non sono certo le uniche funzioni integrabili. Introduciamone brevemente altre classi. Intanto premettiamo la seguente

Proposizione 6.14 *Siano $f, g : [a, b] \to \mathbb{R}$ due funzioni tali che*

$$f(x) = g(x)$$

tranne che per un insieme finito di punti di $[a, b]$. Allora f è integrabile su $[a, b]$ se e soltanto se lo è g, e in caso affermativo si ha che

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b g(x)\,\mathrm{d}x\,.$$

Dimostrazione Diamo solo un'idea della dimostrazione. L'osservazione fondamentale da fare è che le somme superiori (risp. inferiori) di f e g relativamente alle partizioni uniformi δ_n, differiscono tra di loro per un numero finito (soltanto dipendente dal numero di punti nei quali f e g differiscono) di addendi per cui la loro differenza è infinitesima per $n \to +\infty$. Il risultato dunque segue dal Corollario 6.7. ■

Dal risultato precedente e dal Teorema 6.12 segue quindi che ogni funzione che presenta soltanto un numero finito di discontinuità eliminabili ed è per il resto continua, è integrabile. Anche le discontinuità di prima specie non creano problemi all'integrabilità. Consideriamo in effetti una funzione $f : [a, b] \to \mathbb{R}$ continua ovunque tranne in un punto $c \in]a, b[$ dove presenta un salto. Allora la funzione è integrabile su $[a, c]$ e su $[c, b]$ per il Teorema 6.12 e per la Proposizione 6.14. Per l'Osservazione 1 dopo la Proposizione 6.11, segue allora che f è integrabile su tutto quanto $[a, b]$. Quindi funzioni che presentino un numero finito di discontinuità eliminabili e/o di prima specie sono integrabili. Citiamo infine un risultato di tipo diverso del quale non diamo alcun cenno dimostrativo.

Proposizione 6.15 *Sia $f : [a,b] \to \mathbb{R}$ una funzione limitata e monotona. Allora f è integrabile su $[a,b]$*

6.7 IL teorema fondamentale del calcolo integrale

Arriviamo ora al punto cruciale che collega il calcolo delle aree al problema della ricerca delle primitive. Questo collegamento è un risultato di straordinaria importanza che rende ancora più importante il calcolo differenziale studiato nel primo corso.

Supponiamo di avere una funzione continua $f : [a,b] \to \mathbb{R}$. Anzichè considerare il solo integrale

$$\int_a^b f(t)\,\mathrm{d}t\,,$$

consideriamo tutta la famiglia di integrali

$$F(x) = \int_a^x f(t)\,\mathrm{d}t$$

al variare di $x \in [a,b]$. Otteniamo così una funzione F detta *funzione integrale*. Possiamo ora enunciare il seguente fondamentale risultato

Teorema 6.16 *Sia $f : [a,b] \to \mathbb{R}$ una funzione continua. Allora la funzione integrale*

$$F(x) = \int_a^x f(t)\,\mathrm{d}t$$

è una primitiva di f su $[a,b]$, cioè essa è derivabile su $[a,b]$ e si ha

$$F'(x) = f(x), \quad \forall x \in [a,b]\,.$$

Dimostrazione Fissiamo un qualunque $x_0 \in [a,b]$ e consideriamo il rapporto incrementale della F relativo a questo punto. Sfruttando la Proposizione 6.11 e le considerazioni ad essa seguenti si ha che

$$\frac{F(x)-F(x_0)}{x-x_0} = \frac{\int_a^x f(t)\,\mathrm{d}t - \int_a^{x_0} f(t)\,\mathrm{d}t}{x-x_0} = \frac{\int_a^{x_0} f(t)\,\mathrm{d}t + \int_{x_0}^{x} f(t)\,\mathrm{d}t - \int_a^{x_0} f(t)\,\mathrm{d}t}{x-x_0} = \frac{\int_{x_0}^{x} f(t)\,\mathrm{d}t}{x-x_0}\,.$$

Segue ora dal Teorema 6.13 e dall'Osservazione 2 ad esso seguente che esiste c tra x e x_0 tale che

$$\frac{\int\limits_{x_0}^{x} f(t)\,\mathrm{d}t}{x - x_0} = f(c)\,.$$

Si ha quindi che

$$\frac{F(x) - F(x_0)}{x - x_0} = f(c)\,. \tag{6.11}$$

Vogliamo far vedere che il rapporto incrementale converge, per $x \to x_0$ a $f(x_0)$. Fissiamo dunque $\epsilon > 0$. In virtù della continuità di f si ha che esiste $\delta > 0$ tale che

$$|x - x_0| < \delta \;\Rightarrow\; |f(x) - f(x_0)| < \epsilon\,. \tag{6.12}$$

D'altra parte, se x è tale che $|x - x_0| < \delta$, poichè c sta tra x e x_0 si ha anche che $|c - x_0| < \delta$ e quindi per la (6.12) si ha che $|f(c) - f(x_0)| < \epsilon$. Usando infine la (6.11) si ha quindi che

$$\left|\frac{F(x) - F(x_0)}{x - x_0} - f(x_0)\right| < \epsilon\,.$$

Questo implica la tesi. ■

Come si può utilizzare questo risultato? Esso è utile per due motivi distinti. Da una parte ci dice che ogni funzione continua ammette primitive rispondendo così ad un quesito lasciato aperto nella prima sezione di questo capitolo. D'altra parte offre un modo per calcolare gli integrali definiti. In effetti si ha il seguente:

Corollario 6.17 *Sia $f : [a, b] \to \mathbb{R}$ una funzione continua e sia $G : [a, b] \to \mathbb{R}$ una primitiva qualunque di f. Allora, si ha*

$$\int\limits_{a}^{b} f(t)\,\mathrm{d}t = G(b) - G(a)\,. \tag{6.13}$$

Dimostrazione Poichè anche la funzione integrale

$$F(x) = \int\limits_{a}^{x} f(t)\,\mathrm{d}t$$

è una primitiva della f per il Teorema precedente, segue dal Teorema 6.1 che esiste una costante k tale che

$$\int\limits_{a}^{x} f(t)\,\mathrm{d}t = G(x) + k\,.$$

Valutando l'espressione sopra per $x = a$ si ottiene

$$0 = \int_a^a f(t)\,\mathrm{d}t = G(a) + k$$

da cui segue che, necessariamente, $k = -G(a)$. Dunque, sostituendo, si ha che

$$\int_a^x f(t)\,\mathrm{d}t = G(x) - G(a)\,.$$

Valutando ora per $x = b$ si ottiene la tesi. ■

Per ragioni di compattezza di scrittura in seguito useremo sovente la notazione

$$G(x)\Big|_a^b = G(b) - G(a)$$

6.8 Tecniche di integrazione

Motivati dai risultati della sezione precedente che collegano il calcolo integrale alla ricerca delle primitive, da ora in poi useremo la notazione

$$\int f(x)\,\mathrm{d}x$$

per indicare l'insieme delle primitive di una funzione (in genere continua) $f(x)$. Tale insieme verrà detto *integrale indefinito* di $f(x)$. Se $F(x)$ è una primitiva di $f(x)$, scriveremo dunque,

$$\int f(x)\,\mathrm{d}x = F(x) + k$$

ad indicare proprio che tutte le primitive di $f(x)$ si ottengono da $F(x)$ sommando costanti. Se non si specifica su quale intervallo stiamo considerando $f(x)$ e le sue primitive, si supporrà in seguito di essere su tutto $\mathbb{R}$. Nel caso in cui ci restringessimo a particolari intervalli, lo diremo esplicitamente. Molti integrali indefiniti si calcolano immediatamente ricordando semplicemente le regole di derivazione. Si ha ad esempio:

$\int x^n \,\mathrm{d}x = \frac{x^{n+1}}{n+1} + k \;\; (n \in \mathbb{N})$ $\qquad$ $\int x^\alpha \,\mathrm{d}x = \frac{x^{\alpha+1}}{\alpha+1} + k \;\; (\alpha \neq -1, x > 0)$

$\int x^{-1} \,\mathrm{d}x = \ln x + k \;\; (x > 0)$ $\qquad$ $\int e^x \,\mathrm{d}x = e^x + k$

$\int \cos x \,\mathrm{d}x = \sin x + k$ $\qquad$ $\int \sin x \,\mathrm{d}x = -\cos x + k$

$\int \frac{1}{1+x^2} \,\mathrm{d}x = \arctan x + k$ $\qquad$ $\int \frac{1}{\sqrt{1-x^2}} \,\mathrm{d}x = \arcsin x + k \;\; (|x| < 1)$

Osservazione. Se consideriamo la funzione $1/x$ sulla semiretta negativa $]-\infty, 0[$ è immediato verificare che il suo integrale è dato da

$$\int \frac{1}{x} \,\mathrm{d}x = \ln(-x) + k \quad x \in]-\infty, 0[\,.$$

Si noti che si può quindi dire che la funzione $\ln |x|$ è una primitiva della funzione $1/x$ su $\mathbb{R} \setminus \{0\}$. Si noti tuttavia che sarebbe scorretto scrivere

$$\int \frac{1}{x} \,\mathrm{d}x = \ln |x| + k \quad x \in \mathbb{R} \setminus \{0\}$$

in quanto ci sono altre primitive che non differiscono da $\ln |x|$ per una costante (si veda a questo proposito l'Osservazione che segue il Teorema 6.1.

L'integrale indefinito gode di tutta una serie di proprietà. Essa è, parlando un po' impropriamente, l'operazione inversa della derivazione nel senso che

$$\frac{d}{\mathrm{d}x} \left(\int f(x) \,\mathrm{d}x \right) = f(x) \tag{6.14}$$

o anche, se f è una funzione di classe C^1

$$\int \frac{d}{\mathrm{d}x} f(x) \,\mathrm{d}x = f(x) + k\,. \tag{6.15}$$

Dalla definizione stessa di integrale definito segue poi che ad ogni regola di derivazione corrisponde una regola di integrazione.

Proprietà di linearità: Sfruttando la linearità della derivazione si ha che

$$\int (\lambda f(x) + \mu g(x)) \,\mathrm{d}x = \lambda \int f(x) \,\mathrm{d}x + \mu \int g(x) \,\mathrm{d}x\,. \tag{6.16}$$

Integrazione per parti: Consideriamo ora la regola di derivazione del prodotto:
$$\frac{d}{\mathrm{d}x}[f(x)g(x)] = f(x)g'(x) + f'(x)g(x)\,.$$
Considerandone gli integrali indefiniti e utilizzando (6.15) e (6.16) si ottiene

$$f(x)g(x) + k = \int f(x)g'(x)\,\mathrm{d}x + \int f'(x)g(x)\,\mathrm{d}x$$

che può anche essere scritta come

$$\int f(x)g'(x)\,\mathrm{d}x = f(x)g(x) - \int f'(x)g(x)\,\mathrm{d}x \tag{6.17}$$

(essendoci ora integrali indefiniti ad entrambi i membri abbiamo omesso la costante k). La (6.17) è nota come *regola di integrazione per parti* ed è molto utile per calcolare integrali.

Integrazione per sostituzione: Consideriamo ora una funzione continua $f(x)$ e una sua primitiva qualsiasi $F(x)$. Supponiamo ora di introdurre ua nuova variabile t legata dalla x da una relazione $x = \phi(t)$ con ϕ di classe C^1. Dalla regola di derivazione della funzione composta si ha che

$$\frac{d}{\mathrm{d}t}F(\phi(t)) = F'(\phi(t)) \cdot \phi'(t) = f(\phi(t)) \cdot \phi'(t)$$

il che significa che $F(\phi(t))$ è una primitiva della funzione $f(\phi(t))\phi'(t)$. Dunque possiamo scrivere

$$\int f(x)\,\mathrm{d}x = \int f(\phi(t))\phi'(t)\,\mathrm{d}t \qquad [x = \phi(t)]\,. \tag{6.18}$$

Questa è detta *regola di integrazione per sostituzione* e può essere utilizzata in due modi distinti. Da una parte se si riconosce che la funzione che vogliamo integrare è del tipo $f(\phi(t))\phi'(t)$ possiamo integrare $f(x)$ (se ci riusciamo) e poi semplicemente sostituire alla x la $\phi(t)$. Talvolta può invece essere utile ragionare nel senso opposto: non sappiamo integrare $f(x)$, consideriamo la sostituzione $x = \phi(t)$, facciamo l'integrale a secondo membro di (6.18) e poi torniamo indietro sostituendo indietro $t = \phi^{-1}(x)$, (questo ovviamente a patto di considerare sostituzioni ϕ che siano invertibili).

Osservazione. La formula (6.18) motiva il simbolismo adottato per l'integrale: se $x = \phi(t)$ si ha $\mathrm{d}x/\,\mathrm{d}t = \phi'(t)$ che formalmente può anche essere scritto come

$\mathrm{d}x = \phi'(t)\,\mathrm{d}t$ che fornisce esattamente il modo per passare dal primo integrale al secondo.

Utilizzando la formula (6.13), le precedenti regole di integrazione indefinita diventano subito anche regole di integrazione definita. Quella per parti può scriversi come:

$$\int_a^b f(x)g'(x)\,\mathrm{d}x = [f(x)g(x)]\Big|_a^b - \int_a^b f'(x)g(x)\,\mathrm{d}x\,. \tag{6.19}$$

Per la regola di integrazione per sostituzione è invece necessaria qualche ulteriore ipotesi. Il risultato è il seguente:

Teorema 6.18 (del cambiamento di variabile) *Supponiamo che $f : [a,b] \to \mathbb{R}$ sia continua e che $\phi : [\alpha,\beta] \to [a,b]$ sia di classe C^1 e invertibile. Allora si ha*

$$\int_a^b f(x)\,\mathrm{d}x = \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f(\phi(t))\phi'(t)\,\mathrm{d}t\,. \tag{6.20}$$

Dimostrazione Conseguenza immediata delle formule (6.18) e (6.13). ■

Mostriamo ora con una serie di esempi e di esercizi come si possono applicare le varie tecniche di integrazione descritte.

Cominciamo col mostrare alcune semplici applicazioni della tecnica di integrazione per sostituzione.

Esempio 103 Sia $f(x) = e^{\alpha x}$ dove $\alpha \neq 0$. La sostituzione $\alpha x = t$ corrispondente a $x = t/\alpha$ trasforma, utilizzando la (6.18), l'integrale della $f(x)$ in

$$\int e^t \frac{1}{\alpha}\,\mathrm{d}t = \frac{1}{\alpha}e^t + k\,.$$

Si ha quindi,

$$\int e^{\alpha x}\,\mathrm{d}x = \frac{1}{\alpha}e^{\alpha x} + k\,.$$

Esercizio 6.1 Calcolare gli integrali indefiniti

$$\int \sin\beta x\,\mathrm{d}x\,, \qquad \int \cos\beta x\,\mathrm{d}x\,.$$

R: $-\frac{\cos(\beta x)}{\beta} + k$, $\frac{\sin(\beta x)}{\beta} + k$.

Altri esempi di integrazione per sostituzione sono i seguenti:

Esempio 104 Sia $f(x) = \tan x$ definita su di un intervallo del tipo

$$]-\pi/2 + k\pi, \pi/2 + k\pi[$$

dove $k \in \mathbb{Z}$. Si ha

$$\int \tan x \, \mathrm{d}x = \int \frac{\sin x}{\cos x} \, \mathrm{d}x = -\int \frac{-\sin x}{\cos x} \, \mathrm{d}x \, .$$

Poichè $D \cos x = -\sin x$, questo è un integrale di una funzione del tipo $f(\phi(x))\phi'(x)$ con $f(t) = 1/t$ e $\phi(x) = \cos x$. Utilizzando quindi la (6.18) e la forma dell'integrale di $1/x$ si ottiene, considerando $t = \cos x$,

$$\int \frac{-\sin x}{\cos x} \, \mathrm{d}x = \int \frac{D \cos x}{\cos x} \, \mathrm{d}x = \int \frac{1}{t} \, \mathrm{d}t = \ln |t| + k = \ln |\cos x| + k$$

e quindi

$$\int \tan x \, \mathrm{d}x = -\ln |\cos x| + k \, .$$

Esempio 105 Consideriamo la funzione $f(x) = x^{-1} \ln x$ definita su $]0, +\infty[$. Con le stesse considerazioni del caso precedente, notando che $D \ln x = x^{-1}$ e utilizzando la (6.18), si ha (considerando $t = \ln x$)

$$\int x^{-1} \ln x \, \mathrm{d}x = \int \ln x (D \ln x) \, \mathrm{d}x = \int t \, \mathrm{d}t = \frac{t^2}{2} + k = \frac{(\ln x)^2}{2} + k \, .$$

Esercizio 6.2 Calcolare gli integrali indefiniti

$$\int \frac{\sin x}{1 + \cos^2 x} \, \mathrm{d}x \, , \qquad \int \frac{x}{1 + x^2} \, \mathrm{d}x \, , \qquad \int \frac{1}{1 + 2x} \, \mathrm{d}x \, .$$

R. $-\arctan(\cos x) + k$, $\frac{\ln(1+x^2)}{2} + k$, $\frac{\ln |1+2x|}{2} + k$ $(x \neq -\frac{1}{2})$.

Passiamo ora ad illustrare vari esempi di integrazione per parti:

Esempio 106 Consideriamo la funzione xe^x. Applichiamo la (6.17) con $f(x) = x$ e $g(x) = e^x$. Poichè $f'(x) = 1$ e $g'(x) = e^x$ si ha che

$$\int xe^x \, \mathrm{d}x = xe^x - \int e^x \, \mathrm{d}x = xe^x - e^x + k \, .$$

Esercizio 6.3 Calcolare gli integrali indefiniti

$$\int x\sin x\,\mathrm{d}x\,,\qquad \int x\cos x\,\mathrm{d}x\,,\qquad \int x^\alpha \ln x\,\mathrm{d}x\,,\quad (\alpha\neq -1,\ x\in]0,+\infty[)\,.$$

R. $-x\cos x+\sin x+k$, $x\sin x+\cos x+k$, $\frac{x^{\alpha+1}}{\alpha+1}\ln x-\frac{x^{\alpha+1}}{(\alpha+1)^2}+k$.

Esempio 107 Consideriamo la funzione $\arcsin x$ su $]-1,1[$. Applichiamo la (6.17) con $f(x)=\arcsin x$ e $g(x)=x$ si ha

$$\int \arcsin x\,\mathrm{d}x = x\arcsin x-\int \frac{x}{\sqrt{1-x^2}}\,\mathrm{d}x\,.$$

Sostituendo ora $t=\phi(x)=1-x^2$ nell'ultimo integrale e sfruttando la (6.18), si ottiene

$$\int \frac{x}{\sqrt{1-x^2}}\,\mathrm{d}x=-\frac{1}{2}\int \frac{\phi'(x)}{\sqrt{\phi(x)}}\,\mathrm{d}x=-\frac{1}{2}\int\frac{\mathrm{d}t}{\sqrt{t}}=-\sqrt{t}+k=-\sqrt{1-x^2}+k\,.$$

Si ha dunque

$$\int \arcsin x\,\mathrm{d}x = x\arcsin x+\sqrt{1-x^2}+k\,.$$

Esercizio 6.4 Calcolare l'integrale indefinito

$$\int \arctan x\,\mathrm{d}x\,.$$

R. $x\arctan x-\frac{1}{2}\log(1+x^2)+k$.

Talvolta è necessaria l'applicazione di più integrazioni per parti come illustra il seguente:

Esempio 108 Consideriamo la funzione $e^{\alpha x}\sin\beta x$ con $\beta\neq 0$. Applichiamo due volte la (6.17):

$$\begin{aligned}\textstyle\int e^{\alpha x}\sin\beta x\,\mathrm{d}x &= -\tfrac{1}{\beta}e^{\alpha x}\cos\beta x+\tfrac{\alpha}{\beta}\textstyle\int e^{\alpha x}\cos\beta x\,\mathrm{d}x\\ &= -\tfrac{1}{\beta}e^{\alpha x}\cos\beta x+\tfrac{\alpha}{\beta^2}e^{\alpha x}\sin\beta x-\tfrac{\alpha^2}{\beta^2}\textstyle\int e^{\alpha x}\sin\beta x\,\mathrm{d}x\,.\end{aligned}\tag{6.21}$$

Abbiamo così ottenuto un integrale uguale a quello di partenza. Si potrebbe così pensare che la doppia integrazione per parti non abbia sortito alcun effetto. Tuttavia se nella relazione (6.21) portiamo al primo membro l'integrale che si trova al secondo membro, un semplice passaggio algebrico mostra che

$$\int e^{\alpha x}\sin\beta x\,\mathrm{d}x=\frac{1}{\alpha^2+\beta^2}e^{\alpha x}(\alpha\sin\beta x-\beta\cos\beta x)+k\,.$$

Esercizio 6.5 Calcolare gli integrali indefiniti

$$\int \sin nx \sin mx \, \mathrm{d}x \,, \qquad \int \cos nx \cos mx \, \mathrm{d}x \,, \qquad \int \sin nx \cos mx \, \mathrm{d}x \,.$$

dove m e n sono due interi positivi diversi tra loro.

R.
$\dfrac{n \cos nx \sin mx - m \sin nx \cos mx}{m^2 - n^2} + k$, $\dfrac{n \sin nx \cos mx - m \cos nx \sin mx}{n^2 - m^2} + k$,
$\dfrac{m \sin nx \sin mx + n \cos nx \cos mx}{m^2 - n^2} + k$.

Si noti come nell'esercizio precedente abbiamo supposto $n \neq m$. Il caso $n = m$ in effetti si risolve in altro modo, come mostra il seguente:

Esempio 109 Consideriamo la funzione $\sin^2 x$. Si può scrivere, usando le formule trigonometriche di duplicazione, $\sin^2 x = (1 - \cos 2x)/2$. Dunque si ha

$$\int \sin^2 x \, \mathrm{d}x = \frac{1}{2}\left(x - \frac{\sin 2x}{2}\right) = \frac{1}{2}(x - \sin x \cos x) \,. \tag{6.22}$$

Esercizio 6.6 Calcolare l'integrale indefinito

$$\int \cos^2 x \, \mathrm{d}x \,.$$

R. $\frac{1}{2}(x + \sin x \cos x) + k$.

Si possono calcolare in realtà tutti gli integrali del tipo

$$\int \sin^n x \, \mathrm{d}x \,, \qquad \int \cos^n x \, \mathrm{d}x \,.$$

Vi ritorneremo nella sezione sull'integrazione di funzioni complesse.

Naturalmente non va dimenticata la proprietà di linearità degli integrali che permette di integrare tutte quelle funzioni che sono esprimibili come combinazione lineare di funzioni delle quali sappiamo già calcolare i rispettivi integrali. Vediamo un semplice esempio:

Esempio 110 Consideriamo la funzione

$$5xe^x + \sin 7x + 4x^2 \,.$$

Utilizzando la proprietà di linearità dell'integrale, l'Esempio 103 e l'Esercizio 6.1 si ottiene:

$$\int (5xe^x + \sin 7x + 4x^2) \, \mathrm{d}x = 5xe^x - 5e^x - \frac{1}{7}\cos 7x + \frac{4}{3}x^3 + k \,.$$

Esercizio 6.7 Calcolare gli integrali indefiniti

$$\int (3x^4 - xe^{-x} + x\sin x)\,\mathrm{d}x\,, \qquad \int x(\cos x + \sin x)\,\mathrm{d}x\,.$$

R. $\frac{3}{4}x^3 + xe^{-x} + e^{-x} - x\cos x + \sin x + k$, $(1+x)\sin x + (1-x)\cos x + k$.

Esercizio 6.8 Calcolare gli integrali indefiniti

$$\int \sinh x\,\mathrm{d}x\,, \qquad \int \cosh x\,\mathrm{d}x$$

R. $\cosh x + k$, $\sinh x + k$.

Non sempre può essere chiaro quale tecnica si debba usare per portare a termine un procedimento di integrazione e talvolta è necessario anche usarne più di una per arrivare in fondo. Presentiamo qui qualche esempio finale in cui qualche 'trucco' più o meno visibile è necessario per integrare.

Esempio 111 Consideriamo la funzione

$$\frac{1}{x^2+3x+2}\,.$$

Come possiamo integrarla? Si noti che se avessimo avuto invece

$$\frac{1}{x^2+2}\,,$$

non avremmo avuto alcun problema. In effetti,

$$\int \frac{1}{x^2+2}\,\mathrm{d}x = \frac{1}{2}\int \frac{1}{\frac{x^2}{2}+1}\,\mathrm{d}x$$

e con la sostituzione $x = \sqrt{2}t$ otterremmo

$$\int \frac{1}{\frac{x^2}{2}+1}\,\mathrm{d}x = \sqrt{2}\int \frac{1}{t^2+1}\,\mathrm{d}t = \sqrt{2}\arctan t + k = \sqrt{2}\arctan\left(\frac{x}{\sqrt{2}}\right) + k\,.$$

Quindi,

$$\int \frac{1}{x^2+2}\,\mathrm{d}x = \frac{1}{\sqrt{2}}\arctan\left(\frac{x}{\sqrt{2}}\right) + k\,.$$

Per integrare invece la nostra funzione originale, si noti che $x^2+3x+2 = (x+1)(x+2)$. Scriviamo

$$\frac{1}{x^2+3x+2} = \frac{A}{x+1} + \frac{B}{x+2} \tag{6.23}$$

dove A e B sono due parametri da determinare. Imponendo che (6.23) sia un'identità, cioè sia vera per qualunque valore di x, si vede subito, facendo il calcolo a destra del numeratore che deve valere $1 = A(x+2) + B(x+1)$ come identità, che implica $A = 1$ e $B = -1$. Si può dunque scrivere

$$\frac{1}{x^2+3x+2} = \frac{1}{x+1} - \frac{1}{x+2}\,.$$

Per linearità si ha dunque

$$\int \frac{1}{x^2+3x+2}\,\mathrm{d}x = \ln|x+1| - \ln|x+2| + k\,.$$

La decomposizione (6.23) è nota come decomposizione in fratti semplici e permette di calcolare gli integrali di ogni funzione razionale del tipo

$$\frac{1}{x^2+ax+b}\,.$$

ogni volta che il denominatore abbia due radici reali distinte, cioè valga

$$x^2 + ax + b = (x - x_1)(x - x_2)$$

con $x_1 \neq x_2$. Basta in effetti ripetere gli argomenti dell'esempio precedente. La decomposizione in fratti semplici ha in realtà anche molte altre applicazioni e la incontreremo di nuovo in futuro.

Esercizio 6.9 Calcolare l'integrale

$$\int \frac{1}{x^2+x-2}\,\mathrm{d}x\,.$$

R. $\frac{1}{3}[\ln|x-1| - \ln|x+2|]$.

Esempio 112 Consideriamo ora la funzione

$$\frac{1}{x^2+x+1}\,.$$

Si noti che il polinomio $x^2 + x + 1$ non ha zeri reali e quindi la tecnica precedente della decomposizione in fratti semplici non si può applicare in questo caso. Si usa invece un'altra tecnica anch'essa molto importante e che ha molte applicazioni nota come 'il completamento del quadrato'. Funziona in questo modo: chiediamoci quale sarebbe stata la costante giusta anzichè 1 da avere nel nostro polinomio per avere un quadrato perfetto. La risposta è $1/4$. In effetti, come è facile verificare, $x^2 + x + 1/4 = (x + 1/2)^2$. Possiamo dunque scrivere

$$x^2 + x + 1 = x^2 + x + \frac{1}{4} + \frac{3}{4} = \left(x + \frac{1}{2}\right)^2 + \frac{3}{4} = \frac{3}{4}\left[\left(\frac{\sqrt{3}}{2}\left(x + \frac{1}{2}\right)\right)^2 + 1\right].$$

Si ha dunque,

$$\int \frac{1}{x^2 + x + 1}\, dx = \frac{4}{3}\int \frac{1}{\left(\frac{2}{\sqrt{3}}\left(x + \frac{1}{2}\right)\right)^2 + 1}\, dx$$

Si procede a questo punto per sostituzione, ponendo

$$t = \frac{2}{\sqrt{3}}\left(x + \frac{1}{2}\right)$$

ottenendo così

$$\int \frac{1}{\left(\frac{2}{\sqrt{3}}\left(x + \frac{1}{2}\right)\right)^2 + 1}\, dx = \frac{\sqrt{3}}{2}\int \frac{1}{t^2 + 1}\, dt = \frac{\sqrt{3}}{2}\arctan t + k = \frac{\sqrt{3}}{2}\arctan\left(\frac{2}{\sqrt{3}}\left(x + \frac{1}{2}\right)\right) + k$$

Otteniamo così alla fine

$$\int \frac{1}{x^2 + x + 1}\, dx = \frac{2}{\sqrt{3}}\arctan\left(\frac{2}{\sqrt{3}}\left(x + \frac{1}{2}\right)\right) + k$$

Come nel caso della decomposizione in fratti semplici, anche questa tecnica del completamento del quadrato è del tutto generale e può essere utilizzata ogni volta che dobbiamo integrare una funzione del tipo

$$\frac{1}{x^2 + ax + b}.$$

nel caso in cui il polinomio a denominatore non abbia zeri reali.

Esercizio 6.10 Calcolare l'integrale

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

R. $\frac{\sqrt{2}}{\sqrt{3}} \arctan\left(\frac{sqrt2}{\sqrt{3}}\left(x - \frac{1}{2}\right)\right) + k$.

Esercizio 6.11 Calcolare gli integrali

$$\int \frac{x+7}{x^2+2x-3}\,\mathrm{d}x\,, \qquad \int \frac{x-2}{x^2-3x-4}\,\mathrm{d}x$$

Gli esercizi sugli integrali definiti normalmente si risolvono riducendoli al calcolo di integrali indefiniti. Vi sono tuttavia alcuni 'trucchi' che vale la pena di ricordare. Se $f : [-a,a] \to \mathbb{R}$ è una funzione dispari e integrabile, si ha che

$$\int_{-a}^{a} f(x)\,\mathrm{d}x = 0\,.$$

Se invece è una funzione pari, si ha che

$$\int_{-a}^{a} f(x)\,\mathrm{d}x = 2\int_{0}^{a} f(x)\,\mathrm{d}x\,.$$

Esercizio 6.12 Calcolare gli integrali definiti

$$\int_{-1}^{1} |x|\,\mathrm{d}x\,, \qquad \int_{-2}^{2} x^8 \sin x\,\mathrm{d}x\,, \qquad \int_{-1}^{3} x\sin|x|\,\mathrm{d}x\,.$$

R.1, 0, $\sin 2 + \cos 1 - -2\cos 2 - \sin 1$.

Naturalmente gli integrali definiti possono esere utilizzati per calcolare aree. Il prossimo esempio mostra come calcolare l'area del cerchio.

Esempio 113 (Area del cerchio) Consideriamo la funzione $\sqrt{r^2-x^2}$ con $x \in [-r,r]$. Con la sostituzione $x = \cos t$, equivalente a $t = \arccos x$ $(0 \le t \le \pi)$, si ottiene

$$\int \sqrt{r^2-x^2}\,\mathrm{d}x = \int \sqrt{r^2-r^2\cos^2 t}\,(-\sin t)\,\mathrm{d}t = -r^2\int \sin^2 t\,\mathrm{d}t\,.$$

Sfruttando (6.22), si ha dunque

$$\int \sqrt{r^2-x^2}\,\mathrm{d}x = -\frac{r^2}{2}(t-\sin t\cos t)+k = \frac{r^2}{2}\left[x\sqrt{1-x^2}-\arccos x\right]+k\,. \tag{6.24}$$

Si può ora calcolare l'area S_r del cerchio di raggio r. Si ha in effetti

$$\begin{aligned} S_r &= 2\int_{-1}^{1} \sqrt{r^2-x^2}\,\mathrm{d}x = 2\tfrac{r^2}{2}[x\sqrt{1-x^2}-\arccos x]_{-1}^{1} \\ &= r^2[-\arccos 1 + \arccos(-1)] = \pi r^2\,. \end{aligned}$$

Il calcolo delle aree dei domini normali si riporta sempre ad un calcolo di integrali definiti. In effetti se abbiamo che

$$A = \{(x, y) \in \mathbb{R} \times \mathbb{R} \,|\, a \le x \le b,\ f_1(x) \le y \le f_2(x)\}$$

dove $f_1 : [a, b] \to \mathbb{R}$ e $f_2 : [a, b] \to \mathbb{R}$ sono due funzioni continue tali che $f_1(x) \le f_2(x)$ per ogni $x \in [a, b]$, si ha che l'area di A, $S(A)$ si calcola come

$$S(A) = \int_a^b (f_2(x) - f_1(x))\,\mathrm{d}x\,. \tag{6.25}$$

Il motivo per cui la (6.25) vale è il seguente: se le due funzioni sono non negative, risulta chiaro che essendo A la differenza tra la regione sottesa da f_2 e quella sottesa da f_1, si deve avere

$$S(A) = \int_a^b f_2(x)\,\mathrm{d}x - \int_a^b f_2(x)\,\mathrm{d}x = \int_a^b (f_2(x) - f_1(x))\,\mathrm{d}x\,.$$

In generale, esiste sempre $L \in \mathbb{R}$ tale che $f_1(x) + L$ e quindi anche $f_2(x) + L$ è sempre non negativa. Poichè le traslazioni non modificano le aree, lavorando con le due funzioni traslate si ottiene nuovamente la tesi.

Esercizio 6.13 Calcolare l'area del dominio normale

$$A = \{(x, y) \in \mathbb{R} \times \mathbb{R} \,|\, \pi/4 \le x \le \pi,\ \cos x \le y \le \sin x\}\,.$$

R. $1 + \sqrt{2}$.

6.9 Integrazione per serie

Ci sono funzioni continue per le quali non è possibile determinarne esplicitamente una primitiva. In questi casi le tecniche di integrazione illustrate nella sezione precedente sono inutilizzabili. Un esempio di funzioni di questo tipo è e^{-x^2}. Si vedrà in seguito, in particolar modo nei corsi di Probabilità, che tale funzione riveste tuttavia un'importanza fondamentale per la quale è importante riuscire a calcolare, magari in modo approssimato, integrali definiti del tipo

$$\int_0^b e^{-x^2}\,\mathrm{d}x$$

(si noti che per la parità della funzione, ogni integrale definito della funzione assegnata si riporta al calcolo di questi sopra). Come possiamo fare senza primitive a disposizione?

Ci vengono in aiuto, come in altri problemi di approssimazione (si pensi al calcolo approssimato di funzioni trascendenti), gli sviluppi di Taylor. In effetti se vogliamo calcolare $\int_a^b f(x)\,\mathrm{d}x$ e conosciamo lo sviluppo di Taylor di f con il resto di Lagrange in a fino ad un certo ordine

$$f(x) = \sum_{k=0}^{N-1} f^{(k)}(a)\frac{(x-a)^k}{k!} + f^{(N)}(c)\frac{(x-a)^N}{N!},$$

possiamo pensare di approssimare l'integrale di f con l'integrale del polinomio di Taylor

$$\int_a^b \left(\sum_{k=0}^{N-1} f^{(k)}(a)\frac{(x-a)^k}{k!}\right) \mathrm{d}x = \sum_{k=0}^{N-1} f^{(k)}(a)\frac{(b-a)^{k+1}}{(k+1)!}.$$

L'errore commesso è dato da

$$\int_a^b f^{(N)}(c)\frac{(x-a)^N}{N!}\,\mathrm{d}x.$$

Si presti attenzione al fatto che c dipende da x nella formula di Taylor quindi non è affatto da pensarsi come una costante. Per stimare l'errore commesso è dunque necessario avere delle stime di $f^{(N)}(x)$ per $x \in [a,b]$. Per illustrare questa tecnica, consideriamo proprio la funzione $f(x) = e^{-x^2}$. Non è semplicissimo scrivere la formula di Taylor con il resto di Lagrange per questa funzione; conviene passare attraverso la funzione e^x per la quale invece la formula è molto semplice. Abbiamo

$$e^x = \sum_{k=0}^{N-1} \frac{x^k}{k!} + e^c\frac{x^N}{N!}$$

dove c è compreso tra 0 e x. Da questa si ottiene:

$$e^{-x^2} = \sum_{k=0}^{N-1} (-1)^k\frac{x^{2k}}{k!} + (-1)^N e^c\frac{x^{2N}}{N!}$$

dove ora $c \in]-x^2, 0[$. Si ha quindi, ad esempio

$$\int_0^1 e^{-x^2}\,\mathrm{d}x = \int_0^1 \left[\sum_{k=0}^{N-1} (-1)^k \frac{x^{2k}}{k!}\right] \mathrm{d}x + \int_0^1 (-1)^N e^c \frac{x^{2N}}{N!}\,\mathrm{d}x\,. \tag{6.26}$$

L'errore che si commette trascurando l'ultimo integrale a destra si può stimare come segue

$$\left|\int_0^1 (-1)^N e^c \frac{x^{2N}}{N!}\,\mathrm{d}x\right| \le \int_0^1 e^c \frac{x^{2N}}{N!}\,\mathrm{d}x \le$$

$$e\frac{1}{N!}\int_0^1 x^{2N}\,\mathrm{d}x = \frac{e}{N!}\left[\frac{x^{2N+1}}{2N+1}\right]\Big|_0^1 = \frac{e}{(2N+1)N!}\,.$$

Fissato $\epsilon > 0$, per esser certi che l'errore che si commette trascurando l'ultimo integrale a destra della (6.26) non superi ϵ è sufficiente scegliere N in modo tale che

$$\frac{e}{(2N+1)N!} \le \epsilon\,.$$

Esercizio 6.14 Calcolare, con un errore minore di 10^{-3}, gli integrali

$$\int_0^1 \frac{e^x - 1}{x}\,\mathrm{d}x\,, \qquad \int_0^{\sqrt{\pi/2}} \sin x^2\,\mathrm{d}x\,.$$

R. 1.318, 0.549.

6.10 Altri esercizi

Esercizio 6.15 Calcolare i seguenti integrali indefiniti

$$\int xe^{-x^2}\,\mathrm{d}x\,, \qquad \int x \arctan x^2\,\mathrm{d}x\,, \qquad \int \cot x\,\mathrm{d}x\,,$$

$$\int \sin(3x+7)\,\mathrm{d}x\,, \qquad \int \sin^3 x\,\mathrm{d}x\,, \qquad \int x^2 e^{-2x}\,\mathrm{d}x\,,$$

$$\int \tfrac{x-1}{x+1}\,\mathrm{d}x\,, \qquad \int \tfrac{1}{3x^2+5}\,\mathrm{d}x\,, \qquad \int \tfrac{1}{2x^2-7}\,\mathrm{d}x\,,$$

$$\int \tfrac{\ln x}{x(1+\ln x)}\,\mathrm{d}x\,, \qquad \int \tfrac{x+1}{\sqrt{1-x^2}}\,\mathrm{d}x\,, \qquad \int \tfrac{1}{\cos x}\,\mathrm{d}x\,.$$

R. (da sinistra a destra e dall'alto in basso) $-\frac{1}{2}e^{-x^2}+k$, $\frac{1}{2}x^2\arctan x^2-\frac{1}{4}\ln(1+x^4)+k$, $\ln|\sin x|+k$ $(x\neq n\pi, n\in\mathbb{Z})$, $-\frac{1}{3}\cos(3x+7)+k$, $\frac{1}{3}\cos^3 x-\cos x+k$, $-\frac{1}{4}(2x^2+2x+1)e^{-2x}+k$, $x-\ln(x+1)^2+k$ $(x\neq -1)$, $\frac{1}{\sqrt{15}}\arctan\left(\sqrt{\frac{3}{5}}x\right)+k$, $\frac{1}{2\sqrt{14}}\ln\left|\frac{x-\sqrt{\frac{7}{2}}}{x+\sqrt{\frac{7}{2}}}\right|+k$ $(x\neq\pm\sqrt{\frac{7}{2}})$, $\ln x-\ln|1+\ln x|+k$ $(x>0, x\neq 1/e)$, $\arcsin x-\sqrt{1-x^2}+k$ $(|x|<1)$, $\ln\left|\frac{1+\sin x}{\cos x}\right|+k$ $(x\neq n\pi+\pi/2, n\in\mathbb{Z})$.

Esercizio 6.16 Calcolare i seguenti integrali definiti

$$\int_0^{2\pi}\sin^3 x\,\mathrm{d}x\,,\qquad \int_{-\pi}^{\pi}x^4\sin^3 x+x^2\,\mathrm{d}x\,,\qquad \int_0^{\ln 2}\sqrt{e^x-1}\,\mathrm{d}x,\qquad \int_{\ln 4}^{\ln 9}\frac{1}{e^{2x}-2e^x}\,\mathrm{d}x\,.$$

R. 0, $\frac{2}{3}\pi^3$, $2-\frac{\pi}{2}$, $-\frac{5}{72}+\frac{1}{4}\ln\frac{14}{9}$.

Esercizio 6.17 Calcolare l'area delimitata dall'ellisse

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1\,.$$

R. πab.

Capitolo 7

Integrali generalizzati

7.1 Integrali su semirette

L'integrale di Riemann trattato sino a questo momento ha riguardato funzioni limitate su intervalli limitati. Scopo di questa sezione è mostrare delle utili estensioni a situazioni dove la funzione può essere illimitata e/o definita su di un intervallo illimitato. Cominceremo con il secondo caso.

Consideriamo una funzione $f : [a, +\infty[\to \mathbb{R}$ integrabile su ogni intervallo finito del tipo $[a, b]$ con $b \geq a$. Poniamo la seguente:

Definizione 7.1 La funzione f si dice *integrabile (in senso improprio)* su $[a, +\infty[$ se esiste finito il limite

$$\lim_{b \to +\infty} \int_a^b f(x)\,\mathrm{d}x$$

e in tal caso si pone

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = \lim_{b \to +\infty} \int_a^b f(x)\,\mathrm{d}x\,.$$

Esempio 114 Consideriamo la funzione

$$f(x) = \frac{1}{x^\gamma}, \qquad x \in [1, +\infty[$$

dove $\gamma \in \mathbb{R}$. Si ha che

$$\int_1^b \frac{1}{x^\gamma}\, \mathrm{d}x = \begin{cases} \dfrac{b^{-\gamma+1}}{-\gamma+1} - \dfrac{1}{-\gamma+1} & \gamma \neq 1\,, \\ \ln b & \gamma = 1\,. \end{cases}$$

Per $b \to +\infty$ l'espressione sopra diverge a $+\infty$ se $\gamma \leq 1$ mentre converge a $1/(\gamma-1)$ se $\gamma > 1$. Quindi, la funzione $1/x^\gamma$ è integrabile su $[1, +\infty[$ se e soltanto se $\gamma > 1$ e in tal caso si ha

$$\int_1^{+\infty} \frac{1}{x^\gamma}\, \mathrm{d}x = \frac{1}{\gamma - 1} \qquad (\gamma > 1)\,.$$

Esempio 115 Consideriamo la funzione

$$f(x) = \frac{1}{x(\ln x)^\gamma}, \qquad x \in [2, +\infty[$$

dove $\gamma \in \mathbb{R}$. Si ha che

$$\int_2^b \frac{1}{x(\ln x)^\gamma}\, \mathrm{d}x = \int_{\ln 2}^{\ln b} \frac{1}{t^\gamma}\, \mathrm{d}t = \begin{cases} \dfrac{(\ln b)^{-\gamma+1}}{-\gamma+1} - \dfrac{\ln 2}{-\gamma+1} & \gamma \neq 1,, \\ \ln(\ln b) - \ln(\ln 2) & \gamma = 1\,. \end{cases}$$

Per $b \to +\infty$ l'espressione sopra diverge a $+\infty$ se $\gamma \leq 1$ mentre converge a $\ln 2/(\gamma - 1)$ se $\gamma > 1$. Quindi, la funzione $1/1/[x(\ln x)^\gamma]$ è integrabile su $[2, +\infty[$ se e soltanto se $\gamma > 1$ e in tal caso si ha

$$\int_2^{+\infty} \frac{1}{x(\ln x)^\gamma}\, \mathrm{d}x = \frac{\ln 2}{\gamma - 1} \qquad (\gamma > 1)\,.$$

Esempio 116 Consideriamo la funzione

$$f(x) = e^{-\lambda x}, \qquad x \in [0, +\infty[$$

con $\lambda > 0$. Si ha che

$$\int_0^b e^{-\lambda x}\, \mathrm{d}x = \frac{1}{\lambda}[1 - e^{-\lambda b}]\,.$$

Quando $b \to +\infty$, l'espressione sopra converge (poichè $\lambda > 0$) a $1/\lambda$. Quindi $e^{-\lambda x}$ è integrabile su $[0, +\infty[$ e si ha

$$\int_0^{+\infty} e^{-\lambda x}\,\mathrm{d}x = \frac{1}{\lambda} \qquad (\lambda > 0)\,.$$

Osservazione. Nel caso in cui $f(x)$ sia una funzione a valori non negativi, l'integrale definito sopra ammette ancora l'interpretazione di area sottesa al grafico della funzione come nel caso degli integrali ordinari; l'unica differenza è che ora la regione considerata è illimitata.

Naturalmente non c'è niente di speciale riguardo alle semirette destre e tutto si può ripetere per le semirette sinistre. Precisamente, se abbiamo una funzione $f :]-\infty, b] \to \mathbb{R}$ integrabile su ogni intervallo finito del tipo $[a, b]$ con $a \leq b$, si pone la seguente

Definizione 7.2 La funzione f si dice *integrabile (in senso improprio)* su $]-\infty, b]$ se esiste finito il limite

$$\lim_{a \to -\infty} \int_a^b f(x)\,\mathrm{d}x$$

e in tal caso si pone

$$\int_{-\infty}^b f(x)\,\mathrm{d}x = \lim_{a \to -\infty} \int_a^b f(x)\,\mathrm{d}x\,.$$

Osservazione. Nel caso in cui si abbia una $f : \mathbb{R} \to \mathbb{R}$ integrabile sia sulla semiretta $[0, +\infty[$ che sulla semiretta $]-\infty, 0]$, diremo anche che la funzione è integrabile su $]-\infty, +\infty[$ e si pone

$$\int_{-\infty}^{+\infty} f(x)\,\mathrm{d}x = \int_{-\infty}^0 f(x)\,\mathrm{d}x + \int_0^{+\infty} f(x)\,\mathrm{d}x\,.$$

Esercizio 7.1 Dire se le seguenti funzioni sono integrabili nei domini specificati e in caso affermativo calcolarne gli integrali impropri relativi:

$$f(x) = \sin x \quad x \in [0, +\infty[\,, \qquad\qquad f(x) = xe^{3x} \quad x \in]-\infty, 0]\,,$$

$$f(x) = e^{-x} \sin x \quad x \in [0, +\infty[\,, \qquad\qquad f(x) = \tfrac{1}{1+x^2} \quad x \in]-\infty, +\infty[\,.$$

R. no, si $(-\frac{1}{9})$, si $(\frac{1}{2})$, si (π).

Gli esempi sopra potrebbero indurre a pensare che non ci sia, da un punto di vista operativo, nessuna novità nel dover trattare integrali su semirette anzichè su intervalli finiti: si integra la funzione sugli intervalli finiti e poi si calcola un limite. Tutto questo può funzionare se intanto sappiamo calcolare esplicitamente l'integrale sugli intervalli finiti. Ci sono molti casi tuttavia in cui questo calcolo non è fattibile; come fare dunque per sapere se la funzione è integrabile o meno? Si noti come per la definizione di integrale improprio che abbiamo dato sia necessario a priori che la funzione sia integrabile sugli intervalli limitati, questo ad esempio sappiamo essere automaticamente vero se la funzione è continua. Gli esempi mostrano tuttavia che la continuità non è una condizione sufficiente per l'integrabilità in senso improprio; è chiaro che tutto dipende dal comportamento asintotico della funzione. Daremo nel seguito criteri di integrabilità che si collegano proprio al comportamento all'infinito della funzione.

Cominciamo con l'illustrare alcune proprietà simili a quelle dell'integrale di Riemann ordinario. Saranno illustrate per il caso di semirette destre, lasciando allo studente il compito di riformularle per il caso di semirette sinistre

Proposizione 7.3 *Siano $f : [a, +\infty[\to \mathbb{R}$ e $g : [a, +\infty[\to \mathbb{R}$ due funzioni integrabili su $[a, +\infty[$ e sia $\lambda \in \mathbb{R}$. Allora:*

(i) $f + g$ è integrabile su $[a, +\infty[$ e si ha

$$\int_a^{+\infty} (f + g)(x)\,\mathrm{d}x = \int_a^{+\infty} f(x)\,\mathrm{d}x + \int_a^{+\infty} g(x)\,\mathrm{d}x\,.$$

(ii) λf è integrabile su $[a, +\infty[$ e si ha

$$\int_a^{+\infty} \lambda f(x)\,\mathrm{d}x = \lambda \int_a^{+\infty} f(x)\,\mathrm{d}x\,.$$

Dimostrazione Segue subito dalla Definizione 7.1, dalla Proposizione 6.9 e dalle proprietà dei limiti. ■

Proposizione 7.4 *Sia* $f : [a, +\infty[\to \mathbb{R}$ *una funzione e sia* $c > a$*. Allora,*

(a) Se f *è integrabile su* $[a, +\infty[$*, essa è anche integrabile su* $[c, +\infty[$ *e si ha*

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^{+\infty} f(x)\,\mathrm{d}x\,. \tag{7.1}$$

(b) Se f *è integrabile su* $[a, c]$ *e integrabile su* $[c, +\infty[$*, allora* f *è integrabile su* $[a, +\infty[$ *e vale ancora (7.1).*

Dimostrazione (a): Sia $b > c$. Segue dalla Proposizione 6.11 (vedi Osservazione 1 ad essa seguente) che

$$\int_c^b f(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x - \int_a^c f(x)\,\mathrm{d}x\,.$$

Per $b \to +\infty$, il secondo membro converge a

$$\int_a^{+\infty} f(x)\,\mathrm{d}x - \int_a^c f(x)\,\mathrm{d}x\,.$$

Questo vuol dire che f è integrabile su $[c, +\infty[$ e si ha

$$\int_c^{+\infty} f(x)\,\mathrm{d}x = \int_a^{+\infty} f(x)\,\mathrm{d}x - \int_a^c f(x)\,\mathrm{d}x\,.$$

Portando l'ultimo integrale a primo membro si ottiene proprio la (7.1).

(b): Sia $b > c$. Segue dalla Proposizione 6.11 che f è integrabile su $[a, c]$ e si ha

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x\,.$$

Il secondo membro, per ipotesi, ammette limite per $b \to +\infty$. Quindi f è integrabile su $[c, +\infty[$ e si ottiene ancora la (7.1). ■

7.2 Integrali su semirette di funzioni positive

C'è un'importante classe di funzioni per le quali lo studio dell'integrabilità impropria risulta particolarmente agevole: sono le funzioni a valori non-negativi. In effetti si supponga di avere una funzione $f : [a,+\infty[\to \mathbb{R}$, integrabile su ogni intervallo finito $[a,b]$ e tale che $f(x) \geq 0$ per ogni $x \geq a$. Segue allora che

$$\int_a^b f(x)\,\mathrm{d}x, \quad b \geq a$$

è una funzione crescente di b (più b cresce, più aumenta l'area positiva che sta tra il grafico e l'asse x; si veda la Figura 7.1).

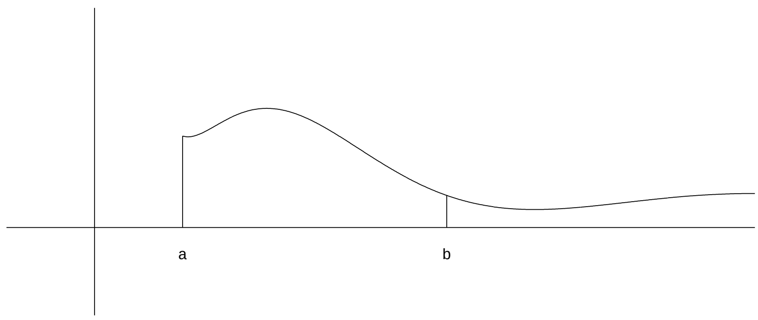


Figura 7.1:

Dunque esiste sempre il limite per $b \to +\infty$, finito o eguale a $+\infty$; si ha in effetti

$$\lim_{b\to+\infty} \int_a^b f(x)\,\mathrm{d}x = \sup_{b\geq a} \int_a^b f(x)\,\mathrm{d}x\,.$$

Il limite è finito se e soltanto se esiste $k \geq 0$ tale che

$$\int_a^b f(x)\,\mathrm{d}x \leq k, \quad \forall b \geq a\,.$$

Nel caso in cui il limite sia $+\infty$ si usa comunque la notazione

$$\int_a^{+\infty} f(x)\,\mathrm{d}x = +\infty$$

e si dice anche che l'integrale improprio *diverge*.

Queste osservazioni mostrano una notevole similarità tra l'integrazione impropria di funzioni non-negative e la somma delle serie a termini non-negativi. In effetti si ha il seguente fondamentale criterio di integrabilità in senso improprio che ricalca strettamente il criterio del confronto per serie.

Teorema 7.5 (del confronto) *Siano* $f : [a, +\infty[\to \mathbb{R}$ *e* $g : [a, +\infty[\to \mathbb{R}$ *due funzioni integrabili su ogni intervallo finito* $[a, b]$ *tali che*

$$0 \le f(x) \le g(x), \quad \forall x \ge a\,.$$

Allora se g *è integrabile su* $[a, +\infty[$, *lo è anche* f *e si ha*

$$\int_a^{+\infty} f(x)\,\mathrm{d}x \le \int_a^{+\infty} g(x)\,\mathrm{d}x\,.$$

Se invece f *non è integrabile (il suo integrale diverge), anche* g *non è integrabile.*

Dimostrazione Basta osservare che per le ipotesi fatte si ha

$$\int_a^{b} f(x)\,\mathrm{d}x \le \int_a^{b} g(x)\,\mathrm{d}x, \quad \forall b \ge a$$

e poi ripetere le stesse considerazioni fatte nella dimostrazione del Teorema 3.23 del confronto tra serie. ■

Esempio 117 Sia

$$f(x) = e^{-x^2}\,.$$

Facciamo vedere che essa è integrabile su $[0, +\infty[$. Si noti che non possiamo calcolare una primitiva esplicita di questa funzione. Cerchiamo invece di utilizzare la tecnica del confronto. Poichè $x^2 \ge x$ se $x \ge 1$, si ha che $e^{-x^2} \le e^{-x}$ per $x \ge 1$. Segue dall'Esempio 116 che e^{-x} è integrabile su $[0, +\infty[$ e quindi, per la Proposizione 7.4, anche su $[1, +\infty[$. Quindi, per il Teorema del confronto, anche e^{-x^2} è integrabile su $[1, +\infty[$ e, nuovamente per la Proposizione 7.4, anche quindi su $[0, +\infty[$. Si noti che per simmetria (la funzione è pari) si ha anche che e^{-x^2} è integrabile su $]-\infty, 0]$ e quindi anche su $]-\infty, +\infty[$.

Esempio 118 Sia

$$f(x) = \frac{1}{x^2 + \sin x}\,, \qquad x \geq 2\,.$$

Si ha che

$$\frac{1}{x^2 + \sin x} \leq \frac{1}{x^2 - 1} \leq \frac{2}{x^2}\,, \quad \forall x \geq 2$$

dove la prima diseguaglianza segue dal fatto che $\sin x \geq -1$, mentre la seconda può essere verificata direttamente. Poichè $1/x^2$ è integrabile su $[1, +\infty[$ (vedi Esempio 114), lo è anche su $[2, +\infty[$. Per confronto anche $f(x) = 1/(x^2 + \sin x)$ è dunque integrabile su $[2, +\infty[$ e vale la diseguaglianza

$$\int_2^{+\infty} \frac{1}{x^2 + \sin x}\,\mathrm{d}x \leq \int_2^{+\infty} \frac{2}{x^2}\,\mathrm{d}x = \lim_{b\to+\infty} \left(-\frac{2}{x}\bigg|_2^b \right) = 1\,.$$

Esempio 119 Sia

$$f(x) = \frac{1}{\ln x}\,, \qquad x \geq 2\,.$$

Poichè $0 < \ln x \leq x$ se $x \geq 2$, si ha che

$$0 < \frac{1}{x} \leq \frac{1}{\ln x}, \quad \forall x \geq 2\,.$$

Poichè $1/x$ non è integrabile su $[1, +\infty[$ (vedi Esempio 114), non lo è neppure su $[2, +\infty[$. Per confronto neanche $f(x) = 1/\ln x$ lo è. Si ha dunque,

$$\int_2^{+\infty} \frac{1}{\ln x}\,\mathrm{d}x = +\infty\,.$$

Come accadeva per le serie, anche il Teorema del confronto per gli integrali ammette un utile variante di tipo asintotico.

Teorema 7.6 (del confronto asintotico) *Siano $f : [a, +\infty[\to \mathbb{R}$ e $g : [a, +\infty[\to \mathbb{R}$ due funzioni integrabili su ogni intervallo finito $[a, b]$, strettamente positive e tali che esista finito il limite*

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)} = l > 0\,.$$

Allora f è integrabile su $[a, +\infty[$ se e soltanto se g lo è.

Dimostrazione Supponiamo che g sia integrabile su $[a,+\infty[$. Per la definizione di limite, esiste $M \geq a$ tale che

$$x \geq M \;\Rightarrow\; \frac{f(x)}{g(x)} \leq l+1$$

o, equivalentemente,

$$x \geq M \;\Rightarrow\; f(x) \leq (l+1)g(x)\,.$$

Ne segue che, per il Teorema del confronto, f è integrabile su $[M,+\infty[$ e quindi anche su $[a,+\infty[$. L'altra implicazione segue dalla precedente scambiando il ruolo di f e g. ■

Osservazione. Il Teorema precedente dice in particolare che se $f,g>0$ ed f è *asintotica* a g per $x\to+\infty$ allora gli integrali impropri di f e di g su $[a,+\infty[$ o convergono entrambi o divergono entrambi. Basterà allora di solito identificare la *parte principale* $P(x)$ della funzione del cui integrale vi si chiede di discutere l'esistenza, e studiare l'integrabilità di $P(x)$, la cui espressione è di norma molto più facilmente indagabile.

Esempio 120 Sia

$$f(x)=\frac{x}{x^2-3}\,,\qquad x\geq 2\,.$$

Consideriamo $g(x)=1/x$. Si ha che

$$\lim_{x\to+\infty}\frac{f(x)}{g(x)}=\lim_{x\to+\infty}\frac{x^2}{x^2-3}=1\,.$$

Ciò può anche essere riscritto, come ben sappiamo, nella forma $f(x)\sim\frac{1}{x}$ se $x\to+\infty$. f quindi *si comporta come* $\frac{1}{x}$ *per* $x\to+\infty$. Poichè g non è integrabile su $[2,+\infty[$ (vedi Esempio 114), non lo è neanche f. Si ha dunque

$$\int_2^{+\infty}\frac{x}{x^2-3}\,\mathrm{d}x=+\infty\,.$$

Esempio 121 Sia

$$f(x)=\sin\left(\frac{1}{x^2}\right)\,,\qquad x\geq 1/2\,.$$

Si noti innanzitutto che $f(x)>0$ per $x>1/\sqrt{\pi}$. Fissiamo un qualunque $a>1/\sqrt{\pi}$, ad esempio possiamo prendere $a=1$. Consideriamo ora $g(x)=1/x^2$. Si ha che

$$\lim_{x\to+\infty}\frac{f(x)}{g(x)}=\lim_{x\to+\infty}x^2\sin\frac{1}{x^2}=\lim_{t\to0}\frac{\sin t^2}{t^2}=1\,.$$

Poichè g è integrabile su $[1,+\infty[$ (vedi Esempio 114), lo è anche f. Per la Proposizione 7.4 f è anche integrabile su $[1/2,+\infty[$.

Esercizio 7.2 Studiare l'integrabilità delle funzioni seguenti sui domini specificati

$$e^{-\sqrt{x}} \quad x \in [0,+\infty[\,, \qquad \sqrt{x+1}-\sqrt{x}\,, \quad x \in [0,+\infty[\,,$$

$$\frac{\arctan x}{x} \quad x \in]-\infty,-1]\,, \qquad \frac{\sqrt{x}+1}{x^2+1}\,, \quad x \in [0,+\infty[\,.$$

R. si, no, no, si.

Il collegamento con la teoria delle serie è reso ancora più stringente dal seguente risultato che collega la convergenza di un integrale alla convergenza di una serie.

Teorema 7.7 *Sia $f : [0,+\infty[\to \mathbb{R}$ una funzione decrescente e a valori non negativi. Sono fatti equivalenti*

(a) f è integrabile su $[0,+\infty[$.

(b) $\sum_{k=0}^{+\infty} f(k)$ converge.

Inoltre si ha

$$\sum_{k=1}^{+\infty} f(k) \le \int_0^{+\infty} f(x)\,\mathrm{d}x \le \sum_{k=0}^{+\infty} f(k)\,. \tag{7.2}$$

Dimostrazione (a) $\Rightarrow$ (b): consideriamo l'intervallo $[0,n]$ con $n \in \mathbb{N}$ e sia δ_n la partizione uniforme di detto intervallo in n sottointervalli, cioè

$$\delta_n = \{0,1,2,\ldots,n\}\,.$$

Poichè f è decrescente, si ha che

$$s_{\delta_n}(f) = \sum_{k=1}^{n} f(k)\,.$$

Dunque abbiamo (si veda la Figura 7.2)

$$\sum_{k=1}^{n} f(k) \le \int_0^{n} f(x)\,\mathrm{d}x \le \int_0^{+\infty} f(x)\,\mathrm{d}x\,. \tag{7.3}$$

La serie a termini positivi $\sum_{k=1}^{+\infty} f(k)$ è dunque limitata e, di conseguenza, convergente. E' quindi anche convergente la serie $\sum_{k=0}^{+\infty} f(k)$. Si noti inoltre che da (7.3) segue che

$$\sum_{k=1}^{+\infty} f(k) \le \int_0^{+\infty} f(x)\,\mathrm{d}x\,. \tag{7.4}$$

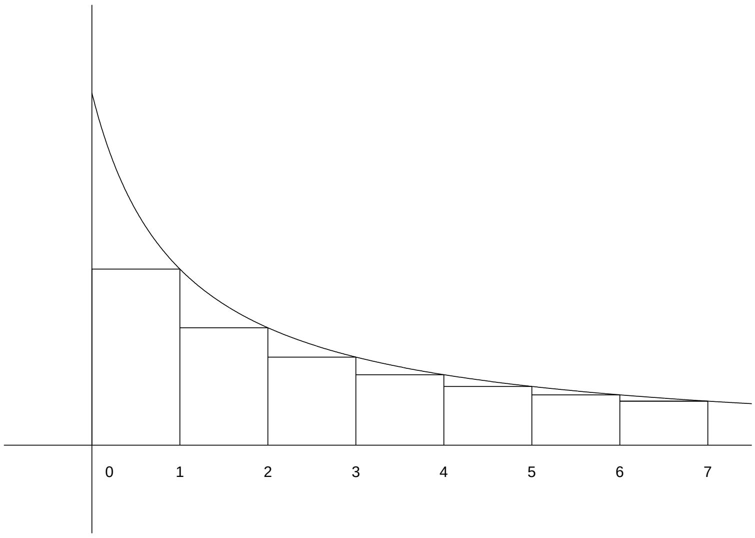


Figura 7.2:

(b) $\Rightarrow$ (a): consideriamo come prima l'intervallo $[0, n]$ con $n \in \mathbb{N}$ e la partizione uniforme δ_n che lo divide in n sottointervalli. Poichè f è decrescente, si ha che

$$S_{\delta_n}(f) = \sum_{k=0}^{n-1} f(k)\,.$$

Dunque abbiamo (si veda la Figura 7.3)

$$\int_0^n f(x)\,\mathrm{d}x \leq \sum_{k=0}^{n-1} f(k) \leq \sum_{k=0}^{+\infty} f(k)\,. \tag{7.5}$$

Sia ora $b \in \mathbb{R}$ qualunque. Esiste sicuramente $n \in \mathbb{N}$ tale che $b \leq n$. Quindi, per la non-negatività della f e la diseguaglianza (7.5) si ha che

$$\int_0^b f(x)\,\mathrm{d}x \leq \int_0^n f(x)\,\mathrm{d}x \leq \sum_{k=0}^{+\infty} f(k)\,.$$

Questo significa che

$$\int_0^{+\infty} f(x)\,\mathrm{d}x = \sup_{b\geq 0} \int_0^b f(x)\,\mathrm{d}x \leq \sum_{k=0}^{+\infty} f(k)\,. \tag{7.6}$$

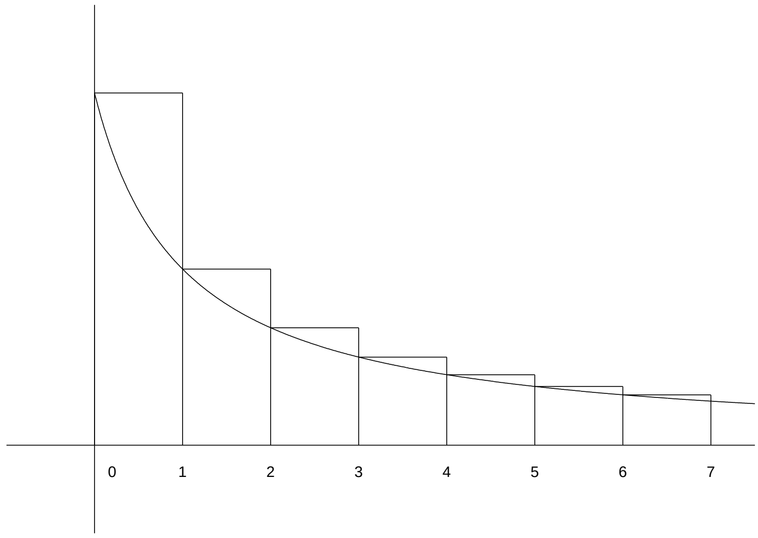


Figura 7.3:

Quindi f è integrabile su $[0, +\infty[$. Si ha inoltre dalla (7.6) che

$$\int_0^{+\infty} f(x)\,\mathrm{d}x \leq \sum_{k=0}^{+\infty} f(k)\,. \tag{7.7}$$

La dimostrazione è dunque completata: si noti che la diseguaglianza (7.2) segue dalle (7.4) e (7.7). ■

Osservazione. Si noti come il risultato precedente continui a valere se f è definita su una semiretta più piccola $[\nu, +\infty[$ con $\nu \in \mathbb{N}$, ed è ivi ancora integrabile su ogni intervallo finito $[\nu, b]$, non-negativa e decrescente. In tal caso la diseguaglianza (7.2) viene rimpiazzata da

$$\sum_{k=\nu+1}^{+\infty} f(k) \leq \int_\nu^{+\infty} f(x)\,\mathrm{d}x \leq \sum_{k=\nu}^{+\infty} f(k)\,. \tag{7.8}$$

L'utilizzo principale del Teorema 7.7 consiste nel riportare lo studio della convergenza di una serie a termini positivi allo studio, in genere più semplice, della convergenza di un integrale improprio. Il prossimo esempio mostra un'applicazione in questo senso, completando lo studio di convergenza delle serie numeriche armoniche.

Esempio 122 Consideriamo $f(x) = 1/x^\gamma$ con $\gamma \geq 0$. Essa risulta integrabile su $[1, +\infty[$ se e soltanto se $\gamma > 1$. Essa soddisfa inoltre alle ipotesi del Teorema 7.7 (vedi Osservazione precedente). Si ha quindi che la serie $\sum_{k=1}^{+\infty} 1/k^\gamma$ converge se e soltanto se $\gamma > 1$ e si ha

$$\sum_{k=2}^{+\infty} \frac{1}{k^\gamma} \leq \int_1^{+\infty} \frac{1}{x^\gamma} \, \mathrm{d}x \leq \sum_{k=1}^{+\infty} \frac{1}{k^\gamma} \, .$$

Utilizzando il risultato dell'Esempio 114 si ottiene così

$$\sum_{k=2}^{+\infty} \frac{1}{k^\gamma} \leq \frac{1}{\gamma - 1} \leq \sum_{k=1}^{+\infty} \frac{1}{k^\gamma} \, .$$

Sommando 1 ad entrambi i membri della prima eguaglianza, si può riscrivere come

$$\frac{1}{\gamma - 1} \leq \sum_{k=1}^{+\infty} \frac{1}{k^\gamma} \leq \frac{1}{\gamma - 1} + 1 \, .$$

Esercizio 7.3 Utilizzando il Teorema 7.7 e l'Esempio 115, si studi la convergenza della serie

$$\sum_{k=2}^{+\infty} \frac{1}{k(\ln k)^\gamma}$$

al variare di $\gamma \in \mathbb{R}$.
R. Vi è convergenza se e solo se $\gamma > 1$.

7.3 Integrazione assoluta

Con la tecnica del confronto si può studiare l'integrabilità impropria di funzioni non negative o anche di funzioni che, pur non essendo sempre nonnegative sulla semiretta di integrazione, lo sono comunque al di fuori di un intervallo finito. La tecnica del confronto si applica naturalmente anche a funzioni f che sono invece sempre negative (o negative fuori di un intervallo finito): basta considerare $-f$ che è quindi a valori positivi. Come accadeva per le serie, le cose cambiano drasticamente se la funzione ha cambiamenti di segno su tutta la semiretta. Per queste c'è un concetto più forte di integrabilità analogo alla convergenza assoluta:

Definizione 7.8 Una funzione $f : [a, +\infty[\to \mathbb{R}$ si dice *integrabile assolutamente (in senso improprio)* su $[a, +\infty[$ se è integrabile su ogni intervallo limitato $[a, b]$ e se è integrabile in senso improprio su tale semiretta la funzione $|f|$, cioè se esiste finito

$$\lim_{b\to+\infty} \int_a^b |f(x)| \, \mathrm{d}x \, .$$

Come nel caso delle serie, l'integrabilità assoluta implica l'integrabilità, come illustra il seguente risultato.

Teorema 7.9 *Se una funzione $f : [a, +\infty[\to \mathbb{R}$ è assolutamente integrabile su $[a, +\infty[$, allora è integrabile su detta semiretta e si ha*

$$\left| \int_a^{+\infty} f(x) \, \mathrm{d}x \right| \leq \int_a^{+\infty} |f(x)| \, \mathrm{d}x \, . \tag{7.9}$$

Dimostrazione Definiamo preliminarmente i concetti di parte positiva e negativa di un numero reale x come:

$$x^+ = \begin{cases} x & \text{se} \quad x \geq 0 \\ 0 & \text{se} \quad x < 0 \, , \end{cases}$$

$$x^- = \begin{cases} 0 & \text{se} \quad x \geq 0 \\ -x & \text{se} \quad x < 0 \, . \end{cases}$$

Si hanno le seguenti relazioni di immediata verifica:

$$\begin{array}{ll} x = x^+ - x^- \, , & |x| = x^+ + x^- \\ x^+ = \frac{|x|+x}{2} \, , & x^- = \frac{|x|-x}{2} \end{array}$$

Veniamo ora alla dimsotrazione vera e propria. Consideriamo la parte positiva e la parte negativa di $f(x)$: si ottengono così due funzioni: $f(x)^+$ e $f(x)^-$ entrambe non negative. Poichè sia $f(x)$ che $|f(x)|$ sono per ipotesi integrabili su ogni intervallo finito $[a, b]$ e poichè

$$f(x)^+ = \frac{|f(x)| + f(x)}{2} \, , \qquad f(x)^- = \frac{|f(x)| - f(x)}{2} \, ,$$

lo sono anche, in virtù della Proposizione 6.9, le funzioni $f(x)^+$ e $f(x)^-$. Inoltre, poichè

$$|f(x)| = f(x)^+ + f(x)^- \, ,$$

si ha

$$f(x)^+ \leq |f(x)| \, , \quad f(x)^- \leq |f(x)| \, .$$

Ne segue, in virtù del Teorema 7.5, che $f(x)^+$ e $f(x)^-$ sono entrambe integrabili su $[a, +\infty[$. Essendo $f(x) = f(x)^+ - f(x)^-$, segue allora dalla Proposizione 7.3 che anche $f(x)$ è integrabile su $[a, +\infty[$. Infine la diseguaglianza (7.9) segue dalla corrispondente diseguaglianza per intervalli finiti in (ii) della Proposizione 6.10 , prendendo poi il limite per $b \to +\infty$.■

Esempio 123 Sia

$$f(x) = \frac{\sin x}{x^2}\,, \qquad x \geq 1\,.$$

Poichè

$$|f(x)| = \frac{|\sin x|}{x^2} \leq \frac{1}{x^2}\,,$$

per confronto $|f(x)|$ è integrabile su $[1, +\infty[$. Dunque, $f(x)$ è integrabile assolutamente su $[1, +\infty[$. In virtù del Teorema 7.9 si ha quindi che f è integrabile su $[1, +\infty[$.

Presentiamo ora un esempio un po' più complicato (noto come *integrale di Fresnel*) che mostra come non sia affatto necessario che la funzione sia infinitesima per essere integrabile in senso improprio; questo fatto marca un'importante differenza rispetto al comportamento delle serie numeriche convergenti (vedi Proposizione 3.21).

Esempio 124 Sia

$$f(x) = \sin(x^2)$$

(vedi Figura 7.4). Studiamone l'integrabilità sulla semiretta $[1, +\infty[$. Scriviamo

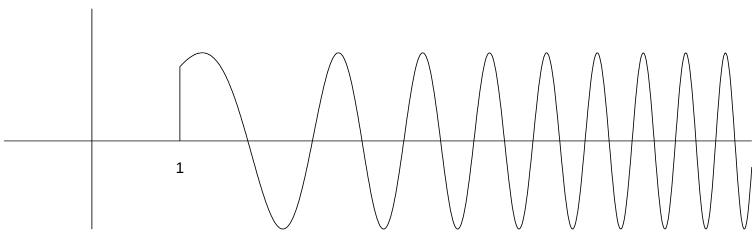


Figura 7.4:

$$\sin(x^2) = \frac{1}{2x} 2x \sin(x^2)$$

e si noti che $2x\sin(x^2)$ è la derivata di $-\cos(x^2)$. Utilizzando la tecnica di integrazione per parti si ha che

$$\int_1^b \sin(x^2)\,\mathrm{d}x = \int_1^b \frac{1}{2x} 2x\sin(x^2)\,\mathrm{d}x = \left[-\frac{1}{2x}\cos(x^2)\right]\Bigg|_1^b - \int_1^b \frac{1}{2x^2}\cos(x^2)\,\mathrm{d}x\,. \tag{7.10}$$

Si noti ora che

$$\lim_{b\to+\infty}\left[-\frac{1}{2x}\cos(x^2)\right]\Bigg|_1^b = \lim_{b\to+\infty}\left[-\frac{1}{2b}\cos(b^2)+\frac{1}{2}\cos 1\right] = \frac{1}{2}\cos 1\,.$$

D'altra parte, con un ragionamento completamente analogo a quello fatto nell'Esempio 123 si vede che la funzione

$$\frac{1}{2x^2}\cos(x^2)$$

è assolutamente integrabile su $[1,+\infty[$; quindi essa è integrabile su $[1,+\infty[$ il che significa che esiste finito il limite, per $b\to+\infty$, dell'espressione

$$\int_1^b \frac{1}{2x^2}\cos(x^2)\,\mathrm{d}x\,.$$

Segue allora da (7.10) che esiste finito il limite, per $b\to+\infty$, di

$$\int_1^b \sin(x^2)\,\mathrm{d}x$$

e questo significa proprio che la funzione $\sin(x^2)$ è integrabile su $[1,+\infty[$. Si presti attenzione al fatto che non abbiamo affatto dimostrato che $\sin(x^2)$ è integrabile assolutamente, ma soltanto che è integrabile. In effetti si può far vedere (Esercizio * per voi) che detta funzione non è assolutamente integrabile. Questo significa che sono proprio le continue cancellazioni tra aree negative e positive a far si che l'integrale converga.

Esercizio 7.4 Studiare l'integrabilità assoluta delle funzioni seguenti sui domini di seguito indicati:

$$e^{-x}\sin\sqrt{x},\quad x\in[0,+\infty[;\qquad \frac{\ln x\sin x}{x^2},\quad x\in[1,+\infty[.$$

R. Entrambe assolutamente integrabili.

Esercizio 7.5 Studiare l'integrabilità della funzione $\cos(x^2)$ su $[1, +\infty[$.
R. Integrabile.

Esercizio 7.6 Studiare l'integrabilità della funzione $(\sin x)/x$ su $[1, +\infty[$. (Questo integrale è noto come *integrale di Dirichlet*).
R. Integrabile.

7.4 Integrazione di funzioni non limitate

Passiamo ora all'altra estensione del concetto di integrale a funzioni non limitate. Consideriamo una funzione $f : [a, b[\to \mathbb{R}$ integrabile su ogni intervallo chiuso $[a, b - \epsilon]$, con $\epsilon > 0$. Non facciamo per il momento ipotesi sul comportamento della f per $x \to b-$: potrebbe non esistere il limite o essere presente un asintoto. Abbiamo la seguente:

Definizione 7.10 La funzione $f : [a, b[\to \mathbb{R}$ integrabile su ogni intervallo chiuso $[a, b - \epsilon]$, con $\epsilon > 0$, si dice *integrabile (in senso improprio)* su $[a, b[$ se esiste finito il limite

$$\lim_{t \to b-} \int_a^t f(x)\,\mathrm{d}x$$

e in tal caso si pone

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{t \to b-} \int_a^t f(x)\,\mathrm{d}x\,.$$

Analogamente si definiscono gli integrali impropri su intervalli aperti a sinistra:

Definizione 7.11 La funzione $f :]a, b] \to \mathbb{R}$ integrabile su ogni intervallo chiuso $[a + \epsilon, b]$, con $\epsilon > 0$, si dice *integrabile (in senso improprio)* su $]a, b]$ se esiste finito il limite

$$\lim_{t \to a+} \int_t^b f(x)\,\mathrm{d}x$$

e in tal caso si pone

$$\int_a^b f(x)\,\mathrm{d}x = \lim_{t \to a+} \int_t^b f(x)\,\mathrm{d}x\,.$$

Esempio 125 Consideriamo la funzione

$$f(x) = \frac{1}{x^\gamma}, \qquad x \in]0,1]$$

dove $\gamma \in \mathbb{R}$. Si ha che

$$\int_t^1 \frac{1}{x^\gamma}\,\mathrm{d}x = \begin{cases} \dfrac{1}{-\gamma+1} - \dfrac{t^{-\gamma+1}}{-\gamma+1} & \gamma \neq 1\,, \\[2ex] -\ln t & \gamma = 1\,. \end{cases}$$

Per $t \to 0+$ l'espressione sopra diverge a $+\infty$ se $\gamma \geq 1$ mentre converge a $1/(1-\gamma)$ se $\gamma < 1$. Quindi, la funzione $1/x^\gamma$ è integrabile su $]0,1]$ se e soltanto se $\gamma < 1$ e in tal caso si ha

$$\int_0^1 \frac{1}{x^\gamma}\,\mathrm{d}x = \frac{1}{1-\gamma} \qquad (\gamma < 1)\,.$$

Nel caso in cui la funzione possa essere estesa per continuità nell'estremo b, l'integrale improprio si riduce ad un integrale ordinario, come mostra la seguente.

Proposizione 7.12 *Sia $f : [a,b[\to \mathbb{R}$ una funzione continua estendibile per continuità a sinistra nel punto b. Indichiamo con $\tilde{f}$ tale estensione; cioè $\tilde{f} : [a,b] \to \mathbb{R}$ è continua e $\tilde{f}(x) = f(x)$ per ogni $x \in [a,b[$. Allora si ha che f è integrabile su $[a,b[$ e si ha*

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^b \tilde{f}(x)\,\mathrm{d}x\,,$$

dove il secondo integrale è il normale integrale di Riemann di $\tilde{f}$ su $[a,b]$.

Dimostrazione Per ogni $t \in [a,b[$ abbiamo che

$$\begin{aligned} \left| \int_a^b \tilde{f}(x)\,\mathrm{d}x - \int_a^t f(x)\,\mathrm{d}x \right| &= \left| \int_a^b \tilde{f}(x)\,\mathrm{d}x - \int_a^t \tilde{f}(x)\,\mathrm{d}x \right| \\ &= \left| \int_t^b \tilde{f}(x)\,\mathrm{d}x \right| \\ &\leq \int_t^b |\tilde{f}(x)|\,\mathrm{d}x \\ &\leq (t-b) \max_{s\in[a,b]} |\tilde{f}(x)|\,. \end{aligned}$$

Poichè quest'ultima espressione tende a 0 per $t \to b-$, si ha che

$$\lim_{t \to b-} \int_a^t f(x)\,\mathrm{d}x = \int_a^b \tilde{f}(x)\,\mathrm{d}x\,.$$

Questo dimostra la tesi. ■

La teoria degli integrali impropri su intervalli si sviluppa in modo del tutto analogo alla teoria illustrata precedentemente sulle semirette. Riformuliamo nel seguito i risultati principali senza aggiungere ulteriori dimostrazioni. Consideriamo il caso di funzioni definite su intervalli aperti a destra, essendo il caso degli intervalli aperti a sinistra del tutto identico.

Proposizione 7.13 *Siano $f : [a,b[\to \mathbb{R}$ e $g : [a,b[\to \mathbb{R}$ due funzioni integrabili su $[a,b[$ e sia $\lambda \in \mathbb{R}$. Allora:*

(i) $f+g$ è integrabile su $[a,b[$ e si ha

$$\int_a^b (f+g)(x)\,\mathrm{d}x = \int_a^b f(x)\,\mathrm{d}x + \int_a^b g(x)\,\mathrm{d}x\,.$$

(ii) λf è integrabile su $[a,b[$ e si ha

$$\int_a^b \lambda f(x)\,\mathrm{d}x = \lambda \int_a^b f(x)\,\mathrm{d}x\,.$$

Proposizione 7.14 *Sia $f : [a,b[\to \mathbb{R}$ una funzione e sia $a < c < b$. Allora,*

(a) Se f è integrabile su $[a,b[$, essa è anche integrabile su $[c,b[$ e si ha

$$\int_a^b f(x)\,\mathrm{d}x = \int_a^c f(x)\,\mathrm{d}x + \int_c^b f(x)\,\mathrm{d}x\,. \tag{7.11}$$

(b) Se f è integrabile su $[a,c]$ e integrabile su $[c,b[$, allora f è integrabile su $[a,b[$ e vale ancora (7.11).

Come nel caso dell'integrazione su semirette, l'integrale improprio su intervalli finiti ha un comportamento più semplice per le funzioni a valori non-negativi. In effetti si supponga di avere una funzione $f : [a, b[\to \mathbb{R}$ integrabile su ogni intervallo chiuso $[a, b - \epsilon]$, con $\epsilon > 0$, tale che $f(x) \geq 0$ per ogni $x \in [a, b[$. Segue allora che

$$\int_a^t f(x)\,\mathrm{d}x, \quad t \in [a, b[$$

è una funzione crescente di t (più t cresce, più aumenta l'area positiva che sta tra il grafico e l'asse x). Dunque esiste sempre il limite per $t \to b-$, finito o eguale a $+\infty$. Più precisamente si ha

$$\lim_{t \to b-} \int_a^t f(x)\,\mathrm{d}x = \sup_{a \leq t < b} \int_a^t f(x)\,\mathrm{d}x\,.$$

Il limite quindi è finito se e soltanto se esiste $k \geq 0$ tale che

$$\int_a^t f(x)\,\mathrm{d}x \leq k, \quad \forall t \in [a, b[\,.$$

Nel caso in cui il limite sia $+\infty$ si usa comunque la notazione

$$\int_a^b f(x)\,\mathrm{d}x = +\infty$$

e si dice anche in questo caso che l'integrale improprio *diverge*. Anche per questo tipo di integrali si ha un importante criterio del confronto.

Teorema 7.15 *Siano $f : [a, b[\to \mathbb{R}$ e $g : [a, b[\to \mathbb{R}$ due funzioni integrabili su ogni intervallo chiuso $[a, b - \epsilon]$, con $\epsilon > 0$, tali che*

$$0 \leq f(x) \leq g(x), \quad \forall x \in [a, b[\,.$$

Allora se g è integrabile su $[a, b[$, lo è anche f e si ha

$$\int_a^b f(x)\,\mathrm{d}x \leq \int_a^b g(x)\,\mathrm{d}x\,.$$

Se invece f non è integrabile su $[a,b[$ (il suo integrale diverge), anche g non è integrabile su $[a,b[$.

Anche in questo caso il Teorema del confronto per gli integrali ammette un utile variante.

Teorema 7.16 *Siano $f:[a,b[\to\mathbb{R}$ e $g:[a,b[\to\mathbb{R}$ due funzioni integrabili su ogni intervallo chiuso $[a,b-\epsilon]$, con $\epsilon>0$, strettamente positive tali che esista finito il limite*

$$\lim_{x\to b-}\frac{f(x)}{g(x)}=l>0\,.$$

Allora f è integrabile su $[a,b[$ se e soltanto se g lo è.

Di nuovo questo implica in particolare che se due funzioni strettamente positive sono tra loro asintotiche per $x\to b$ allora l'una è integrabile in un intorno sinistro di b se e solo se lo è l'altra.

L'integrabilità assoluta in questo caso si definisce come segue:

Definizione 7.17 Una funzione $f:[a,b[\to\mathbb{R}$ si dice *integrabile assolutamente (in senso improprio)* su $[a,b[$ se è integrabile su ogni intervallo $[a,b-\epsilon]$ $(\epsilon>0)$ e se è integrabile in senso improprio su tale intervallo la funzione $|f|$, cioè se esiste finito

$$\lim_{t\to b-}\int_a^t |f(x)|\,\mathrm{d}x\,.$$

Vale il seguente

Teorema 7.18 *Se una funzione $f:[a,b[\to\mathbb{R}$ è assolutamente integrabile su $[a,b[$, allora è integrabile su detto intervallo e si ha*

$$\left|\int_a^b f(x)\,\mathrm{d}x\right|\le\int_a^b |f(x)|\,\mathrm{d}x\,. \tag{7.12}$$

Esempio 126 Sia

$$f(x)=\frac{1}{e^{-x}-1},\quad x\in[-1,0[\,.$$

Si noti che f è, sul dominio specificato, a valori positivi. Consideriamo ora $g(x) = -1/x$ sullo stesso dominio. Si ha che

$$\lim_{x\to 0-} \frac{f(x)}{g(x)} = \lim_{x\to 0-} \frac{-x}{e^{-x}-1} = 1\,.$$

Quindi, per il Teorema 7.16, poichè g non è integrabile su $[-1,0[$, non lo è neppure f.

Esempio 127 Sia

$$f(x) = \frac{\sin(1/x)}{\sqrt{x}}\,, \quad x \in]0,1]\,.$$

Si noti che

$$|f(x)| \le \frac{1}{\sqrt{x}}\,, \quad \forall x \in]0,1]\,.$$

Poichè la funzione $1/\sqrt{x}$ è integrabile su $]0,1]$ (vedi Esempio 125), per il Teorema del confronto 7.15, lo è anche $|f(x)|$. Quindi f è assolutamente integrabile su $]0,1]$.

Esercizio 7.7 Studiare l'integrabilità delle seguenti funzioni sui domini specificati:

$$f(x) = \sin(1/x)\ln x \quad x \in]0,1]\,, \qquad f(x) = \tfrac{\sin x}{x} \quad x \in]0,10]\,,$$

$$f(x) = \tfrac{1}{\sin x} \quad x \in]0,1]\,, \qquad f(x) = \tfrac{(\pi/2-x)^{3/2}}{1-\sin x} \quad x \in [0,\pi/2[\,.$$

R. (da sinistra a destra e dall'alto in basso) integrabile, integrabile, non integrabile, integrabile.

Capitolo 8

Equazioni differenziali ordinarie

8.1 Alcuni esempi

Il mondo fisico è governato da *leggi*. Per capire che cosa si intenda con questa frase pensate a quella di esse che è forse maggiormente nota, la seconda legge della dinamica

$$\mathbf{F} = m\mathbf{a}, \tag{8.1}$$

dove $\mathbf{F}$ è la forza che agisce su un corpo di massa m e $\mathbf{a}$ è l'accelerazione di tale corpo (sia $\mathbf{F}$ che $\mathbf{a}$ sono vettori nello spazio). L'equazione scritta si rivelerà utile se saremo capaci, assegnata la forza $\mathbf{F}$, di calcolare la *legge oraria* del corpo stesso, vale a dire di scrivere esplicitamente la funzione $\mathbf{y}(t)$ che assegna la posizione del corpo al tempo t, in funzione della posizione e della velocità iniziali del corpo stesso. Poiché questo è spesso molto difficile, potrà essere sufficiente a volte avere almeno informazioni qualitative su $\mathbf{y}(t)$.

Per capire come, *in linea di principio*, sia possibile ricavare la legge oraria, ricordiamo dalla Fisica il fatto ben noto che l'accelerazione di un corpo coincide con la derivata seconda della sua posizione:

$$\mathbf{a}(t) = \mathbf{y}''(t), \tag{8.2}$$

dove la derivata è ovviamente fatta rispetto al tempo, e che usualmente le forze più comunemente considerate nella meccanica classica (forza gravitazionale, forza elastica...) sono funzioni di posizione e velocità del corpo cui sono applicate. Ciò vuole dire che

$$\mathbf{F} = \mathbf{F}(\mathbf{y}(t), \mathbf{y}'(t))$$

e per capire cosa si intenda con questo pensate per esempio alla forza gravitazionale cui è soggetto un corpo puntiforme P di massa m a causa dell'attrazione esercitata da un corpo Q, supposto per semplicità fisso, puntiforme e di massa M: essa è data da

$$F(\mathbf{y}) = -\frac{GMm}{r^2}\frac{\mathbf{y}}{r},$$

dove $r := |\mathbf{y}|$ è la distanza di P da Q ($\mathbf{y}$ è orientato da Q a P), supposto quest'ultimo posto nell'origine del sistema di coordinate che si considera, e G è la costante di gravitazione universale. Naturalmente la posizione del punto P dipende dal tempo, e quindi la *legge del moto* $\mathbf{y}(t)$ del punto P soddisfa l'equazione

$$-\frac{GM}{|\mathbf{y}(t)|^2}\frac{\mathbf{y}(t)}{|\mathbf{y}(t)|} = \mathbf{y}''(t). \tag{8.3}$$

Un'equazione di questo tipo si chiama *equazione differenziale* perchè lega una funzione incognita $\mathbf{y}(t)$ con i valori di alcune delle sue derivate. In questo caso (come in tutti i casi che "provengono dalla seconda legge della dinamica) l'equazione si dice *del secondo ordine* perché la derivata di ordine più alto che compare è appunto la derivata seconda. In realtà in questo caso abbiamo a che fare con un *sistema* di equazioni differenziali, perche' prendendo le componenti lungo gli assi $\mathbf{e}_1, \mathbf{e_2}, \mathbf{e_3}$ dei vettori che compaiono in (8.3) si ottengono le equazioni

$$-GM\frac{y_1}{(y_1^2 + y_2^2 + y_3^2)^{3/2}} = y_1''$$

$$-GM\frac{y_2}{(y_1^2 + y_2^2 + y_3^2)^{3/2}} = y_2''$$

$$-GM\frac{y_3}{(y_1^2 + y_2^2 + y_3^2)^{3/2}} = y_3''.$$

Ci aspettiamo (e così è veramente) che, partendo da (8.3) e conoscendo posizione e velocità iniziali di P, sia possibile conoscere la posizione ad ogni tempo t del punto P. Svolgere questa operazione sarà possibile se saremo capaci a *trovare le soluzioni* dell'equazione differenziale data. Da tale equazione seguono in particolare le ben note leggi di Keplero della meccanica Newtoniana. Scopo di questa parte di corso è capire cosa significhi la frase, per ora un po' vaga, "trovare le soluzioni di un'equazione differenziale e, in alcuni casi particolarmente semplici, imparare a calcolare esplicitamente tali soluzioni. Che quest'ultimo obiettivo possa essere in generale molto difficile (o addirittura impossibile per via analitica) ce lo mostra l'esempio che abbiamo fatto poco fa: trovare un metodo generale che ci permetta di calcolare le soluzioni della (8.3) va al di là degli scopi di questo corso ed è comunque tutt'altro che semplice, sebbene possibile.

Vediamo ora qualche esempio più semplice e, soprattutto, *monodimensionale*, caso al quale quasi sempre ci restringeremo in questo corso. Considerate una particella di massa m, vincolata a muoversi su una retta, e legata attraverso una molla a un punto fisso O. La legge del moto e la cosiddetta *legge di Hooke* dell'elasticità lineare ci dicono che la posizione al tempo t della particella deve soddisfare l'equazione differenziale

$$-ky(t) = my''(t)$$

ad ogni tempo, dove k è la costante elastica della molla (essa ci dice in pratica quanto la molla sia "rigida). Potete facilmente convincervi calcolando esplicitamente le derivate che la funzione

$$y(t) = A\cos\left(\sqrt{\frac{k}{m}}t\right) + B\sin\left(\sqrt{\frac{k}{m}}t\right)$$

è soluzione dell'equazione differenziale data quali che siano le costanti arbitrarie reali A e B. È anche chiaro che se fissiamo posizione y_0 e velocità iniziali y'_0 del corpo, le costanti stesse vengono fissate univocamente in funzione di y_0, y'_0. Varie domande sono possibili:

- c'è qualche ragione per cui compaiano proprio dei coseni e dei seni, e quindi il moto sia *oscillatorio*?
- Come mai compaiono delle costanti arbitrarie?
- E poi: siamo proprio sicuri che non ci siano soluzioni di altro tipo all'equazione differenziale data?

Vedremo più avanti che è possibile rispondere a tutte queste domande.

Un'equazione differenziale più generale è la seguente:

$$ay''(t) + by'(t) + cy(t) = f(t),$$

che compare in vari esempi fisici. Eccone due.

- $a = m$, $b = c = 0$, $f(t) = -g$, dove g è l'accelerazione di gravità al suolo. L'equazione descrive allora il moto di un corpo di massa m in caduta libera presso la superficie terrestre. Le funzioni
$$y(t) = y_0 + y_1 t - \frac{1}{2} g t^2$$
sono soluzioni dell'equazione differenziale data per ogni valore delle costanti reali y_0 e y_1 (evidentemente collegate a posizione e velocità iniziale del corpo). Si può mostrare che non ve ne sono altre. Se b fosse diverso da zero si considererebbe un termine "di attrito (proporzionale alla velocità) per tener conto della resistenza dell'aria.

- $a = L$, $b = R$, $c = C^{-1}$, dove L, R, C sono rispettivamente l'induttanza, la resistenza e la capacità proprie di un circuito LRC con potenza applicata $f(t)$. $y(t)$ rappresenta allora l'intensità di corrente al tempo t. L'espressione esplicita della soluzione dipende dai valori di L, R e C e dall'espressione di f: vedremo più avanti come ricavarla.

Va notato che il linguaggio delle equazioni differenziali non è proprio solamente dei fenomeni fisici. Consideriamo per esempio un modello di crescita di una popolazione biologica in cui il numero $N(t)$ (non necessariamente intero!) di individui cresca proporzionalmente a N stesso e all'intervallo di tempo in cui si osserva la popolazione, con un tasso di accrescimento $\lambda(t)$ (in questo modello semplice λ non dipende da N, e l'equazione differenziale che se ne deduce è *lineare*). Allora l'equazione che descrive l'evoluzione della popolazione può essere dedotta dal fatto che
$$\Delta N(t) = \lambda(t) N(t) \Delta t,$$
cosicché, passando formalmente al limite quando Δt tende a zero si ha
$$N'(t) = \lambda(t) N(t).$$
La soluzione di questa equazione differenziale (*lineare del primo ordine*) è
$$N(t) = N(0) \exp\left[\int_0^t \lambda(s) \mathrm{d}s\right].$$
Si vede da questo esempio (ma anche dai precedenti) come le *costanti arbitrarie* che sempre compaiono nell'espressione della soluzione generale sono legate ai dati iniziali del problema, in questo caso al numero iniziale di individui della popolazione studiata.

8.2 Cos'è un'equazione differenziale?

Abbiamo appena visto che un'equazione differenziale consiste in una relazione che coinvolga una funzione incognita e alcune delle sue derivate. Per essere più precisi diamo una definizione, nella quale sarà usato il concetto di funzione di più variabili, che per il momento non abbiamo formalmente introdotto. Tuttavia potete facilmente immaginare di cosa si tratti per analogia con il caso delle funzioni di una sola variabile: ad esempio la funzione

$$f(x, y) = x^2 y$$

è una funzione definita su $\mathbb{R}^2$, cioè sulle *coppie* di numeri reali (x, y), ed è a valori reali. Si scrive in formule $f : \mathbb{R}^2 \to \mathbb{R}$.

Definizione 8.1 Si dice equazione differenziale una relazione del tipo

$$f(t, y, y', y'', \dots, y^{(n)}) = 0, \tag{8.4}$$

dove $f : \mathbb{R}^{n+2} \to \mathbb{R}$ è una funzione di $n + 2$ variabili, t una variabile indipendente e y è una funzione incognita della variabile t. In tal caso l'equazione viene detta di ordine n, dove n è l'ordine massimo delle derivate che compaiono nella relazione (8.4). Una funzione $y :]a, b[\to \mathbb{R}$ si dice *soluzione* dell'equazione differenziale data se essa è derivabile n volte e si ha

$$f(t, y(t), y'(t), y''(t), \dots, y^{(n)}(t)) = 0$$

per ogni $t \in]a, b[$.

Una soluzione di un'equazione differenziale è quindi una funzione della variabile indipendente t (che, negli esempi fatti nell'introduzione, è il tempo) tale che l'equazione in questione sia soddisfatta per ogni valore di t.

Vediamo ora con un esempio che non è sempre possibile risolvere un'equazione differenziale: *è possibile che non esistano soluzioni*!

Esempio 128 Si supponga di cercare soluzioni dell'equazione differenziale

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

dove $f : \mathbb{R} \to \mathbb{R}$ è una funzione con una discontinuità di prima specie in un punto x_0, ad esempio la funzione a gradino

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

vista nell'Osservazione fatta nella Sezione 1.1. Trovare una soluzione all'equazione differenziale data è la stessa cosa che cercare una *primitiva* di f: ma ben sappiamo che le funzioni con discontinuità a salto *non ammettono mai primitive* (si veda l'Osservazione dopo il Teorema 1.1). Questo perché è possibile dimostrare che la derivata di una funzione derivabile f non può ammettere discontinuità di prima specie, quale che sia la funzione f. Possiamo sperare che aggiungere qualche condizione di continuità sul secondo membro possa eliminare questo problema, e in effetti così è, come vedremo in seguito.

Nel prosieguo, vedremo vari casi speciali di equazioni differenziali. Vale la pena di elencare alcuni di questi casi già ora.

Definizione 8.2

- Un'equazione differenziale di ordine n della forma (8.4) si dice *autonoma* se la funzione f non dipende da t, cioè se
$$f(t, y, y', y'', \ldots, y^{(n)}) = g(y, y', y'', \ldots, y^{(n)})$$
per un'opportuna funzione $g : \mathbb{R}^{n+1} \to \mathbb{R}$*;*

- un'equazione differenziale di ordine n si dice *in forma normale* se è della forma
$$y^{(n)} = g(t, y, y', \ldots, y^{(n-1)});$$

- un'equazione differenziale di ordine n della forma (8.4) si dice *lineare* se f è un polinomio di primo grado nelle variabili $y, y', \ldots, y^{(n)}$, cioè se essa si può scrivere come
$$a_0(t)y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y = b(t).$$

Come abbiamo visto dagli esempi della prima sezione, le soluzioni di un'equazione differenziale, *quando esistono*, sono in generale molte: in effetti gli esempi suggeriscono la dipendenza di queste soluzioni da opportune *costanti arbitrarie* e gli esempi stessi mostrano come tali condizioni sembrino legate, per esempio, alle condizioni iniziali del problema che si sta considerando (se esse sono note).

Assegnare congiuntamente un'equazione differenziale e opportuni dati iniziali dovrebbe quindi portare a porre un problema matematico risolubile. Malgrado ciò non sia purtroppo sempre vero è comunque opportuno dare una definizione basilare per il seguito:

Definizione 8.3 *Si dice problema di Cauchy, o dei valori iniziali, il sistema costituito da un'equazione differenziale di ordine n, in forma normale, e da n dati iniziali al tempo t_0:*

$$\begin{aligned}
&y^{(n)} = g(t, y, y', \ldots, y^{(n-1)}),\\
&y(t_0) = y_0\\
&y'(t_0) = y_1\\
&y''(t_0) = y_2\\
&\vdots\\
&y^{(n-1)}(t_0) = y_{n-1}.
\end{aligned} \tag{8.5}$$

Quello che speriamo di poter dimostrare, imponendo oltre alla validità dell'equazione differenziale considerata, anche un numero di *condizioni iniziali* esattamente pari all'ordine dell'equazione, è che il problema di Cauchy considerato abbia esattamente una soluzione. Negli esempi fisici visti all'inizio questo deve essere vero: se assegnamo le forze agenti sul corpo considerato e, in aggiunta, conosciamo posizione e velocità iniziali del corpo stesso, la legge oraria del corpo dovrebbe essere univocamente determinata. Si noti che in questo esempio l'equazione considerata è del secondo ordine, e le condizioni iniziali che è necessario conoscere sono appunto due.

8.3 Un cenno all'analisi in due variabili

Nella sezione successiva ci saranno indispensabili alcuni concetti che tradizionalmente non vengono insegnati durante i corsi del primo anno. Tuttavia senza un cenno a questi non saremmo in grado di capire quanto seguirà.

Cominciamo col prendere il piano cartesiano. I punti del piano sono coppie di numeri reali del tipo

$$P \equiv (x_1, x_2),\ x_1, x_2 \in \mathbb{R}.$$

Per questa ragione si è soliti denominare con $\mathbb{R}^2$ il piano stesso: dunque l'insieme $\mathbb{R}^2$ è costituito dalle coppie $x = (x_1, x_2)$ di numeri reali. I numeri reali x_1 e x_2 si dicono *componenti cartesiane* del punto x. Esattamente come nel caso di funzioni di una variabile visto finora, potremo definire le funzioni di due variabili: una funzione sarà per noi una regola che ci permette di associare al valore assunto da una coppia di variabili indipendenti $(x, y) \in \mathbb{R}^2$, un numero reale univocamente determinato. Scriveremo

$$f : \mathbb{R}^2 \to \mathbb{R}.$$

La prima fondamentale proprietà che vogliamo discutere è la *continuità di una funzione di due variabili*. Prima di fare ciò ricordiamo il seguente ben noto fatto: sul piano è definita una distanza naturale tra due punti $x, y \in \mathbb{R}^2$ di componenti $(x_1, x_2), (y_1, y_2)$ rispettivamente:

$$|x - y| = \sqrt{(x_1 - y_1)^2 + (x_2 - y_2)^2}.$$

Questa formula altro non è che una versione del Teorema di Pitagora: notate che, per evitare di introdurre troppi simboli, la distanza nel piano è denotata con $|x - y|$ esattamente come la distanza tra numeri reali (quest'ultima altro non è naturalmente che il modulo della loro differenza): ma non si tratta ovviamente della medesima operazione, perchè l'una agisce su coppie di punti del piano, l'altra su coppie di numeri reali.

A questo punto ricordiamoci il significato del concetto di continuità per una funzione di una variabile: $f : \mathbb{R} \to \mathbb{R}$ si dice continua in $\overline{x} \in \mathbb{R}$ se, informalmente parlando, i valori di $f(x)$ "distano poco da quelli di $f(\overline{x})$ se x "dista poco da $\overline{x}$. Il concetto di distanza qui usato è quello di cui disponiamo sull'insieme dei numeri reali. Abbiamo però appena ricordato che vi è una distanza naturale anche sull'insieme dei punti del piano. Possiamo dunque dare una definizione analoga:

Definizione 8.4 Una funzione $f : \mathbb{R}^2 \to \mathbb{R}$ si dice continua in un punto $\overline{x}$ se, per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che, se $|x - \overline{x}| \leq \delta$, allora $|f(x) - f(\overline{x})| \leq \varepsilon$. la funzione si dice continua su un insieme $A \subset \mathbb{R}^2$ se essa è continua in ogni punto di tale insieme.

Se provate a riguardare la definizione di continuità per funzioni di una variabile vi accorgerete che essa è *identica* a quella appena data: la differenza sta solo nel diverso significato del simbolo "$|x - y|$ a seconda che x e y siano numeri reali o punti del piano.

Osservazione: il tema della continuità per le funzioni di più variabili richiederebbe una trattazione estesa almeno quanto quella relativa alle funzioni di una variabile. Non essendo però questo tra gli scopi del corso, ci limitiamo a segnalare che larghe classi di funzioni di due variabili sono continue: ad esempio i polinomi in due variabili del tipo

$$P(x, y) = \sum_{i,j=1}^{k} a_{i,j} x^i y^j$$

sono funzioni continue, così come sono continue (ove definite) funzioni del tipo P qui sopra ma in cui compaiano potenze anche non intere, ma positive, di x e y. Un'analogo del Teorema della continuità della funzione composta mostra anche che funzione composte quali ad esempio

$$h(x, y) = \sin(x^2 y^3 - xy^2 + y)$$

sono anch'esse funzioni continue (qui e in seguito useremo le coordinate cartesiane per identificare un punto del piano). Ciò è vero perché la funzione $f : \mathbb{R}^2 \to \mathbb{R}$ data da $f(x, y) = x^2 y^3 - xy^2 + y$ è continua, così come è continua la funzione $g : \mathbb{R} \to \mathbb{R}$ data da $g(w) = \sin w$. Poiché $h = g \circ f$ allora il Teorema cui si accennava sopra mostra che anche h, funzione di due variabili, è continua.

Il prossimo concetto che discuteremo brevemente è quello di *derivata parziale*. L'idea è veramente semplice: prendete una funzione di due variabili e fissate il valore di una delle due, diciamo la x. A questo punto avrete di fatto a che fare con una funzione di una sola variabile, la y, e potrete fare la derivata, se essa esiste, di tale funzione. Tale derivata si chiama derivata parziale rispetto alla y. Poiché la derivata di una funzione di una variabile è definita come il limite del rapporto incrementale, daremo la seguente definizione.

Definizione 8.5 *Sia $f : \mathbb{R}^2 \to \mathbb{R}$ una funzione di due variabili, e sia $(\overline{x}, \overline{y})$ un punto fissato del piano. Si dice derivata parziale di f rispetto a x nel punto $(\overline{x}, \overline{y})$ il numero*

$$\frac{\partial f}{\partial x}(\overline{x}, \overline{y}) := \lim_{h \to 0} \frac{f(\overline{x} + h, \overline{y}) - f(\overline{x}, \overline{y})}{h},$$

qualora il limite esista finito. Analogamente, si dice derivata parziale di f rispetto a y nel punto $(\overline{x}, \overline{y})$ il numero

$$\frac{\partial f}{\partial y}(\overline{x}, \overline{y}) := \lim_{h \to 0} \frac{f(\overline{x}, \overline{y} + h) - f(\overline{x}, \overline{y})}{h},$$

qualora il limite esista finito.

Fare una derivata parziale non è più facile o più difficile di fare una derivata usuale. Basta fissare la variabile rispetto alla quale non si deriva e fare la derivata richiesta guardando solo alla dipendenza dalla variabile rispetto alla quale si deriva. In particolare l'esistenza delle derivate parziali si riduce all'esistenza di una derivata usuale, e quindi si tratta di un problema che sappiamo già affrontare. Vediamo un esempio elementare.

Esempio 129 Consideriamo la funzione vista prima, $h(x,y) = \sin(x^2y^3 - xy^2 + y)$. La sua derivata rispetto alla variabile x vale

$$\frac{\partial h}{\partial x}(x,y) = (2xy^3 - y^2)\cos(x^2y^3 - xy^2 + y).$$

La derivata rispetto alla variabile y vale invece

$$\frac{\partial h}{\partial x}(x,y) = (3x^2y^2 - 2xy + 1)\cos(x^2y^3 - xy^2 + y).$$

L'esistenza di entrambe le derivate non deve essere in questo caso verificata direttamente dalla definizione, ma può essere dedotta da quanto già ben sappiamo sulla derivabilità delle funzioni elementari di una variabile quali polinomi, funzioni trigonometriche e così via.

Osservazione: Sappiamo bene che, per funzioni di una variabile, la derivabilità implica la continuità. Purtroppo così non è per le funzioni di più variabili: esistono funzioni che ammettono in un fissato punto tutte le derivate parziali ma non sono continue. Un esempio semplice è dato dalla funzione

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

Non è difficile convincersi che la funzione non è continua nel punto $(0,0)$. In effetti, se ci si avvicina all'origine sulla retta $x = y$ si ha che la funzione vale sempre 1, e quindi i valori di $f(x,y)$ non si avvicinano necessariamente a $f(0,0) = 0$ quando (x,y) si avvicina a $(0,0)$. Peraltro

$$f(x,0) = 0 \ \ \forall x \neq 0, \quad f(0,y) = 0 \ \ \forall y \neq 0.$$

Per definizione di derivata parziale si ha dunque

$$\begin{aligned} \frac{\partial f}{\partial x}(0,0) \ &:= \lim_{h\to 0} \frac{f(h,0) - f(0,0)}{h} \\ &= \lim_{h\to 0} \frac{0}{h} = 0, \end{aligned}$$

così come $(\partial f/\partial y)(0,0)$. Si può dimostrare che questo non può accadere se le derivate parziali stesse sono a loro volta funzioni continue nel punto in cui si lavora.

L'ultimo concetto che dobbiamo discutere in questa sezione, per ragioni che saranno chiare tra poco, è quello della relazione tra Lipschitzianità di una funzione ed esistenza delle derivate parziali. Abbiamo visto nella Sezione 5.1 delle note di Calcolo I che una funzione con derivata continua in un intervallo è necessariamente Lipschitziana. Un fatto analogo vale anche per funzioni di due variabili, ma sono coinvolte le derivate parziali. Diamo prima alcune definizioni: parleremo di *rettangolo chiuso* nel piano riferendoci a insiemi della forma
$$R := \{(x,y) \in \mathbb{R}^2 : x \in [a,b], y \in [c,d]\}$$
con a,b,c,d numeri reali (*non* consideriamo quindi il caso in cui qualcuno di essi sia $+\infty$ o $-\infty$: questo è molto importante per la sezione successiva). L'aggettivo "chiuso si riferisce, senza entrare nelle motivazione matematiche più dettagliate di tale terminologia, al fatto che si tratta di un rettangolo che "contiene il proprio bordo. Parleremo di rettangolo *aperto* per descrivere un insieme del tipo
$$S = \{(x,y) \in \mathbb{R}^2 : x \in]a,b[, y \in]c,d[\}.$$

Definizione 8.6 Si dice che una funzione $f : \mathbb{R}^2 \to \mathbb{R}$ è Lipschitziana rispetto a y, uniformemente in x, in un rettangolo (chiuso o aperto) R se accade che
$$|f(x_1,y_1) - f(x_2,y_2)| \leq A|y_1 - y_2| \tag{8.6}$$
per ogni coppia di punti (x_1,y_1), (x_2,y_2) entrambi in R, dove $A > 0$ è un'opportuna costante.

Si noti in questa definizione il fatto che A debba essere una costante, non debba in particolare dipendere da x_1, x_2: è per questo che si dice che la Lipschitzianità vale *uniformemente in x*. Naturalmente il ruolo di x e di y potrebbe anche essere scambiato e si potrebbe parlare di funzioni Lipschitziane in x uniformemente rispetto alla y, ma con la convenzioni che useremo nelle prossime sezioni non ci sarà necessario discutere esplicitamente quest'ultimo caso.

Vale il seguente risultato, che non possiamo dimostrare mancandoci vari strumenti di analisi in più variabili. Esso ci sarà utile nella prossima sezione.

Proposizione 8.7 *Si supponga che $f : \mathbb{R}^2 \to \mathbb{R}$ sia una funzione che ammette derivata parziale, rispetto alla variabile y, continua in un un rettangolo chiuso R. Allora f è Lipschitziana in y, uniformemente rispetto alla x, nel rettangolo R.*

8.4 Esistenza e unicità

Abbiamo visto finora che cos'è un'equazione differenziale e cosa sia una sua soluzione. Ci siamo però finora astenuti dal discutere i problemi fondamentali relativi a un'equazione differenziale. Schematicamente le domande cui cercheremo di rispondere sono le seguenti:

- è possibile dare condizioni sull'equazione differenziale che vogliamo studiare sotto le quali un problema di Cauchy ad essa associato abbia senz'altro una soluzione?
- Nei casi nei quali la soluzione esiste, essa è unica?
- Sempre quando la soluzione esiste, sappiamo dire qualcosa sull'intervallo sul quale è definita?

Vedremo nel seguito che sarà possibile rispondere a queste domande, almeno per equazioni differenziali del primo ordine in forma normale. Vedremo in particolare che, per un problema di Cauchy della forma

$$\begin{cases} y' = f(t, y) \\ y(t_0) = y_0 \end{cases}$$

alcune possibili risposte sono le seguenti:

- se f è una funzione continua in un rettangolo contenente il punto (t_0, y_0), allora esiste almeno una soluzione al problema di Cauchy dato;
- se, in più, f è continua in t e Lipschitziana in y, uniformemente rispetto a t, allora la soluzione è unica;
- se inoltre f cresce, come funzione di y, non più che linearmente e ciò accade uniformemente rispetto a t, cioè

  $$|f(t, y)| \leq A|y| + b \quad \forall t, y \in \mathbb{R},$$

  allora la soluzione è definita per tutti i tempi.

Faremo anche vedere che, per ciascuna delle tre questioni poste, rinunciare alle ipotesi richieste fa cadere la conclusione: ad esempio già sappiamo che se f non è continua vi sono casi in cui l'equazione differenziale considerata non ammette alcuna soluzione.

Risponderemo in questa sezione alle prime due domande. Non saremo in grado di fare alcuna dimostrazione perché non ne abbiamo a disposizione gli strumenti tecnici. Ci accontenteremo degli enunciati e di proporre alcuni controesempi per mostrare che le ipotesi fatte sono realmente essenziali. Cominciamo con un Teorema essenziale, il *Teorema di esistenza e unicità locale*, dovuto al grande matematico francese L. Cauchy. Avremo bisogno di una definizione preliminare.

Definizione 8.8 *Sia* $f\ :]a,b[\to \mathbb{R}$ *una funzione. Essa si dice di classe* $C^k(]a,b[)$ *(*$k \geq 1$*) se è derivabile* k *volte in* $]a,b[$ *e le sue derivate sono funzioni continue. Si dice di classe* $C^\infty(]a,b[)$ *se essa è derivabile infinite volte in* $]a,b[$.

Si noti che, affinché una funzione sia di classe $C^k(]a,b[)$ basta che sia derivabile k volte e che la derivata k–esima sia una funzione continua. Questo perché una funzione derivabile è anche continua.

Teorema 8.9 *Si consideri il problema di Cauchy*

$$\begin{cases} y' = f(t,y) \\ y(t_0) = y_0 \end{cases} \tag{8.7}$$

dove si suppone che f *sia definita su un rettangolo aperto* R*,* $f : R \subset \mathbb{R}^2 \to \mathbb{R}$*, che* R *contenga il punto* (t_0, y_0) *e che* f *sia una funzione soddisfacente le seguenti ipotesi:*

- f *è continua su* R*;*
- f *è Lipschitziana in* y*, uniformemente rispetto alla variabile* t*, in* R*.*

Allora il problema di Cauchy (8.7) ammette esattamente una soluzione $y(t)$ *definita su un intervallo del tipo* $]t_0-\varepsilon, t_0+\varepsilon[$*. Tale soluzione è di classe* C^1 *in tale intervallo.*

Il Teorema precedente è fondamentale. Esso garantisce l'esistenza e l'unicità per il problema di Cauchy assegnato, sotto ipotesi non troppo forti su f. Va notato però che il Teorema stesso non dice nulla di preciso sull'intervallo sul quale la soluzione è definita. Questo non è casuale perché, come vedremo tra poco, se anche f fosse una funzione regolare *su tutto il piano*, potrebbe darsi che la soluzione del corrispondente problema di Cauchy sia definita solo su un intervallo, e non su tutto $\mathbb{R}$.

Un corollario immediato di questo Teorema fa uso del concetto di derivata parziale: le ipotesi sono un po' più forti, ma di verifica più immediata.

Corollario 8.10 *Si supponga che, nel teorema 8.9, l'ipotesi di Lipschitzianità della funzione f sia sostituita dalla seguente:*

- *esiste in R la derivata parziale di f rispetto a y, e tale funzione è continua in R.*

Allora valgono le conclusioni del Teorema 8.9.

Osservazione: il fatto che la soluzione $y(t)$ la cui esistenza è stabilita nel Teorema 8.9 sia di classe C^1 è immediato vista l'ipotesi di continuità di f. In effetti per definizione di soluzione la funzione y deve essere derivabile e la sua derivata $y'(t)$, per l'equazione differenziale che è supposta essere soddisfatta, coincide con $f(t, y(t))$ che è una funzione continua in quanto composizione di funzioni continue. Se f ha proprietà di regolarità aggiuntive lo stesso è vero per la soluzione.

Come detto non potremo dimostrare il Teorema di esistenza e unicità; possiamo però accennare all'idea della dimostrazione, contenuta nella prossima Proposizione.

Proposizione 8.11 *Si consideri il problema di Cauchy (8.7), e si supponga che le ipotesi del Teorema 8.9 siano verificate. Si considerino, al variare di $n \in \mathbb{N}$, le funzioni definite per ricorrenza come segue:*

$$\begin{cases} y_0(t) = y_0 \\ y_n(t) = y_0 + \int_{t_0}^{t} f(s, y_{n-1}(s))\mathrm{d}s \quad \text{se } n \geq 1 \end{cases} \tag{8.8}$$

definite per $t \in [t_0 - \varepsilon, t_0 + \varepsilon]$ con ε sufficientemente piccolo. Allora, per ogni tale t, si ha

$$\lim_{n \to \infty} y_n(t) = y(t),$$

dove $y(t)$ è l'unica soluzione al problema di Cauchy assegnato. Inoltre la convergenza è uniforme su $[t_0 - \varepsilon, t_0 + \varepsilon]$ nel senso che

$$\sup_{t \in [t_0-\varepsilon, t_0+\varepsilon]} |y_n(t) - y(t)| \to 0$$

quando $n \to +\infty$.

Il procedimento sopra descritto è *costruttivo.* Esso inoltre permette di implementare un metodo approssimato di soluzione di problemi di Cauchy del primo ordine. Vediamo come in un esempio semplice.

Esempio 130 Si consideri il problema di Cauchy

$$\begin{cases} y' = y \\ y(3) = -2. \end{cases}$$

Non è troppo difficile immaginare la forma delle soluzioni, perché ben sappiamo che la derivata della funzione $y(x) = e^x$ coincide con la funzione stessa, e quindi c'è da aspettarsi che la soluzione del problema di Cauchy dato (unica, perché le condizioni del teorema di esistenza e unicità sono tutte soddisfatte), abbia a che fare con un esponenziale. Vediamo cosa ci dice la proposizione precendente. Poniamo quindi $y_0(t) = -2$. Si ha allora

$$y_1(t) = -2 + \int_3^t (-2)\mathrm{d}s = -2 - 2(t-3).$$

Analogamente:

$$\begin{array}{l} y_2(t) = -2 + \int_3^t [-2 - 2(s-3)]\mathrm{d}s = -2 - 2(t-3) - (t-3)^2 \\ y_3(t) = -2 + \int_3^t [-2 - 2(s-3) - (s-3)^2]\mathrm{d}s \\ = -2 - 2(t-3) - (t-3)^2 - \dfrac{(t-3)^3}{3}. \end{array}$$

Non è difficile dimostrare per induzione che, per ogni $n \in \mathbb{N}$ si ha:

$$y_n(t) = -2 \sum_{k=0}^{n} \frac{(t-3)^k}{k!}.$$

A destra nell'ultima equazione riconoscerete le somme parziali della serie esponenziale, quindi, per $n \to +\infty$ si ha, per definizione di somma di una serie:

$$y_n(t) \to y(t) := -2e^{t-3}$$

per ogni $t \in \mathbb{R}$. La funzione $y(t)$ è quindi l'unica soluzione del problema di Cauchy dato. Si noti che essa è definita per tutti i tempi t: vedremo tra non molto che anche questo poteva essere previsto a priori.

Che questa fosse la soluzione non è peraltro, come già detto, sorprendente. Il punto è che il metodo usato qui in questo caso semplice è *generale*, e mette quindi a disposizione un metodo approssimato di soluzione anche nei numerosissimi casi in cui non si è in grado di trovare una soluzione esplicita.

Concludiamo questa sezione con un risultato dall'enunciato molto semplice, ma dalla dimostrazione tutt'altro che elementare (e quindi omessa): il Teorema di Peano.

Teorema 8.12 *Si consideri il problema di Cauchy (8.7), dove si suppone che f sia definita su un rettangolo aperto R, $f : R \subset \mathbb{R}^2 \to \mathbb{R}$, che R contenga il punto (t_0, y_0) e che f sia una funzione soddisfacente solamente la seguente ipotesi:*

- *f è continua su R;*

Allora il problema di Cauchy (8.7) ammette almeno una soluzione $y(t)$ definita su un intervallo del tipo $]t_0 - \varepsilon, t_0 + \varepsilon[$.

Qual'è la differenza tra il Teorema di Cauchy e quello di Peano? Nel secondo le ipotesi sono più deboli, perché viene meno l'ipotesi di Lipschitzianità della f. Ma sono più deboli anche le conclusioni, perché non si dice nulla sull'*unicità* della soluzione. In effetti l'unicità può venir meno, come mostra l'esempio seguente.

Esempio 131 Si consideri il problema di Cauchy

$$\begin{cases} y' = y^{1/3} \\ y(0) = 0. \end{cases}$$

La funzione $f(x, y) = y^{1/3}$ è certamente continua su tutto $\mathbb{R}^2$: essa è anche Lipschitziana in un rettangolo che contenga l'origine? La risposta è no. In effetti se lo fosse dovrebbe succedere che, per ogni y_1, y_2 in tale rettangolo si abbia

$$|y_1^{1/3} - y_2^{1/3}| \le k|y_1 - y_2|$$

per un opportuno $k > 0$. Questo non può essere: infatti, ponendo $y_2 = 0$, dovrebbe anche valere, per ogni y_1 con $|y_1| < \varepsilon$ opportuno,

$$|y_1^{1/3}| \le k|y_1|,$$

cioè

$$|y_1|^{-2/3} \le k.$$

è però chiaro che nessuna scelta di k può rendere vera quest'ultima disequazione per ogni y_1 vicino a zero. Dunque f non è Lipschitziana vicino all'origine, e quindi le ipotesi del Teorema di Cauchy non sono soddisfatte. Lo sono però quelle del Teorema di Peano: quindi avremo senz'altro una soluzione al problema di Cauchy dato, ma ve ne potrebbero però anche essere molte.

Vedremo più avanti come risolvere equazioni differenziali del tipo qui considerato (le cosiddette *equazioni differenziali a variabili separabili*). Ci accontentiamo qui di notare (lo potete verificare direttamente) che le seguenti funzioni sono *tutte* soluzioni del problema di Cauchy considerato:

$$
\begin{aligned}
&y_0(t) = 0 \quad \forall t \geq 0 \\
&y_{1,c}(t) = \begin{cases} 0 & \text{se } t \leq \dfrac{3}{2}c \\ \left[\dfrac{2t}{3} - c\right]^{3/2} & \text{se } t > \dfrac{3}{2}c \end{cases} \\
&y_{2,c}(t) = \begin{cases} 0 & \text{se } t \leq \dfrac{3}{2}c \\ -\left[\dfrac{2t}{3} - c\right]^{3/2} & \text{se } t > \dfrac{3}{2}c \end{cases}
\end{aligned}
$$

quale che sia il numero reale c. Le soluzioni sono quindi infinite. Notate che le soluzioni $y_{1,c}$ e $y_{2,c}$ sono definite "a pezzi. Tuttavia esse si saldano bene nel punto di contatto tra i due rami da cui sono costituite, nel senso che la funzione risultante è una funzione derivabile. La figura 8.1 mostra alcune di queste soluzioni.

Notiamo inoltre che, se considerassimo un dato di Cauchy della forma $y(0) = y_0$ con $y_0 \neq 0$, la funzione f sarebbe Lipschitziana vicino a y_0 (ad esempio perché la sua derivata, $y^{-2/3}/3$, è limitata se y è vicino a $y_0 \neq 0$). Quindi le ipotesi del Teorema di Cauchy sono verificate e la soluzione, oltre a esistere, sarebbe anche unica.

8.5 Esistenza in grande

Abbiamo visto finora condizioni, sul secondo membro di un'equazione differenziale del primo ordine in forma normale, che ci assicurino l'esistenza ed eventualmente l'unicità della soluzione a un problema di Cauchy associato all'equazione data. Come già abbiamo fatto notare, se anche tali condizioni sono soddisfatte non si può dire nulla su quale sia l'intervallo di definizione della soluzione stessa, *anche se il secondo membro è regolare.*

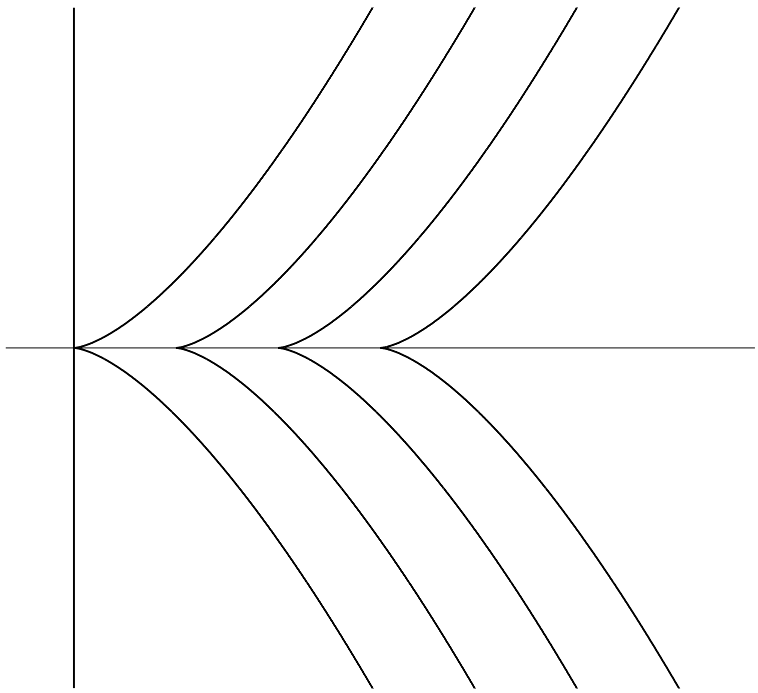


Figura 8.1:

Osserviamo ora che vale il seguente risultato, peraltro non troppo sorprendente (ed enunciato in ipotesi un po' più forti del necessario):

Proposizione 8.13 *Sia R un rettangolo chiuso in $\mathbb{R}^2$, (t_0, y_0) un punto di R che non stia sul bordo di R, e sia $f : R \to \mathbb{R}^2$ una funzione continua su R insieme alla sua derivata parziale $\partial f/\partial y$. Allora la soluzione del problema*

$$\begin{cases} y' = f(t, y) \\ y(x_0) = y_0 \end{cases}$$

che esiste senz'altro per tempi vicini a t_0 per il Teorema di Cauchy, può essere estesa fino ad arrivare a toccare il bordo del rettangolo R.

Osservazione. Il punto da capire è che la soluzione potrebbe raggiungere i bordi *laterali* del rettangolo, e in questo caso si può pensare di poter ulteriormente prolungare la soluzione, considerando un nuovo problema di Cauchy con dato iniziale nel punto del bordo del rettangolo che si è raggiunto (se le ipotesi del Teorema di Cauchy sono verificate vicino a quel punto) ed eventualmente proseguendo nel procedimento: ma la soluzione potrebbe anche raggiungere il bordo *superiore* o *inferiore* del rettangolo e, almeno in linea di principio, se questo dovesse ripetersi quando si cerca di prolungare ulteriormente la soluzione considerando nuovi problemi di Cauchy, potrebbe anche succedere che la soluzione tenda all'infinito a un tempo t finito. Intuitivamente un fenomeno del genere può accadere quando la derivata y' della soluzione è (in modulo) grande e cioè, per l'equazione differenziale che y deve soddisfare, quando $f(t, y)$ cresce rapidamente.

In effetti una situazione del genere si può verificare anche in casi semplicissimi, come mostra il seguente esempio.

Esempio 132 Si consideri il problema di Cauchy

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

Si tratta di un'equazione differenziale all'apparenza innocua, e in effetti essa è facilmente risolubile (si tratta nuovamente di un'equazione a variabili separabili, che impareremo ad affrontare a breve). Si può in effetti mostrare che la soluzione è la seguente:

$$y(t) = \frac{1}{1-t} \quad \forall t < 1.$$

Che questa sia l'unica soluzione del problema di Cauchy dato segue dal fatto che tutte le ipotesi del Teorema di Cauchy sono soddisfatte: $f(t, y) = y^2$ è infatti ovunque una funzione regolare, in particolare ovunque continua e con derivate parziali ovunque continue. Tuttavia si ha

$$\lim_{t \to 1^-} y(t) = +\infty$$

e quindi non c'e' alcun modo di prolungare la soluzione oltre al "tempo $t = 1$, perchè sappiamo che la soluzione di un'equazione differenziale deve essere una funzione derivabile, dunque continua. Peraltro *la stessa funzione* $y(t)$, definita però per $t > 1$, cioè su un intervallo illimitato dall'alto, è anche soluzione (ad esempio) del problema

$$\begin{cases} y' = y^2 \\ y(2) = -1, \end{cases}$$

mentre prima la soluzione era definita solo su un insieme di tempi limitato dall'alto. Non sembra quindi esserci alcun modo semplice per prevedere, dato un problema di Cauchy, dove sia definita l'eventuale soluzione: abbiamo inoltre appena visto l'importanza del dato iniziale nel problema.

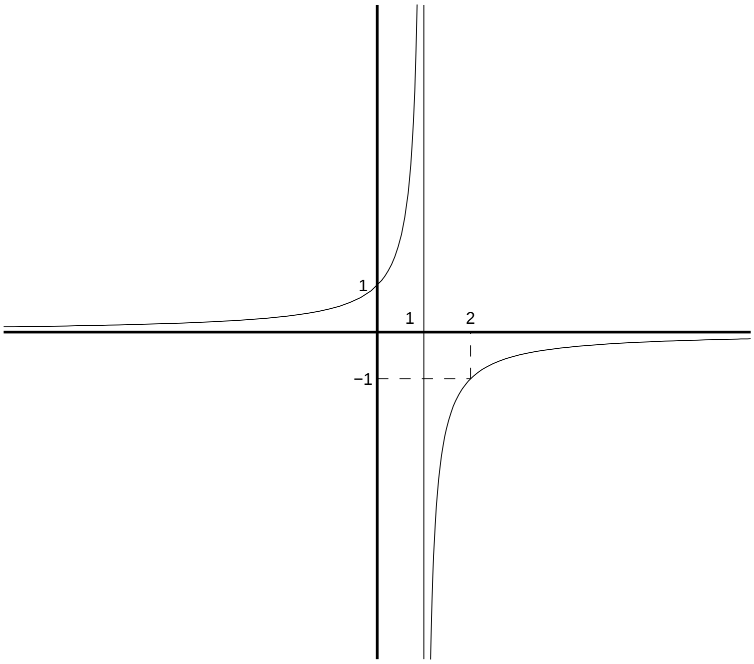


Figura 8.2:

Che un fenomeno di questo tipo abbia luogo per un'equazione differenziale del tipo visto sopra fa capire che il problema *non è nella regolarità della funzione* f che compare a secondo membro dell'equazione differenziale: $f(t, y) = y^2$ è quanto di più regolare si possa immaginare. Vediamo nel prossimo Teorema che l'ipotesi che serve è di tutt'altro tipo: f non deve crescere troppo rapidamente come funzione di y.

Teorema 8.14 *Si consideri il problema di Cauchy*

$$\begin{cases} y' = f(t, y) \\ y(t_0) = y_0 \end{cases}, \tag{8.9}$$

dove si suppone che f sia definita sulla striscia

$$S := \{(t, y) \in \mathbb{R}^2 : t \in [a, b], y \in \mathbb{R}\},$$

con $t_0 \in]a, b[$ e $f : S \subset \mathbb{R}^2 \to \mathbb{R}$. Si supponga inoltre che f sia una funzione continua su S e che, su ogni rettangolo chiuso

$$R = \{(t, y) \in \mathbb{R}^2 : t \in [\alpha, \beta], y \in [\gamma, \delta]\}$$

con $a < \alpha < \beta < b$ la funzione f sia Lipschitziana in y, uniformemente rispetto alla t, così che la soluzione a ogni problema di Cauchy del tipo (8.9) esista e sia unica. Si supponga inoltre che valga la seguente condizione aggiuntiva: esistono costanti non negative c_1 e c_2 tali che

$$|f(t,y)| \leq c_1 + c_2|y| \tag{8.10}$$

per ogni $(t,y) \in S$. Allora la soluzione al problema di Cauchy assegnato è definita su $[a,b]$.

Ecco quindi la condizione che cercavamo: se f cresce, come funzione di y, non più che linearmente (cioè non più velocemente di una retta), allora la soluzione è definita sul massimo intervallo "temporale possibile. Se si riesce ad applicare il Teorema precedente per *ogni* scelta di $[a,b]$ se ne potrà in particolare dedurre l'esistenza per *ogni* valore della variabile t.

Esempio 133 Il problema di Cauchy

$$\begin{cases} y' = \sqrt{y^2 + \dfrac{1}{1+t^2}} \\ y(0) = 1 \end{cases}$$

non è risolubile esplicitamente. Possiamo però dimostrare che la soluzione al problema dato esiste, è unica ed è definita sull'intera retta reale.

Per prima cosa la funzione

$$f(t,y) = \sqrt{y^2 + \frac{1}{1+t^2}}$$

è continua su tutto $\mathbb{R}^2$, in quanto composta di funzioni continue. Quindi una soluzione al problema di Cauchy esiste, per il Teorema di Peano. Per vedere che f soddisfa una condizione di Lipschitzianità potremmo anche procedere con la verifica della definizione, ma è più facile calcolare la derivata parziale

$$\frac{\partial f}{\partial y}(t,y) = \frac{y}{\sqrt{y^2 + \dfrac{1}{1+t^2}}}..$$

Senza entrare in troppi calcoli, notiamo solo che la derivata parziale è ovviamente positiva, e che inoltre essa è limitata, nel senso che esiste $k > 0$ tale che

$$0 < \frac{\partial f}{\partial y}(t,y) < k \quad \forall t, y \in \mathbb{R}.$$

Ciò accade, informalmente parlando, perché la funzione (di una variabile) $g(t) = 1/(1+t^2)$ è limitata e perché, fissato t, le parti principali di numeratore e denominatore (come funzioni della variabile y) sono dello stesso grado. Quindi la soluzione al problema di Cauchy è unica. Siccome poi

$$0 < f(t,y) \leq \sqrt{y^2+1} \leq |y|+1 \quad \forall t, y \in \mathbb{R}$$

possiamo applicare anche il Teorema 8.14 per concludere che la soluzione è definita per ogni valore di t.

Non abbiamo naturalmente alcuna idea di quale la soluzione sia esplicitamente, né vi è alcun modo per saperlo: tuttavia il metodo delle approssimazioni successive ci fornisce un algoritmo approssimato di calcolo, e vedremo più avanti che il metodo dello *studio qualitativo* ci può dare informazioni ulteriori.

8.6 Alcuni tipi particolari di equazioni differenziali del primo ordine

Affronteremo in questa sezione l'analisi di alcuni tipi di equazioni differenziali del primo ordine per i quali è noto un procedimento generale di soluzione. Molti altri sono i casi in cui sarebbe possibile individuare un procedimento generale di soluzione, ma ci limiteremo qui ad alcuni dei più importanti. Il primo dei casi che affronteremo è già stato accennato in precedenza.

8.6.1 Equazioni differenziali a variabili separabili

Consideriamo un'equazione differenziale della forma

$$y' = f(y)g(t) \tag{8.11}$$

dove f e g sono funzioni continue della loro variabile. Facciamo questa ipotesi perché, in sua assenza, non abbiamo a disposizione alcun risultato di esistenza per le soluzioni di un problema di Cauchy assegnato, e sappiamo in effetti che l'esistenza può anche venir meno.

Vi potrebbe essere una prima classe di soluzioni particolari alla (8.11). Se infatti esiste un valore $\overline{y}$ tale che $f(\overline{y}) = 0$, è chiaro che la funzione *costante*

$$y(t) = \overline{y} \quad \forall t$$

è una soluzione dell'equazione differenziale data, perchè con tale scelta il membro di sinistra dell'equazione data vale zero così come, per costruzione, il membro di destra. Le eventuali soluzioni così costruite si chiamano a volte *integrali singolari* dell'equazione considerata. Ad esempio l'equazione differenziale

$$y' = y^{1/3}$$

vista in precedenza ammette l'integrale singolare $y(t) = 0$.

Per vedere come trovare altre soluzioni, procediamo dapprima formalmente; in effetti gli integrali singolari corrispondono al caso in cui $f(y) = 0$, e se così non è possiamo dividere entrambi i membri per $f(y)$, ottenendo

$$\frac{y'}{f(y)} = g(t).$$

A sinistra ora abbiamo un oggetto dipendente solo dalla funzione y, e non esplicitamente dalla variabile t. A destra abbiamo una funzione solo di t: ecco perché queste equazioni si chiamano *a variabili separabili.* Procedendo per il momento senza troppa pretesa di rigore e usando la notazione di Leibniz

$$y' = \frac{\mathrm{d}y}{\mathrm{d}t}$$

si arriva a scrivere che

$$\frac{\mathrm{d}y}{f(y)} = g(t)\mathrm{d}t :$$

questa scrittura di per sé non ha senso, ma è solo un modo *formale* di riscrivere la formula ad essa precedente. Integrando, possiamo quindi pensare che valga la relazione

$$\int \frac{\mathrm{d}y}{f(y)} = \int g(t)\mathrm{d}t. \tag{8.12}$$

Malgrado il modo un po' approssimativo in cui siamo arrivati alla (8.12), la conclusione è corretta. In effetti si ha il seguente risultato.

Teorema 8.15 *Se f e g sono funzioni continue laddove definite, e se inoltre F è una primitiva di $1/f$ e G è una primitiva di g, cioè*

$$F'(y) = \frac{1}{f(y)}$$

laddove f è definita e diversa da zero, e

$$G'(t) = g(t)$$

laddove g è definita, allora tutte le funzioni $y(t)$ di classe C^1 tali che

$$F(y(t)) = G(t) + c \tag{8.13}$$

con $c \in \mathbb{R}$ costante, sono soluzioni dell'equazione differenziale (8.11).

Dimostrazione è sufficiente derivare rispetto a t entrambi i membri della 8.13 per ottenere, usando il Teorema di derivazione della funzione composta, che

$$F'(y(t))y'(t) = G'(t).$$

Per come sono costruite F e G si ha allora

$$\frac{y'(t)}{f(y(t))} = g(t),$$

che è l'equazione differenziale di partenza. ■

Non dovete pensare che, in generale, le soluzioni appena determinate siano *tutte* quelle possibili. In effetti l'equazione differenziale studiata nell'esempio 131 è a variabili separabili (ora potete capire come sono state ricavate le soluzioni di tale equazioni) e ammette delle soluzioni che sono definite "a pezzi, cioè congiungendo tratti di più di una soluzione (quando questa operazione dà luogo a una funzione derivabile).

Potrete forse pensare che per equazioni del genere sia tutto molto semplice. Purtroppo così non è, perchè dobbiamo pur sempre calcolare delle primitive, e questo può non essere semplice. Ma non solo: anche una volta calcolate le primitive richieste potrebbe non essere possibile esplicitare le soluzioni, cioè scrivere esplicitamente l'espressione di y, come mostra il seguente esempio

Esempio 134 Si consideri l'equazione differenziale a variabili separabili

$$y' = \frac{2t}{2 + \cos y}.$$

Non vi sono integrali singolari. Riscrivendo l'equazione come

$$(2 + \cos y)y' = 2t$$

la si può integrare ottenendo

$$2y + \sin y = t^2 + c \tag{8.14}$$

con $c \in \mathbb{R}$ costante arbitraria, ovvero

$$h(y) = t^2 + c$$

con

$$h(y) := 2y + \sin y.$$

La funzione h è certamente invertibile come funzione da $\mathbb{R}$ in $\mathbb{R}$: infatti la sua derivata $h'(y) = 2 + \cos y$ è sempre strettamente positiva, quindi h è monotona strettamente crescente, quindi è iniettiva ed è inoltre suriettiva se vista come funzione da $\mathbb{R}$ alla propria immagine. Tuttavia tale immagine coincide con $\mathbb{R}$ stesso, poichè h tende a $\pm\infty$ se $y \to + \pm \infty$ (rispettivamente) e h è continua. Quindi potremmo scrivere

$$y(t) = h^{-1}(t^2 + c),$$

dove abbiamo denotato come di consueto con h^{-1} la funzione inversa di h. Tuttavia non è possibile scrivere un'espressione analitica per h^{-1}. Diremo che y è *definita implicitamente* dall'equazione (8.14). Sebbene esistano metodi molto generali per poter studiare funzioni definite implicitamente, non se potrà parlare in questo corso, e quindi ci dovremo accontentare in casi come questo dell'espressione implicita.

Bisogna prestare molta attenzione al *dominio di definizione* della soluzione a un'equazione differenziale a variabili separabili, come mostra il seguente semplice esempio.

Esempio 135 Consideriamo il problema di Cauchy seguente:

$$\begin{cases} y' = -\dfrac{t}{y} \\ y(1) = -2. \end{cases}$$

Per prima cosa risolviamo l'equazione differenziale per separazione di variabili. Non vi sono integrali singolari. Inoltre possiamo scrivere

$$yy' = -t$$

e quindi

$$y^2(t) = c - t^2$$

con $c \in \mathbb{R}$ costante arbitraria. Ne segue anche che deve valere

$$y(t) = \sqrt{c - t^2}$$

oppure

$$y(t) = -\sqrt{c - t^2}$$

ogni qualvolta la radice abbia senso, cioè *quando c è strettamente positiva* (altrimenti il radicando è sempre negativo, se $c < 0$, ed è uguale a zero solo per $t = 0$ quando $t = 0$) e quando, in tale ipotesi, $|t| < \sqrt{c}$. Si noti il segno di minore stretto: in effetti si farebbe un errore definendo la soluzione anche per $|t| = \sqrt{c}$ (punti nei quali $y(t)$ è ben definita e vale zero), perché in tali punti essa non risulterebbe derivabile (da destra o da sinistra). Vediamo ora il problema di Cauchy. Il fatto che il dato iniziale abbia un ordinata *negativa* obbliga a scegliere la soluzione con il segno meno. Posto $t = 1$ si ottiene

$$-2 = -\sqrt{c - 1}$$

e quindi $c = 5$. Quindi la soluzione è

$$y(t) = -\sqrt{5 - t^2} \quad \forall t \in]-\sqrt{5}, \sqrt{5}[.$$

Esempio 136 Modifichiamo un poco l'esempio precedente considerando l'equazione differenziale

$$y' = -\frac{t^2}{y^2}$$

che, con metodi del tutto analoghi a quelli dell'esempio precedente, si mostra ammettere *formalmente* le soluzioni seguenti:

$$y(t) = \sqrt[3]{c - t^3}$$

con $c \in \mathbb{R}$ costante arbitraria. Apparentemente non ci sono problemi di dominio di definizione, perché la radice cubica è sempre ben definita quale che ne sia l'argomento. Tuttavia il punto $t = \sqrt[3]{c}$ deve essere escluso, perché la funzione non è ivi derivabile. Questa non è una sottigliezza, come si vede considerando ad esempio il problema di Cauchy relativo al dato iniziale $y(-1) = 2$, che corrisponde alla scelta $c = 7$. La soluzione è

$$y(t) = \sqrt[3]{7 - t^3},$$

ma soltanto per $t < \sqrt[3]{7}$. Che per $t = \sqrt[3]{7}$ ci siano dei problemi ce lo dice peraltro l'equazione differenziale di partenza stessa: per tale valore di t la y si annulla e quindi il secondo membro dell'equazione differenziale non ha senso.

Esercizio 8.1 Trovare, eventualmente in forma implicita, le soluzioni delle seguenti equazioni differenziali:

$$t(1+y^2)y' = 3;$$

$$y' = \sqrt[3]{y} + y;$$

$$(1+e^{2t})y' = ye^{2t};$$

$$y' = \frac{t^2 y}{t^2-1}.$$

Non è richiesta la discussione del dominio della soluzione se non è possibile porre quest'ultima in forma esplicita.
R. (si indica con c, se non diversamente specificato, una costante reale arbitraria).
$y(t) + \frac{y^3(t)}{3} = 3\log|t| + c$;
$\frac{1}{2}\ln(1+y^2(t)) + \ln(1+y^{2/3}(t)) - \frac{1}{2}\ln(y^{4/3}(t) - y^{2/3}(t) + 1) = t + c$ (vi è inoltre la soluzione particolare $y(t) = 0 \ \forall t$);
$y(t) = c\sqrt{1+e^{2t}}$ $(t \in \mathbb{R})$;
$y(t) = ce^t\sqrt{\frac{t-1}{t+1}}$ separatamente su ognuno degli intervalli $]-\infty,-1[$, $]-1,1[$, $]1,+\infty[$.

Esercizio 8.2 Trovare le soluzioni dei seguenti problemi di Cauchy, precisandone l'intervallo di definizione:

$$\begin{cases} y' = y^{2/3}\sin t \\ y\left(\frac{\pi}{2}\right) = -8; \end{cases}$$

$$\begin{cases} y' = -\dfrac{\ln t}{t\cos y} \\ y(2) = \pi; \end{cases}$$

$$\begin{cases} e^{t^2+y}y' = -2t \\ y(0) = -1. \end{cases}$$

R. $y(t) = -\left(\frac{6+\cos t}{3}\right)^3$ $(t \in \mathbb{R})$;
$y(t) = \arcsin\left[\ln\left(\frac{\ln 2}{\ln t}\right)\right]$ $(t \in]2^{1/e}, 2^e[)$;
$y(t) = \ln\left(e^{-t^2} + \frac{1}{e} - 1\right)$ $(|t| < \sqrt{\ln\frac{e}{e-1}})$.

8.6.2 Equazioni differenziali omogenee

Il prossimo tipo di equazioni differenziali è costituito da equazioni differenziali che, tramite un'opportuna *sostituzione* si possono ricondurre a un'equazione differenziale a variabili separabili. Si tratta di equazioni differenziali della forma

$$y' = f\left(\frac{y}{t}\right)$$

dove f è una funzione continua sul proprio dominio. L'espressione stessa dall'equazione ci dice che, in linea di principio, le soluzioni non possono essere definite in $t = 0$ perché in tale punto il secondo membro perde senso. Vedremo con un esempio cosa può accadere per $t \to 0$. Notiamo subito che possono esistere integrali singolari: se infatti $c \in \mathbb{R}$ è un numero tale che $c = f(c)$ si ha subito che la funzione

$$y(t) = ct$$

è una soluzione dell'equazione differenziale data. Naturalmente non è detto che esistano valori di c siffatti. Notiamo anche che ogni soluzione dell'equazione differenziale data *non può essere definita per* $t = 0$, ma gli eventuali integrali singolari possono essere senz'altro prolungati in tale punto ottenendo una funzione regolare (ciò non è detto per le altre soluzioni).

Per trovare altre soluzioni dell'equazione differenziale data, occorre come detto fare un *cambiamento di variabili.* Va infatti posto

$$z = \frac{y}{t}.$$

la funzione z dovrà allora soddisfare una nuova equazione differenziale: vediamo quale. Si ha

$$z' = \frac{y'}{t} - \frac{y}{t^2}$$

e quindi, ricordando la definizione di z e l'equazione differenziale soddisfatta da y, deve valere

$$z' = \frac{f(z) - z}{t}.$$

Questa è un'equazione differenziale a variabili separabili. In linea di principio la si può risolvere, e una volta fatto questo si potrà agevolmente ricavare la y scrivendo $y(t) = tz(t)$. Vediamo un esempio.

Esempio 137 Risolviamo l'equazione differenziale

$$y' = \frac{t^2 + y^2}{ty}.$$

L'equazione non ha senso se $t = 0$ e se $y = 0$. Ricordiamoci di questo fatto per discuterlo al termine dello studio. L'equazione non sembra proprio omogenea, ma basta dividere numeratore e denominatore del secondo membro per t^2 (lo possiamo fare, visto che chiediamo $t \neq 0$), per ottenere

$$y' = \frac{1 + \frac{y^2}{t^2}}{\frac{y}{t}}.$$

Quindi $f(x) = (1 + x^2)/x$. Non vi sono integrali singolari perché l'equazione

$$c = \frac{1+c}{c}$$

non ammette soluzioni reali.

Sappiamo poi che, posto $z = y/t$, l'equazione differenziale diventa

$$z' = \frac{\frac{1+z^2}{z} - z}{t}$$

cioè

$$z' = \frac{1}{zt}.$$

Questa equazione differenziale a variabili separabili si integra facilmente. Si ottiene infatti

$$z^2 = c + 2 \ln |t|$$

dove c è una costante reale. Fate attenzione però: il membro di destra non è definito in $t = 0$ e quindi l'ultima l'equazione scritta va considerata *separatamente* per $t > 0$ e $t < 0$, e si potrebbero anche scegliere costanti *diverse* nei due casi. A questo punto si ottiene, ad esempio per $t > 0$,

$$z(t) = \sqrt{c + 2 \ln t}$$

oppure

$$z(t) = -\sqrt{c + 2 \ln t}$$

ma solo per i tempi t per i quali il secondo membro è definito cioè, dato c, per $t > e^{-c/2}$ (il valore $t = e^{-c/2}$ va escluso per le solite considerazioni sulla non derivabilità della funzione in tale punto che rifaremo tra poco). Tornando alla variabile y si ottiene

$$y(t) = t\sqrt{c + 2 \ln t}$$

oppure

$$y(t) = -t\sqrt{c + 2\ln t},$$

sempre per tali tempi. La figura (8.3) mostra la forma di queste soluzioni. Si noti che $y(t) \to 0$ se $t \to (e^{-c/2})^+$, e che l'equazione differenziale perdeva senso quando $y = 0$. Ci chiediamo quindi se sia possibile o meno prolungare la soluzione oltre il valore limite: ma ciò non si può fare, perché tutte le soluzioni scritte, che di per sé sarebbero definite per $t \geq e^{-c/2}$, non sono derivabili in $t = e^{-c^2}$. A maggior ragione non si può prolungare alcuna soluzione per tempi *più piccoli* di tale valore. Si può procedere analogamente per tempi $t < 0$.

Esercizio 8.3 Risolvere l'equazione differenziale

$$y' = \frac{t^3 + y^3}{ty^2}$$

precisando il dominio delle soluzioni.
R. $y(t) = t\sqrt[3]{c + 3\ln|t|}$, separatamente per $t > 0$ e per $t < 0$.

Le equazioni differenziali ora discusse hanno una loro importanza, ma forse ancora più importante è che si sia capito il procedimento di cambiamento di variabili in un'equazione differenziale, che spesso può semplificare di molto il problema affrontato. Commentiamo solo per chiudere che il cambiamento di variabili effettuato ha coinvolto la variabile *dipendente* y, mentre è possibile effettuare cambiamenti di variabili anche sulla variabile *indipendente* t. Senza entrare in dettagli, questo vuol dire porre

$$s = a(t)$$

con a di classe C^1 e invertibile, e notare che, con la notazione di Leibniz

$$\frac{\mathrm{d}y}{\mathrm{d}s} = \frac{\mathrm{d}y}{\mathrm{d}t}\frac{\mathrm{d}t}{\mathrm{d}s}.$$

A questo punto si può usare, per esplicitare il secondo membro, l'equazione differenziale soddisfatta da y (vista come funzione di t) oltre all'espressione esplicita di t come funzione di s. Come nel calcolo integrale procedimenti di questo tipo possono semplificare molto la situazione, ma le sostituzioni vanno scelte con cura. Non ci occuperemo qui di trattare la questione in dettaglio.

8.6.3 Equazioni differenziali lineari del primo ordine

Studieremo qui come risolvere esplicitamente, a meno del calcolo delle primitive di opportune funzioni, una classe particolarmente semplice di equazioni differenziali: le equazioni differenziali lineari del primo ordine. Ricordiamo (si veda la Definizione 8.2) che si tratta di equazioni della forma

$$y' = p(t)y + q(t), \tag{8.15}$$

dove assumeremo senz'altro che le funzioni p e q siano continue su un intervallo $[a,b]$.

Osserviamo dapprima che, assegnato un dato iniziale $y(t_0) = y_0$ con $t_0 \in]a,b[$, $y_0 \in \mathbb{R}$, il corrispondente problema di Cauchy ammette esattamente una soluzione e che tale soluzione è definita su tutto l'intervallo $[a,b]$.

Teorema 8.16 *Siano $p, q : [a,b] \to \mathbb{R}$ funzioni continue, e si consideri il problema di Cauchy*

$$\begin{cases} y' = p(t)y + q(t) \\ y(t_0) = y_0 \end{cases} \tag{8.16}$$

dove $t_0 \in]a,b[$ e $y_0 \in \mathbb{R}$. Allora tale problema di Cauchy ammette esattamente una soluzione $y(t)$ e tale soluzione è definita sull'intero intervallo $[a,b]$.

Dimostrazione Valgono le ipotesi del Teorema di Cauchy in quanto la funzione

$$f(t,y) = p(t)y + q(t)$$

è continua ovunque definita per le ipotesi fatte sulle funzioni p e q, ed essa è inoltre Lipschitziana in y uniformemente rispetto alla t, perché

$$|f(t,y_1) - f(t,y_2)| = |p(t)|\,|y_1 - y_2| \leq \left[\max_{t\in[a,b]} |p(t)|\right] |y_1 - y_2|$$

per ogni $t \in [a,b]$, $y_1, y_2 \in \mathbb{R}$, in quanto per il Teorema di Weierstrass il massimo di $|p(t)|$ su $[a,b]$ esiste finito essendo p, e quindi $|p|$, una funzione continua su $[a,b]$. Dunque la soluzione al problema assegnato esiste localmente ed è unica. Vale però anche la condizione (8.10), in quanto la funzione f in questione ha crescita esattamente lineare come funzione di y e i coefficienti p e q sono funzioni continue, quindi limitate su $[a,b]$, di nuovo per il Teorema di Weierstrass. Valgono quindi le ipotesi del Teorema di esistenza in grande, e la soluzione è quindi definita su $[a,b]$. ■

Corollario 8.17 *Se, in aggiunta alle condizioni del Teorema precedente, si suppone che p e q siano definite e continue su tutto $\mathbb{R}$, allora la soluzione al problema di Cauchy considerato esiste, è unica ed è definita su tutto $\mathbb{R}$.*

Dimostrazione Basta osservare che si può applicare il Teorema precedente su *ogni* intervallo $[a, b]$ con $a, b \in \mathbb{R}$. ■

Vediamo ora come costruire esplicitamente le soluzioni di (8.15) (a meno del calcolo di opportune primitive). Valendo il Teorema precedente per la soluzione del problema di Cauchy ci aspettiamo che esse dipendano da una sola costante arbitraria (ad esempio il valore del dato di Cauchy y_0 al tempo t_0).

Per prima cosa cominciamo con il considerare un'equazione ancora più semplice, *l'equazione omogenea associata* alla (8.15):

$$y' = p(t)y. \tag{8.17}$$

Si tratta di un'equazione a variabili separabili, che può essere risolta immediatamente (a meno del calcolo di una primitiva) come sappiamo. Risulta infatti che le soluzioni dell'equazione differenziale data sono le seguenti:

$$y(t) = Ce^{\int p(t)\mathrm{d}t}$$

dove come di consueto $\int p(t)\mathrm{d}t$ indica una primitiva di $p(t)$. Indipendentemente dall'espressione esplicita della soluzione, si ha il seguente Teorema

Teorema 8.18 *Siano $y_1, y_2 : [a, b] \to \mathbb{R}$ due qualunque soluzioni dell'equazione (8.15). Allora la loro differenza*

$$\overline{y}(t) = y_1(t) - y_2(t), \quad t \in [a, b]$$

soddisfa l'equazione (8.17). In particolare la soluzione generale $y(t)$ dell'equazione (8.15) si ottiene sommando a una particolare soluzione $\widetilde{y}$ dell'equazione (8.15) la soluzione generale y_0 dell'omogenea associata:

$$\underbrace{y(t)}_{\substack{\text{Soluzione generale}\\ \text{equazione completa}}} = \underbrace{\widetilde{y}(t)}_{\substack{\text{Soluzione particolare}\\ \text{equazione completa}}} + \underbrace{y_0(t)}_{\substack{\text{Soluzione generale}\\ \text{equazione omogenea}}}.$$

Dimostrazione Basta calcolare

$$\begin{aligned}\overline{y}'(t) &= p(t)y_1(t) + q(t) - (p(t)y_2(t) + q(t))\\ &= p(t)(y_1(t) - y_2(t))\\ &= p(t)\overline{y}(t).\end{aligned}$$

Per la seconda parte basta quindi osservare che, come appena mostrato, la differenza tra $y(t)$ e $\widetilde{y}(t)$ è una soluzione dell'equazione differenziale omogenea. ■

Ci rimane quindi soltanto da costruire esplicitamente almeno una soluzione all'equazione data. Questo si fa con il cosiddetto *metodo di variazione delle costanti arbitrarie*. Questo metodo prende il nome dal fatto che si considera la soluzione dell'omogenea

$$y(t) = Ce^{\int p(t)\mathrm{d}t}$$

e si cerca una soluzione *dell'equazione completa* del tipo

$$y(t) = C(t)e^{\int p(t)\mathrm{d}t},$$

cioè facendo "variare la costante arbitraria, per quanto la terminologia sia in effetti un po' contraddittoria.

Teorema 8.19 *La funzione*

$$\widetilde{y}(t) = \int q(t)e^{-\int p(t)\mathrm{d}t}\mathrm{d}t$$

è una soluzione dell'equazione (8.15). Quindi la soluzione generale dell'equazione stessa è la seguente:

$$y(t) = e^{\int p(t)\mathrm{d}t}\left[C + \int q(t)e^{-\int p(t)\mathrm{d}t}\right]$$

con $C \in \mathbb{R}$ costante arbitraria.

Dimostrazione Come detto procediamo cercando una soluzione del tipo

$$\widetilde{y}(t) = C(t)e^{\int p(t)\mathrm{d}t}$$

con $C(t)$ funzione di classe C^1. La derivata di $\widetilde{y}$ vale, per definizione di primitiva,

$$\begin{aligned}\widetilde{y}'(t) \quad &= C'(t)e^{\int p(t)\mathrm{d}t} + C(t)p(t)e^{\int p(t)\mathrm{d}t} \\ &= C'(t)e^{\int p(t)\mathrm{d}t} + p(t)\widetilde{y}(t).\end{aligned}$$

Se allora chiediamo che valga l'identità

$$C'(t)e^{\int p(t)\mathrm{d}t} = q(t)$$

l'equazione differenziale (8.15) sarà soddisfatta. Quindi dobbiamo avere

$$C'(t) = e^{-\int p(t)\mathrm{d}t}q(t)$$

cioè

$$C(t) = \int e^{-\int p(t)\mathrm{d}t} q(t)\mathrm{d}t.$$

■

L'aver determinato la soluzione generale dell'equazione differenziale studiata ci permette immediatamente di trovare la soluzione del problema di cauchy associato.

Corollario 8.20 *Siano $p, q : [a, b] \to \mathbb{R}$ funzioni continue: La soluzione del problema di Cauchy*

$$\begin{cases} y' = p(t)y + q(t) \\ y(t_0) = y_0 \end{cases}$$

dove $t_0 \in]a, b[$ e $y_0 \in \mathbb{R}$ è data da:

$$y(t) = e^{\int_{t_0}^{t} p(s)\mathrm{d}s} \left[y_0 + \int_{t_0}^{t} q(s) e^{-\int_{t_0}^{s} p(u)\mathrm{d}u} \mathrm{d}s \right].$$

Dimostrazione La funzione scritta è, per il Teorema precedente, una soluzione dell'equazione differenziale studiata. Che $y(t_0)$ valga y_0 è certamente vero, e quindi l'asserto segue dall'unicità della soluzione al problema di Cauchy. ■

L'avere una formula esplicita per la soluzione è estremamente utile, ma non esime dalla necessità di calcolare primitive.

Esempio 138 Risolviamo il problema di Cauchy

$$\begin{cases} y' = y \tan t + \sin t \\ y(0) = 1. \end{cases}$$

Notiamo che la tangente non è definita nei punti $k\pi + (\pi/2)$ con $k \in \mathbb{Z}$, quindi l'intervallo di definizione della soluzione sarà $I :=] - \pi/2, \pi/2[$, in quanto il punto $t = 0$ cade appunto in I. Si avrà

$$\int \tan t \mathrm{d}t = -\ln \cos t,$$

per $t \in] - \pi/2, \pi/2[$, in quanto $\cos x$ è positivo in tale intervallo (altrimenti come sappiamo sarebbe stato necessario prendere la funzione coseno in valore assoluto). Si noti l'assenza della costante arbitraria perche' ci basta, per determinare una soluzione, trovare *una* primitiva di $\tan t$. Inoltre

$$\begin{aligned} \int \sin t e^{-\int \tan t \mathrm{d}t} \mathrm{d}t &= \int \sin t e^{\ln \cos t} \mathrm{d}t \\ &= \int \sin t \cos t \mathrm{d}t \\ &= \frac{\sin^2 t}{2}. \end{aligned}$$

Quindi la soluzione generale *nell'intervallo* $]-\pi/2, \pi/2[$ è:

$$y(t) = \frac{1}{\cos t}\left(C + \frac{\sin^2 t}{2}\right).$$

Per $t = 0$ si ottiene $y(0) = C$ e quindi il dato di Cauchy impone la condizione $C = 1$. In conclusione il problema di Cauchy ammette la soluzione (unica)

$$y(t) = \frac{1}{\cos t}\left(1 + \frac{\sin^2 t}{2}\right)$$

per ogni $t \in]-\pi/2, \pi/2[$.

Esercizio 8.4 Risolvere le seguenti equazioni differenziali lineari del primo ordine, precisando il dominio di definizione delle soluzioni:

$$y' = \frac{1}{t}y + \frac{1}{t};$$

$$y' = ty - t;$$

$$y' = -\frac{t}{1+t^2}y + \frac{1}{t(1+t^2)}.$$

R. $y(t) = ct + 1$ (prolungabile anche a $t = 0$ malgrado l'equazione differenziale sia singolare in tal punto);
$y(t) = 1 + ce^{t^2/2}$ (su $\mathbb{R}$);
$y(t) = \frac{c}{\sqrt{1+t^2}} + 1 - \frac{1}{2}\ln\left(\frac{\sqrt{1+t^2}+1}{\sqrt{1+t^2}-1}\right)$ (su $\mathbb{R}$).

8.6.4 Equazioni differenziali di Bernoulli

Accenniamo brevemente a uno dei molti tipi di equazioni differenziali che si possono ricondurre con un cambiamento di variabili a un'equazione lineare. Si tratta delle cosiddette equazioni di Bernoulli, cioè di equazioni differenziali del primo ordine della forma

$$y' = p(t)y + q(t)y^\alpha \tag{8.18}$$

con $\alpha \in \mathbb{R}$ e p, q funzioni continue su un intervallo $[a, b]$. Si noti che, se $\alpha > 0$, vi è l'integrale singolare $y(t) = 0$ $\forall t \in [a, b]$. Se $\alpha = 0$ o $\alpha = 1$ abbiamo a che fare con un'equazione lineare (rispettivamente omogenea e non omogenea) e dunque non vi è nulla da dire in tali casi.

Teorema 8.21 *Se $\alpha \neq 1$ la soluzione generale all'equazione (8.18) è la seguente:*

$$y(t) = e^{\int p(t)\mathrm{d}t}\left(c + (1-\alpha)\int q(t)e^{-(1-\alpha)\int p(t)\mathrm{d}t}\right)^{1/(1-\alpha)}.$$

per ogni $t \in [a,b]$, con $C \in \mathbb{R}$ costante arbitraria. Se $\alpha > 0$ vi è anche l'integrale singolare $y(t) = 0$ per ogni t.

Dimostrazione Si operi il cambiamento di variabili

$$z(t) = y(t)^{1-\alpha}.$$

Se si deriva si ottiene

$$\begin{aligned} z'(t) &= (1-\alpha)y(t)^{-\alpha}y'(t) \\ &= (1-\alpha)p(t)y(t)^{1-\alpha} + (1-\alpha)q(t) \\ &= (1-\alpha)p(t)z(t) + (1-\alpha)q(t). \end{aligned}$$

Quindi z è soluzione dell'equazione lineare

$$z' = (1-\alpha)p(t)z + (1-\alpha)q(t).$$

Quindi

$$z(t) = e^{(1-\alpha)\int p(t)\mathrm{d}t}\left(c + (1-\alpha)\int q(t)e^{-(1-\alpha)\int p(t)\mathrm{d}t}\right).$$

Ritornando alla variabile y si ha l'asserto. ■

Si noti che se $1/(1-\alpha)$ è un razionale con denominatore pari (ad esempio se $\alpha = -1$, cosicché $1/(1-\alpha) = 1/2$) si deve intendere che vi è anche un'analoga famiglia di soluzioni che differiscono da quelle scritte per un segno meno. Va anche notato che la soluzione scritta va intesa *laddove essa è ben definita*, poiché una potenza a esponente reale è in generale ben definita solo se la base è positiva.

Esercizio 8.5 * Trovare le soluzioni delle equazioni differenziali (non è richiesto di indicare esplicitamente i domini di definizione):

$$y' = 2ty + t^3y^3;$$

$$y' = \frac{t + ty^3}{y^2};$$

$$(1-t^2)y' - 3ty - \frac{t}{y^2} = 0;$$

$$ty' + y(1 - ty^n) = 0 \quad n \in \mathbb{N},\ n \geq 2.$$

R. $y(t) = \pm\left(ce^{-2t^2} - \frac{t^2}{2} - \frac{1}{4}\right)^{-1/2}$;
$y(t) = \left(ce^{3t^2/2} - 1\right)^{1/3}$;
$y(t) = \frac{1}{|1-t^2|^{3/2}}\left(c - \frac{1}{3}|1-t^2|^{9/2}\right)^{1/3}$ (separatamente su $|t| > 1$ e $|t| < 1$);
$y(t) = \pm\frac{1}{|t|}\left(c - \frac{n}{n+1}t^{n+1}\right)^{-1/n}$ se n è pari, $y(t) = \frac{1}{|t|}\left(c - \frac{n}{n+1}t|t|^n\right)^{-1/n}$ se n è dispari.

8.7 Equazioni lineari

In questo capitolo affronteremo lo studio delle cosiddette *equazioni differenziali lineari*, su cui è particolarmente semplice dare risultati generali anche quando l'ordine dell'equazione è maggiore o uguale a due. Ricordiamo (si veda la terza parte della definizione 8.2) che un'equazione differenziale si dice lineare se è della forma

$$a_0(t)y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y = f(t).$$

Tale equazione si dice *non omogenea* per la presenza del termine $f(t)$; se $f(t)$ fosse uguale a zero parleremmo invece di equazione *omogenea*. In realtà studieremo d'ora in poi il caso in cui l'equazione può essere posta in forma normale:

$$y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y = f(t). \tag{8.19}$$

Il primo punto è capire se l'equazione data ammette o meno soluzioni. Il fatto che sia lineare dovrebbe farci sospettare, per analogia con le equazioni del primo ordine, che sia effettivamente così. Non potremo dimostrarlo, ma notiamo solo che l'equazione (8.19) può essere riscritta come *sistema di equazioni differenziali del primo ordine*. In effetti, posto

$$y_1 = y,\ y_2 = y', \ldots, y_n = y^{(n-1)}$$

si ha che

$$\begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ \vdots \\ y_{n-1}' = y_n \\ y_n' = -a_1(t)y_n - a_2(t)y_{n-1} - \ldots - a_{n-1}(t)y_2 - a_n(t)y_1 + f(t). \end{cases} \tag{8.20}$$

Un eventuale insieme di dati di Cauchy per l'equazione (8.19) del tipo

$$y(t_0) = \tilde{y}_0,\ y'(t_0) = \tilde{y}_1, \ldots, y^{(n-1)}(t_0) = \tilde{y}_{n-1} \tag{8.21}$$

diventa, con le nuove variabili:

$$y_1(t) = \tilde{y}_0, y_2(t_0) = \tilde{y}_2, \ldots, y_n(t_0) = \tilde{y}_{n-1}. \tag{8.22}$$

Vi sarete naturalmente chiesti cosa abbiamo guadagnato in questa operazione. In effetti si vede immediatamente che le due formulazioni sono tra loro equivalenti, nel senso che y è una soluzione dell'equazione (8.19) con i dati di Cauchy (8.21) se e soltanto se il vettore di funzioni $(y_1, \ldots, y_n)$ costruito come sopra soddisfa le equazioni (8.20) con i dati (8.22).

Il punto è che è molto più facile dimostrare un risultato di esistenza per il sistema del primo ordine (8.20) che per l'equazione differenziale di partenza. Si può in effetti notare che la dipendenza dalle incognite $y_1, \ldots, y_n$ è *lineare* in ognuna delle equazioni che compaiono nel sistema (8.20). Già sappiamo, per il problema di Cauchy relativo a un'equazione lineare del primo ordine con coefficienti continui, che esiste una e una sola soluzione al problema di Cauchy studiato e che la soluzione è definita per tutti i t nell'intervallo di continuità dei coefficienti: lo stesso si dimostra, con metodi del tutto analoghi, anche per *sistemi* di equazioni differenziali lineari del primo ordine.

Teorema 8.22 *Se le funzioni $a_1, \ldots, a_n, f :]a, b[\to \mathbb{R}$ sono continue su $]a, b[$ e $t_0 \in]a, b[$, allora il problema di Cauchy (8.20), (8.22) ammette una e una sola soluzione su tutto $]a, b[$. Quindi anche il problema di Cauchy (8.19), (8.21) ammette una sola soluzione.*

Corollario 8.23 *Se le funzioni $a_1(t), \ldots, a_n(t), f(t)$ sono definite e continue su tutto $\mathbb{R}$, allora le soluzioni determinate nel Teorema 8.22 sono definite su tutto $\mathbb{R}$.*

Noi studieremo in dettaglio solo il caso di equazioni differenziali lineari del secondo ordine ($n = 2$ nell'equazione appena scritta), particolarmente significativo nelle applicazioni, ma notiamo però che vale, quale che sia l'ordine n dell'equazione, un'analogo del principio di sovrapposizione già visto per equazioni differenziali lineari del primo ordine.

Teorema 8.24 (Principio di sovrapposizione) *Si supponga che una funzione $y_1(t)$ sia soluzione, in un intervallo $]a, b[$, dell'equazione differenziale*

$$y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y = f_1(t)$$

e che una funzione $y_2(t)$ sia soluzione, in $]a, b[$ dell'equazione differenziale

$$y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y = f_2(t)$$

con $a_0, \ldots, a_n, f_1, f_2 :]a, b[\to \mathbb{R}$ funzioni assegnate. Allora la funzione

$$y(t) = c_1 y_1(t) + c_2 y_2(t)$$

con $c_1, c_2 \in \mathbb{R}$ è soluzione in $]a, b[$ dell'equazione differenziale

$$y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y = c_1 f_1(t) + c_2 f_2(t).$$

In particolare se $f_1(t) = f_2(t)$ per ogni t, cioè se y_1 e y_2 sono soluzioni della medesima equazione differenziale lineare non omogenea, allora la funzione

$$y(t) = y_1(t) - y_2(t)$$

è soluzione dell'equazione omogenea

$$y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y = 0.$$

Quindi la soluzione generale dell'equazione non omogenea si ottiene sommando ad una soluzione particolare di tale equazione la soluzione generale dell'equazione omogenea.

Dimostrazione Basta calcolare

$$\begin{aligned}
&y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y \\
&= [c_1 y_1^{(n)} + c_2 y_2^{(n)}] + a_1(t)[c_1 y_1^{(n-1)} + c_2 y_2^{(n-1)}] + \ldots \\
&+ \ldots + a_{n-1}(t)[c_1 y_1' + c_2 y_2'] + a_n(t)[c_1 y_1(t) + c_2 y_2(t)] \\
&= c_1 \left[y_1^{(n)} + a_1(t)y_1^{(n-1)} + \ldots + a_{n-1}(t)y_1' + a_n(t)y_1\right] \\
&+ c_2 \left[y_2^{(n)} + a_1(t)y_2^{(n-1)} + \ldots + a_{n-1}(t)y_2' + a_n(t)y_2\right] \\
&= c_1 f_1(t) + c_2 f_2(t).
\end{aligned}$$

La seconda parte si ottiene notando che, come nel caso di equazioni del primo ordine, la differenza di due soluzioni della medesima equazione non omogenea deve essere, per il punto precedente, soluzione dell'equazione omogenea. ■

Il Teorema precedente ci dà varie informazioni. La prima è il principio di sovrapposizione vero e proprio. Per capire di cosa si tratti, pensate all'equazione differenziale lineare non omogenea che stiamo considerando come all'equazione di evoluzione nel tempo di un sistema soggetto alla forzante esterna $f(t)$. Se si considerano la soluzione y_1 corrispondente alla forzante f_1 e la soluzione y_2 corrispondente alla forzante f_2, il Teorema dice che la soluzione corrispondente alla forzante $f_1 + f_2$ è proprio $y_1 + y_2$, cioè gli effetti delle due forzanti si "sommano in modo da far sì che il sistema soggetto alla forzante somma evolva secondo la somma delle due soluzioni. Questo accade perché il primo membro dell'equazione dipende linearmente da y e dalle sue derivate.

La seconda informazione fondamentale è la seguente

Corollario 8.25 *L'insieme delle soluzioni dell'equazione lineare omogenea*

$$y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y = 0.$$

con $a_1, \ldots, a_n$ funzioni continue su un intervallo $]a, b[$, ha la struttura di spazio vettoriale, cioè se y_1 e y_2 sono soluzioni dell'equazione differenziale data allora anche $c_1y_1 + c_2y_2$ lo è, per ogni scelta delle costanti c_1, c_2.

In effetti l'enunciato varrebbe anche senza chiedere che le funzioni $a_j(t)$ fossero continue, ma il richiederlo serve a garantirsi, come abbiamo visto, che esistano effettivamente soluzioni.

Ci occuperemo d'ora in poi solo di equazioni differenziali lineari *del secondo ordine* a coefficienti continui. Questo solo per risparmiarci alcune complicazioni tecniche e di notazione: non vi sono infatti differenze sostanziali nell'impostazione logica.

Partiremo dall'equazione *omogenea*, e considereremo quindi un'equazione differenziale del tipo

$$y'' + a_1(t)y' + a_2(t)y = 0 \tag{8.23}$$

con a_1, a_2 funzioni continue su un intervallo $]a, b[$. Sappiamo che l'insieme delle soluzioni di questa equazione differenziale è uno spazio vettoriale e ci aspettiamo che la generica soluzione sia caratterizzata da *due* costanti arbitrarie, in quanto l'equazione è del secondo ordine. Se questo fosse vero (e lo è, come vedremo), si direbbe in termini matematici che lo spazio delle soluzioni è uno *spazio vettoriale di dimensione due*.

Diamo per cominciare una definizione

Definizione 8.26 Siano y_1, y_2 soluzioni dell'equazione (8.23) nell'intervallo $]a, b[$. Esse si dicono linearmente indipendenti se le uniche costanti c_1, c_2 per le quali si ha

$$c_1 y_1(t) + c_2 y_2(t) = 0 \quad \forall t \in]a, b[\tag{8.24}$$

sono $c_1 = 0, c_2 = 0$. Si dicono invece linearmente dipendenti se esistono costanti c_1 c_2 non entrambe nulle tali che (8.24) valga.

Due soluzioni sono quindi linearmente *dipendenti* se l'una è un multiplo dell'altra: in realtà quindi quando questo accade si è trovata una famiglia di soluzioni che dipende da *un solo* parametro di soluzioni, e ce ne dovrebbero essere altre.

I prossimi Lemmi ci forniscono una condizione tecnica utilissima per vedere se due soluzioni sono linearmente indipendenti.

Lemma 8.27 *Siano y_1, y_2 soluzioni dell'equazione (8.23) nell'intervallo $]a, b[$. Si definisca il Wronskiano $W(t)$ di y_1 e y_2 come segue:*

$$W(t) = \det \begin{pmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{pmatrix}$$

Allora

$$W(t) = Ce^{-\int a_1(t) \mathrm{d}t}.$$

In particolare il Wronskiano o è uguale a zero su tutto l'intervallo $]a, b[$ o non si annulla mai su tale intervallo.

Dimostrazione Si noti che

$$W(t) = y_1(t)y_2'(t) - y_2(t)y_1'(t)$$

e quindi

$$W'(t) = y_1(t)y_2''(t) - y_2(t)y_1''(t).$$

Si scriva ora l'equazione differenziale soddisfatta da y_1 e la si moltiplichi per y_2, quella soddisfatta da y_2 e la si moltiplichi per y_1. Se ne ottiene:

$$\begin{cases} y_2 y_1'' + a_1(t) y_2 y_1' + a_2(t) y_2 y_1 = 0 \\ y_1 y_2'' + a_1(t) y_1 y_2' + a_2(t) y_2 y_1 = 0. \end{cases}$$

Sottraendo la prima equazione dalla seconda si ottiene

$$y_1 y_2'' - y_2 y_1'' + a_1(t)(y_1 y_2' - y_2 y_1') = 0.$$

Questo implica, per l'espressione di W e W' che

$$W' + a_1(t) W = 0.$$

Integrando questa equazione a variabili separabili si ha la tesi. ■

Lemma 8.28 *Due soluzioni y_1, y_2 dell'equazione (8.23) nell'intervallo $]a,b[$ sono ivi linearmente dipendenti se e solo se il loro Wronskiano è diverso da zero in almeno un punto di $]a,b[$, e quindi in tutto $]a,b[$. Sono ivi linearmente dipendenti se il loro Wronskiano è nullo in almeno un punto di $]a,b[$, e quindi in tutto $]a,b[$.*

Dimostrazione Basta dimostrare che il Wronskiano è nullo se e solo se le soluzioni sono linearmente dipendenti. Supponiamo che $W(t_0) = 0$ con $t_0 \in]a,b[$. Allora il sistema lineare, nelle variabili c_1, c_2,

$$\begin{cases} c_1 y_1(t_0) + c_2 y_2(t_0) = 0 \\ c_1 y_1'(t_0) + c_2 y_2'(t_0) = 0 \end{cases} \tag{8.25}$$

ammette soluzioni $\tilde{c}_1, \tilde{c}_2$ non entrambe nulle, visto che il determinante della matrice dei coefficienti è nullo poiché coincide con il Wronskiano. Quindi la funzione

$$y(t) = \tilde{c}_1 y_1(t) + \tilde{c}_2 y_2(t)$$

è una soluzione dell'equazione differenziale che verifica le condizioni di Cauchy

$$y(t_0) = 0,\ y'(t_0) = 0.$$

Per l'unicità della soluzione a tale problema di Cauchy si ha quindi che $y(t) = 0$ per ogni $t \in]a,b[$, e quindi esiste una combinazione lineare di y_1 e y_2 con coefficienti non entrambi nulli che si annulla identicamente su $]a,b[$. Quindi y_1 e y_2 sono linearmente dipendenti.

Viceversa, se y_1 e y_2 sono linearmente dipendenti, allora esistono $\tilde{c}_1, \tilde{c}_2$ non entrambe nulle tali che

$$\tilde{c}_1 y_1(t) + \tilde{c}_2 y_2(t) = 0 \quad \forall t \in]a,b[.$$

Derivando si ha anche

$$\tilde{c}_1 y_1'(t) + \tilde{c}_2 y_2'(t) = 0 \quad \forall t \in]a,b[.$$

Quindi il sistema lineare (8.25) ammette, quale che sia t_0, la soluzione $(\tilde{c}_1, \tilde{c}_2)$ con almeno una delle due componenti non nulla, e ciò può accadere solo se il determinante della matrice dei coefficienti è zero. Dunque il Wronskiano in ogni punto t è nullo. ■

Il prossimo punto sarà verificare, usando i risultati precedenti, che effettivamente *esistono* due soluzioni linearmente indipendenti.

Teorema 8.29 *L'equazione differenziale (8.23), con a_1, a_2 funzioni continue su un intervallo $]a,b[$, ammette due soluzioni linearmente indipendenti su $]a,b[$.*

Dimostrazione Per il Teorema di esistenza e unicità ci è noto che i problemi di Cauchy

$$\begin{cases} y'' + a_1(t)y' + a_2(t)y = 0 \\ y(t_0) = 0 \\ y'(t_0) = 1 \end{cases}$$

e

$$\begin{cases} y'' + a_1(t)y' + a_2(t)y = 0 \\ y(t_0) = 1 \\ y'(t_0) = 0 \end{cases}$$

ammettono soluzioni (uniche) y_1 e y_2 su $]a, b[$. Per definizione di Wronskiano si ha inoltre

$$\begin{aligned} W(t_0) \quad &:= y_1(t_0)y_2'(t_0) - y_2(t_0)y_1'(t_0) \\ &= -1 \neq 0. \end{aligned}$$

Quindi y_1 e y_2 sono linearmente indipendenti.

■

L'ultima questione da risolvere è la seguente: possiamo per caso trovare più di due soluzioni linearmente indipendenti? La risposta è negativa: effettivamente quindi lo spazio delle soluzioni ha dimensione due.

Teorema 8.30 *Siano y_1 e y_2 due soluzioni linearmente indipendenti dell'equazione (8.23), con a_1, a_2 funzioni continue su $]a, b[$. Allora ogni soluzione y dell'equazione (8.23) si può scrivere come combinazione lineare di y_1 e y_2, cioè esistono costanti $\tilde{c}_1, \tilde{c}_2$ tali che*

$$y(t) = \tilde{c}_1 y_1(t) + \tilde{c}_2 y_2(t) \quad \forall t \in]a, b[.$$

Dimostrazione Il Wronskiano di y_1 e y_2 è diverso da zero perché le soluzioni sono per ipotesi linearmente indipendenti. Detta y una qualunque soluzione dell'equazione studiata e fissato $t_0 \in]a, b[$, si consideri il sistema lineare, nelle variabili c_1, c_2

$$\begin{cases} c_1 y_1(t_0) + c_2 y_2(t_0) = y(t_0) \\ c_1 y_1'(t_0) + c_2 y_2'(t_0) = y'(t_0). \end{cases} \tag{8.26}$$

Il determinante della matrice dei coefficienti coincide con con Wronskiano $W(t_0)$ ed è quindi diverso da zero. Quindi si possono trovare costanti $\tilde{c}_1, \tilde{c}_2$ non entrambe nulle per cui il sistema lineare considerato è soddisfatto. Consideriamo la funzione

$$\tilde{y}(t) = \tilde{c}_1 y_1(t) + \tilde{c}_2 y_2(t) \quad \forall t \in]a, b[.$$

Essa è una soluzione dell'equazione differenziale di partenza e soddisfa per costruzione le condizioni

$$\tilde{y}(t_0) = y(t_0), \;\; \tilde{y}'(t_0) = y'(t_0).$$

Per l'unicità della soluzione al problema di Cauchy si deve quindi avere

$$y(t) = \tilde{y}(t) := \tilde{c}_1 y_1(t) + \tilde{c}_2 y_2(t) \quad \forall t \in]a, b[$$

■

Il problema che sarebbe bello saper risolvere ora è la risoluzione esplicita dell'equazione data. Tuttavia questo in generale non si sa fare se non in pochi casi particolari, il principale del quale vedremo nella prossima sezione. Vediamo però nella prossima osservazione che se si conosce una soluzione dell'equazione differenziale è possibile determinarne una seconda linearmente indipendente dalla prima in modo esplicito.

Osservazione (il metodo della riduzione dell'ordine). Supponiamo che y_1 sia una soluzione dell'equazione differenziale lineare omogenea del secondo ordine, con coefficienti continui

$$y'' + a_1(t)y'(t) + a_2(t)y = 0.$$

Cerchiamo una nuova soluzione del tipo

$$y_2(t) = A(t)y_1(t)$$

con $A(t)$ funzione di classe C^1 da determinarsi. Calcolando le derivate si ha

$$\begin{aligned} y_2' &= A'y_1 + Ay_1' \\ y_2'' &= A''y_1 + 2A'y_1' + Ay_1''. \end{aligned}$$

Imponendo che y_2 sia una soluzione dell'equazione considerata si ottiene subito, usando il fatto che y_1 è soluzione di tale equazione, che deve valere l'equazione differenziale

$$A''y_1 + A'(2y_1' + a_1y_1) = 0$$

dove dovete pensare a y_1 come a una funzione nota, e ad A come all'incognita. Questa è un'equazione differenziale lineare omogenea per A' ed ha come soluzione

$$A'(t) = Ce^{-\int(2y_1'+a_1y_1)\mathrm{d}t}$$

cioè

$$A(t) = C\int e^{-\int(2y_1'+a_1y_1)\mathrm{d}t}\mathrm{d}t + D$$

con C, D costanti arbitrarie. È facile vedere che la soluzione y_2 è linearmente indipendente da y_1 (perché si vede subito che $A(t)$ non può essere costante). Quindi una volta nota una soluzione, una seconda linearmente indipendente si trova facilmente.

Il procedimento ora descritto funziona anche per equazioni differenziali lineari di ordine $n > 2$.

Vediamo ora, senza dimostrazione, come si generalizzano alcuni dei risultati precedenti al caso di equazioni differenziali lineari omogenee di ordine n.

Teorema 8.31 *L'equazione differenziale*

$$y^{(n)} + a_1(t)y^{(n-1)} + \ldots + a_{n-1}(t)y' + a_n(t)y = 0$$

con $a_1, \ldots, a_n$ funzioni continue su un intervallo $]a, b[$, ammette esattamente n soluzioni linearmente indipendenti su $]a, b[$, cioè n soluzioni $y_1, \ldots, y_n$ tali che l'equazione

$$c_1y_1(t) + \ldots, c_ny_n(t) = 0$$

è vera per tutti i $t \in]a, b[$ se e soltanto se $c_1, \ldots, c_n = 0$. Ogni soluzione y dell'equazione data si può scrivere come combinazione lineare di y_1, y_n, cioè esistono costanti $d_1, \ldots, d_n$ tali che si può scrivere

$$y(t) = d_1y_1(t) + \ldots + d_ny_n(t) \quad \forall t \in]a, b[.$$

8.7.1 Equazioni non omogenee

Abbiamo visto che un'equazione differenziale lineare omogenea del tipo

$$y^{(n)} + a_1(t)y^{(n-1)} + \ldots a_n(t)y = 0 \tag{8.27}$$

con a_j funzioni continue per tutti gli indici j, ammette esattamente n soluzioni linearmente indipendenti $y_1, \ldots, y_n$, e che ogni soluzione di tale equazione si può scrivere come combinazione lineare di queste. Il problema sta nel fatto che tali n soluzioni non sono in generale esplicitamente calcolabili. Tuttavia, quando per qualche motivo fossero note le espressioni esplicite di $y_1, \ldots, y_n$ allora è possibile risolvere esplicitamente anche l'equazione *non omogenea*

$$y^{(n)} + a_1(t)y^{(n-1)} + \ldots a_n(t)y = b(t) \tag{8.28}$$

con b funzione continua. Il prossimo Teorema ci dice infatti come trovare, a meno del calcolo di alcune primitive, un integrale particolare della (8.28): si tratta del procedimento di *variazione delle costanti arbitrarie*, dovuto a Lagrange, che abbiamo già visto per le equazioni lineari del primo ordine.

Teorema 8.32 *Siano $y_1, \ldots, y_n$ n soluzioni linearmente indipendenti dell'equazione (8.27), con $a_1(t), \ldots, a_n(t)$ funzioni continue in un intervallo. Si supponga che il vettore*

$$\mathbf{c}(t) = \begin{pmatrix} c_1(t) \\ c_2(t) \\ \vdots \\ c_n(t) \end{pmatrix}$$

soddisfi l'equazione

$$\mathbf{c}'(t) = \mathbf{W}^{-1}(t) \begin{pmatrix} 0 \\ 0 \\ \vdots \\ b(t) \end{pmatrix} \tag{8.29}$$

dove $b(t)$ è una funzione continua, $\mathbf{W}(t)$ è la matrice Wronskiana delle soluzioni $y_1, \ldots, y_n$:

$$\mathbf{W}(t) := \begin{pmatrix} y_1(t) & y_2(t) & \ldots & y_n(t) \\ y_1'(t) & y_2'(t) & \ldots & y_n'(t) \\ \vdots & \vdots & & \vdots \\ y_1^{(n-1)}(t) & y_2^{(n-1)}(t) & \ldots & y_n^{(n-1)}(t) \end{pmatrix},$$

$\mathbf{W}^{-1}(t)$ ne indica la matrice inversa, e il prodotto a membro di destra è l'usuale prodotto "righe per colonne. Allora la funzione

$$\overline{y}(t) = c_1(t)y_1(t) + \ldots c_n(t)y_n(t)$$

è una soluzione dell'equazione non omogenea (8.28). In particolare la soluzione generale di tale equazione è la seguente:

$$y(t) = \overline{y}(t) + \alpha_1 y_1(t) + \ldots \alpha_n y_n(t)$$

con $\alpha_1, \ldots, \alpha_n$ costanti arbitrarie.

Dimostrazione Cerchiamo una soluzione della forma

$$y(t) = c_1(t)y_1(t) + \ldots c_n(t)y_n(t).$$

Abbiamo introdotto n funzioni arbitrarie e vogliamo che $y(t)$ sia soluzione di un'opportuna equazione differenziale: possiamo pensare quindi di poter imporre $(n-1)$ condizioni sulle funzioni $c_1(t), \ldots, c_n(t)$, anche se dovremo verificarne poi la consistenza con l'equazione differenziale che deve essere soddisfatta. Chiediamo allora che valgano le condizioni

$$\begin{array}{l} c_1'y_1 + \ldots + c_n'y_n = 0 \\ c_1'y_1' + \ldots + c_n'y_n^1\prime = 0 \\ \vdots \\ c_1'y_1^{(n-2)} + \ldots + c_n'y_n^{(n-2)} = 0. \end{array} \tag{8.30}$$

Se esse sono soddisfatte si ha allora

$$\begin{aligned} y'(t) &= c_1'(t)y_1(t) + c_1(t)y_1'(t) + \ldots + c_n'(t)y_n(t) + c_n(t)y_n'(t) \\ &= c_1(t)y_1'(t) + \ldots + c_n(t)y_n'(t) \end{aligned} \tag{8.31}$$

e, con calcoli analoghi,

$$y^{(k)}(t) = c_1(t)y_1^{(k)}(t) + \ldots + c_n(t)y_n^{(k)}(t) \quad \forall k = 1, \ldots, n-1.$$

Si ha anche

$$\begin{aligned} y^{(n)}(t) &= c_1(t)y_1^{(n)}(t) + \ldots + c_n(t)y_n^{(n)}(t) \\ &+ c_1'(t)y_1^{(n-1)}(t) + \ldots + c_n'(t)y_n^{(n-1)}(t). \end{aligned}$$

Calcoliamo allora, utilizzando le formule appena scritte per le derivate di y,

$$\begin{aligned} & y^{(n)}(t) + a_1(t)y^{(n-1)}(t) + \ldots + a_{n-1}(t)y'(t) + a_n(t)y(t) \\ & = c_1(t)y_1^{(n)}(t) + \ldots + c_n(t)y_n^{(n)}(t) \\ & + c_1'(t)y_1^{(n-1)}(t) + \ldots + c_n'(t)y_n^{(n-1)}(t) \\ & + a_1(t)\left[c_1(t)y_1^{(n-1)}(t) + \ldots + c_n(t)y_n^{(n-1)}(t)\right] \\ & \vdots \\ & + a_{n-1}(t)\left[c_1(t)y_1'(t) + \ldots + c_n(t)y_n'(t)\right] \\ & + a_n(t)\left[c_1(t)y_1(t) + \ldots c_n(t)y_n(t)\right]. \end{aligned}$$

Ricordiamo ora che $y_1, \ldots, y_n$ sono soluzioni dell'equazione differenziale omogenea (8.27). Quindi:

$$\begin{aligned} & y^{(n)}(t) + a_1(t)y^{(n-1)}(t) + \ldots + a_{n-1}(t)y'(t) + a_n(t)y(t) \\ & = c_1'(t)y_1^{(n-1)}(t) + \ldots + c_n'(t)y_n^{(n-1)}(t). \end{aligned}$$

Per far sì che $y(t)$ sia soluzione dell'equazione (8.28), basta quindi chiedere che

$$c_1'(t)y_1^{(n-1)}(t) + \ldots + c_n'(t)y_n^{(n-1)}(t) = b(t).$$

Aggiungendo quest'ultima equazione alle condizioni (8.30) già imposte e ricordando la definizione di matrice Wronskiana, dovremo dunque avere

$$\mathbf{W}(t)\mathbf{c}'(t) = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ b(t) \end{pmatrix}$$

La matrice Wronskiana è invertibile, essendo $y_1, \ldots y_n$ soluzioni linearmente indipendenti dell'equazione omogenea (abbiamo visto questo per equazioni del secondo ordine, ma non ci sono differenze di rilievo anche nel caso di equazioni di ordine $n > 2$). Quindi, se vale la (8.29), la funzione $\overline{y}(t)$ è soluzione dell'equazione non omogenea, come affermato. ■

La situazione è particolarmente semplice per equazioni del secondo ordine.

Corollario 8.33 *Siano y_1, y_2 sono soluzioni linearmente indipendenti dell'equazione differenziale*

$$y'' + a_1(t)y' + a_2(t)y = 0$$

con a_1, a_2 funzioni continue su un intervallo, e si supponga che le funzioni c_1, c_2 soddisfino le equazioni

$$c_1'(t) = -\frac{y_2(t)b(t)}{y_1(t)y_2'(t) - y_2(t)y_1'(t)}$$
$$c_2'(t) = \frac{y_1(t)b(t)}{y_1(t)y_2'(t) - y_2(t)y_1'(t)}.$$

Allora la funzione

$$\overline{y}(t) = c_1(t)y_1(t) + c_2(t)y_2(t)$$

è soluzione dell'equazione differenziale

$$y'' + a_1(t)y' + a_2(t)y = b(t).$$

Quindi l'integrale generale di quest'ultima equazione è

$$y(t) = \overline{y}(t) + \alpha_1 y_1(t) + \alpha_2 y_2(t)$$

con α_1, α_2 costanti arbitrarie.

Dimostrazione È sufficiente notare che la matrice Wronskiana è per definizione

$$\mathbf{W}(t) = \begin{pmatrix} y_1(t) & y_2(t) \\ y_1'(t) & y_2'(t) \end{pmatrix}$$

e che quindi, calcolando esplicitamente l'inversa con le ben note regole:

$$W^{-1}(t) = \frac{1}{y_1(t)y_2'(t) - y_2(t)y_1'(t)} \begin{pmatrix} y_2'(t) & -y_2(t) \\ -y_1'(t) & y_1(t) \end{pmatrix}$$

■

Dunque basta calcolare l'inversa di una matrice e alcune primitive per poter trovare una soluzione particolare dell'equazione non omogenea, sempre naturalmente che siano note n soluzioni linearmente indipendenti dell'equazione omogenea. Sottolineiamo che la cosa migliore *non è* imparare a memoria le formule di cui sopra, ma apprendere il procedimento e saper ricostruire le formule necessarie.

Vediamo ora un esempio estremamente importante: *l'oscillatore armonico con termine forzante*

Esempio 139 Consideriamo l'equazione differenziale

$$y'' + \omega^2 y(t) = b(t)$$

dove ω è una costante positiva e $b(t)$ una funzione continua. Se $b(t) \equiv 0$ allora abbiamo a che fare con *l'oscillatore armonico libero*, e cioè (riguardate le considerazioni all'inizio di questo capitolo) con l'equazione del moto di un corpo soggetto a una forza elastica in assenza di attrito e di termini forzanti. Vedremo tra non molto (ma ciò è già stato detto negli esempi iniziali) che le soluzioni dell'equazione *omogenea*

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

sono le funzioni

$$y(t) = A\sin(\omega t) + B\cos(\omega t)$$

con A, B costanti arbitrarie. Il moto è quindi oscillatorio di frequenza ω. Prendiamo dunque

$$y_1(t) = \sin(\omega t), \quad y_2(t) = \cos(\omega t)$$

e applichiamo la procedura vista nei risultati precedenti. La matrice Wronskiana è:

$$\mathbf{W}(t) = \begin{pmatrix} \sin(\omega t) & \cos(\omega t) \\ \omega\cos(\omega t) & -\omega\sin(\omega t) \end{pmatrix}$$

e tale matrice ha determinante pari a $-\omega$. La matrice inversa quindi è la seguente:

$$\mathbf{W}^{-1}(t) = \begin{pmatrix} \sin(\omega t) & \omega^{-1}\cos(\omega t) \\ \cos(\omega t) & -\omega^{-1}\sin(\omega t). \end{pmatrix}.$$

Quindi si ha

$$\mathbf{W}^{-1}(t)\begin{pmatrix} 0 \\ b(t) \end{pmatrix} = \begin{pmatrix} \omega^{-1}b(t)\cos(\omega t) \\ -\omega^{-1}b(t)\sin(\omega t) \end{pmatrix}.$$

Fissato $t_0 \in \mathbb{R}$, si possono dunque scegliere $c_1(t)$ e $c_2(t)$ come segue:

$$c_1(t) = \frac{1}{\omega}\int_{t_0}^{t} b(s)\cos(\omega s)\mathrm{d}s$$

$$c_2(t) = -\frac{1}{\omega}\int_{t_0}^{t} b(s)\sin(\omega s)\mathrm{d}s.$$

In conclusione abbiamo che una soluzione dell'equazione considerata è la seguente:

$$\overline{y}(t) = \frac{\sin(\omega t)}{\omega}\int_{t_0}^{t} b(s)\cos(\omega s)\mathrm{d}s - \frac{\cos(\omega t)}{\omega}\int_{t_0}^{t} b(s)\sin(\omega s)\mathrm{d}s.$$

Si noti che vale allora anche la semplice formula

$$\overline{y}(t) = \frac{1}{\omega}\int_{t_0}^{t} \sin[\omega(t-s)]b(s)\mathrm{d}s.$$

Quest'ultimo risultato è notevole in quanto ci dà esplicitamente, quale che sia la forzante esterna $b(t)$, una soluzione esplicita all'equazione considerata (a meno del calcolo di una primitiva). L'integrale a secondo membro prende a volte il nome di *convoluzione* tra la funzione b e la funzione $g(u) = \sin(\omega u)$. Rivedrete certamente questo concetto, che non può essere qui discusso oltre, in corsi più avanzati.

8.8 Equazioni lineari a coefficienti costanti

La principale classe di equazioni differenziali lineari per cui è possibile trovare esplicitamente soluzioni è quella delle equazioni a coefficienti costanti. Considereremo quindi qui dapprima equazioni differenziali *omogenee* della forma

$$y^{(n)} + a_1 y^{(n-1)} + \ldots + a_{n-1}y' + a_n y = 0$$

con $a_1, \ldots a_n$ costanti reali. Sappiamo già, per i risultati generali delle sezioni precedenti, che ogni problema di Cauchy associato all'equazione considerata ammette una sola soluzione definita per tutti i valori di t. Per determinare esplicitamente le soluzioni l'idea è semplicissima: cerchiamo soluzioni di tipo esponenziale:

$$y(t) = e^{\lambda t}$$

e vediamo se si può trovare qualche costante λ per cui la funzione assegnata è davvero una soluzione. Facendo le derivate e imponendo che y sia una soluzione si ottiene

$$\left[\lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n\right] e^{\lambda t} = 0$$

cioè, visto che $e^{\lambda t} \neq 0$ per ogni t l'equazione, nella variabile λ:

$$\lambda^n + a_1\lambda^{n-1} + \ldots + a_{n-1}\lambda + a_n = 0. \tag{8.32}$$

Questa equazione si dice *equazione caratteristica* associata all'equazione differenziale data. Ben sappiamo che un'equazione di grado n ammette sempre, nel campo complesso, esattamente n radici (se contate con la loro molteplicità). Vediamo i vari casi che si possono presentare.

Radici reali e distinte. In questo caso ci saranno n numeri reali $\lambda_1, \ldots, \lambda_n$ tali che le funzioni

$$y_1(t) = e^{\lambda_1 t}, y_2(t) = e^{\lambda_2 t}, \ldots, y_n(t) = e^{\lambda_n t}$$

sono soluzioni dell'equazione data. Si vede (esercizio $**$) che le soluzioni scritte sono linearmente indipendenti. Farlo direttamente non è semplice, ma si può verificare che il Wronskiano non si annulla. L'integrale generale sarà dunque

$$y(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t} + \ldots + c_n e^{\lambda_n t}$$

con $c_1, \ldots c_n \in \mathbb{R}$.

Esempio 140 Si consideri l'equazione differenziale

$$y''' - 2y' = 0.$$

L'equazione caratteristica è $\lambda^3 - 2\lambda = 0$ che ammette le soluzioni $\lambda = 0$, $\lambda = \pm\sqrt{2}$. La soluzione generale dell'equazione differenziale data è allora

$$y(t) = c_1 + c_2 e^{\sqrt{2}t} + c_3 e^{-\sqrt{2}t}.$$

Radici reali, alcune delle quali con molteplicità maggiore di uno. Se un radice λ ha molteplicità maggiore di uno allora abbiamo trovato “troppo poche soluzioni. Si può però vedere con un calcolo diretto che, se la molteplicità della radice λ è $k > 1$, allora le funzioni

$$y_1(t) = e^{\lambda t},\ y_2(t) = te^{\lambda t}, \ldots y_k(t) = t^{k-1}e^{\lambda t}$$

sono k soluzioni dell'equazione considerata. Si può anche mostrare, dette $m_1, \ldots, m_k$ le molteplicità delle radici $\lambda_1, \ldots, \lambda_k$ rispettivamente (con $m_1 + \ldots + m_k = n$), allora le n soluzioni così trovate:

$$e^{\lambda_1 t},\ te^{\lambda_1 t}, \ldots, t^{m_1 - 1}e^{\lambda_1 t}, \ldots, e^{\lambda_k t},\ te^{\lambda_k t}, \ldots, t^{m_k - 1}e^{\lambda_k t}$$

sono linearmente indipendenti. Quindi la soluzione generale dell'equazione differenziale data è:

$$y(t) = c_{1,1}e^{\lambda_1 t} + \ldots + c_{1,m_1}t^{m_1-1}e^{\lambda_1 t} + \ldots + c_{k,1}e^{\lambda_k t} + \ldots + c_{k,m_k}t^{m_k-1}e^{\lambda_k t}.$$

Esempio 141 Consideriamo l'equazione differenziale

$$y^{(4)} - 2y'' + y = 0.$$

L'equazione caratteristica è $\lambda^4 - 2\lambda^2 + 1 = 0$, che ha come soluzioni $\lambda_1 = 1, \lambda_2 = -1$. Entrambe le radici hanno molteplicità due, e quindi l'integrale generale dell'equazione data è:

$$y(t) = c_1 e^t + c_2 te^t + c_3 e^{-t} + c_4 te^{-t}.$$

Alcune delle radici hanno parte immaginaria diversa da zero. In questo caso sia $\lambda = a + ib$ una di tali radici ($a, b \in \mathbb{R}$). Poiché l'equazione caratteristica ha coefficienti reali vi deve anche essere la radice complessa coniugata $\overline{\lambda} = a - ib$. Vi sarebbero quindi le soluzioni

$$y_1(t) = e^{(a+ib)t},\ y_2(t) = e^{(a-ib)t}$$

che possono essere riscritte, per la formula di Eulero, come

$$y_1(t) = e^{at}(\cos(bt) + i\sin(bt)),\ y_2(t) = e^{at}(\cos(bt) - i\sin(bt)).$$

Sebbene queste funzioni siano effettivamente soluzioni dell'equazione differenziale considerata, il fatto che abbiano una parte immaginaria diversa da zero è un po' fastidioso, visto che siamo partiti da un'equazione a coefficienti reali. C'è però un semplice rimedio: visto che sappiamo che le combinazioni lineari di soluzioni sono ancora soluzioni, sommiamo le due soluzioni (e dividiamo per due) ottenendo la nuova soluzione

$$\tilde{y}_1(t) = e^{at}\cos(bt).$$

Sottraendo le due soluzioni (e dividendo per $-2i$) si ricava poi la nuova soluzione

$$\tilde{y}_1(t) = e^{at}\sin(bt).$$

Si può dimostrare che queste soluzioni sono linearmente indipendenti e che lo è il sistema complessivo delle soluzioni individuate.

Questo permette di trovare le due soluzioni indipendenti associate alla coppia di radici complesse coniugate $a \pm ib$. Se tali radici sono semplici, non vi è altro da dire perché a tale coppia di radici sono associate esattamente due soluzioni indipendenti. Se invece ognuna di tali radici ha molteplicità $k > 1$ allora occorre come in precedenza considerare le seguenti funzioni

$$\begin{array}{l} e^{at}\cos(bt), te^{at}\cos(bt), t^2e^{at}\cos(bt), \ldots, t^{k-1}e^{at}\cos(bt) \\ e^{at}\sin(bt), te^{at}\sin(bt), t^2e^{at}\sin(bt), \ldots, t^{k-1}e^{at}\sin(bt). \end{array}$$

Si tratta di $2k$ funzioni: tale numero coincide col numero esatto di soluzioni linearmente indipendenti associate alla coppia di radici considerata. Si può mostrare in effetti che ognuna di tali funzioni è una soluzione dell'equazione data, e che si tratta di soluzioni tra loro linearmente indipendenti, così come è linearmente indipendente il sistema di soluzioni costituito, oltre che dalle funzioni appena scritte, dalle soluzioni corrispondenti ad eventuali altre radici dell'equazione caratteristica.

Esempio 142 Studiamo ora un esempio fondamentale cui abbiamo già accennato nella sezione precedente: *l'oscillatore armonico.* Consideriamo in effetti l'equazione di Newton per un corpo soggetto a una forza elastica:

$$y'' + \omega^2 y = 0.$$

L'equazione caratteristica è $\lambda^2 + \omega^2 = 0$ che ha come radici $\lambda = \pm i\omega$. Quindi, con la notazione precedente, $a = 0$ e $b = \omega$ e si ha l'integrale generale

$$y(t) = A\cos(\omega t) + B\sin(\omega t).$$

Si tratta di un moto periodico con frequenza ω e periodo $T = 2\pi/\omega$.

Esempio 143 Naturalmente i casi precedenti possono verificarsi anche contemporaneamente. Si consideri per esempio l'equazione

$$y^{(4)} + 2y^{(3)} + 2y'' + 2y' + 2y = 0.$$

L'equazione caratteristica è $\lambda^4 + 2\lambda^3 + 2\lambda^2 + 2\lambda + 1 = 0$. Risolvere un'equazione di quarto grado non è in generale così agevole: tuttavia in questo caso si vede subito che l'equazione caratteristica si può scrivere come

$$(\lambda + 1)^2(\lambda^2 + 1) = 0$$

che ha come radici $\lambda = -1$ con molteplicità due e $\lambda = \pm i$. Quindi la soluzione generale all'equazione differenziale data è la seguente:

$$y(t) = Ae^{-t} + Bte^{-t} + C\sin t + D\cos t.$$

8.8.1 Equazioni non omogenee

Una volta determinata la soluzione all'equazione omogenea il metodo di variazione delle costanti arbitrarie, già visto in precedenza, ci permette in linea di principio di trovare una soluzione dell'equazione non omogenea, e quindi la soluzione generale di questa. In alcuni casi questa procedura non è però necessaria. Si consideri in effetti un'equazione differenziale del tipo

$$y^{(n)} + c_1 y^{(n-1)} + \ldots + c_{n-1} y' + c_n y = p_h(t) e^{\lambda t} \tag{8.33}$$

con p_h polinomio di grado h. Si può dimostrare che, se λ non è soluzione dell'equazione caratteristica relativa all'equazione omogenea, allora esiste *sempre* una soluzione di (8.33) del tipo

$$\overline{y}(t) = q_h(t) e^{\lambda t}$$

dove q_h è un polinomio (da determinarsi) di grado h.

Esempio 144 Consideriamo l'equazione differenziale

$$y'' + y = te^{-t}.$$

Si tratta di un oscillatore armonico con termine forzante. $\lambda = -1$ non è soluzione dell'equazione caratteristica dell'omogenea (che ha, ricordiamolo, le soluzioni $\lambda_{1,2} = \pm i$). Dunque esiste senz'altro una soluzione particolare della non omogenea del tipo

$$\overline{y}(t) = (At + B)e^{-t}$$

con A e B da determinarsi. Si calcola

$$\overline{y}''(t) = (B - 2A + At)e^{-t}$$

e quindi affinché l'equazione richiesta sia soddisfatta deve valere

$$e^{-t}(B - 2A + At) + e^{-t}(At + B) = te^{-t}.$$

Semplificando gli esponenziali e uguagliando i coefficienti di medesimo grado a primo e a secondo membro si ottiene $A = 1/2$, $B = 1/2$. Quindi una soluzione particolare dell'equazione considerata è

$$\overline{y}(t) = \frac{t+1}{2} e^{-t}.$$

La soluzione generale dell'equazione è quindi

$$y(t) = \frac{t+1}{2} e^{-t} + \alpha_1 \cos t + \alpha_2 \sin t$$

con α_1, α_2 costanti arbitrarie.

Se invece λ è radice del polinomio caratteristico di molteplicità k, si può mostrare che esiste sempre una soluzione del tipo

$$\overline{y}(t) = t^k q_h(t) e^{\lambda t}$$

con q_h polinomio di grado h da determinarsi.

Esempio 145 Consideriamo l'equazione differenziale

$$y'' + y' = t.$$

Qui non compaiono esponenziali, ma possiamo descrivere questo caso ponendo $\lambda = 0$. Siccome $\lambda = 0$ è radice semplice ($k = 1$) dell'equazione caratteristica, allora occorrerà trovare una soluzione particolare del tipo

$$\overline{y}(t) = t(At + B)$$

con A e B costanti da determinarsi. Calcoli elementari mostrano che $\overline{y}$ è soluzione se e solo se

$$2At + 2A + B = t \quad \forall t.$$

Quindi deve essere $A = 1/2$, $B = -1$ e la soluzione particolare è

$$\overline{y}(t) = \frac{t^2}{2} - t.$$

Calcolando anche le soluzioni dell'equazione omogenea ($\lambda = 0, \lambda = -1$) si ottiene l'integrale generale

$$y(t) = \frac{t^2}{2} - t + \alpha_1 + \alpha_2 e^{-t}$$

con α_1, α_2 costanti reali.

La procedura precedente può essere utilmente utilizzata anche quando, invece di esponenziali, a secondo membro compaiano seni o coseni. La cosa non è sorprendente: tali funzioni sono parti immaginarie o reali di esponenziali complessi. In effetti, quando si considerano equazioni differenziali del tipo

$$y^{(n)} + c_1 y^{(n-1)} + \ldots + c_{n-1} y' + c_n y = p_h(t) e^{at} \cos(bt) \tag{8.34}$$

oppure

$$y^{(n)} + c_1 y^{(n-1)} + \ldots + c_{n-1} y' + c_n y = p_h(t) e^{at} \sin(bt) \tag{8.35}$$

e p_h è, come prima, un polinomio di grado h, si può dimostrare che, se

$$\lambda := a + ib$$

non è radice del polinomio caratteristico, allora esiste sempre una soluzione del tipo
$$\overline{y}(t) = q_{1,h}(t)e^{at}\cos(bt) + q_{2,h}(t)e^{at}\sin(bt),$$
mentre se λ è radice del polinomio caratteristico di molteplicità r allora esiste una soluzione del tipo

$$\overline{y}(t) = t^r \left[q_{1,h}(t)e^{at}\cos(bt) + q_{2,h}(t)e^{at}\sin(bt)\right]$$

dove in entrambi i casi $q_{1,h}, q_{2,h}$ sono polinomi di grado h da determinarsi.

Si può però procedere in un altro modo, spesso più semplice dal punto di vista dei calcoli. Cominciamo col ricordare il *principio di sovrapposizione*: se $y_1(t)$ è una soluzione dell'equazione

$$y^{(n)} + c_1 y^{(n-1)} + \ldots + c_{n-1}y' + c_n y = b_1(t) \tag{8.36}$$

e $y_2(t)$ è una soluzione dell'equazione

$$y^{(n)} + c_1 y^{(n-1)} + \ldots + c_{n-1}y' + c_n y = b_2(t) \tag{8.37}$$

allora $y_1(t) + y_2(t)$ è soluzione dell'equazione

$$y^{(n)} + c_1 y^{(n-1)} + \ldots + c_{n-1}y' + c_n y = b_1(t) + b_2(t). \tag{8.38}$$

Vediamo su un esempio come questo fatto può essere applicato allo studio di equazioni con termini forzanti che coinvolgono seni e coseni.

Esempio 146 Consideriamo l'equazione differenziale

$$y'' + y = \cos(at),$$

con $a > 0$. Si tratta di un *oscillatore armonico forzato*, con forzante periodica di periodo $T = 2\pi/a$. Abbiamo posto per semplicità uguale a uno la pulsazione propria ω dell'oscillatore libero. Notiamo ora che, come ben sappiamo

$$\cos(at) = \frac{e^{iat} + e^{-iat}}{2}.$$

Cerchiamo quindi separatamente una soluzione y_1 all'equazione

$$y'' + y = \frac{e^{iat}}{2}$$

e una soluzione y_2 all'equazione

$$y'' + y = \frac{e^{-iat}}{2}.$$

Si devono distinguere due casi: in effetti $\lambda := \pm ia$ sono soluzioni dell'equazione caratteristica se e solo se $a = 1$ (a è positivo per ipotesi). Se $a \neq 1$ si può quindi trovare una soluzione della prima equazione del tipo $y_1(t) = Ae^{iat}$ e una soluzione della seconda del tipo $y_2(t) = Be^{-iat}$. Con calcoli immediati si trova, imponendo che le funzioni scelte siano effettivamente soluzioni,

$$A = B = \frac{1}{2(1-a^2)}.$$

Si noti che quanto appena scritto perde senso per $a = 1$: questo significa solamente che in tal caso, come già detto, non ci sono soluzioni di tale tipo. Quindi, se $a \neq 1$:

$$\begin{aligned} \overline{y}(t) &= y_1(t) + y_2(t) \\ &= \frac{1}{2(1-a^2)} \left(e^{iat} + e^{-iat}\right) \\ &= \frac{\cos t}{(1-a^2)}. \end{aligned}$$

Che la soluzione contenga solo la funzione coseno è in un certo senso casuale, o meglio è legato al fatto che nell'equazione differenziale cercata non ci sia il termine y'. In generale ci si deve aspettare che compaia anche la funzione seno *anche quando nel termine forzante compaia solo la funzione coseno.* La soluzione generale dell'equazione non omogenea è quindi

$$y(t) = \alpha_1 \cos t + \alpha_2 \sin t + \frac{\cos t}{(1-a^2)}$$

con α_1, α_2 costanti arbitrarie. Si noti che la soluzione è una funzione periodica e quindi, essendo anche continua, una funzione limitata.

Se invece $a = 1$ si devono cercare soluzioni del tipo

$$y_1(t) = Ate^{it}, \quad y_2(t) = Bte^{-it}.$$

Calcolando esplicitamente le derivate e imponendo che le corrispondenti equazioni differenziali siano soddisfatte si ottiene

$$A = \frac{1}{4i}, \quad B = -\frac{1}{4i}.$$

Quindi una soluzione particolare dell'equazione differenziale studiata è la seguente:

$$\begin{aligned}\overline{y}(t) &= \frac{t}{4i}\left(e^{it} - e^{-it}\right) \\ &= \frac{t\sin t}{2}.\end{aligned}$$

La soluzione generale è quindi

$$y(t) = \alpha_1 \cos t + \alpha_2 \sin t + \frac{t \sin t}{2}.$$

Le soluzioni così ottenute *non* sono periodiche e non sono neppure funzioni limitate. Informalmente parlando si può dire che l'ampiezza delle oscillazioni cresce col crescere del tempo. È questo il ben noto fenomeno fisico della *risonanza*: se la forzante è una funzione periodica di frequenza pari a quella propria del sistema, l'ampiezza delle oscillazioni cresce col tempo fino ad effetti potenzialmente distruttivi.

Esercizio 8.6 Trovare le soluzioni generali delle seguenti equazioni differenziali:

$$\begin{aligned}&y'' + y' + y = 1;\\ &y^{(3)} + y' = \sin t + \cos t;\\ &y'' - 2y' + y = e^t;\\ &\tfrac{1}{4}y^{(4)} + y^{(3)} + 5y'' = 4.\end{aligned}$$

R. $y(t) = e^{-t/2}\left[\alpha_1 \sin\left(\frac{\sqrt{3}}{2}t\right) + \alpha_2 \cos\left(\frac{\sqrt{3}}{2}t\right)\right] + 1$;
$y(t) = \alpha_1 + \alpha_2 \sin t + \alpha_3 \cos t - \frac{t}{2}(\sin t + \cos t)$;
$y(t) = e^t\left(\alpha_1 + \alpha_2 t + \frac{1}{2}t^2\right)$;
$y(t) = \alpha_1 + \alpha_2 t + e^{-2t}[\alpha_3 \sin(2\sqrt{2}t) + \alpha_4 \cos(2\sqrt{2}t)] + \frac{2}{5}t^2$.

8.9 Integrazione per serie

Consideriamo in questa sezione equazioni differenziali lineari del secondo ordine *a coefficienti non costanti* del tipo

$$y'' + a_1(t)ty' + a_2(t)y = b(t).$$

Mostreremo con alcuni esempi un possibile procedimento di risoluzione nel caso si abbiano condizioni di regolarità abbastanza forti sui coefficienti. Richiederemo che tali coefficienti siano *sviluppabili in serie di potenze* in un opportuno intervallo. Abbiamo detto in precedenza che condizioni di regolarità sui coefficienti si ripercuotono in condizioni di regolarità sulle soluzioni, e in questo caso si ha in effetti il seguente Teorema, che descrive un risultato profondo e che non potremo dimostrare.

Teorema 8.34 *Se per ogni $x_0 \in]a, b[$ le funzioni a_1, a_2, b sono sviluppabili in serie di potenze centrata in x_0 e con raggio di convergenza non nullo, allora ogni soluzione dell'equazione differenziale considerata è anch'essa, per ogni $x_0 \in]a, b[$, sviluppabile in serie di potenze centrata in x_0 e con raggio di convergenza non nullo.*

Le funzioni con la proprietà richiesta nel Teorema si dicono *funzioni analitiche (in senso reale) su* $]a, b[$. Il Teorema dice allora che se i coefficienti dell'equazione sono analitici in un intervallo, anche le soluzioni lo sono.

Vorremo qui mostrare come, in casi particolarmente semplici, sia possibile scrivere in modo abbastanza esplicito lo sviluppo in serie. L'idea è molto semplice: scrivere la soluzione in serie di potenze, calcolare le derivate con il Teorema di derivazione delle serie di potenze, e identificare i coefficienti in base all'equazione che deve essere soddisfatta. Bisognerà usare il cosiddetto *principio di identità delle serie di potenze*, che già conoscete e che dice che l'uguaglianza

$$\sum_{n=0}^{+\infty} a_n(t-t_0)^n = \sum_{n=0}^{+\infty} b_n(t-t_0)^n$$

vale per tutti i t per i quali le serie convergono *se e soltanto se $a_n = b_n$ per ogni n.*

Vediamo come si lavora in pratica su due esempi particolarmente significativi.

Esempio 147 Consideriamo *l'equazione di Hermite*

$$y'' - ty' + \lambda y = 0$$

con $\lambda \in \mathbb{R}$. Poiché i coefficienti sono analitici su $\mathbb{R}$, anche la soluzione lo sarà. Cerchiamo quindi una soluzione del tipo

$$y(t) = \sum_{n=0}^{+\infty} a_n t^n$$

ma dovremo controllare a posteriori quale sia il raggio di convergenza della serie stessa. Il fatto che la serie sia centrata in $t = 0$ è dettato dalla forma dei coefficienti: siccome $a_1(t) = t$ sarà più semplice fare i calcoli centrando la serie nell'origine. Se la serie ha raggio di convergenza non nullo, come peraltro sappiamo essere vero per il Teorema precedente, possiamo usare il Teorema di derivazione per serie per mostrare che

$$y'(t) = \sum_{n=1}^{+\infty} n a_n t^{n-1}$$
$$y''(t) = \sum_{n=2}^{+\infty} n(n-1) a_n t^{n-2}.$$

Si ha quindi

$$ty'(t) = \sum_{n=1}^{+\infty} n a_n t^n.$$

Notiamo poi che nella serie appena scritta l'indice n può anche partire da $n = 0$ (il termine aggiunto è nullo), e che un cambiamento degli indici di sommatoria ci consente di scrivere

$$y''(t) = \sum_{n=0}^{+\infty} (n+2)(n+1) a_{n+2} t^n.$$

Quindi l'equazione data diventa

$$\sum_{n=0}^{+\infty} \left[(n+2)(n+1)a_{n+2} + (\lambda - n)a_n\right] t^n.$$

Per il principio di identità delle serie di potenze questo può aver luogo se e soltanto se i coefficienti della serie sono tutti nulli. Quindi dovrà valere la relazione di ricorrenza

$$a_{n+2} = \frac{(n-\lambda)}{(n+2)(n+1)} a_n.$$

Questa relazione permette di ricavare tutti i coefficienti della serie conoscendone i primi due. Sappiamo che dobbiamo trovare due soluzioni linearmente indipendenti: un'idea sarebbe di cercare una soluzione pari e una dispari, in modo che l'indipendenza lineare sia ovvia. Questo si può fare scegliendo dapprima $a_0 = 1, a_1 = 0$, ottenendo così che tutti i termini con indice dispari sono nulli mentre quelli con indice pari valgono (tralasciamo i semplici calcoli)

$$a_{2n} = -a_0 \frac{\lambda(2-\lambda)(4-\lambda)\cdots(2n-2-\lambda)}{(2n)!} \quad \forall n \geq 1.$$

Il criterio del rapporto mostra che la serie

$$y_1(t) := \sum_{n=0}^{+\infty} a_{2n} t^{2n}$$

ha raggio di convergenza infinito: essa rappresenta una prima soluzione all'equazione differenziale data, ed è una funzione pari. La seconda si può ottenere scegliendo $a_0 = 0, a_1 = 1$ e ottenendo con un procedimento analogo che tutti i termini di indice pari sono nulli mentre quelli di indice dispari valgono

$$a_{2n+1} = a_1 \frac{(1-\lambda)(3-\lambda)\cdots(2n-1-\lambda)}{(2n)!} \quad \forall n \geq 1.$$

Di nuovo la serie

$$y_2(t) := \sum_{n=0}^{+\infty} a_{2n+1} t^{2n+1}$$

ha raggio di convergenza infinito e rappresenta una seconda soluzione, questa volta dispari, all'equazione differenziale di partenza. Le due soluzioni sono per costruzione linearmente indipendenti: per vederlo formalmente notate che se

$$c_1 y_1(t) + c_2 y_2(t) = 0 \quad \forall t \in \mathbb{R}$$

allora si ha anche, scrivendo l'equazione precedente per il tempo $-t$ e usando le simmetrie appena notate, che

$$c_1 y_1(t) - c_2 y_2(t) = 0 \quad \forall t \in \mathbb{R}.$$

Sommando le due equazioni si ha dunque che per tutti i tempi vale $2c_1 y_1(t) = 0$. Questo significa, visto che $y_1(t)$ non è identicamente nulla, che $c_1 = 0$. Analogamente si mostra che $c_2 = 0$ e quindi che y_1 e y_2 sono linearmente indipendenti. Tutte le soluzioni all'equazione data sono quindi del tipo

$$y(t) = \alpha_1 y_1(t) + \alpha_2 y_2(t)$$

con α_1, α_2 costanti arbitrarie. Si noti che y_1 e y_2 *non sono scrivibili in termini di funzioni elementari*: di esse è noto solo lo sviluppo in serie.

Esempio 148 Il prossimo esempio mostrerà come trovare una soluzione della cosiddetta *equazione di Bessel*

$$y'' + \frac{y'}{t} + \left(1 - \frac{m^2}{t^2}\right) y = 0$$

dove $m \in \mathbb{N}$. I coefficienti dell'equazione *non sono analitici* (in realtà non sono nemmeno definiti per $t = 0$...) e quindi non possiamo certo usare il Teorema 8.34. Possiamo tuttavia *tentare*, senza alcuna certezza di successo, di trovare egualmente una soluzione che si possa scrivere nella forma

$$y(t) = t^k \sum_{n=0}^{+\infty} a_n t_n$$

con la costante k e i coefficienti a_n da determinarsi. Svolgendo i calcoli come nell'esempio precedente si ottiene

$$\begin{aligned} t^{k-2} \Big[& (k^2 - m^2) a_0 + \big((k+1)^2 - m^2\big) a_1 t \\ & + \sum_{n=2}^{+\infty} \left[((k+n)^2 - m^2) a_n + a_{n-2} \right] t^n \Big] = 0. \end{aligned}$$

Poiché i coefficienti di tutte le potenze di t devono essere nulli se ne ottiene

$$k = m, \ a_1 = 0, \ a_n = -\frac{a_{n-2}}{(k+n)^2 - m^2} \quad \forall n \geq 2.$$

Rimane quindi un solo grado di libertà, il valore di a_0. È abbastanza semplice vedere che le relazioni appena scritte sono soddisfatte se si pongono uguali a zero tutti i coefficienti di indice dispari e, posto $a_0 = 1/(2^m m!)$:

$$a_{2n} = \frac{(-1)^n}{2^{m+2n} n! (m+n)!} \quad \forall n \geq 0.$$

Se ne ottiene la soluzione

$$J_m(t) = \sum_{n=0}^{+\infty} (-1)^n \frac{(t/2)^{m+2n}}{n!(m+n)!}$$

Questa funzione analitica (il raggio di convergenza della serie è chiaramente $+\infty$) si chiama *funzione di Bessel di prima specie.* La soluzione è definita e analitica anche per $t = 0$, che era un punto in cui i coefficienti dell'equazione di partenza non erano neppure definiti. Questo mostra che tale soluzione può essere *prolungata* fino a $t = 0$ incluso, e quindi che possono esistere soluzioni analitiche anche di equazioni con coefficienti singolari. Naturalmente tutti i multipli di J_m sono anch'essi soluzioni.

Abbiamo trovato in questa maniera una sola soluzione dell'equazione differenziale data. Un'eventuale seconda soluzione nel'intorno dell'origine potrebbe anche in linea di principio non esistere, visto che i coefficienti dell'equazione non sono funzioni continue. Tuttavia si mostra che una tale soluzione (linearmente indipendente dalla prima) esiste, ma non è analitica e quindi non può essere trovata con questi metodi.

Esercizio 8.7 * Stabilire se l'equazione differenziale

$$(1 - t^2)y'' - 2ty' + \lambda y = 0$$

ammette, al variare del parametro $\lambda \in \mathbb{R}$, soluzioni esprimibili come serie di potenze centrata nell'origine. Stabilire poi se esistono valori di λ per i quali l'equazione data ammette come soluzioni degli opportuni polinomi.
R. Posto $y(t) = \sum_{n=0}^{+\infty} a_n t^n$ e fissati $a_0, a_1 \in \mathbb{R}$ si deve avere $a_{n+2} = \frac{n^2-n-\lambda}{(n+2)(n+1)} a_n$ per ogni $n \in \mathbb{N}$. La serie risultante ha raggio di convergenza $R = 1$. Se $\lambda = n^2 - n$ per qualche $n \in \mathbb{N}$ allora esistono soluzioni polinomiali.

8.10 Un cenno allo studio qualitativo di equazioni differenziali

Scopo di questa sezione è dare un cenno a un metodo molto generale per studiare le proprietà delle soluzioni di equazioni differenziali: il metodo dell'*analisi qualitativa*. Come avete visto è assai raro riuscire a trovare esplicitamente le soluzioni di un'equazione differenziale. Abbiamo visto che perfino nel caso di equazioni differenziali del primo ordine a variabili separabili può essere impossibile calcolare esplicitamente le primitive necessarie oppure, anche quando le primitive si possono calcolare, scrivere esplicitamente y in funzione di t. Tuttavia in alcuni casi è possibile egualmente ottenere informazioni *qualitative* sul comportamento delle soluzioni.

Cominciamo con l'osservare che, data un'equazione differenziale del primo ordine in forma normale

$$y' = f(t, y)$$

con f soddisfacente le ipotesi del Teorema di Cauchy, il valore $f(t_0, y_0)$ coincide per costruzione con il valore della derivata y' della soluzione (unica!) che al tempo t_0 vale y_0. Possiamo quindi, esaminando i valori della f nel piano (t, y), avere un'idea di come si comportano le soluzioni dell'equazione differenziale data. Ci accontenteremo di mostrare come su due esempi.

Consideriamo l'equazione differenziale

$$y' = \frac{y^2 - 1}{e^y}.$$

Si tratta di una semplice equazione a variabili separabili, ma purtroppo la funzione
$$g(y) = \frac{e^y}{y^2 - 1}$$
non ammette una primitiva che si possa scrivere in termini di funzioni elementari. I metodi studiati finora quindi non ci permettono di scrivere esplicitamente la soluzione, ma il problema più grave è che, se anche indichiamo con $G(y)$ la primitiva di $g(y)$ e scriviamo la soluzione in forma implicita come

$$G(y) = t + c$$

non siamo in grado di dire nulla sul dominio, sull'immagine e sull'*invertibilità* di G, o almeno non abbiamo imparato alcun metodo per farlo. Per ricavare y come $G^{-1}(t+c)$ dovremmo essere in grado di sapere almeno qual'è il dominio della funzione inversa G^{-1}: senza questa informazione non possiamo neppure iniziare a studiare le soluzioni.

Vediamo ora però che ci è possibile determinare quasi tutte le principali caratteristiche *qualitative* delle soluzioni esaminando la forma del secondo membro dell'equazione differenziale considerata. Suggeriamo di verificare una per una le osservazioni che faremo nel grafico delle soluzioni che compare di seguito. Per prima cosa vi sono due integrali singolari: le funzioni

$$y_1(t) = 1 \quad \forall t \in \mathbb{R}, \qquad y_2(t) = -1 \quad \forall t \in \mathbb{R}.$$

I grafici di queste due funzioni dividono il piano (t, y) in tre regioni disgiunte. L'equazione differenziale che deve essere soddisfatta ci mostra che tutte le soluzioni che si trovano nelle regioni del piano caratterizzate dalle condizioni $y > 1$ oppure $y < -1$ *hanno derivata positiva*, quindi sono monotone crescenti. Invece le soluzioni che si trovano nella striscia $-1 < y < 1$ hanno derivata negativa, quindi sono monotone decrescenti.

Ci si potrebbe chiedere se una soluzione che a un certo tempo t si trova, diciamo, nella regione $y < -1$ possa "oltrepassare la retta $y = -1$ e passare nella regione $-1 < y < 1$. Questo però non è possibile perché allora il grafico di tale soluzione dovrebbe intersecare quello dell'altra soluzione $y_2(t) = -1$. Questo però sarebbe in contraddizione con l'unicità della soluzione al problema di Cauchy con dato iniziale nell'eventuale punto di intersezione. Lo stesso vale per soluzioni che stanno, a un certo tempo t, in una delle altre due regioni del piano sopra descritte: quindi ogni soluzione "vive in una sola di tali tre regioni. Questo mostra anche, *indipendentemente dalla verifica delle ipotesi del Teorema di esistenza in grande delle soluzioni*, che ogni soluzione che sta nella striscia $-1 < y < 1$ è definita per tutti i valori di t.

Calcoliamo ora la derivata seconda delle soluzioni. Si ha, derivando rispetto a t l'equazione differenziale di partenza,

$$\begin{aligned} y'' &= [(y^2-1)e^{-y}]' \\ &= y'e^{-y}(2y - y^2 + 1) \\ &= e^{-2y}(y^2-1)(2y-y^2+1). \end{aligned}$$

Quindi con un semplice studio del segno del secondo membro si vede che le soluzioni hanno concavità rivolta verso l'alto se

$$-1 < y < 1 - \sqrt{2} \quad \text{oppure} \quad 1 < y < 1 + \sqrt{2}$$

e rivolta verso il basso altrimenti.

Osserviamo ora che ogni soluzione è, come già detto, monotona (crescente o decrescente). Quindi esse ammettono limite agli estremi dell'insieme dei tempi su cui sono definite. Consideriamo dapprima le soluzioni y_3 che stanno nella striscia $-1 < y < 1$. Esse sono tutte monotone decrescenti. Quando $t \to +\infty$ allora il limite ℓ di ognuna di tali soluzioni esiste, è finito e appartiene all'intervallo $[-1, 1]$ per costruzione. Si vede subito che

$$\lim_{t\to+\infty} y_3'(t) = 0 :$$

se infatti così non fosse non sarebbe possibile che la funzione $y_3(t)$ avesse concavità rivolta verso l'alto quando t è abbastanza grande. Allora, passando al limite nell'equazione differenziale di partenza si ottiene

$$0 = (\ell^2 - 1)e^{-\ell}. \tag{8.39}$$

Quindi $\ell = \pm 1$, ma può aver luogo solo il caso $\ell = -1$ visto che si tratta di soluzioni decrescenti e che stanno dentro la striscia $-1 < y < 1$. Dunque

$$\lim_{t\to+\infty} y_3(t) = -1.$$

Analogamente si vede che

$$\lim_{t\to-\infty} y_3(t) = 1$$

Si può procedere analogamente e mostrare ora che, per le soluzioni y_4 che stanno nella regione $y > 1$ si ha

$$\lim_{t \to -\infty} y_4(t) = 1.$$

Non possiamo invece essere certi che le soluzioni esistano per tutti i tempi, in particolare che si possa far scorrere il tempo fino a $+\infty$. Questo però è vero; in effetti sul semipiano $y > 1$ la funzione $f(y) = (y^2 - 1)e^{-y}$ soddisfa certamente la condizione

$$0 < f(y) \leq C \quad \forall y > 1$$

per un'opportuna costante C, visto che f una funzione continua, positiva sull'insieme considerato e che tende a zero a $+\infty$. Siccome y' coincide con $f(y)$ se ne deduce che ogni soluzione sta sempre sotto una retta di coefficiente angolare C e quindi è definita per tutti i tempi.

Per vedere qual'è il limite a $+\infty$ notiamo solo che come prima, siccome le soluzioni sono monotone, detto ℓ tale limite (che esiste per la monotonia) dovrebbe valere la (8.39). Poiché la soluzione è crescente e sempre maggiore di uno tale equazione non può essere soddisfatta. Quindi

$$\lim_{t \to +\infty} y_4(t) = +\infty$$

(ve ne convincerete agevolmente mettendo formalmente $\ell = +\infty$ nella (8.39)). Si vede poi che non esistono asintoti obliqui per le soluzioni considerate, perché l'equazione differenziale che deve essere soddisfatta impone, visto che le soluzioni y_4 considerate tendono a $+\infty$ per $t \to +\infty$, che

$$\lim_{t \to +\infty} y_4'(t) = \lim_{t \to +\infty} (y_4^2(t) - 1))e^{-y_4(t)} = 0.$$

Si può infine procedere analogamente per le soluzioni y_5 che stanno nel semipiano $y < -1$. Con considerazioni analoghe a quelle già fatte si ottiene che

$$\lim_{t \to +\infty} y_5(t) = -1.$$

Resta da stabilire se le soluzioni possano essere definite per tutti i tempi. La risposta stavolta è *negativa*. In effetti si vede facilmente che, definendo come sopra $f(y) = (y^2 - 1)e^{-y}$ si ha

$$f(y) \geq C_\varepsilon y^2 \quad \forall y < -1 - \varepsilon$$

dove C_ε è un'opportuna costante (dipendente da ε). Conosciamo già l'equazione differenziale $y' = y^2$ e sappiamo che le soluzioni *non sono definite per tutti i tempi*. Lo stesso avrà quindi luogo per le soluzioni dell'equazione che stiamo considerando quando l'ordinata del dato iniziale è minore di -1: si dovrà quindi avere che, fissato il dato iniziale, esiste un t_0 tale che

$$\lim_{t \to t_0^+} y_5(t) = -\infty.$$

Concludiamo osservando che se $y(t)$ è una soluzione allora anche $y(t+c)$ lo è: questo è sempre vero quando si consideri un'equazione a variabili separabili *autonoma*, cioè con secondo membro indipendente da t. Tali soluzioni sono ottenute da quella da cui si parte semplicemente traslando il grafico di quest'ultima lungo l'asse delle t.

Siamo quindi in grado di disegnare qualitativamente il grafico qualitativo di tutte le soluzioni dell'equazione considerata come segue.

Il prossimo esempio riguarderà un'equazione differenziale del primo ordine a variabili separabili ma *non autonoma*, cioè con secondo membro dipendente esplicitamente dal tempo. L'equazione è la seguente:

$$y' = \frac{y}{y+1} t.$$

Tale equazione a variabili separabili si può integrare esplicitamente e si possono scrivere in forma implicita le soluzioni così :

$$g(y) := y + \ln |y| = \frac{t^2}{2} + c. \tag{8.40}$$

La funzione g non è invertibile tra il suo dominio e la sua immagine: provate a fare il semplice studio di funzione necessario per mostrarlo. Occorrerebbe restringere dominio e immagine e studiare i vari casi che si possono presentare, studiando poi dominio e immagine delle funzioni inverse che corrispondono a tali diversi casi. Sebbene ciò sia possibile, questa non è una via particolarmente agevole e conviene procedere come nell'esempio precedente.

Per prima cosa notiamo che le soluzioni sono tutte funzioni *pari*. Ciò si può vedere dalla (8.40), perche' se $y(t)$ è soluzione dell'equazione considerata in un intervallo del tipo $[0, A]$ la si può estendere all'intervallo $[-A, 0]$ ponendo, per ogni $t > 0$, $y(-t) := y(t)$: i due tratti di soluzione si *congiungono bene* per $t = 0$ perché l'equazione differenziale che deve essere soddisfatta impone che $y'(0) = 0$. Si potrebbe peraltro notare questo fatto anche semplicemente dalla forma dell'equazione differenziale che si sta studiando (Esercizio $*$ per voi).

Notiamo ora per prima cosa che vi è un integrale singolare dell'equazione data, la funzione $y_1(t) = 0$ per ogni t. Si nota anche subito che il secondo membro dell'equazione non è definito per $y = -1$, quindi in prima battuta dobbiamo aspettarci che non vi siano soluzioni che attraversino tale retta del piano (t, y) (a meno che qualche soluzione non si possa *prolungare* oltre tale retta, cosa in linea di principio possibile ma che non accade qui).

Notiamo poi subito che il segno di $yt/(y+1)$ (e quindi di y') si può studiare senza difficolta: in effetti il piano viene diviso in sei regioni, in tre delle quali il segno è positivo mentre nelle altre tre il segno è negativo. Possiamo, per la simmetria delle soluzioni, restringerci al caso $t > 0$ e quindi notare che in tale regione le soluzioni sono monotone crescenti se e solo se

$$t, y > 0 \quad \text{oppure } t > 0, y < -1.$$

Se $t > 0$ e $-1 < y < 0$ le corrispondenti soluzioni sono decrescenti.

Cominciamo a studiare le soluzioni y_2 che si trovano nella regione $t, y > 0$. Per far questo conviene calcolare la derivata seconda. Si ottiene:

$$\begin{aligned} y'' &= \frac{y}{y+1} + ty' \frac{y+1-y}{(y+1)^2} \\ &= \frac{1}{y+1}\left[y + \frac{y't}{y+1}\right] \\ &= \frac{1}{y+1}\left[y + \frac{yt^2}{(y+1)^2}\right] \end{aligned} \tag{8.41}$$

Si vede subito che, se $y > 0$, si ha sempre $y'' > 0$. Quindi le soluzioni che stanno in tale regione sono funzioni convesse, oltre che monotone crescenti come già sappiamo. Resta da vedere se esse sono o meno definite per ogni tempo t: ma si vede subito che

$$0 < f(t, y) = \frac{y}{y+1} t < A$$

per ogni $y > 0$ e per ogni $t \in [0, A]$. Questo vuole dire, come prima, che su ogni intervallo di tempo limitato la soluzione sta sotto una retta, e quindi in particolare è ben definita su ogni tale intervallo (non può cioè tendere a $+\infty$ ad un tempo t_0 finito). Quindi le soluzioni considerate sono definite per tutti i tempi. Poichè ogni soluzione è, come detto, crescente e convessa, dovrà necessariamente valere che

$$\lim_{t \to +\infty} y_2(t) = +\infty.$$

Inoltre le soluzioni date non ammettono asintoto obliquo per $t \to +\infty$, in quanto, dato che $y_2(t) \to +\infty$ in tale limite si deve avere anche, per l'equazione differenziale che si suppone soddisfatta,

$$\lim_{t \to +\infty} y_2'(t) = +\infty.$$

Nella regione $t > 0$, $y < -1$ il comportamento delle soluzioni è del tutto diverso. Le soluzioni sono crescenti e *convesse*, come si vede subito dall'espressione di y'' calcolata prima. Questo implica per costruzione (si veda il grafico più sotto) che esiste un tempo $t_0 > 0$ tale che

$$\lim_{t \to t_0^-} y_2(t) = -1.$$

Tali soluzioni non possono essere prolungate oltre perché l'equazione differenziale da cui siamo partiti ci dice che, quando una soluzione si avvicina alla retta $y = -1$, la sua derivata prima tende a $+\infty$ o a $-\infty$.

La situazione è poi del tutto simile per le soluzioni y_3 che stanno nella regione $-1 < y < 0$, visto che già sappiamo che esse sono decrescenti, mentre l'espressione di y'' ci dice che esse sono *concave*. Questo implica che, come prima, esiste un tempo $t_0 > 0$ tale che

$$\lim_{t \to t_0^-} y_3(t) = -1$$

e che le soluzioni non si possono estendere oltre, visto che la loro derivata prima tende, per $t \to t_0^-$, a $-\infty$. In conclusione il grafico qualitativo delle soluzioni è il seguente:

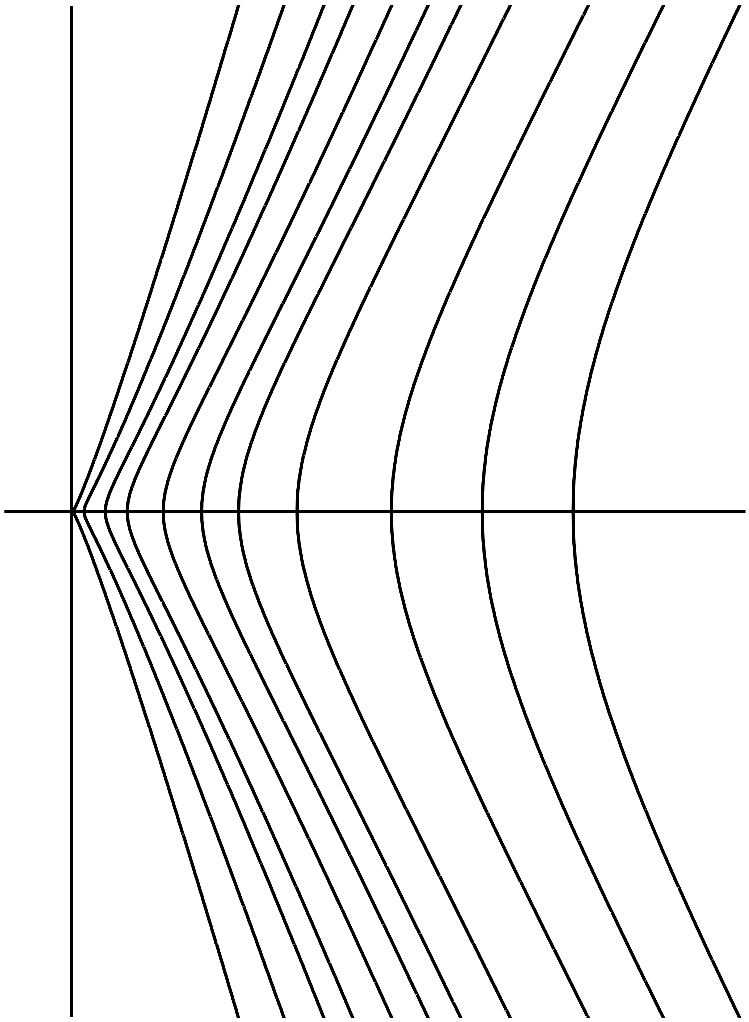


Figura 8.3:

Appendice A

Elementi di insiemistica e di calcolo combinatorio

A.1 Notazioni insiemistiche e quantificatori

Il concetto di *insieme* viene assunto come primitivo, da intendersi intuitivamente come *collezione di elementi*. Gli insiemi verranno solitamente denotati con lettere maiuscole A, B, X, Ω, ecc. mentre gli elementi degli insiemi con lettere minuscole a, b, x, ecc. $a \in A$ significa che a è *un elemento dell'insieme* A. $a \notin A$ significa che a non è un elemento dell'insieme A. Insiemi numerici rilevanti sono $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{Q}$, $\mathbb{R}$. Un insieme è in generale descritto enumerando esplicitamente, tra parentesi graffe, gli elementi che esso contiene, ad esempio $A_1 = \{1, 2, 3\}$ oppure, attraverso una proprietà caratterizzante gli elementi che lo compongono; ad esempio, l'insieme dei numeri naturali pari, $A_2 = \{n \in \mathbb{N} \,|\, n \text{ è pari}\}$. Il simbolo $A \subseteq B$ significa che tutti gli elementi di A sono anche elementi di B e si dice che A è un *sottoinsieme* di B o che A è incluso in B. Gli insiemi A_1 e A_2 introdotti prima sono entrambi sottoinsiemi di $\mathbb{N}$. Tra gli insiemi numerici valgono le inclusioni

$$\mathbb{N} \subseteq \mathbb{Z} \subseteq \mathbb{Q} \subseteq \mathbb{R}\,.$$

Un insieme speciale è l'*insieme vuoto* indicato con il simbolo $\emptyset$: esso è l'insieme privo di elementi ed è quindi incluso in qualunque insieme. Se A è un insieme con un numero finito di elementi, il simbolo $|A|$ indica il numero di tali elementi ed è detto la *cardinalità* di A.

Tra insiemi si possono fare una serie di operazioni che riportiamo brevemente qui di seguito:

- *Unione di insiemi* $A \cup B$ è costituito dagli elementi di A e dagli elementi di B.
- *Intersezione di insiemi* $A \cap B$ è costituito dagli elementi che stanno simultaneamente in A e in B.
- *Differenza di insiemi* $A \setminus B$ è costituito dagli elementi di A che non sono elementi di B.
- *Prodotto cartesiano di insiemi* $A \times B$ è costituito dalle coppie ordinate (a, b) dove $a \in A$ e $b \in B$.

Ricordiamo l'importante proprietà, di facile dimostrazione (esercizio)

$$A \cap (B \cup C) = (A \cap C) \cup (B \cap C)\,.$$

Introduciamo ora un concetto di importanza fondamentale in tutta la matematica, quello di funzione. Dati due insiemi A e B, una *funzione* (o *applicazione*) f da A in B, denotata con il simbolo

$$f : A \to B\,,$$

è una legge che associa ad ogni elemento di $a \in A$ uno ed un solo elemento $b \in B$: si usa la notazione $b = f(a)$ e b è detto immagine di a. A viene detto *dominio* e B *codominio* della funzione f. Il sottoinsieme di B costituito dalle immagini degli elementi del dominio viene detto *immagine* della funzione f e denotato $\mathrm{Im} f$. In simboli

$$\mathrm{Im} f = \{b \in B \mid \text{esiste } a \in A,\ f(a) = b\}\,.$$

Introduciamo ora alcune importanti proprietà delle quali può godere una funzione. Una funzione $f : A \to B$ si dice *iniettiva* se elementi distinti nel dominio hanno immagini distinte nel codominio, cioè se:

$$a_1 \neq a_2 \;\Rightarrow\; f(a_1) \neq f(a_2)\,.$$

Una funzione $f : A \to B$ si dice *surgettiva* se la sua immagine coincide con il codominio, cioè se

$$\text{per ogni } b \in B,\ \ \text{esiste } a \in A,\ \text{tale che } f(a) = b\,.$$

Se una funzione $f : A \to B$ è iniettiva e surgettiva, essa determina una corrispondenza biunivoca tra i due insiemi A e B: ad ogni elemento di A corrisponde uno ed un solo elemento di B e viceversa. In tal caso si può definire la cosidetta *funzione inversa* $f^{-1} : B \to A$ che associa ad ogni elemento di B, il corrispondente elemento di A da cui proviene, cioè:

$$f^{-1}(b) = a \;\Leftrightarrow\; f(a) = b.$$

(attenzione: f^{-1} è un *simbolo* per l'inversa di una funzione e non ha niente a che vedere con l'elevare f alla potenza -1!) Una funzione iniettiva e surgettiva sarà anche detta, per i motivi sopra, funzione *invertibile* come sinonimo di corrispondenza biunivoca. Si noti che se A e B sono finiti ed in corrispondenza biunivoca, allora necessariamente $|A| = |B|$.

Commentiamo ancora sull'importante concetto di invertibilità che abbiamo sopra introdotto. Porsi il problema di invertire una funzione $f : A \to B$, vuole dire porsi i seguenti due problemi: 1) se b appartiene al codominio B di f, è o non è possibile trovare un $a \in A$ tale che $f(a) = b$ (surgettività)? In caso affermativo, tale a è unico (iniettività)? Le risposte dipendono ovviamente dalla funzione, dove per funzione intendiamo non solo la "regola che dato a ci fornisce $f(a)$, ma anche il dominio e il codominio di f. Se la risposta alle precedenti domande è sì , allora non solo potremo associare ad ogni a del dominio un elemento $b = f(a)$ del codominio, come da definizione di funzione; ma potremo anche tornare indietro, nel senso che ad ogni b del codominio si potrà anche associare uno e un solo a del dominio. È immediato dalla definizione che

$$\begin{aligned} f(f^{-1}(b)) &= b \qquad \forall b \in B\,, \\ f^{-1}(f(a)) &= a \qquad \forall a \in A\,. \end{aligned}$$

e che inoltre tali proprietà caratterizzano la funzione inversa, una volta accertatisi che essa esista.

Le funzioni possono essere composte tra di loro.

Definizione A.1 Siano $f : A \to B$ e $g : C \to D$ due funzioni. Si supponga che $f(A) \subset C$. Allora la funzione

$$x \mapsto g(f(x))$$

definita su A si dice *funzione composta* di g ed f (nell'ordine) e si indica con $[g \circ f](x) = g(f(x))$.

Osserviamo per concludere che $g : C \to D$ è la funzione inversa di $f : A \to B$ se e soltanto se $D = A$ (cioè il codominio di g coincide col dominio di f), $B = C$ (cioè il codominio di f coincide col dominio di g), cosicché sono ben definite $f \circ g : C \to B = C$ e $g \circ f : A \to D = A$ e vale

$$\begin{array}{ll} [f \circ g](x) = x & \forall x \in C \\ [g \circ f](x) = x & \forall x \in A. \end{array}$$

Ricordiamo infine alcuni importanti simboli matematici con accanto il loro significato:

$$\begin{array}{rcl@{\qquad\qquad}rcl} \exists & : & \text{esiste}\,, & \exists! & : & \text{esiste uno ed un solo}\,, \\ \nexists & : & \text{non esiste}\,, & \forall & : & \text{per ogni}\,, \\ \Rightarrow & : & \text{implica}\,, & \Leftrightarrow & : & \text{se e solo se}\,. \end{array}$$

A.2 Il Principio di Induzione

I numeri naturali, pur essendo così antichi, hanno una ricchissima teoria che ancora oggi mostra importanti sviluppi. Ne sfioriamo alcuni interessanti aspetti in questa sezione, e nelle rimanenti di questa Appendice.

Iniziamo con una proprietà di cui gode l'insieme dei numeri naturali che va sotto il nome di principio di induzione. Supponiamo di avere una proposizione logica (cioè una proprietà che può essere o vera o falsa) che riguarda i numeri naturali. La indicheremo con il simbolo $P(n)$ ad enfatizzare il fatto che a seconda del numero n che consideriamo essa potrebbe essere o vera o falsa. Per chiarire facciamo un paio di esempi:

Esempio 149 Sia $P(n)$: n è un multiplo di 3. Allora $P(n)$ sarà vera se $n = 3, 6, 9, 12, \ldots$, falsa in tutti gli altri casi.

Esempio 150 Sia $P(n)$: $n^2 > n + 7$. In tal caso, risolvendo esplicitamente la diseguaglianza di secondo grado si vede che

$$n^2 > n + 7 \;\Leftrightarrow n > 4\,.$$

Quindi $P(n)$ è vera esattamente per $n > 4$ e falsa per $n = 1, 2, 3, 4$.

Principio di induzione: *Si consideri una proprietà $P(n)$ e supponiamo che*

(i) $P(1)$ *è vera.*

(ii) Dato un qualunque naturale n, se $P(n)$ è vera, allora anche $P(n+1)$ risulta vera.

Allora, $P(n)$ è vera per ogni n.

Pensandoci bene il principio di induzione è molto intuitivo. In effetti sia $P(n)$ una proprietà per la quale (i) e (ii) valgono. Allora $P(1)$ è vera. D'altra parte la (ii) applicata con $n = 1$ dice che se $P(1)$ è vera, allora è anche vera $P(2)$. Quindi $P(2)$ è vera. Applicando ora la (ii) con $n = 2$ si ha che se $P(2)$ è vera, allora è anche vera $P(3)$. Quindi $P(3)$ è vera. Continuando ad applicare la (ii) con $n = 3, 4, \cdots$ si otterrà che $P(4)$ è vera, $P(5)$ è vera e così via. Quindi $P(n)$ è vera per tutti i naturali n. Si potrebbe pensare che questa sia addirittura una dimostrazione del principio di induzione. In effetti non è così , quell'apparentemente innocuo 'e così via' con il quale passavamo da $P(4)$, $P(5)$ a $P(n)$ per ogni n non è formalizzabile. Il principio di induzione è un postulato e non può essere dimostrato. E' intuitivo come mostra il ragionamento sopra, come lo sono in genere i postulati.

Il principio di induzione è uno strumento potente, utilissimo nel dimostrare proposizioni che riguardano i numeri naturali. Facciamo un esempio

Esempio 151 Sia $n \in \mathbb{N}$. Allora vale la formula

$$1 + 2 + 3 + \cdots + n = \frac{n(n+1)}{2}\,. \tag{A.1}$$

Dimostriamolo per induzione. Sia $P(n)$ l'eguaglianza sopra.

$$P(1) : 1 = \frac{1(1+1)}{2} \text{ vera}\,.$$

Fissato n, supponiamo che $P(n)$ sia vera e dimostriamo che $P(n+1)$ è vera. Si ha

$$\begin{aligned} 1 + 2 + 3 + \cdots + n + (n+1) &= (1 + 2 + 3 + \cdots n) + (n+1) \\ &= \tfrac{n(n+1)}{2} + (n+1) \\ &= \tfrac{n(n+1) + 2(n+1)}{2} \\ &= \tfrac{(n+1)(n+2)}{2} \end{aligned}$$

(la prima eguaglianza è ovvia, la seconda usa l'ipotesi che $P(n)$ sia vera cioè (A.1), la terza e la quarta sono passaggi algebrici). Guardando il primo e l'ultimo membro della catena di eguaglianze sopra si vede che abbiamo proprio ottenuto la formula (A.1) per $n + 1$.

Esercizio A.1 Dimostrare, per induzione, che per ogni $n \in \mathbb{N}$ si ha

$$\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}\,. \tag{A.2}$$

Esercizio A.2 Sia $q \neq 1$. Dimostrare, per induzione, che per ogni $n \in \mathbb{N}$ si ha

$$\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q}\,. \tag{A.3}$$

A.3 Elementi di combinatoria

Supponiamo di avere un insieme (un'urna più concretamente) A contenente n elementi distinti che per semplicità possiamo pure pensare siano identificati dai primi n numeri naturali:

$$A = \{1, 2, \cdots, n\}.$$

Estraiamo un numero da A, annotiamo da una parte il risultato e quindi lo rimettiamo in A; poi ne estraiamo un secondo, annotiamo accanto al primo il risultato e lo rimettiamo in A. Andiamo avanti compiendo k di queste estrazioni. Il risultato ottenuto sarà una k-upla ordinata di numeri $(i_1, i_2, \ldots, i_k)$ dove ogni i_j sta in A. L'insieme di tali k-uple ordinate è appunto il prodotto cartesiano $A \times A \times \cdots \times A = A^k$; per questo motivo i suoi elementi sono anche detti *disposizioni di k elementi di A con ripetizione.*

Cambiamo ora leggermente esperimento. Estraiamo un elemento da A e mettiamolo da una parte, poi ne estraiamo un secondo e lo mettiamo accanto al primo estratto e così via fino ad averne estratti un numero k. I k numeri estratti formano ancora una k-upla ordinata (l'ordine è naturalmente quello di estrazione) di elementi distinti di A, detta una *disposizione di k elementi di A (senza ripetizione)*. L'insieme di tali disposizioni si indica $D^{n,k}$ e forma un sottoinsieme del prodotto cartesiano A^k. Le k-uple che stanno in $D^{n,k}$ sono quelle nelle quali non ci sono ripetizioni. Vediamo un semplice esempio per $n = 3$ ($A = \{1, 2, 3\}$) e $k = 2$:

$$A^2 = \{(11), (12), (13), (21), (22), (23), (31), (32), (33)\}\,;$$

$$D^{3,2} = \{(12), (13), (21), (23), (31), (32)\}\,.$$

A^k contiene n^k elementi distinti. Quanti elementi distinti contiene $D^{n,k}$? Ecco la risposta:

Proposizione A.2

$$|D^{n,k}| = n(n-1)\cdots(n-k+1) \quad k = 1, \dots, n\,. \tag{A.4}$$

Dimostrazione Poichè A contiene n elementi, la prima estrazione si può fare in n modi distinti. Fatta la prima estrazione, vi sono rimasti in A soltanto $n-1$ elementi e quindi la seconda estrazione si può fare in $(n-1)$ modi distinti e, così via, la r-esima estrazione si potrà fare in $n-r+1$ modi distinti. Per contare le disposizione di lunghezza k basta ora moltiplicare i modi in cui si può fare la prima estrazione, per i modi in cui si può fare la seconda, e così via, fino all'ultima e si ottiene così la formula (A.4). In realtà questa non è una dimostrazione completamente rigorosa, l'uso di locuzioni del tipo 'e così via' suggeriscono infatti che ci sia dietro un procedimento di induzione mascherato. In effetti si può rendere rigorosa nel modo seguente. Pensiamo che n sia fissato e dimostriamo (A.4) per induzione su k. Per $k = 1$ si ottiene

$$|D^{n,1}| = n$$

che è vero. Supponiamo ora che (A.4) sia vera per un certo $k < n$ e dimostriamola per $k+1$. Per fare $k+1$ estrazioni da A, devo prima farne k e questo posso farlo in $|D^{n,k}|$ modi distinti, poi mi rimane da fare la $k+1$ estrazione da un insieme che è ridotto ad avere $n-k$ elementi: quest'ultima estrazione posso quindi farla in $n-k$ modi distinti. Si ha quindi

$$|D^{n,k+1}| = |D^{n,k}| \cdot (n-k) = n(n-1)\cdots(n-k+1)(n-k)$$

che è esattamente la (A.4) con $k+1$ al posto di k. ■

Osservazione: Vale la pena di notare come il principio di induzione serva per dimostrare la validità di una certa formula dipendente da naturali come la (A.4), ma non aiuti a trovarle queste formule! Ricordiamo che noi siamo arrivati alla (A.4) per altra via con un ragionamento che non era del tutto rigoroso. Questa situazione è abbastanza tipica nel mondo dei numeri naturali: si 'indovina' una formula in qualche modo, poi la si dimostra rigorosamente per induzione.

Torniamo alle disposizioni. Abbiamo definito $D^{n,k}$ con $k > 0$. E' utile amche poter parlare di 0 estrazioni e considerare quindi anche $D^{n,0}$: esso, per convenzione avrà un solo elemento (intuitivamente c'è solo un modo di fare 0 estrazioni). Si noti un altro caso particolare della (A.4), quando $k = n$. Si ottiene

$$|D^{n,n}| = n(n-1)\cdots 3 \cdot 2 \cdot 1\,.$$

A destra vi sono tutti i numeri naturali fino ad n moltiplicati tra di loro. Si usa un simbolo speciale per questa espressione

$$n! = 1 \cdot 2 \cdot 3 \cdots (n-1) \cdot n$$

e tale numero si chiama *n-fattoriale.* Per convenzione poniamo anche $0! = 1$. Gli elementi di $D^{n,n}$ sono le disposizioni di lunghezza n degli n elementi di A senza ripetizione: in altre parole sono i modi distinti in cui posso mettere in fila gli n elementi di A: sono anche dette *permutazioni* di A. Un'insieme di n elementi ha dunque $n!$ permutazioni distinte.

Osservazione: Una permutazione può essere pensata come una funzione invertibile (cioè una corrispondenza biunivoca) da A in A nel modo seguente. Sia $A = \{1, 2, \cdots, n\}$ e consideriamo una permutazione di A data dalla n-upla ordinata $(a_1, a_2, \ldots, a_n)$; si può ad essa associare la funzione $f : A \to A$ data da $f(i) = a_i$ al variare di $i = 1, 2, \ldots n$. Essendo una permutazione, tutti gli a_i sono tra loro distinti e questo implica che f è effettivamente iniettiva e surgettiva, cioè invertibile . Inoltre è facile rendersi conto che ogni funzione invertibile da A in A può essere costruita come nel modo precedente a partire da una permutazione. Vi sono dunque $n!$ funzioni invertibili distinte da un insieme di n elementi in se stesso.

Se quando operiamo le k estrazioni da un insieme A, ci 'dimentichiamo' dell'ordine in cui le varie estrazioni sono avvenute, quello che otteniamo alla fine non è più una k-upla ordinata, bensì semplicemente un sottoinsieme di A costituito da k elementi distinti. Tali estrazioni non ordinate, a differenza delle prime, vengono dette *combinazioni di k elementi di A, senza ripetizione.* L'insieme di esse viene indicato con $C^{n,k}$. Per illustrare questo nuovo concetto presentiamo un semplice esempio. Nel caso $n = 3$ e $k = 2$ si ha:

$$D^{3,2} = \{(12), (13), (21), (23), (31), (32)\},$$

$$C^{3,2} = \{\{12\}, \{13\}, \{23\}\}.$$

Calcoliamo ora la cardinalità di $C^{n,k}$ che sarà indicata con il simbolo

$$\binom{n}{k}$$

detto, per motivi che saranno chiari in seguito, *coefficiente binomiale.*

Proposizione A.3

$$\binom{n}{k} = \frac{|D^{n,k}|}{k!} = \frac{n(n-1)\cdots(n-k+1)}{k!} \qquad k = 1, \ldots, n. \tag{A.5}$$

Dimostrazione In virtù di (A.4) si tratta solo di dimostrare la prima eguaglianza. Si noti che ogni disposizione determina una combinazione: basta considerare il sottoinsieme dei k elementi estratti dimenticando l'ordine di estrazione. D'altra parte, dovrebbe essere chiaro che due disposizioni determinano la stessa combinazione se e soltanto se le due k-uple ordinate di elementi estratti differiscono per una permutazione. Poichè il numero delle permutazioni di un insieme di k elementi è dato da $k!$, si ha che $k!$ diverse disposizioni determinano la stessa combinazione. Questo dimostra il risultato. ■

Come nel caso delle disposizioni parliamo anche di combinazioni di 0 elementi ponendo, per convenzione,

$$\binom{n}{0} = 1\,.$$

Si noti che, in questo modo, la prima eguaglianza in (A.5) continua a valere anche per $k = 0$.

Vale il seguente importante fatto

Proposizione A.4

$$\binom{n}{k} = \binom{n}{n-k} \qquad k = 0, 1, \ldots, n\,.$$

Dimostrazione Scelto un sottoinsieme di k elementi da un insieme di n, si determina automaticamente un sottoinsieme di $n-k$ elementi, gli esclusi dal sottoinsieme precedente. Questo ragionamento mostra come $C^{n,k}$ e $C^{n,n-k}$ siano in corrispondenza biunivoca: le loro cardinalità devono quindi essere uguali. ■

Esercizio A.3 Dimostrare che

$$\binom{n+1}{k} = \binom{n}{k} + \binom{n}{k-1} \tag{A.6}$$

dove $n \in \mathbb{N}$ e $1 \le k \le n$. Si noti poi che la relazione continua a valere per $k = 0$ interpretando $\binom{n}{-1} = 0$.

A.4 Il binomio di Newton

Arriviamo ora ad un'importante applicazione dei concetti di combinatoria fin qui introdotti. Siano a e b due numeri reali e sia n un numero naturale. Vorremmo calcolare $(a+b)^n$ o meglio svilupparlo nella somma di tutti i monomi che si ottengono moltiplicando $(a+b)$ per se stesso n volte. Per $n = 2, 3$ si hanno i prodotti notevoli che dovreste conoscere:

$$(a+b)^2 = a^2 + 2ab + b^2 \,, \qquad (a+b)^3 = a^3 + 3a^2b + 3ab^2 + b^3 \,.$$

Quale sarà l'espressione generale per $(a+b)^n$? Ecco la risposta:

Proposizione A.5

$$(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} \,. \tag{A.7}$$

Dimostrazione E' chiaro che nello sviluppo di $(a+b)^n$ compaiono soltanto monomi in a e b di grado n, cioè:

$$a^n,\ a^{n-1}b,\ a^{n-2}b^2,\ \cdots,\ ab^{n-1},\ b^n \,.$$

Dobbiamo determinare con quali coefficienti i vari monomi compaiono. Quando sviluppiamo il prodotto $(a+b)^n$, vengono fuori 2^n monomi dei tipi sopra; le 2^n scelte corrispondono a scegliere a oppure b negli n fattori $(a+b)$. Si noti che si ottiene $a^k b^{n-k}$ se si è scelto k volte a (e conseguentemente $n-k$ volte b). I modi distinti in cui posso scegliere k volte a corrispondono ai modi distinti in cui posso scegliere un sottoinsieme di k elementi da un insieme di n (le posizioni in cui scelgo a), cioè alle combinazioni di k elementi di un insieme di n. Quindi il coefficiente di $a^k b^{n-k}$ è $\binom{n}{k}$. Si noti che tutto torna anche nel caso $k = 0$: c'è un solo modo in cui posso scegliere 0 volte a e per convenzione $\binom{n}{0} = 1$.
■

Lo sviluppo presentato dalla (A.7) è noto come il *binomio di Newton*. Prendendo $a = b = 1$ nel binomio di Newton, si ottiene la formula

$$2^n = \sum_{k=0}^{n} \binom{n}{k} \,. \tag{A.8}$$

Che significa? A destra sto sommando tutti i coefficienti binomiali $\binom{n}{k}$. Poichè $\binom{n}{k}$ è il numero di sottoinsiemi di k elementi di un insieme di n, si ha che la somma mi da il numero complessivo di sottoinsiemi di un insieme di n elementi. La formula sopra mi dice che questo numero è semplicemente 2^n.

Il binomio di Newton ha moltissime applicazioni. Una è la seguente

Proposizione A.6 *Sia $x \geq 0$ e sia $n \in \mathbb{N}$. Allora si ha*

$$(1+x)^n \geq 1 + nx\,.$$

Dimostrazione Segue da (A.7) che

$$(1+x)^n = \sum_{k=0}^{n} \binom{n}{k} x^k = 1 + nx + \binom{n}{2} x^2 + \cdots\,.$$

Poichè tutti gli addendi della formula sopra sono non negativi, si ha, in particolare, che $(1+x)^n \geq 1 + nx$ come volevamo dimostrare. ■

Esercizio A.4 Dimostrare la Proposizione A.5 con il metodo di induzione.

Esercizio A.5 Dimostrare la Proposizione A.6 con il metodo di induzione.

Esercizio A.6 *Utilizzando la relazione (A.6) si dimostri, per induzione, direttamente l'eguaglianza del primo e terzo termine in (A.5).

Esercizio A.7 *Si noti che dalla formula

$$k = \frac{(k+1)^2 - k^2 - 1}{2}$$

si ottiene

$$\sum\nolimits_{k=1}^{n} k = \frac{1}{2}\left[\sum_{k=1}^{n}((k+1)^2 - k^2) - \sum_{k=1}^{n} 1\right] = \frac{(n+1)^2 - 1 - n}{2} = \frac{n(n+1)}{2}\,.$$

Abbiamo così ottenuto, direttamente, la formula (A.1). Dopo aver studiato e compreso i passaggi precedenti, utilizzando lo sviluppo del binomio di Newton $(k+1)^3$, si dimostri, direttamente, (A.2).

Esercizio A.8 * Determinare formule esplicite, in termini di $n \in \mathbb{N}$, per le sommatorie

$$\sum_{k=1}^{n} k^3\,, \qquad \sum_{k=1}^{n} k^4\,.$$