ELIO CABIB

Lezioni di Analisi 1

Alla memoria del mio babbo Renzo
che ha illuminato i sentieri della Matematica
ai miei primi passi incerti

ELIO CABIB
cabib@uniud.it
professore di Analisi Matematica
Università di Udine

Lezioni di Analisi 1

Indice

Introduzione

Questo libro è il risultato di parecchi anni di esperienza didattica in Analisi Matematica e contiene gli argomenti classici del corso di Analisi 1, così come i programmi tradizionali dei corsi di Laurea in Matematica, Fisica e Ingegneria prevedono. Credo però che ci sia qualcosa di non tradizionale nello stile. Innanzitutto trattandosi, come tanti altri, di un testo elettronico a libera diffusione (con qualche collegamento extratestuale che aiuta a renderlo un po' più vivace), è svincolato dalle pretese, sia grafiche, sia contrattuali, degli editori. Poi c'è la tentazione del taglia e cuci sempre in agguato, che invita a migliorarlo (o ... a peggiorarlo ☹) quando si vuole. L'esposizione della materia non è molto omogenea e coerente, a volte è rigorosa e formale, altre, specialmente nelle prime battute di un capitolo nuovo, sembra procedere in modo intuitivo tra domande e ingenui tentativi di dare risposte, tanto da assomigliare più alla lezione orale. Ma ha un pregio non irrilevante: non esistono "note a margine", "rubriche di approfondimento" o altre sciocchezze e smanie didattiche di moda oggi, come il riassuntino memorizzatore delle regolette appena spiegate, a mio avviso un ostacolo alla libertà di apprendimento da parte dello studente. È vero che non bisogna mettere tutto sullo stesso piano, che ci sono risultati centrali, fondamentali, e risultati minori, ma non sta al docente doverli distinguere sottolineando le prime e scrivendo "in piccolo" le seconde, per due motivi: primo, perché le cose scritte in piccolo non le legge nessuno; secondo, se si toglie l'opportunità allo studente di usare le sue capacità critiche e la libertà di personalizzare la materia, stabilendo da solo cosa ritiene più importante e cosa no, non gli resterà altro da fare se non cercare di imparare (per l'esame) invece di capire. E non è con la memorizzazione, passiva per sua natura, che si impara la Matematica. Lo sforzo che la riflessione autonoma richiede per comprendere le idee e le motivazioni che stanno dietro a concetti ed esempi, dietro a definizioni e teoremi, tiene vivo lo spirito critico, esercita la capacità di confrontare situazioni, di stabilire collegamenti e di scoprire analogie, costringe a porsi domande e quindi stimola la costruzione di esempi e controesempi, ed è infine anche essenziale per l'acquisizione del linguaggio corretto. La Matematica, ma non solo, si apprende con le domande e i tentativi di trovare risposte, una palestra in cui le intuizioni dello spirito creativo trovano una sistemazione logica. Il modo corretto di studiare la Matematica è dunque leggere l'enunciato di un teorema, chiudere il libro e tenerlo chiuso finché non si è costruita passo passo la dimostrazione con le proprie forze.

L'Analisi Matematica è la disciplina dei passaggi al limite: oltre al concetto di limite in sé, per successioni, funzioni e serie, anche gli estremi inferiore e superiore di una funzione o di un insieme di numeri, su cui si basa anche la definizione di integrale, sono (quasi) dei passaggi al limite. La nozione di continuità per le funzioni è strettamente legata a quella di limite e la derivata è per definizione un limite, se vogliamo limitarci ai soli argomenti di questo corso. Per questo un peso notevole viene dato alla completezza dell'insieme dei numeri reali, non solo quando tale concetto viene introdotto e illustrato, ma per l'uso che se ne fa durante tutto il corso. Si pensi alla funzione $\operatorname{sen} x$ che se fosse definita sull'insieme dei soli numeri razionali non avrebbe né massimo né minimo. Ben più in generale, possiamo senz'altro affermare che tutti i teoremi di esistenza in Analisi, da quella del limite nelle varie situazioni all'esistenza di zeri e punti fissi per le funzioni continue, all'esistenza di soluzioni delle equazioni differenziali, si basano sulla completezza. L'importanza di questo strumento, o meglio di questa ipotesi sui reali, ci dà la possibilità di costruire teorie utili e sufficientemente generali.

Vediamo adesso i contenuti divisi per capitolo.

Cap. 1 - È spendibile in ogni settore della Matematica perché riguarda le nozioni fondamentali della logica, fissando i termini del linguaggio matematico e le regole del ragionamento, e dell'insiemistica con le operazioni tra gli insiemi, le relazioni e il concetto generale di funzione.

Cap. 2 - Si illustrano gli insiemi numerici dal punto di vista algebrico e dell'ordinamento partendo da una presentazione assiomatica dell'insieme $\mathbf{R}$ dei numeri reali che si distingue per l'assioma di completezza. Si passa poi a trattare i sottoinsiemi numerici dei naturali, degli interi e dei razionali. La teoria degli insiemi introdotta nel Cap. 1 prosegue in questo intrecciandosi con il calcolo combinatorio e la questione della cardinalità. Il valore assoluto come funzione reale ci permette di introdurre una distanza in $\mathbf{R}$ e ci permette di cominciare a prendere familiarità con la nozione di spazio metrico.

Cap. 3 - È dedicato alla costruzione delle *funzioni elementari* che sono quelle funzioni reali di variabile reale (estendibili anche al campo complesso) che ammettono una definizione in *forma chiusa*, cioè data mediante una formula esplicita, riconducibile alle operazioni elementari. A partire dalle potenze ad esponente intero, che sono semplicemente dei prodotti, con la somma già presente in $\mathbf{R}$ si passa alle funzioni polinomiali e poi alle funzioni algebriche razionali. Con una prima versione del teorema degli zeri si riescono a costruire anche i radicali e quindi le funzioni algebriche irrazionali. Infine con la completezza di $\mathbf{R}$ si passa alle potenze con esponente reale, alle funzioni esponenziali e logaritmiche e alle funzioni iperboliche ad esse strettamente legate. Per via geometrica si costruiscono anche le funzioni trigonometriche. Le tre proprietà fondamentali che ricorrono per tutta la durata del corso: *limitatezza*, *monotonia* e *convessità*, vengono qui introdotte per testarle subito con le funzioni elementari. Le funzioni trigonometriche ci offrono anche l'opportunità di trattare in generale il concetto di funzione periodica.

Cap. 4 - Vengono definiti i numeri complessi con le loro proprietà algebriche e geometriche. Dai polinomi complessi si passa ad analizzare le proprietà di alcune funzioni elementari di variabile complessa per estensione di quelle reali. Tipica di questa teoria è l'esistenza di funzioni multivoche come la radice n-esima e l'*argomento*.

Cap. 5 - Viene trattato il concetto di limite per le successioni di numeri reali con una certa attenzione, per quanto riguarda il linguaggio usato, alla possibilità di generalizzarlo al caso degli spazi metrici. Se ne illustrano le proprietà generali, quelle legate all'ordinamento e quelle di natura algebrica. Si studiano i limiti delle funzioni elementari, si introduce il numero di Nepero e le relative funzioni esponenziale e logaritmica con la loro estensione al campo complesso. In quest'ambito si scopre un legame sorprendente tra le funzioni trigonometriche e la funzione esponenziale, tanto che, abbandonando la via geometrica, le prime possono essere definite dalla seconda per via algebrica. Di una certa importanza per gli sviluppi successivi sono le successioni di medie.

Cap. 6 - Questo capitolo è dedicato alle serie numeriche che sono in realtà successioni di somme che difficilmente possono essere riscritte in termini di un'unica espressione in forma chiusa. La somma di una serie è nient'altro che il limite di una successione di somme, calcolabile esplicitamente solo in pochissimi casi. Vengono allora sviluppati dei criteri, cioè delle condizioni sufficienti, che ci permettono di stabilire se la somma di una serie esiste, finita o meno, o se non esiste proprio. Una categoria importante è costituita dalle serie di potenze reali e complesse. Esse allargano l'orizzonte sulle funzioni elementari dal momento che definiscono la classe più ampia delle funzioni analitiche di cui quelle elementari sono casi particolari.

Cap. 7 - Qui viene sviluppata e generalizzata la nozione di distanza e di spazio metrico già introdotta in $\mathbf{R}$. Alla distanza sono legati i concetti di intorno, di insieme aperto, di insieme chiuso e altre nozioni metriche come la completezza e la compattezza. Si analizza il ruolo che hanno le successioni e i loro punti limite nel descrivere queste proprietà, in particolare trattiamo anche il massimo e il minimo limite in $\mathbf{R}$.

Cap. 8 - Si riprende la teoria delle funzioni analizzando nella prima parte il concetto di limite che viene presentato, insieme alle sue proprietà, cercando di preservare l'analogia con la teoria delle successioni. Si applica al calcolo dei limiti la teoria del confronto tra infinitesimi e tra infiniti nella quale si sfruttano sistematicamente gli sviluppi in serie di potenze del Cap. 6. Nella seconda parte vengono trattate le funzioni continue con le loro proprietà direttamente discendenti da quelle dei limiti per poi illustrare quelle legate alla scelta del dominio. Le funzioni continue godono infatti di certe proprietà speciali legate al dominio: trasformano insiemi connessi in insiemi connessi che in $\mathbf{R}$ sono solo gli intervalli. In quest'ordine di idee rientrano alcune conseguenze importanti come il teorema di esistenza degli zeri e la continuità della funzione inversa. Altro filone di risultati riguarda il caso in cui il dominio è un insieme compatto, situazione in cui si può dedurre l'esistenza del minimo e del massimo e, ancora una volta, la continuità della funzione inversa. Si passa poi a studiare le proprietà delle funzioni uniformemente continue, il cui "grado" di continuità è, in un senso che preciseremo, omogeneo su tutto il dominio.

Cap. 9 - Il concetto di derivata nasce dall'esigenza geometrica di trovare l'equazione della retta tangente al grafico di una funzione in un dato punto e dall'esigenza, spesso di origine fisica, di misurare il tasso di variazione di una grandezza rispetto ad un'altra di cui è funzione, come nel caso della velocità che è una misura di quanto varia la posizione di un punto nel tempo. Dopo aver illustrato le regole di derivazione e alcuni teoremi fondamentali di carattere generale, si passa alle applicazioni che riguardano la determinazione dei punti di massimo e di minimo relativo, la monotonia e la convessità, sia nel senso puntuale, sia nel senso globale intese come proprietà legate alla scelta di un intervallo come dominio della funzione. In termini delle derivate si calcolano i coefficienti dell'approssimazione polinomiale di una funzione, utilissima nel calcolo dei limiti. Le serie di Taylor per le funzioni analitiche vengono riprese a completamento della teoria delle serie di potenze.

Cap. 10 - Dopo aver introdotto l'idea di misura, l'area per una regione piana, si passa a definire l'integrale di una funzione costante a tratti e quindi l'integrale di Riemann per una funzione limitata e nulla al di fuori di un intervallo. Se ne illustrano le proprietà principali e si individuano alcune classi notevoli di funzioni integrabili. Si passa poi alla nozione di funzione integrale e al calcolo di un integrale per mezzo delle primitive. Si estende poi l'integrale alla classe più ampia delle funzioni non limitate o definite su un intervallo non limitato. Si illustrano infine i più comuni metodi di integrazione, per parti e per sostituzione, e si offre una casistica di funzioni elementari cui sono applicabili procedimenti standard.

Con piacere esprimo la mia riconoscenza al prof. Luciano Battaia per il suo aiuto prezioso e i suoi consigli sull'uso del LaTeX, al dott. Stephane Matiz PhD e a Gerrit Oomens, brillante studente di Matematica presso il Korteweg-de Vries Institute for Mathematics dell'Università di Amsterdam, che con pazienza e gentilezza mi hanno aiutato nell'uso del Mathematica. Con gratitudine chiedo ai miei studenti, e a coloro che hanno voglia di leggerlo, di contribuire alla buona riuscita del testo con suggerimenti e consigli e di segnalarmi errori, carenze e stonature che di certo non mancano.

Capitolo 1

Logica e insiemistica

1.1 Proposizioni e predicati

Accettiamo come primitiva la nozione di *proposizione*. Ogni proposizione può essere solo *vera* oppure solo *falsa*, è escluso che possa essere sia vera che falsa o assumere un terzo valore ancora diverso. Due proposizioni si dicono *equivalenti* se sono entrambe vere o entrambe false. Ad esempio le due proposizioni $\pi > 3$ e $2 < 3$ sono equivalenti in quanto entrambe vere. Più avanti rivedremo questa definizione da un altro punto di vista.

Nella realtà quotidiana ci troviamo spesso di fronte a meccanismi a due valori: un motore può essere solo acceso o solo spento, un circuito elettrico può essere solo chiuso o solo aperto, situazioni che si possono associare ai due valori 0 e 1 di una variabile numerica. Ad esempio, si può stabilire che ad una proposizione vera o ad una lampadina accesa, che rivela quindi il passaggio della corrente attraverso il circuito, venga associato il valore 1 e alle situazioni opposte, proposizione falsa o lampadina spenta, il valore 0. Tener presente l'analogia tra le proposizioni e i circuiti elettrici può essere utile per interpretare il senso di certe operazioni logiche. È possibile infatti combinare tra loro più proposizioni, in modo da formarne di nuove, mediante operazioni che in questo contesto vengono chiamate *connettivi logici*.

I connettivi logici fondamentali sono la *disgiunzione* $\vee$, la *congiunzione* $\wedge$ e la *negazione* $\neg$. Date le due proposizioni p e q, la proposizione $p \vee q$, p *oppure* q, è vera quando almeno una delle due è vera, la proposizione $p \wedge q$, p *e* q, è vera quando entrambe le proposizioni sono vere e infine $\neg p$, *non* p, è vera quando p è falsa. Le seguenti *tabelle di verità* mostrano i valori di $p \vee q$, di $p \wedge q$ e di $\neg p$ in termini di quelli di p e di q

$p \vee q$:

$p \backslash q$	V	F
V	V	V
F	V	F

$p \wedge q$:

$p \backslash q$	V	F
V	V	F
F	F	F

$\neg p$:

p	$\neg p$
V	F
F	V

Se ad una proposizione vera facciamo corrispondere un interruttore chiuso e ad una falsa un interruttore aperto, nel circuito relativo alla disgiunzione i due interruttori vanno disposti in parallelo mentre in quello della congiunzione in serie.

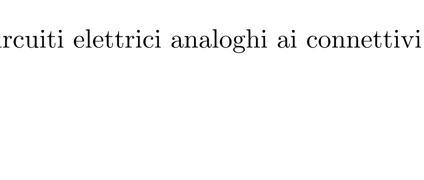


Figura 1.1: Circuiti elettrici analoghi ai connettivi logici $\vee$ e $\wedge$

Usando opportune tabelle di verità e/o circuiti si possono capire facilmente le seguenti proprietà di *reticolo* caratteristiche dell'*Algebra di Boole*

$(\vee)1.\ p \vee p = p$	$(\wedge)1.\ p \wedge p = p$	*idempotente*
$(\vee)2.\ p \vee q = q \vee p$	$(\wedge)2.\ p \wedge q = q \wedge p$	*commutativa*
$(\vee)3.\ (p \vee q) \vee r = p \vee (q \vee r)$	$(\wedge)3.\ (p \wedge q) \wedge r = p \wedge (q \wedge r)$	*associativa*
$(\vee)4.\ p \vee (q \wedge r) = (p \vee q) \wedge (p \vee r)$	$(\wedge)4.\ p \wedge (q \vee r) = (p \wedge q) \vee (p \wedge r)$	*distributiva*
$(\neg)5.\ \neg(\neg p) = p$		*legge della doppia negazione*
$(\neg)6.\ \neg(p \vee q) = \neg p \wedge \neg q$	$(\neg)6.\ \neg(p \wedge q) = \neg p \vee \neg q$	*leggi di De Morgan.*

Combinando tra loro i connettivi logici fondamentali se ne ottengono altri per i quali conviene introdurre nuovi simboli; di grande importanza per noi sono l'*implicazione* e l'*equivalenza*

$p \Rightarrow q = \neg p \vee q$ *p implica q, se p allora q,*

$p \Leftrightarrow q = (p \Rightarrow q) \wedge (q \Rightarrow p)$ *p equivale a q, p se e solo se q.*

L'implicazione si può esprimere anche dicendo che *p è condizione sufficiente affinché valga q*, oppure che *q è condizione necessaria affinché valga p.* Naturalmente nel caso dell'equivalenza si dice che *condizione necessaria e sufficiente affinché valga p (o q) è che valga q (o p).* Si osservi che $p \Rightarrow q$ è vera non solo quando p e q sono vere entrambe, ma anche quando p è falsa, qualunque sia la q, non si deve confondere la verità dell'implicazione con la verità di q. Invece nell'equivalenza, poiché le due proposizioni si implicano a vicenda, dovranno essere entrambe vere oppure entrambe false. Si ritrova così la nozione di equivalenza data all'inizio di questo paragrafo.

Di immediata verifica sono le seguenti proprietà dell'implicazione

$(\Rightarrow)1.\ p \Rightarrow p$ *riflessiva*

$(\Rightarrow)2.\ [(p \Rightarrow q) \wedge (q \Rightarrow p)] \Rightarrow (p \Leftrightarrow q)$ *antisimmetrica*

$(\Rightarrow)3.\ [(p \Rightarrow q) \wedge (q \Rightarrow r)] \Rightarrow (p \Rightarrow r)$ *transitiva.*

Inoltre combinando la definizione di implicazione con la $(\neg)1$ si ottiene anche

$(\Rightarrow)4.\ (p \Rightarrow q) \Leftrightarrow (\neg q \Rightarrow \neg p)$ *legge di contrapposizione*

sulla quale, come vedremo, si basa il metodo di dimostrazione per assurdo dei teoremi.

Per l'equivalenza si ha

$(\Leftrightarrow)1.\ p \Leftrightarrow p$ *riflessiva*

$(\Leftrightarrow)2.\ (p \Leftrightarrow q) \Rightarrow (q \Leftrightarrow p)$ *simmetrica*

$(\Leftrightarrow)3.\ [(p \Leftrightarrow q) \wedge (q \Leftrightarrow r)] \Rightarrow (p \Leftrightarrow r)$ *transitiva.*

Abbiamo visto che il senso di una proposizione sta nel poter stabilire se è vera o falsa. Invece l'affermazione $x > 3$ non è una proposizione perché il suo carattere dipende da x. Se ad x si attribuisce un valore numerico particolare si ottiene una proposizione che può essere vera o falsa a seconda del valore scelto e facendo variare tutti questi valori si ottengono altrettante proposizioni, alcune vere, altre false. La dipendenza $p(x)$ che in questo modo viene stabilita tra la proposizione e la variabile si chiama *predicato.* Anche se per un predicato non ha senso essere vero o falso, alcuni di essi li diciamo *sempre veri*, o *sempre falsi*, ma intendendoli più giustamente come proposizioni. Per esempio il predicato $x^2 \geqslant 0$, dove x è un numero reale, è per ogni scelta di x una proposizione vera, per questo diciamo che è sempre vero, ma intendiamo semplicemente che ad ogni valore della variabile x corrisponde una proposizione sempre vera.

Rispetto al concetto di proposizione, quello di predicato è più generale, infatti se si considera, come caso particolare, un predicato che associa ad ogni valore di x sempre la stessa proposizione p, cioè $p(x) = p$, ogni proposizione può essere vista come un predicato *costante.*

Predicati che dipendono da più variabili, come $x > y$, $x + y + z > 0$, $x_1^2 + x_2^2 + \ldots + x_n^2 = 1$, si chiamano anche *relazioni.* Una di queste è l'*uguaglianza*, $x = y$, che si definisce come quel predicato binario, cioè in due variabili, che diventa una proposizione vera in tutti e soli i casi in cui x e y sono lo stesso oggetto. Nel linguaggio corrente l'uguaglianza può avere un significato un po' diverso, simile a quello che si usa in geometria quando si parla di due figure sovrapponibili con una trasformazione rigida, ma proprio per non abusare dell'aggettivo *uguali* è invalso più di recente l'uso di chiamarle *congruenti*.

Come vengono scelte le variabili? Il predicato "*n è pari*" diventa ogni volta una proposizione vera se ad n si attribuiscono i valori 0, 2, 4, ecc., falsa se $n = 1, 3, 5, \ldots$, ecc., ma perde ogni significato come predicato se ad n si attribuiscono valori non interi. Non avere senso non va confuso con l'essere falsa. La proposizione "$0,5$ *è pari*" non è ben definita come proposizione, non ha alcun senso perché non è né falsa né vera. Neanche il predicato "*n è verde*" ha significato, a meno di non precisare che gli n vanno intesi come numeri telefonici.

Nell'introdurre un predicato è dunque essenziale stabilire in quale ambiente va scelta la variabile, o le variabili, a quali oggetti il predicato si applica. Chiameremo questo ambiente *classe* e lo accetteremo come concetto *primitivo*, ben distinto, peraltro, da quello di insieme che vedremo tra poco. Per un predicato usiamo dunque la notazione

$$x \in X\,, \quad p(x)\,,$$

dove viene specificata, mediante il simbolo $\in$ di *appartenenza*, la classe degli *elementi* ai quali il predicato p si applica. Scegliendo un particolare elemento $x_0 \in X$, $p(x)$ diventa la proposizione $p(x_0)$.

Più predicati si possono combinare tra loro mediante due simboli chiamati *quantificatori*, in modo da formare nel complesso un'affermazione per cui abbia senso stabilire se è vera o falsa, una proposizione dunque. Essi sono

$\forall$ *quantificatore universale* si legge: *per ogni*

$\exists$ *quantificatore esistenziale* si legge: *esiste (almeno uno)*

e in più si usa anche il simbolo $\exists_1$ per indicare che *esiste almeno uno e non più di uno*, o, brevemente, *esiste uno ed uno solo, esiste ed è unico.* Per fare un esempio, consideriamo la classe $\mathbf{R}$ dei numeri reali, sebbene ne parleremo estesamente più avanti. Con i due predicati

$$y \in \mathbf{R}\,,\; y \geqslant 0 \qquad \text{e} \qquad x, y \in \mathbf{R}\,,\; x^2 = y$$

si può costruire la proposizione

$$\forall y \in \mathbf{R} \;:\; y \geqslant 0 \quad \exists x \in \mathbf{R} \;:\; x^2 = y$$

che va letta nel modo seguente:

$$\textit{per ogni } y \in \mathbf{R} \textit{ tale che } y \geqslant 0 \textit{ esiste } x \in \mathbf{R} \textit{ tale che } x^2 = y$$

e il fatto che si tratti di una proposizione è evidente perché è nel complesso un'affermazione che può solo essere vera oppure falsa (in questo caso è vera). Bisogna porre attenzione all'ordine in cui vengono disposti i quantificatori perché cambiarlo dà luogo a proposizioni differenti. È ad esempio ovviamente falsa

$$\exists x \in \mathbf{R} \;:\; \forall y \in \mathbf{R}\,,\; y \geqslant 0\,, x^2 = y\,.$$

Se vogliamo negare che una certa proprietà valga per ogni $x \in X$ dobbiamo affermare che per almeno un $x \in X$ non vale, cioè vale la sua negazione. Viceversa, negare che esista un $x \in X$ che soddisfa una certa proprietà significa affermare che la proprietà contraria vale per tutti gli $x \in X$. Dunque la negazione di una proposizione del tipo

$$\forall x \in X\; \exists y \in Y \;:\; p(x, y)$$

è data da

$$\exists x \in X \ : \ \forall y \in Y \quad \neg p(x,y)$$

in cui si può osservare che, portando il simbolo di negazione sulla p, il $\forall$ va sostituito con l'$\exists$ e viceversa, e ciò avviene sempre, in qualunque ordine si trovino i due simboli e indipendentemente da quante volte compaiono. Pertanto la negazione della proposizione enunciata nell'esempio precedente diventa

$$\exists y \in \mathbf{R} \ : \ y \geqslant 0 \ : \ \forall x \in \mathbf{R} \quad x^2 \neq y\,.$$

1.2 Struttura del ragionamento matematico

Il procedimento ipotetico-deduttivo, con il quale si costruisce una teoria matematica, consiste in primo luogo nel formulare un sistema di proposizioni indipendenti e non contraddittorie (nessuna di esse deve implicarne un'altra né la relativa negazione) che si chiamano *assiomi* o *postulati*, riguardanti alcuni enti che non vengono definiti, i *concetti primitivi*. Si cerca poi di dedurne il maggior numero possibile di conseguenze, o meglio, delle conseguenze di qualche interesse e utilità. L'indagine comporta l'introduzione di nuovi concetti attraverso delle *definizioni* e la formulazione di nuove proposizioni, i *teoremi*, che richiedono invece una dimostrazione, cioè la verifica della loro verità, un termine che in questo contesto ha un significato relativo in quanto riferito alla verità degli assiomi. Un teorema si presenta generalmente sotto la forma di implicazione: se $x \in X$ viene scelto in modo che $p(x)$, *l'ipotesi*, sia vera (altrimenti l'implicazione sarebbe banalmente vera!), allora anche $q(x)$, la *tesi*, deve essere vera. La dimostrazione va sviluppata attraverso un ragionamento formalmente rigoroso, nel rispetto delle leggi della logica illustrate in questo paragrafo. Ma la logica da sola non basta, un ruolo fondamentale viene svolto dall'intuizione liberata da schemi rigidi e pregiudizi, dalla capacità di confrontare situazioni diverse e di stabilire analogie, dallo spirito critico che ci permette di valutare il senso di una proprietà o la portata del teorema mediante la costruzione di esempi e controesempi, in definitiva dal buon senso. Applicando sistematicamente questo tipo di indagine, che ci pone mille domande per ogni risposta (ma sono più le domande che le risposte a tener viva una materia), possiamo cogliere gli aspetti essenziali, lo scopo e il senso di una certa teoria. Attorno ad essa si formano così, in modo naturale, delle idee che ci dovrebbero guidare in una dimostrazione o nel risolvere problemi.

Vediamo adesso come si articola formalmente la dimostrazione di un teorema. Siano p l'ipotesi e q la tesi. L'implicazione $p \Rightarrow q$ non è di solito così immediata da riconoscere. Allora cerchiamo, come primo passo, di individuare delle conseguenze immediate di p per scegliere poi tra queste quella che ci sembra portare verso q. Da questa se ne troveranno altre che sperabilmente ci avvicinano alla tesi. Talvolta accade, nel corso di una dimostrazione, di scoprire che il teorema è falso se non si aggiungono altre ipotesi o se certe ipotesi non vengono modificate, oppure se ci si accorge dell'esistenza di un *controesempio*, cioè di una situazione vera che contraddice la tesi. Se tutto va bene la tesi viene raggiunta come ultima fase del ragionamento.

Si ottiene in questo modo una catena di implicazioni del tipo

$$p \Rightarrow r_1 \Rightarrow r_2 \Rightarrow \ \ldots \ \Rightarrow r_n \Rightarrow q\,,$$

dove il ragionamento è riconoscibile nel complesso delle varie proposizioni intermedie r_i. La correttezza di questa procedura sta nella proprietà ($\Rightarrow$)3.

Un altro metodo, che si basa sostanzialmente sulla ($\Rightarrow$)4, è il ragionamento *per assurdo*. Si prova a negare l'implicazione, cioè si nega il teorema assumendo per vera $\neg(\neg p \vee q)$ che, per la ($\neg$)6, significa assumere accanto alla p, come ipotesi aggiuntiva, anche la negazione della tesi $\neg q$. Poi si costruisce un ragionamento

$$\neg(\neg p \vee q) \Rightarrow r_1 \Rightarrow r_2 \Rightarrow \ \ldots \ \Rightarrow r_n$$

con una proposizione finale r_n che viene riconosciuta come falsa perché in contraddizione con p o con risultati precedentemente dimostrati o con gli assiomi. Poiché le implicazioni del ragionamento sono tutte vere, per la ($\Rightarrow$)4 son vere anche le implicazioni

$$\neg r_n \Rightarrow \neg r_{n-1} \Rightarrow \ldots \Rightarrow \neg r_2 \Rightarrow \neg r_1 \Rightarrow \neg p \vee q$$

ed essendo vera $\neg r_n$ deve essere vera anche $\neg p \vee q$ che per definizione coincide con $p \Rightarrow q$.

Se nell'enunciato di un teorema si afferma che due proposizioni p e q sono equivalenti, allora siamo in realtà di fronte ai due teoremi $p \Rightarrow q$ e $q \Rightarrow p$.

Talvolta un teorema viene chiamato *lemma* oppure *corollario*. Se si vuole dare un certo risalto ad un teorema ritenuto centrale nella teoria, ma la sua dimostrazione appare troppo lunga e pedante, conviene spezzare il ragionamento in teoremi minori, i lemmi, che preparano il terreno al teorema fondamentale. I corollari invece vengono dopo, sono conseguenze di qualche interesse del teorema centrale.

1.3 Insiemi

Assegnato un predicato p sulla classe X, gli elementi $x \in X$ che rendono vera la proposizione $p(x)$ formano un nuovo ente che si chiama *insieme* e che si indica con la notazione

$$A = \{x \in X \mid p(x)\}.$$

Per indicare che x è un elemento di A scriviamo $x \in A$. Se gli elementi di A sono in numero finito (ma non troppi!) si possono anche elencare

$$A = \{x_1, x_2, \ldots, x_n\},$$

ma la notazione precedente, che fa uso diretto del predicato, è senz'altro più comoda e di uso più frequente. Per esempio, se X è la classe $\mathbf{N}$ dei numeri naturali, l'insieme dei numeri pari

$$\mathbf{P} = \{0, 2, 4, \ldots\}$$

può essere indicato in uno dei seguenti modi

$$\mathbf{P} = \{n \in \mathbf{N} \mid n \text{ è pari}\} = \{n \in \mathbf{N} \mid \exists k \in \mathbf{N} \ : \ n = 2k\} = \{2n \mid n \in \mathbf{N}\}$$

dove compaiono dei predicati. Naturalmente ad una famiglia di predicati tutti tra loro equivalenti, cioè che diventano proposizioni vere per gli stessi elementi della classe, corrisponde lo stesso insieme A e tra questi vi è anche il predicato $x \in A$. Viceversa, l'insieme A è formato da tutti e soli gli elementi $x \in X$ per i quali il predicato $x \in A$, insieme a tutti quelli ad esso equivalenti, è una proposizione vera. In particolare, due insiemi A e B coincidono se $x \in A$ e $x \in B$ sono proposizioni equivalenti per ogni $x \in X$. Ad esempio i due predicati su $\mathbf{R}$: $x(1-x) \geqslant 0$ e $0 \leqslant x \leqslant 1$ definiscono lo stesso insieme, l'intervallo $[0,1]$ perché diventano proposizioni entrambe vere per tutti e soli gli elementi di quell'intervallo.

Se il predicato $p(x)$, $x \in X$, definisce l'insieme

$$A = \{x \in X \mid p(x)\},$$

$\neg p(x)$ definisce l'insieme *complementare* di A

$$\complement A = \{x \in X \mid \neg p(x)\},$$

formato da tutti gli elementi di X che non appartengono ad A perché rendono vera la negazione di $p(x)$. L'insieme individuato da un predicato sempre vero non è altro che l'intera classe X e un predicato sempre falso definisce l'*insieme vuoto* $\emptyset$, pertanto $\complement X = \emptyset$ e $\complement \emptyset = X$.

L'*unione* e l'*intersezione* di due insiemi $A = \{x \in X \mid p(x)\}$ e $B = \{x \in X \mid q(x)\}$ sono rispettivamente

$$\begin{aligned} A \cup B &= \{x \in X \mid p(x) \vee q(x)\} = \{x \in X \mid x \in A \vee x \in B\}\,, \\ A \cap B &= \{x \in X \mid p(x) \wedge q(x)\} = \{x \in X \mid x \in A \wedge x \in B\}\,. \end{aligned}$$

Dunque $x \in A \cup B$ se e solo se almeno uno dei due insiemi, A o B, ha x come elemento, mentre $x \in A \cap B$ se x appartiene sia ad A che a B. È immediato riconscere che $A \cup \complement A = X$ e che $A \cap \complement A = \emptyset$. Due insiemi A e B si dicono *disgiunti* se $A \cap B = \emptyset$, cioè se la proposizione $x \in A \wedge x \in B$ è falsa per ogni $x \in X$.

Definiamo *complementare di B rispetto ad A*, o *differenza* $A - B$, l'insieme

$$\complement_A B = A - B = A \cap \complement B = \{x \in X \mid x \in A \wedge x \notin B\}\,.$$

L'analogia tra la logica delle proposizioni e la teoria degli insiemi è così forte da renderci evidente il passaggio dai connettivi logici alle operazioni insiemistiche.

Esercizio 1.1 *A partire dalle proprietà* $(\vee)1$—$(\neg)3$ *per le proposizioni, ricostruire le analoghe per gli insiemi*

$(\cup)1.\ A \cup A = A$	$(\cap)1.\ A \cap A = A$
$(\cup)2.\ A \cup B = B \cup A$	$(\cap)2.\ A \cap B = B \cap A$
$(\cup)3.\ (A \cup B) \cup C = A \cup (B \cup C)$	$(\cap)3.\ (A \cap B) \cap C = A \cap (B \cap C)$
$(\cup)4.\ A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$	$(\cap)4.\ A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
$(\complement)5.\ \complement(\complement A) = A$	
$(\complement)6.\ \complement(A \cup B) = \complement A \cap \complement B$	$(\complement)6.\ \complement(A \cap B) = \complement A \cup \complement B$

La $(\cup)3$ e la $(\cap)3$ ci dicono che le espressioni $A \cup B \cup C$ e $A \cap B \cap C$ non presentano alcuna ambiguità, neanche nell'ordine in cui compaiono i tre insiemi per le $(\cup)2$ e $(\cap)2$. Si possono pertanto considerare unioni e intersezioni di famiglie arbitrarie, finite o infinite, di insiemi

$$\mathscr{A} = \{A_k\}_{k=1,\ldots,n}\,, \qquad \mathscr{A} = \{A_n\}_{n \in \mathbf{N}}\,, \qquad \mathscr{A} = \{A_i\}_{i \in I}\,,$$

dove I è un insieme non meglio definito di indici, ponendo

$$\bigcup_{k=1}^{n} A_k = A_1 \cup A_2 \cup \ldots \cup A_n = \{x \in X \mid \exists k = 1, \ldots, n \ :\ x \in A_k\}\,,$$

$$\bigcap_{k=1}^{n} A_k = A_1 \cap A_2 \cap \ldots \cap A_n = \{x \in X \mid \forall k = 1, \ldots, n\ x \in A_k\}\,,$$

$$\bigcup_{n=1}^{\infty} A_n = \{x \in X \mid \exists n \in \mathbf{N} \ :\ x \in A_n\}, \qquad \bigcap_{i \in I} A_i = \{x \in X \mid x \in A_i\ \forall i \in I\}.$$

Diciamo che l'insieme A è *sottoinsieme* dell'insieme B, o che A è *contenuto*, o *incluso*, in B, e si scrive $A \subset B$ o $A \subseteq B$, se

$$\forall x \in X\,, \quad x \in A \Rightarrow x \in B\,.$$

In questo caso si dice anche che B *contiene* A e si scrive $B \supset A$ o $B \supseteq A$. Ad esempio l'insieme

$$P' = \{n \in \mathbf{N} \mid \exists k \in \mathbf{N} \ :\ n = k(k+1)\}$$

è contenuto nell'insieme $\mathbf{P}$ dei numeri pari perché il prodotto di due numeri naturali consecutivi è necessariamente pari, ma ci sono dei numeri pari che non sono di quel tipo come 4, 8, 10, 14 e tanti altri. Diciamo che A è *sottoinsieme proprio* di B, o che A è strettamente contenuto in B, se $A \subset B$ e $A \neq B$. Se avviene che

$$\forall x \in X\,, \quad x \in A \Leftrightarrow x \in B\,,$$

cioè $A \subset B$ e $B \subset A$, allora $A = B$, l'appartenenza ad uno o all'altro sono equivalenti e i due insiemi coincidono perché hanno gli stessi elementi.

Si faccia attenzione a non confondere l'appartenenza con l'inclusione, la prima è una relazione tra elementi e insiemi, mentre la seconda è una relazione tra insiemi, sono due concetti ben distinti contrariamente a quanto si possa ritenere. Se x è un elemento dell'insieme $A \subset X$, il simbolo $\{x\}$ indica il sottoinsieme di A formato dal solo elemento x; così è corretto scrivere $\{x\} \subset A$, ma $\{x\} \in A$ non ha senso. Ciò non toglie che sia lecito considerare insiemi i cui elementi sono a loro volta insiemi. Per esempio l'*insieme delle parti* di X, che si indica con $\mathscr{P}(X)$, è l'insieme che ha per elementi tutti i sottoinsiemi di X e quindi è lecito scrivere $\{x\} \in \mathscr{P}(X)$, così come è lecito scrivere $\{x\} \in \mathscr{P}(A)$ se $x \in A \subset X$.

Si noti che $\mathscr{P}(A \cup B) = \mathscr{P}(A) \cup \mathscr{P}(B)$ e che $\mathscr{P}(A) \subset \mathscr{P}(B)$ se $A \subset B$.

Sia p un predicato tale che la proposizione $p(x)$ sia falsa per ogni $x \in X$, come $x \neq x$, $x^2 < 0$ o $\operatorname{sen} x = 2$ se $x \in \mathbf{R}$ ecc.. Poiché l'implicazione $p(x) \Rightarrow q(x)$ è in questo caso sempre vera qualunque sia q, $\emptyset = \{x \in X \mid p(x)\}$, che diremo *insieme vuoto*, è contenuto in ogni insieme della classe X, ovvero $\emptyset \in \mathscr{P}(X)$. Di conseguenza l'insieme vuoto è unico: se ve ne fosse un altro, $\emptyset'$, si avrebbe $\emptyset' \subset \emptyset$ e $\emptyset \subset \emptyset'$, da cui $\emptyset' = \emptyset$.

Esercizio 1.2 - *Dimostrare che $A \cup \emptyset = A$, $A \cap \emptyset = \emptyset$ e $A \subset B \Leftrightarrow \complement A \supset \complement B$ per ogni coppia di insiemi $A, B \subset X$.*

Esercizio 1.3 - *Sapendo che X è un insieme formato da 3 elementi, $X = \{a, b, c\}$, quali e quanti sono gli elementi di $\mathscr{P}(X)$? E quelli di $\mathscr{P}(\mathscr{P}(X))$?*

La distinzione concettuale tra classe, che è un concetto primitivo, e insieme, che è formato dagli elementi della classe per i quali un predicato diventa una proposizione vera, è importante nell'ambito astratto della teoria degli insiemi, ma per i nostri scopi è irrilevante, tanto che ogni classe è anche il massimo insieme in essa contenuto, precisamente quello relativo ad un predicato sempre vero. Tuttavia continueremo talvolta ad usare il termine *classe*, ma nel senso di un insieme ragionevolmente grande da scegliere come ambiente nel quale certi problemi trovano una collocazione naturale.

1.4 Prodotto cartesiano e relazioni

Dati due insiemi X e Y e due elementi $x \in X$ e $y \in Y$, l'insieme $\{x, y\}$ è indistinguibile da $\{y, x\}$, l'ordine dei suoi elementi è del tutto irrilevante. Ci interessa invece definire un nuovo ente che sia formato dagli stessi elementi e dove abbia importanza anche l'ordine in cui vengono disposti, ad esempio possiamo considerare l'insieme $\{x, \{y\}\}$, il quale coincide ovviamente con $\{\{y\}, x\}$, ma non con $\{\{x\}, y\} = \{y, \{x\}\}$. Poniamo allora

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

in modo che $(x, y) \neq (y, x)$. Il nuovo elemento (x, y) si chiama *coppia ordinata*, o semplicemente coppia. Il *prodotto cartesiamo* dei due insiemi è l'insieme

$$X \times Y = \{(x, y) \mid x \in X\,,\ y \in Y\}$$

formato da tutte le coppie ordinate (x, y), al variare di $x \in X$ e di $y \in Y$. Nel caso particolare $X = Y$ viene indicato anche con X^2. Il prodotto cartesiano di un insieme $A \subset X$ con un insieme $B \subset Y$ è dato da

$$A \times B = \{(x, y) \in X \times Y \mid x \in A\,,\ y \in B\}\,.$$

Si può estendere la nozione di prodotto a famiglie qualsiasi, finite o infinite, di insiemi. Nel caso finito si parla di terne, quaterne, n-uple e il prodotto cartesiano diventa

$$X_1 \times X_2 \times \ldots \times X_n = \prod_{k=1}^{n} X_k = \{(x_1, x_2, \ldots, x_n) \mid x_k \in X_k\}\,.$$

Esercizio 1.4 - *Dimostrare che se $A, B \subset X$ e $C, D \subset Y$*

$$\begin{aligned}(A \cup B) \times (C \cup D) &= (A \times C) \cup (B \times D) \cup (A \times D) \cup (B \times C)\,,\\ (A \cap B) \times (C \cap D) &= (A \times C) \cap (B \times D) \cap (A \times D) \cap (B \times C)\,.\end{aligned}$$

Esempi

1.1 *Se $A = \{a, b\}$ e $B = \{1, 2, 3\}$ allora*

$$A \times B = \{(a, 1), (a, 2), (a, 3), (b, 1), (b, 2), (b, 3)\}\,.$$

1.2 *Dopo aver introdotto opportuni sistemi di riferimento, il piano e lo spazio della geometria euclidea possono essere identificati con i prodotti cartesiani*

$$\mathbf{R} \times \mathbf{R} = \mathbf{R}^2 = \{(x, y) \mid x, y \in \mathbf{R}\} \qquad e \qquad \mathbf{R}^3 = \{(x, y, z) \mid x, y, z \in \mathbf{R}\}\,.$$

Più in generale, lo spazio euclideo n-dimensionale si identifica con l'insieme delle n-uple ordinate $x = (x_1, x_2, \ldots\ldots, x_n)$ di numeri reali

$$\mathbf{R}^n = \{(x_1, x_2, \ldots\ldots, x_n) \mid x_i \in \mathbf{R}\ \forall i = 1, 2, \ldots\ldots, n\}$$

1.3 *Un rettangolo $R \subset \mathbf{R}^2$ è il prodotto cartesiano di due intervalli*

$$R = \{(x, y) \in \mathbf{R}^2 \mid a \leqslant x \leqslant b\,,\ c \leqslant y \leqslant d\} = I \times J\,,$$

dove $I = \{x \in \mathbf{R} \mid a \leqslant x \leqslant b\}$ e $J = \{y \in \mathbf{R} \mid c \leqslant y \leqslant d\}$. Ogni lato di R è il prodotto cartesiano di un intervallo per un numero, per esempio $I \times \{d\}$ è il lato orizzontale superiore. Quindi se nella definizione di R le disuguaglianze sono strette, rimane solo la parte interna di R, che è il rettangolo senza il suo bordo.

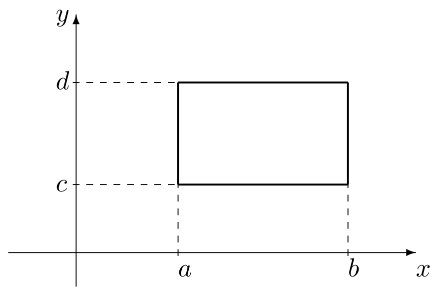


Figura 1.2: Il rettangolo come prodotto cartesiano di due intervalli

1.4 *Un parallelepipedo di $\mathbf{R}^3$ è il prodotto cartesiano di tre intervalli, o anche di un rettangolo per un intervallo. Ogni faccia è il prodotto di due intervalli per un numero e ogni spigolo è il prodotto di un intervallo per due numeri.*

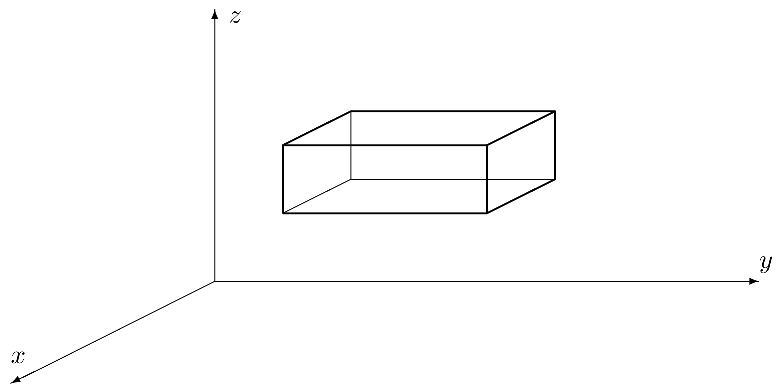


Figura 1.3: Il parallelepipedo come prodotto cartesiano di tre intervalli

1.5 *La superficie di un cilindro circolare retto è il prodotto cartesiano di una circonferenza per una retta, mentre il prodotto di un cerchio per una retta è il cilindro pieno. Il cerchio ne è la sezione che rimane costante lungo tutto il cilindro e viene ottenuta per traslazione della base lungo l'asse. Se al posto del cerchio si considera una regione piana D questa genera, sempre per traslazione, un cilindro a sezione D che risulta quindi dato come prodotto cartesiano fra D e una retta. Il cilindro è dunque illimitato, ma sarebbe limitato se al posto della retta si considerasse un segmento.*

Consideriamo due classi X e Y, due rispettivi insiemi $A \subset X$ e $B \subset Y$ e il loro prodotto cartesiano $A \times B \subset X \times Y$. Si chiama *relazione binaria* sul *dominio* $A \times B$, o brevemente *relazione* su $A \times B$, un predicato $\Re(x, y)$ con $x \in A$ e $y \in B$. Due relazioni coincidono se hanno lo stesso dominio e sono equivalenti come predicati, cioè se diventano proposizioni equivalenti per le stesse coppie $(x, y) \in A \times B$. Il grafico di $\Re$ è l'insieme $G(\Re) \subset X \times Y$ definito da

$$G(\Re) = \{(x, y) \in X \times Y \mid x \in A\,,\ y \in B\,,\ \Re(x, y)\}\,.$$

I casi più comuni di relazioni sono le equazioni e le disequazioni in $\mathbf{R}^2$, $\mathbf{R}^3$, $\mathbf{R}^n$, i loro grafici sono le relative soluzioni che, in quanto insiemi di coppie, terne ecc. di numeri reali, si possono immaginare, o disegnare, come curve, superfici, punti isolati, regioni del piano o dello spazio.

Esempi

1.6 *La ben nota relazione cartesiana $ax + by = c$ definita per ogni $(x, y) \in \mathbf{R}^2$ ha per grafico una retta se i coefficienti (dati) a e b in $\mathbf{R}$ non sono entrambi nulli, altrimenti rappresenta tutto il piano se $a = b = c = 0$ e il vuoto se $a = b = 0$ e $c \neq 0$. Se $b \neq 0$, posto $m = -a/b$ e $q = c/a$, si ottiene la forma equivalente $y = mx + q$. Questa relazione viene detta equazione in forma esplicita della retta perché rispetto alla precedente, che è quella implicita, una delle due variabili è esprezza in funzione dell'altra. Considerazioni analoghe valgono per l'equazione di un piano nello spazio nelle forme implicita $ax + by + cz = d$ ed esplicita $z = mx + ny + q$.*

La stessa relazione in due variabili $ax + by = c$ ha significato anche nello spazio, basta pensarla così $ax + by + 0z = c$, come un'identità rispetto a z. Essa ha per grafico il piano parallelo all'asse z, sempre che a e b non siano entrambi nulli, che interseca il piano $z = 0$ lungo la retta precedente.

Una rappresentazione diversa, detta *parametrica* può essere ottenuta introducendo una nuova variabile $t \in \mathbf{R}$, il *parametro*, rispetto alla quale far dipendere le coordinate del punto sulla curva. Per la retta nel piano si può porre $x = t$, e di conseguenza

$y = mt + q$, nell'equazione esplicita ($y = t$ nell'altro caso di forma esplicita) ma vi sono infiniti altri modi. Se si interpreta t come la variabile temporale si ottiene la legge con cui la posizione di un punto $P(t)$ dipende dal tempo t. Ad esempio nel moto *rettilineo e uniforme* si ha $P(t) = P_0 + vt = (x_0 + v_1 t, y_0 + v_2 t, z_0 + v_3 t)$ dove $v = (v_1, v_2, v_3)$ è il vettore velocità.

Rispetto a quella cartesiana, la relazione in forma parametrica contiene in più l'informazione sul *modo* in cui la curva viene percorsa. Sostituendo ad esempio t con $2t$ nel moto precedente raddoppia la velocità.

1.7 *Dato $r > 0$, la relazione $x^2 + y^2 = r^2$ con $(x, y) \in \mathbf{R}^2$, anche questa ben nota, ha per grafico la circonferenza di centro $O = (0, 0)$ e raggio r.*

Una descrizione parametrica particolarmente semplice della stessa curva è la seguente

$$\begin{cases} x = r\cos\vartheta \\ y = r \operatorname{sen}\vartheta\,, \qquad \vartheta \in \mathbf{R}\,, \end{cases} \tag{1.1}$$

ma ve ne sono infinite altre, ognuna corrispondente ad una particolare legge di percorrenza $\vartheta(t)$. Ad esempio per il moto circolare e uniforme basta porre $\vartheta = \omega t$, dove ω, detta *frequenza*, è costante. L'intero giro ha durata temporale $T = 2\pi/\omega$, detto *periodo*. Per i moti pendolari, oscillatori o progressivi che siano, si avranno altre funzioni $\vartheta(t)$. Fissando invece il valore di ϑ e lasciando variare r, si ottengono i punti della semiretta uscente dall'origine e formante l'angolo ϑ con l'sse x. Infine, facendo variare entrambi, si ottengono tutti i punti del piano e la (1.1) stabilisce la relazione con cui le coordinate cartesiane (x, y) dipendono dalle *coordinate polari* (r, ϑ).

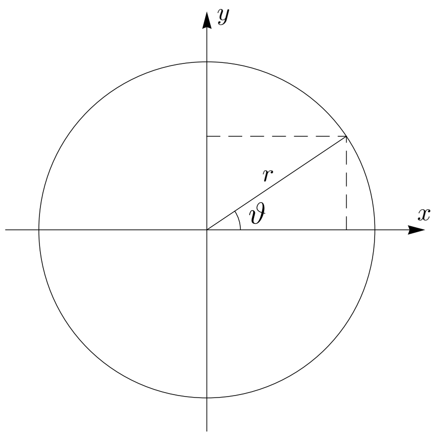


Figura 1.4: La circonferenza con le variabili che la descrivono

1.8 *Immaginando che nella relazione dell'esempio precedente sia presente anche la variabile $z \in \mathbf{R}$, pensata ad esempio come $x^2 + y^2 + 0z^2 = r^2$ con $(x, y, z) \in \mathbf{R}^3$, il relativo grafico sarà un cilindro circolare retto con l'asse z come asse di simmetria. Il fatto che sia un'identità rispetto a z significa che è invariante per traslazioni lungo l'asse z, cioè tutte le sezioni ottenute tagliando la superficie con i vari piani $z =$ costante sono curve sovrapponibili e congruenti alla stessa circonferenza. Trattandosi di una superficie, ci aspettiamo l'uso di due parametri indipendenti in ogni sua rappresentazione parametrica. La più naturale è questa*

$$\begin{cases} x = r\cos\vartheta \\ y = r \operatorname{sen}\vartheta \\ z = t \end{cases} \tag{1.2}$$

dove ϑ è lo stesso di prima e t varia in $\mathbf{R}$ o in un intervallo, a seconda del tipo di cilindro che si vuole descrivere. Se lasciamo variare anche $r \geqslant 0$ otteniamo tutto $\mathbf{R}^3$ descritto nelle coordinate (polari) cilindriche (r, ϑ, t).

1.9 *L'ellisse di semiassi $a > 0$ e $b > 0$ è il grafico di una relazione simile a quella della circonferenza*

$$\begin{cases} x = a\cos\varphi \\ y = b\,\mathrm{sen}\,\varphi\,, \qquad \varphi \in \mathbf{R}\,. \end{cases}$$

Eliminando φ si perviene alla ben nota equazione cartesiana $(x/a)^2 + (y/b)^2 = 1$.

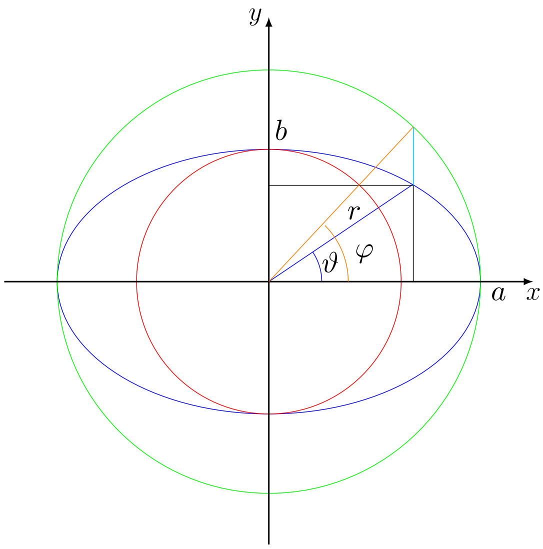


Figura 1.5: L'ellisse, si noti che $a\cos\varphi = r\cos\vartheta$ e $b\,\mathrm{sen}\,\varphi = r\,\mathrm{sen}\,\vartheta$

1.10 *La posizione di un punto P su una superficie sferica di centro O e raggio r dipende da una coppia di coordinate angolari, la latitudine ϑ rispetto al piano equatoriale e la longitudine φ rispetto al meridiano di Greenwich, usando il linguaggio della geografia terrestre.*

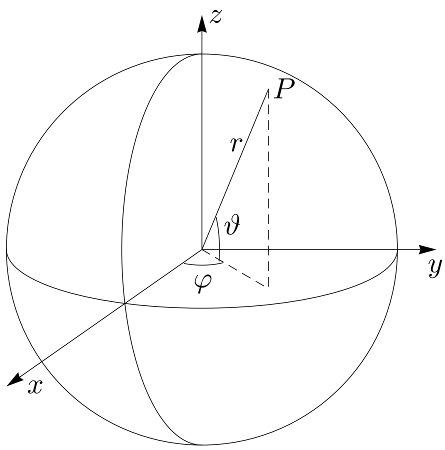


Figura 1.6: La sfera

Le relazioni fra le tre coordinate cartesiane e i due angoli sono

$$\begin{cases} x = r\cos\vartheta\cos\varphi \\ y = r\cos\vartheta\,\mathrm{sen}\,\varphi \\ z = r\,\mathrm{sen}\,\vartheta\,, \end{cases} \tag{1.3}$$

dove $\vartheta \in [-\pi/2, \pi/2]$ *e* $\varphi \in [0, 2\pi]$ *per cui la sfera può essere identificata col prodotto cartesiano di una circonferenza per una semicirconferenza.*

Se si fa variare anche $r \geqslant 0$ si ottengono tutti i punti dello spazio $\mathbf{R}^3$ e i parametri (r, ϑ, φ) si chiamano *coordinate sferiche.*

Esercizio 1.5 - *Scrivere le coordinate di un punto della circonferenza che rotola senza strisciare su una retta.*

1.11 *Una circonferenza di raggio r che, insieme al piano che la contiene, compie una rivoluzione completa attorno ad una retta di questo piano, posta a distanza* $R > r$ *dal centro, genera una superficie che si chiama* ***toro****. Per descriverla servono due coordinate angolari che variano da 0 a* 2π*, per questo possiamo identificare il toro con il prodotto cartesiano di due circonferenze.*

Se $\vartheta \in [0, 2\pi]$ è l'angolo che varia lungo la circonferenza e $\varphi \in [0, 2\pi]$ quello di rotazione, supponiamo attorno all'asse z, le relazioni tra le coordinate cartesiane e i due parametri sono date da

$$\begin{cases} x = (R + r\cos\vartheta)\cos\varphi \\ y = (R + r\cos\vartheta)\,\mathrm{sen}\,\varphi \\ z = r\,\mathrm{sen}\,\vartheta\,. \end{cases}$$

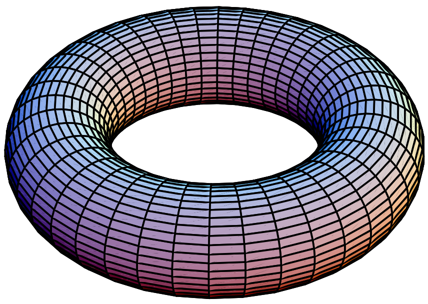


Figura 1.7: Il toro

Vedremo più avanti altri esempi, adesso consideriamo due importanti tipi di relazione.

Una relazione $\Re(x, y)$ su X^2 che gode delle proprietà

($\leqslant$)1. $\forall x \in X\,,\ \Re(x, x)$ *riflessiva*
($\leqslant$)2. $\forall x, y \in X\,,\ \Re(x, y) \wedge \Re(y, x) \Rightarrow x = y$ *antisimmetrica*
($\leqslant$)3. $\forall x, y, z \in X\,,\ \Re(x, y) \wedge \Re(y, z) \Rightarrow \Re(x, z)$ *transitiva*

si chiama *relazione d'ordine*, o anche *ordinamento*, su X e al posto di $\Re(x, y)$ si usa scrivere $x \leqslant y$. La notazione $(X, \leqslant)$ indica la presenza di una relazione d'ordine su X che per questo viene detto *insieme ordinato.* Se in più vale la proprietà

($\leqslant$)4. $\forall x, y \in X\,,\ \Re(x, y) \vee \Re(y, x)$

allora si tratta di una *relazione d'ordine totale* e in questo caso X viene detto *totalmente ordinato.* Nella ($\leqslant$)4 si afferma

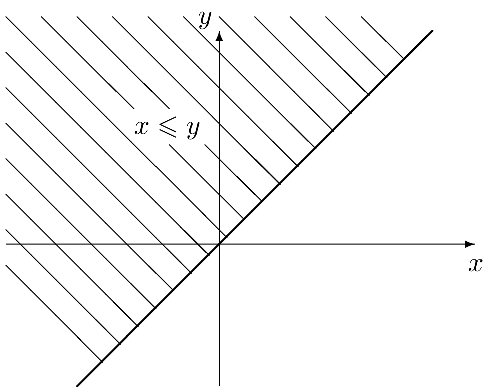


Figura 1.8: La relazione d'ordine su **R**

che due elementi comunque scelti nell'insieme sono sempre confrontabili, cioè uno dei due precede (è in relazione con) l'altro. Se la ($\leqslant$)4 non vale si tratta di una *relazione d'ordine parziale* e l'insieme viene detto *parzialmente ordinato*, l'inclusione tra insiemi ne è un esempio. Si pone $x \geqslant y$ se $y \leqslant x$, $x < y$ se $x \leqslant y$ e $x \neq y$ e infine $x > y$ se $y < x$. Il grafico della relazione $x \leqslant y$ in $\mathbf{R}^2$ è il semipiano tratteggiato nella Figura 1.8.

Una relazione $\Re(x, y)$ su X^2 che gode delle proprietà

($\cong$)1. $\forall x \in X\,, \Re(x, x)$ *riflessiva*
($\cong$)2. $\forall x, y \in X\,, \Re(x, y) \Rightarrow \Re(y, x)$ *simmetrica*
($\cong$)3. $\forall x, y, z \in X\,, \Re(x, y) \wedge \Re(y, z) \Rightarrow \Re(x, z)$ *transitiva*

si chiama *relazione di equivalenza* su X. Per questa relazione useremo il simbolo $x \sim y$, ma si incontrano di frequente anche $x \equiv y$ e $x \cong y$. Relazioni di equivalenza sono, oltre all'uguaglianza e all'equivalenza tra proposizioni, anche il parallelismo tra rette, la congruenza e la similitudine tra figure geoemtriche. La relazione di uguaglianza su $\mathbf{R}$ ha come grafico la retta in $\mathbf{R}^2$ che nella figura precedente separa il piano tratteggiato dal suo opposto.

Data la coppia $(X, \sim)$, che indica un insieme X munito della relazione di equivalenza $\sim$, consideriamo per ogni $x \in X$ tutti gli elementi equivalenti a x. L'insieme che essi formano si chiama *classe di equivalenza* di x e si indica con $[x]$.

Esercizio 1.6 - *Dimostrare che valgono le seguenti proprietà delle classi di equivalenza*

[$\sim$]1. *se* $y \in [x]$ *allora* $[x] = [y]$

[$\sim$]2. $[x]$ *e* $[y]$ *hanno intersezione vuota oppure coincidono. In altre parole*

$$[x] \cap [y] \neq \emptyset \Rightarrow [x] = [y]\,.$$

L'elemento x stesso, ma anche qualunque altro ad esso equivalente, può essere scelto per indicare l'intera sua classe e viene chiamato *rappresentante* della classe $[x]$. L'insieme che ha per elementi le classi $[x]$ al variare di $x \in X$ si chiama *insieme quoziente* di X rispetto alla relazione di equivalenza $\sim$, si indica col simbolo $X/_\sim$ e si legge X *modulo* $\sim$. Ad esempio, se nello spazio euclideo si considerano equivalenti due segmenti orientati che hanno la stessa lunghezza e lo stesso verso, l'insieme quoziente è quello dei vettori; ogni vettore è una classe di equivalenza di segmenti orientati e disegnare una freccia significa scegliere un vettore usando un particolare rappresentante.

Il quoziente $X/_\sim$ è una *partizione* di X, cioè una famiglia di sottoinsiemi non vuoti di X, a due a due disgiunti, la cui unione ricopre tutto X. Infatti nessun insieme $S \in X/_\sim$ può essere vuoto perché essendo la classe di qualche elemento x deve contenere almeno x. Che sono a due a due disgiunti deriva dall'Esercizio 1.6. Infine, ogni elemento $x \in X$ deve appartenere almeno alla sua stessa classe $[x]$.

Poiché nell'insieme quoziente tutti gli elementi tra loro equivalenti vengono identificati in un solo elemento, la relazione di equivalenza può essere interpretata anche come operazione di "incollamento", come è evidente nel primo dei seguenti esempi.

Esempi

1.12 *Definiamo sul quadrato* $Q = [0, 1] \times [0, 1]$ *la relazione di equivalenza*

$$(x, y) \sim (x', y') \Leftrightarrow \begin{cases} x = x' \text{ e } y = y' \\ x = 0\,,\ x' = 1 \text{ e } \ y = y'\,. \end{cases}$$

Il quoziente $Q/_\sim$ può essere immaginato come un cilindro circolare di altezza unitaria.

1.13 *Modificando la relazione dell'esempio precedente in questo modo*

$$(x,y) \sim (x',y') \Leftrightarrow \begin{cases} x = x' \text{ e } y = y' \\ x = 0\,,\ x' = 1 \text{ e } \ y = 1 - y'\,. \end{cases}$$

i due lati verticali del quadrato vengono identificati rovesciati uno rispetto all'altro e l'insieme quoziente diventa il nastro di Möbius.

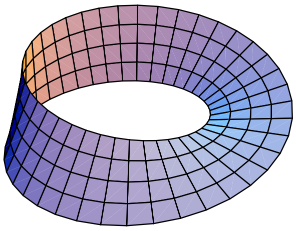


Figura 1.9: Il nastro di Möbius

Esercizio 1.7 - *Riconoscere il toro nella relazione di equivalenza su Q*

$$(x,y) \sim (x',y') \Leftrightarrow \begin{cases} x = x' \text{ e } y = y' \\ x = 0\,,\ x' = 1 \text{ e } \ y = y' \\ y = 0\,,\ y' = 1 \text{ e } \ x = x'\,. \end{cases}$$

Esercizio 1.8 - *Riconoscere la bottiglia di Klein della Figura 1.10 nella relazione di equivalenza su Q*

$$(x,y) \sim (x',y') \Leftrightarrow \begin{cases} x = x' \text{ e } y = y' \\ x = 0\,,\ x' = 1 \text{ e } \ y = y' \\ y = 0\,,\ y' = 1 \text{ e } \ x = 1 - x'\,. \end{cases}$$

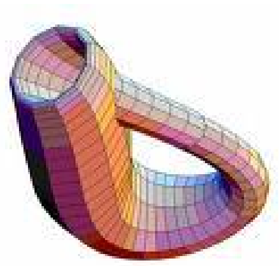


Figura 1.10: La bottiglia di Klein

1.14 *Nell'insieme $\mathbf{Z}$ dei numeri interi relativi la scrittura $m \equiv n \bmod k$, che si legge m congruo ad n modulo k, significa che la differenza $m - n$ è un multiplo intero di k. È immediato verificare che si tratta di una relazione di equivalenza. L'insieme quoziente*

$$\mathbf{Z}_k = \{[0], [1], \ldots, [k-1]\}$$

ha k elementi che possono essere identificati con k numeri interi consecutivi come rappresentanti.

Esercizio 1.9 - *Dato un numero reale T, verificare che*

$$x \sim y \Leftrightarrow \exists n \in \mathbf{Z} \ : \ x - y = nT$$

è una relazione di equivalenza in $\mathbf{R}$. Descrivere l'insieme quoziente.

Esercizio 1.10 - *Le seguenti relazioni in $\mathbf{R}$ sono di equivalenza?*

$$x \sim y \Leftrightarrow x - y \in \mathbf{Q}\,, \qquad \text{e} \qquad x \sim y \Leftrightarrow x - y \in \mathbf{R} - \mathbf{Q}\,.$$

1.5 Funzioni

Siano X e Y due insiemi e $A \subset X$.

Definizione 1.1 - *Una relazione $\Re$ su $A \times Y$ con la seguente proprietà*

$$\forall x \in A\,, \quad \forall y, y' \in Y\,, \qquad \Re(x,y) \wedge \Re(x,y') \Rightarrow y = y' \tag{1.4}$$

si chiama ***funzione****, o anche* ***applicazione*** *o* ***mappa****, definita sul* ***dominio*** *A e a valori sul* ***codominio*** *Y, si indica con la notazione*

$$f : A \to Y$$

e si legge f manda, o trasforma, A in Y.

Da quanto detto sulle relazioni segue che due funzioni $f, g : A \to Y$ coincidono, e si scrive $f = g$, se e solo se $f(x) = g(x)$ per ogni $x \in A$. Il grafico di f viene ad essere l'insieme

$$G(f) = \{(x,y) \in A \times Y \mid y = f(x)\ \forall x \in A\} = \{(x, f(x)) \mid x \in A\}\,.$$

Per meglio chiarire la (1.4) osserviamo che ogni elemento $x \in A$ deve essere in relazione con uno ed un solo elemento di Y, ma è consentito che due o più elementi di A lo siano con uno stesso elemento $y \in Y$. Ad esempio, una funzione *costante* associa a *tutti* gli elementi di A uno stesso elemento di Y, cosa che si può esprimere nella forma $f(x) = y$ per ogni $x \in A$, oppure affermando che $f(x_1) = f(x_2)$ per ogni $x_1, x_2 \in A$.

Una funzione $f : A \to Y$ va interpretata come una legge che, procedendo da A verso Y, fa corrispondere ad ogni elemento $x \in A$ uno ed un solo $y \in Y$ che dunque dipende da x e viene indicato con $f(x)$. Vista come *trasformazione*, la f opera su ogni elemento $x \in A$ e lo trasforma nell'elemento $y = f(x) \in Y$.

Come abbiamo detto per i predicati e le classi in generale, ogni volta che si vuole definire una funzione è essenziale precisarne il dominio e il codominio, cioè l'ambiente in cui deve operare, e dare la regola con la quale deve trasformare gli elementi di A nei corrispondenti elementi di Y. La legge $x \to f(x) = x^2$ va considerata una funzione ogni volta diversa a seconda degli insiemi numerici che si scelgono come dominio e come codominio anche se la legge, che associa ad ogni numero il suo quadrato, è sempre quella.

Definizione 1.2 - *Si chiama* ***restrizione*** *della funzione $f : A \to Y$ al dominio $B \subset A$ la funzione $f_{|B} : B \to Y$ definita da*

$$f_{|B}(x) = f(x) \qquad \forall x \in B\,.$$

Definizione 1.3 - *Si chiama* ***prolungamento****, o* ***estensione****, della funzione $f : A \to Y$ al dominio $B \supset A$ una qualunque funzione $\tilde{f} : B \to Y$ tale che $\tilde{f}_{|A} = f$.*

Definizione 1.4 - *Si chiama* ***immagine di*** *A* ***secondo*** *f, o immagine di f, il sottoinsieme di Y*

$$f(A) = \{y \in Y \mid \exists x \in A \ :\ f(x) = y\} = \{f(x) \mid x \in A\}\,.$$

L'immagine di A è formata da tutti e soli gli elementi di $y \in Y$ che corrispondono ad elementi $x \in A$. Ad esempio, se $f : \mathbf{R} \to \mathbf{R}$ è la funzione $f(x) = x^2$, l'immagine $f(\mathbf{R})$ è $[0, +\infty[$. Se si lascia inalterata la legge $x \to x^2$, ma si sceglie un altro dominio come ad esempio $[0,1]$, si ottiene un'altra funzione, la $f_{|[0,1]}$, che ha per immagine l'intervallo $[0,1]$, manda cioè $[0,1]$ in se stesso.

Definizione 1.5 - *Una funzione $f : A \to Y$ si dice **surgettiva** se $f(A) = Y$, cioè se*

$$\forall y \in Y \quad \exists x \in A \quad : \quad f(x) = y\,.$$

Ogni funzione $f : A \to Y$ può essere resa surgettiva modificandone opportunamente il codominio, basta definirla come funzione $f : A \to f(A)$.

Esercizio 1.11 - *Dare la definizione di funzione non surgettiva.*

Definizione 1.6 - *Dato un insieme $B \subset Y$, si chiama **immagine inversa**, o **controimmagine**, di B secondo f il sottoinsieme di X definito da*

$$f^{-1}(B) = \{x \in A \mid f(x) \in B\}\,.$$

La controimmagine di B è formata da tutti e soli gli elementi di A che vengono trasformati in B. In particolare $f^{-1}(Y) = A$.

Esercizio 1.12 - *Dimostrare che se $S \subset Y$ e $S \cap f(A) = \emptyset$ allora $f^{-1}(S) = \emptyset$.*

Esercizio 1.13 - *Dimostrare che se $f : A \to Y$ e $S_1, S_2 \subset A$ allora*

$$f(S_1 \cup S_2) = f(S_1) \cup f(S_2) \;\; e \;\; f(S_1 \cap S_2) \subset f(S_1) \cap f(S_2)$$

e che se $S_1, S_2 \subset f(A)$ allora

$$f^{-1}(S_1 \cup S_2) = f^{-1}(S_1) \cup f^{-1}(S_2) \;\; e \;\; f^{-1}(S_1 \cap S_2) = f^{-1}(S_1) \cap f^{-1}(S_2)\,.$$

Perché la seconda è così diversa dalle altre? Basta costruire un controesempio da cui risulti evidente che l'inclusione $f(S_1 \cup S_2) \supset f(S_1) \cup f(S_2)$ è falsa. Se $f(x) = x^2$, con $x \in [-1, 1]$, si ha $f[-1, 0] = f[0, 1] = [0, 1]$, per cui $f[-1, 0] \cap f[0, 1] = [0, 1]$, mentre $f([-1, 0] \cap [0, 1]) = f\{0\} = \{0\}$. Ci si rende conto facilmente che l'inclusione diventa vera se f soddisfa la seguente definizione.

Definizione 1.7 - *Una funzione $f : A \to Y$ si dice **iniettiva** se ad elementi distinti di A associa elementi distinti di Y, cioè se*

$$\forall x_1, x_2 \in A \qquad x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)\,.$$

Altre definizioni equivalenti sono

$$\forall x_1, x_2 \in A \qquad f(x_1) = f(x_2) \Rightarrow x_1 = x_2\,,$$

oppure

$$\forall y \in f(A) \quad \exists_1 x \in A \quad : \quad f(x) = y\,.$$

Esercizio 1.14 - *Dare la definizione di funzione non iniettiva.*

Esercizio 1.15 - *Studiare le inclusioni tra $f(\complement S)$ e $\complement f(S)$ e tra $f^{-1}(\complement S)$ e $\complement f^{-1}(S)$.*

Esercizio 1.16 - *Dimostrare che $S \subset f^{-1}(f(S))$ per ogni $S \subset A$ e che $f(f^{-1}(S)) \subset S$ per ogni $S \subset Y$. Vale l'uguaglianza nella prima inclusione se e solo se f è iniettiva e vale nella seconda se e solo se f è surgettiva.*

Ogni funzione può essere resa iniettiva imponendo opportune restrizioni. Per esempio, la solita $f(x) = x^2$, $x \in \mathbf{R}$, non è iniettiva perché $f(x) = f(-x)$, ma le sue restrizioni agli intervalli $[0, +\infty[$ e $]-\infty, 0]$ lo sono.

Definizione 1.8 - *Una funzione $f : A \to Y$ si dice **bigettiva** se è sia iniettiva che surgettiva, cioè se*

$$\forall y \in Y \quad \exists_1 x \in A \quad : \quad f(x) = y\,.$$

In questo caso è possibile definire la funzione $g : Y \to A$ che associa ad ogni $y \in Y$ quell'unico elemento $x \in A$ tale che $f(x) = y$. Una sola funzione soddisfa questa proprietà e si chiama *funzione inversa* di f e si indica con f^{-1}. Ovviamente $(f^{-1})^{-1} = f$. Per il fatto di ammettere funzione inversa, una funzione bigettiva viene detta anche *invertibile*.

Siano $A \subset X$, $B \subset Y$, $f : A \to Y$ tale che $f(A) \subset B$ e $g : B \to Z$.

Definizione 1.9 - *La funzione $h : A \to Z$ definita da*

$$h(x) = g(f(x))\,, \qquad \forall x \in A\,,$$

*si chiama **prodotto di composizione** o **funzione composta** di f con g e si indica con $g \circ f$.*

In modo ovvio si può generalizzare la composizione al caso di più funzioni. Per esempio, se $f : A \to Y$, $g : B \to Z$ e $h : C \to W$, con la condizione di prima sull'inclusione tra domini e immagini, la loro composizione è la funzione $h \circ g \circ f : A \to W$ che associa ad ogni punto $x \in A$ il punto $h(g(f(x))) \in W$. Oppure si può ricondurre la definizione di composizione di tre o più funzioni al caso delle due funzioni dimostrando preliminarmente la proprietà associativa $h \circ (g \circ f) = (h \circ g) \circ f$. Riguardo alla proprietà commutativa, il fatto stesso che funzioni e domini coinvolti possano essere di natura completamente diversa ci basta per concludere che in generale non ha neanche senso. Ma anche quando avesse senso invertire il loro ordine, è facile costruire esempi dove $g \circ f \neq f \circ g$. Se ad esempio $f(x) = x + 1$ e $g(y) = y^2$, le due composizioni possibili sono $g(f(x)) = (x + 1)^2$ e $f(g(x)) = x^2 + 1$ che coincidono solo per $x = 0$, non sono la stessa funzione.

Si chiama *identità* o *funzione identica* su $A \subset X$ la funzione $i_A : A \to X$ definita da $i_A(x) = x$ per ogni $x \in A$. Ovviamente si ha $f \circ i_A = f$ e $i_Y \circ f = f$ per ogni $f : A \to Y$. Con l'identità si può ridefinire anche la restrizione: se $A \subset B \subset X$ e $f : B \to Y$ allora $f_{|A} = f \circ i_A$. Osserviamo infine che se $f : A \to Y$ è invertibile la sua inversa $f^{-1} : Y \to A$ è caratterizzata da $f^{-1} \circ f = i_A$ e $f \circ f^{-1} = i_Y$.

Esercizio 1.17 - *Siano $f : A \to Y$, $g : B \to Z$, $f(A) \subset B \subset Y$. Dimostrare che*

(a) se f e g sono iniettive allora $g \circ f$ è iniettiva. Viceversa se $g \circ f$ è iniettiva lo è anche f, ma se in più $f(A) = B$ allora è iniettiva anche g;

(b) se $f : A \to B \subset Y$ e $g : B \to C \subset Z$ sono surgettive allora anche $g \circ f : A \to C$ è surgettiva. Viceversa se $g \circ f : A \to C$ è surgettiva allora g è surgettiva, ma non è detto che lo sia anche f in quanto può ancora accadere che $f(A)$ sia contenuto strettamente in B;

(c) se f e g sono bigettive allora $g \circ f$ è bigettiva e $(g \circ f)^{-1} = f^{-1} \circ g^{-1}$.

Esempi

1.15 *Vediamo per quali $a, b \in \mathbf{R}$ la funzione $f : \mathbf{R} \to \mathbf{R}$ definita da $f(x) = ax + b$ è iniettiva o surgettiva. L'implicazione*

$$x_1 \neq x_2 \Rightarrow f(x_1) \neq f(x_2)$$

è equivalente a

$$f(x_1) = f(x_2) \Rightarrow x_1 = x_2$$

che sembra un pochino più maneggevole. Allora partiamo dall'uguaglianza

$$ax_1 + b = ax_2 + b$$

e vediamo che equivale a

$$a(x_1 - x_2) = 0\,.$$

Da questa si può dedurre $x_1 = x_2$, cioè che f è iniettiva, se $a \neq 0$, altrimenti è la funzione costante con valore b. È surgettiva negli stessi casi, infatti per ogni $y \in \mathbf{R}$ l'equazione $ax+b = y$ ammette come unica soluzione $x = (y-b)/a$, ma solo se $a \neq 0$. Per questo tipo di funzioni essere iniettiva equivale ad essere surgettiva, quindi anche bigettiva con inversa $f^{-1} : \mathbf{R} \to \mathbf{R}$ definita da

$$f^{-1}(y) = \frac{y-b}{a}\,.$$

1.16 *Siano $A, B \subset \mathbf{R}$ e $f : A \to B$ la funzione $f(x) = x^2$. Se $A = B = \mathbf{R}$ non è né iniettiva né surgettiva. Non è iniettiva perché un numero reale e il suo opposto hanno lo stesso quadrato e non è surgettiva perché nessun numero negativo è il quadrato di un numero reale. Definita invece come funzione $f : \mathbf{R} \to [0, +\infty[$ diventa surgettiva. Per renderla iniettiva basta restringerla ad un sottoinsieme che non contenga un numero reale e anche il suo opposto, ad esempio $f : [0, +\infty[\to \mathbf{R}$. Come funzione $f : [0, +\infty[\to [0, +\infty[$ è sia iniettiva che surgettiva, quindi bigettiva.*

Esercizio 1.18 - *Stabilire se la funzione $f : A \to B$, $f(x) = x^2$, è iniettiva e/o surgettiva nei seguenti casi:*
- $A = B = \mathbf{N}$,
- $A = B = \mathbf{Z}$,
- $A = B = \mathbf{Q} \cap [0, +\infty[$,
- $A = B = \mathbf{Q}$,
- $A = \mathbf{Q}$ e $B = \mathbf{R}$.

Esercizio 1.19 - *Stabilire se la funzione $f : A \to B$, $f(x) = \operatorname{sen} x$, è iniettiva e/o surgettiva nei seguenti casi:*
- $A = B = \mathbf{R}$,
- $A = \mathbf{R}$ e $B = [-1, 1]$,
- $A = \mathbf{Z}$ e $B = [-1, 1]$,
- $A = \mathbf{N}$ e $B = [-1, 1]$,
- $A = \mathbf{Q}$ e $B = [-1, 1]$,
- $A = \mathbf{Q}$ e $B = \mathbf{Q} \cap [-1, 1]$.

Capitolo 2

I numeri reali

Facciamo l'ipotesi che esista un insieme $\mathbf{R}$, l'insieme dei *numeri reali*, munito di certe strutture che in questa sede vogliamo trattare e discutere insieme ad alcune conseguenze importanti. Si tratta di un punto di vista assiomatico che permette poi di riconoscere, come sottoinsiemi particolari, $\mathbf{N}$, $\mathbf{Z}$ e $\mathbf{Q}$, rispettivamente l'insieme dei *numeri naturali*, dei *numeri interi* e dei *numeri razionali*.

2.1 Struttura algebrica

Un'*operazione* su un insieme X è un'applicazione $* : X \times X \to X$ con certe proprietà da stabilire volta per volta. Il *risultato* dell'operazione, $*(x, y)$, viene indicato con $x * y$ e la coppia $(X, *)$ prende il nome di *struttura algebrica*.

Le due operazioni fondamentali in $\mathbf{R}$ sono l'*addizione* e la *moltiplicazione* i cui risultati sono rispettivamente

$$(x, y) \to x + y = \textit{somma} \qquad \text{e} \qquad (x, y) \to x \cdot y = \textit{prodotto}\,.$$

Le loro proprietà sono caratteristiche di una struttura di *gruppo commutativo*, o *abeliano* dal matematico Abel, e le riportiamo qui di seguito.

- Il gruppo abeliano $(\mathbf{R}, +)$:

(+)1. $(x + y) + z = x + (y + z) \quad \forall x, y, z \in \mathbf{R}$ *proprietà associativa*

(+)2. $x + y = y + x \qquad \forall x, y \in \mathbf{R}$ *proprietà commutativa*

(+)3. $\exists\, 0 \in \mathbf{R} \ : \ x + 0 = x \quad \forall x \in \mathbf{R}$ *esistenza dell'elemento neutro*

(+)4. $\forall x \in \mathbf{R} \quad \exists - x \in \mathbf{R} \ : \ -x + x = 0$ *esistenza dell'opposto.*

La *differenza* di due numeri reali x e y può essere definita ricorrendo alla somma

$$x - y = x + (-y)\,.$$

Alcune conseguenze importanti sono

(+)a. *unicità dell'elemento neutro:* se 0 e $0'$ sono due elementi neutri allora

$$0' = 0' + 0 = 0\,;$$

(+)b. *unicità dell'opposto:* se r e s sono due opposti dello stesso $x \in \mathbf{R}$, cioè $x + r = 0$ e $x + s = 0$, allora

$$r = r + 0 = r + (x + s) = (r + x) + s = 0 + s = s\,;$$

(+)c. *ogni numero coincide con l'opposto del suo opposto:* essendo $-x$ l'opposto di x, dalla relazione $x + (-x) = 0$ si deduce anche che x è l'opposto di $-x$, cioè $x = -(-x)$.

Passiamo adesso alle proprietà della moltiplicazione che sono del tutto analoghe.
- Il gruppo abeliano $(\mathbf{R}-\{0\},\cdot)$:

$(\cdot)$1. $(x\cdot y)\cdot z = x\cdot(y\cdot z) \quad \forall x,y,z\in\mathbf{R}$ *proprietà associativa*

$(\cdot)$2. $x\cdot y = y\cdot x \qquad \forall x,y\in\mathbf{R}$ *proprietà commutativa*

$(\cdot)$3. $\exists\, 1\in\mathbf{R} \ :\ x\cdot 1 = x \quad \forall x\in\mathbf{R}$ *esistenza dell'elemento neutro*

$(\cdot)$4. $\forall x\in\mathbf{R}-\{0\} \quad \exists\, x^{-1}\in\mathbf{R} \ :\ x^{-1}x = 1$ *esistenza dell'inverso.*

L'inverso, detto anche *reciproco*, di un numero $x\neq 0$ si indica anche con $1/x$ e se $y\neq 0$ è possibile definire l'operazione di *divisione*, o *rapporto*, tra x e y ponendo

$$\frac{x}{y} = x\cdot\frac{1}{y}\,.$$

Esercizio 2.1 - *Verificare che dalle* $(\cdot)$1.—$(\cdot)$4. *discendono conseguenze analoghe a quelle ottenute per la somma.*

Vi è un'ulteriore proprietà che lega le due strutture, di *gruppo additivo* $(\mathbf{R},+)$ e di *gruppo moltiplicativo* $(\mathbf{R}-\{0\},\cdot)$, con la quale $\mathbf{R}$ acquista la struttura di *corpo commutativo* o *campo.*
- Il campo $(\mathbf{R},+,\cdot)$:

$(+,\cdot)$1. $(\mathbf{R},+)$ *gruppo abeliano*

$(+,\cdot)$2. $(\mathbf{R}-\{0\},\cdot)$ *gruppo abeliano*

$(+,\cdot)$3. $(x+y)\cdot z = x\cdot z + y\cdot z \quad \forall x,y,z\in\mathbf{R}$ *proprietà distributiva.*

Alcune conseguenze importanti sono

$(+,\cdot)$a. $x\cdot 0 = 0$ per ogni $x\in\mathbf{R}$, infatti

$$x\cdot 0 = x\cdot(0+0) = x\cdot 0 + x\cdot 0 \ \Rightarrow\ x\cdot 0 = 0\,,$$

in particolare $0\neq 1$ altrimenti si avrebbe $x = x\cdot 1 = x\cdot 0 = 0$ per ogni $x\in\mathbf{R}$ e l'intero insieme $\mathbf{R}$ si ridurrebbe ad un solo elemento;

$(+,\cdot)$b. *legge di annullamento del prodotto:*

$$x\cdot y = 0 \ \Rightarrow\ x = 0 \vee y = 0\,,$$

(in particolare $x^2 = 0 \Rightarrow x = 0$) infatti, se ad esempio $x\neq 0$, si ricava, dividendo per x, che $y = x^{-1}0 = 0$;

$(+,\cdot)$c. $x\cdot(-y) = -xy$ per ogni $x,y\in\mathbf{R}$, infatti

$$x\cdot(-y) + x\cdot y = x\cdot(-y+y) = 0 \ \Rightarrow\ x\cdot(-y) = -x\cdot y\,.$$

Le regole del calcolo discendono tutte dai nove assiomi relativi alla struttura di campo. Da ora in poi il puntino del prodotto verrà omesso.

Esercizio 2.2 - *Dimostrare che se* $xy = 0$ *per ogni* $y\in\mathbf{R}$ *allora* $x = 0$.

Più in generale si ha

$$x_1y_1 + x_2y_2 = 0 \qquad \forall y_1,y_2\in\mathbf{R} \ \Rightarrow\ x_1 = x_2 = 0\,.$$

E ancora più in generale

$$x_1y_1 + x_2y_2 + \ldots + x_ny_n = 0 \quad \forall y_i\in\mathbf{R} \ \Rightarrow\ x_i = 0 \quad \forall i = 1,\ldots,n\,.$$

2.2 Ordinamento

Su $\mathbf{R}$ è definita una relazione d'ordine totale (v. § 1.4). Ma ciò che fa di $\mathbf{R}$ un *campo ordinato*, $(\mathbf{R}, +, \cdot, \leqslant)$, sono le due proprietà

$(+, \leqslant)1.$ $\quad x \leqslant y \Rightarrow x + z \leqslant y + z \quad \forall x, y, z \in \mathbf{R}$

$(\,\cdot\,, \leqslant)2.$ $\quad x \leqslant y \wedge z \geqslant 0 \Rightarrow xz \leqslant yz \quad \forall x, y, z \in \mathbf{R}$

che rendono l'ordinamento compatibile con le operazioni. Di conseguenza valgono le seguenti:

$(+, \leqslant)$a. $x \geqslant 0 \Rightarrow -x \leqslant 0$. Infatti $x \geqslant 0$ e la $(+, \leqslant)$ implicano

$$0 = x + (-x) \geqslant -x\,;$$

$(\,\cdot\,, \leqslant)$b. $x \leqslant y \wedge z \leqslant 0 \Rightarrow xz \geqslant yz$. Infatti per la $(\cdot, \leqslant)$ si ha

$$-xz = x(-z) \leqslant y(-z) = -yz\,,$$

a questo punto basta aggiungere ai due membri $xz + yz$.

Osserviamo infine che $x^2 \geqslant 0$ per ogni $x \in \mathbf{R}$. Infatti se $x \geqslant 0$ allora, moltiplicando per x, $x^2 = xx \geqslant 0$; se invece $x < 0$, moltiplicando questa per $-x > 0$ si ottiene $-x^2 = x(-x) < 0$, cioè $x^2 > 0$. In particolare $1 = 1^2 \geqslant 0$, ma essendo $1 \neq 0$, si ha $1 > 0$.

Il *segno* è la funzione $\operatorname{sign} : \mathbf{R} \to \mathbf{R}$ definita da

$$\operatorname{sign} x = \begin{cases} -1 & \text{se } x < 0 \\ 0 & \text{se } x = 0 \\ 1 & \text{se } x > 0\,. \end{cases}$$

Se $x > 0$ [< 0] diciamo che è *positivo* [*negativo*], se $x \geqslant 0$ [$\leqslant 0$] diciamo che è *non negativo* [*non positivo*]. Infine x e y si dicono *concordi* [*discordi*] se

$$\operatorname{sign} x = \operatorname{sign} y \qquad [\operatorname{sign} x = -\operatorname{sign} y]$$

che equivale a dire $xy > 0$ [$xy < 0$]. In genere però x e y vengono detti concordi [discordi] anche se $xy \geqslant 0$ [$xy \leqslant 0$].

Esercizio 2.3 - *Verificare che* $\operatorname{sign} x = \operatorname{sign}(1/x)$ *per ogni* $x \in \mathbf{R} - \{0\}$ *e che, se* $x, y \in \mathbf{R} - \{0\}$*, allora* $x < y$ *equivale a* $1/y < 1/x$ *se* x *e* y *sono concordi, a* $1/y > 1/x$ *se sono discordi.*

Esercizio 2.4 - *Per* $x, y \in \mathbf{R}$ *si ha*

$$x^2 + y^2 = 0 \Leftrightarrow x = y = 0\,.$$

L'implicazione "$\Leftarrow$" è banale. Per verificare l'altra basta osservare che

$$0 \leqslant x^2 = -y^2 \leqslant 0\,.$$

Esercizio 2.5 - *Se* $xy \geqslant 0$ *per ogni* $y \in \mathbf{R}$ *allora* $x = 0$.

Esercizio 2.6 - *Se* $x_1y_1 + x_2y_2 \geqslant 0$ *per ogni* $(y_1, y_2) \in \mathbf{R}^2$ *allora*

$$(x_1, x_2) = (0, 0)\,.$$

Esercizio 2.7 - *Se* $x \geqslant 0$ *e* $x \leqslant y$ *per ogni* $y > 0$ *allora* $x = 0$.

Esercizio 2.8 - *Per ogni* $a, x, y \in \mathbf{R}$*,* $a \neq 0$*, si ha*

$$a^2x^2 + \frac{y^2}{a^2} \geqslant 2xy\,.$$

Basta osservare che $(ax - y/a)^2 \geqslant 0$.

Esercizio 2.9 - *Siano $x, y \in \mathbf{R}$ tali che $x \neq 0$, $y \neq 0$ e $x + y \neq 0$. Allora si ha*

$$\frac{1}{x} + \frac{1}{y} \neq \frac{1}{x+y}\,.$$

Se valesse l'uguaglianza, sarebbe equivalente a

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

ma ciò è assurdo perché si otterrebbe

$$0 = (x+y)^2 - xy = x^2 + y^2 + xy = x^2 + y^2 + (x+y)^2 > 0\,.$$

Usando la relazione d'ordine possiamo definire i seguenti sottoinsiemi di $\mathbf{R}$ che chiameremo *intervalli* di estremi $a, b \in \mathbf{R}$

$[a,b] = \{x \in \mathbf{R} \mid a \leqslant x \leqslant b\}$ *intervallo chiuso*

$]a,b[= \{x \in \mathbf{R} \mid a < x < b\}$ *intervallo aperto*

$[a,b[= \{x \in \mathbf{R} \mid a \leqslant x < b\}$ *intervallo aperto a destra*

$]a,b] = \{x \in \mathbf{R} \mid a < x \leqslant b\}$ *intervallo aperto a sinistra.*

Se $a = b$ il primo si riduce a un punto e gli altri all'insieme vuoto (anche $\emptyset$ è un intervallo!) mentre se $a > b$ sono tutti vuoti. Se $a \leqslant b$ assumiamo il numero reale non negativo $b - a$ come *lunghezza* o *misura* dell'intervallo. Il *punto medio*, o *centro*, di un intervallo di estremi a e b è il numero $(a+b)/2$. Un intervallo di centro $x_0 \in \mathbf{R}$ e raggio $r \geqslant 0$ ha lunghezza $2r$ e ammette x_0 come punto medio, in questo caso i suoi estremi sono $a = x_0 - r$ e $b = x_0 + r$.

Oltre ai precedenti, che vengono detti *limitati*, gli intervalli *non limitati* sono $\mathbf{R} =]-\infty, +\infty[$ e le *semirette*

$$\begin{matrix}]-\infty, a] = \{x \in \mathbf{R} \mid x \leqslant a\} \\ [a, +\infty[= \{x \in \mathbf{R} \mid x \geqslant a\} \end{matrix} \ \textit{chiuse} \ \text{e} \ \begin{matrix}]-\infty, a[= \{x \in \mathbf{R} \mid x < a\} \\ ]a, +\infty[= \{x \in \mathbf{R} \mid x > a\} \end{matrix} \ \textit{aperte}$$

dove $a \in \mathbf{R}$.

Si chiama *combinazione convessa* dei numeri $a, b \in \mathbf{R}$ la funzione $\varphi : [0,1] \to [a,b]$ definita da

$$\varphi(\lambda) = a + \lambda(b-a) = \lambda b + (1-\lambda)a \qquad \forall \lambda \in [0,1]\,. \tag{2.1}$$

Esercizio 2.10 - *Dimostrare che φ è un'applicazione bigettiva da $[0,1]$ in $[a,b]$.*

Esercizio 2.11 - *Nell'intervallo di tempo $[0,T]$ un punto P in moto rettilineo uniforme percorre l'intervallo $[a,b]$ di una retta nello spazio. Determinare la legge oraria e la velocità di P nei due moti da a verso b e da b verso a e verificare che in ogni caso si tratta di un'applicazione bigettiva $[0,T]$ su $[a,b]$ del tipo della (2.1).*

Esercizio 2.12 - *Riconoscere i seguenti sottoinsiemi di $\mathbf{R}$*

$$\bigcup_{n=2}^{\infty} \left]\frac{1}{n}, 1 - \frac{1}{n}\right[\,, \ \bigcup_{n=2}^{\infty} \left[\frac{1}{n}, 1 - \frac{1}{n}\right]\,, \ \bigcap_{n=1}^{\infty} \left]-\frac{1}{n}, 1 + \frac{1}{n}\right[\,, \ \bigcap_{n=1}^{\infty} \left[-\frac{1}{n}, 1 + \frac{1}{n}\right]\,,$$

dove l'indice n varia nell'insieme $\mathbf{N}$ dei numeri naturali.

2.3 Completezza

Vogliamo adesso studiare un altro aspetto dell'insieme $\mathbf{R}$ che riguarda come sono distribuiti i numeri. Si tratta di formalizzare l'idea di un aggregato di punti che, essendo privo di spazi vuoti, buchi e lacune, va pensato come un "continuo". Questa proprietà può essere stabilita mediante un opportuno *postulato di continuità* di cui vi sono diverse versioni equivalenti, tra cui quella ben nota delle sezioni di Dedekind. Un modo particolarmete semplice di trattare la continuità ci sembra quello che si basa sull'*assioma di completezza*. Premettiamo la definizione di insiemi separati.

Definizione 2.1 - *Due sottoinsiemi A e B di $\mathbf{R}$ si dicono separati se*

$$x \leqslant y \qquad \forall x \in A\,,\ \forall y \in B\,.$$

In $\mathbf{R}$ vale il seguente

Assioma 2.2 (di completezza) - *Se $A, B \subset \mathbf{R}$ sono separati e non vuoti allora*

$$\exists \xi \in \mathbf{R}\ :\quad x \leqslant \xi \leqslant y \qquad \forall x \in A \text{ e } \forall y \in B\,. \tag{2.2}$$

*Ogni $\xi \in \mathbf{R}$ che soddisfa la (2.2) viene detto **elemento separatore** dei due insiemi A e B.*

Sarebbe errato sostituire, nell'assioma di completezza, gli insiemi A e B con due numeri a e b affermando:

$$a \leqslant b \Rightarrow \exists c \in \mathbf{R}\ :\quad a \leqslant c \leqslant b$$

che è infatti meno restrittiva. Ad esempio $\mathbf{Q}$ soddisfa questa proprietà, ma non l'assioma di completezza: tra due numeri razionali c'è sempre un terzo numero razionale, ad esempio la media, ma l'insieme A dei razionali minori di $\sqrt{2}$ e l'insieme B di quelli maggiori, che sono sottoinsiemi separati di $\mathbf{Q}$, hanno come unico elemento separatore $\sqrt{2}$ che non appartiene a $\mathbf{Q}$.

Nell'assioma di completezza l'elemento separatore non è in generale unico.

Definizione 2.3 - *Due classi separate di numeri A e B si dicono **contigue** se*

$$\forall \varepsilon > 0\ \exists x \in A\ \exists y \in B\ :\quad y - x < \varepsilon\,. \tag{2.3}$$

È questo il caso in cui l'elemento separatore è unico, infatti se A e B sono contigue e vi fossero due elementi separatori distinti, $\xi < \eta$, si troverebbe subito un $\varepsilon > 0$ per il quale la (2.3) non vale, basta scegliere $\varepsilon < \eta - \xi$ e si avrebbe

$$y - x \geqslant \eta - \xi > \varepsilon \qquad \forall x \in A\,,\ \forall y \in B\,.$$

Viceversa, se le classi separate A e B hanno un unico elemento separatore $\xi \in \mathbf{R}$ allora devono essere contigue, altrimenti per un $\varepsilon > 0$ sufficientemente piccolo tutti i numeri reali dell'intervallo $]\xi - \varepsilon/2, \xi + \varepsilon/2[$ sarebbero elementi separatori.

Le due classi

$$\begin{array}{llcll}
A: & x_1 = 0,1 & \qquad & B: & y_1 = 0,2 \\
 & x_2 = 0,12 & & & y_2 = 0,13 \\
 & x_3 = 0,121 & & & y_3 = 0,122 \\
 & x_4 = 0,1211 & & & y_4 = 0,1212 \\
 & x_5 = 0,12112 & & & y_5 = 0,12113 \\
 & x_6 = 0,121122 & & & y_6 = 0,121123 \\
 & \quad\vdots & & & \quad\vdots
\end{array}$$

sono contigue ed ammettono come unico elemento separatore il numero reale

$$\xi = 0,12112211122211112222\dots .$$

Che sono contigue lo si vede dal fatto che $y_i - x_i = 10^{-i}$, numero che può essere reso piccolo quanto si vuole, minore cioè di un $\varepsilon > 0$ arbitrario, pur di prendere i abbastanza grande.

L'intersezione di due insiemi separati A e B può essere vuota o contenere al più un solo elemento. È vuota quando vi sono più elementi separatori, e quindi infiniti, (A e B non contigui) e quando ve ne è uno solo (A e B contigui) che non appartiene né ad A, né a B o appartiene ad uno solo dei due. Altrimenti è non vuota, ma non può contenere più di un elemento (A e B contigui) che è anche l'unico elemento separatore.

2.4 Sottoinsiemi limitati di numeri reali

Usando l'ordinamento naturale di $\mathbf{R}$ possiamo definire le nozioni di insieme limitato, di massimo e di minimo. Vi sono però degli insiemi limitati che non ammettono massimo o minimo. Nasce allora l'esigenza di introdurre concetti più generali, quelli di estremo superiore e inferiore, che richiedono anche la completezza.

Definizione 2.4 - *Dato un insieme $A \subset \mathbf{R}$, diciamo che $k \in \mathbf{R}$ è un* ***maggiorante*** *[minorante] di A se $x \leqslant k$ $[x \geqslant k]$ per ogni $x \in A$.*

Un maggiorante [minorante] viene detto anche *limitazione superiore* [*inferiore*].

Definizione 2.5 - *L'insieme $A \subset \mathbf{R}$ viene detto* ***limitato superiormente*** *[**inferiormente**] se l'insieme dei maggioranti [minoranti] è non vuoto. Viene detto* ***limitato*** *se è limitato sia superiormente che inferiormente.*

In particolare, gli intervalli limitati introdotti nel § 2.2 sono limitati anche secondo la Definizione 2.5. In essa si definisce limitato un insieme A arbitrario che sia interamente contenuto in un intervallo limitato.

Rivediamo le definizioni date in termini simbolici:

- A è limitato superiormente [inferiormente] se e solo se
$$\exists k \in \mathbf{R} \;:\; x \leqslant k \; [x \geqslant k] \; \forall x \in A\,,$$
- A è limitato se e solo se $\exists k_1, k_2 \in \mathbf{R} \;:\; k_1 \leqslant x \leqslant k_2 \; \forall x \in A$.

Per negare che un insieme sia limitato, in un senso o nell'altro, basta usare le regole della negazione viste nel Cap. 1, precisamente

- A è non limitato superiormente [inferiormente] se e solo se
$$\forall k \in \mathbf{R} \; \exists x \in A \;:\; x > k \; [x < k]\,,$$
- A è non limitato se e solo se $\forall k_1, k_2 \in \mathbf{R} \; \exists x, y \in A \;:\; x < k_1 \, \mathrm{e} \, y > k_2$.

Ad esempio $\mathbf{R}$ è non limitato, le semirette del tipo $[a, +\infty[$ non lo sono superiormente e le rette del tipo $]-\infty, a]$ non lo sono inferiormente. Guardando alle negazioni delle definizioni precedenti, un insieme non limitato superiormente [inferiormente] non ha maggioranti [minoranti].

Definiamo adesso il massimo e il minimo di un insieme

Definizione 2.6 - *Se esiste un numero reale x_0 tale che*

$$x_0 \in A \quad e \quad x \leqslant x_0\ [x \geqslant x_0]\ \forall x \in A$$

allora diciamo che x_0 è il ***massimo*** *[**minimo**] di A e si scrive $x_0 = \max A$ [$x_0 = \min A$].*

Esercizio 2.13 - *Dimostrare che il massimo [minimo], se esiste, è unico.*

Esercizio 2.14 - *Un sottoinsieme finito di $\mathbf{R}$, cioè del tipo $A = \{a_1, a_2, \dots, a_n\}$, ha massimo e minimo.*

Esercizio 2.15 - *Un intervallo chiuso a destra [sinistra] ha massimo [minimo].*

Esercizio 2.16 - *Verificare che l'insieme $\{x+1/x \mid x \in \mathbf{R}\,,\ x > 0\}$ non è limitato superiormente ma ammette minimo, calcolarlo.*

Esercizio 2.17 - *Fra tutti i quadrilateri di dato perimetro, qual è quello di area massima?*

Esercizio 2.18 - *Uno dei raggi di luce emessi da una sorgente posta nel punto A colpisce B dopo una riflessione su uno specchio. Dedurre l'uguaglianza tra gli angoli di incidenza e di riflessione assumendo che il tempo impiegato sia minimo.*

Teorema 2.7 - *L'insieme dei maggioranti di un insieme $A \neq \emptyset$ limitato superiormente ha minimo.*

Dimostrazione. Indichiamo con M l'insieme

$$M = \{k \in \mathbf{R} \mid k \geqslant x\ \forall x \in A\}$$

di tutti i maggioranti di A. Per costruzione A e M sono separati e per l'assioma di completezza esiste $\xi \in \mathbf{R}$ tale che

$$x \leqslant \xi \leqslant k \qquad \forall x \in A \text{ e } \forall k \in M\,.$$

La disuguaglianza di sinistra dice che ξ è un maggiorante di A e quindi $\xi \in M$, quella di destra che tutti i maggioranti di A superano ξ. Dunque ξ è il minimo di M.

□

Analogamente si dimostra che l'insieme dei minoranti di A ha massimo se A è limitato inferiormente. Osserviamo che il numero reale $\xi \in M$ del Teorema 2.7, la cui esistenza è garantita dall'assioma di completezza, non è detto che appartenga anche ad A, a meno che A non abbia il massimo, in questo caso $\xi = \min M = \max A$ e $M \cap A = \{\xi\}$.

Definizione 2.8 - *Se A è limitato superiormente [inferiormente] il minimo dei maggioranti [il massimo dei minoranti] si chiama* ***estremo superiore*** *[**estremo inferiore**] e si indica con*

$$\sup A \qquad\qquad [\inf A]\,.$$

Definiamo *retta reale estesa* l'unione tra l'insieme dei numeri reali e due nuovi elementi che indichiamo con $+\infty$ e $-\infty$:

$$\overline{\mathbf{R}} = \mathbf{R} \cup \{-\infty, +\infty\}$$

assumendo che sia munito della stessa relazione d'ordine di $\mathbf{R}$ con l'estensione

$$-\infty < x < +\infty \qquad \forall x \in \mathbf{R}\,.$$

Questo nuovo ordinamento non può rimanere compatibile con alcuna estensione a $\overline{\mathbf{R}}$ delle operazioni algebriche. In altre parole, simboli come $+\infty + (-\infty)$, $x + \infty$, $1/+\infty$ ecc., non hanno alcun significato, non indicano alcun elemento di $\overline{\mathbf{R}}$, e ammesso che si possa attribuire loro qualche significato, $\overline{\mathbf{R}}$ non sarà un corpo ordinato.

Definizione 2.9 - *Se A è non limitato superiormente [inferiormente] si pone*

$$\sup A = +\infty \qquad [\inf A = -\infty]\,.$$

Esercizio 2.19 - *Dimostrare che* $\inf A \leqslant \sup A$ *per ogni* $A \subset \mathbf{R}$.

Esercizio 2.20 - *Dimostrare che se* $A \subset B \subset \mathbf{R}$ $\inf B \leqslant \inf A \leqslant \sup A \leqslant \sup B$.

Esercizio 2.21 - *Dimostrare che A e B sono contigue se e solo se* $\sup A = \inf B$.

Diamo adesso una caratterizzazione degli estremi superiore e inferiore molto utile nelle applicazioni.

Teorema 2.10 - *Se $A \subset \mathbf{R}$ è non vuoto e limitato superiormente [inferiormente],* $L = \sup A$ $[L = \inf A]$ *se e solo se valgono insieme le condizioni*

(sup)1. $x \leqslant L \quad \forall x \in A\,,$ (inf)1. $x \geqslant L \quad \forall x \in A\,,$

(sup)2. $\forall \varepsilon > 0\ \exists x \in A\ :\ x > L - \varepsilon\,,$ (inf)2. $\forall \varepsilon > 0\ \exists x \in A\ :\ x < L + \varepsilon\,.$

L'analoga condizione per un insieme non limitato superiormente coincide con quella data nella negazione della Definizione 2.5.

Dimostrazione. Nella (sup)1. si afferma che L è un maggiorante di A. Nella (sup)2. si nega che $L - \varepsilon$ sia un maggiorante di A qualunque sia $\varepsilon > 0$, quindi L è il minimo. Le due condizioni insieme equivalgono ad affermare che L è il minimo dei maggioranti. Il ragionamento è del tutto analogo per l'estremo inferiore.

□

2.5 Struttura metrica

La funzione $|\cdot| : \mathbf{R} \to \mathbf{R}$ definita da

$$|x| = \begin{cases} x & \text{se } x \geqslant 0 \\ -x & \text{se } x < 0 \end{cases}$$

si chiama *valore assoluto* o *modulo*.

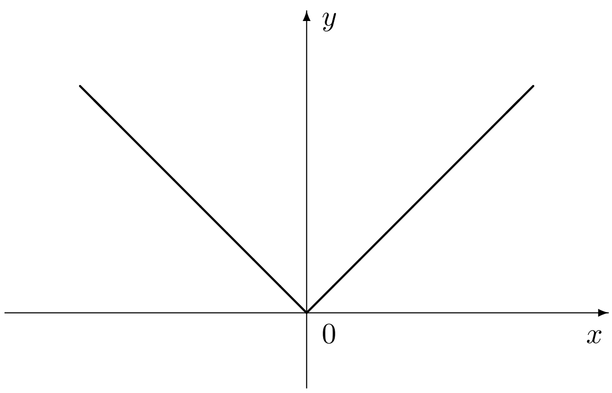


Figura 2.1: Il grafico della funzione $x \to |x|$.

Affermazioni come "$|x|$ è il numero x privato del segno", oppure "$|x| = \pm x$" non hanno alcun significato. Altre definizioni equivalenti sono

$$|x| = \max\{x, -x\}\,, \qquad |x| = x \operatorname{sign} x\,.$$

Di verifica immediata sono le seguenti proprietà:

$|1|$. $|x| \geqslant 0 \quad \forall x \in \mathbf{R}$ e $|x| = 0 \Leftrightarrow x = 0$,

$|2|$. $|xy| = |x||y| \quad \forall x, y \in \mathbf{R}$,

$|3|$. $\forall a \in \mathbf{R}\,,\ a \geqslant 0\,,\ \forall x \in \mathbf{R} \quad |x| \leqslant a \Leftrightarrow -a \leqslant x \leqslant a \quad$ e $\quad |x| \geqslant a \Leftrightarrow x \leqslant -a \vee x \geqslant a$,

$|4|$. $|x| = |-x| \quad \forall x \in \mathbf{R}$,

$|5|$. $x \leqslant |x| \quad \forall x \in \mathbf{R}$,

$|6|$. $|x|^2 = x^2 \quad \forall x \in \mathbf{R}$,

$|7|$. $x^2 \leqslant y^2 \Leftrightarrow |x| \leqslant |y| \quad \forall x, y \in \mathbf{R}$,

$|8|$. $\frac{1}{2}\,(x + y - |x - y|) = \min\{x, y\}\,, \qquad \frac{1}{2}\,(x + y + |x - y|) = \max\{x, y\} \quad \forall x, y \in \mathbf{R}$.

Dimostriamo le seguenti

$|9|$. $|x + y| \leqslant |x| + |y| \quad \forall x, y \in \mathbf{R}$,

$|10|$. $||x| - |y|| \leqslant |x - y| \quad \forall x, y \in \mathbf{R}$.

$|9|$. Per la $|6|$ basta dimostrare che $(x + y)^2 \leqslant (|x| + |y|)^2$. Si ha

$$(x + y)^2 = x^2 + 2xy + y^2 \leqslant |x|^2 + 2|x||y| + |y|^2 = (|x| + |y|)^2\,.$$

$|10|$. Per la $|3|$ basta dimostrare che $-|x - y| \leqslant |x| - |y| \leqslant |x - y|$. Dalla $|9|$

$$|x| = |x - y + y| \leqslant |x - y| + |y|\,,$$

da cui

$$|x| - |y| \leqslant |x - y|$$

e scambiando x con y si ottiene l'altra disuguaglianza.

Esercizio 2.22 - *$A \subset \mathbf{R}$ è limitato se e solo se*

$$\exists k \in \mathbf{R}\,,\ k \geqslant 0\,,\ :\ |x| \leqslant k\ \forall x \in A\,.$$

Una *distanza*, o *metrica*, su un insieme X è un'applicazione $d : X \times X \to \mathbf{R}$ che soddisfa le seguenti proprietà:

$(d)1$. $d(x, y) \geqslant 0\ \forall x, y \in X \quad$ e $\quad d(x, y) = 0 \Leftrightarrow x = y$ — *positività della distanza*

$(d)2$. $d(x, y) = d(y, x)\ \forall x, y \in X$ — *proprietà simmetrica*

$(d)3$. $d(x, y) \leqslant d(x, z) + d(z, y)\ \forall x, y, z \in X$ — *disuguaglianza triangolare.*

L'insieme X munito di questa nuova struttura si chiama *spazio metrico* e si indica con (X, d). Ogni sottoinsieme A di X è a sua volta uno spazio metrico, e viene detto *sottospazio* di X, con la distanza $d_{|A\times A}$, cioè con la stessa metrica di X ristretta ad A, si dice anche "con la metrica *indotta* da X".

In $\mathbf{R}$ quella più comune, e per noi la più utile, è la *distanza euclidea*

$$d(x, y) = |x - y| \qquad \forall x, y \in \mathbf{R}\,. \tag{2.4}$$

Esercizio 2.23 - *Dimostrare che la* (2.4) *soddisfa le proprietà* $(d)1-(d)3$ *e quindi definisce effettivamente una metrica su* $\mathbf{R}$.

Esercizio 2.24 - *Dimostrare che anche la distanza euclidea in* $\mathbf{R}^2$

$$d(x,y) = \sqrt{(x_1-y_1)^2+(x_2-y_2)^2} \qquad \forall x=(x_1,x_2), y=(y_1,y_2) \in \mathbf{R}^2$$

soddisfa le proprietà $(d)1-(d)3$.

Di distanze se ne possono definire tante, le più diverse, ognuna delle quali conferisce ad $\mathbf{R}$, o all'insieme in questione, una struttura metrica diversa. Alcune sono tra loro *simili*, come d e kd con $k>0$, ad esempio $d_k(x,y)=k|x-y|$ rimanendo nel caso euclideo. Altre possono essere perfino un po' bizzarre come la distanza *atomica* del seguente esercizio.

Esercizio 2.25 - *Dimostrare che la funzione su* X^2

$$d(x,y) = \begin{cases} 0 & se\, x=y \\ 1 & se\, x \neq y\,, \end{cases} \tag{2.5}$$

con $x,y\in X$, *è una distanza su* X.

Vi sono distanze "curvilinee", cioè distanze tra punti di curve o di superfici la cui misura dipende dalla loro forma. Talvolta anche il tempo impiegato durante uno spostamento e costi di vario tipo possono essere usati come misure di distanze (si pensi al traffico stradale che in un certo senso influisce sulle distanze).

Esercizio 2.26 - *Dimostrare che la distanza euclidea, su* $\mathbf{R}$ *o su* $\mathbf{R}^2$, *è* ***invariante per traslazioni*** *nel senso che* $d(x,y)=d(x+\tau,y+\tau)$ *per ogni* $x,y,\tau\in\mathbf{R},\mathbf{R}^2$. *È l'unica distanza con questa proprietà?*

Altre distanze variano da una "zona" all'altra dell'insieme. Per fare un esempio, ne possiamo costruire una legata alla nostra esperienza psicologica. Quando pensiamo al passato la nostra mente adotta distanze diverse: riguardo ai giorni appena trascorsi possiamo distinguere così chiaramente gli eventi accaduti da collocarli con buona approssimazione a distanza (temporale) reale; se pensiamo agli anni della nostra vita, o ai primi anni della nostra vita, più gli eventi sono lontani, più le distanze ci sembrano accorciate fino a non poter distinguere un giorno, un mese, un anno dall'altro. Se cerchiamo di collocare al loro posto gli eventi storici arriviamo a fare confusione addirittura tra secoli, millenni e ere geologiche. Sul futuro c'è solo il buio, ma supponiamo che anche pensando al futuro succeda la stessa cosa. Disponiamo allora gli eventi secondo la loro distanza reale sulla retta, l'asse x, dal passato più lontano, $-\infty$, al futuro $+\infty$ e facciamo corrispondere ad ogni evento, ad ogni punto x, la sua collocazione $f(x)$ nella nostra mente, che essendo limitata non distingue bene eventi lontani. Se quella da noi percepita è la distanza $|f(x_1)-f(x_2)|$ tra due eventi x_1 e x_2, un tipo di funzione che schematizza abbastanza bene questa situazione potrebbe avere l'andamento illustrato in figura.
Secondo questo modello la distanza tra due eventi x_1 e x_2 che la nostra mente percepisce non è quella usuale $|x_1-x_2|$, ma la *distanza indotta* da f

$$d_f(x_1,x_2)=|f(x_1)-f(x_2)|\,. \tag{2.6}$$

Si noti che f deve essere iniettiva (perché?). Più in generale, se X è un insieme, (Y,d) uno spazio metrico e $f:X\to Y$ una funzione iniettiva, si può definire la distanza indotta su X dalla f ponendo

$$d_f(x_1,x_2)=d(f(x_1),f(x_2)) \qquad \forall x_1,x_2\in X\,. \tag{2.7}$$

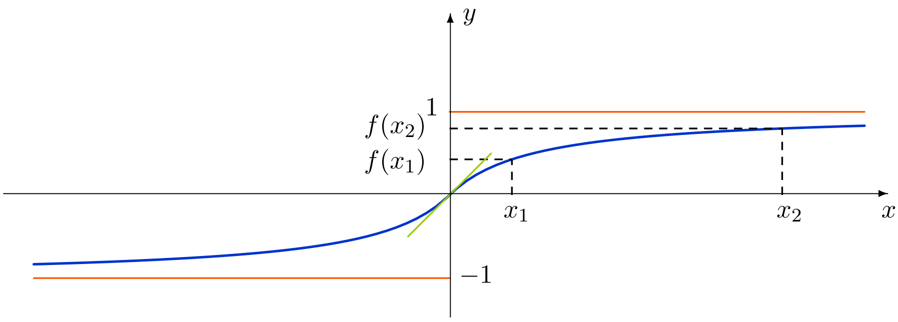


Figura 2.2: Il grafico della funzione $x \to \dfrac{x}{1+|x|}$.

Esercizio 2.27 - *Con riferimento alla (2.6), dimostrare che $(\mathbf{R}, d_f)$ è uno spazio metrico limitato con diametro pari a 2.*

Prolunghiamo la f ad una nuova funzione $\bar{f} : \overline{\mathbf{R}} \to \mathbf{R}$ ponendo

$$\bar{f}(x) = \begin{cases} -1 & \text{se } x = -\infty \\ f(x) & \text{se } x \in \mathbf{R} \\ 1 & \text{se } x = +\infty\,. \end{cases}$$

Definizione 2.11 - *Si chiama* ***diametro*** *dello spazio metrico X l'elemento di $\overline{\mathbf{R}}$*

$$\operatorname{diam} X = \sup_{x,y \in X} d(x,y)\,.$$

Diciamo che X è ***limitato*** *se* $\operatorname{diam} X < +\infty$.

Esercizio 2.28 - *Dimostrare che $(\mathbf{R}, d_f)$ e $(\overline{\mathbf{R}}, d_{\bar{f}})$ sono spazi metrici limitati con diametro pari a 2 e che nel secondo l'estremo superiore è un massimo.*

In uno spazio metrico X è possibile dare un senso alla nozione di "vicinanza" tra punti ricorrendo alla nozione di *palla*. La *palla di centro x_0 e raggio $r > 0$* è l'insieme

$$B_r(x_0) = \{x \in X \mid d(x, x_0) < r\}\,.$$

Possiamo ad esempio stabilire che i punti x "vicini" ad un certo x_0 sono quelli che distano da x_0 meno di r e che appartengono quindi a $B_r(x_0)$. Naturalmente il tipo di distanza scelta è determinante per la forma delle palle, la cui estensione dipenderà dal raggio r.

Esercizio 2.29 - *In $\mathbf{R}$ determinare $B_r(x_0)$ rispetto alla distanza euclidea e alla distanza (2.6).*

Esercizio 2.30 - *In $\overline{\mathbf{R}}$ determinare $B_r(x_0)$ rispetto alla distanza (2.7).*

Esercizio 2.31 - *Determinare $B_r(x_0)$ in un insieme X con la distanza atomica (2.5) e spiegare il motivo di questo nome.*

Definizione 2.12 - *In uno spazio metrico X si chiama* ***intorno*** *del punto $x_0 \in X$ ogni insieme $U \in X$ contenente almeno una palla $B_r(x_0)$ di centro x_0 e raggio $r > 0$.*

In particolare anche $B_r(x_0)$ è un intorno di x_0. Indicheremo con $\mathscr{I}(x_0)$ la totalità degli intorni di x_0. Osserviamo che se A è un sottospazio metrico di X gli intorni di x_0 in A sono tutti e soli gli insiemi $U \cap A$ con $U \in \mathscr{I}(x_0)$.

2.6 I numeri naturali

Per costruire con il dovuto rigore i vari insiemi numerici due sono le vie possibili. Quella che stiamo percorrendo è iniziata con la presentazione assiomatica dei numeri reali, nell'ambito dei quali dobbiamo adesso individuare e riconoscere importanti classi numeriche più ristrette: i *numeri naturali*, i *numeri interi* e i *numeri razionali*. Altrimenti si poteva partire dall'insieme dei numeri naturali, il più ristretto, presentandolo come concetto primitivo soggetto agli assiomi di Peano per procedere poi per ampliamenti successivi, fino ai numeri reali. Le due strade sono equivalenti e sceglierne una piuttosto che l'altra è solo questione di gusti, ma dato che i fondamenti dell'Analisi Matematica si trovano nei numeri reali, abbiamo preferito il primo metodo.

I numeri naturali sono nati per *contare* gli oggetti, ed è per questo scopo che abbiamo cominciato ad usarli nei primi anni di vita, ne abbiamo acquisito facilmente e molto presto gli algoritmi di calcolo e non abbiamo mai cessato di usarli nell'esperienza quotidiana. Certamente questa esperienza acquisita è essenziale e irrinunciabile per i nostri scopi, ma per non rischiare di commettere errori logici è giunto il momento di dare un fondamento rigoroso all'esistenza stessa dei numeri naturali e di riconoscerli all'interno del sistema dei numeri reali.

Definizione 2.13 - *Un insieme $N \subset \mathbf{R}$ si dice **induttivo** se*

(N)1. $0 \in N$,

(N)2. $x \in N \Rightarrow x + 1 \in N$.

Ad esempio sono induttivi gli insiemi $\mathbf{R}$, $[0, +\infty)$ e tanti altri come $[-1/3, 1/3] \cup [1 - 1/3, 1 + 1/3] \cup [2 - 1/3, 2 + 1/3] \cup \ldots$. Quale di questi è l'insieme più piccolo?

Definizione 2.14 - *L'intersezione di tutti i sottoinsiemi induttivi di $\mathbf{R}$ viene detto insieme dei **numeri naturali** ed è indicato con $\mathbf{N}$.*

Proposizione 2.15 - $\mathbf{N}$ *è non vuoto e induttivo.*

Dimostrazione. Ricordiamo che un elemento appartiene ad ogni insieme di una famiglia se e solo se sta nella loro intersezione. Dunque $0 \in \mathbf{N}$ e, in particolare, $\mathbf{N} \neq \emptyset$.

Scelto un elemento $x \in \mathbf{N}$, si ha $x \in N$ e quindi $x + 1 \in N$ per ogni insieme induttivo N, quindi $x + 1 \in \mathbf{N}$.

□

Di notevole portata è la seguente conseguenza della Proposizione 2.15.

Corollario 2.16 (Principio d'induzione) - *Se $A \subset \mathbf{N}$ è induttivo $A = \mathbf{N}$.*

Dimostrazione. Basta osservare che siccome $\mathbf{N}$ è il più piccolo insieme induttivo, $\mathbf{N} \subset A$, ma se per ipotesi $A \subset \mathbf{N}$ allora $A = \mathbf{N}$.

□

La seguente versione equivalente si ottiene subito dal Corollario 2.16 scrivendo A nella forma $A = \{n \in \mathbf{N} \mid p(n)\}$.

Principio d'induzione - *Se p è un predicato su $\mathbf{N}$ tale che*

$(\mathbf{N})$1. $p(0)$ *è vera,*

$(\mathbf{N})$2. $p(n) \Rightarrow p(n + 1)$ $\forall n \in \mathbf{N}$ *(p è induttiva),*

allora $p(n)$ è vera per ogni $n \in \mathbf{N}$.

Se una certa affermazione, espressa da un predicato p su $\mathbf{N}$, è vera per $n = 0$ e $p(n) \Rightarrow p(n + 1)$ per ogni $n \in \mathbf{N}$, si genera un meccanismo "automatico" di tipo, appunto, *induttivo* o *iterativo*, per cui $p(1)$, $p(2)$, $p(3)$ ecc. dovranno risultare vere perché

$$p(0) \Rightarrow p(1) \Rightarrow p(2) \ \dots\ ,$$

le varie $p(n)$ sono tutte conseguenze di $p(0)$ che è vera. Il passaggio da questa catena di implicazioni all'affermazione che $p(n)$ è vera per ogni $n \in \mathbf{N}$ è intuitivo ma non banale ed ha richiesto per questo una vera dimostrazione, il Corollario 2.16, sebbene molto semplice. Ora, se al posto di 0 si considera un altro numero iniziale $k \in \mathbf{N}$, usando la traslata $q(n) = p(n+k)$, che inizia da $q(0) = p(k)$, possiamo applicare il principio d'induzione nella forma seguente

$$\textit{se } p(k) \textit{ è vera e } p(n) \Rightarrow p(n+1)\ \forall n \geqslant k \textit{ allora } p(n) \textit{ è vera per ogni } n \geqslant k\,.$$

Questo principio è uno strumento molto utile di indagine delle proprietà di $\mathbf{N}$ e serve, come vedremo, a dimostrare numerose proposizioni che riguardano i numeri naturali.

In quanto insieme induttivo, ad $\mathbf{N}$ appartengono i numeri 0, $1 = 0+1$, $2 = 1+1$, $3 = 2+1$ ecc., o più rigorosamente

$$0 \in \mathbf{N} \quad \text{e} \quad n \in \mathbf{N} \Rightarrow n+1 \in \mathbf{N}\,,$$

ma essendo il più piccolo insieme induttivo non può contenerne altri. Essi sono ordinati nel modo che conosciamo, $0 < 1 < 2 < \dots$, perché in $\mathbf{R}$ sappiamo che $n < n+1$, in particolare $0 = \min \mathbf{N}$.

Definizione 2.17 - *Un insieme totalmente ordinato $(X, \leqslant)$ viene detto* ***bene ordinato*** *se ogni sottoinsieme non vuoto di X ammette minimo.*

Esercizio 2.32 - *Dimostrare che $\mathbf{N}$ è bene ordinato. (Se $A \subset \mathbf{N}$ e $0 \in A$ la tesi è ovvia. Altrimenti si consideri l'estremo superiore dei $k \in \mathbf{N}$ tali che $k \notin A$.....)*

Teorema 2.18 (Postulato di Archimede) - *Se $a > 0$ e $b \geqslant 0$ sono due numeri reali qualunque esiste $n \in \mathbf{N}$ tale che $na > b$.*

Dimostrazione. Supponiamo $na \leqslant b$ per ogni $n \in \mathbf{N}$ e sia $L = \sup\{na \mid n \in \mathbf{N}\}$. Per la prima proprietà dell'estremo superiore $na \leqslant L$ per ogni $n \in \mathbf{N}$ e per la seconda esiste un $k \in \mathbf{N}$ tale che $ka > L - a$. Ma allora $(k+1)a > L$ che è assurdo.

□

Ne segue che i numeri naturali "attraversano" tutto $\mathbf{R}$.

Corollario 2.19 - *L'insieme $\mathbf{N}$ non è limitato superiormente in $\mathbf{R}$.*

Dimostrazione. Basta scegliere nel Teorema 2.18 $a = 1$.

□

Vediamo le operazioni algebriche che $\mathbf{N}$ eredita da $\mathbf{R}$.

Somma - *Per ogni $m, n \in \mathbf{N}$ si ha $m + n \in \mathbf{N}$.* Fissiamo ad arbitrio m e ragioniamo per induzione rispetto ad n. Dobbiamo dimostrare che la proposizione $p(n)$ che afferma "$m + n \in \mathbf{N}$" è vera per ogni $n \in \mathbf{N}$. Ora per $n = 0$ è vera per ipotesi e se è vera per n, cioè $m+n \in \mathbf{N}$, allora $(m+n)+1 \in \mathbf{N}$ perché $\mathbf{N}$ è induttivo, ma $(m+n)+1 = m+(n+1)$, quindi $p(n)$ è induttiva ed è vera per ogni $n \in \mathbf{N}$.

Osserviamo che $\mathbf{N}$ non contiene gli opposti dei suoi elementi (eccetto ovviamente l'opposto di 0) perché $n > 0 \Rightarrow -n < 0$, ma 0 non sarebbe il minimo di $\mathbf{N}$ se fosse $-n \in \mathbf{N}$. In particolare $(\mathbf{N}, +)$ non è un gruppo. Tuttavia la differenza $m - n$ (che in $\mathbf{N}$ non ha senso scrivere nella forma $m + (-n)$) può essere definita come quel numero $k \in \mathbf{N}$ tale che $k + n = m$, a condizione che sia $n \leqslant m$.

Prodotto - *Per ogni $m, n \in \mathbf{N}$ si ha $mn \in \mathbf{N}$.* Analogamente alla somma, $m \cdot 1 \in \mathbf{N}$ per ipotesi e se $mn \in \mathbf{N}$ allora $m(n+1) = mn + m \in \mathbf{N}$ in quanto somma di due numeri naturali.

Eccetto il numero 1, l'inverso di un numero naturale $n \neq 0$ non è un numero naturale (perché?) e in generale, come sopra, $m/n \notin \mathbf{N}$.

Definizione 2.20 - *Diciamo che $m \in \mathbf{N}$ è **divisibile** per $n \in \mathbf{N} - \{0\}$, o che m è **multiplo** di n (si dice anche n **sottomultiplo** o **divisore** di m), se esiste $k \in \mathbf{N}$ tale che $kn = m$. In questo caso si scrive $k = m/n$.*

Si noti che 0 è divisibile per ogni numero naturale. Si suppone che il lettore abbia familiarità con i concetti di *massimo comune divisore* e di *minimo comune multiplo* di un insieme di numeri naturali e disinvoltura nei calcoli. Ricordiamo che $p \in \mathbf{N}$ viene detto *numero primo* se $p \geqslant 2$ e non ammette divisori oltre a 1 e p stesso e che due numeri m e n sono *mutuamente primi* se non hanno divisori comuni eccetto l'1. Infine si dimostra che ogni numero naturale ammette un'unica scomposizione in fattori primi.

Non abbiamo ancora definito la potenza x^y con *base* $x \in \mathbf{R}$, $x > 0$, ed *esponente* $y \in \mathbf{R}$, ci arriveremo più avanti, ma intanto cominciamo col definire m^n con $m, n \in \mathbf{N}$.

Potenza - *Per ogni $m, n \in \mathbf{N}$ con $m > 0$ definiamo per induzione*

$$\begin{cases} m^0 = 1 \\ m^n = m \cdot m^{n-1} \quad \forall n > 0\,. \end{cases}$$

Se $m = 0$, 0^0 non ha nessun significato, poniamo invece $0^n = 0$ per ogni $n > 0$.

Ogni numero naturale può essere scritto nella forma

$$n = \sum_{i=0}^{k} a_i 10^i$$

dove i coefficienti a_i sono dei numeri naturali tali che $0 \leqslant a_i \leqslant 9$ e si chiamano *cifre* (a_0 è la cifra delle unità, a_1 quella delle decine, a_2 quella delle centinaia ecc.). Naturalmente si otterrebbero simili rappresentazioni scegliendo basi diverse $b \in \mathbf{N}$ con $0 \leqslant a_i \leqslant b - 1$. Ma una volta fissata la base, per indicare il numero è molto vantaggioso considerare la notazione posizionale

$$n = a_k a_{k-1} a_{k-2} \dots a_2 a_1 a_0$$

che in base 10 viene detta *rappresentazione decimale* di n, altrimenti binaria, ternaria ecc.. Ad esempio 1954 è la rappresentazione decimale del numero $10^3 + 9 \cdot 10^2 + 5 \cdot 10 + 4$, ma in base 16 (sistema esadecimale) significa $16^3 + 9 \cdot 16^2 + 5 \cdot 16 + 4 = 6484$. Nella notazione posizionale ogni cifra è associata ad una potenza della base b a seconda della sua posizione, a partire dall'ultima a destra che è associata a $b^0 = 1$.

Il principio d'induzione garantisce la correttezza, dal punto di vista logico, del modo *ricorsivo*, o *iterativo*, di definire una sequenza di numeri, come abbiamo appena fatto per definire la potenza. Si sceglie il valore del primo elemento e si definisce la legge che permette di passare da un elemento al successivo. Tra le tante sono ricorsive le seguenti successioni

$$\begin{cases} a_0 = 1 \\ a_{n+1} = \sqrt{1 + a_n^2} \quad \forall n \geqslant 1\,, \end{cases} \qquad \begin{cases} a_0 = \alpha \\ a_{n+1} = 4a_n(1 - a_n) \quad \forall n \geqslant 1\,, \end{cases}$$

le *frazioni continue*

$$x = a_0 + \cfrac{1}{a_1 + \cfrac{1}{a_2 + \cfrac{1}{a_3 + \dots}}}\,, \qquad a_k \in \mathbf{N} - \{0\}\,,$$

che costituiscono una rappresentazione alternativa, rispetto a quella decimale, di ogni numero reale x, e altre in cui vi è dipendenza esplicita di a_{n+1} anche da n, oltreché da a_n.

Un esempio notevole è l'algoritmo di Erone per approssimare (per eccesso) la radice di 2

$$\begin{cases} a_0 = 2 \\ a_{n+1} = \dfrac{1}{2}\left(a_n + \dfrac{2}{a_n}\right) \quad \forall n \geqslant 1\,, \end{cases} \tag{2.8}$$

Che sia $a_n > \sqrt{2}$ per ogni $n \in \mathbf{N}$ discende da

$$a_n - \sqrt{2} = \frac{1}{2a_{n-1}}(a_{n-1} - \sqrt{2})^2 > 0\,.$$

Allora $2/a_n < \sqrt{2}$ e

$$a_{n+1} - \sqrt{2} = \frac{1}{2}\left(a_n + \frac{2}{a_n}\right) - \sqrt{2} < \frac{1}{2}(a_n + \sqrt{2}) - \sqrt{2} = \frac{1}{2}(a_n - \sqrt{2}) \qquad \forall n \in \mathbf{N}\,.$$

Ne segue che la distanza di a_n dalla $\sqrt{2}$ si riduce ad ogni passo con andamento almeno esponenziale

$$a_n - \sqrt{2} < \frac{1}{2}(a_{n-1} - \sqrt{2}) < \frac{1}{2^2}(a_{n-2} - \sqrt{2}) < \ldots < \frac{1}{2^n}(a_0 - \sqrt{2}) = \frac{1}{2^n}(2 - \sqrt{2})\,.$$

Per ultima, ma non meno importante, citiamo la successione di Leonardo Pisano detto Fibonacci

$$\begin{cases} a_0 = 1 \\ a_1 = 1 \\ a_n = a_{n-1} + a_{n-2} \quad \forall n \geqslant 2\,, \end{cases} \tag{2.9}$$

strettamente legata alla sezione aurea, il numero che soddisfa la *divina proportione* $1 : x = x : (1-x)$, alla spirale logaritmica e all'armonia musicale.

Fattoriale - *Definiamo per induzione il fattoriale di n ponendo*

$$\begin{cases} 0! = 1 \\ n! = (n-1)!n\,. \end{cases}$$

Pertanto $1! = 1,\ 2! = 1 \cdot 2 = 2,\ \ldots,\ n! = 1 \cdot 2 \cdot 3 \cdot \ldots \cdot (n-1) \cdot n$.

Vediamo qualche esempio sull'uso del principio d'induzione.

Esempi

2.1 - *Formula per il calcolo della somma dei primi n numeri naturali*

$$1 + 2 + 3 + \ldots + n = \frac{n(n+1)}{2}\,, \qquad \forall n \in \mathbf{N}\,. \tag{2.10}$$

Sia $p(n)$ il predicato che esprime l'uguaglianza. La proposizione $p(1)$ è banalmente vera, ma anche $p(0)$ se a primo membro si parte da 0. Dimostriamo che $p(n)$ è induttiva, cioè che $p(n) \Rightarrow p(n+1)$ per ogni $n \in \mathbf{N}$. Se vale la $p(n)$ allora

$$1+2+3+\ldots+n+(n+1) = \frac{n(n+1)}{2}+n+1 = \frac{n(n+1)+2(n+1)}{2} = \frac{(n+1)(n+2)}{2}$$

che è la $p(n+1)$, dunque $p(n)$ è vera per ogni $n \in \mathbf{N}$.

Osserviamo che se si raddoppia la (2.10) si ottiene una formula per la somma dei primi n numeri pari

$$2 + 4 + 6 + \ldots + 2n = n(n+1)\,,$$

che naturalmente può anche essere dimostrata indipendentemente per induzione, e se da questa si toglie n membro a membro si ottiene quella per la somma per i primi n numeri dispari
$$1+3+5+\ldots+(2n-1)=n^2\,.$$
La (2.10) può essere dimostrata senza il ricorrere al principio d'induzione

$$\begin{array}{ccccccccccccc} 1 & + & 2 & + & 3 & + & \ldots & + & n & + & \\ n & + & n-1 & + & n-2 & + & \ldots & + & 1 & = & \\ \hline n+1 & + & n+1 & + & n+1 & + & \ldots & + & n+1 & = & n(n+1) \end{array}$$

o in simboli
$$\sum_{k=1}^{n} k=\frac{1}{2}\left[\sum_{k=1}^{n} k+\sum_{k=1}^{n}(n+1-k)\right]=\frac{1}{2}\sum_{k=1}^{n}(n+1)=\frac{n(n+1)}{2}\,.$$

2.2 - *Per ogni $x\in\mathbf{R}$, $x\neq 1$, si ha*
$$\sum_{k=0}^{n} x^k=\frac{1-x^{n+1}}{1-x}\qquad \forall n\in\mathbf{N}\,.$$

È evidente che la somma vale $n+1$ se $x=1$.

Senza usare il principio d'induzione, poniamo
$$s_n(x)=\sum_{k=0}^{n} x^k=1+x+x^2+\ldots+x^n$$
e osserviamo che
$$(1-x)s_n(x)=s_n(x)-xs_n(x)=1-x^{n+1}\,.$$
Col principio d'induzione, se $n=0$ è banalmente vera, e se è vera per n
$$\sum_{k=0}^{n+1} x^k=x^{n+1}+\sum_{k=0}^{n} x^k=x^{n+1}+\frac{1-x^{n+1}}{1-x}=\frac{1-x^{n+2}}{1-x}\,.$$

2.3 - *Dimostriamo la disuguaglianza*
$$3^n+4^n\leqslant 5^n\qquad \forall n\geqslant 2\,.$$
Per $n=2$ è banalmente vera valendo come uguaglianza (ed è falsa per $n=0$ e $n=1$). Verifichiamo che è induttiva
$$5^{n+1}=5^n\cdot 5\geqslant(3^n+4^n)\cdot 5=3^n\cdot 5+4^n\cdot 5>3^n\cdot 3+4^n\cdot 4=3^{n+1}+4^{n+1}\,.$$

2.4 *Per ogni $x>-1$, vale la disuguaglianza di Bernoulli*
$$(1+x)^n\geqslant 1+nx\qquad \forall n\in\mathbf{N}\,.$$

Per $n=0$ è banale (come uguaglianza). Verifichiamo che è induttiva
$$(1+x)^{n+1}=(1+x)(1+x)^n\geqslant(1+x)(1+nx)=1+(n+1)x+nx^2\geqslant 1+(n+1)x\,.$$

Esercizio 2.33 *Dimostrare l'identità di Catalan*
$$\sum_{k=1}^{n}\frac{1}{n+k}=\sum_{h=1}^{2n}\frac{(-1)^{h-1}}{h}\qquad \forall n\geqslant 1\,.\tag{2.11}$$

2.7 Insiemi finiti e calcolo combinatorio

Contare gli elementi di un insieme X significa stabilire una corrispondenza biunivoca tra X e un sottoinsieme di $\mathbf{N}$. Questa osservazione, tradotta in termini matematici, dà origine alla seguente definizione.

Definizione 2.21 - *Due insiemi X e Y si dicono **equipotenti** se esiste una funzione bigettiva $\varphi : X \to Y$.*

Esercizio 2.34 - *Dimostrare che essere equipotenti è una relazione di equivalenza tra gli insiemi.*

Definizione 2.22 - *Una classe di equivalenza rispetto alla relazione di equipotenza tra insiemi si chiama **cardinalità** o numero cardinale. Il numero cardinale di un insieme X si indica con $c(X)$.*

Definizione 2.23 - *Un insieme X si dice **finito** se esiste $n \in \mathbf{N}$ tale che X è equipotente all'insieme $I_n = \{1, 2, \dots, n\}$. In questo caso si scrive $c(X) = n$.*

Nel caso finito la cardinalità coincide col numero degli elementi. Ovviamente

$$c(I_n) = n \;\text{ e } m \neq n \quad \Rightarrow \quad c(I_m) \neq c(I_n)$$

quindi nessun insieme finito può essere equipotente ad un suo sottoinsieme proprio. L'esistenza di una corrispondenza biunivoca $\varphi : X \to I_n$ nel caso $c(X) = n$ equivale alla possibilità di *contare* gli elementi di X, che significa attribuire a ciascuno di essi l'indice che gli viene assegnato dalla φ come un'etichetta. Di fatto è già presente nella notazione $X = \{x_1, x_2, \dots, x_n\}$.

Esercizio 2.35 - *Dimostrare che se X e Y sono finiti $c(X \times Y) = c(X)c(Y)$ e che se $X \cap Y = \emptyset$ allora $c(X \cup Y) = c(X) + c(Y)$.*

Esercizio 2.36 - *Dimostrare che se X è finito e $\sim$ è una relazione di equivalenza su X tale che $c[x] = r$ per ogni $x \in X$ allora $c(X/\sim) = c(X)/r$. Ciò spiega perché $X/\sim$ si chiama insieme quoziente.*

Disposizioni - Dato un insieme di n elementi, si chiama *disposizione di ordine k*, $k \leqslant n$, ogni sottoinsieme ordinato formato da k elementi. Due disposizioni possono differire sia per la scelta degli elementi, sia per il loro ordine. Vogliamo determinare il numero $D_{n,k}$ di tutte le disposizioni di ordine k.

Ovviamente $D_{n,1} = n$. Inoltre, scelto un sottoinsieme ordinato di k elementi, il $(k+1)$-esimo può essere scelto tra gli $n-k$ del complementare. Abbiamo dunque $n-k$ possibilità per ognuna delle $D_{n,k}$ disposizioni, cioè $D_{n,k+1} = (n-k)D_{n,k}$. Pertanto

$$\begin{aligned}
D_{n,2} &= (n-1)D_{n,1} = n(n-1) \\
D_{n,3} &= (n-2)D_{n,2} = n(n-1)(n-2) \\
&\;\vdots \\
D_{n,k} &= (n-(k-1))D_{n,k-1} = n(n-1)(n-2)\dots(n-k+1)
\end{aligned} .$$

Moltiplicando e dividendo per $(n-k)!$ si può anche scrivere

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

È facile rendersi conto che $D_{n,k}$ è anche il numero di tutte le funzioni iniettive definite su un insieme di k elementi e a valori in un insieme di n elementi.

Permutazioni - In particolare, per $k = n$ si ottiene $D_{n,n} = n!$ che rappresenta il numero dei modi in cui è possibile ordinare n elementi. Ovviamente $n!$ è il numero di funzioni bigettive tra due insiemi di n elementi.

Combinazioni - Dato un insieme di n elementi, si chiama *combinazione di ordine* k, $k \leqslant n$, un sottoinsieme di k elementi, a prescindere dunque dall'ordine. Poiché ognuno di essi può essere ordinato in $k!$ modi, il numero $C_{n,k}$ di tutte le combinazioni si ottiene dividendo $D_{n,k}$ per $k!$

$$C_{n,k} = \frac{D_{n,k}}{k!} = \frac{n!}{k!(n-k)!}\,.$$

Questo numero si chiama *coefficiente binomiale* e si indica con $\binom{n}{k}$. Si osservi che

$$\binom{n}{k} = \binom{n}{n-k}\,.$$

La lista dei coefficienti binomiali forma il noto *triangolo di Tartaglia*

$$\begin{array}{ccccccccccc}
 & & & & 1 & & 1 & & & & \\
 & & & 1 & & 2 & & 1 & & & \\
 & & 1 & & 3 & & 3 & & 1 & & \\
 & 1 & & 4 & & 6 & & 4 & & 1 & \\
1 & & 5 & & 10 & & 10 & & 5 & & 1 \\
 & & & & & \vdots & & & & &
\end{array}$$

dove n varia per righe e k per colonne. L'algoritmo che ci permette di costruire ogni riga dalla precedente è proposto nel seguente esercizio.

Esercizio 2.37 - *Verificare che per* $1 \leqslant k \leqslant n$ *si ha*

$$\binom{n}{k} = \binom{n-1}{k} + \binom{n-1}{k-1}\,. \tag{2.12}$$

Il seguente esercizio giustifica il nome che viene dato a questi numeri.

Esercizio 2.38 - *Per ogni* $a, b \in \mathbf{R}$ *e per ogni* $n \in \mathbf{N}$ *vale la seguente formula per lo sviluppo del* ***binomio di Newton***

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

Si può procedere per induzione. Per $n = 0$ l'uguaglianza è banalmente vera. Dimostriamo che è induttiva

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

Come applicazione di questa formula calcoliamo il numero di tutti i sottoinsiemi di un insieme di n elementi. Se i sottoinsiemi con k elementi, le combinazioni di ordine k, sono $\binom{n}{k}$, il numero di tutti i sottoinsiemi, compreso l'intero e il vuoto, è

$$\sum_{k=0}^{n}\binom{n}{k}=\sum_{k=0}^{n}\binom{n}{k}1^k 1^{n-k}=(1+1)^n=2^n\,. \tag{2.13}$$

Esercizio 2.39 - *Applicando ripetutamente la (2.12) ai termini del II membro della (2.12) stessa, si ottiene l'identità di Vandermonde*

$$\binom{n}{k}=\sum_{i=0}^{h}\binom{h}{i}\binom{n-h}{k-i}\qquad \forall h,k,n\in\mathbf{N}\ :\ h\leqslant k\leqslant n\,.$$

Esercizio 2.40 - *Dedurre dall'esercizio precedente che*

$$\frac{(2n)!}{n!^2}\geqslant 2^n\qquad \forall n\in\mathbf{N}$$

dopo aver sostituito n con $2n$ e scelto $h=k=n$.

Esercizio 2.41 - *Dimostrare che*

$$\sum_{k=1}^{n}k\binom{n}{k}=n2^{n-1}\qquad \forall n\geqslant 1\,.$$

Il numero 2^n è anche il numero di funzioni da un insieme X di n elementi in un insieme di due elementi come $\{0,1\}$. Infatti ad ogni sottoinsieme A di X possiamo associare la *funzione caratteristica* di A

$$\chi_A(x)=\begin{cases}1 & \text{se } x\in A\\ 0 & \text{se } x\in \complement A\,.\end{cases}$$

Viceversa, ad ogni funzione $\chi : X\to\{0,1\}$ corrisponde l'insieme $A=\{x\in X : \chi(x)=1\}$. Vediamo allora quante sono le funzioni definite su un insieme di n elementi a valori in un insieme di k elementi.

Disposizioni con ripetizione - Le *disposizioni con ripetizione di k elementi e ordine n* sono le n-uple ordinate che si possono costruire con k elementi. Il loro numero, $F_{n,k}$, coincide col numero delle funzioni definite su un insieme X di n elementi a valori in un insieme Y di k elementi. Dimostriamo per induzione rispetto a n, fissato k, che $F_{n,k}=k^n$.

Per $n=1$ vi sono evidentemente k funzioni, all'unico elemento di X corrisponde uno dei k elementi di Y. Supponiamo di conoscere $F_{n,k}$ e calcoliamo $F_{n+1,k}$ relativo al caso $c(X)=n+1$. Ogni funzione definita su un sottoinsieme di X con n elementi ammette k possibili estensioni per l'$(n+1)$-esimo elemento, il rimanente, ottenibili associando a questo ogni volta uno dei k elementi di Y. Dunque $F_{n+1,k}=kF_{n,k}$ e di conseguenza $F_{n,k}=k^n$.

Esercizio 2.42 - *Dimostrare che il numero di relazioni su $X\times Y$ è pari a 2^{kn}.*

Calcolata la somma dei primi n numeri naturali, v. l'Esempio 2.1, la formula del binomio ci permette di calcolare anche la somma dei primi n quadrati perfetti. Si ha

$$\sum_{k=1}^{n}[(k+1)^3-k^3]=(2^3-1)+(3^3-2^3)+\ldots+(n^3-(n-1)^3)+((n+1)^3-n^3)=(n+1)^3-1\,.$$

D'altra parte

$$\sum_{k=1}^{n}[(k+1)^3 - k^3] = \sum_{k=1}^{n}(3k^2+3k+1) = 3\sum_{k=1}^{n}k^2 + 3\frac{n(n+1)}{2} + n$$

e per confronto delle due relazioni

$$\sum_{k=1}^{n}k^2 = \frac{n(n+1)(2n+1)}{6}$$

che può essere facilmente dimostrata anche per induzione.

Per la somma dei primi n cubi si ragiona allo stesso modo

$$\sum_{k=1}^{n}[(k+1)^4 - k^4] = (n+1)^4 - 1\,,$$

inoltre

$$\sum_{k=1}^{n}[(k+1)^4 - k^4] = \sum_{k=1}^{n}(4k^3+6k^2+4k+1) = 4\sum_{k=1}^{n}k^3 + 6\frac{n(n+1)(2n+1)}{6} + 4\frac{n(n+1)}{2} + n\,,$$

da cui

$$\sum_{k=1}^{n}k^3 = \frac{n^2(n+1)^2}{4} = \left(\sum_{k=1}^{n}k\right)^2 .$$

Riguardo il caso generale, la somma delle prime n potenze con esponente $m \in \mathbf{N}$, si dimostra che

$$\sum_{k=1}^{n}k^m = \frac{1}{m+1}\sum_{h=0}^{m}\binom{m+1}{h}B_h(n+1)^{m+1-h}$$

dove i B_h sono i *numeri di Bernoulli*, definiti induttivamente da

$$\begin{cases} B_0 = 1 \\ \displaystyle\sum_{k=0}^{n}\binom{n+1}{k}B_k = 0 \quad \forall n \in \mathbf{N}\,. \end{cases} \tag{2.14}$$

2.8 Gli interi e i razionali

Percorrendo $\mathbf{N}$ da m a n bisogna fare esattamente $k = |n-m|$ passi. Così definito, k è un numero naturale, ma se vogliamo che contenga anche l'informazione sul verso di percorrenza dobbiamo tener conto del segno e definire $k = n - m$ che è negativo se $n < m$. Le due differenze $n - m$ e $q - p$ definiscono lo stesso *numero intero (relativo)* k di passi (con segno) se e solo se $n + p = q + m$. L'insieme $\mathbf{Z}$ dei *numeri interi* è stato introdotto apposta per descrivere situazioni di questo tipo, molto comuni tra l'altro anche nella nostra esperienza quotidiana, si pensi ad un termometro graduato, o ad un estratto conto bancario i cui movimenti in entrata sono considerati positivi, mentre quelli in uscita negativi, anche queste operazioni contabili hanno un verso. In definitiva l'insieme $\mathbf{Z}$ è il sottoinsieme di $\mathbf{R}$ formato da tutti i numeri naturali insieme ai loro opposti in $\mathbf{R}$

$$\mathbf{Z} = \mathbf{N} \cup \{-n \mid n \in \mathbf{N}\} = \{0, 1, -1, 2, -2, \ldots\}$$

e, tenuto conto anche del § 2.6, eredita da $\mathbf{R}$ l'ordinamento naturale

$$\ldots -3 < -2 < -1 < 0 < 1 < 2 < 3 \ldots .$$

Ne segue che $(\mathbf{Z}, +, \leqslant)$ è un gruppo totalmente ordinato, è un sottogruppo di $(\mathbf{R}, +)$.

Esercizio 2.43 - *Dimostrare che* $\mathbf{Z}$ *non è limitato e non è bene ordinato, ma ogni sottoinsieme limitato inferiormente [superiormente] di* $\mathbf{Z}$ *ammette minimo [massimo].*

Rispetto al prodotto $\mathbf{Z} - \{0\}$ non è un gruppo per l'assenza dell'inverso, comunque valgono le proprietà associativa e commutativa e possiede l'elemento neutro. In più vale la proprietà distributiva, per questo $(\mathbf{Z}, +, \cdot)$, che non può essere un corpo, viene detto *anello*. La teoria della divisibilità, cui abbiamo accennato nel caso di $\mathbf{N}$, si applica allo stesso modo anche a $\mathbf{Z}$ con le ovvie varianti legate alle regole sui segni nei prodotti.

Consideriamo il numero reale m/n con $m, n \in \mathbf{N}$ e $n \neq 0$. Per il Postulato di Archimede del Teorema 2.18 l'insieme $\{h \in \mathbf{N} \mid m < hn\}$ è non vuoto e quindi ammette minimo. Ne segue che esiste $k \in \mathbf{N}$ tale che

$$kn \leqslant m < (k+1)n\,.$$

Il numero k così definito si chiama *quoziente* della divisione $m : n$ e coincide col risultato della divisione, cioè col rapporto m/n, se m è divisibile per n. Altrimenti $kn < m$ e il numero naturale $r = m - kn$ si chiama *resto*. Ovviamente $0 \leqslant r < n$. Infatti $r = m - kn \geqslant 0$ perché $kn \leqslant m$, inoltre se fosse $r = n + r'$ con $r' \geqslant 0$ allora $r' = m - (k+1)n$, in contraddizione con la definizione di k. Se adesso si considerano *tutti* i numeri interi k le relative differenze $r_k = m - kn$ formano un sottoinsieme di $\mathbf{Z}$ che si chiama *classe dei resti modulo* k e si indica con $\mathbf{Z}_k$. Si tratta dello stesso insieme dell'Esempio 1.14 del Capitolo 1, d'altra parte $m - r_k = kn$ è proprio la definizione della relazione di equivalenza $m \equiv r_k \bmod k$.

Esercizio 2.44 - *Dimostrare che* $\mathbf{Z}_k$ *è un anello per ogni* $k \in \mathbf{Z}$ *ed è un corpo se e solo se* k *è primo.*

Definizione 2.24 - *Si chiama* ***parte intera*** *del numero reale* x *il numero intero*

$$[x] = \max\{h \in \mathbf{Z} \mid h \leqslant x\}\,.$$

Il numero $\{x\} = x - [x] \in [0, 1[$, *si chiama* ***parte decimale*** *di* x.

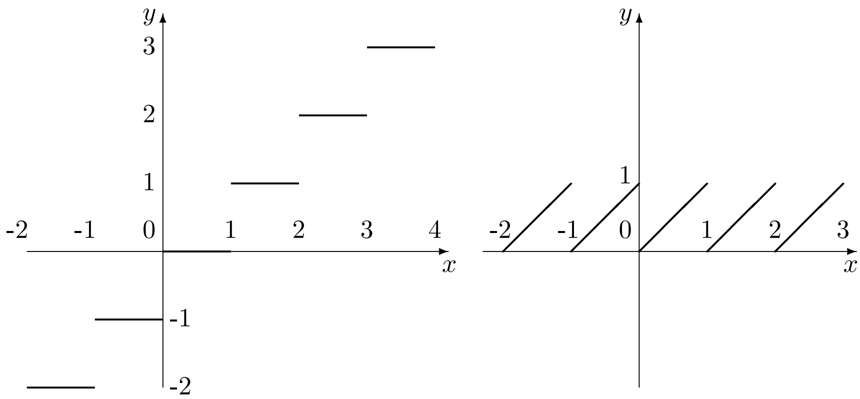


Figura 2.3: I grafici delle funzioni $[x]$ e $\{x\}$.

I numeri razionali sono gli elementi del sottoinsieme di $\mathbf{R}$

$$\mathbf{Q} = \left\{ \frac{m}{n} \;\middle|\; m \in \mathbf{Z}\,,\ n \in \mathbf{N} - \{0\} \right\}.$$

Poiché tra questi figurano anche i rapporti in cui m è multiplo di n, $\mathbf{Z} \subset \mathbf{Q}$. Ricordiamo che moltiplicando il *numeratore* m e il *denominatore* n per uno stesso numero intero h il valore del rapporto non cambia, per questo possiamo supporre in generale che la frazione con cui si vuole indicare un numero razionale sia sempre *ridotta ai minimi termini*, cioè con m e n mutuamente primi, a meno che non si debbano confrontare o sommare due frazioni con denominatori diversi com'è ben noto.

Esercizio 2.45 - *Dimostrare che* $\mathbf{Q}$ *è un campo totalmente ordinato, è un sottocampo di* $\mathbf{R}$.

Definizione 2.25 - *Un sottoinsieme* A *di* $\mathbf{R}$ *viene detto* ***denso*** *in* $\mathbf{R}$ *se per ogni coppia di numeri reali* a *e* b*, con* $a < b$*, esiste* $x \in A$ *tale che* $a < x < b$.

Teorema 2.26 - *L'insieme* $\mathbf{Q}$ *è denso in* $\mathbf{R}$.

Dimostrazione. Per il Postulato di Archimede esiste $n \in \mathbf{N}$ tale che $n(b-a) > 1$. Se due numeri reali, come na e nb, differiscono per una quantità maggiore di 1 tra di essi deve cadere un numero naturale, quindi esiste $m \in \mathbf{N}$ tale che $na < m < nb$. Dividendo per n si ottiene la tesi. □

Esercizio 2.46 - *Dimostrare che se* $A \subset B \subset \mathbf{R}$ *e* A *è denso in* $\mathbf{R}$ *allora anche* B *è denso in* $\mathbf{R}$.

Particolari numeri razionali sono quelli dell'insieme

$$\mathbf{D} = \left\{ \frac{h}{10^i} \;\middle|\; h \in \mathbf{Z}\,,\ i \in \mathbf{N} \right\} .$$

Anch'essi ammettono una comoda rappresentazione decimale, analoga a quella degli interi, che ora richiede l'uso della virgola da posizionare a seconda del valore di i. Precisamente, dato $h = a_k a_{k-1} \dots a_2 a_1 a_0$, l'allineamento decimale che rappresenta $h/10^i$ è

$$\frac{h}{10^i} = a_k a_{k-1} \dots a_i, a_{i-1} \dots a_2 a_1 a_0 .$$

Gli elementi di $\mathbf{D}$ vengono detti *numeri decimali finiti.* Si osservi che

$$\left[\frac{h}{10^i}\right] = \begin{cases} a_k a_{k-1} \dots a_i & \text{se } h \geqslant 0 \\ a_k a_{k-1} \dots a_i - 1 & \text{se } h < 0\,. \end{cases}$$

Esercizio 2.47 - *Dimostrare che i numeri decimali finiti sono tutti e soli quelli della forma* $h/2^i 5^j$ *con* $h \in \mathbf{Z}$ *e* $i, j \in \mathbf{N}$.

Esercizio 2.48 - *Dimostrare che anche* $\mathbf{D}$ *è denso in* $\mathbf{R}$.

Veniamo alla rappresentazione decimale di un numero razionale qualunque che per semplicità supponiamo positivo. Ricordiamo che calcolare m/n significa trovare due numeri naturali q e r, il quoziente e il resto, tali che $m = nq + r$ con $r < n$. Essendo $r/n < 1$, al primo passo si ritrova la decomposizione in parte intera + parte decimale

$$\frac{m}{n} = q_0 + \frac{r_0}{n} = \left[\frac{m}{n}\right] + \left\{\frac{m}{n}\right\} .$$

Come secondo passo procediamo alla divisione di $10r_0$ per n

$$\frac{10r_0}{n} = q_1 + \frac{r_1}{n}\,,$$

ricaviamo r_0/n e lo inseriamo nella precedente

$$\frac{m}{n} = q_0 + \frac{q_1}{10} + \frac{r_1}{10n} .$$

Adesso dividiamo $10r_1$ per n

$$\frac{10r_1}{n} = q_2 + \frac{r_2}{n}\,,$$

ricaviamo r_1/n e lo inseriamo nella precedente

$$\frac{m}{n} = q_0 + \frac{q_1}{10} + \frac{q_2}{100} + \frac{r_2}{100n}\,.$$

Procedendo sempre nello stesso modo si ottiene al k-esimo passo

$$\frac{m}{n} = q_0 + \frac{q_1}{10} + \frac{q_2}{10^2} + \ldots + \frac{q_k}{10^k} + \frac{r_k}{10^k n}\,.$$

Non appena $r_k = 0$ per un certo k, caso in cui m/n è decimale finito, l'algoritmo si ferma e fornisce la rappresentazione

$$\frac{m}{n} = q_0, q_1 q_2 \ldots q_k\,,$$

altrimenti non ha termine, ma in questo caso siamo di fronte ad una successione di resti (r_k), $k \in \mathbf{N}$, i cui valori variano all'interno dell'insieme finito $0, 1, 2, \ldots, n-1$. Allora uno di essi deve ripetersi dopo qualche passo e quando un certo r_i assume un valore già incontrato l'algoritmo si ritrova nelle stesse condizioni e genera nella rappresentazione decimale un gruppo di cifre che si ripetono periodicamente:

ogni numero razionale m/n ammette una rappresentazione decimale ***periodica***

$$\frac{m}{n} = q_0, q_1 q_2 \ldots q_h \overline{p_1 p_2 \ldots p_j} \tag{2.15}$$

dove q_0 è la parte intera, il gruppo di cifre $q_1 q_2 \ldots q_h$ si chiama *antiperiodo* e $p_1 p_2 \ldots p_j$ *periodo*. Come caso particolare ritroviamo i numeri decimali finiti che sono periodici con periodo nullo. Nella (2.15) si riconosce un primo esempio di somma formata da infiniti termini che ha come risultato un numero reale, finito quindi. Ricordando la regola che ci permette di passare dalla rappresentazione decimale a quella frazionaria (che nel seguito sembrerà meno misteriosa), possiamo affermare che vale anche il viceversa, dunque i numeri razionali sono tutti e soli i numeri decimali periodici.

Vediamo per concludere la rappresentazione decimale dei numeri reali.

Teorema 2.27 - *Ogni allineamento decimale rapprenta un numero reale. Viceversa, ogni numero reale ammette una rappresentazione decimale.*

Dimostrazione. Consideriamo un generico allineamento decimale

$$x_0, x_1 x_2 x_3 \ldots x_h \ldots \tag{2.16}$$

dove x_0 è un numero naturale e le x_i una famiglia qualunque di cifre. Se x_i è una cifra tra 0 e 9, indichiamo con $\hat{x}_i$ la cifra $x_i + 1$ se $x_i < 9$, altrimenti, se $x_i = 9$, poniamo $\hat{x}_i = 0$, ma a condizione di incrementare di 1 la cifra precedente. Se anche questa è 9 si sostituisce con 0 e si incrementa di 1 la cifra ancora precedente. Le due classi di numeri

$$\begin{array}{ll}
A & B \\
x_0 & \hat{x}_0 \\
x_0, x_1 & x_0, \hat{x}_1 \\
x_0, x_1 x_2 & x_0, x_1 \hat{x}_2 \\
\vdots & \vdots \\
x_0, x_1 x_2 \ldots x_h & x_0, x_1 x_2 \ldots \hat{x}_h \\
\vdots & \vdots
\end{array}$$

sono ovviamente separate e ammettono (2.16) come elemento separatore. Ma A e B sono anche contigue perché

$$x_0, x_1 x_2 \dots \hat{x}_h - x_0, x_1 x_2 \dots x_h = \frac{1}{10^h}$$

che può essere reso arbitrariamente piccolo, cioè minore di ε, pur di prendere h sufficientemente grande. Quindi (2.16) è l'unico elemento separatore.

Viceversa, sia x un numero reale. Per ogni $n \in \mathbf{N}$ il numero intero $m = [10^n x]$ soddisfa

$$\frac{m}{10^n} \leqslant x < \frac{m+1}{10^n} \tag{2.17}$$

in quanto $[10^n x] \leqslant 10^n x < [10^n x] + 1$. Per $n = 0$ la (2.17) diventa $x_0 \leqslant x < x_0 + 1$ essendo $x_0 = [x]$, per $n = 1$

$$[10x] \leqslant 10x < [10x] + 1\,, \tag{2.18}$$

d'altra parte $[10x] = 10x_0 + [10(x - x_0)] = 10x_0 + x_1$ dove $0 \leqslant x_1 < 1$ e dividendo la (2.18) per 10 si ottiene

$$x_0 \leqslant x_0 + \frac{x_1}{10} \leqslant x < x_0 + \frac{x_1}{10} + \frac{1}{10} \leqslant x_0 + 1$$

dove $0 \leqslant x_1 \leqslant 9$ in quanto $0 \leqslant x - x_0 < 1$. Per $n = 2$ si ha

$$[10^2 x] \leqslant 10^2 x < [10^2 x] + 1\,,$$

d'altra parte

$$[10^2 x] = 10^2 x_0 + 10 x_1 + \left[10^2\left(x - x_0 - \frac{x_1}{10}\right)\right]\,,$$

quindi se $x_2 = [10^2(x - x_0 - x_1/10)]$ si ottiene

$$x_0 + \frac{x_1}{10} + \frac{x_2}{10^2} \leqslant x < x_0 + \frac{x_1}{10} + \frac{x_2}{10^2} + \frac{1}{10^2}$$

dove $0 \leqslant x_2 \leqslant 9$ in quanto $0 \leqslant x - x_0 - x_1/10 < 1/10$. Proseguendo sempre nello stesso modo, all'n-esimo passo si ottiene

$$x_0 + \frac{x_1}{10} + \frac{x_2}{10^2} + \dots + \frac{x_n}{10^n} \leqslant x < x_0 + \frac{x_1}{10} + \frac{x_2}{10^2} + \dots + \frac{x_n}{10^n} + \frac{1}{10^n} \tag{2.19}$$

o, in notazione decimale,

$$x_0, x_1 x_2 \dots x_n \leqslant x < x_0, x_1 x_2 \dots \hat{x}_n\,.$$

L'allineamento decimale che le due classi di numeri così costruite al crescere di n definiscono non può che coincidere con x. □

Proposizione 2.28 - *L'allineamento decimale $x_0, x_1 x_2 \dots x_n \dots$ relativo ad un numero $x \in \mathbf{R}$ non può contenere solo la cifra 9 da un certo x_k in poi.*

Dimostrazione. Supponiamo $x_i = 9$ per ogni $i > k$. Fissato ad arbitrio $n > k$, si ha

$$\begin{aligned} x \geqslant x_0, x_1 x_2 \dots x_k x_{k+1} \dots x_n &= x_0, x_1 x_2 \dots x_k + 9 \sum_{i=k+1}^{n} \frac{1}{10^i} \\ &= x_0, x_1 x_2 \dots x_k + \frac{1}{10^k} - \frac{1}{10^n} \end{aligned}$$

che dovendo valere per ogni $n > k$ equivale a

$$x \geqslant x_0, x_1 x_2 \dots x_k + \frac{1}{10^k}\,,$$

in contraddizione con la disuguaglianza a destra nella (2.19).

□

2.9 Insiemi infiniti, cardinalità di N, Q e R

Definizione 2.29 - *Un insieme X non vuoto e non finito, cioè non equipotente ad alcun $I_n = \{1, 2 \dots, n\}$, viene detto* ***infinito****.*

Teorema 2.30 - *L'insieme* **N** *dei numeri naturali è infinito.*

Dimostrazione. Supponiamo per assurdo che per un certo $n \in \mathbf{N}$ esista una funzione bigettiva $\varphi : \mathbf{N} \to I_n$. La restrizione $\varphi_{|I_n}$ è ancora bigettiva, ma ammette come immagine un sottoinsieme proprio di I_n. Quindi I_n, pensato come dominio di questa restrizione, è equipotente a se stesso e al suo sottoinsieme proprio $\varphi(I_n)$. Ciò è assurdo perché nessun insieme finito è equipotente ad un suo sottoinsieme proprio. □

Definizione 2.31 - *Diremo* ***potenza del numerabile*** *il numero cardinale di* **N**. *Ogni insieme X equipotente a* **N** *viene detto* ***numerabile*** *e si scrive $c(X) = \aleph_0$(*)*

In particolare **N** è numerabile. Osserviamo tra l'altro che **N** è equipotente ad un suo sottoinsieme proprio, l'applicazione $n \to 2n$ per esempio è bigettiva da **N** nei numeri pari. Se esiste una funzione iniettiva da X in Y si dice che $c(X) \leqslant c(Y)$. Ovviamente questo accade sempre se $X \subset Y$ perché l'identità $i_X : X \to Y$ è iniettiva.

Esercizio 2.49 - *Dimostrare che un sottoinsieme di* **N** *non limitato superiormente è numerabile e dedurne che non esiste in* **N** *un insieme il cui numero cardinale sia compreso strettamente tra il finito e $\aleph_0$. Si consideri anche il caso generale: se $c(X) \leqslant \aleph_0$ allora X è finito oppure $c(X) = \aleph_0$.*

In base al seguente teorema, di cui omettiamo la dimostrazione, ha senso definire $c(X) < c(Y)$ se $c(X) \leqslant c(Y)$ e $c(X) \neq c(Y)$.

Teorema 2.32 (di Cantor- Bernstein) - *Se $c(X) \leqslant c(Y)$ e $c(X) \geqslant c(Y)$ allora $c(X) = c(Y)$.*

Esercizio 2.50 - *Dimostrare che se $c(X) = c(Y) = \aleph_0$ allora $c(X \cup Y) = \aleph_0$, dedurne che* **Z** *è numerabile. Inoltre $c(X \times Y) = \aleph_0$, dedurne che* **Q** *è numerabile.*

Col Teorema 2.32 si introduce una gerarchia tra le cardinalità, come fossero numeri, che è l'oggetto dell'*aritmetica transfinita*. Questo aspetto della teoria degli insiemi, scoperto da Cantor verso la fine del XIX secolo, trovò inizialmente la ferma opposizione di molti matematici dell'epoca per l'eccessiva disinvoltura con cui viene trattato l'*infinito*. Egli stesso non credeva ai risultati apparentemente paradossali delle sue ricerche, fino a dubitare della correttezza delle sue stesse dimostrazioni. Non è facile accettare, per esempio, che un piccolo segmento possa essere equipotente all'intero spazio! Con Hilbert, successivamente, si è comunque riconosciuto il grande valore della sua opera, meritevole di aver ampliato gli orizzonti della Matematica e di aver stimolato indagini assai profonde nel campo della logica e della teoria degli insiemi.

Con un semplice (e geniale) procedimento diagonale dovuto a Cantor dimostriamo che la cardinalità di **R**, $c(\mathbf{R})$, è superiore al numerabile. Prendendo a pretesto la completezza, detta anche continuità, cosa che ha poco a che fare con quanto stiamo trattando, si è scelto di attribuire ad **R**, e ad ogni insieme ad esso equipotente, la *potenza del continuo* e di indicare $c(\mathbf{R})$ con $\aleph_1$.

---

*si legge *alef*, è la prima lettera dell'alfabeto ebraico, quindi *alef con* 0 in questo caso.

Teorema 2.33 (di Cantor) - $\aleph_1 > \aleph_0$.

Dimostrazione. Se fosse $\aleph_1 = \aleph_0$ (escludiamo $\aleph_1 < \aleph_0$ perché $\mathbf{R} \supset \mathbf{N}$), sarebbe possibile fare un elenco (numerabile per sua natura) di *tutti* i suoi elementi

$$\begin{array}{l} x_0 = x_{00}, x_{01}x_{02}x_{03}\dots x_{0k}\dots \\ x_1 = x_{10}, x_{11}x_{12}x_{13}\dots x_{1k}\dots \\ x_2 = x_{20}, x_{21}x_{22}x_{23}\dots x_{2k}\dots \\ x_3 = x_{30}, x_{31}x_{32}x_{33}\dots x_{3k}\dots \\ \quad\vdots \qquad \vdots \qquad \vdots \\ x_k = x_{k0}, x_{k1}x_{k2}x_{k3}\dots x_{kk}\dots \\ \quad\vdots \qquad \vdots \qquad \vdots \end{array}$$

dove il primo indice di ogni cifra si riferisce al numero e il secondo al posto che occupa. Ma adesso costruiamo un allineamento decimale $x = x_0, x_1x_2x_3\dots x_k\dots$, quindi un numero reale, che non appartiene alla lista. Basta scegliere $x_0 \neq x_{00}$, $x_1 \neq x_{11}$, $x_2 \neq x_{22}$, ..., $x_k \neq x_{kk}$ ecc.. Questo nuovo numero differisce da x_0 per la parte intera, differisce da x_1 per la prima cifra decimale, da x_2 per la seconda, da x_k per la k-esima e così via. In definitiva differisce da tutti i numeri della lista, la quale quindi non è completa.

□

Esercizio 2.51 - *Dimostrare che tutti gli intervalli,* $\mathbf{R}$ *compreso, sono tra loro equipotenti, in particolare hanno potenza del continuo.*

Esercizio 2.52 - *Dimostrare che il quadrato* $Q = [0,1] \times [0,1]$ *ha potenza del continuo costruendo una corrispondenza biunivoca con l'intervallo* $[0,1]$ *di* $\mathbf{R}$.

Esercizio 2.53 - *Dimostrare che tutti gli spazi* $\mathbf{R}^n$*, al variare di* n*, sono tra loro equipotenti e* $c(\mathbf{R}^n) = \aleph_1$.

Da quanto visto possiamo concludere che $\aleph_0 \leqslant \aleph_1$ perché $\mathbf{N} \subset \mathbf{R}$ e che $\aleph_0 \neq \aleph_1$ perché $\mathbf{R}$ non è numerabile, quindi $\aleph_0 < \aleph_1$, ma non abbiamo detto nulla su quanto siano tra loro "distanti" questi due numeri cardinali. In altre parole, mentre è ormai ben chiara in tutti i particolari la relazione tra gli insiemi $\mathbf{N}$ e $\mathbf{R}$, nulla sappiamo della relazione tra $\aleph_0$ e $\aleph_1$, oltre al fatto che la prima è minore della seconda. Per farci un'idea più precisa dimostriamo adesso che $\mathbf{R}$ è equipotente all'insieme $\mathscr{P}(\mathbf{N})$ delle parti di $\mathbf{N}$. Per analogia con gli insiemi finiti, v. (2.13), per indicare $c(\mathscr{P}(\mathbf{N}))$ si usa anche la notazione $2^{\aleph_0}$.

Teorema 2.34 - $\aleph_1 = 2^{\aleph_0}$.

Dimostrazione. Consideriamo la famiglia $\mathscr{Q}$ di tutte le applicazioni $\chi : \mathbf{N} \to \{0,1\}$. Ad ognuna di esse corrisponde l'insieme $A = \{n \in \mathbf{N} \mid \chi(n) = 1\} \subset \mathbf{N}$, viceversa ad ogni sottoinsieme $A \in \mathbf{N}$ corrisponde la funzione caratteristica di A $\chi \in \mathscr{Q}$ definita da

$$\chi_A(n) = \begin{cases} 1 & \text{se } n \in A \\ 0 & \text{se } n \notin A\,. \end{cases}$$

È evidente che si tratta di una corrispondenza biunivoca, quindi $c(\mathscr{Q}) = 2^{\aleph_0}$. Adesso mostriamo che $\mathscr{Q}$ è equipotente a $\mathbf{R}$. Se esprimiamo i numeri reali in forma binaria ad ogni scelta della funzione $\chi \in \mathscr{Q}$ corrisponde l'allineamento binario

$$\chi(0), \chi(1)\chi(2)\chi(3)\dots\chi(k)\dots$$

Viceversa ad ogni allineamento binario

$$x_0, x_1x_2x_3 \dots x_k \dots$$

corrisponde la funzione $\chi \in \mathscr{Q}$ che assume nel loro ordine tutte le cifre x_i. Trattandosi anche in questo caso di una corrispondenza biunivoca, si ottiene che $2^{\aleph_0}$ è anche la cardinalità degli allineamenti binari. Per giungere alla tesi dobbiamo ancora superare un piccolo ostacolo, precisamente dobbiamo escludere come numeri reali tutti gli allineamenti binari che non siano propri, i quali corrispondono ai sottoinsiemi di **N** che contengono tutti i numeri naturali da un certo elemento in poi. Ma questo insieme è ovviamente numerabile e se viene tolto da un insieme con potenza del continuo ciò che rimane è ancora un continuo, quindi la tesi è dimostrata.

□

Il fatto che il numero cardinale $\aleph_0$ sia più basso di $\aleph_1 = 2^{\aleph_0}$ è in realtà una proprietà generale del passaggio da un insieme all'insieme delle parti. Ad esempio l'insieme delle parti di **R** ha potenza superiore al continuo che è come dire $\aleph_2 = 2^{\aleph_1} > \aleph_1$. Dunque per ogni numero cardinale $c(X)$ ne esiste sempre uno maggiore, ad esempio $2^{c(X)} = c(\mathscr{P}(X))$, e i numeri cardinali sono infiniti.

Teorema 2.35 - *Per ogni insieme X si ha $c(X) < 2^{c(X)}$.*

Dimostrazione. Nel caso finito di un insieme di n elementi la tesi si riduce banalmente a $n < 2^n$. Nel caso generale, la disuguaglianza $c(X) \leqslant 2^{c(X)}$ discende dal fatto che l'applicazione $x \to \{x\}$ è iniettiva da X in $\mathscr{P}(X)$, dobbiamo allora dimostrare che non esiste nessuna applicazione surgettiva tra questi insiemi.

Supponiamo per assurdo che esista una funzione surgettiva $\varphi : X \to \mathscr{P}(X)$ e definiamo l'insieme

$$H = \{x \in X \mid x \notin \varphi(x)\} .$$

Poiché $H \in \mathscr{P}(X)$ e φ è surgettiva, esiste $h \in X$ tale che $\varphi(h) = H$. L'elemento h appartiene all'insieme H? Per la definizione di H si perviene all'assurdo

$$h \in H \Leftrightarrow h \notin \varphi(h) = H .$$

□

La situazione in cui una proposizione implica la sua negazione viene detta *paradosso*.

Dal Teorema 2.35 segue subito che i numeri cardinali sono infiniti e soggetti all'ordinamento totale $\aleph_0 < \aleph_1 < \dots \aleph_h < \dots$. Ma ve ne sono altri oltre a questi? Non esiste un numero cardinale strettamente compreso tra $\aleph_0$ e $\aleph_1$? La questione ha tenuto impegnati a lungo molti matematici fin dalla dimostrazione di Cantor sulla cardinalità di **R**, ma per l'impossibilità di trovare esempi concreti si era andata formando la convinzione, nota come l'*ipotesi del continuo*, che un insieme con potenza strettamente compresa tra il numerabile e il continuo non potesse esistere. Si tratta di un caso particolare dell'*ipotesi generalizzata del continuo*, secondo cui non esisterebbe nessun numero cardinale strettamente compreso tra $c(X)$ e $2^{c(X)}$. Anche il seguente assioma ha avuto una storia controversa, all'inizio non era chiaro se poteva essere dimostrato nell'ambito della Teoria degli Insiemi (ben più profonda di come l'abbiamo trattata noi), poi si è capito che era del tutto indipendente ed equivalente all'ipotesi generalizzata del continuo, oltre che a varie altre proposizioni utili. Quindi oggi viene comunemente accettato.

Assioma 2.36 (della scelta) - *Data una famiglia $\mathscr{A}$ di insiemi, esiste una funzione, la **funzione di scelta**, che associa ad ogni insieme $A \in \mathscr{A}$ un elemento $a \in A$.*

Dall'assioma della scelta discende il seguente risultato.

Teorema 2.37 - *Ogni insieme infinito ammette un sottoinsieme numerabile.*

Dimostrazione. Si procede per induzione. L'assioma della scelta ci permette di scegliere un punto $x_0 \in X$ e di considerare l'insieme $X_1 = X - \{x_0\}$. Poi si sceglie un punto $x_1 \in X_1$ e si definisce l'insieme $X_2 = X - \{x_0, x_1\}$ dal quale si sceglie un punto x_2 e così via, scelti i primi $n-1$ punti, si sceglie l'n-esimo nell'insieme $X - \{x_0, x_1, \ldots, x_n\}$. I punti x_i sono a due a due distinti per costruzione e l'insieme $M = \{x_i \mid i \in \mathbf{N}\}$ è numerabile.

□

Corollario 2.38 - *Il numero cardinale di un insieme infinito non può essere inferiore a* $\aleph_0$.

Questa conseguenza immediata del teorema ci conferma anche che tra il finito e il numerabile non vi sono altri numeri cardinali. Un'altra conseguenza consiste nel fatto che ogni insieme infinito è equipotente ad una sua parte propria (l'implicazione inversa l'abbiamo già vista, gli insiemi finiti non godono di questa proprietà). Per dimostrarlo si considera un sottoinsieme numerabile M di X e lo si mette in corrispondenza biunivoca con un suo sottoinsieme M'. Tolti da X gli elementi di $M - M'$, ciò che rimane è equipotente a X.

Esercizio 2.54 - *Dimostrare che se* X *è infinito e* $x_0 \in X$ *allora* X *e* $X - \{x_0\}$ *sono equipotenti. Il risultato vale ancora se si toglie da* X *un insieme finito.*

Esercizio 2.55 - *Dimostrare che se* $c(X) = \aleph_0$ *e* $c(Y) = \aleph_1$ *allora* $c(X \cup Y) = \aleph_1$. *Più in generale, se* $c(X) = \aleph_h$, $c(Y) = \aleph_k$ *e* $h < k$ *allora* $c(X \cup Y) = \aleph_k$.

Tornando ai numeri reali ci si può chiedere: in che relazione sta la cardinalità di un insieme con la sua misura? Senza pretendere di trattare qui la teoria della misura, se accettiamo la lunghezza $b-a$ come la misura di un intervallo di estremi $a, b \in \mathbf{R}$, $a < b$, che li contenga o meno, un insieme esprimibile come unione di intervalli disgiunti avrà per misura la somma delle relative lunghezze. Se un insieme con un solo punto (come un intervallo chiuso con gli estremi coincidenti), o con un numero finito di punti, ha misura nulla, che cosa possiamo dire di un insieme infinito? Secondo la teoria più accreditata, quella di Lebesgue, anche ogni insieme numerabile di $\mathbf{R}$, come $[a, b] \cap \mathbf{Q}$ o $\mathbf{Q}$ stesso, ha misura nulla. Per cominciare a vedere insiemi di misura positiva è evidente che dobbiamo passare al continuo, come nel caso di un intervallo "pieno". Tuttavia, è alquanto sorprendente scoprire che esistono insiemi continui, ma di misura nulla, ne è un esempio l'insieme che ora costruiamo. Si elimina dall'intervallo $[0, 1[$ il "terzo medio", ossia $[1/3, 2/3[$, si fa poi la stessa cosa con ciascuno dei due intervalli rimanenti, $[0, 1/3[$ e $[2/3, 1[$ eliminando $[1/9, 2/9[$ e $[7/9, 8/9[$, e poi la stessa cosa anche dai quattro intervalli rimanenti eliminando $[1/27, 2/27[$, $[7/27, 8/27[$, $[19/27, 20/27[$ e $[25/27, 26/27[$ e avanti così. Se ad ogni passo si procede all'eliminazione di intervalli sempre questa regola, ciò che rimane è un intervallo tutto bucherellato, un *frattale*, che si chiama *insieme di Cantor.* Calcoliamo la misura complessiva, la somma L delle lunghezze, di tutti gli intervalli che abbiamo eliminato

$$L = \frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \ldots + \frac{2^n}{3^{n+1}} + \ldots = \frac{1}{3}\sum_{k=0}^{n}\left(\frac{2}{3}\right)^k = \frac{1}{3} \cdot \frac{1-(2/3)^{n+1}}{1-2/3}$$

che tende verso 1 per $n \to \infty$ (almeno intuitivamente, anche se non abbiamo trattato i limiti) essendo $(2/3)^{n+1} \to 0$. Ciò che abbiamo tolto ha misura pari a quella dell'intervallo iniziale, quindi l'insieme di Cantor ha misura nulla. Ma di "quanti" punti è composto? Per capirlo conviene rappresentare i numeri dell'intervallo $[0, 1[$ in base 3 usando solo le cifre 0, 1 e 2; la parte intera è sempre 0 e dopo la virgola possono comparire solo queste cifre. Al primo passo vengono eliminati i numeri che hanno la cifra 1 al primo posto dopo la virgola, che sono quelli tra 1/3 (compreso) e 2/3 (escluso), al secondo passo quelli che hanno la cifra 1 al secondo posto, al terzo quelli che l'hanno al terzo posto e così via. Alla fine saranno eliminati tutti i numeri che hanno la cifra 1 ovunque sia. Allora l'insieme di Cantor è formato dai numeri di $[0,1[$ che hanno solo le cifre 0 e 2, ma questi sono in corrispondenza biunivoca con tutti i numeri di $[0,1[$ se scritti in forma binaria e siccome questo insieme ha cardinalità del continuo lo stesso si può dire dell'insieme di Cantor, si tratta di un insieme continuo che ha misura nulla.

2.10 La retta reale

La retta euclidea può essere scelta come modello geometrico di $\mathbf{R}$ ricorrendo ad una corrispondenza biunivoca tra punti e numeri reali, detta *sistema di ascisse sulla retta*, che ne conserva le varie strutture. La continuità nella Geometria Euclidea è già presente quando si afferma che due rette non parallele del piano si incontrano in un punto (se le ascisse di ciascuna fossero solo razionali potrebbero non incontrarsi anche se non parallele). Ricorrendo ai concetti primitivi di distanza tra punti e di trasformazione rigida, è possibile in geometria confrontare e sommare due segmenti a e b. Si dicono *uguali* o *isometrici* se sono *sovrapponibili*, cioè se esiste una *trasformazione rigida* che li fa coincidere, sarà invece $a < b$ se a è uguale ad un segmento strettamente contenuto in b. Infine la loro somma, $a + b$, è il segmento che si ottiene allineando a e b uno di seguito all'altro con un solo punto in comune. In particolare $2a$, $3a$ ... ma sono ben definiti come somme di segmenti tutti uguali (anche per $m = 0$, con ma che si riduce ad un punto), detti *multipli* (interi) di a.

Definiamo dapprima il *rapporto* tra a e b ponendo

$$\frac{a}{b} = m \in \mathbf{N} \text{ se } a = mb \quad \text{e} \quad \frac{a}{b} = \frac{1}{m}\,, m \in \mathbf{N} - \{0\}\,,\ \text{se } b = ma\,,$$

inoltre

$$\frac{a}{b} = \frac{m}{n} \in \mathbf{Q} \ \textit{se}\ a = \frac{1}{n}(mb)\,,$$

in questo caso diciamo che a e b sono *commensurabili* e m/n ne esprime il rapporto. Se per b viene scelto un segmento u convenzionale come unità di misura, diciamo *misura* di a il rapporto a/u . Se invece non esiste un numero razionale che esprime il rapporto tra a e b diciamo che i due segmenti sono *incommensurabili.* Non conoscendo i numeri irrazionali, l'incommensurabilità tra il lato del quadrato e la sua diagonale rimase misteriosa e inspiegabile per i pitagorici, dato che la diagonale era perfettamente costruibile, quindi esistente, ma era impossibile da misurare prendendo il lato come unitario. Infatti dalla relazione attribuita allo stesso Pitagora $a^2 + b^2 = c^2$, in particolare $d^2 = 2l^2$ per il quadrato, discende che se esistesse un numero razionale m/n, con m e n primi tra loro, tale che $d = (m/n)l$, si avrebbe a $m^2 = 2n^2$, quindi m sarebbe pari, ma se $m = 2k$ sarebbe pari anche n, in contraddizione con l'ipotesi che siano primi tra loro.

La definizione di rapporto tra i segmenti a e b va dunque completata e lo facciamo usando la completezza di $\mathbf{R}$. Se $a \leqslant b$ consideriamo le due classi di numeri razionali

$$A = \{p \in \mathbf{Q} \mid p \geqslant 0\,,\ a \geqslant pb\} \qquad \text{e} \qquad B = \{q \in \mathbf{Q} \mid q \geqslant 0\,,\ a \leqslant qb\}\,.$$

Evidentemente A precede B, sono contigue in $\mathbf{R}$ ed esiste un unico $\xi \in \mathbf{R}$ tale che

$$p \leqslant \xi \leqslant q \quad \forall p \in A\ \forall q \in B$$

che per definizione è il rapporto tra a e b. Questo numero coincide col valore comune a $\sup A$ e $\inf B$. Come sopra, il numero reale a/u esprime la misura di a se u è il segmento unitario.

Quanto detto fin qui si può ripetere nel caso generale del rapporto tra due grandezze omogenee di qualsiasi tipo.

Adesso possiamo costruire la *retta reale* r come modello geometrico di $\mathbf{R}$. Si fissano due punti arbitrari e distinti $O, U \in r$, un'*origine* e un *punto unità*, e si assume il segmento $u = OU$ come unitario, orientato da O verso U. Questa scelta induce su r un ordinamento totale e quindi un verso positivo di percorrenza. Se r^+ è la semiretta OU e r^- la sua opposta, si definisce il sistema di ascisse $P \to x$ su r ponendo

$$x = \begin{cases} \dfrac{OP}{OU} & \text{se } P \in r^+ \\ -\dfrac{OP}{OU} & \text{se } P \in r^- \, . \end{cases}$$

Si tratta di un'applicazione bigettiva che conserva l'ordinamento, la distanza e rispetta le operazioni, è un *isomorfismo isometrico* che permette di identificare $\mathbf{R}$ con r. Naturalmente su questa costruzione si basano tutti gli altri sistemi di coordinate, anche curvilinee come abbiamo già visto in qualche esempio (si ricordi quelle sferiche), che possono richiedere più *copie* di $\mathbf{R}$ a seconda della dimensione.

Capitolo 3

Funzioni di una variabile

Questo capitolo è dedicato alla costruzione di una lista di funzioni reali di una variabile di uso comune, che chiameremo *funzioni elementari*, di cui tratteremo alcune proprietà importanti. Non sono molte e si riducono essenzialmente alle seguenti:

$$\bullet\ funzioni\ algebriche \begin{cases} razionali\ intere\ (polinomi) \\ razionali\ fratte\ (rapporti\ di\ polinomi) \\ irrazionali\ (radicali) \end{cases}$$

$$\bullet\ funzioni\ trascendenti \begin{cases} esponenziali\ e\ logaritmiche \\ trigonometriche\ e\ loro\ inverse \end{cases}$$

Combinandole tra loro mediante operazioni algebriche e prodotti di composizione, se ne possono costruire tante altre che però costituiscono solo una parte irrisoria di tutte le funzioni di una variabile esistenti e definibili nei modi più disparati, le quali sono invece "tante di più", nel senso che hanno la stessa cardinalità dell'insieme delle parti di $\mathbf{R}$, $\aleph_2$, che è maggiore del continuo. Pensiamo ad esempio a funzioni empiriche dedotte da fenomeni reali; non c'è nessun motivo per cui debbano potersi scrivere in *forma chiusa*, usando cioè funzioni elementari. Lo stesso si può dire di fronte a svariati problemi, anche semplici da formulare, che ammettono come soluzioni delle funzioni, come avviene con le equazioni differenziali. Nella grande maggioranza dei casi le funzioni soluzioni di questi problemi non si possono calcolare in modo esplicito, esistono teoremi che ne garantiscono l'esistenza, se ne possono studiare tante proprietà qualitative e anche quantitative se pensiamo ai metodi di approssimazione numerica, tutte strategie di indagine che a seconda dei casi ci costringono a mettere in discussione il senso stesso del termine "risolvere" e a scoprire nuovi significati.

Un altro aspetto che va considerato, soprattutto oggi nell'era degli elaboratori elettronici, è che in molti problemi applicativi che possono sorgere in tutti i campi delle Scienze, dalla Fisica all'Ingegneria, all'Economia ecc., il trattamento numerico delle funzioni può essere meno costoso in termini di tempo e di memoria ed è generalmente in grado di dare risposte più chiare e utili di qualsiasi rappresentazione esplicita. Anche nei rari casi in cui un problema può essere risolto usando le sole funzioni elementari, talvolta la soluzione si presenta in una forma così complessa da risultare del tutto inservibile.

Cominciamo col presentare alcune questioni riguardanti le funzioni in generale. Nel primo paragrafo trattiamo le operazioni algebriche tra le funzioni e la relazione d'ordine che ci permette di confrontarle. I successivi riguardano le proprietà fondamentali di cui una funzione può godere: limitatezza, monotonia e convessità. I problemi ai quali esse ci conducono riguardano rispettivamente la ricerca dei massimi e dei minimi, più in generale degli estremi inferiore e superiore, e la determinazione degli insiemi in cui una funzione cresce o decresce, o in cui è convessa o concava. Queste proprietà vengono definite e commentate mentre si introducono le funzioni elementari, in questo modo ci procuriamo una grande varietà di esempi su cui lavorare.

3.1 Struttura algebrica e ordinamento

Due funzioni a valori reali si possono sommare, moltiplicare e confrontare puntualmente purché siano definite sullo stesso dominio A. Per gli scopi di questo corso assumeremo che A sia contenuto in $\mathbf{R}$, ma in via di principio potrebbe anche essere di natura qualunque per il fatto che le funzioni ereditano le operazioni da $\mathbf{R}$ come codominio.

Ricordiamo che affinché una funzione sia *ben definita* è **indispensabile** specificarne con chiarezza il dominio A e la legge che attribuisce a ciascun punto $x \in A$ il valore corrispondente $f(x) \in \mathbf{R}$. In accordo con quanto visto sulle relazioni, di cui le funzioni sono un caso particolare, due funzioni $f, g : A \to \mathbf{R}$ coincidono, e scriviamo $f = g$, se e solo se

$$f(x) = g(x) \qquad \forall x \in A\,.$$

È sufficiente che in un solo punto $x \in A$ si abbia $f(x) \neq g(x)$ per concludere che $f \neq g$.

L'insieme delle funzioni definite su A e a valori in $\mathbf{R}$, che indicheremo con $\mathbf{R}^A$, è uno spazio vettoriale con le operazioni

$$\begin{aligned} (f+g)(x) &= f(x) + g(x) \quad && \forall x \in A \\ (\lambda f)(x) &= \lambda f(x) && \forall x \in A \end{aligned}$$

ed essendo presente anche un prodotto

$$(fg)(x) = f(x)g(x) \qquad \forall x \in A$$

diciamo che $\mathbf{R}^A$ è un'*algebra*.

Una funzione $f : A \to \mathbf{R}$ viene detta *costante* se per ogni coppia di punti $x_1, x_2 \in A$ si ha $f(x_1) = f(x_2)$. In modo equivalente f è costante se esiste $k \in \mathbf{R}$ tale che $f(x) = k$ per ogni $x \in A$. Anche se le notazioni possono generare un po' di confusione, si tenga ben presente fin da ora che una funzione costante e il valore che essa assume in ogni punto sono due enti ben distinti e di natura diversa: la prima è una funzione e il secondo è un numero. Oppure, più in generale, f è una funzione mentre $f(x)$ è il valore che assume nel punto x. Due esempi di funzioni costanti sono l'elemento neutro della somma, che assume in ogni punto il valore 0, e quello del prodotto che vale ovunque 1.

L'opposta e la reciproca (la funzione inversa è altra cosa) di f sono le funzioni $-f$ e $1/f$ definite da

$$(-f)(x) = -f(x) \ \ \forall x \in A \quad \text{e} \quad \left(\frac{1}{f}\right)(x) = \frac{1}{f(x)} \ \ \forall x \in A' = \{x \in A \mid f(x) \neq 0\}\,.$$

Riguardo al prodotto di composizione e il passaggio all'inversa non c'è niente da aggiungere a quanto detto nel § 1.5, ricordiamo solo che affinché la funzione composta $g \circ f$ sia ben definita è necessario che l'immagine $f(A)$ della f sia contenuta nel dominio B della g. L'identità è la funzione $i_A(x) = x$, $x \in A$, e il suo grafico è la parte della retta di equazione $y = x$ i cui punti hanno ascissa $x \in A$.

Su $\mathbf{R}^A$ è definita la relazione d'ordine

$$f \leqslant g \quad \Leftrightarrow \quad f(x) \leqslant g(x)\ \forall x \in A\,,$$

che è solo parziale perché non sempre due funzioni sono confrontabili, può essere $f(x) < g(x)$ per alcuni valori di x, ma non per altri, comunque è compatibile con le operazioni algebriche. Diciamo che f è *non negativa* ($f \geqslant 0$), *positiva* ($f > 0$), e similmente gli altri casi, se $f(x) \geqslant 0$ o $f(x) > 0$ per ogni $x \in A$. Ha senso parlare del segno di una funzione solo se è lo stesso in tutti punti, altrimenti si dice che la funzione non è definita in segno e bisogna specificare in quali punti è positiva e in quali negativa. La ormai nota funzione $x \to |x|$, $x \in \mathbf{R}$, è positiva su $\mathbf{R} - \{0\}$ e nulla per $x = 0$.

Esercizio 3.1 *Disegnare i grafici di funzioni del tipo $x \to a|x - \alpha| + b$, $x \in \mathbf{R}$, scegliendo un po' di valori particolari, negativi e positivi, per a, b e α.*

Il modulo di $f : A \to \mathbf{R}$ è la funzione $|f| : A \to \mathbf{R}$ definita da

$$|f|(x) = |f(x)| \qquad \forall x \in A\,,$$

il suo grafico coincide con quello di f nei punti nei punti $x \in A$ dove $f(x) \geqslant 0$, la parte restante, dove $f < 0$, si ottiene per simmetria attorno all'asse x.

Operazioni di reticolo tra le funzioni reali sono $f \vee g$ e $f \wedge g$

$$(f \vee g)(x) = \max\{f(x), g(x)\} \ \text{ e } \ (f \wedge g)(x) = \min\{f(x), g(x)\} \qquad \forall x \in A\,,$$

per cui ad esempio $|f| = -f \vee f$. Di verifica immediata sono le seguenti

$$f \vee g = \frac{f + g + |f - g|}{2} \ \text{ e } \ f \wedge g = \frac{f + g - |f - g|}{2}\,.$$

Le funzioni positive

$$f^+ = f \vee 0 \ \text{ e } \ f^- = -(f \wedge 0)$$

si chiamano rispettivamente *parte positiva* e *parte negativa* di f, per cui

$$f = f^+ - f^- \ \text{ e } \ |f| = f^+ + f^-\,.$$

Esercizio 3.2 - *Se $i_{\mathbf{R}}$ indica l'identità su $\mathbf{R}$, riconoscere e tracciare il grafico delle funzioni $i^+_{\mathbf{R}}$, $i^-_{\mathbf{R}}$ e $|i_{\mathbf{R}}|$.*

Esercizio 3.3 - *Consideriamo il biliardo $[0, a] \times [0, b]$, $a, b > 0$. Una boccia viene lanciata dal punto $(0, p)$ di una sponda e, dopo un rimbalzo su una delle due sponde contigue, raggiunge il punto (a, q) di quella opposta. Di quale funzione è grafico la traiettoria percorsa?*

Assumendo che il moto sia rettilineo prima e dopo il rimbalzo, per la legge della riflessione la funzione richiesta deve essere del tipo $f(x) = k|x - x_0|$, $x \in [0, a]$. Non resta allora che scegliere il punto $x_0 \in [0, a]$ e il coefficiente $k \geqslant 0$ in modo che siano soddisfatte le condizioni

$$f(0) = p \qquad \text{e} \qquad f(a) = q$$

che si traducono nel sistema

$$\begin{cases} kx_0 = p \\ k(a - x_0) = q\,. \end{cases}$$

La soluzione è $k = (p + q)/a$ e $x_0 = ap/(p + q)$.

3.2 Potenze e polinomi

Definiamo induttivamente le *potenze* con esponente $n \in \mathbf{N}$ ponendo per ogni $x \in \mathbf{R}$

$$\begin{cases} p_0(x) = 1 \\ p_n(x) = x p_{n-1}(x) \quad \forall n \in \mathbf{N}\,. \end{cases}$$

Riconosciamo in questo modo $p_1(x) = x \cdot 1 = x$ che è la funzione identità su $\mathbf{R}$, $p_2(x) = x \cdot x = x^2$, $p_3(x) = x \cdot x^2 = x^3 \ldots p_n(x) = x \cdot x^{n-1} = x^n$ e così via. La p_0 prolunga nel modo più ovvio la potenza $x \to x^0$ da $\mathbf{R} - \{0\}$ a tutto $\mathbf{R}$. Escluso la p_0, tutte le p_n si annullano solo in 0 e $p_n(1) = 1$ per ogni $n \in \mathbf{N}$.

Esercizio 3.4 - *Dalla definizione dedurre le proprietà $x^{m+n} = x^m x^n$ e $(x^m)^n = x^{mn}$ per ogni $m, n \in \mathbf{N}$.*

Esercizio 3.5 - *Moltiplicando per x membro a membro ripetutamente (cioè per induzione) la relazione $x > 1$, mostrare che $x^n > x^{n-1}$ per ogni $n \in \mathbf{N}$ e che quindi $m < n \Rightarrow x^m < x^n$. Cosa succede se $0 < x < 1$?*

Conseguenza di questo esercizio è che le potenze, esclusa la p_0, assumono valori arbitrariamente grandi, non sono cioè limitate superiormente. Infatti per ogni $M > 1$, basta scegliere $x = M$ e si ha $p_n(x) = M^n > M$.

La p_0 assume il valore costante 1 e non c'è molto da dire. Riguardo le potenze ad esponente positivo, nel caso pari si ha

$$p_{2n}(x) = (x^n)^2 \geqslant 0 \; \forall x \in \mathbf{R} \qquad \text{e} \qquad p_{2n}(x) = 0 \;\text{ se e solo se } x = 0\,,$$

quindi il valore minimo di queste funzioni è 0 e $x = 0$ è l'unico punto in cui viene raggiunto, è il punto di minimo. Nel caso dispari si ha

$$\operatorname{sign} x^{2n+1} = \operatorname{sign} x^{2n} \operatorname{sign} x = \operatorname{sign} x$$

Le relazioni

$$p_n(x) = p_n(-x) \text{ se } n \text{ è pari} \qquad \text{e} \qquad p_n(x) = -p_n(-x) \text{ se } n \text{ è dispari} \qquad \forall x \in \mathbf{R}$$

ci permettono di dedurre il comportamento delle potenze sull'intervallo $]-\infty, 0]$ da quello su $[0, +\infty[$. Ad esempio quelle dispari non sono limitate neanche inferiormente su $\mathbf{R}$, oppure, se tutte le potenze sull'intervallo $[0, +\infty[$ sono crescenti, su $]-\infty, 0]$ sono decrescenti quelle con esponente pari e crescenti quelle con esponente dispari, le quali sono in realtà crescenti su $\mathbf{R}$.

Crescenti su $[0, +\infty[$ significa: $\forall x_1, x_2 \in [0, +\infty[\quad x_1 < x_2 \Rightarrow p_n(x_1) < p_n(x_2)$. Per verificarlo basta semplice scomposizione in fattori

$$x_1^n - x_2^n = (x_1 - x_2)(x_1^{n-1} + x_1^{n-2} x_2 + \ldots + x_2^{n-1})$$

e se $0 \leqslant x_1 < x_2$ si ha $p_n(x_1) < p_n(x_2)$ perché la somma dentro la seconda parentesi è positiva. Ora, per n pari o dispari si ha rispettivamente

$$x_1 < x_2 < 0 \Leftrightarrow 0 < -x_2 < -x_1 \Rightarrow x_2^n = (-x_2)^n < (-x_1)^n = x_1^n$$

$$x_1 < x_2 < 0 \Leftrightarrow 0 < -x_2 < -x_1 \Rightarrow \; -x_2^n = (-x_2)^n < (-x_1)^n = -x_1^n \Leftrightarrow x_1^n < x_2^n\,,$$

inoltre per n dispari p_n è crescente su $\mathbf{R}$ perché in più

$$x_1 < 0 < x_2 \Rightarrow x_1^n < 0 < x_2^n\,.$$

I grafici delle potenze sono illustrati nella Figura 3.1.

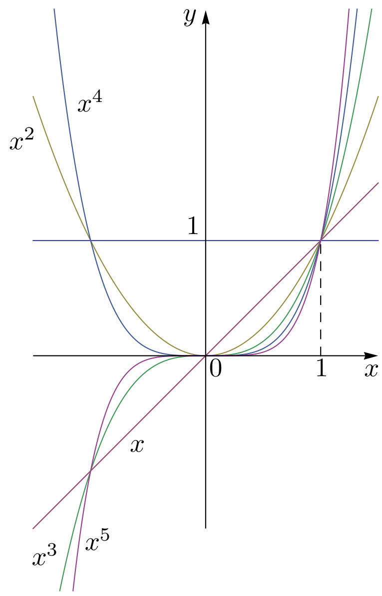


Figura 3.1: I grafici di alcune potenze.

Le considerazioni fatte sulla monotonia implicano la limitatezza delle potenze su ogni intervallo limitato $[a,b]$, infatti

$$a \leqslant x \leqslant b \Rightarrow |x| \leqslant \max\{|a|,|b|\} = k \Rightarrow |x^n| \leqslant k^n \quad \forall x \in [a,b]\,.$$

Dopo aver considerato le proprietà di limitatezza, che comprende lo studio dei massimi e dei minimi, e di monotonia, trattiamo adesso la convessità. La potenza $p_2(x) = x^2$, $x \in \mathbf{R}$, è convessa perché per ogni $x_1, x_2 \in \mathbf{R}$ la corda che unisce i punti del grafico (x_1, x_1^2) e (x_2, x_2^2) si trova al di sopra del grafico su $[x_1, x_2]$. In altre parole

$$p_2(x) \leqslant p_2(x_1) + \frac{p_2(x_2) - p_2(x_1)}{x_2 - x_1}(x - x_1) \qquad \forall x \in [x_1, x_2]\,. \tag{3.1}$$

Poiché $x \in [x_1, x_2]$ se e solo se esiste $t \in [0,1]$ tale che $x = x_1 + t(x_2 - x_1) = tx_2 + (1-t)x_1$, la (3.1) può essere scritta nella forma

$$p_2(tx_2 + (1-t)x_1) \leqslant tp_2(x_2) + (1-t)p_2(x_1)\,. \tag{3.2}$$

Per dimostrare la (3.2) possiamo ricorrere alla disuguaglianza $2ab \leqslant a^2 + b^2$

$$\begin{aligned}(tx_2 + (1-t)x_1)^2 &= t^2x_2^2 + 2t(1-t)x_1x_2 + (1-t)^2x_1^2 \\ &\leqslant t^2x_2^2 + t(1-t)(x_1^2 + x_2^2) + (1-t)^2x_1^2 = tx_2^2 + (1-t)x_1^2\,.\end{aligned}$$

Anche la funzione $x \to |x|$ è convessa su $\mathbf{R}$

$$|tx_2 + (1-t)x_1| \leqslant t|x_2| + (1-t)|x_1|$$

per la proprietà triangolare del modulo. Da questa ricaviamo per induzione la convessità di tutte le funzioni $x \to |x|^n$ con $n \in \mathbf{N}$ ricorrendo alla disuguaglianza $a^nb + ab^n \leqslant a^{n+1} + b^{n+1}$, $a, b > 0$, che essendo equivalente a $(a^n - b^n)(a - b) > 0$

(le potenze sono crescenti su $[0,+\infty[$) è vera. Se dunque $|x|^n$ è convessa dimostriamo che lo è anche $|x|^{n+1}$

$$\begin{aligned} |tx_2+(1-t)x_1|^{n+1} &= |tx_2+(1-t)x_1|^n|tx_2+(1-t)x_1| \\ &\leqslant (t|x_2|^n+(1-t)|x_1|^n)(t|x_2|+(1-t)|x_1|) \\ &= t^2|x_2|^{n+1}+(1-t)^2|x_1|^{n+1}+t(1-t)(|x_2|^n|x_1|+|x_1|^n|x_2|) \\ &\leqslant t^2|x_2|^{n+1}+(1-t)^2|x_1|^{n+1}+t(1-t)(|x_2|^{n+1}+(|x_1|^{n+1}) \\ &= t|x_2|^{n+1}+(1-t)|x_1|^{n+1}\,. \end{aligned}$$

Una *funzione polinomiale* reale in una variabile reale, detta anche *funzione razionale intera*, o brevemente un *polinomio*, è una funzione $P:\mathbf{R}\to\mathbf{R}$ del tipo

$$(3.3)\qquad P(x)=a_nx^n+a_{n-1}x^{n-1}+\ldots+a_1x+a_0=\sum_{k=0}^{n}a_kx^k\qquad\forall x\in\mathbf{R}\,,$$

dove i numeri reali $a_0,a_1,a_2,\ldots,a_n$ si chiamano *coefficienti*. Si tratta dunque di una combinazione lineare finita di potenze ad esponente naturale. Diciamo che P ha *grado* nullo se $a_i=0$ per ogni $i\geqslant 1$, caso in cui P si riduce alla costante a_0, altrimenti, se almeno uno di essi è non nullo, il grado di P è il numero

$$\mathrm{gr}(P)=\max\{i\geqslant 1\mid a_i\neq 0\}\,.$$

Naturalmente in una rappresentazione del tipo (3.3) supporremo sempre $a_n\neq 0$, in modo che il grado coincida con l'indice più alto.

I polinomi formano un sottospazio vettoriale delle funzioni con la somma e il prodotto per scalari

$$(P+Q)(x)=\sum_{k=0}^{m}(a_k+b_k)x^k+\sum_{k=m+1}^{n}a_kx^k\qquad\text{e}\qquad(\lambda P)(x)=\sum_{k=0}^{n}\lambda a_kx^k\,,$$

essendo

$$P(x)=\sum_{k=0}^{n}a_kx^k\,,\quad Q(x)=\sum_{k=0}^{m}b_kx^k\,,\quad m\leqslant n\,.$$

È evidente che $\mathrm{gr}(P+Q)=n$ se $n>m$ e $\mathrm{gr}(P+Q)\leqslant n$ se $n=m$. Infine il prodotto è il polinomio

$$(PQ)(x)=\sum_{h=0}^{m}\sum_{k=0}^{n}a_hb_kx^{h+k}=\sum_{k=0}^{m+n}\left(\sum_{i=0}^{k}a_ib_{k-i}\right)x^k\,,$$

per cui $\mathrm{gr}(PQ)=\mathrm{gr}(P)+\mathrm{gr}(Q)$. Gli elementi neutri della somma e del prodotto sono rispettivamente il polinomio costante nullo e quello costante pari a 1, tuttavia, mentre l'opposto di un polinomio esiste, basta cambiare di segno a tutti i suoi coefficienti, il reciproco no (come polinomio s'intende), altro che nel caso costante. Per questo i polinomi formano un *anello*, esattamente come i numeri interi. L'analogia si spinge oltre includendo la teoria della divisibilità.

Definizione 3.1 - *Diciamo che un polinomio D non costante* ***divide*** *il polimonio P, oppure che P è* ***divisibile*** *per D, se $\mathrm{gr}(D)<\mathrm{gr}(P)$ ed esiste un polinomio Q tale che $P=DQ$.*

Nell'ordine, P, D e Q si chiamano *dividendo*, *divisore* e *quoziente*. Ovviamente anche Q divide P se la definizione è soddisfatta e $\mathrm{gr}(D)+\mathrm{gr}(Q)=\mathrm{gr}(P)$.

Una *radice*, o *zero*, del polinomio P è un numero reale α tale che $P(\alpha)=0$. Poiché $P(\alpha)=D(\alpha)Q(\alpha)$, α è una radice di P se e solo se lo è di D o di Q. Un polinomio non costante viene detto *irriducibile* se è di primo grado o non ha radici in $\mathbf{R}$. Ovviamente moltiplicare un polinomio per uno stesso numero reale non nullo non ha alcuna influenza sulle radici.

Consideriamo il caso in cui P non sia divisibile per D. Allora esiste un polinomio R, detto *resto*, tale che $\mathrm{gr}(R) < \mathrm{gr}(D)$ e

$$P(x) = D(x)Q(x) + R(x) \qquad \forall x \in \mathbf{R}\,. \tag{3.4}$$

Se in particolare D è di primo grado R è costante e può essere pensato come un numero. Supponiamo allora che D sia del tipo $D(x) = x - \alpha$, caso a cui ci si può sempre ricondurre moltiplicando Q per una costante, la precedente relazione diventa

$$P(x) = (x - \alpha)Q(x) + R \qquad \forall x \in \mathbf{R}\,, \tag{3.5}$$

dalla quale, per $x = \alpha$, discende $P(\alpha) = R$. Vale dunque il seguente famoso teorema.

Teorema 3.2 (Ruffini) - *Il numero reale α è una radice del polinomio P se e solo se P è divisibile per $x - \alpha$.*

Siano P un polinomio e $\alpha \in \mathbf{R}$ una sua radice, per cui esiste un polinomio Q_1 tale che

$$P(x) = (x - \alpha)Q_1(x) \qquad \forall x \in \mathbf{R}\,.$$

Se $Q_1(\alpha) = 0$ esiste un altro polinomio Q_2 tale che $Q_1(x) = (x - \alpha)Q_2(x)$ e di conseguenza

$$P(x) = (x - \alpha)^2 Q_2(x) \qquad \forall x \in \mathbf{R}$$

e se anche $Q_2(\alpha) = 0$ si ottiene, per qualche polinomio Q_3, $P(x) = (x - \alpha)^3 Q_3(x)$ e così via, finché non si arriva ad un esponente m tale che

$$P(x) = (x - \alpha)^m Q_m(x) \qquad \forall x \in \mathbf{R}\,, \qquad Q_m(\alpha) \neq 0\,.$$

Allora diciamo che α è radice di P con molteplicità m, semplice per $m = 1$, doppia per $m = 2$, tripla per $m = 3$ ecc. Ragionando allo stesso modo per Q_m si deduce il seguente risultato sulla fattorizzazione:

Siano $x_1, x_2, \ldots, x_h \in \mathbf{R}$ le radici distinte di un polinomio P non identicamente nullo. Allora esistono h numeri naturali $m_1, m_2, \ldots, m_h$ ed eventualmente un polinomio P_0 privo di zeri tali che

$$P(x) = a_n (x - x_1)^{m_1} (x - x_2)^{m_2} \ldots (x - x_h)^{m_h} P_0(x) \qquad \forall x \in \mathbf{R}\,, \tag{3.6}$$

Per la (3.6) si ha

$$\mathrm{gr}(P) = m_1 + m_2 + \ldots + m_h + \mathrm{gr}(P_0) > m_1 + m_2 + \ldots + m_h\,,$$

quindi un polinomio di grado $n \geqslant 1$ non può avere più di n radici distinte. Dalla (3.6) discende anche il seguente teorema.

Teorema 3.3 (Principio d'identità dei polinomi) - *Due polinomi coincidono se e solo se hanno uguali i coefficienti dei rispettivi termini simili.*

Dimostrazione. L'enunciato equivale, per differenza, ad affermare che

$$P(x) = \sum_{k=0}^{n} a_k x^k = 0 \quad \forall x \in \mathbf{R} \quad \Leftrightarrow \quad a_k = 0 \quad \forall k = 1, 2, \ldots, n\,.$$

L'implicazione "$\Leftarrow$" è ovvia. Per la "$\Rightarrow$", se P è il polinomio nullo e fosse $a_0 \neq 0$ allora $P(0) \neq 0$ contro l'ipotesi; se non fosse nullo uno degli altri coefficienti allora

P avrebbe grado positivo e un numero finito di radici, ma anche questo contraddice l'ipotesi che P sia ovunque nullo.

□

La (3.5) suggerisce rappresentazioni diverse dello stesso polinomio a seconda di dove viene "centrato". La (3.3) per esempio è centrata in 0 (si pensi a x come $x-0$), ma scegliendo come centro qualunque altro punto $x_0 \in \mathbf{R}$ lo stesso polinomio assume la forma

$$P(x) = \sum_{k=0}^{n} a'_k (x-x_0)^k \tag{3.7}$$

con nuovi coefficienti a'_i. In questo caso diciamo che P è centrato in x_0. Un'applicazione di questa osservazione è illustrata nel seguente esercizio.

Esercizio 3.6 - *Trovare $x_0 \in \mathbf{R}$ tale che, scrivendo il polinomio (3.3) nella forma (3.7), si abbia $a'_{n-1} = 0$.*

Per ricavare la (3.7) basta usare ripetutamente la (3.5). Al primo passo

$$P(x) = P(x_0) + P_1(x)(x-x_0)\,,$$

d'altra parte

$$P_1(x) = P_1(x_0) + P_2(x)(x-x_0)\,,$$

che, inserita nella precedente, ci dà

$$P(x) = P(x_0) + P_1(x_0)(x-x_0) + P_2(x)(x-x_0)^2$$

e continuando allo stesso modo per P_2, P_3 ecc., il cui grado diminuisce ad ogni passo, si perviene ad un ultimo polinomio P_n, ultimo perché n è il grado di P, che assume il valore costante $P_n(x_0)$ per cui si ha

$$P(x) = \sum_{k=0}^{n} P_k(x_0)(x-x_0)^k \qquad \forall x \in \mathbf{R}\,.$$

Più avanti, usando le derivate, saremo in grado di calcolare i numeri $P_k(x_0)$.

Il problema della ricerca di formule risolutive per le equazioni algebriche è durato centinaia di anni. Ben nota è quella dell'equazione di secondo grado, forse meno note sono quelle relative al terzo e quarto grado, ma sulle equazioni di grado superiore a 4 è stato dimostrato da Galois che non sono risolubili per radicali. Nella pratica questa carenza può essere parzialmente compensata cercando di "indovinare", con un po' di fortuna, una prima soluzione, poi un'altra ecc., col vantaggio di abbassare ogni volta il grado dell'equazione. Per i polinomi a coefficienti interi i tentativi da fare per la ricerca delle eventuali radici razionali sono "pochi", nel senso del seguente esercizio.

Esercizio 3.7 - *Dimostrare che ogni radice razionale di un polinomio a coefficienti interi (se sono razionali non fa differenza) è il rapporto, a meno del segno, tra un divisore di a_0 e un divisore di a_n.*

Informazioni aggiuntive, ad esempio sui segni dei coefficienti e/o sui valori del polinomio in qualche punto speciale, possono limitare ulteriormente i numeri da provare. Non tutti i numeri reali sono radici di polinomi a coefficienti interi, ma quelli che lo sono si chiamano *numeri algebrici.*

Esercizio 3.8 - *Dimostrare che i numeri algebrici formano un insieme numerabile che contiene strettamente* $\mathbf{Q}$.

L'uso dei radicali nelle formule risolutive ha portato storicamente alla costruzione dei numeri complessi, argomento cui sarà dedicato il prossimo capitolo. Vedremo che nell'insieme $\mathbf{C}$ dei numeri complessi l'estrazione della radice n-esima è sempre possibile e questo fa di esso l'ambiente naturale in cui la ricerca delle radici di un polinomio trova un fondamento teorico. Mentre in $\mathbf{R}$ non si dispone di un risultato generale, d'altra parte siamo a conoscenza di polinomi come x^2+1, x^4+1 e x^2+x+1 che non hanno radici reali, in $\mathbf{C}$ vale il *Teorema fondamentale dell'Algebra* per il quale ogni polinomio complesso ammette almeno una radice, di conseguenza (ovvia) ne ammette tante quant'è il grado se contate con la loro molteplicità e gli unici polinomi irriducibili sono quelli di primo grado. Vi sono numerose dimostrazioni di questo risultato, ma richiedono tutte delle conoscenze che esulano dagli scopi di questo corso. Ora ci siamo limitati a citarlo per la sua importanza e ce ne serviremo più avanti anche per dedurre ulteriori proprietà dei polinomi reali, come il fatto che quelli di grado superiore al secondo non sono mai irriducibili.

Volendo rimanere in quest'ambito, che cosa si può dire sull'esistenza di radici reali reali? Una condizione sufficiente si basa sull'idea, ragionevole ed intuitiva, che se un polinomio assume valori di segno opposto in due punti allora deve annullarsi da qualche parte in un punto intermedio. Vedremo più avanti che questa proprietà è tipica delle funzioni continue.

Lemma 3.4 - *Ogni polinomio è localmente limitato.*

Dimostrazione. L'abbiamo già osservato per le potenze e ogni polinomio è una combinazione finita di potenze.

□

Lemma 3.5 - *Per ogni $a, b \in \mathbf{R}$ esiste $k \geqslant 0$ tale che*

$$|P(x) - P(x_0)| \leqslant k|x - x_0| \qquad \forall x, x_0 \in [a, b]\,.$$

Dimostrazione. Scelto $x_0 \in [a, b]$, sappiamo che esiste un polinomio Q tale che

$$P(x) = P(x_0) + Q(x)(x - x_0) \qquad \forall x \in \mathbf{R}\,.$$

Per il Lemma 3.4 esiste $k \geqslant 0$ tale che $|Q(x)| \leqslant k$ per ogni $x \in [a, b]$, pertanto

$$|P(x) - P(x_0)| = |Q(x)||x - x_0| \leqslant k|x - x_0| \qquad \forall x \in [a, b]\,.$$

□

Lemma 3.6 - *Per ogni $x_0 \in \mathbf{R}$ tale che $P(x_0) \neq 0$ esiste un numero reale $\delta > 0$ tale che $P(x)$ ha lo stesso segno di $P(x_0)$ su $I_\delta(x_0) =]x_0 - \delta, x_0 + \delta[$.*

Dimostrazione. Supponiamo $P(x_0) > 0$. Scegliendo nel Lemma 3.5 $a = x_0 - 1$ e $b = x_0 + 1$, esiste $k \geqslant 0$ tale che

$$P(x) \geqslant P(x_0) - k|x - x_0| \qquad \forall x \in [x_0 - 1, x_0 + 1]\,.$$

Allora per tutti gli x sufficientemente vicini a x_0, precisamente per $|x - x_0| < P(x_0)/k$, si ha $P(x) > 0$, quindi la tesi è verificata se $\delta = \min\{1, P(x_0)/k\}$.

Se $P(x_0) < 0$ si ragiona allo stesso modo con la disuguaglianza

$$P(x) \leqslant P(x_0) + k|x - x_0|\,,$$

oppure ci si può ricondurre al caso precedente considerando $-P(x_0)$.

□

Teorema 3.7 (degli zeri) - *Siano $a, b \in \mathbf{R}$, $a < b$, tali che $P(a)P(b) < 0$. Allora esiste $\xi \in]a, b[$ tale che $P(\xi) = 0$.*

Dimostrazione. Supponiamo $P(a) < 0$ e $P(b) > 0$ e poniamo

$$A = \{x \in [a, b] \mid P(x) < 0\} .$$

Tale insieme è non vuoto, dato che contiene a, e limitato superiormente da b. Dimostriamo che $\xi = \sup A$ è uno zero di P. Se fosse $P(\xi) < 0$, per il Lemma 3.6 $P(x)$ rimarrebbe negativo in tutti i punti di un intervallo $[\xi, \xi + \delta[$ e ξ non sarebbe un maggiorante di A. Se invece fosse $P(\xi) > 0$, $P(x)$ rimarrebbe positivo in tutti i punti di un intervallo $]\xi - \delta, \xi]$ e ξ non sarebbe il minimo dei maggioranti di A.

□

Se in più il polinomio è su $[a, b]$ una funzione strettamente monotona, e quindi iniettiva, ammette un unico zero. Il Teorema degli zeri ci dà soltanto una condizione sufficiente, polinomi come x^2, $x^2(x-1)^2$ hanno degli zeri senza presentare alcun cambiamento di segno. Inoltre, come sempre, può essere letto al contrario: un polinomio privo di zeri assume ovunque valori con lo stesso segno, in particolare è limitato da una parte.

Sul Teorema degli zeri si basano molti algoritmi numerici per l'approssimazione delle radici. Ma per un'implementazione efficace mediante programmi di calcolo è necessario, ma anche utile per risparmiare tempo e memoria, fornire qualche buona stima a priori di un intervallo che le contenga. È facile rendersi conto, scrivendo P nella forma

$$P(x) = x^n \left(a_n + \frac{a_{n-1}}{x} + \frac{a_{n-2}}{x^2} + \ldots + \frac{a_1}{x^{n-1}} + \frac{a_0}{x^n} \right) , \tag{3.8}$$

che per $|x|$ sufficientemente grande $P(x)$ è definitivamente concorde con a_n per n pari, caso in cui può anche non avere radici reali, ma se n è dispari almeno una radice deve esistere perché per $|x|$ sufficientemente grande $P(x)$ sarà definitivamente concorde con a_n se $x > 0$ e discorde se $x < 0$. Vediamo più precisamente come determinare una costante $c > 0$ tale che l'intervallo $[-c, c]$ contenga tutte le radici reali. Supponiamo $a_n > 0$ nella (3.8). I termini con indice k tale che $1 \leqslant k \leqslant n$ soddisfano $a_{n-k}/x^k > -|a_{n-k}|/x$ per $x > 1$, quindi

$$P(x) \geqslant x^n \left(a_n - \frac{Mn}{x} \right) \qquad \forall x > 1 , \tag{3.9}$$

dove $M = \max\{|a_k| \mid 0 \leqslant k \leqslant n-1\}$. Ne segue che se n è pari $P(x) > 0$ per $|x| > c = \max\{1, Mn/a_n\}$, se invece è dispari $P(x) > 0$ per $x > c$ e $P(x) < 0$ per $x < -c$, in ogni caso le radici di P possono stare solo in $[-c, c]$.

Dalla (3.9) si deduce anche la stima dal basso

$$P(x) \geqslant \frac{a_n}{2} x^n \qquad \forall x \geqslant \max\left\{1, \frac{2Mn}{a_n}\right\} \tag{3.10}$$

in cui P maggiora una funzione non limitata superiormente. A seconda che sia n pari o dispari vale la stessa disuguaglianza o l'opposta per $x \leqslant \min\{-1, -2Mn/a_n\}$, quindi un polinomio di grado dispari ammette sempre una radice reale e un polinomio di grado pari è limitato inferiormente se $a_n > 0$ e superiormente se $a_n < 0$.

Esercizio 3.9 - *Dimostrare che un polinomio $P(x) = a_n x^n + \ldots + a_0$ di grado pari tale che $a_n > 0$ e $a_0 < 0$ ha almeno una radice positiva e una radice negativa. Se invece ha grado dispari, sempre con $a_n > 0$, ammette una radice positiva se $a_0 < 0$ e una radice negativa se $a_0 > 0$. In particolare ogni polinomio di grado dispari ha almeno una radice reale.*

Più avanti vedremo un paio di algoritmi ricorsivi per l'approssimazione degli zeri di una funzione, il metodo delle corde e delle tangenti.

3.3 Funzioni limitate

La definizione di funzione limitata si riconduce a quella di insieme limitato in $\mathbf{R}$.

Definizione 3.8 - *Una funzione $f : A \to \mathbf{R}$ si dice* ***limitata superiormente*** *[**inferiormente**] se $f(A)$ è un sottoinsieme di $\mathbf{R}$* ***limitato superiormente*** *[**inferiormente**]. Se f è limitata sia superiormente che inferiormente diciamo che è* ***limitata****. Gli elementi $\sup f(A)$ e $\inf f(A)$ esistono sempre in $\bar{\mathbf{R}}$, si chiamano* ***estremo superiore*** *ed* ***estremo inferiore*** *di f e si indicano con*

$$\sup\{f(x) \mid x \in A\}\,,\quad \sup_{x\in A} f(x)\,,\quad \sup_A f \quad e \quad \inf\{f(x) \mid x \in A\}\,,\quad \inf_{x\in A} f(x)\,,\quad \inf_A f\,.$$

La definizione è equivalente alle seguenti condizioni:

- f è limitata superiormente se $\exists k \in \mathbf{R}\ :\quad f(x) \leqslant k\ \forall x \in A$,
- f è limitata inferiormente se $\exists k \in \mathbf{R}\ :\quad f(x) \geqslant k\ \forall x \in A$,
- f è limitata se $\exists h, k \in \mathbf{R}\ :\quad h \leqslant f(x) \leqslant k\ \forall x \in A$.

L'ultima è esprimibile nella forma

$$\exists M \in \mathbf{R}\ :\quad |f(x)| \leqslant M \quad \forall x \in A\,.$$

Se f è limitata superiormente si ha

$$\sup_{x\in A} f(x) = L \Leftrightarrow \begin{cases} f(x) \leqslant L \quad \forall x \in A & L \text{ è un maggiorante} \\ \forall \varepsilon > 0\ \exists x \in A\ :\ f(x) > L - \varepsilon & L \text{ è il minimo dei maggioranti}\,, \end{cases}$$

altrimenti

$$\sup_{x\in A} f(x) = +\infty \Leftrightarrow \forall M \in \mathbf{R}\ \exists x \in A\ :\ f(x) > M\,.$$

Esercizio 3.10 - *Scrivere le proprietà analoghe per l'estremo inferiore.*

Esercizio 3.11 - *Se $f : A \to \mathbf{R}$ e $B \subset A$ allora*

$$\inf_A f \leqslant \inf_B f \leqslant \sup_B f \leqslant \sup_A f\,.$$

Esercizio 3.12 - *Una funzione $f : A \to \mathbf{R}$ è costante se e solo se $\inf_A f \geqslant \sup_A f$.*

Esercizio 3.13 - *Se $f : A \to \mathbf{R}$ allora*

$$\sup_A f = -\inf_A(-f) \qquad e \qquad \inf_A f = -\sup_A(-f)\,,$$

in particolare f è limitata superiormente se e solo se $-f$ lo è inferiormente e viceversa.

Esercizio 3.14 - *Se $f : A \to \mathbf{R}$ e $\lambda \in \mathbf{R}$ si ha*

$$\sup_{x\in A}(\lambda + f(x)) = \lambda + \sup_A f \qquad e \qquad \inf_{x\in A}(\lambda + f(x)) = \lambda + \inf_A f\,.$$

Inoltre

$$\sup_A(\lambda f) = \begin{cases} \lambda \sup_A f & se\ \lambda \geqslant 0 \\ \lambda \inf_A f & se\ \lambda < 0 \end{cases} \qquad e \qquad \inf_A(\lambda f) = \begin{cases} \lambda \inf_A f & se\ \lambda \geqslant 0 \\ \lambda \sup_A f & se\ \lambda < 0\,. \end{cases}$$

Esercizio 3.15 - *Se $f, g : A \to \mathbf{R}$ valgono le seguenti disuguaglianze*

$$\sup_A f + \inf_A g \leqslant \sup_A (f+g) \leqslant \sup_A f + \sup_A g \quad \textit{se } \sup_A f < +\infty \textit{ oppure } \inf_A g > -\infty$$
$$\inf_A f + \inf_A g \leqslant \inf_A (f+g) \leqslant \inf_A f + \sup_A g \quad \textit{se } \inf_A f > -\infty \textit{ oppure } \sup_A g < +\infty$$

con le convenzioni $\pm\infty + (\pm\infty) = \pm\infty$ e $c \pm \infty = \pm\infty$ per ogni $c \in \mathbf{R}$.

Indichiamo con $\mathscr{L}(A)$ l'insieme delle funzioni reali limitate su A.

Esercizio 3.16 - *Da quanto detto dedurre che $\mathscr{L}(A)$ è uno spazio vettoriale.*

Passiamo adesso agli estremi del prodotto di due funzioni adottando le convenzioni $\pm\infty \cdot (\pm\infty) = +\infty$, $\pm\infty \cdot (\mp\infty) = -\infty$ e $\pm\infty \cdot c = \pm\infty \cdot \operatorname{sign} c$ per ogni $c \in \mathbf{R}$.

Esercizio 3.17 - *Se $f, g : A \to [0, +\infty[$ valgono le disuguaglianze*

$$\inf_A f \cdot \inf_A g \leqslant \inf_A (fg) \leqslant \sup_A (fg) \leqslant \sup_A f \cdot \sup_A g$$

ed altre analoghe (quali?) con $f, g \leqslant 0$ e con $f \geqslant 0$ e $g \leqslant 0$ su A.

Esercizio 3.18 - *Dedurre che $\mathscr{L}(A)$ è un'algebra: $f, g \in \mathscr{L}(A) \Rightarrow fg \in \mathscr{L}(A)$.*

Esercizio 3.19 - *Se $f : A \to B$ e $g \in \mathscr{L}(B)$ allora $g \circ f \in \mathscr{L}(A)$ e $\sup_A g \circ f \leqslant \sup_B g$; vale l'uguaglianza se f è surgettiva.*

Si chiama *oscillazione* di f su A la quantità in $\overline{\mathbf{R}}$

$$\operatorname{osc}(f) = \sup_A f - \inf_A f\,.$$

Esercizio 3.20 - *Dimostrare che*

$$\operatorname{osc}(f) = \sup_{x,y \in A} |f(x) - f(y)|\,.$$

Definizione 3.9 - *Una funzione $f :]a, b[\to \mathbf{R}$ viene detta* ***localmente limitata*** *se è limitata su ogni intervallo chiuso e limitato $[\alpha, \beta]$ contenuto in $]a, b[$.*

In altre parole, una funzione su un intervallo aperto è localmente limitata se e solo se per ogni $\alpha, \beta \in \mathbf{R}$ tali che $a < \alpha < \beta < b$ esiste una costante $k(\alpha, \beta) \geqslant 0$ tale che

$$|f(x)| \leqslant k(\alpha, \beta) \quad \forall x \in [\alpha, \beta]\,.$$

Ovviamente se f è limitata lo è anche localmente, ma non è detto il viceversa. Ad esempio la funzione $f(x) = 1/x$ su $]0, +\infty[$ è non limitata ma localmente limitata.

Definizione 3.10 - *Un punto $x_0 \in A$ tale che $f(x_0) \leqslant f(x)$ per ogni $x \in A$ viene detto* ***punto di minimo*** *per f in A. Il* ***minimo*** *di f in A è il valore di f nel punto di minimo*

$$\min_{x \in A} f(x) = f(x_0)\,.$$

Analogamente $x_0 \in A$ è di ***massimo*** *per f in A se $f(x_0) \geqslant f(x)$ per ogni $x \in A$ e*

$$\max_{x \in A} f(x) = f(x_0)\,.$$

Esercizio 3.21 - *Dimostrare che $f : A \to \mathbf{R}$ ammette minimo in A se e solo se è limitata inferiormente ed esiste $x_0 \in A$ tale che $f(x_0) = \inf_A f$, per cui $\inf_A f = \min_A f$. Analogamente per il massimo.*

Esercizio 3.22 - *Se f ammette massimo e minimo su A e $\min_A f = \max_A f$ allora f è costante su A.*

Abbiamo visto che la somma di due funzioni limitate inferiormente e/o superiormente è limitata nello stesso senso. Tuttavia due funzioni possono avere minimo e/o massimo, ma non la loro somma. Ad esempio

$$f(x) = \begin{cases} (x+1)^2 & \text{se } x \in \mathbf{R} - \{0\} \\ 2 & \text{se } x = 0 \end{cases} \qquad \text{e} \qquad g(x) = \begin{cases} (x-1)^2 & \text{se } x \in \mathbf{R} - \{0\} \\ 2 & \text{se } x = 0 \end{cases}$$

hanno minimo rispettivamente in $x = -1$ e $x = 1$, ma la loro somma non ce l'ha.

Indichiamo con $I_\delta(x_0)$ l'intervallo aperto di centro x_0 e raggio δ, cioè $]x_0 - \delta, x_0 + \delta[$.

Definizione 3.11 - *Diciamo che $x_0 \in A$ è un punto di **minimo [massimo] relativo**, o **locale**, per la funzione $f : A \to \mathbf{R}$ se esiste un $\delta > 0$ tale che*

$$f(x_0) \leqslant f(x) \quad [f(x_0) \geqslant f(x)] \quad \forall x \in I_\delta(x_0) \cap A\,. \tag{3.11}$$

*Se le disuguaglianze valgono in senso stretto per $x \neq x_0$ diciamo che si tratta di un punto di minimo [massimo] relativo **isolato** o **proprio**.*

Ovviamente ogni punto di minimo [massimo] per f su A lo è anche in senso relativo.

Esercizio 3.23 - *Se ogni punto del dominio è di minimo relativo la funzione è costante? Se ogni punto del dominio è sia di massimo che di minimo relativo la funzione è costante?*

Più avanti useremo anche la seguente versione equivalente alla (3.11)

$$\frac{f(x) - f(x_0)}{x - x_0} \begin{cases} \geqslant 0 \ [\leqslant 0] & \forall x \in I_\delta^+(x_0) \cap A \\ \leqslant 0 \ [\geqslant 0] & \forall x \in I_\delta^-(x_0) \cap A \end{cases}$$

dove $I_\delta^-(x_0) = \{x \in I_\delta \mid x < x_0\}$ e $I_\delta^+(x_0) = \{x \in I_\delta \mid x > x_0\}$.

Per concludere con un esempio, vogliamo studiare i massimi e i minimi assoluti e relativi del polinomio $P(x) = x^4 - 4x^3 + 3x^2$. Ovviamente P ammette la radice doppia $x_0 = 0$ e le radici semplici $x_1 = 1$ e $x_2 = 3$. Poiché P assume valori positivi per $x < 1$ ed essendo $P(0) = 0$, x_0 è di minimo relativo, ma non assoluto in quanto $P(x) < 0$ per $1 < x < 3$. Inoltre P non è limitato superiormente in quanto $P(x) > x^4/2$ per ogni $x > 8$. Se $k < 0$ è il minimo, per determinarlo bisogna imporre la condizione che il polinomio $P(x) - k$ abbia una radice doppia, cioè

$$x^4 - 4x^3 + 3x^2 - k = (x - \alpha)^2(x^2 + \beta x + \gamma)$$

dove α è il punto di minimo. Per il principio d'identità dei polinomi si trova, oltre ad $\alpha = 0$ che non serve, anche $\alpha = (3 + \sqrt{3})/2$.

3.4 Funzioni monotone

La definizione di funzione monotona, crescente o decrescente, esige che dominio e codominio siano entrambi muniti di un ordinamento. Le regioni piane, ad esempio, sono ordinate per inclusione e la *misura*, che associa a ciascuna di esse la relativa area, è una funzione crescente: $A \subset B \Rightarrow area(A) \leqslant area(B)$. Comunque noi continueremo per il momento ad occuparci delle funzioni reali definite su domini di $\mathbf{R}$.

Definizione 3.12 - *Una funzione $f : A \to \mathbf{R}$, $A \subset \mathbf{R}$, viene detta* ***monotòna crescente*** *[**decrescente**] se*

$$\forall x_1, x_2 \in A \qquad x_1 < x_2 \Rightarrow f(x_1) \leqslant f(x_2) \quad [f(x_1) \geqslant f(x_2)] \tag{3.12}$$

e ***strettamente monotòna****, crescente o decrescente, se le disuguaglianze tra $f(x_1)$ e $f(x_2)$ valgono in senso stretto, caso in cui nella (3.12) vale ovviamente anche l'implicazione inversa.*

Una funzione allo stesso tempo crescente e decrescente è costante.

Esercizio 3.24 - *Scrivere la negazione della (3.12).*

Talvolta conviene scrivere la proprietà di f di essere crescente nelle forme equivalenti

$$[f(x_1) - f(x_2)](x_1 - x_2) \geqslant 0\,, \quad \frac{f(x_1) - f(x_2)}{x_1 - x_2} \geqslant 0 \quad \forall x_1, x_2 \in A\,,\ x_1 \neq x_2\,,$$

e analogamente nel caso decrescente.

Esercizio 3.25 - *Dimostrare che ogni funzione crescente [decrescente] $f : [a, b] \to \mathbf{R}$ ammette minimo [massimo] $f(a)$ e massimo [minimo] $f(b)$. Vale la stessa conclusione se al posto di $[a, b]$ si considera un dominio limitato qualunque $A \subset \mathbf{R}$ che ha per minimo a e per massimo b.*

Esercizio 3.26 - *Dimostrare che se f è monotona nello stesso verso sugli intervalli $[a, b]$ e $[b, c]$ allora è monotona su $[a, c]$.*

Esercizio 3.27 - *Costruire un esempio di funzione $f : A \cup B \to \mathbf{R}$ che sia crescente su A e crescente su B, ma non monotona su $A \cup B$.*

Esercizio 3.28 - *Verificare che se A e B sono, nell'ordine, due insiemi separati, $f_{|A}$ e $f_{|B}$ sono crescenti e $\sup_A f \leqslant \inf_B f$ (ad esempio se $f \leqslant 0$ su A e $f \geqslant 0$ su B) allora f è crescente su $A \cup B$.*

Esercizio 3.29 - *Verificare che ogni funzione $f : [a, b] \to \mathbf{R}$ monotona ammette massimo e minimo. Lo stesso si può dire se $[a, b]$ viene sostituito con un insieme A qualunque che sia munito di massimo e minimo.*

Esercizio 3.30 - *Dimostrare che ogni funzione $f :]a, b[\to \mathbf{R}$ monotona è localmente limitata.*

Esercizio 3.31 - *Se $f, g : A \to \mathbf{R}$ sono crescenti lo sono anche $f + g$, fg e, se $f : g(A) \to \mathbf{R}$, $f \circ g$? Che ipotesi bisogna fare perché lo siano? Costruire esempi e controesempi.*

Esercizio 3.32 - *Dimostrare che una funzione è monotona (indipendentemente dal verso della monotonia) se e solo se*

$$\forall x_1, x_2, x_3 \in A \qquad x_1 < x_2 < x_3 \Rightarrow [f(x_1) - f(x_2)][f(x_2) - f(x_3)] \geqslant 0\,.$$

Per negare questa proprietà dobbiamo trovare tre punti $x_1, x_2, x_3 \in A$ tali che $x_1 < x_2 < x_3$ e $[f(x_1) - f(x_2)][f(x_2) - f(x_3)] < 0$, per cui $f(x_1) < f(x_2)$ e $f(x_2) > f(x_3)$ oppure $f(x_1) > f(x_2)$ e $f(x_2) < f(x_3)$.

Ovviamente una funzione strettamente monotona è anche invertibile in quanto iniettiva e $f^{-1} : f(A) \to A$ è monotona nello stesso senso. La funzione

$$f(x) = \begin{cases} x & \text{se } 0 \leqslant x < 1 \\ 3 - x & \text{se } 1 \leqslant x < 2 \end{cases}$$

è iniettiva ma non monotona.

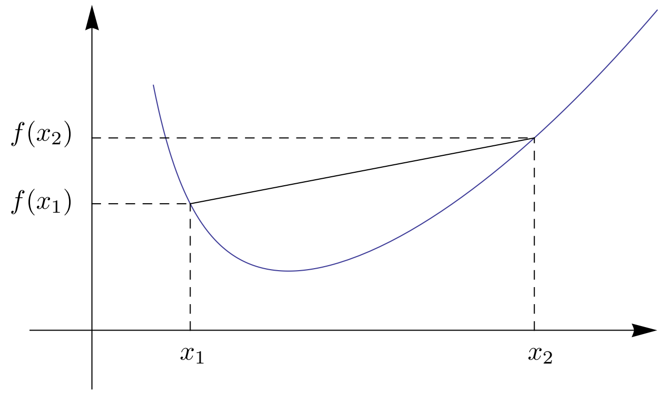
3.5 Funzioni convesse

Abbiamo già introdotto la nozione di funzione convessa in relazione alle potenze e l'abbiamo applicata a qualche esempio. In questo paragrafo la riprendiamo da un punto di vista generale trattandone le principali proprietà. Sia I un intervallo qualsiasi.

Definizione 3.13 - *Una funzione $f : I \to \mathbf{R}$ viene detta* ***convessa*** *se*

$$\forall x_1, x_2 \in I\,,\ x_1 < x_2\,,\quad f(x) \leqslant f(x_1) + \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1) \quad \forall x \in [x_1, x_2] \tag{3.13}$$

e ***strettamente convessa*** *se la disuguaglianza è stretta, con $x_1 \neq x_2$, per ogni $x \in]x_1, x_2[$. Se vale la disuguaglianza opposta, che equivale ad affermare che $-f$ è convessa, f si dice f* ***concava****, o* ***strettattamente concava****.*



Il grafico di una funzione convessa o concava, ma non nel senso stretto, presenta dei tratti rettilinei. Ovviamente una funzione è convessa su I se e solo se lo è la sua restrizione ad ogni intervallo contenuto in I. È facile invece costruire esempi di funzioni non convesse con restrizioni convesse, come $(x+1)^2 \wedge (x-1)^2$. Inoltre, essere concava non è la negazione di essere convessa, esistono funzioni non convesse, né concave, mentre quelle che sono simultaneamente convesse e concave, dovendo soddisfare entrambe le disuguaglianze nella (3.13), sono esattamente le funzioni del tipo $x \to ax + b$.

Usando l'applicazione $\lambda \to x(\lambda) = \lambda x_1 + (1-\lambda)x_2$, che in quanto strettamente monotona è bigettiva da $[0,1]$ in $[x_1, x_2]$, si verifica banalmente che f è convessa se e solo se

$$\forall x_1, x_2 \in I \quad f(\lambda x_1 + (1-\lambda)x_2) \leqslant \lambda f(x_1) + (1-\lambda) f(x_2) \quad \forall \lambda \in [0,1]\,, \tag{3.14}$$

oppure, in modo equivalente,

$$\forall x_1, x_2 \in I \quad f(\lambda x_1 + \mu x_2) \leqslant \lambda f(x_1) + \mu f(x_2) \quad \forall \lambda, \mu \in [0,1]\,,\ \lambda + \mu = 1\,, \tag{3.15}$$

con le ovvie modifiche per una funzione concava, anche nei casi in cui tali condizioni valgano nel senso stretto.

Ricordiamo che una figura piana E viene detta convessa se per ogni $P, Q \in E$ si ha $\lambda P + (1-\lambda)Q \in E$ per ogni $\lambda \in [0,1]$. In $\mathbf{R}$ gli insiemi convessi sono tutti e soli gli intervalli. Un'altra condizione ancora, necessaria e sufficiente per la convessità di f, è che il suo *epigrafico*

$$E^+(f) = \{(x,y) \in \mathbf{R}^2 \mid x \in I\,,\ y \geqslant f(x)\}$$

sia una regione convessa. Ovviamente f sarà concava se il suo *ipografico*

$$E^-(f) = \{(x,y) \in \mathbf{R}^2 \mid x \in I\,,\ y \leqslant f(x)\}$$

è convesso.

Esercizio 3.33 - *Dimostrare che $f : I \to \mathbf{R}$ è convessa se e solo se*

$$\forall x_1, x_2, \dots, x_n \in I \quad f\left(\sum_{i=1}^{n} \lambda_i x_i\right) \leqslant \sum_{i=1}^{n} \lambda_i f(x_i) \quad \forall \lambda_1, \lambda_2, \dots, \lambda_n > 0\,,\ \sum_{i=1}^{n} \lambda_i = 1\,.$$

Interpretando gli x_i come dei punti materiali di masse $m_i > 0$, posto $\lambda_i = m_i/M$, dove $M = \sum_i m_i$ è la massa totale, nell'esercizio si afferma che f è convessa se e solo se l'immagine del baricentro non supera il baricentro delle immagini.

Esercizio 3.34 - *Dimostrare che la media aritmetica μ_n di n numeri positivi $x_1, \dots, x_n$ è inferiore alla media quadratica*

$$\chi_n = \sqrt{\frac{x_1^2 + x_2^2 + \ldots + x_n^2}{n}}$$

Esercizio 3.35 - *Dimostrare che se $f : I \to \mathbf{R}$ è convessa allora per ogni $k \in \mathbf{R}$ il sottolivello*

$$f^{-1}]-\infty, k] = \{x \in I \mid f(x) \leqslant k\}$$

è un insieme convesso, quindi un intervallo. Vale l'implicazione contraria?

Esercizio 3.36 - *Dimostrare che una funzione lineare a tratti, cioè definita da $f(x) = a_h x + b_h$ su ogni intervallo del tipo $I_h = [x_{h-1}, x_h[$ di una famiglia finita, quindi con $h = 1, \dots, n$, è convessa su $\cup_h I_h$ se e solo se $a_h \leqslant a_{h+1}$ e $a_h x_h + b_h = a_{h+1} x_h + b_{h+1}$.*

Esercizio 3.37 - *Dimostrare che se $f, g : I \to \mathbf{R}$ sono convesse allora $f+g$ e $f \vee g$ sono convesse (in particolare è vero se g è lineare) e e kf è convessa se $k > 0$.*

Le funzioni fg e $f \wedge g$, invece, non sono in generale convesse, basta prendere $f(x) = (x-1)^2$ e $g(x) = (x+1)^2$.

Supporre che $f : I \to \mathbf{R}$ e $g : \mathbf{R} \to \mathbf{R}$ siano entrambe convesse non basta a garantire che la composizione $g \circ f$ sia convessa, basta prendere $f(x) = x^2 - 1$ e $g(y) = y^2$. Vi sono peraltro dei casi con f convessa, g concava e $g \circ f$ convessa.

Esercizio 3.38 - *Dimostrare che se f è convessa e g è convessa e crescente allora $g \circ f$ è convessa. Studiare nello stesso modo gli altri casi che si ottengono combinando la convessità/concavità con la crescenza/decrescenza.*

Un argomento importante in questo corso è lo studio dei massimi e minimi delle funzioni e la convessità, almeno intuitivamente, sembra una buona proprietà per i minimi (le funzioni concave per i massimi), sebbene da sola non ne garantisce l'esistenza.

Esercizio 3.39 - *Dimostrare che se un punto è di minimo relativo per una funzione convessa allora è di minimo assoluto.*

Sull'unicità abbiamo il seguente.

Proposizione 3.14 - *Una funzione $f : I \to \mathbf{R}$ strettamente convessa ammette al più un punto di minimo.*

Dimostrazione. Siano $x_1, x_2 \in I$ due punti di minimo distinti e $m = f(x_1) = f(x_2)$ il minimo della funzione. Se nella (3.14), intesa in senso stretto, si considera il punto medio, corrispondente a $\lambda = 1/2$, si ottiene

$$f\left(\frac{x_1 + x_2}{2}\right) < \frac{f(x_1) + f(x_2)}{2} = m\,,$$

ma questa disuguaglianza contraddice che m sia il minimo di f.

□

Vediamo un'altra caratterizzazione delle funzioni convesse che ha notevole importanza sia per lo studio dei minimi, sia per certe conseguenze che vedremo sulla regolarità. Per una $f : A \to \mathbf{R}$ qualunque, la funzione $R_{x_0} : A - \{x_0\} \to \mathbf{R}$ definita da

$$R_{x_0}(x) = \frac{f(x) - f(x_0)}{x - x_0} \qquad \forall x \in A - \{x_0\}$$

si chiama *rapporto incrementale* di f nel punto $x_0 \in A$.

Teorema 3.15 - *Una funzione $f : I \to \mathbf{R}$ è convessa se e solo se il rapporto incrementale di f è crescente su $I - \{x_0\}$.*

Dimostrazione. La tesi equivale ad affermare che per tre punti $x_1 < x_2 < x_3$ in I, detti m_{ij} i rapporti incrementali

$$m_{ij} = \frac{f(x_i) - f(x_j)}{x_i - x_j}\,,$$

si ha $m_{12} \leqslant m_{13} \leqslant m_{23}$. Per ottenerle basta applicare la definizione. Per esempio da

$$f(x_2) \leqslant f(x_1) + \frac{f(x_3) - f(x_1)}{x_3 - x_1}(x_2 - x_1)$$

segue

$$m_{12} = \frac{f(x_2) - f(x_1)}{x_2 - x_1} \leqslant \frac{f(x_3) - f(x_1)}{x_3 - x_1} = m_{13}$$

Analogamente la seconda discende da

$$f(x_2) \leqslant f(x_3) + \frac{f(x_1) - f(x_3)}{x_1 - x_3}(x_2 - x_3)\,.$$

Viceversa, se R_{x_1} è crescente, per ogni $x_1 < x < x_2$ si ha $R_{x_1}(x) \leqslant R_{x_1}(x_2)$ da cui segue subito la (3.13).

□

Corollario 3.16 - *Una funzione convessa $f : I \to \mathbf{R}$ che ammette $x_0 \in I$ come punto di minimo è decrescente per $x < x_0$ se $x_0 > \inf I$ e crescente per $x > x_0$ se $x_0 < \sup I$.*

Dimostrazione. Supponiamo $x_0 < \sup I$ e siano $x_1, x_2 \in I$ tali che $x_0 < x_1 < x_2$. Dalla definizione di punto di minimo e dal Teorema 3.15 segue

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} \geqslant \frac{f(x_1) - f(x_0)}{x_1 - x_0} \geqslant 0\,.$$

Sul complementare si ragiona in modo analogo.

□

Osserviamo che una funzione convessa su un intervallo può essere solo crescente o solo decrescente oppure prima decrescente e poi crescente. Nell'ultimo caso ammette minimo e un punto di minimo è $x_0 = \sup\{\alpha \mid f \text{ è decrescente per } x \leqslant \alpha\}$.

Corollario 3.17 - *Siano $f, g : I \to \mathbf{R}$ due funzioni convesse che ammettono rispettivamente x_1 e x_2 come punti di minimo. Se $x_1 < x_2$ la funzione convessa $f+g$ ammette minimo in almeno un punto $x_0 \in [x_1, x_2]$.*

Dimostrazione. Per $x < x_1$ f e g sono entrambe decrescenti e quindi lo è anche $f+g$ e per lo stesso motivo $f+g$ è crescente per $x > x_2$. Ne segue che ogni punto di minimo deve appartenere all'intervallo $[x_1, x_2]$.

□

Diamo adesso la definizione formale di una condizione che abbiamo già incontrato.

Definizione 3.18 - *Diciamo che $f : A \to \mathbf{R}$ soddisfa la condizione di **Lipschitz**, o che è **lipschitziana**, se esiste una costante $k \geqslant 0$ tale che*

$$|f(x_1) - f(x_2)| \leqslant k|x_1 - x_2| \qquad \forall x_1, x_2 \in A\,.$$

*Diciamo che $f :]a, b[\to \mathbf{R}$ è **localmente lipschitziana** se la sua restrizione ad ogni intervallo chiuso e limitato contenuto in $]a, b[$ è lipschitziana, cioè se per ogni $\alpha, \beta \in]a, b[$ esiste un numero reale $k(\alpha, \beta) \geqslant 0$ tale che*

$$|f(x_1) - f(x_2)| \leqslant k(\alpha, \beta)|x_1 - x_2| \qquad \forall x_1, x_2 \in [\alpha, \beta]\,.$$

Esercizio 3.40 - *Dimostrare che una funzione lipschitziana su un insieme limitato è limitata.*

Teorema 3.19 - *Ogni funzione convessa $f : I =]a, b[\to \mathbf{R}$ è localmente lipschitziana.*

Dimostrazione. Scelti $\alpha, \beta \in I$ e $x_1, x_2 \in [\alpha, \beta]$ tali che $x_1 < x_2$, dobbiamo stimare il valore assoluto del rapporto incrementale in x_1 e x_2. Se in esso facciamo variare x_2 fino ad α con x_1 fisso e x_1 fino a β con x_2 fisso e passiamo poi all'estremo inferiore a sinistra e superiore a destra si ottiene

$$\begin{aligned} k_1 = \inf_{x_1 > \alpha} \frac{f(\alpha) - f(x_1)}{\alpha - x_1} \leqslant \frac{f(\alpha) - f(x_1)}{\alpha - x_1} &\leqslant \frac{f(x_1) - f(x_2)}{x_1 - x_2} \\ &\leqslant \frac{f(\beta) - f(x_2)}{\beta - x_2} \leqslant \sup_{x_2 < \beta} \frac{f(\beta) - f(x_2)}{\beta - x_2} = k_2\,. \end{aligned}$$

I valori di k_1 e k_2 sono finiti perché gli intervalli $]a, \alpha[$ e $]\beta, b[$ hanno lunghezza positiva. Ne segue la tesi con $k = \max\{|k_1|, |k_2|\}$.

□

Il fatto che non valga in generale la condizione di Lipschitz fino gli estremi, nel caso $I = [a, b]$, è evidente ad esempio considerando la seguente funzione convessa

$$f(x) = \begin{cases} x^2 & \text{se } |x| < 1 \\ 2 & \text{se } |x| = 1\,. \end{cases}$$

Esercizio 3.41 - *Dimostrare che se $f : I \to J$ è bigettiva, essendo I un intervallo aperto e J un altro intervallo, allora*

$$f \text{ convessa e crescente} \Rightarrow f^{-1} \text{ concava e crescente}$$

$$f \text{ convessa e decrescente} \Rightarrow f^{-1} \text{ convessa e decrescente}\,.$$

Ricavare di conseguenza le analoghe proprietà con f concava.

3.6 Potenze e funzioni razionali fratte

Vogliamo definire le potenze con esponente $n \in \mathbf{Z}$. Se $n \geqslant 0$ ritroviamo quelle del § 3.2, se $n < 0$ poniamo

$$p_n(x) = \frac{1}{x^{-n}} \qquad \forall x \in \mathbf{R} - \{0\}\,.$$

Qui $\mathbf{R} - \{0\}$ è il *dominio naturale* di queste funzioni, inteso come il più grande sottoinsieme di $\mathbf{R}$ dove l'espressione che le definisce ha significato.

Nessuna potenza ad esponente negativo è limitata su $\mathbf{R} - \{0\}$, né su $]0, +\infty[$, né su $]-\infty, 0[$, né su un intervallo qualsiasi che ammette 0 come estremità. Ad esempio p_{-1} non è limitata su $]0, 1[$ perché per ogni $M > 1$ si ha $x < 1/M \Rightarrow 1/x > M$. Comunque quelle con esponente pari, in quanto positive, sono limitate inferiormente e ammettono 0 come estremo inferiore perché per ogni $\varepsilon \in \mathbf{R}$ tale che $0 < \varepsilon < 1$, scelto un $x > 1/\varepsilon$ si ha per $n > 0$

$$\frac{1}{x^{2n}} < \varepsilon^{2n} < \varepsilon\,.$$

D'altra parte non esiste nessun numero reale x tale che $p_{-2n}(x) = 0$, quindi nessuna di queste potenze ha minimo.

Sulla monotonia, dalle considerazioni fatte nel § 3.2 segue che

$$n \text{ pari} \Rightarrow \begin{cases} p_{n|]0,+\infty[} \text{ decrescente} \\ p_{n|]-\infty,0[} \text{ crescente} \end{cases} \qquad \text{e} \qquad n \text{ dispari} \Rightarrow \begin{cases} p_{n|]0,+\infty[} \text{ decrescente} \\ p_{n|]-\infty,0[} \text{ decrescente}\,, \end{cases}$$

ma nessuna di esse è monotona su $\mathbf{R} - \{0\}$.

Esercizio 3.42 - *Dimostrare che ogni potenza ad esponente intero negativo è localmente limitata su* $]-\infty, 0[$ *e su* $]0, +\infty[$.

Vediamo la convessità. Per dimostrare che la funzione p_{-1} è strettamente convessa su $]0, +\infty[$, basta osservare che per ogni $x_1, x_2 > 0$, con $x_1 \neq x_2$, e per ogni $\lambda \in]0, 1[$ la relazione

$$\frac{1}{\lambda x_1 + (1-\lambda)x_2} < \frac{\lambda}{x_1} + \frac{1-\lambda}{x_2}$$

è equivalente alla ben nota $2x_1x_2 < x_1^2 + x_2^2$. Componendo poi questa funzione con le varie potenze p_n, con $n > 0$, che sono convesse e crescenti su $]0, +\infty[$, si ottiene per l'Esercizio 3.38 che le $p_n(p_{-1}(x)) = 1/x^n$ sono tutte convesse su $]0, +\infty[$. Si vede poi facilmente che sono convesse anche sulla semiretta complementare quelle con n pari e concave quelle con n dispari.

Tutte queste funzioni sono anche localmente lipschitziane per il Teorema 3.19, ma possiamo verificarlo anche direttamente. Su $]0, +\infty[$, fissato $a > 0$, se $x_2 > x_1 \geqslant a$ si ha

$$\begin{aligned} \left|\frac{1}{x_1^n} - \frac{1}{x_2^n}\right| &= \frac{x_2^n - x_1^n}{x_1^n x_2^n} = \frac{(x_2 - x_1)(x_2^{n-1} + x_2^{n-2}x_1 + \ldots + x_1^{n-1})}{x_1^n x_2^n} \\ &\leqslant \frac{(x_2 - x_1)n x_2^{n-1}}{x_1^n x_2^n} = \frac{n(x_2 - x_1)}{x_1^n x_2} \leqslant \frac{n}{a^{n+1}}(x_2 - x_1)\,. \end{aligned}$$

La funzione

$$f(x) = \frac{1}{x - x^2} \qquad 0 < x < 1$$

è convessa e non limitata superiormente in quanto somma delle funzioni convesse $1/x$ e $1/(1-x)$ non limitate superiormente. È positiva e ammette minimo nel punto in cui

$x-x^2$ ha massimo, cioè per $x = 1/2$ e il minimo di f è $f(1/2) = 4$. Verifichiamo anche in questo caso la locale lipschitzianità. Se $0 < \alpha < \beta < 0$, per ogni $x_1, x_2 \in [\alpha, \beta]$ si ha

$$|f(x_1) - f(x_2)| = \frac{|x_1 + x_2 - 1|}{x_1 x_2 (1 - x_1)(1 - x_2)} |x_1 - x_2| \leqslant \frac{1}{\alpha^2 (1 - \beta)^2} |x_1 - x_2|$$

Siano P e Q due polinomi reali dei quali il secondo non identicamente nullo e sia $A = \{x \in \mathbf{R} \mid Q(x) \neq 0\}$, insieme che eventualmente coincide con $\mathbf{R}$ se Q non ha radici reali. La funzione $f : A \to \mathbf{R}$ definita da

$$f(x) = \frac{P(x)}{Q(x)} \qquad \forall x \in A$$

si chiama *funzione razionale*. Le potenze ad esponente intero sono casi particolari di funzioni razionali. Se Q è costante o un divisore di P si riduce ad un polinomio se opportunamente prolungata agli eventuali zeri di Q, quindi anche i polinomi sono funzioni razionali. Una volta ridotta ai minimi termini, gli zeri di Q sono i *punti singolari* e si chiamano *poli*.

Esercizio 3.43 - *Verificare che l'insieme delle funzioni razionali è uno spazio vettoriale e un campo.*

Come immediata conseguenza della (3.4), una funzione razionale si può sempre scrivere come somma di un polinomio con un'altra funzione razionale nella quale il grado del numeratore è inferiore a quello del denominatore. Esistono cioè due polinomi D e R, univocamente determinati, con $\mathrm{gr}(R) < \mathrm{gr}(Q)$, tali che

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

Consideriamo la funzione razionale

$$f(x) = \frac{ax + b}{cx + d} \qquad \forall x \neq -\frac{d}{c}$$

con $c \neq 0$. Se i numeri a, b, c e d sono direttamente proporzionali, condizione equivalente a $ad - bc = 0$, f si riduce ad una costante. Altrimenti

$$f(x) = \frac{a}{c} + \frac{bc - ad}{c(cx + d)} \qquad \forall x \neq -\frac{d}{c} \, .$$

Si tratta in questo caso di una potenza "riscalata" e "traslata" con esponente -1 e il suo grafico è un'iperbole equilatera centrata nel punto $(-d/c, a/c)$.

A titolo di esempio studiamo adesso qualche grafico. La funzione

$$f(x) = \frac{1 + x^3}{x^2} = x + \frac{1}{x^2}$$

ha come dominio naturale $\mathbf{R} - \{0\}$. Le restrizioni di f agli intervalli $]-\infty, 0[$ e $]0, +\infty[$ sono convesse in quanto ognuna è somma di una convessa con una lineare (ma di certo non è convessa su tutto il dominio!). Su $]-\infty, 0[$ è strettamente crescente essendo somma di funzioni strettamente crescenti e si annulla quindi solo per $x = -1$. Non è limitata inferiormente perché se $x < -2$ anche $x^3 < -2$, quindi $1/x^2 < -x/2$ e

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

ma non lo è neanche superiormente vicino a 0 perché per $x > -1/2$ anche $x^3 > -1/2$ e quindi

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

Su $]0, +\infty[$ è limitata inferiormente in quanto positiva, ma non lo è superiormente in quanto

$$f(x) > x \qquad \text{e} \qquad f(x) > \frac{1}{x^2}$$

dove a destra compaiono funzioni non limitate, quindi è ragionevole ritenere che abbia minimo (e sarà possibile dimostrarlo più avanti). Vediamo allora se esiste e cerchiamo eventualmente di calcolarlo. Se il minimo esiste, e se esiste è unico per la stretta convessità, tra le rette orizzontali $y = k$ ve ne deve essere una che tocca il grafico di f in un solo punto. Poiché l'equazione che ne deriva

$$1 + x^3 - kx^2 = 0$$

è di III grado, una soluzione reale esiste sempre ed è l'ascissa del punto d'incontro della retta con il ramo del grafico $x < 0$, per il punto di minimo basta imporre che sia una soluzione doppia. Indichiamo questa con $\alpha > 0$ e l'altra con $\beta < 0$. Dall'identità

$$x^3 - kx^2 + 1 = (x - \alpha)^2(x - \beta) \qquad \forall x \in \mathbf{R}$$

si deducono i valori $\alpha = 2/\sqrt[3]{4}$, $k = f(\alpha) = 3/\sqrt[3]{4}$, il valore di β non ci interessa.

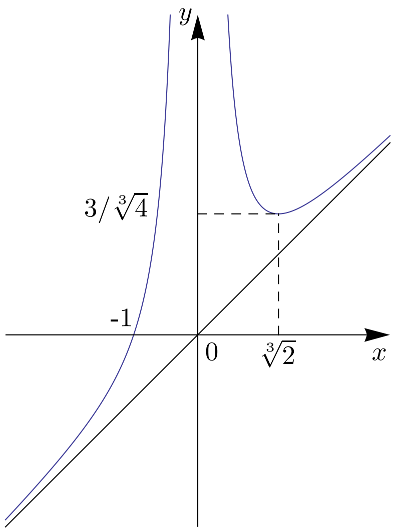


Figura 3.2: Il grafico della funzione $x \to \dfrac{x^3 + 1}{x^2}$.

Consideriamo adesso la coppia di funzioni

$$f(x) = \frac{2x}{1 + x^2} \qquad \text{e} \qquad g(x) = \frac{1 - x^2}{1 + x^2}\,, \qquad x \in \mathbf{R}\,.$$

La prima è dispari e l'altra pari, quindi basta studiarle su $[0, +\infty[$ e siccome soddisfano identicamente $f(x)^2 + g(x)^2 = 1$, con semplici considerazioni sul segno si può dedurre il grafico di g da quello di f.

Osserviamo che $f(0) = 0$ e $f(x) > 0$ per ogni $x > 0$, inoltre $f(x) \leqslant 1$ per la solita disuguaglianza $2x \leqslant 1+x^2$ e raggiunge il valore 1 per $x = 1$. Inoltre è crescente su $[0, 1]$ in quanto la disuguaglianza $f(x_1) - f(x_2) < 0$ è equivalente a $(x_1 - x_2)(1 - x_1x_2) < 0$ che è vera per $0 \leqslant x_1 < x_2 \leqslant 1$. Su $[1, +\infty[$ decresce fino a 0 asintoticamente essendo $f(x) = f(1/x)$. Nelle due figure sottostanti sono rappresentati i grafici di f e g.

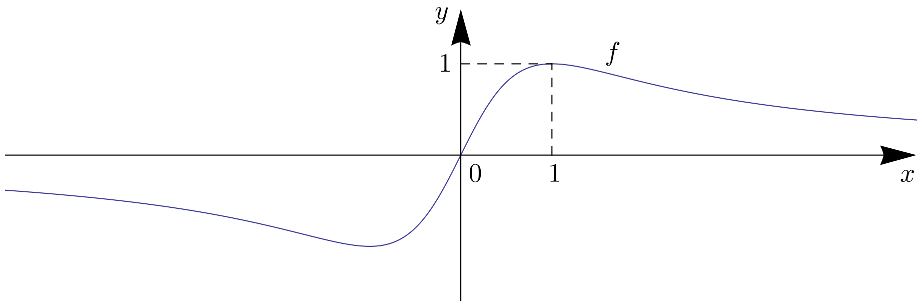


Figura 3.3: Il grafico della funzione f.

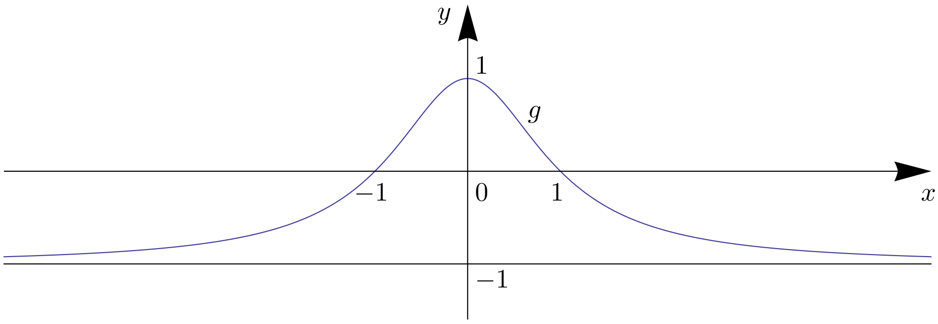


Figura 3.4: Il grafico della funzione g.

Esercizio 3.44 - *Tracciare i grafici delle funzioni reciproche di quelle viste in questi tre esempi.*

Una funzione razionale può non essere limitata per due motivi:

- perché il grado del numeratore supera quello del denominatore,
- perché il denominatore ha degli zeri (sono detti *poli* della funzione $f = P/Q$).

Se il denominatore non ha zeri ed ha grado superiore a quello del numeratore f è limitata. Per dimostrare la prima affermazione, supponiamo $P(x) = a_n x^n + \ldots + a_0$, con $a_n > 0$, e $Q(x) = b_m x^m + \ldots + b_0$. Se $n > m$, posto $\beta = m \max\{b_i\}$, dalla (3.10) segue

$$\left|\frac{P(x)}{Q(x)}\right| \geqslant \frac{a_n |x|^{n-m}}{2\beta} \qquad \forall x \in \mathbf{R} \ : \quad |x| \geqslant \max\left\{1, \frac{2Mn}{a_n}\right\}$$

e la nostra funzione non è limitata per $|x|$ "grande". Per la seconda, supponiamo che x_0 sia una radice di Q con molteplicità m_0. Allora $Q(x) = (x - x_0)^{m_0} Q_1(x)$ e siccome $P(x_0) \neq 0$ e $Q(x_0) \neq 0$, per la proprietà della permanenza del segno esiste $\delta > 0$ tale che $|P/Q|$ ha estremo inferiore $\gamma > 0$ su $]x_0 - \delta, x_0 + \delta[$. Su questo intervallo si ha

$$\left|\frac{P(x)}{Q(x)}\right| \geqslant \frac{\gamma}{|x - x_0|^{m_0}}$$

dove il II membro non è limitato.

3.7 Potenze e funzioni irrazionali algebriche

Problema 3.20 (della radice n-esima) - *Sia $n \geqslant 1$ un numero naturale. Dato un numero reale $a \geqslant 0$, trovare un numero reale $x \geqslant 0$ tale che $x^n = a$.*

Teorema 3.21 - *Il problema 3.20 ammette una ed una sola soluzione.*

Dimostrazione. L'unicità è dovuta al fatto che $p_n :]0, +\infty[\to]0, +\infty[$ è strettamente crescente, quindi iniettiva, su $]0, +\infty[$. Rimane da dimostrare che è surgettiva. Se $a = 0$ la soluzione è $x = 0$ e se $a = 1$ $x = 1$. Consideriamo il polinomio $P(x) = x^n - a$. Se $0 < a < 1$

$$P(0) = -a < 0 \quad \text{e} \quad P(1) = 1 - a > 0\,,$$

quindi esiste $x \in]0, 1[$ tale che $P(x) = 0$. Se $a > 1$

$$P(1) = 1 - a < 0 \quad \text{e} \quad P(a) = a^n - a > 0$$

e vale la stessa conclusione con $x > 1$.

□

La funzione p_n è dunque invertibile e la sua inversa $p_n^{-1} :]0, +\infty[\to]0, +\infty[$ si chiama *radice n-esima*. Oltre alla notazione $\sqrt[n]{x}$, si usa anche $x^{1/n}$ perché, essendo $p_n \circ p_n^{-1} = p_n^{-1} \circ p_n = i$, è coerente con la regola del prodotto degli esponenti $(x^{1/n})^n = (x^n)^{1/n} = x$. Si tratta dunque di un primo esempio di potenza con esponente razionale.

Se si affronta il problema con $a \in \mathbf{R}$ e ci chiediamo quali $x \in \mathbf{R}$ soddisfano l'equazione $x^n = a$ si vede subito che per n pari anche $-\sqrt[n]{a}$ è soluzione se $a > 0$ in quanto $(-\sqrt[n]{a})^n = a$, ma per $a < 0$ non vi sono soluzioni. Se invece n è dispari la soluzione è ancora unica, precisamente la stessa di prima se $a > 0$ e $-\sqrt[n]{-a}$ se $a < 0$, infatti

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

Per questo possiamo definire la radice n-esima di un numero $a < 0$, con n dispari, ponendo $\sqrt[n]{a} = -\sqrt[n]{-a}$.

Esercizio 3.45 - *Dimostrare che $\sqrt{x^2} = |x|$ per ogni $x \in \mathbf{R}$.*

Esercizio 3.46 - *Dimostrare che se $x, y \in \mathbf{R}$ sono concordi $\sqrt[n]{xy} = \sqrt[n]{|x|}\sqrt[n]{|y|}$ e mostrare di conseguenza che $\sqrt[n]{x^m} = (\sqrt[n]{x})^m$ per ogni $x > 0$ e per ogni $m \in \mathbf{Z}$ e $n \in \mathbf{N} - \{0\}$.*

Se dunque $(x^m)^{1/n} = (x^{1/n})^m$, questo valore comune è per definizione la *potenza ad esponente razionale* $p_r(x) = x^r = x^{m/n}$ essendo $r = m/n$ con $m, n \in \mathbf{Z}$ e $n \neq 0$.

Esercizio 3.47 - *Dimostrare che per ogni $r, s \in \mathbf{Q}$ si ha $x^r x^s = x^{r+s}$ che insieme a $x^0 = 1$ implica $x^{-r} = 1/x^r$. Dimostrare inoltre che $(x^r)^s = x^{rs}$.*

Essendo le potenze p_n tutte crescenti e convesse su $]0, +\infty[$, le loro inverse, le radici n-esime, devono essere crescenti e concave. Più in generale, p_r è crescente per $r > 0$ e decrescente per $r < 0$ ed è convessa per $r > 1$ e per $r < 0$ e concava per $0 < r < 1$. Infine, se $r < s$ allora $x^r < x^s$ se $x > 1$ e $x^r > x^s$ se $0 < x < 1$. I grafici di queste potenze sono simili a quelli con esponente intero.

Esercizio 3.48 - *Dimostrare che $\sqrt{x+y} \leqslant \sqrt{x} + \sqrt{y}$ per ogni $x, y \geqslant 0$ e si ha l'uguaglianza se e solo se uno dei numeri è nullo.*

Le espressioni

$$\nu = \left[\frac{1}{2}\left(\frac{1}{x} + \frac{1}{y}\right)\right]^{-1} = \frac{2xy}{x+y}\,, \quad \gamma = \sqrt{xy} \quad \text{e} \quad \mu = \frac{x+y}{2}$$

si chiamano rispettivamente *media armonica*, *media geometrica* e *media aritmetica* dei numeri reali $x, y > 0$.

Esercizio 3.49 - *Dimostrare che* $\nu \leqslant \gamma \leqslant \mu$.

Esercizio 3.50 - *Dimostrare che per ogni* $x, y \in \mathbf{R}$ *tali che* $x \neq y$ *si ha*

$$1 + xy < \sqrt{(1+x^2)(1+y^2)}\,.$$

Esercizio 3.51 - *Dimostrare che la funzione* $x/\sqrt{1-x^2}$ *è crescente su* $]-1,1[$.

Una funzione *algebrica irrazionale* è una combinazione finita, mediante somme e prodotti, di composizioni di potenze ad esponente razionale con funzioni razionali. Funzioni come

- $x \to \sqrt{1-x^2}\,, \quad x \in [0,1]\,,$ che ha per grafico l'intersezione della circonferenza unitaria col semipiano $y \geqslant 0$,
- $x \to \sqrt{1+x^2}\,, \quad x \in \mathbf{R}\,,$ che ha per grafico il ramo dell'iperbole equilatera $y^2 - x^2 = 1$ contenuto nel semipiano $y \geqslant 0$,
- $x \to \sqrt{x^2-1}\,, \quad |x| \geqslant 1\,,$ che ha per grafico i due mezzi rami dell'iperbole equilatera $x^2 - y^2 = 1$ del semipiano $y \geqslant 0$,

sono esempi di funzioni irrazionali algebriche. Di queste consideriamo la seconda (il grafico è ben noto dalla geometria analitica) e dimostriamo che è convessa e lipschitziana su $\mathbf{R}$. Dalla disuguaglianza dell'Esercizio3.50, scelti in $\mathbf{R}$ $x \neq y$ e $\lambda, \mu > 0$ con $\lambda + \mu = 1$, si ha

$$\begin{aligned}
\sqrt{1+(\lambda x+\mu y)^2} &= \sqrt{\lambda^2(1+x^2)+\mu^2(1+y^2)+2\lambda\mu(1+xy)} \\
&< \sqrt{\lambda^2(1+x^2)+\mu^2(1+y^2)+2\lambda\mu\sqrt{(1+x^2)(1+y^2)}} \\
&= \sqrt{(\lambda\sqrt{1+x^2}+\mu\sqrt{1+y^2})^2} = \lambda\sqrt{1+x^2}+\mu\sqrt{1+y^2}\,.
\end{aligned}$$

Per vedere che è lipschitziana, una via geometrica è suggerita dalla Figura 3.5 dove la differenza dei due lati $\sqrt{1+y^2} - \sqrt{1+x^2}$ del triangolo è minore del terzo $y - x$.

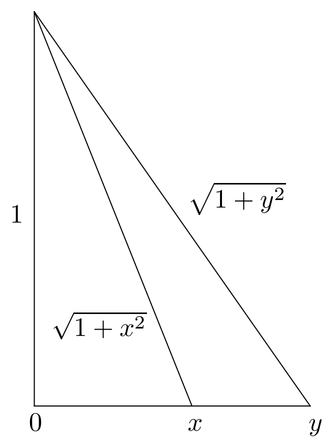


Figura 3.5: Vedere il triangolo di base $|x-y|$.

Altrimenti per via algebrica

$$\left|\sqrt{1+y^2} - \sqrt{1+x^2}\right| = \frac{|y^2-x^2|}{\sqrt{1+y^2}+\sqrt{1+x^2}} < \frac{|x+y|}{|x|+|y|}|x-y| \leqslant |x-y|\,, \quad x \neq y\,.$$

Esercizio 3.52 - *Dimostrare che $x \to \sqrt{x}$ è lipschitziana su ogni intervallo $[a, +\infty[$ con $a > 0$.*

Esercizio 3.53 - *Dimostrare la disuguaglianza*

$$|\sqrt{x} - \sqrt{y}| \leqslant \sqrt{|x - y|} \qquad \forall x, y \geqslant 0\,.$$

Nell'Esercizio 3.53 si afferma che la funzione $x \to \sqrt{x}$ soddisfa la condizione di Hölder con esponente $\alpha = 1/2$ nel senso della seguente definizione.

Definizione 3.22 - *Una funzione $f : A \in \mathbf{R}$ viene detta **hölderiana** con esponente $\alpha \in]0, 1[$ se esiste una costante $k \geqslant 0$ tale che*

$$|f(x_1) - f(x_2)| \leqslant k|x_1 - x_2|^\alpha \qquad \forall x_1, x_2 \in A\,. \tag{3.16}$$

*La condizione (3.16) vale **localmente** su un intervallo $]a, b[$ se è soddisfatta su ogni intervallo $[\alpha, \beta] \subset]a, b[$ con $k = k(\alpha, \beta)$.*

Ogni potenza p_r è hölderiana su $[0, +\infty[$ se $0 < r < 1$ e per ogni $r \in \mathbf{Q}$ è localmente lipschitziana su $]0, +\infty[$.

3.8 Potenze, funzioni esponenziali e logaritmiche

Vogliamo adesso procedere con un'ultima estensione (non si può mai dire, più avanti vedremo le potenze complesse): dagli esponenti razionali agli esponenti reali. Nella definizione di x^y con $x, y \in \mathbf{R}$ supporremo sempre $x > 0$, mentre su y non vi saranno restrizioni.

Poniamo $1^y = 1$ per ogni $y \in \mathbf{R}$. Se $x \neq 1$ e $y \in \mathbf{R}$ le considerazioni del § 3.7 sul comportamento delle potenze al variare degli esponenti ci dicono che le due classi di numeri

$$A = \{x^r \mid r \in \mathbf{Q}\,,\ r \leqslant y\} \qquad \text{e} \qquad B = \{x^s \mid s \in \mathbf{Q}\,,\ s \geqslant y\}$$

sono certamente separate in $\mathbf{R}$, in particolare A precede B se $x > 1$ e B precede A se $x < 1$. Tra poco dimostriamo che A e B sono anche contigue, pertanto ammettono un unico elemento separatore che naturalmente verrà indicato con x^y. Che si tratti di un'estensione è evidente, se $y \in Q$ $A \cap B = \{x^y\}$. Tanto per fissare un caso, supporremo $x > 1$, per l'altro si può ragionare in modo analogo oppure ricondursi al precedente ponendo $x^y = (1/x)^{-y}$.

Lemma 3.23 - *Per ogni $x > 1$ si ha*

$$\inf_{n \geqslant 1} \sqrt[n]{x} = 1\,.$$

Dimostrazione. Già sappiamo che 1 è una limitazione inferiore, quindi

$$t_n = \sqrt[n]{x} - 1 > 0 \qquad \forall n \geqslant 1\,.$$

D'altra parte per la disuguaglianza di Bernoulli si ha

$$x = (1 + t_n)^n \geqslant 1 + nt_n \qquad \forall n \geqslant 1$$

da cui

$$t_n \leqslant \frac{x - 1}{n} \qquad \forall n \geqslant 1$$

dove il II membro ha estremo inferiore nullo. Allora

$$\inf_{n \geqslant 1} t_n = \inf_{n \geqslant 1} (\sqrt[n]{x} - 1) = 0\,.$$

□

Teorema 3.24 *Le due classi A e B sono contigue.*

Dimostrazione. Sia M un maggiorante dell'insieme A (che è limitato superiormente) e fissiamo un $\varepsilon > 0$ ad arbitrio. Per il Lemma 3.23 esiste $k \in \mathbf{N}$ tale che $x^{1/k} - 1 < \varepsilon/M$. Scegliamo allora due numeri $r, s \in \mathbf{Q}$ tali che $r \leqslant y \leqslant s$ e $s - r < 1/k$. Così $x^r \in A$, $x^s \in B$ e

$$x^s - x^r = x^r(x^{s-r} - 1) \leqslant M(x^{1/k} - 1) < M\frac{\varepsilon}{M} = \varepsilon\,.$$

Dunque A e B sono contigue.

□

Applicando direttamente la definizione dimostriamo la proprietà fondamentale $x^{y_1+y_2} = x^{y_1}x^{y_2}$, $y_1, y_2 \in \mathbf{R}$. Presi due numeri razionali $r_1 \leqslant y_1$ e $r_2 \leqslant y_2$, poiché $x^{r_1} \leqslant x^{y_1}$ e $x^{r_2} \leqslant x^{y_2}$, si ha

$$x^{r_1+r_2} = x^{r_1}x^{r_2} \leqslant x^{y_1}x^{y_2}$$

e tenendo presente che $r_1 + r_2 \leqslant y_1 + y_2$, passando all'estremo superiore a sinistra si ottiene

$$x^{y_1+y_2} \leqslant x^{y_1}x^{y_2}\,.$$

L'altra disuguaglianza si ottiene con lo stesso ragionamento passando all'estremo inferiore con $s_1 \geqslant y_1$ e $s_2 \geqslant y_2$. Mettendo insieme questa proprietà con $x^0 = 1$ si ricava $x^{-y} = 1/x^y$.

L'espressione x^y dipende dalle due variabili x e y, ma fissandone una dipende solo dall'altra. Sono pertanto ben definite le due funzioni

$$x \to p_y(x) = x^y\,,\ x \in]0, +\infty[\,,\ \text{fissato } y \in \mathbf{R}$$

$$y \to \exp_x y = x^y\,,\ y \in \mathbf{R}\,,\ \text{fissato } x \in]0, +\infty[\,.$$

La prima è la *potenza ad esponente reale y* e la seconda è l'*esponenziale di base x*. Le proprietà qualitative della prima sono simili a quelle delle potenze con esponente razionale, monotonia, convessità e per $y \neq 0$ $p_y^{-1} = p_{1/y}$. Riguardo la seconda, è evidente che cresce se $x > 1$ e decresce se $0 < x < 1$ (ed è costante se $x = 1$). Infatti, con $x > 1$, se $y_1 < y_2$ le corrispondenti coppie di classi contigue (A_1, B_1) e (A_2, B_2) con esponenti razionali soddisfano le relazioni di inclusione in senso stretto $A_1 \subset A_2$ e $B_1 \supset B_2$, quindi

$$x^{y_1} = \sup A_1 = \inf B_1 < \sup A_2 = \inf B_2 = x^{y_2}\,.$$

Inoltre si verifica facilmente che le funzioni esponenziali sono surgettive e più avanti vedremo che sono convesse e di conseguenza localmente lipschitziane.

Poiché nella funzione esponenziale si considera fissata la base, da ora in poi la indicheremo con a^x e talvolta con $\exp_a x$. Per ogni $a > 0$, $a \neq 1$, la funzione esponenziale $\exp_a : \mathbf{R} \to]0, +\infty[$ è invertibile, la funzione inversa si chiama *logaritmo in base a* e si indica con $x \to \log_a x$. La funzione $\log_a :]0, +\infty[\to \mathbf{R}$ è pertanto crescente e concava per $a > 1$ e decrescente e convessa per $0 < a < 1$, inoltre

$$\log_a x_1x_2 = \log_a x_1 + \log_a x_2\,, \quad \log_a 1 = 0\,, \quad \log_a a = 1\,.$$

Esercizio 3.54 - *Dimostrare le identità*
(log)1 $\log_a b \log_b c = \log_a c \quad \forall a, b, c > 0\,,\ a, b \neq 1\,,$
(log)2 $\log_a b^c = c \log_a b \quad \forall a, b > 0\,,\ a \neq 1\,.$

I grafici delle funzioni a^x e $\log_a x$ sono illustrati nella Figura 3.6.

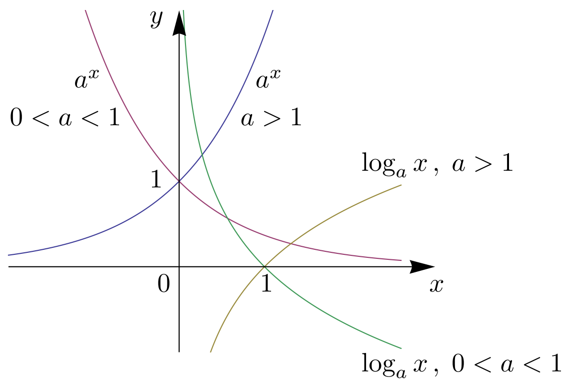


Figura 3.6: I grafici di a^x e $\log_a x$ a seconda di a.

3.9 Funzioni iperboliche

Ricordiamo che se $A \subset \mathbf{R}$ è un insieme simmetrico rispetto a 0 una funzione $f : A \to \mathbf{R}$ è pari se $f(x) = f(-x)$ e dispari se $f(x) = -f(-x)$ per ogni $x \in A$. Ogni funzione su A può essere decomposta in modo unico nella somma di una funzione pari f_p, la *parte pari* di f, e di una funzione dispari f_d che ne è la *parte dispari*

$$f_p(x) = \frac{f(x) + f(-x)}{2} \qquad \text{e} \qquad f_d(x) = \frac{f(x) - f(-x)}{2} \qquad \forall x \in A\,.$$

Per l'unicità, se $f = \varphi + \psi$ con φ pari e ψ dispari allora

$$\begin{cases} f(x) = \varphi(x) + \psi(x) \\ f(-x) = \varphi(x) + \psi(-x)\,, \end{cases}$$

da cui, per somma e differenza membro a membro, si ottiene $\varphi = f_p$ e $\psi = f_d$.

Dato $a > 0$, le parti pari e dispari di a^x

$$\cosh_a x = \frac{a^x + a^{-x}}{2} \qquad \text{e} \qquad \operatorname{senh} x = \frac{a^x - a^{-x}}{2} \qquad \forall x \in \mathbf{R}$$

si chiamano *coseno iperbolico* e *seno iperbolico* di x in base a ($\cosh_1 x \equiv 1$ e $\operatorname{senh}_1 x \equiv 0$). I nomi di questi funzioni sono dovuti a certe relazioni, di verifica immediata, che richiamano quelle delle funzioni trigonometriche cos e sen, anch'esse una pari e l'altra dispari. Vale ad esempio l'identità

$$\cosh_a^2 x - \operatorname{senh}_a^2 x = 1 \qquad \forall x \in \mathbf{R}$$

e le formule di addizione

$$\begin{aligned} \cosh_a(x+y) &= \cosh_a x \cosh_a y + \operatorname{senh}_a x \operatorname{senh}_a y \\ \operatorname{senh}_a(x+y) &= \operatorname{senh}_a x \cosh_a y + \cosh_a x \operatorname{senh}_a y \qquad \forall x, y \in \mathbf{R} \end{aligned}$$

con tutte le conseguenze (duplicazione, bisezione ecc.). L'aggettivo “iperbolico” deriva dal fatto che le curve piane $(\cosh_a t, \operatorname{senh}_a t)$ e $(\operatorname{senh}_a t, \cosh_a t)$, $t \in \mathbf{R}$, sono rami di iperbole.

Osserviamo che $\cosh_a x$ è una funzione positiva, non limitata superiormente e strettamente convessa in quanto somma di funzioni di questo tipo, ed essendo pari, ha minimo in 0 dove assume il valore 1. La funzione $\operatorname{senh}_a x$ non è limitata né in un senso, né nell'altro ed è strettamente crescente se $a > 1$, decrescente se $0 < a < 1$, è dispari e si annulla solo in 0. Messe a confronto, la differenza tra le due funzioni è

$$\cosh_a x - \operatorname{senh}_a x = a^{-x} > 0 \text{ e decrescente [crescente] se } a > 0\ [a < 0]\,,$$

infine

$$\cosh_{1/a} x = \cosh_a x \quad \text{e} \quad \operatorname{senh}_{1/a} x = -\operatorname{senh}_a x \qquad \forall x \in \mathbf{R}\,.$$

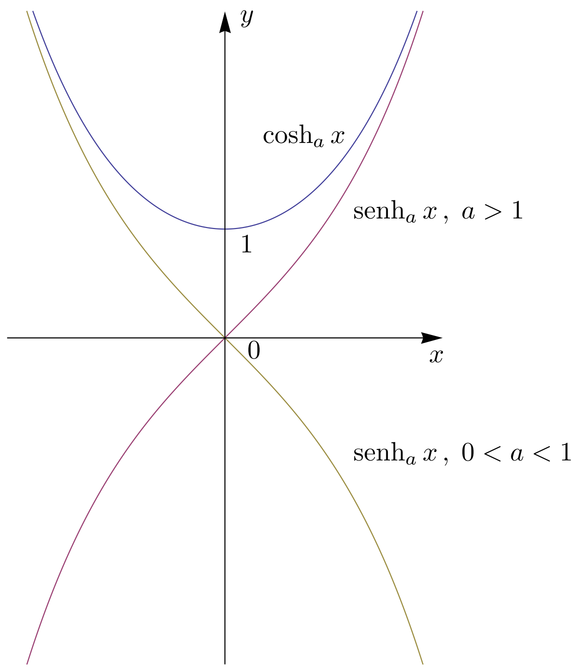


Figura 3.7: I grafici di $\cosh_a x$ e $\operatorname{senh}_a x$ a seconda di a.

La *tangente iperbolica* è il rapporto

$$\operatorname{tangh}_a x = \frac{\operatorname{senh}_a x}{\cosh_a x}$$

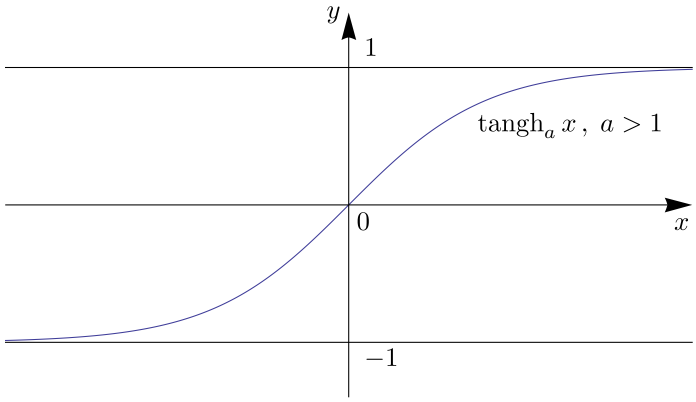


Figura 3.8: Il grafico della $\operatorname{tangh}_a x$ con $a > 1$.

Esercizio 3.55 - *Disegnare il grafico delle funzioni* $\operatorname{tangh}_a x$, *con* $0 < a < 1$, *e* $\coth_a x = \cosh_a x / \operatorname{senh}_a x$.

Abbiamo osservato che $\operatorname{senh}_a : \mathbf{R} \to \mathbf{R}$ è strettamente monotona e dato che è anche surgettiva (questo deriva dal fatto che lo è la funzione esponenziale) è invertibile.

L'inversa si trova risolvendo l'equazione $\operatorname{senh}_a x = y$ rispetto a x che è equivalente a

$$a^{2x} - 2ya^x - 1 = 0\,.$$

Si tratta di un'equazione di II grado in a^x con $\Delta = y^2 + 1 > 0$ che presenta una variazione e una permanenza. Essendo $a^x > 0$, ammette solo la soluzione

$$a^x = y + \sqrt{y^2+1}\,,$$

da cui

$$x = \log_a(y + \sqrt{y^2+1})\,.$$

Si chiama *settore iperbolico* la funzione inversa del seno iperbolico

$$\operatorname{settsenh}_a x = \log(x + \sqrt{x^2+1}) \qquad x \in \mathbf{R}\,.$$

Il nome deriva dal fatto che assume il valore dell'area del settore di iperbole.

Esercizio 3.56 - *Trovare le funzioni inverse del* $\cosh_a$ *su* $[0, +\infty[$ *e su* $]-\infty, 0]$ *separatamente.*

Per una lista completa delle funzioni iperboliche e delle loro inverse si può consultare wikipedia.

3.10 Funzioni trigonometriche e funzioni periodiche

È ben noto dalla Geometria che la circonferenza è rettificabile, la sua lunghezza è l'elemento separatore tra i perimetri dei poligoni ad essa inscritti e quelli dei poligoni circoscritti e il rapporto tra un arco e il raggio dipende solo dall'angolo al centro su cui l'arco insiste, ma non dalla scelta della circonferenza. Ne segue che è possibile misurare un angolo ϑ usando la lunghezza del relativo arco s

$$\vartheta = \frac{s}{r}$$

a partire da un punto della circonferenza a cui si fa corrispondere per convenzione il valore $s = 0$. La variabile $\vartheta \in \mathbf{R}$ è quindi adimensionale e viene a coincidere con la lunghezza d'arco (anch'essa adimensionale) nel caso della circonferenza unitaria ($r = 1$). Il numero 2π esprime il rapporto tra l'intera circonferenza e il suo raggio. Le prime tracce nella storia di questo numero risalgono ai babilonesi e agli egizi e fu poi Archimede di Siracusa il primo a calcolarne le prime due cifre decimali. Ma fu Lambert a dimostrare nel 1761 che si tratta di un numero irrazionale e successivamente, nel 1882, Lindemann che è trascendente.

Consideriamo nel piano cartesiano la circonferenza unitaria $\mathscr{C}$ centrata in O, detta anche *circonferenza goniometrica*, con l'arco/angolo $s = \vartheta$ orientato nel senso antiorario a partire dal punto $(1, 0)$. Ad ogni valore di $\vartheta \in \mathbf{R}$ corrisponde un solo punto di $P = (x, y) \in \mathscr{C}$ di intersezione della semiretta uscente da O che forma l'angolo ϑ col semiasse positivo delle ascisse. Le coordinate di P sono una coppia di funzioni di ϑ, il *coseno* e il *seno* di ϑ

$$\vartheta \to (x(\vartheta, y(\vartheta)) = (\cos\vartheta, \operatorname{sen}\vartheta)\,,$$

la prima è una funzione pari e l'altra dispari. Su una circonferenza ad essa concentrica di raggio r si ha

$$\begin{cases} x(\vartheta) = r\cos\vartheta \\ y(\vartheta) = r\operatorname{sen}\vartheta \end{cases} \tag{3.17}$$

dove $r > 0$ è la distanza (costante) del punto P da O. Incrementando ϑ di un multiplo intero di 2π si ottiene nuovamente lo stesso punto di $\mathscr{C}$, quindi le funzioni introdotte sono 2π-*periodiche*, vediamo che cosa vuol dire.

Definizione 3.25 - *Siano $T > 0$ un numero reale e $A \subset \mathbf{R}$ tale che*

$$x \in A \Rightarrow x + T \in A\,.$$

*Una funzione $f : A \to \mathbf{R}$ è detta T-**periodica**, o **periodica di periodo** $T > 0$, se*

$$f(x+T) = f(x) \qquad \forall x \in A \tag{3.18}$$

e se questa condizione diventa falsa sostituendo T con un qualsiasi numero reale τ tale che $0 < \tau < T$.

Ovviamente se $f : A \to \mathbf{R}$ è T-periodica si ha $f(x) = f(x + nT)$ per ogni $n \in \mathbf{Z}$ e per ogni $x \in A$. Osserviamo che le funzioni costanti su $\mathbf{R}$ non sono periodiche perché soddisfano sì la (3.18), ma con qualunque $T > 0$. Neanche la funzione

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

può essere accettata come periodica perché l'estremo inferiore dei $T \in \mathbf{Q}$ per cui è soddisfatta la (3.18) è nullo.

Per studiare una funzione T-periodica su $\mathbf{R}$, o su un suo sottoinsieme, basta studiarne la restrizione ad un intervallo qualunque di ampiezza T, per esempio $[0,T]$ o $[-T/2, T/2]$, e replicare il suo grafico per traslazioni di entità nT con $n \in \mathbf{Z}$. Questa operazione permette di ottenere il prolungamento periodico di ogni funzione definita inizialmente su un dominio limitato, il periodo sarà pari all'ampiezza del più piccolo intervallo che lo contiene.

Esercizio 3.57 - *Dimostrare che se f è T-periodica allora $x \to f(\lambda x)$ è, per ogni $\lambda \in \mathbf{R} - \{0\}$, periodica di periodo $T/|\lambda|$. Calcolare il periodo delle funzioni $\cos \pi x$ e $\operatorname{sen} \pi x$, e, più in generale, delle $\cos \omega x$ e $\operatorname{sen} \omega x$.*

La somma di due funzioni periodiche non è detto che sia periodica, non lo sono ad esempio $\operatorname{sen} x + \operatorname{sen} \pi x$ e $\operatorname{sen}^2 x + \cos^2 x$ per motivi diversi, e se lo è può avere periodo più piccolo di ciascun addendo, come nel caso delle funzioni 2π-periodiche $\operatorname{sen} x$ e $\operatorname{sen} 2x - \operatorname{sen} x$ la cui somma ha periodo π, o anche più grande come $\operatorname{sen} x/3 + \operatorname{sen} x/2$ che è 12π-periodica. Lo stesso avviene per il prodotto: $\operatorname{sen} x \cos x$ ha periodo π, ma $\operatorname{sen} x(1 - \cos x)$ ha periodo 2π. Somme e prodotti non costanti di funzioni con lo stesso periodo T sono periodici con periodo non superiore a T, ma in generale non si possono fare previsioni. È facile invece dimostrare che se le due funzioni hanno periodi commensurabili T_1 e T_2 distinti la loro somma è periodica di periodo pari al minimo comune multiplo fra T_1 e T_2. Precisamente se f è T_1-periodica e g è T_2-periodica, per semplicità definite su tutto $\mathbf{R}$, e $T_1/T_2 = m/n \in \mathbf{Q}$, con n, m primi tra loro, la funzione $h = f + g$ è periodica di periodo al più $T = nT_1 = mT_2$, infatti per ogni $x \in \mathbf{R}$

$$h(x) = f(x) + g(x) = f(x + nT_1) + g(x + mT_2) = f(x + T) + g(x + T) = h(x + T)\,.$$

Lo stesso vale per il prodotto e il quoziente. Le funzioni $\operatorname{sen} 3x$ e $\operatorname{sen} 5x$ hanno periodi rispettivi $2\pi/3$ e $2\pi/5$ il cui minimo comune multiplo è 2π, ma il loro prodotto è la funzione π-periodica $(\cos 2x - \cos 8x)/2$.

Esercizio 3.58 - *Dimostrare che se $f : \mathbf{R} \to \mathbf{R}$ è periodica allora $g \circ f$ è periodica per ogni $g : \mathbf{R} \to \mathbf{R}$ non costante.*

Tornando alle nostre funzioni trigonometriche, non intendiamo ricavare in questa sede le usuali formule trigonometriche, le diamo per scontate e pronte per l'uso, ci limitiamo invece a richiamare qualche loro proprietà rilevante e ad illustrarne alcuni aspetti forse non del tutto noti. Per come sono state introdotte, diciamo per costruzione, si ha subito la relazione fondamentale

$$\cos^2 x + \operatorname{sen}^2 x = 1 \qquad \forall x \in \mathbf{R}$$

da cui segue

$$|\cos x| \leqslant 1 \quad \text{e} \quad |\operatorname{sen} x| \leqslant 1 \qquad \forall x \in \mathbf{R}\,,$$

si tratta di funzioni dunque limitate a valori in $[-1,1]$. Per via geometrica si vede subito che sono anche surgettive (se è vero che ad ogni punto della circonferenza goniometrica corrisponde un angolo). Riguardo l'iniettività, e quindi l'esistenza delle funzioni inverse, dobbiamo individuare intervalli in cui sono monotone perché a causa della periodicità non saranno certamente invertibili su $\mathbf{R}$. Si è convenuto allora di definire le inverse, *arcocoseno* e *arcoseno*, su dei particolari intervalli

$$\operatorname{arcos} : [-1,1] \to [0, 2\pi] \qquad \text{e} \qquad \operatorname{arcsen} : [-1,1] \to \left[-\frac{\pi}{2}, \frac{\pi}{2}\right] .$$

I grafici di tutte le funzioni trigonometriche e delle loro inverse sono reperibili ovunque e non li riportiamo. Un'altra relazione

$$\cos\left(x - \frac{\pi}{2}\right) = \operatorname{sen} x$$

ci mostra che una può essere ottenuta dall'altra con una semplice traslazione.

La *tangente* è la funzione trigonometrica

$$\operatorname{tang} x = \frac{\operatorname{sen} x}{\cos x} .$$

Essa è π-periodica su $\mathbf{R} - \left\{\frac{\pi}{2} + n\pi \mid n \in \mathbf{Z}\right\}$ e strettamente crescente sull'intervallo $]-\pi/2, \pi/2[$ dove ammette come inversa la funzione *arcotangente*

$$\operatorname{arctg} : \mathbf{R} \to \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[.$$

Il nome deriva da una semplice considerazione geometrica. Si consideri la retta r tangente alla circonferenza unitaria nel punto $(1,0)$. La retta passante per O e $P = (1,t) \in r$ ha coefficiente angolare t e interseca $\mathscr{C}$ nei due punti

$$\begin{cases} x(t) = \dfrac{\pm 1}{\sqrt{1+t^2}} \\ y(t) = \dfrac{\pm t}{\sqrt{1+t^2}}\,, \end{cases} \quad t \in \mathbf{R}\,,$$

equazioni che parametrizzano $\mathscr{C} - \{(0, \pm 1)\}$ e che riconosciamo come le formule che legano $\operatorname{sen} x$ e $\cos x$ a $\operatorname{tang} x$.

Analogamente la funzione *cotangente* è definita da

$$\operatorname{cotg} x = \frac{\cos x}{\operatorname{sen} x} = \operatorname{tang}\left(\frac{\pi}{2} - x\right) = -\operatorname{tang}\left(\frac{\pi}{2} + x\right)\,, \qquad x \in \mathbf{R} - \{n\pi \mid n \in \mathbf{Z}\}\,,$$

strettamente decrescente su $]0, \pi[$ e con inversa la funzione $\operatorname{arcotg} : \mathbf{R} \to]0, \pi[$.

Esercizio 3.59 - *Servendosi di un foglio a quadretti dimostrare che*

$$\operatorname{arctg} 1 + \operatorname{arctg} 2 + \operatorname{arctg} 3 = \pi .$$

Esercizio 3.60 - *Dimostrare che la funzione* $x \to x/\sqrt{1+x^2}$ *è strettamente crescente (è composizione di funzioni crescenti, quali?).*

Confrontando corde archi e tangenti, si ricavano le seguenti importanti relazioni

$$|\operatorname{sen} x| \leqslant |x| \quad \forall x \in \mathbf{R} \qquad \text{e} \qquad |\operatorname{sen} x| \leqslant |x| \leqslant |\operatorname{tang} x| \quad \forall x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\,.$$

Dalla prima segue che la funzione $\operatorname{sen} x$, e quindi anche $\cos x$, è lipschitziana su $\mathbf{R}$

$$|\operatorname{sen} x - \operatorname{sen} y| = 2\left|\cos\frac{x+y}{2}\operatorname{sen}\frac{x-y}{2}\right| \leqslant 2\left|\operatorname{sen}\frac{x-y}{2}\right| \leqslant |x-y|\,.$$

Una conseguenza è che la funzione $x + \operatorname{sen} x$ è strettamente crescente su $\mathbf{R}$, infatti

$$x < y \Rightarrow x - y < \operatorname{sen} x - \operatorname{sen} y < y - x$$

e dalla disuguaglianza a destra si ottiene $x + \operatorname{sen} x < y + \operatorname{sen} y$.

Esercizio 3.61 - *Disegnare i grafici delle funzioni non periodiche* $\operatorname{sen} x^2$, $(\operatorname{sen} x)/x$, $e^{-x}\operatorname{sen} x$, $x \operatorname{sen} x$, $\operatorname{sen} 1/x$, $x \operatorname{sen} 1/x$ *e* $x^2 \operatorname{sen} 1/x$.

Esercizio 3.62 - *Disegnare le funzioni* $\operatorname{arcsen}\operatorname{sen} x$, $\arccos\cos x$ *e* $\operatorname{arctg}\operatorname{tang} x$.

Esercizio 3.63 - *Dimostrare che*

$$\operatorname{arctg} x + \operatorname{arctg}\frac{1}{x} = \begin{cases} \pi/2 & \textit{se } x > 0 \\ -\pi/2 & \textit{se } x < 0\,. \end{cases}$$

Proiettiamo i punti della retta $y = 0$ su $\mathscr{C}$ dal punto $N = (0, 1)$. Ogni semiretta uscente da N e passante per $T = (t, 0)$ incontra $\mathscr{C}$ nel punto P di coordinate

$$\begin{cases} x(t) = \dfrac{2t}{t^2+1} \\[2ex] y(t) = \dfrac{t^2-1}{t^2+1}\,, \end{cases}$$

che sono le stesse funzioni studiate nel § 3.6. In queste relazioni si riconoscono le ben note *formule parametriche*.

Concludiamo il paragrafo illustrando un nuovo sistema di coordinate nel piano che è molto utile in numerosi problemi. Se nelle equazioni parametriche (3.17) si fa variare anche la distanza $r = \rho$ da O, oltre all'angolo ϑ, il punto corrispondente può muoversi liberamente nel piano. Ogni circonferenza ha equazione $\rho = \textit{costante}$ e ogni semiretta uscente da O ha equazione $\vartheta = \textit{costante}$. Ogni coppia (ρ, ϑ) rappresenta ancora un punto del piano e le due variabili si chiamano *coordinate polari*, ρ è la *distanza polare* e ϑ si chiama *argomento* o *anomalia*, una funzione multivoca che ritroveremo, insieme al suo grafico, nel Cap. 4. Invertendo la relazione funzionale con cui le coordinate cartesiane dipendono dalle polari, si ottiene il sistema

$$\begin{cases} \rho = \sqrt{x^2+y^2} \\ \vartheta = \arg(x, y) \end{cases}$$

dove la seconda può essere resa una vera funzione se ne limitiamo l'immagine ad un angolo giro, scegliendo ad esempio l'*argomento principale* $[-\pi, \pi[$. Limitandone ulteriormente l'immagine al semipiano $x > 0$, viene a coincidere con la funzione $\operatorname{arctg} y/x$.

Capitolo 4

I numeri complessi

4.1 Struttura algebrica

Il prodotto cartesiano $\mathbf{R}^2 = \mathbf{R} \times \mathbf{R}$, l'insieme delle coppie (x, y) di numeri reali, è munito dell'operazione di somma

$$(x, y) + (x', y') = (x + x', y + y') \qquad \forall x, y, x', y' \in \mathbf{R}$$

rispetto alla quale è un gruppo abeliano con $0 = (0, 0)$ come elemento neutro e $(-x, -y)$ come opposto di (x, y). Con semplici considerazioni geometriche riconosciamo in questa operazione la ben nota somma vettoriale secondo la regola del parallelogrammo.

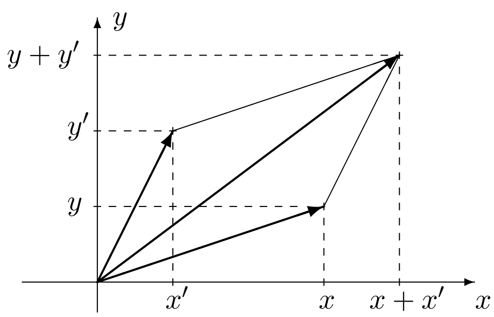


Figura 4.1: La somma in $\mathbf{R}^2$.

È definita anche un'operazione di prodotto tra gli elementi di $\mathbf{R}^2$ e i numeri reali che insieme alla somma fa di $\mathbf{R}^2$ uno spazio vettoriale. Ad ogni coppia (x, y) e ad ogni numero reale t corrisponde il prodotto $t(x, y)$ definito da

$$t(x, y) = (tx, ty)\,.$$

I punti $P = (x, y)$ e $Q = t(x, y)$ sono allineati con $O = (0, 0)$, in particolare si trovano nell'ordine OPQ se $t > 1$, OQP se $0 < t < 1$ e QOP se $t < 0$, inoltre $P = Q$ se $t = 1$ e $Q = O$ se $t = 0$. La distanza tra due punti $P = (x, y)$ e $Q = (x', y')$ è la solita, quella euclidea

$$d(P, Q) = |PQ| = \sqrt{(x - x')^2 + (y - y')^2}\,.$$

Introducendo anche l'operazione di prodotto

$$(x, y) \cdot (x', y') = (xx' - yy', xy' + x'y) \qquad \forall x, y, x', y' \in \mathbf{R}$$

le coppie $z = (x, y)$ vengono dette *numeri complessi* e $\mathbf{R}^2$ viene indicato con $\mathbf{C}$.

Si vede facilmente che $\mathbf{C}$, l'insieme dei numeri complessi, è un campo, infatti, banalmente, il prodotto è associativo, commutativo e ammette $(1, 0)$ come elemento neutro, inoltre vale la proprietà distributiva. Per calcolare l'inverso $z^{-1} = 1/z = (x', y')$ del numero complesso $z = (x, y) \neq 0$ dobbiamo imporre la condizione $zz' = (1, 0)$ che si traduce nel sistema

$$\begin{cases} xx' - yy' = 1 \\ xy' + x'y = 0\,. \end{cases}$$

Moltiplicando la prima per x e la seconda per y, per somma membro a membro si ottiene

$$(x^2 + y^2)x' = x$$

e moltiplicando la prima per y e la seconda per x, si ottiene per differenza

$$(x^2 + y^2)y' = -y\,,$$

da cui

$$z^{-1} = \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2} \right)\,. \tag{4.1}$$

Il rapporto tra due numeri complessi z e $w \neq 0$ rimane definito dal prodotto del primo per l'inverso del secondo

$$\frac{z}{w} = z \cdot w^{-1} = z \cdot \frac{1}{w}\,.$$

Esercizio 4.1 *Dimostrare la legge dell'annullamento del prodotto.*

La formula ottenuta per l'inverso ha un'interessante interpretazione geometrica. Indichiamo con $\mathbf{U}$ la circonferenza unitaria di centro 0. Se dal punto $P = (x, y)$ si traccia una retta tangente alla circonferenza, detto T il punto di tangenza, il triangolo OTP è rettangolo in T.

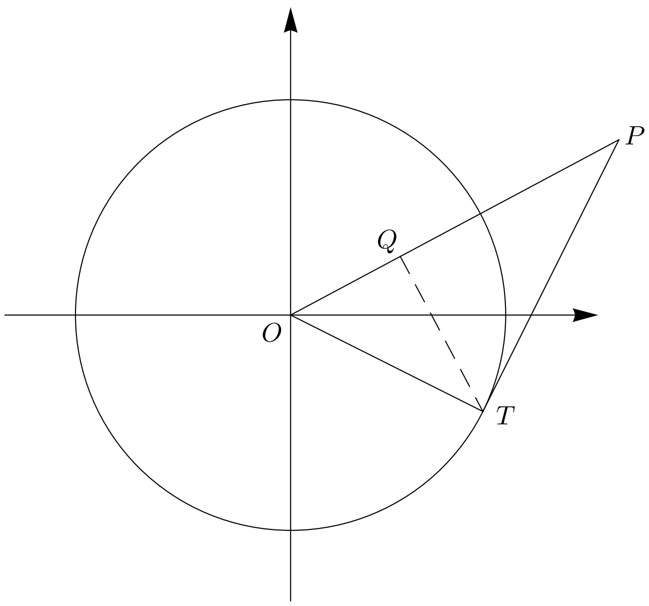


Figura 4.2: Costruzione geometrica dell'inverso.

Per il I Teorema di Euclide, detta Q la proiezione di T su OP, si ha la relazione di proporzionalità tra segmenti

$$OP : OT = OT : OQ$$

e passando alle lunghezze

$$|OP||OQ| = 1\,. \tag{4.2}$$

Allora

$$|OQ| = \frac{1}{\sqrt{x^2+y^2}} \quad \text{e} \quad Q = \left(\frac{x}{x^2+y^2}, \frac{y}{x^2+y^2}\right).$$

Confrontando Q con z^{-1} si vede che sono simmetrici rispetto all'asse x.

L'inversione, cioè il passaggio da z a z^{-1}, porta punti esterni alla circonferenza unitaria in punti interni e viceversa e ad ogni punto di $\mathbf{U}$ associa il suo simmetrico rispetto all'asse x, anch'esso su $\mathbf{U}$.

Esercizio 4.2 - *Quali sono gli inversi dei numeri $z \in \mathbf{C}$ tali che $y = 1$?*

4.2 Modulo e coniugato

Non vi è in $\mathbf{C}$ una relazione d'ordine di qualche utilità che sia compatibile con le operazioni precedenti, quindi per noi scritture del tipo $z \leqslant w$ o $z > 0$ non avranno alcun senso, non esistono numeri complessi positivi o negativi o maggiori di altri numeri complessi. Il *modulo* del numero complesso $z = (x, y)$ è il numero reale non negativo

$$|z| = \sqrt{x^2+y^2}$$

che è poi la distanza di z da 0. Assumiamo infatti come metrica di $\mathbf{C}$ quella della distanza euclidea

$$d(z, z') = |z - z'| = \sqrt{(x-x')^2 + (y-y')^2}\,.$$

Il *coniugato* di z è il numero complesso

$$\bar{z} = (x, -y)\,.$$

Esercizio 4.3 *Verificare che $\overline{z+w} = \bar{z} + \bar{w}$ e $\overline{zw} = \bar{z}\bar{w}$.*

La (4.1) assume la forma

$$\frac{1}{z} = \frac{\bar{z}}{|z|^2}\,. \tag{4.3}$$

Del resto questa può essere ottenuta direttamente osservando che

$$z\bar{z} = (x, y)(x, -y) = x^2 + y^2 = |z|^2\,.$$

Esercizio 4.4 *Verificare che $z \in U \Rightarrow z^{-1} = \bar{z}$.*

Il modulo in $\mathbf{C}$ soddisfa formalmente le stesse proprietà del modulo in $\mathbf{R}$, dobbiamo però precisare che la relazione $|z| \leqslant r$ non definisce un intervallo, ma il disco di centro 0 e raggio r essendo $r \geqslant 0$.

Esercizio 4.5 *Verificare che* $|z| \geqslant 0$ *per ogni* $z \in \mathbf{C}$ *e* $|z| = 0$ *se e solo se* $z = 0$.

Esercizio 4.6 *Dimostrare che* $|\bar{z}| = |z|$ *e che* $|zw| = |z||w|$ *per ogni* $z, w \in \mathbf{C}$.

Tra le proprietà rilevanti ricordiamo quella triangolare

$$\begin{aligned}|z+w|^2 &= (z+w)\overline{(z+w)} = |z|^2 + z\bar{w} + \bar{z}w + |w|^2 \\ &\leqslant |z|^2 + 2|z||w| + |w|^2 = (|z| + |w|)^2 ,\end{aligned}$$

dove la disuguaglianza ha senso perché di fatto $z\bar{w} + \bar{z}w$ è reale, da cui

$$|z+w| \leqslant |z| + |w| .$$

Esercizio 4.7 *Verificare che* $||z| - |w|| \leqslant |z - w|$.

Anche in $\mathbf{C}$ il modulo può essere usato per definire la distanza

$$d(z,w) = |z-w| \qquad \forall z, w \in \mathbf{C} , \tag{4.4}$$

ma coincide esattamente con la distanza euclidea di $\mathbf{R}^2$ se scritta in termini delle componenti cartesiane di z e w. La condizione $|z - z_0| < r$ definisce il cerchio aperto di centro z_0 e raggio $r > 0$.

Analogamente alle potenze del campo reale definiamo per ogni $n \in \mathbf{N}$

$$\begin{cases} p_0(z) = 1 & \text{se } n = 0 \\ p_n(z) = zp_{n-1}(z) & \text{se } n \geqslant 1 \end{cases}$$

e se $n \in \mathbf{Z}$ e $n < 0$ poniamo $z^n = 1/z^{-n}$ per ogni $z \in \mathbf{C} - \{0\}$.

Esercizio 4.8 *Per quali* $z \in \mathbf{C}$ *si ha* $|z|^2 = z^2$*?*

4.3 Il sottocampo dei numeri reali, parti reale e immaginaria

Sommando e moltiplicando tra loro due numeri complessi con ordinata nulla si ottiene un numero complesso con ordinata nulla, infatti

$$(x,0) + (x',0) = (x+x',0) \qquad \text{e} \qquad (x,0)\cdot(x',0) = (xx',0) .$$

Allora l'insieme $\boldsymbol{R} = \{z = (x,y) \in \mathbf{C} \mid y = 0\}$, essendo chiuso rispetto a queste operazioni, è un sottocampo di $\mathbf{C}$. L'applicazione $\varphi : \mathbf{R} \to \boldsymbol{R}$ definita da

$$\varphi(x) = (x,0)$$

conserva le operazioni nel senso che

$$\varphi(x+x') = \varphi(x) + \varphi(x') \qquad \text{e} \qquad \varphi(xx') = \varphi(x)\varphi(x')$$

per cui viene detta *omomorfismo*, in più è ovviamente bigettiva e conserva il modulo, cioè $|\varphi(x)| = |x|$, pertanto è un *isomorfismo isometrico* che ci permette di identificare $\boldsymbol{R}$ con $\mathbf{R}$ e di considerarli come lo stesso insieme. In pratica i numeri complessi della forma $(x,0)$ sono di fatto i numeri reali x, in particolare l'elemento neutro $(1,0)$ di $\mathbf{C}$ non è altro che l'elemento neutro 1 di $\mathbf{R}$.

Per l'insieme $\mathbf{I} = \{z = (x, y) \in \mathbf{C} \mid x = 0\}$ non possiamo dire la stessa cosa. Certamente $\mathbf{I}$ è chiuso rispetto alla somma e per questo costituisce un sottogruppo additivo di $\mathbf{C}$, ma il prodotto di due suoi elementi

$$(0, y)(0, y') = (-yy', 0)$$

è, come si vede, un numero reale, ad esempio $(0, y)^2 = (-y^2, 0)$ oppure $(0, 1)^2 = (-1, 0)$. I numeri di questo tipo vengono detti *immaginari*. Sfruttando la struttura vettoriale di $\mathbf{C}$, se si pone $i = (0, 1)$ che viene detta *unità immaginaria*, si ottiene per un numero complesso $z = (x, y)$ la rappresentazione

$$(x, y) = x(1, 0) + y(0, 1) = x \cdot 1 + y \cdot i = x + iy\,,$$

dove

$$x = \operatorname{Re} z = \textit{parte reale di } z \qquad \text{e} \qquad y = \operatorname{Im} z = \textit{parte immaginaria di } z\,.$$

Di verifica immediata sono le relazioni

$$\operatorname{Re} z = \frac{z + \bar{z}}{2} \qquad \text{e} \qquad \operatorname{Im} z = \frac{z - \bar{z}}{2i}\,,$$
$$|\operatorname{Re} z| \leqslant |z| \qquad \text{e} \qquad |\operatorname{Im} z| \leqslant |z|\,.$$

Esercizio 4.9 *Usando la (4.3) scrivere il numero complesso $z = 1/(x + iy)$, con $y \neq 0$, nella forma $z = \operatorname{Re} z + i \operatorname{Im} z$.*

Esercizio 4.10 *Verificare che il prodotto per i di un numero complesso z, considerato come vettore, ha l'effetto di ruotarlo di un angolo retto attorno a 0. Studiare la successione i^n, $n \in \mathbf{Z}$.*

4.4 Un sottogruppo moltiplicativo, formula di De Moivre

La circonferenza unitaria $\mathbf{U} = \{z \in \mathbf{C} \mid |z| = 1\}$ è un sottogruppo moltiplicativo di $\mathbf{C}$ perché

$$|z| = |w| = 1 \Rightarrow |zw| = |z||w| = 1\,.$$

Naturalmente $z \in \mathbf{U}$ se e solo se esiste $\vartheta \in \mathbf{R}$, detto *argomento* di z, tale che $z = \cos\vartheta + i \operatorname{sen}\vartheta$ e se anche $w \in \mathbf{U}$, per cui $w = \cos\varphi + i \operatorname{sen}\varphi$, il loro prodotto è dato da

$$zw = (\cos\vartheta + i \operatorname{sen}\vartheta)(\cos\varphi + i \operatorname{sen}\varphi) = \cos(\vartheta + \varphi) + i \operatorname{sen}(\vartheta + \varphi)\,.$$

Il prodotto di due numeri complessi unitari, cioè di modulo 1, è il numero complesso unitario che ha per argomento la somma degli argomenti dei due numeri. Possiamo aggiungere che al rapporto corrisponde la differenza degli argomenti

$$\frac{z}{w} = z\bar{w} = (\cos\vartheta + i \operatorname{sen}\vartheta)(\cos\varphi - i \operatorname{sen}\varphi) = \cos(\vartheta - \varphi) + i \operatorname{sen}(\vartheta - \varphi)\,.$$

Come immediata conseguenza si ha

$$z^2 = (\cos\vartheta + i \operatorname{sen}\vartheta)^2 = \cos 2\vartheta + i \operatorname{sen} 2\vartheta$$

e per induzione

$$z^n = (\cos\vartheta + i \operatorname{sen}\vartheta)^n = \cos n\vartheta + i \operatorname{sen} n\vartheta \qquad \forall n \in \mathbf{Z}\,,$$

da cui risulta evidente che mentre z compie un giro sulla circonferenza unitaria, z^n compie n giri a velocità n volte più grande, in senso antiorario se $n > 0$ e orario se $n < 0$.

Esercizio 4.11 - *Dedurne le formule di duplicazione, triplicazione, quadruplicazione,..., n-plicazione.*

Ora, se $z \neq 0$ è un numero complesso qualsiasi $z/|z|$ è unitario e quindi del tipo

$$z/|z| = \cos\vartheta + i \operatorname{sen}\vartheta$$

per qualche $\vartheta \in \mathbf{R}$. Posto $\rho = |z|$, si ottiene la *forma polare* o *trigonometrica*

$$z = \rho(\cos\vartheta + i \operatorname{sen}\vartheta)\,.$$

I parametri $\rho \geqslant 0$ e ϑ si chiamano *coordinate polari* di z, per cui è lecito scrivere $z = (\rho, \vartheta)$, ma dobbiamo tener presente che la corrispondenza tra coppie e punti non è biunivoca perché ϑ è determinato a meno di multipli interi di 2π se $\rho > 0$, addirittura indeterminato se $\rho = 0$, cioè se $z = 0$.

Indichiamo con $\arg z$ l'argomento di z. Da quanto visto, discendono immediate le relazioni

$$|zw| = |z||w| \qquad \text{e} \qquad \arg zw = \arg z + \arg w + 2k\pi\,,\ k \in \mathbf{Z}\,,\ \ \forall z, w \in \mathbf{C} \tag{4.5}$$

e le *formule di De Moivre*

$$z^n = \rho^n(\cos n\vartheta + i \operatorname{sen} n\vartheta) \qquad \forall n \in \mathbf{Z} \tag{4.6}$$

che possono essere riscritte nella forma

$$|z^n| = |z|^n \qquad \text{e} \qquad \arg z^n = n \arg z + 2k\pi\,,\ k \in \mathbf{Z}\,,\ \ \forall z \in \mathbf{C}\,.$$

Vediamo come il sottogruppo $\mathbf{U}$ opera su $\mathbf{C}$, in altre parole vogliamo vedere qual è l'effetto sui numeri complessi del prodotto per un elemento di $\mathbf{U}$. Scelto un certo $u \in \mathbf{U}$, l'applicazione $\psi : \mathbf{C} \to \mathbf{C}$

$$\psi(z) = uz \qquad \forall z \in \mathbf{C}$$

è un'isometria in quanto

$$|\psi(z)| = |uz| = |u||z| = |z|\,.$$

Se $u = 1$ ψ è l'identità su $\mathbf{C}$, se invece $u \neq 1$ ψ ammette come unico punto fisso lo 0, come si vede risolvendo l'equazione ai punti fissi $\psi(z) = z$. Si tratta dunque, in ogni caso, di una rotazione di centro 0. A conferma di ciò, possiamo infatti verificare direttamente che, dato $u = (\cos\alpha, \operatorname{sen}\alpha) \in \mathbf{U}$, per ogni $z = \rho(\cos\vartheta + i \operatorname{sen}\vartheta)$ si ha

$$uz = \rho(\cos(\vartheta + \alpha) + i \operatorname{sen}(\vartheta + \alpha))\,.$$

Come caso particolare ritroviamo la rotazione dell'angolo retto nel caso del prodotto per $u = i$ come abbiamo già detto.

Se si considera il prodotto di z per un numero complesso c qualunque, la rotazione che z subisce per l'incremento del suo argomento con l'argomento di c è composta con una omotetia di centro 0 del fattore $|c|$, come risulta evidente calcolando il prodotto.

La corrispondenza $z \to \arg z$, che ha come dominio $\mathbf{C} - \{0\}$, non è una vera funzione, se guardiamo al significato che abbiamo dato a questo termine, dato che associa ad ogni $z \neq 0$ un'infinità numerabile di valori, anziché uno soltanto, che differiscono uno dall'altro per un multiplo intero di 2π.

Diciamo allora che si tratta di una *funzione multivoca*, a più valori. A tutti i numeri complessi allineati su una semiretta uscente da 0 associa lo stesso valore, l'angolo che essa forma col semiasse reale positivo; mentre la semiretta ruota in senso antiorario [in senso orario] attorno a 0 l'angolo cresce [decresce]. Il grafico di arg nello spazio è una superficie che ha la stessa forma della rampa di accesso ad un garage multipiano e si chiama *elicoide*. Indicata con z la terza coordinata, lungo cui riportiamo il valore di $\arg(x+iy)$, le sue equazioni parametriche sono date da

$$\begin{cases} x = \rho\cos\vartheta \\ y = \rho \operatorname{sen}\vartheta \\ z = \vartheta\,, \qquad\qquad \rho > 0\,,\ \vartheta \in \mathbf{R}\,. \end{cases}$$

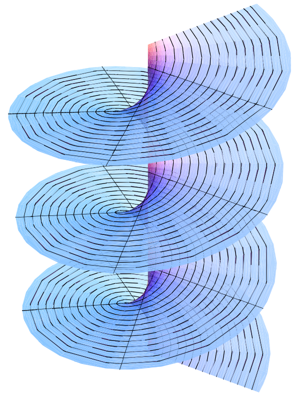


Figura 4.3: Il grafico dell'arg a elicoide.

4.5 La radice n-esima in C

In questo paragrafo vogliamo definire la radice n-esima nel campo complesso partendo dall'equazione $z^n = w$, in cui sono dati il numero complesso w e il numero intero n. Ovviamente possiamo supporre $w \neq 0$ dato che il caso $w = 0$ non è significativo, o perché ci dà la soluzione banale $z = 0$ se $n > 0$, o perché non ha soluzioni se $n \leqslant 0$. Per $n = 0$ l'equazione è soddisfatta da tutti i numeri complessi non nulli solo per $w = 1$, altrimenti non ha soluzione. Osserviamo infine che $z = w$ per $n = 1$ e $z = 1/w$ per $n = -1$. Comunque basta formulare e affrontare il problema con $n > 0$, in **N** quindi, dato che il caso $n < 0$ può essere ricondotto al precedente scrivendo l'equazione nella forma $(1/z)^{-n} = w$.

Problema 4.1 (della radice) - *Dati $w \in \mathbf{C} - \{0\}$ e $n \geqslant 1$, determinare $z \in \mathbf{C}$ tale che*

$$z^n = w\,.$$

Scriviamo il dato w e l'incognita z in forma polare

$$w = r(\cos\varphi + i \operatorname{sen}\varphi) \qquad \text{e} \qquad z = \rho(\cos\vartheta + i\operatorname{sen}\vartheta)$$

con $r, \rho > 0$. Dalle (4.6) si ottengono le condizioni

$$\rho^n = r \qquad \text{e} \qquad n\vartheta = \varphi + 2k\pi\,,\ k \in \mathbf{Z}\,.$$

Ne discendono i valori delle incognite ρ e ϑ

$$\rho = \sqrt[n]{r} \qquad \text{e} \qquad \vartheta_k = \frac{\varphi + 2k\pi}{n}\,,\ k \in \mathbf{Z}\,.$$

Il modulo ρ di z è univocamente determinato come l'unica radice reale positiva del numero reale positivo r, ma di argomenti possibili ne risulta un'infinità numerabile: alla frazione φ/n dell'angolo iniziale φ vanno sommate le frazioni $2k\pi/n$ di tutti i multipli interi dell'angolo giro 2π. Ai vari $k \in \mathbf{Z}$ corrispondono dunque gli angoli

$$\begin{aligned}
&\vdots\\
&\vartheta_{-1} = \frac{\varphi}{n} - \frac{2\pi}{n}\\
&\vartheta_0 = \frac{\varphi}{n}\\
&\vartheta_1 = \frac{\varphi}{n} + \frac{2\pi}{n}\\
&\vartheta_2 = \frac{\varphi}{n} + 2\frac{2\pi}{n}\\
&\vdots\\
&\vartheta_k = \frac{\varphi}{n} + k\frac{2\pi}{n}\\
&\vdots\\
&\vartheta_{n-1} = \frac{\varphi}{n} + (n-1)\frac{2\pi}{n}\\
&\vartheta_n = \frac{\varphi}{n} + n\frac{2\pi}{n} = \frac{\varphi}{n} + 2\pi\\
&\vdots
\end{aligned}$$

Osserviamo che dopo n valori consecutivi di k, per esempio da $k = 0$ a $k = n - 1$, si ottengono gli stessi angoli incrementati di 2π, poi ancora gli stessi incrementati di 4π e così via. Di conseguenza i numeri complessi distinti che soddisfano l'equazione $z^n = w$ sono solo n e sono disposti ai vertici di un poligono regolare di n lati. Per determinarli tutti basterà attribuire a k n valori interi consecutivi, ad esempio $k = 0, 1, 2, \dots, n-1$.

Definizione 4.2 *Si chiama **radice** n**-esima** del numero $w \in \mathbf{C}$ l'insieme*

$$\sqrt[n]{w} = \{z \in \mathbf{C} \mid z^n = w\}\,.$$

Da quanto appena visto risulta

$$\sqrt[n]{w} = \left\{ z_k = \sqrt[n]{|w|}\left(\cos\frac{\varphi + 2k\pi}{n} + i \operatorname{sen}\frac{\varphi + 2k\pi}{n}\right) \mid k = 0, 1, 2, \dots, n-1 \right\}.$$

Osserviamo che ogni z_k è dato da $z_k = z_0 u^k$ dove $|u| = 1$ e $\arg u = 2\pi/n$.

Calcoliamo a titolo di esempio $\sqrt[3]{i}$, $\sqrt[6]{-1}$ e $\sqrt[4]{1-i}$. Prima di tutto si scrivono i radicandi in forma polare

$$i = \cos\frac{\pi}{2} + i \operatorname{sen}\frac{\pi}{2}\,, \qquad -1 = \cos\pi + i\operatorname{sen}\pi\,, \qquad 1 - i = \cos\frac{-\pi}{4} + i\operatorname{sen}\frac{-\pi}{4}$$

dei quali i primi due hanno modulo 1 mentre il terzo ha modulo $\sqrt{2}$. Allora

$$z \in \sqrt[3]{i} \Rightarrow |z| = 1\,, \qquad z \in \sqrt[6]{-1} \Rightarrow |z| = 1\,, \qquad z \in \sqrt[4]{1-i} \Rightarrow |z| = \sqrt[8]{2}\,.$$

Rimangono da calcolare gli argomenti. Nei vari casi si ha rispettivamente

$$\vartheta_k = \frac{1}{3}\left(\frac{\pi}{2} + 2k\pi\right)\,, \qquad \vartheta_k = \frac{1}{6}(\pi + 2k\pi)\,, \qquad \vartheta_k = \frac{1}{4}\left(-\frac{\pi}{4} + 2k\pi\right)\,.$$

Nel primo, $n = 3$, ai tre valori consecutivi $k = 0, 1, 2$ corrispondono gli angoli $\vartheta_0 = \pi/6$, $\vartheta_1 = 5\pi/6$ e $\vartheta_2 = 3\pi/2$, da cui

$$\sqrt[3]{i} = \left\{ \frac{\sqrt{3}}{2} + \frac{i}{2}\,,\ -\frac{\sqrt{3}}{2} + \frac{i}{2}\,,\ -i \right\}$$

che è l'insieme dei vertici di un triangolo equilatero inscritto nella circonferenza unitaria. Nel secondo, $n = 6$, si scelgono gli interi k tra 0 e 5, a partire dalla soluzione $z_0 = \pi/6$ si ottengono le altre per rotazioni successive dell'angolo $\pi/3$ che sono i vertici di un esagono regolare. Nel terzo caso si ottiene un quadrato.

La corrispondenza $z \to \sqrt[n]{z}$ è un altro esempio di funzione multivoca, ogni intervallo di ampiezza 2π contiene una determinazione di tutti gli argomenti delle n radici n-esime di z. Osservando la Figura 4.4 che rappresenta i 3 rami in cui si decompone l'argomento passando alla radice cubica, si spiega facilmente la diversa struttura geometrica delle radici di numeri complessi che formano un insieme, ad esempio una curva, contenente o meno lo 0 al suo interno.

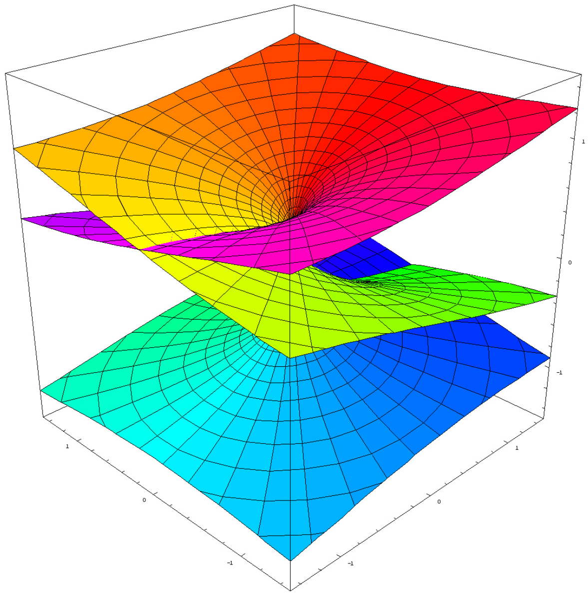


Figura 4.4: Il grafico di arg $\sqrt[3]{z}$.

Se una curva chiusa, come il cerchio della Figura 4.5, circonda lo 0 attraversando tutti i valori tra 0 e 2π le corrispondenti radici formano una un'unica curva. Se invece lascia fuori lo 0, rimanendo tutta contenuta in un angolo minore di 2π con vertice in 0, le radici si sconnettono per formare tre curve ben separate.

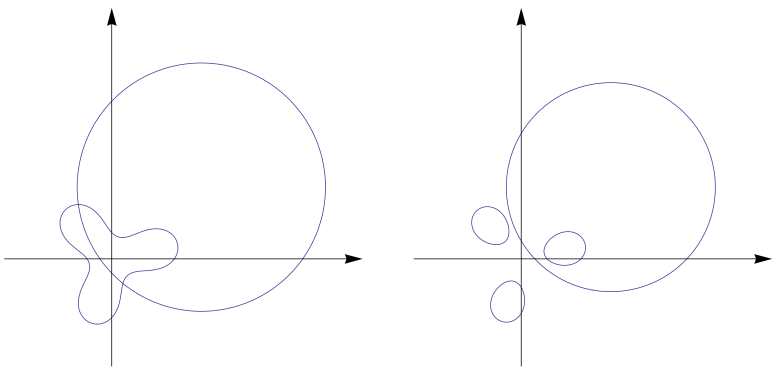


Figura 4.5: Le due diverse strutture di connessione.

4.6 Equazioni algebriche in C

Ogni funzione $f : \mathbf{C} \to \mathbf{C}$ può sempre scriversi nella forma $f(z) = u(x, y) + iv(x, y)$, separando la parte reale dalla parte immaginaria, e identificarsi di fatto con una coppia di funzioni di due variabili reali a valori in $\mathbf{R}$. Questa decomposizione risulta utile ad esempio nella ricerca degli zeri di un polinomio $P(z)$, problema che si traduce nella determinazione dei punti d'incontro di due curve algebriche

$$\begin{cases} u(x, y) = 0 \\ v(x, y) = 0 \, . \end{cases}$$

Le radici di un numero complesso c possono essere calcolate con questo metodo, basta riconoscerle come gli zeri del polinomio $P(z) = z^n - c$. Le radici quadrate di -1 sono le soluzioni dell'equazione

$$z^2 + 1 = 0$$

che diventa il sistema

$$\begin{cases} x^2 - y^2 + 1 = 0 \\ xy = 0 \, . \end{cases}$$

Se $x = 0$ $y = \pm 1$ e se $y = 0$ $x = \pm 1$, quindi $\sqrt{-1} = \{(0, 1), (0, -1)\} = \{i, -i\}$.

Più in generale, sapendo che ogni numero complesso non nullo ammette sempre esattamente n radici n-esime, c'è da aspettarsi che non solo $z^n - c$, ma ogni polinomio di grado n debba ammettere n radici. In effetti le cose stanno proprio così, la storia stessa dei numeri complessi nasce dall'esigenza di costruire una classe numerica nella quale poter garantire l'esistenza di soluzioni alle equazioni algebriche. Il celebre *Teorema Fondamentale dell'Algebra*, per la cui dimostrazione non abbiamo strumenti adeguati, assicura che ogni polinomio $P(z)$ di grado $n \geqslant 1$ possiede almeno una radice z_1, da cui segue

$$P(z) = (z - z_1)P_1(z) \, ,$$

ma applicando lo stesso risultato a P_1 e poi a P_2 ecc. si ottiene che di radici in realtà ne ammette tante quant'è il suo grado, contate ciascuna con la sua molteplicità. In altre parole vale la fattorizzazione

$$P(z) = a_n(z - z_1)^{m_1}(z - z_2)^{m_2} \dots (z - z_k)^{m_k}$$

dove $m_1 + m_2 + \dots + m_k = n$.

Per trovare le radici di un polinomio di secondo grado

$$P(z) = az^2 + bz + c \, , \qquad a \neq 0 \, ,$$

conviene scriverlo nella forma

$$P(z) = a \left[\left(z + \frac{b}{2a} \right)^2 + \frac{4ac - b^2}{4a^2} \right] .$$

La condizione di annullamento equivale all'equazione

$$\left(z + \frac{b}{2a} \right)^2 = \frac{\Delta}{4a^2} \, , \qquad \Delta = b^2 - 4ac \, .$$

Ora, se $\Delta = 0$ abbiamo l'unica radice doppia $z = -b/2a$, se $\Delta \neq 0$ questo numero complesso ammette sempre due radici distinte, $\sqrt{\Delta} = \{\delta_1 = \delta \, , \ \delta_2 = -\delta\}$, opposte una all'altra perché la differenza dei loro argomenti vale π. L'insieme delle radici z_1 e z_2 è dato da

$$\{z_1 \, , \ z_2\} = \frac{-b + \sqrt{\Delta}}{2a} = \left\{ \frac{-b \pm \delta}{2a} \right\} .$$

Per una simpatica e interessante discussione sulla risoluzione dell'equazione di terzo grado e la sua storia si veda il documento di Gianluca Gorni del Dipartimento di Matematica dell'Università di Udine. Come già detto, non esistono formule risolutive se il grado supera 4.

Un'importante osservazione riguarda il caso dei coefficienti reali: se i coefficienti del polinomio $P(z)$, con $z \in \mathbf{C}$, sono reali le radici sono a due a due coniugate. In altre parole, se α è radice di P anche $\bar{\alpha}$ lo è. Infatti $P(\bar{z}) = \overline{P(z)}$ e se $P(\alpha) = 0$

$$P(\bar{\alpha}) = \overline{P(\alpha)} = 0\,.$$

Ne segue che se P ha grado dispari una radice deve essere reale. Questo spiega, insieme al Teorema Fondamentale dell'Algebra, perché un polinomio in $\mathbf{R}$ di grado superiore a 2 non è mai irriducibile. Si provi ad esempio a scomporre in fattori $z^4 + 1$.

Esercizio 4.12 - *Determinare tutte le radici in* **C** *del polinomio*

$$P(z) = z^5 - 2z^4 + 2z^3 - 8z^2 + 16z - 16$$

sapendo che una sua radice è il numero complesso $1 + i$.

Sono molto utili le relazioni tra le radici e i coefficienti del polinomio: se $P(z) = c_2 z^2 + c_1 z + c_0$ allora $z_1 + z_2 = -c_1/c_2$ e $z_1 z_2 = c_0/c_2$, se $P(z) = c_3 z^3 + c_2 z^2 + c_1 z + c_0$ allora $z_1 + z_2 + z_3 = -c_2/c_3$, $z_1 z_2 + z_2 z_3 + z_3 z_1 = c_1/c_3$ e $z_1 z_2 z_3 = -c_0/c_3$ e così via. Esse vanno tenute presenti quando sono assegnate delle condizioni che devono essere soddisfatte da espressioni scrivibili in termini di somme e prodotti delle soluzioni, come le seguenti

$$z_1^2 + z_2^2 = (z_1 + z_2)^2 - 2z_1 z_2\,, \qquad z_1^3 + z_2^3 = (z_1 + z_2)^3 - 3z_1 z_2 (z_1 + z_2)\,,$$

$$(z_1 - z_2)^2 = (z_1 + z_2)^2 - 4z_1 z_2\,, \qquad \frac{1}{z_1} + \frac{1}{z_2} = \frac{z_1 + z_2}{z_1 z_2}\,.$$

Vogliamo ad esempio determinare tutti i numeri $\lambda \in \mathbf{C}$ tali che una delle due soluzioni dell'equazione

$$z^2 - \lambda z + 1 = 0$$

abbia modulo pari a 2.

Poiché $z_1 z_2 = 1$, se $|z_1| = 2$ deve essere $|z_2| = 1/2$, quindi

$$z_1 = 2(\cos\vartheta + i\,\mathrm{sen}\,\vartheta) \qquad \text{e} \qquad z_2 = \frac{1}{2}(\cos\varphi + i\,\mathrm{sen}\,\varphi)\,,$$

inoltre $\arg(z_1 z_2) = \vartheta + \varphi = 0$. Posto $\lambda = x + iy$, la relazione $z_1 + z_2 = \lambda$ equivale al sistema

$$\begin{cases} x = \dfrac{5}{2}\cos\vartheta \\[2ex] y = \dfrac{3}{2}\,\mathrm{sen}\,\vartheta\,. \end{cases}$$

I numeri λ richiesti sono i punti dell'ellisse di centro 0 e di semiassi 5/2 lungo l'asse x e 3/2 lungo l'asse y.

L'Esercizio 4.2 tratta il caso di una retta che viene trasformata in una circonferenza attraverso l'inversione. Facendo attenzione a come esse vengono percorse al variare del parametro, si noterà che mentre il punto sulla retta si allontana all'infinito la sua immagine sulla circonferenza si avvicina a 0. Si può immaginare che se il piano complesso si dotasse di un punto aggiuntivo ∞, considerato come punto all'infinito, ad esso corrisponderebbe lo 0 nella trasformazione dell'esercizio. Un modo per realizzare questo ampliamento consiste nell'identificare il piano complesso con la sfera unitaria, la *sfera di Riemann*, mediante la *proiezione stereografica* illustrata nella Figura 4.6.

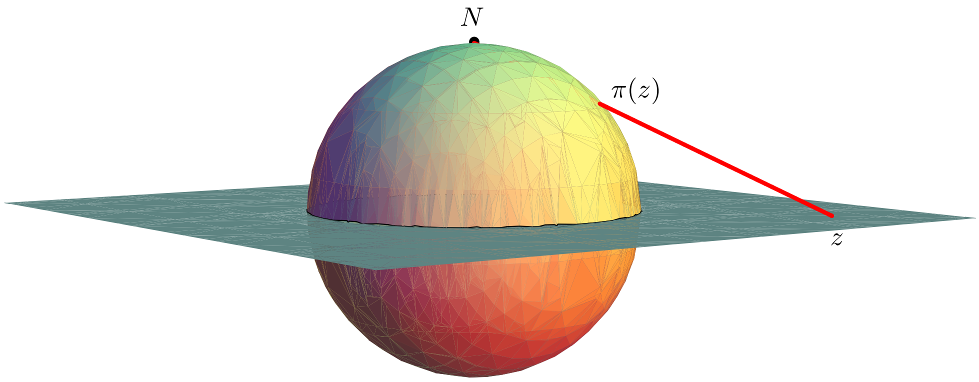


Figura 4.6: La proiezione stereografica della sfera.

Le semirette uscenti dal "polo nord" N stabiliscono una corrispondenza biunivoca $z \to \pi(z)$ che ci permette di identificare N col punto all'∞. In un piano pensato "richiuso" come una sfera viene meno ogni distinzione tra rette e circonferenze venendo a far parte della stessa categoria.

La classe delle *trasformazioni di Möbius*

$$z \to \frac{az+b}{cz+d} = \alpha + \frac{\beta}{z-z_0}\,, \qquad ad - bc \neq 0\,,$$

è un gruppo transitivo*, rispetto alla composizione, di applicazioni bigettive su $\mathbf{C}^* = \mathbf{C} \cup \{\infty\}$ che manda rette o circonferenze in rette o circonferenze, di cui l'inversione $z \to 1/z$, la traslazione $z \to z + c$ e l'omotetia con rotazione $z \to cz$ sono casi particolari. Vediamo concretamente, generalizzando un po' l'Esercizio 4.2, come opera ad esempio l'inversione lasciando al lettore il compito di esercitarsi in modo simile con le altre due. Una retta in $\mathbf{C}$ ha equazione

$$ax + by = a\frac{z+\bar{z}}{2} + b\frac{z-\bar{z}}{2i} = pz + \bar{p}\bar{z} = c$$

e ponendo $z = 1/w$ si trasforma in

$$p\bar{w} + \bar{p}w = c|w|^2\,.$$

Se $c = 0$ la retta data passa per 0 e viene trasformata in un'altra retta per 0, se $c \neq 0$ non passa per 0 e ha per immagine

$$|w|^2 - \frac{p}{c}\bar{w} - \frac{\bar{p}}{c}w = x^2 + y^2 + ax + by = 0$$

che rappresenta una circonferenza passante per 0.

Esercizio 4.13 - *Verificare che l'inversione manda rette ortogonali in rette ortogonali o in circonferenze che si attraversano ortogonalmente.*

Si può dimostrare che tutte le trasformazioni di Möbius sono *conformi*, conservano cioè gli angoli tra le tangenti a due curve in un punto comune.

E adesso un po' di svago con questo bel video musicale.

*Un gruppo G di trasformazioni da X in se stesso viene detto *transitivo* se per ogni $x_1, x_2 \in X$ esiste $\varphi \in G$ tale che $\varphi(x_1) = x_2$

Capitolo 5

Limiti di successioni

5.1 Che cosa sono le successioni

Una *successione* a valori in un insieme X, o, brevemente, una successione in X, è una funzione che associa ad ogni numero naturale n un elemento $x_n \in X$. Invece di usare la notazione $x(n)$, tipica delle funzioni, si preferisce scrivere $(x_n)_{n\in\mathbf{N}}$, o semplicemente (x_n), facendo però attenzione a non fare confusione con l'insieme $\{x_n \mid n \in \mathbf{N}\}$ che ne è l'immagine, l'insieme dei valori della successione. Potremmo dire, per intenderci, che le successioni sono *stringhe ordinate con un'infinità numerabile di elementi*, come le coppie, le terne e le n-uple ordinate sono stringhe finite. La variabile n si chiama *indice*, e va considerata come un'etichetta associata all'elemento della successione che occupa l'n-esimo posto. Una successione può anche non essere definita su tutto $\mathbf{N}$, ad esempio solo per n maggiore di un certo k, oppure per n variabile in un certo insieme numerabile $\mathbf{N}' \subset \mathbf{N}$, ma mai in un sottoinsieme finito di $\mathbf{N}$. I valori che (x_n) assume possono essere infiniti o un numero finito, caso in cui almeno uno di essi viene ripetuto infinite volte, ma i valori di n devono essere infiniti. Se (x_n) assume sempre lo stesso valore, $x_n = x$ per ogni $n \in \mathbf{N}$ diciamo che la successione è *costante.*

Esempi

5.1 *Sia S_n la sfera di $\mathbf{R}^3$ di centro $O = (0,0,0)$ e raggio $1/n$, $n \geqslant 1$. L'applicazione $n \to S_n$ è una successione di sfere in $\mathscr{P}(\mathbf{R}^3)$.*

5.2 *Per ogni $n \in \mathbf{N}$ sia $f_n : [0,1] \to \mathbf{R}$ la funzione $f_n(x) = x^n$. L'applicazione $n \to f_n$ è una successione di funzioni definite su $[0,1]$.*

5.3 *Le successioni $(-1)^n$, $1+1/n$, $\operatorname{sen} n$, $z^n/n!$ sono esempi di successioni numeriche, a valori interi, razionali, reali e complessi.*

5.4 *Una successione (x_n) può essere definita per induzione, si dice anche per ricorrenza o in modo iterativo, assegnando il primo elemento e la legge che esprime la dipendenza di ogni elemento dal precedente*

$$\begin{cases} x_0 = \alpha \\ x_{n+1} = f(x_n) \qquad \forall n \in \mathbf{N}\,. \end{cases}$$

Per adesso ci occupiamo delle successioni di numeri reali. Queste ereditano da $\mathbf{R}$ le operazioni fondamentali e l'ordinamento: due successioni numeriche si possono sommare e moltiplicare "termine a termine"

$$(a_n) + (b_n) = (a_n + b_n) \qquad (a_n)\cdot(b_n) = (a_n b_n) \qquad \forall n \in \mathbf{N}$$

e confrontare

$$(a_n) \leqslant (b_n) \quad \text{se e solo se} \quad a_n \leqslant b_n \qquad \forall n \in \mathbf{N}$$

che è una relazione d'ordine parziale.

Dal punto di vista algebrico, l'insieme $\mathbf{R}^{\mathbf{N}}$ delle successioni reali è munito, oltre che della somma, anche del *prodotto per scalari* $\lambda(a_n) = (\lambda a_n)$ per ogni $\lambda \in \mathbf{R}$. Si tratta pertanto di uno spazio vettoriale.

Una successione $(a_n) \subset \mathbf{R}$ è *limitata superiormente* se ammette un maggiorante, cioè se esiste $M \in \mathbf{R}$ tale che $a_n \leqslant M$ per ogni $n \in \mathbf{N}$; sappiamo che il minimo α dei maggioranti, l'estremo superiore, esiste in $\mathbf{R}$ ed è caratterizzato dalle due proprietà

$$\alpha = \sup_{n \in \mathbf{N}} a_n \Leftrightarrow \begin{cases} a_n \leqslant \alpha \quad \forall n \in \mathbf{N}\,, \\ \forall \varepsilon > 0\ \exists n \in \mathbf{N} \,:\, a_n > \alpha - \varepsilon\,. \end{cases}$$

È invece *non limitata superiormente* se l'insieme dei maggioranti è vuoto e si ha

$$\sup_{n \in \mathbf{N}} a_n = +\infty \Leftrightarrow \forall M \in \mathbf{R}\, \exists n \in \mathbf{N} \,:\, a_n > M\,.$$

Se qualche elemento di una successione limitata superiormente assume il valore dell'estremo superiore questo ne è il massimo. Per l'estremo inferiore valgono analoghe caratterizzazioni.

Una successione $(a_n) \subset \mathbf{R}$ è *crescente* [*decrescente*] se $a_n \leqslant a_{n+1}$ [$a_n \geqslant a_{n+1}$] per ogni $n \in \mathbf{N}$ e *strettamente crescente* [*strettamente crescente*] se le disuguaglianze precedenti valgono in senso stretto. In ognuno di questi casi viene detta *monotòna*. Ovviamente (a_n) ammette il suo valore iniziale, mettiamo a_1 se $n \geqslant 1$, come minimo se è crescente e come massimo se decrescente.

Esercizio 5.1 - *Dimostrare che la somma di due successioni a_n e b_n crescenti è crescente. Se sono anche positive dimostrare che $a_n b_n$ è crescente e che $1/a_n$ è decrescente. Infine, se a_n è crescente e b_n decrescente allora a_n/b_n è crescente.*

Esempi

5.5 *La successione $(1/n)$, $n \geqslant 1$, è decrescente perché essendo $n < n+1$ si ha $1/n > 1/(n+1)$. Allora assume il suo valore massimo per $n = 1$. Essendo positiva è limitata inferiormente, inoltre $\inf 1/n = 0$ perché per ogni $\varepsilon > 0$ basta scegliere $n > 1/\varepsilon$ affinché $1/n < \varepsilon$.*

5.6 *La successione*

$$n \to a_n = \frac{n^2+1}{n+2}\,, \qquad n \in \mathbf{N}\,,$$

è limitata inferiormente in quanto positiva, ma non lo è superiormente. Infatti

$$\frac{n^2+1}{n+2} = \frac{n^2-4}{n+2} + \frac{5}{n+2} = n - 2 + \frac{5}{n+2}$$

e se per ogni $M \in \mathbf{R}$ si sceglie $n > M+2$ si ha

$$n - 2 > M \Rightarrow n - 2 + \frac{5}{n+2} > M\,.$$

Infine questa successione è crescente perché

$$a_{n+1} - a_n = \frac{n^2+5n+1}{(n+3)(n+2)} > 0\,,$$

perciò raggiunge il suo minimo, che vale 1/2, per $n = 0$.

5.7 *La successione a primo membro dell'identità di Catalan* (2.11)*, indichiamola con A_n, ammette* 1 *come maggiorante (vedremo che ha* $\log 2$ *come estremo superiore) ed è crescente a partire dal valore* $1/2$*. Infatti, essendo la somma di n termini tutti inferiori al primo di essi che è* $1/n+1$*, è minore di* $n/n+1$*. Inoltre*

$$\begin{aligned} A_{n+1} &= \sum_{k=1}^{n+1} \frac{1}{n+1+k} = \sum_{h=2}^{n+2} \frac{1}{n+h} = \frac{1}{2n+2} + \frac{1}{2n+1} - \frac{1}{n+1} + \sum_{h=1}^{n} \frac{1}{n+h} \\ &= \frac{1}{2n+1} - \frac{1}{2n+2} + A_n > A_n \,. \end{aligned}$$

Si chiama *successione estratta* da a_n, o *sottosuccessione* di (a_n), una successione ottenuta scegliendo infiniti a_n senza alterarne l'ordine. In altre parole, si costruisce una successione strettamente crescente (k_n) di indici in $\mathbf{N}$ e poi si considera la successione (a_{k_n}). Ad esempio (a_{2n}) è la sottosuccessione di (a_n) ottenuta selezionando solo gli indici pari. Naturalmente ogni successione ammette se stessa come sottosuccessione. Si noti che $k_n \geqslant n$ per ogni $n \in \mathbf{N}$ e non appena $k_r > r$ per un certo $r \in \mathbf{N}$ si avrà anche $k_n > n$ per ogni $n > r$. La successione $a_n = \sqrt[n]{n}$, $n \geqslant 1$, ammette, tra le infinite possibili, $a_{2n} = \sqrt[2n]{2n}$, $a_{n^2} = \sqrt[n^2]{n^2}$ e $a_{n!} = \sqrt[n!]{n!}$ come sottosuccessioni. Le successioni $(\log 2 + \log n)$ e $(2 \log n)$ sono estratte dalla $(\log n)$.

5.2 Proprietà generali dei limiti

Vogliamo tradurre in termini rigorosi l'idea di avvicinamento progressivo di una successione numerica (a_n) verso un numero a, il suo limite, al variare di n. Ogni a_n è un'approssimazione di a e l'accuratezza con cui a viene approssimato dagli a_n può essere resa arbitraria pur di scegliere n sufficientemente grande. L'ammettere limite è una proprietà della successione che riguarda il suo comportamento *asintotico*, *definitivo*, e non dipende da alcun insieme finito di elementi a_n.

Definizione 5.1 - *Una successione* $(a_n) \subset \mathbf{R}$ ***converge*** *ad* $a \in \mathbf{R}$ *se*

$$\forall \varepsilon > 0 \quad \exists \nu \in \mathbf{N} \;:\quad |a_n - a| < \varepsilon \quad \forall n > \nu \tag{5.1}$$

e si scrive

$$\lim_{n\to\infty} a_n = a \quad \textit{oppure} \quad a_n \to a \,.$$

Per le proprietà del valore assoluto la condizione $|a_n - a| < \varepsilon$ si può scrivere in una delle forme seguenti

$$-\varepsilon < a_n - a < \varepsilon \,, \qquad a - \varepsilon < a_n < a + \varepsilon \,, \tag{5.2}$$

ma la scrittura usata nella (5.1) pone l'accento sul ruolo della distanza, mentre nella (5.2) si usa l'ordinamento che è strettamente legato al caso particolare di $\mathbf{R}$. Questa osservazione suggerisce che il limite può essere trattato in ogni spazio metrico, come vedremo tra poco.

Il senso della Definizione 5.1 è il seguente:

ogni intervallo aperto $I_\varepsilon(a) =]a-\varepsilon, a+\varepsilon[$*, qualunque sia la sua ampiezza, contiene tutti gli elementi della successione da un certo indice in poi, al più un numero finito di essi possono rimanere al di fuori di* $I_\varepsilon(a)$*.*

Chiaramente l'indice ν dipende da ε. In generale, più è piccolo ε, più grande va scelto il valore di ν oltre il quale $a_n \in I_\varepsilon(a)$. Se la (5.1) è soddisfatta da un certo indice $\nu \in \mathbf{N}$ allora è ancora soddisfatta da tutti quelli che lo seguono, quindi ν non è unico. Se un indice ν che soddisfa la (5.1) esiste per un certo ε, si può scegliere lo stesso ν per ogni $\varepsilon' > \varepsilon$, per questo alcuni usano dire "per ogni ε piccolo a piacere...". È evidente che la definizione di limite ha carattere definitivo: modificando o eliminando un insieme finito di termini il limite non cambia. Per questo si dice anche $|a_n - a| < \varepsilon$ *definitivamente*, per brevità, al posto dell'intera proposizione.

Una successione viene detta *infinitesima* se converge a 0. Ovviamente $a_n \to a$ se e solo se la successione delle distanze $d(a_n, a) = |a_n - a|$ di a_n dal suo limite a è infinitesima.

Definiamo adesso il limite nel caso generale di una successione (x_n) di punti di uno spazio metrico (X, d).

Definizione 5.2 - *Diciamo che* $(x_n) \subset X$ *converge a* $x \in X$*, e si scrive*

$$\lim_{n\to\infty} x_n = x\,, \quad \textit{oppure} \quad x_n \to x\,,$$

se $\lim_{n\to\infty} d(x_n, x) = 0$.

In altre parole $x_n \to x$ se e solo se

$$\forall \varepsilon > 0\ \exists \nu \in \mathbf{N}\ :\quad d(x_n, x) < \varepsilon \quad \forall n > \nu\,,$$

cioè se $x_n \in B_\varepsilon(x)$ definitivamente.

Esercizio 5.2 - *Dimostrare che se* $x_n \to x$ *allora* $d(x_n, x_{n+1}) \to 0$.

Esercizio 5.3 - *Dimostrare che se* $x_n, y_n \to x$ *allora* $d(x_n, y_n) \to 0$.

Esempi

5.8 *In ogni spazio metrico una successione* (x_n) *costante,* $x_n = x$ *per ogni* $n \in \mathbf{N}$*, è convergente e* $x_n \to x$*. La definizione di limite è soddisfatta in questo caso da qualunque* $\nu \in \mathbf{N}$.

5.9 *La successione di numeri reali* $a_n = 1/n$*,* $n \geqslant 1$*, è infinitesima. Infatti la definizione di limite è soddisfatta per* $a = 0$ *e* $\nu > 1/\varepsilon$.

5.10 *La successione di numeri reali* $a_n = n/(n+1)$ *converge a 1. Infatti, scelto un* $\varepsilon > 0$*, la disequazione*

$$\left|\frac{n}{n+1} - 1\right| < \varepsilon$$

è soddisfatta per ogni $n > 1/\varepsilon$.

Definizione 5.3 - *La successione* $(a_n) \subset \mathbf{R}$ *è* ***positivamente*** [***negativamente***] ***divergente****, e si scrive*

$$\lim_{n\to\infty} a_n = +\infty \quad o \quad a_n \to +\infty \qquad [\lim_{n\to\infty} a_n = -\infty \quad o \quad a_n \to -\infty]\,,$$

se

$$\forall M \in \mathbf{R}\ \exists \nu \in \mathbf{N}\ :\quad a_n > M\ [a_n < M] \quad \forall n > \nu\,.$$

Una successione che non ammette limite, non convergente quindi, né divergente, viene detta ***indeterminata****.*

Esempi

5.11 *La successione* $a_n = an + b$ *converge solo se* $a = 0$*, in quanto costante, ed ha per limite* b*, altrimenti diverge positivamente se* $a > 0$ *e negativamente se* $a < 0$.

5.12 *La successione dell'Esempio 5.6 è positivamente divergente perché, essendo crescente, se la disuguaglianza* $a_n > M$ *vale per un certo* $\nu > M + 2$ *vale anche per tutti gli* $n > \nu$.

5.13 *Se $\alpha > 0$ $n^\alpha \to +\infty$ perché la disuguaglianza $n^\alpha > M$ vale definitivamente per $n > M^{1/\alpha}$. Il rapporto tra somme di potenze*

$$\frac{n^\alpha + \dots}{n^\beta + \dots},$$

dove compaiono solo i termini di grado massimo, ha lo stesso limite di $n^{\alpha-\beta}$.

Vediamo adesso alcune proprietà generali dei limiti di successioni. Usiamo il termine "generali" per distinguerle da quelle di carattere algebrico o da quelle legate all'ordinamento, strutture in genere assenti in uno spazio metrico astratto.

Teorema 5.4 (Unicità del limite) - *Se il limite di una successione $(a_n) \subset \mathbf{R}$ esiste allora è unico.*

Dimostrazione. Se (a_n) diverge, mettiamo a $+\infty$, non può soddisfare definitivamente una condizione come $a_n < a + \varepsilon$, necessaria per avere limite finito a, o come $a_n < M$ per divergere a $-\infty$. Se $a_n \to a$ e $a_n \to b$ con $a < b$ finiti, scelto $\varepsilon = (b-a)/2$, esistono indici ν_1 e ν_2 tali che

$$a_n < a + \varepsilon = \frac{a+b}{2} \quad \forall n > \nu_1 \qquad \text{e} \qquad a_n > b - \varepsilon = \frac{a+b}{2} \quad \forall n > \nu_2\,.$$

Ma queste due relazioni sono tra loro in contraddizione, risulta infatti

$$\frac{a+b}{2} < a_n < \frac{a+b}{2} \quad \forall n > \max\{\nu_1, \nu_2\}\,.$$

□

L'unicità del limite dipende dalla possibilità, presente in tutti gli spazi metrici, di separare due punti con palle disgiunte. Supponiamo infatti che una successione (x_n) converga simultaneamente a x e x' tra loro distinti. Scelto $\varepsilon < d(x, x')/2$ non può accadere che gli elementi x_n abbiano distanza da x e da x' minore di ε da un certo indice in poi. In altre parole, non possono essere entrambe vere le due condizioni $x_n \in B_\varepsilon(x)$ e $x_n \in B_\varepsilon(x')$ definitivamente se $B_\varepsilon(x) \cap B_\varepsilon(x') = \emptyset$.

Nella seguente definizione si introduce la *condizione di Cauchy*, una proprietà che descrive il reciproco avvicinamento asintotico tra gli elementi di una data successione.

Definizione 5.5 - *Diciamo che $(a_n) \subset \mathbf{R}$ è una* ***successione di Cauchy*** *se*

$$\forall \varepsilon > 0\; \exists \nu \in \mathbf{N} \;:\quad |a_m - a_n| < \varepsilon \quad \forall m, n > \nu\,.$$

Più in generale in uno spazio metrico (X, d), una successione (x_n) è di Cauchy se, come nella definizione precedente, $d(x_m, x_n) < \varepsilon$ definitivamente. Si usa anche la notazione $\lim\limits_{m,n\to\infty} d(x_m, x_n) = 0$.

Esercizio 5.4 - *Usando la disuguaglianza triangolare della distanza, dimostrare che ogni successione convergente è di Cauchy.*

Come sempre di fronte ad una condizione necessaria, possiamo usare la negazione della condizione di Cauchy come condizione sufficiente di non convergenza. Per esempio la successione $a_n = (-1)^n$ non converge perché non è di Cauchy dal momento che $|(-1)^n - (-1)^{n+1}| = 2$ (v. l'Esercizio 5.2).

Teorema 5.6 - *Ogni successione di Cauchy è limitata (quindi ogni successione convergente è limitata). Ogni successione divergente è non limitata (quindi non è di Cauchy).*

Dimostrazione. Sia $(a_n) \subset \mathbf{R}$ di Cauchy e scegliamo $\nu \in \mathbf{N}$ in modo che $|a_n - a_\nu| < 1$ per ogni $n \geqslant \nu$. Poiché

$$||a_n| - |a_\nu|| \leqslant |a_n - a_\nu| < 1 \quad \Rightarrow \quad |a_n| < 1 + |a_\nu|\,,$$

la successione è limitata da ν in poi, d'altra parte i rimanenti $a_1, a_2, \ldots, a_{\nu-1}$ sono un numero finito, quindi è tutta limitata e

$$|a_n| \leqslant \max\{|a_1|, |a_2|, \ldots, |a_{\nu-1}|, 1 + |a_\nu|\}\,.$$

La seconda parte della tesi è banale.

□

Esercizio 5.5 - *Generalizzare il teorema precedente al caso di uno spazio metrico.*

Per affermare che una certa successione (a_n) non converge ad a basta negare la definizione di limite usando le leggi della logica:

$$\exists \varepsilon > 0 \;:\; \forall \nu \in \mathbf{N}\, \exists n > \nu \;:\quad |a_n - a| \geqslant \varepsilon\,.$$

Poiché in questa proposizione l'indice n dipende da ν, si stabilisce un'applicazione $\nu \to n_\nu \geqslant \nu$ da $\mathbf{N}$ in sé, che possiamo supporre strettamente crescente, alla quale corrisponde la sottosuccessione (a_{n_ν}) non convergente ad a. Per negare dunque che a sia il limite di (a_n) basta estrarre da questa una sottosuccessione che non converge ad a. Può accadere in tal caso che (a_n) abbia un altro limite, diverso da a, oppure che sia indeterminata. Di queste, la seconda si verifica se (a_n) ammette almeno due sottosuccessioni con limiti distinti. L'enunciato del seguente teorema può infatti essere interpretato in questo modo (la negazione della tesi implica la negazione dell'ipotesi). La questione, ben più delicata, dell'*esistenza* di sottosuccessioni convergenti verrà analizzata più avanti.

Esercizio 5.6 - *Dimostrare che ogni successione non limitata superiormente ammette una sottosuccessione positivamente divergente.*

Teorema 5.7 - *Se $(a_n) \subset \mathbf{R}$ ammette limite $a \in \bar{\mathbf{R}}$ ogni sua sottosuccessione tende ad a.*

Dimostrazione. Basta osservare che il comportamento definitivo di (a_{k_n}) è lo stesso di quello della (a_n) in quanto $k_n \geqslant n > \nu$ nella definizione di limite.

□

La successione $a_n = (-1)^n$ è indeterminata perché la sottosuccessione con gli indici pari tende a 1 e quella con gli indici dispari tende a -1.

Esercizio 5.7 - *Dimostrare che se (a_n) ha due sottosuccessioni (a_{h_n}) e (a_{k_n}) che ammettono lo stesso limite e $\{h_n\} \cup \{k_n\} = \mathbf{N}$ allora anche tutta la successione (a_n) tende a quel limite.*

Esercizio 5.8 - *Dimostrare che le successioni*

$$\left(\frac{(-1)^n n}{n+1}\right), \quad \left(n \operatorname{sen} \frac{n\pi}{2}\right) \quad e \quad \left(2^{(-1)^n n}\right)$$

sono indeterminate sapendo che se $a > 1$ $a^n \to +\infty$ e $a^{-n} \to 0$.

Passiamo adesso a trattare una serie di teoremi generali sulle successioni reali distinguendo quelli legati all'ordinamento da quelli legati alle operazioni. Dovendo fare una scelta sull'ordine di questi argomenti, per certi esempi utili e interessanti, che vanno affrontati "a caldo" nell'ambito della prima categoria, può essere inevitabile dover applicare anche risultati non ancora dimostrati della seconda. Succederebbe la stessa cosa se presentassimo questi argomenti nell'altro ordine, d'altra parte è preferibile vedere gli esempi insieme alla teoria e non dopo.

5.3 Teoremi di confronto e successioni monotone

In questo paragrafo vediamo come le disuguaglianze tra successioni siano stabili rispetto ai passaggi al limite e, viceversa, in che misura disuguaglianze tra i limiti possano dare informazioni sul confronto tra le successioni. Infine dimostriamo l'esistenza del limite per le successioni monotone.

Teorema 5.8 (I del confronto) - *Siano (a_n) e (b_n) due successioni di numeri reali tali che $a_n \leqslant b_n$ definitivamente. Se $a_n \to a$ e $b_n \to b$, con $a, b \in \bar{\mathbf{R}}$, allora $a \leqslant b$.*

Dimostrazione. Se $b = +\infty$ oppure $a = -\infty$ non c'è niente da dimostrare. Se $a = +\infty$ $\forall M \in \mathbf{R}$ $a_n > M$ definitivamente ed essendo $b_n \geqslant a_n$ anche $b_n > M$ definitivamente. Analogamente $b = -\infty \Rightarrow a = -\infty$.

Se $a, b \in \mathbf{R}$, per ogni $\varepsilon > 0$ si ha definitivamente

$$a - \varepsilon < a_n < a + \varepsilon \quad \text{e} \quad b - \varepsilon < b_n < b + \varepsilon\,,$$

da cui

$$a - \varepsilon < a_n \leqslant b_n < b + \varepsilon\,.$$

Allora $a - \varepsilon < b + \varepsilon$, cioè $a - b < 2\varepsilon$, per ogni $\varepsilon > 0$, quindi $a - b \leqslant 0$ che è la tesi.

□

Corollario 5.9 (I della permanenza del segno) - *Se $b_n \to b$ e definitivamente $b_n \geqslant 0$ allora $b \geqslant 0$.*

Dimostrazione. Basta applicare il teorema precedente con $a_n = 0$ per ogni $n \in \mathbf{N}$.

□

È possibile dedurre il segno della successione dal segno del suo limite? Bisognerebbe invertire l'implicazione del Corollario 5.9. Certamente ogni informazione sul limite, avendo carattere asintotico, non può dirci niente sul comportamento di *tutta* la successione, ma solo sui suoi elementi da un certo indice in poi. Inoltre, trattandosi della determinazione del segno della successione, se il suo limite fosse 0 non si potrebbe dedurre nulla, come è evidente con la $((-1)^n/n)$ che è a segni alterni e infinitesima. Dobbiamo dunque assumere come ipotesi che il limite non sia nullo.

Teorema 5.10 (II della permanenza del segno) - *Se $a_n \to a > 0$ allora $a_n > 0$ definitivamente.*

Dimostrazione. Se $a_n \to +\infty$, scelto $M = 0$ nella definizione di successione divergente si ottiene $a_n > 0$ definitivamente. Se $a_n \to a \in \mathbf{R}$ scegliamo $\varepsilon = a/2$. Allora da un certo indice in poi si ha

$$a_n > a - \frac{a}{2} = \frac{a}{2} > 0\,.$$

□

Esempi

5.14 *Se $a > 1$ la successione (a^n) diverge a $+\infty$. Infatti, posto $a = 1 + h$ con $h > 0$, per la disuguaglianza di Bernoulli si ha*

$$a^n = (1+h)^n \geqslant nh \to +\infty\,.$$

Se $a < -1$ (a^n) è indeterminata perché $a^{2n} \to +\infty$ mentre $a^{2n+1} = a^{2n}a \to -\infty$. Se $a = 1$ è costante e se $a = -1$ è indeterminata. Se infine $|a| < 1$, $a^n \to 0$ perché $|a|^{-1} > 1$ e per ogni $\varepsilon > 0$ si ha, da quanto appena visto, $|1/a|^n > 1/\varepsilon$ definitivamente, quindi $|a^n| < \varepsilon$ definitivamente.

5.15 *Se $a > 1$ la successione (a^n/n) diverge a $+\infty$. Ciò segue dalla formula del binomio per $(1+t)^n$, che, se troncata al I grado, fornisce la disuguaglianza di Bernoulli e il limite del caso precedente, se troncata al II fa al caso nostro*

$$a^n = \sum_{k=0}^{n} \binom{n}{k} t^k > \binom{n}{2} t^2 = \frac{n(n-1)}{2} t^2 \,,$$

da cui $a^n/n > (n-1)t^2/2 \to +\infty$.

5.16 *Più in generale, se $a > 1$ e $\alpha > 0$ è un numero reale, si ha $a^n/n^\alpha \to +\infty$, basta scrivere la successione nel modo seguente*

$$\frac{a^n}{n^\alpha} = \left(\frac{(a^{1/\alpha})^n}{n} \right)^\alpha$$

e ci si riconduce al caso precedente. Altrimenti basta usare la formula del binomio fino al termine in t^m con $m > \alpha$

$$a^n = \sum_{k=0}^{n} \binom{n}{k} t^k > \binom{n}{m} t^m \,,$$

dove il coefficiente binomiale a destra è un polinomio in n di grado m, quindi

$$\frac{a^n}{n^\alpha} > \frac{t^m}{n^\alpha} \binom{n}{m} \sim n^{m-\alpha} \to +\infty \,.$$

Teorema 5.11 (II del confronto) - *Siano (a_n), (b_n) e (c_n) tre successioni di numeri reali tali che $a_n \leqslant b_n \leqslant c_n$ definitivamente. Se (a_n) e (c_n) hanno lo stesso limite $l \in \mathbf{R}$ allora anche (b_n) tende a l.*

Dimostrazione. Si ha definitivamente

$$l - \varepsilon < a_n \leqslant b_n \leqslant c_n < l + \varepsilon \,,$$

quindi $b_n \to l$.

$\square$

Questo risultato, che molti con un po' di senso dell'humor chiamano *teorema dei carabinieri*, è estremamente utile quando la successione da studiare ha un aspetto apparentemente complicato, ma soddisfa delle stime dall'alto e dal basso con successioni note, o più semplici, convergenti allo stesso limite.

Esempi

5.17 *La successione $\sqrt[n]{n}$ converge a 1. Essendo infatti $\sqrt[n]{n} > 1$, posto $a_n = \sqrt[n]{n} - 1$, si ha*

$$n = (a_n + 1)^n > \frac{n(n-1)}{2} a_n^2 \,,$$

da cui, per $n > 1$, si ricava

$$a_n^2 < \frac{2}{n-1} \to 0 \,.$$

Essendo poi $\sqrt[n]{n^\alpha} = (\sqrt[n]{n})^\alpha$, anche questa successione tende a 1 e se, infine, (a_n) è una qualunque successione tale che $n^\alpha \leqslant a_n \leqslant n^\beta$, si ha $\sqrt[n]{a_n} \to 1$.

5.18 *La successione $(\sqrt[n]{a})$ tende a 1 per ogni $a \geqslant 1$. Infatti $1 \leqslant \sqrt[n]{a} \leqslant \sqrt[n]{n}$. Si capisce che il limite è ancora 1 anche se $0 < a < 1$, basta ragionare con $1/a$ al posto di a e applicare il teorema algebrico sulla succssione degli inversi che vedremo nel prossimo paragrafo.*

5.19 *Sebbene* $(\operatorname{sen} n)$ *e* $(\cos n)$ *siano successioni indeterminate, dalle disuguaglianze*

$$\frac{n-1}{n+1} \leqslant \frac{n+\operatorname{sen} n}{n+\cos n} \leqslant \frac{n+1}{n-1}$$

segue che

$$\lim_{n\to\infty} \frac{n+\operatorname{sen} n}{n+\cos n} = 1\,.$$

5.20 *Se* $a, b \geqslant 0$ $\sqrt[n]{a^n+b^n} \to \max\{a,b\}$. *Posto* $a \geqslant b$, *basta osservare che*

$$\sqrt[n]{a} \leqslant \sqrt[n]{a^n+b^n} \leqslant a\sqrt[n]{2}\,.$$

Qual è il limite, se esiste, della successione $(\sqrt[n]{a^n-b^n})$ *con* $a \geqslant |b|$*?*

5.21 *Per ogni successione* $(x_n) \subset \mathbf{R}$ *convergente a* $x \in \mathbf{R}$ *si ha*

$$\lim_{n\to\infty} \operatorname{sen} x_n = \operatorname{sen} x \qquad e \qquad \lim_{n\to\infty} \cos x_n = \cos x\,.$$

Il primo limite si ottiene dal confronto

$$0 \leqslant |\operatorname{sen} x_n - \operatorname{sen} x| = 2\left|\cos\frac{x_n+x}{2}\right|\left|\operatorname{sen}\frac{x_n-x}{2}\right| \leqslant |x_n - x|$$

essendo $|\cos t| \leqslant 1$ *e* $|\operatorname{sen} t| \leqslant |t|$ *per ogni* $t \in \mathbf{R}$. *Il secondo si può ottenere in modo simile, oppure ricordando che*

$$\cos t = \operatorname{sen}\left(\frac{\pi}{2} - t\right)\,.$$

5.22 *Da semplici considerazioni geometriche si ricavano le ben note disuguaglianze*

$$|\operatorname{sen} x| \leqslant |x| \leqslant |\operatorname{tang} x| \qquad \forall x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\,.$$

Tenendo presente che in questo intervallo x *e* $\operatorname{sen} x$ *sono concordi e che* $\cos x > 0$, *dividendo per* $|\operatorname{sen} x|$ *si ottiene*

$$1 \leqslant \frac{x}{\operatorname{sen} x} \leqslant \frac{1}{\cos x}\,.$$

Dunque per ogni successione $x_n \to 0$, *poiché* $\cos x_n \to 1$, *si ha*

$$\lim_{n\to\infty} \frac{x_n}{\operatorname{sen} x_n} = 1\,.$$

Concludiamo questo paragrafo dimostrando che le successioni monotone hanno sempre limite, possono essere convergenti o divergenti, ma mai indeterminate. Questa proprietà, come vedremo, ha notevoli applicazioni al calcolo dei limiti.

Teorema 5.12 - *Una successione* (a_n) *crescente* [*decrescente*] *di numeri reali ammette limite e*

$$\lim_{n\to\infty} a_n = \sup_{n\in\mathbf{N}} a_n \qquad [\lim_{n\to\infty} a_n = \inf_{n\in\mathbf{N}} a_n]\,.$$

Dimostrazione. Trattiamo solo il caso (a_n) crescente, l'altro è analogo, e sia $\alpha = \sup a_n \in \bar{\mathbf{R}}$. Se $\alpha \in \mathbf{R}$, per ogni $\varepsilon > 0$ scegliamo $k \in \mathbf{N}$ tale che $a_k > \alpha - \varepsilon$. In quanto crescente, la (a_n) soddisfa

$$\alpha - \varepsilon < a_k \leqslant a_n \leqslant \alpha \quad \forall n > k \tag{5.3}$$

che è la definizione di limite per la (a_n). La (5.3) mostra tra l'altro che (a_n) approssima α dal basso.

Se $\alpha = +\infty$ (certamente non può essere $\alpha = -\infty$ dato che (a_n) ha per minimo a_1) si ragiona come prima: brevemente, da $a_k > M$ per un certo $k \in \mathbf{N}$ segue $a_n \geqslant a_k > M$ per ogni $n > k$.

□

Esercizio 5.9 - *Se $a_n > 0$ e $a_n/a_{n-1} \leqslant c < 1$ per ogni $n \in \mathbf{N}$ allora $a_n \to 0$.*

Esercizio 5.10 - *Dimostrare che se $a_n \geqslant 0$ per ogni $n \in \mathbf{N}$ la successione delle relative somme $s_n = a_0 + a_1 + a_2 + \ldots + a_n$ è convergente o positivamente divergente, ma non può essere indeterminata.*

5.4 Il numero e di Nepero, esponenziale e logaritmo in base e

Applichiamo il Teorema 5.12 all'importante successione

$$a_n = \left(1 + \frac{1}{n}\right)^n, \qquad n \geqslant 1, \tag{5.4}$$

per dimostrarne la convergenza verso il numero e di Nepero. Precisamente dimostriamo che

($\nearrow$) (a_n) è crescente,

$(2-3)$ $2 \leqslant a_n < 3$ per ogni $n \geqslant 1$,

$(\mathbf{R} - \mathbf{Q})$ e è irrazionale.

Una stima approssimata per difetto del numero e è 2,718281828459045, di questo numero si sa non solo che è irrazionale, ma trascendente.

Per dimostrare la ($\nearrow$) un metodo particolarmente semplice si basa sulla disuguaglianza tra la media geometrica e la media aritmetica

$$\sqrt[n]{a_1 a_2 \ldots a_n} \leqslant \frac{a_1 + a_2 + \ldots + a_n}{n}$$

che ricaveremo più avanti. Se applicata agli $n+1$ numeri

$$1\,,\ \underbrace{1 + \frac{1}{n}\,,\ \ 1 + \frac{1}{n}\,,\ \ \ldots\,,\ \ 1 + \frac{1}{n}}_{n}$$

si ottiene

$$\sqrt[n+1]{\left(1 + \frac{1}{n}\right)^n} \leqslant \frac{1}{n+1}\left[(n+1)\cdot 1 + n \cdot \frac{1}{n}\right] = \left(1 + \frac{1}{n+1}\right)$$

ed elevando alla $n+1$

$$\left(1 + \frac{1}{n}\right)^n \leqslant \left(1 + \frac{1}{n+1}\right)^{n+1}.$$

In particolare la successione ha minimo, pari a 2, per $n = 1$.

Esercizio 5.11 - *Verificare con lo stesso ragionamento che anche la successione $(1 - 1/n)^n$ è crescente.*

Per la $(2-3)$ usiamo la formula del binomio

$$\begin{aligned}\left(1 + \frac{1}{n}\right)^n &= \sum_{k=0}^{n} \binom{n}{k} \frac{1}{n^k} = \sum_{k=0}^{n} \frac{n(n-1)\ldots(n-k+1)}{k!n^k} \\ &< \sum_{k=0}^{n} \frac{1}{k!} < 1 + \sum_{k=1}^{n} \frac{1}{2^{k-1}} < 1 + \sum_{k=0}^{\infty} \frac{1}{2^k} = 3\,.\end{aligned}$$

Ciò mostra non solo che converge la successione di Nepero, ma anche la successione di somme

$$s_n = \sum_{k=0}^{n} \frac{1}{k!}$$

verso un numero $s \leqslant 3$ in quanto anch'essa crescente e limitata. Posto allora

$$s = \lim_{n\to\infty} s_n = \sum_{n=0}^{\infty} \frac{1}{n!}\,,$$

vale la stima $e \leqslant s$. In realtà vediamo subito che vale anche la disuguaglianza opposta, per cui il numero di Nepero è anche somma della *serie esponenziale*

$$e = \sum_{n=0}^{\infty} \frac{1}{n!}\,. \tag{5.5}$$

Fissiamo ad arbitrio un certo $m \in \mathbf{N}$. Per ogni $n > m$ si ha

$$\begin{aligned} e &\geqslant \sum_{k=0}^{n} \frac{1}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right) \\ &\geqslant \sum_{k=0}^{m} \frac{1}{k!}\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\dots\left(1-\frac{k-1}{n}\right). \end{aligned}$$

Passando al limite per $n \to \infty$ nella seconda sommatoria si ottiene

$$e \geqslant \sum_{k=0}^{m} \frac{1}{k!} = s_m$$

e passando di nuovo al limite su m si ottiene

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

Rimane da dimostrare che e è irrazionale. Questo si basa sulla stima

$$e - s_n < \frac{1}{n!n} \tag{5.6}$$

che può essere ottenuta in questo modo

$$\begin{aligned} e - s_n &= \sum_{k=n+1}^{\infty} \frac{1}{k!} = \frac{1}{(n+1)!}\sum_{k=n+1}^{\infty}\frac{(n+1)!}{k!} = \frac{1}{(n+1)!}\sum_{h=0}^{\infty}\frac{(n+1)!}{(n+1+h)!} \\ &= \frac{1}{(n+1)!}\sum_{h=0}^{\infty}\frac{1}{(n+1+h)(n+h)\dots(n+2)} < \frac{1}{(n+1)!}\sum_{h=0}^{\infty}\frac{1}{(n+1)^h} \\ &= \frac{1}{(n+1)!}\cdot\frac{1}{1-\dfrac{1}{n+1}} = \frac{1}{(n+1)!}\cdot\frac{n+1}{n} = \frac{1}{n!n}\,. \end{aligned}$$

Ora, se e fosse razionale, $e = p/q$ con $p, q \geqslant 1$ primi tra loro, il numero $q!(e - s_q)$ sarebbe naturale, ma ciò contraddice la (5.6) che richiede

$$q!(e - s_q) < \frac{1}{q}\,.$$

Dimostriamo adesso che anche se si sostituisce la variabile intera n con una qualunque successione divergente (x_n) si ha comunque

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

Per ricondurre questo caso generale a quello di una variabile intera si ricorre ancora una volta al Teorema 5.11. Ora, dall'Esercizio 5.11 segue che la successione $(1+1/n)^{n+1}$ decresce, infatti

$$\left(1+\frac{1}{n}\right)^{n+1}=\frac{1}{\left(\dfrac{n}{n+1}\right)^{n+1}}=\frac{1}{\left(1-\dfrac{1}{n+1}\right)^{n+1}}\,,$$

e, poiché

$$\left(1+\frac{1}{n}\right)^{n+1}=\left(1+\frac{1}{n}\right)^{n}\left(1+\frac{1}{n}\right)\to e\,,$$

deve tendere ad e dall'alto. D'altra parte è evidente che

$$\left(1+\frac{1}{n+1}\right)^{n}=\frac{\left(1+\dfrac{1}{n+1}\right)^{n+1}}{\left(1+\dfrac{1}{n+1}\right)}\to e$$

dal basso in quanto crescente. In definitiva

$$\left(1+\frac{1}{n+1}\right)^{n}<e<\left(1+\frac{1}{n}\right)^{n+1} \tag{5.7}$$

dove le due successioni sono monotone e convergenti ad e.

Se adesso $x_n\to+\infty$, indicando con $[x_n]$ la sua parte intera, si ha

$$\left(1+\frac{1}{1+[x_n]}\right)^{[x_n]}\leqslant\left(1+\frac{1}{x_n}\right)^{x_n}\leqslant\left(1+\frac{1}{[x_n]}\right)^{1+[x_n]}\,.$$

Per confronto, da quanto detto sopra, quella centrale deve convergere ad e. Se $x_n\to-\infty$ poniamo $y_n=-x_n\to+\infty$. Si ha

$$\left(1+\frac{1}{x_n}\right)^{x_n}=\left(1-\frac{1}{y_n}\right)^{-y_n}=\left(1+\frac{1}{y_n-1}\right)^{y_n}\to e\,.$$

Esercizio 5.12 - *Per ogni successione $x_n\to0$ (che comporta $x_n>-1$ almeno definitivamente) si ha $(1+x_n)^{1/x_n}\to e$.*

Osserviamo che per ogni $x\in\mathbf{R}$ si ha

$$\lim_{n\to\infty}\left(1+\frac{x}{n}\right)^{n}=\lim_{n\to\infty}\left[\left(1+\frac{x}{n}\right)^{n/x}\right]^{x}=e^{x}\,. \tag{5.8}$$

Da ora in poi per *funzione esponenziale* intenderemo sempre quella con base e, la $x\to e^x$, a meno di avviso contrario. Il numero e verrà preso da ora in poi anche come base dei logaritmi, la funzione $x\to\log x$ indicherà il logaritmo in base e, detto anche *logaritmo naturale*. Della funzione esponenziale conosciamo già alcune proprietà, che è positiva, crescente e non limitata superiormente. In più, adesso, usando la rappresentazione (5.8) possiamo far vedere che è convessa. Definiamo la successione di funzioni $f_n : \mathbf{R} \to [0, +\infty[$

$$f_n(x) = \begin{cases} \left(1 + \dfrac{x}{n}\right)^n & \text{se } x \geqslant -n \\ 0 & \text{se } x < -n\,, \end{cases}$$

per ogni $n \geqslant 1$. Per ogni $x \in \mathbf{R}$ $f_n(x) \to e^x$ e da note proprietà di convessità delle potenze segue che ogni f_n è convessa su $\mathbf{R}$, cioè

$$f_n(\lambda x_1 + (1-\lambda)x_2) \leqslant \lambda f_n(x_1) + (1-\lambda) f_n(x_2) \qquad \forall x_1, x_2 \in \mathbf{R}, \lambda \in [0,1]\,,$$

e passando al limite su n si ottiene

$$e^{\lambda x_1 + (1-\lambda)x_2} \leqslant \lambda e^{x_1} + (1-\lambda)e^{x_2}\,.$$

Vediamo alcune conseguenze importanti dell'Esercizio 5.12. Essendo convergente ad e la successione $(1+x_n)^{1/x_n}$ nell'ipotesi $x_n \to 0$, certamente è limitata superiormente da una costante $c \geqslant e$

$$(1+x_n)^{1/x_n} \leqslant c$$

e passando al logaritmo, che è crescente, si ottiene

$$\frac{\log(1+x_n)}{x_n} \leqslant \log c = C\,, \qquad C \geqslant 1\,.$$

Ne segue che

$$|\log(1+x_n)| \leqslant C|x_n|\,, \tag{5.9}$$

pertanto

$$\lim_{n\to\infty} \log(1+x_n) = 0$$

per ogni successione $x_n \to 0$. Allora $\log x_n \to \log x$ per ogni successione $x_n \to x > 0$, infatti $x_n/x \to 1$, quindi

$$|\log x_n - \log x| = \left|\log \frac{x_n}{x}\right| \to 0\,.$$

In particolare

$$(1+x_n)^{1/x_n} \to e \Rightarrow \frac{\log(1+x_n)}{x_n} \to 1\,. \tag{5.10}$$

Da questo deriva un altro limite da tenere presente. Si pone $y_n = \log(1+x_n)$ che è infinitesima, per cui il passaggio al limite precedente diventa

$$\frac{y_n}{e^{y_n} - 1} \to 1$$

per ogni $y_n \to 0$. Allora anche per la funzione esponenziale vale la proprietà $e^{x_n} \to e^x$ per ogni $x_n \to x$, se infatti si pone nella precedente $y_n = |x_n - x|$, per cui

$$\frac{e^{|x_n - x|} - 1}{|x_n - x|} \to 1\,,$$

allora

$$|e^{x_n} - e^x| = e^x |e^{x_n - x} - 1| \leqslant e^x (e^{|x_n - x|} - 1) \leqslant Ce^x |x_n - x| \tag{5.11}$$

che è infinitesima per $x_n \to x$.

Dovendo calcolare un limite del tipo

$$\lim_{n\to\infty} (1 + x_n)^{y_n} ,$$

dove (x_n) è infinitesima e (y_n) divergente, è utile passare alla forma esponenziale

$$\lim_{n\to\infty} e^{y_n \log(1+x_n)}$$

e calcolare poi il limite dell'esponente sfruttando la (5.10)

$$\lim_{n\to\infty} y_n \log(1 + x_n) = \lim_{n\to\infty} x_n y_n \frac{\log(1 + x_n)}{x_n} = \lim_{n\to\infty} x_n y_n .$$

Se questo limite esiste, indichiamolo con l, per la (5.11) il limite richiesto sarà e^l se l è finito, altrimenti $+\infty$ se $l = +\infty$ e 0 se $l = -\infty$.

Esempi

5.23 *Si ha* $(\log n)/n \to 0$*. È un caso particolare del logaritmo di una successione che tende a* 1*, per cui*

$$\frac{\log n}{n} = \log \sqrt[n]{n} \to 0 .$$

Sostituendo a n una qualunque successione che tende a $+\infty$ il limite non cambia. Ne segue che per ogni $\alpha, \beta > 0$, posto $\gamma = \beta/\alpha$, si ha

$$\frac{\log^\alpha n}{n^\beta} = \left(\frac{\log n}{n^\gamma}\right)^\alpha = \left(\frac{\log n^\gamma}{\gamma n^\gamma}\right)^\alpha \to 0 .$$

Per $\alpha \leqslant 0$ è ovvio e per $\beta \leqslant 0$ è falso.

5.24 *Dalla* (5.9) *segue* $\log(n + 1)/\log n \to 1$*. Infatti*

$$\frac{\log(n+1)}{\log n} = \frac{\log \dfrac{n+1}{n} + \log n}{\log n} = 1 + \frac{1}{\log n} \log \frac{n+1}{n} \to 1 .$$

5.5 Proprietà algebriche dei limiti

In questo paragrafo facciamo vedere in che modo i passaggi al limite commutano con le operazioni tra le successioni. Per la stabilità della somma l'insieme delle successioni convergenti è uno spazio vettoriale e con in più quella del prodotto è un'algebra. Rientrano nell'ambito dei risultati algebrici anche alcune proprietà delle successioni divergenti.

Teorema 5.13 - *Siano $(a_n), (b_n) \subset \mathbf{R}$ due successioni convergenti ad $a, b \in \mathbf{R}$ rispettivamente. Allora la successione $(a_n + b_n)$ converge ad $a + b$.*

Dimostrazione. Per ogni $\varepsilon > 0$ si ha definitivamente

$$a - \frac{\varepsilon}{2} < a_n < a + \frac{\varepsilon}{2} \qquad \text{e} \qquad b - \frac{\varepsilon}{2} < b_n < b + \frac{\varepsilon}{2}$$

e sommando membro a membro

$$a + b - \varepsilon < a_n + b_n < a + b + \varepsilon$$

definitivamente.

□

Teorema 5.14 - *Se (a_n) è infinitesima e (b_n) limitata allora $(a_n b_n)$ è infinitesima.*

Dimostrazione. Se $|b_n| \leqslant k$ per ogni $n \in \mathbf{N}$, scelto un $\varepsilon > 0$ si ha definitivamente $|a_n| < \varepsilon/k$, quindi

$$|a_n b_n| < k \cdot \frac{\varepsilon}{k} = \varepsilon$$

definitivamente.

□

Teorema 5.15 - *Se (a_n) converge ad $a \in \mathbf{R}$ e $\lambda \in \mathbf{R}$ la successione (λa_n) converge a λa.*

Dimostrazione. Basta osservare che

$$|\lambda a_n - \lambda a| = |\lambda||a_n - a|$$

dove il secondo membro è il prodotto di $|a_n - a|$ che è infinitesima per la successione limitata che assume il valore costante λ.

□

Teorema 5.16 - *Siano $(a_n), (b_n) \subset \mathbf{R}$ due successioni convergenti ad $a, b \in \mathbf{R}$ rispettivamente. Allora la successione $(a_n b_n)$ converge ad ab.*

Dimostrazione. Basta scrivere la successione prodotto nella forma

$$a_n b_n = (a_n - a) b_n + a b_n$$

e applicare i teoremi precedenti.

□

Per le operazioni inverse, la differenza rientra nel caso della somma, mentre il rapporto è oggetto del prossimo teorema.

Lemma 5.17 - *Se $a_n \to a \in \mathbf{R}$ allora $|a_n| \to |a|$.*

Dimostrazione. Basta osservare che

$$||a_n| - |a|| \leqslant |a_n - a| \,.$$

□

L'implicazione contraria vale in generale solo quando $a = 0$, se $|a_n| \to a > 0$ può esistere una sottosuccessione di (a_n) convergente ad a ed un'altra convergente a $-a$.

Lemma 5.18 - *Se (a_n) converge ad un numero reale $a \neq 0$ allora*

$$\lim_{n\to\infty} \frac{1}{a_n} = \frac{1}{a} \,.$$

Dimostrazione. Non è strettamente necessario supporre $a_n \neq 0$, lo sarà certamente in modo definitivo per il Teorema 5.10 e trattandosi di una questione riguardante il limite ci basta che la successione degli inversi sia definita da un certo indice in poi.

Per il Lemma 5.17 $|a_n| > |a|/2$ definitivamente, quindi se, sempre definitivamente, $|a_n - a| < \varepsilon$ allora

$$\left| \frac{1}{a_n} - \frac{1}{a} \right| = \frac{|a - a_n|}{|a_n a|} < \frac{2\varepsilon}{a^2} \,.$$

□

Teorema 5.19 - *Siano $(a_n), (b_n) \subset \mathbf{R}$ due successioni convergenti ad $a, b \in \mathbf{R}$ rispettivamente con $b \neq 0$. Allora la successione dei rapporti (a_n/b_n) converge ad a/b.*

Dimostrazione. Basta usare insieme il Teorema 5.16 e il Lemma 5.18.

□

Vediamo cosa si può dire sui passaggi al limite per le successioni divergenti.

Esercizio 5.13 - *Se (a_n) diverge, oppure se $a_{h_n} \to -\infty$ e $a_{k_n} \to +\infty$ con $(a_n) = (a_{h_n}) \cup (a_{k_n})$, allora $|a_n| \to +\infty$ e $1/a_n \to 0$. Viceversa $1/|a_n| \to +\infty$ se $a_n \to 0$.*

Esercizio 5.14 - *Costruire esempi di successioni divergenti o indeterminate con somma e prodotto convergente.*

Teorema 5.20 - *Se $a_n \to +\infty$ e (b_n) è limitata inferiormente (ad esempio convergente) allora $a_n + b_n \to +\infty$.*

Dimostrazione. Sia $\beta \in \mathbf{R}$ un minorante della (b_n). Per ogni $M \in \mathbf{R}$ $a_n > M - \beta$ definitivamente, quindi

$$a_n + b_n > (M - \beta) + \beta = M$$

definitivamente.

□

Analogamente $a_n + b_n \to -\infty$ se $a_n \to -\infty$ e (b_n) è limitata superiormente. Ad esempio, la successione $(n + \operatorname{sen} n)$ diverge a $+\infty$ perché, essendo $\operatorname{sen} n > -1$, e $n + \operatorname{sen} n > M$ è vera definitivamente, per $n > M - (-1) = M + 1$. La situazione in cui due successioni divergono in senso opposto non rientra nelle ipotesi del teorema precedente, quindi ogni caso richiede uno studio specifico. Ad esempio $n^2 - n \to +\infty$, $n - n \to 0$ e $n - n^2 \to -\infty$.

Teorema 5.21 - *Se $a_n \to +\infty$ e $\inf\{b_n \mid n \in \mathbf{N}\} > 0$ (che è come dire che $b_n \geqslant \beta > 0$ per ogni $n \in \mathbf{N}$) allora $a_n b_n \to +\infty$.*

Dimostrazione. Sia $\beta > 0$ un minorante della (b_n), ad esempio il suo estremo inferiore. Per ogni $M \in \mathbf{R}$ $a_n > M/\beta$ definitivamente, quindi

$$a_n b_n > \frac{M}{\beta}\beta = M$$

definitivamente.

□

Se la (b_n) ha infiniti termini di un segno e dell'altro, ma $|b_n| \geqslant \beta > 0$ per ogni $n \in \mathbf{N}$, e $a_n \to +\infty$ allora $a_n b_n$ è indeterminata. Se entrambe divergono il prodotto diverge. Infine va osservato che se si lascia cadere l'ipotesi che la (b_n), di segno costante, che rimanga "lontana" da 0, (b_n) *uniformemente positiva*, allora potrebbe avere una sottosuccessione infinitesima, ma il comportamento del prodotto di due successioni, una divergente e l'altra infinitesima, non è contemplato to da nessun teorema algebrico, ogni caso va studiato a parte.

Esercizio 5.15 - *Siano (a_n) e (b_n) due successioni di numeri reali tali che*

$$0 < a_n \leqslant c \leqslant b_n \quad \forall n \in \mathbf{N}\,,$$

dove $c \in \mathbf{R}$ è una costante assegnata. Dimostrare che

$$\lim_{n\to\infty} \frac{a_n}{b_n} = 1 \Leftrightarrow \lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = c\,.$$

L'implicazione "$\Leftarrow$" è conseguenza immediata del Teorema 5.19. Per la "$\Rightarrow$", se $b_n/a_n \to 1$ la successione

$$\sigma_n = \frac{b_n}{a_n} - 1 \quad \forall n \in \mathbf{N}$$

è infinitesima. Moltiplicando per a_n, cosa possibile per ipotesi, si ottiene

$$b_n - a_n = \sigma_n a_n$$

che tende a 0 perché prodotto di una infinitesima per una limitata. La tesi segue allora per confronto

$$0 < b_n - c \leqslant b_n - a_n\,, \qquad 0 < c - a_n \leqslant b_n - a_n\,.$$

Va osservato che senza quella sorta di "barriera", c, che separa le due successioni è facile costruire coppie di successioni divergenti o addirittura indeterminate il cui rapporto tende a 1.

Un altro esercizio interessante è il seguente.

Esercizio 5.16 - *Data una successione* $(t_n) \subset \mathbf{R}$ *tale che* $0 < t_n < 1$*, come si comporta la successione* (a_n) *dei prodotti*

$$a_n = t_1 t_2 \dots t_n\,?$$

Certamente possiamo dire che decresce strettamente in quanto $a_n = t_n a_{n-1} < a_{n-1}$ e dal suo valore iniziale $t_1 < 1$ converge verso un limite $a \in [0,1[$. Se $a > 0$, segue dal Teorema 5.19 che $t_n \to 1$, quindi in tutti i casi in cui (t_n) non tende a 1 si ha $a_n \to 0$. Il viceversa però non è detto, $t_n = 1 - 1/n \to 1$, ma $a_n = 1/n \to 0$.

La successione $t_n = |\operatorname{sen} nx|$ non ha limite se $x \neq 0, \pi$, quindi per tali x

$$a_n = \operatorname{sen} x \cdot \operatorname{sen} 2x \cdots \operatorname{sen} nx \to 0\,.$$

Sono detti di *indeterminazione* tutti quei casi in cui le successioni coinvolte nelle operazioni non soddisfano le ipotesi dei teoremi algebrici. Ogni caso richiede uno studio specifico, prima di tutto sulla natura della successione data, ma che tenga anche ben presenti tutti i risultati di questo capitolo tra i quali vanno scelti volta per volta i più utili. Nei casi di indeterminazione ci siamo già imbattuti trattando vari esempi e sono i seguenti

$$[\infty - \infty]\,, \ \left[\frac{0}{0}\right]\,, \ \left[\frac{\infty}{\infty}\right]\,, \ [0 \cdot \infty]\,, \ [0^0]\,, \ [\infty^0]\,, \ [1^\infty]\,,$$

i quali vanno presi come puri simboli, non certo come valori di limiti, e sono legati tra loro in modo evidente. Per stabilire il comportamento delle potenze del tipo $a_n^{b_n}$ è sempre utile passare alla forma esponenziale $e^{b_n \log a_n}$, in questo modo si vede come l'ultimo caso sia in realtà lo stesso del quarto. I restanti casi di indeterminazione sono solo quelli riguardanti successioni di cui non sappiamo dire nulla.

5.6 La funzione esponenziale complessa

Viene spontaneo definire la funzione esponenziale $z \to e^z$, con $z \in \mathbf{C}$, in questo modo

$$e^z = \lim_{n\to\infty} \left(1 + \frac{z}{n}\right)^n \qquad \forall z \in \mathbf{C}\,. \tag{5.12}$$

Ricordiamo che nello spazio metrico $\mathbf{C}$ la convergenza di una successione (z_n) a z è così definita:

$$\lim_{n\to\infty} z_n = z \Leftrightarrow \lim_{n\to\infty} |z_n - z| = 0\,.$$

Tornando alla (5.12) dobbiamo dimostrare che quel limite esiste e calcolarlo. A questo scopo conviene separare modulo e argomento. Posto $z = x + iy$, si ha

$$|e^z| = \left|1 + \frac{x+iy}{n}\right|^n = \left(1 + \frac{2x}{n} + \frac{x^2+y^2}{n^2}\right)^{n/2} \to e^x \quad \text{per} \quad n \to \infty\,.$$

Per quanto riguarda l'argomento, facciamo riferimento all'*argomento principale*

$$\pi < \operatorname{Arg} z < \pi$$

da cui dipendono gli altri tramite le relazioni

$$\arg z = \operatorname{Arg} z + 2k\pi\,, \qquad k \in \mathbf{Z}\,.$$

Nel nostro caso

$$\arg\left(1+\frac{z}{n}\right)^n = n\arg\left(1+\frac{z}{n}\right) + 2k\pi = n\operatorname{Arg}\left(1+\frac{z}{n}\right) + 2k\pi\,.$$

Osserviamo che per n sufficientemente grande

$$-\frac{\pi}{2} < \operatorname{Arg}\left(1+\frac{z}{n}\right) < \frac{\pi}{2}$$

dato che $z/n \to 0$, quindi

$$n\operatorname{Arg}\left(1+\frac{z}{n}\right) = n\operatorname{arctg}\frac{y/n}{1+x/n}$$

e passando al limite

$$\lim_{n\to\infty} n\operatorname{Arg}\left(1+\frac{z}{n}\right) = \lim_{n\to\infty} n\operatorname{arctg}\frac{y/n}{1+x/n} = y\,.$$

Allora

$$|e^z| = e^x \qquad \text{e} \qquad \arg e^z = y + 2k\pi\,,\; k \in \mathbf{Z}\,,$$

quindi

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

che per $x = 0$ porta alla celebre e sorprendente relazione tra l'esponenziale sui numeri immaginari e le funzioni trigonometriche

$$e^{iy} = \cos y + i\operatorname{sen} y\,. \tag{5.13}$$

Esercizio 5.17 *Dimostrare che* $e^{z+z'} = e^z e^{z'}$ *e di conseguenza riconoscere le formule di De Moivre* $(e^{iy})^n = e^{iny}$.

Combinata la (5.13) con l'analoga $e^{-iy} = \cos y - i\operatorname{sen} y$, otteniamo le *formule di Eulero*

$$\cos y = \frac{e^{iy}+e^{-iy}}{2} \qquad \text{e} \qquad \operatorname{sen} y = \frac{e^{iy}-e^{-iy}}{2i}\,.$$

Ponendo $y = \pi$ nella (5.13) si ottiene la bella uguaglianza

$$e^{i\pi} + 1 = 0$$

che mette in relazione tra loro i numeri più famosi della matematica: e, i, π, 1, 0. Il fisico Richard Feynman l'ha chiamata il "gioiello di Eulero".

Esercizio 5.18 *Dimostrare che* $\overline{e^z} = e^{\bar z}$.

Esercizio 5.19 *Dimostrare che* $|e^z| \le e^{|z|}$.

Esercizio 5.20 *Usando la* (5.12) *dimostrare che* $|e^z - 1| \le e^{|z|} - 1$.

Esercizio 5.21 *Dimostrare che per ogni successione* $(z_n) \subset \mathbf{C}$ *e per ogni* $z \in \mathbf{C}$ *se* $z_n \to z$ *allora* $e^{z_n} \to e^z$. *Dimostrare inoltre che se* $z_n \to 0$ *si ha*

$$\lim_{n\to\infty} \frac{e^{z_n} - 1}{z_n} = 1\,.$$

Le formule di Eulero richiamano le espressioni, contenenti l'esponenziale reale, che definiscono le funzioni iperboliche e tutte quante ci suggeriscono le seguenti estensioni al campo complesso

$$(5.14)\qquad \begin{cases} \cosh z = \dfrac{e^z + e^{-z}}{2} \\ \operatorname{senh} z = \dfrac{e^z - e^{-z}}{2} \end{cases} \qquad \text{e} \qquad \begin{cases} \cos z = \dfrac{e^{iz} + e^{-iz}}{2} \\ \operatorname{sen} z = \dfrac{e^{iz} - e^{-iz}}{2i} \end{cases}.$$

Si osservi che per ogni $z \in \mathbf{C}$

$$(5.15)\qquad \cosh iz = \cos z \qquad \text{e} \qquad \operatorname{senh} iz = i \operatorname{sen} z\,,$$

inoltre

$$\cos^2 z + \operatorname{sen}^2 z = 1 \qquad \text{e} \qquad \cosh^2 z - \operatorname{senh}^2 z = 1\,,$$

ma, a differenza del caso reale, cos e sen non sono limitate, basta osservare che $\cos iy = \cosh y$ lungo tutto l'asse immaginario.

La funzione esponenziale complessa è surgettiva da $\mathbf{C}$ in $\mathbf{C} - \{0\}$, ma non è invertibile come nel caso reale perché $y \to e^{x+iy}$ è periodica per ogni $x \in \mathbf{R}$. Però, così come abbiamo definito l'argomento come funzione multivoca, altrettanto si può fare per il logaritmo, come inversa dell'esponenziale:

$$w = e^z \Rightarrow z = \log w\,.$$

In accordo con le usuali proprietà del logaritmo e tenendo presente che $|w| = e^x$ e $\arg w = y + 2k\pi$, dovrà essere

$$\log w = x + iy = \log e^{x+iy} = \log e^x + \log e^{iy} = \log|w| + i \arg w \quad \forall w \in \mathbf{C} - \{0\}\,.$$

Ad esempio

$$\log(-1) = \log|-1| + i\arg(-1) = \log 1 + i(\pi + 2k\pi) = (2k+1)\pi i\,, \quad k \in \mathbf{Z}\,,$$

$$\log(1+i) = \log\sqrt{2} + i(\pi/4 + 2k\pi)\,, \quad k \in \mathbf{Z}\,.$$

Una particolare determinazione del logaritmo è la funzione *logaritmo principale*

$$\operatorname{Log} w = \log|w| + i \operatorname{Arg} w\,,$$

definita su $\mathbf{C} - \{0\}$, esclusa la semiretta dei w reali negativi.

Con il logaritmo possiamo calcolare ogni potenza, in generale anch'essa a più valori

$$z^\alpha = e^{\alpha \log z}\,.$$

Ad esempio

$$i^i = e^{i\log i} = e^{i(i\pi/2 + i2k\pi)} = e^{-\pi(2k+1/2)}\,, \quad k \in \mathbf{Z}\,,$$

$$i^\pi = e^{\pi \log i} = e^{i\pi(\pi/2 + 2k\pi)} = e^{i\pi^2(1/2+2k)}\,, \quad k \in \mathbf{Z}\,,$$

$$\pi^i = e^{i(\log\pi + i2k\pi)} = e^{-2k\pi}(\cos\log\pi + i\operatorname{sen}\log\pi)\,, \quad k \in \mathbf{Z}\,,$$

$$(-1)^i = e^{i\log(-1)} = e^{-(\pi+2k\pi)}\,, \quad k \in \mathbf{Z}\,.$$

Esercizio 5.22 - *Verificare che*

$$i^{i^i} = e^{ie^{-\frac{\pi}{2}-2h\pi}\left(\frac{\pi}{2}+2k\pi\right)}\,, \quad h, k \in \mathbf{Z}\,.$$

5.7 Successioni di medie e applicazioni

A partire da una successione (a_n) di numeri reali, possiamo definire quella delle *medie aritmetiche*

$$\mu_n = \frac{a_1 + a_2 + \ldots + a_n}{n}.$$

Teorema 5.22 - *Se $a_n \to a \in \mathbf{R}$ allora $\mu_n \to a$.*

Dimostrazione. Supponiamo dapprima che (a_n) sia infinitesima. Scelto arbitrariamente $\varepsilon > 0$, sia $k \in \mathbf{N}$ tale che $|a_n| < \varepsilon/2$ per ogni $n > k$. Osserviamo che

$$\begin{aligned}|\mu_n| &\leqslant \frac{|a_1 + a_2 + \ldots + a_k|}{n} + \frac{|a_{k+1}| + |a_{k+2}| + \ldots + |a_n|}{n} \\ &< \frac{|a_1 + a_2 + \ldots + a_k|}{n} + \frac{(n-k)\varepsilon}{2n} < \frac{|a_1 + a_2 + \ldots + a_k|}{n} + \frac{\varepsilon}{2}.\end{aligned}$$

D'altra parte esiste un altro indice, $k' \in \mathbf{N}$, tale che

$$\frac{|a_1 + a_2 + \ldots + a_k|}{n} < \frac{\varepsilon}{2}$$

per ogni $n > k'$. Per ogni $n > \max\{k, k'\}$ si ha allora $|\mu_n| < \varepsilon$.

Se adesso supponiamo che il limite di (a_n) sia un numero reale a, $a_n - a \to 0$, quindi

$$\mu_n - a = \frac{(a_1 - a) + (a_2 - a) + \ldots + (a_n - a)}{n} \to 0,$$

cioè $\mu_n \to a$.

$\square$

Esercizio 5.23 - *Dimostrare che la media aritmetica è inferiore alla media quadratica*

$$\chi_n = \sqrt{\frac{a_1^2 + a_2^2 + \ldots + a_n^2}{n}}.$$

Per le successioni positive hanno senso anche le medie

$$\textit{geometrica}: \ \gamma_n = \sqrt[n]{a_1 a_2 \ldots a_n} \quad \text{e} \quad \textit{armonica}: \ \nu_n = \frac{1}{\frac{1}{n}\left(\frac{1}{a_1} + \frac{1}{a_2} + \ldots + \frac{1}{a_n}\right)}.$$

Vogliamo dimostrare che se $a_n > 0$ per ogni $n \in \mathbf{N}$ vale la relazione

$$\nu_n \leqslant \gamma_n \leqslant \mu_n$$

e che hanno tutte lo stesso limite nel caso lo abbia a_n.

Lemma 5.23 - *Dati n numeri reali $x_1, x_2, \ldots, x_n$ non negativi tali che*

$$x_1 + x_2 + \ldots + x_n = n,$$

si ha

$$x_1 x_2 \cdots x_n \leqslant 1. \tag{5.16}$$

Dimostrazione. La tesi è banalmente verificata se uno di questi numeri è nullo oppure se sono tutti uguali a 1, l'unico caso in cui la (5.16) vale con l'uguaglianza. D'altra parte non possono essere tutti maggiori o tutti minori di 1, quindi se uno di essi è minore di 1 ne esiste un altro che lo supera.

La (5.16) è ovvia per $n = 1$. Supponiamola vera per n e dimostriamola per $n+1$ numeri $x_1, x_2, \ldots, x_{n+1}$ tali che

$$x_1 + x_2 + \ldots + x_{n+1} = n + 1\,. \tag{5.17}$$

Ordiniamo gli $n+1$ numeri in modo che $x_1 < 1$ e $x_2 > 1$. Portando l'1 a sinistra nella (5.17) si ottengono n numeri con somma n

$$(x_1 + x_2 - 1) + x_3 + \ldots + x_{n+1} = n$$

dove $x_1 + (x_2 - 1) > 0$, quindi

$$(x_1 + x_2 - 1)x_3 \cdots x_{n+1} \leqslant 1\,. \tag{5.18}$$

D'altra parte vale la disuguaglianza $(1 - x_1)(x_2 - 1) > 0$ equivalente alla

$$x_1 + x_2 - 1 > x_1 x_2\,,$$

quindi la (5.18) continua a valere con $x_1 x_2$ al posto di $x_1 + x_2 - 1$.

□

Teorema 5.24 - $\gamma_n \leqslant \mu_n$.

Dimostrazione. Dalla definizione di μ_n si ha

$$\frac{a_1}{\mu_n} + \frac{a_2}{\mu_n} + \ldots + \frac{a_n}{\mu_n} = n$$

e dal Lemma 5.23 si ottiene

$$\frac{a_1 a_2 \cdots a_n}{\mu_n^n} \leqslant 1$$

da cui la tesi.

□

A questo punto la disuguaglianza $\nu_n \leqslant \gamma_n$ si ottiene subito applicando quella appena ottenuta agli inversi di a_n. Va osservato che le disuguaglianze ottenute, nel Teorema 5.24 così come nel Lemma 5.23, diventano uguaglianze se e solo se gli n numeri considerati sono tutti uguali.

Teorema 5.25 - *Se $a_n \to a \in \mathbf{R}$, $a_n > 0$ per ogni $n \in \mathbf{N}$, allora $\nu_n, \gamma_n, \mu_n \to a$.*

Dimostrazione. La successione $(1/\nu_n)$ è la media aritmetica degli inversi di a_n, per cui

$$\begin{aligned}
&a_n \to +\infty \Rightarrow \frac{1}{a_n} \to 0 \Rightarrow \frac{1}{\nu_n} \to 0 \Rightarrow \nu_n \to +\infty \Rightarrow \gamma_n \to +\infty \\
&a_n \to 0 \Rightarrow \mu_n \to 0 \Rightarrow \nu_n, \gamma_n \to 0 \\
&a_n \to a \in \mathbf{R} \Rightarrow \frac{1}{a_n} \to \frac{1}{a} \Rightarrow \frac{1}{\nu_n} \to \frac{1}{a} \Rightarrow \nu_n \to a \Rightarrow \gamma_n \to a\,.
\end{aligned}$$

□

La relazione fra le 3 medie della successione (n) è la seguente

$$n\left(1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{n}\right)^{-1} \leqslant \sqrt[n]{n!} \leqslant \frac{n+1}{2}\,. \tag{5.19}$$

Tutte quante divergono con crescita al più lineare rispetto a n e in particolare

$$\lim_{n\to\infty} \frac{1}{n}\left(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}\right)=0\,,$$

del resto è la media aritmetica di $1/n$. Se ne deduce che la successione $1+1/2+1/3+\ldots+1/n$ ha crescita strettamente meno rapida di quella lineare, ma ancora non è chiaro se converge o diverge. Verificheremo che ha lo stesso andamento di $\log n$. Un'altra conseguenza della (5.19) è che $\sqrt[n]{n!}/n \leqslant 1/2$ e fra poco vedremo che converge.

Come conseguenza valgono i seguenti risultati dovuti a Cesàro.

Proposizione 5.26 - *Se $a_n - a_{n-1} \to a \in \bar{\mathbf{R}}$ allora $a_n/n \to a$.*

Dimostrazione. Basta osservare che

$$\frac{a_n}{n} = \frac{(a_n-a_{n-1})+(a_{n-1}-a_{n-2})+\ldots+(a_2-a_1)}{n}+\frac{a_1}{n}\,.$$

□

Alla luce di questo risultato si può rivedere ad esempio il caso della successione (a^n/n) con $a>1$

$$\lim_{n\to\infty}\frac{a^n}{n}=\lim_{n\to\infty}(a^{n+1}-a^n)=\lim_{n\to\infty}a^n(a-1)=+\infty\,.$$

Proposizione 5.27 - *Se $a_n > 0$ per ogni $n \in \mathbf{N}$ e $a_n/a_{n-1} \to a \in \bar{\mathbf{R}}$ allora $\sqrt[n]{a_n} \to a$.*

Dimostrazione. Basta osservare che

$$\sqrt[n]{a_n} = \sqrt[n]{\frac{a_n}{a_{n-1}}\cdot\frac{a_{n-1}}{a_{n-2}}\cdots\frac{a_2}{a_1}}\,\sqrt[n]{a_1}\,.$$

□

Ad esempio

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

Possiamo stimare la rapidità con cui questa successione converge a $1/e$ in termini di una successione di confronto? Citiamo a tal proposito la *formula di Stirling*

$$\lim_{n\to\infty}\frac{n!e^n}{n^n\sqrt{n}}=\sqrt{2\pi}$$

da cui risulta

$$\frac{\sqrt[n]{n!}}{n}\sim\frac{\sqrt[2n]{2\pi n}}{e}\,.$$

Un risultato più preciso, reperibile ad esempio qui, è il seguente

$$\frac{n!e^n}{n^n\sqrt{n}}\sim e^{\vartheta_n}\sqrt{2\pi}\,,\qquad 0<\vartheta_n<\frac{1}{12n}\,.$$

5.8 Punti limite di una successione

Nel trattare la teoria dei limiti per le successioni reali un paio di questioni, almeno, sono sorte e meritano una risposta. Le conseguenze sono notevoli nell'ambito dell'Analisi e per farsi un'idea chiara e completa sarebbe necessario analizzarle nell'ambito più generale degli spazi metrici, cosa che faremo senz'altro, ma per ora stiamo in $\mathbf{R}$ o, al massimo, $\mathbf{C}$ o in $\mathbf{R}^n$, con la distanza euclidea. Nel Teorema 5.6 del § 5.2 abbiamo dimostrato che ogni successione convergente è limitata. Il viceversa è ovviamente falso, ma siccome nello stesso paragrafo abbiamo anche osservato che ogni successione non limitata ammette una sottosuccessione divergente, è lecito chiedersi

- *una successione limitata ammette sottosuccessioni convergenti?*

Sempre in quel paragrafo, è stato chiesto in un esercizio di dimostrare che ogni successione convergente è di Cauchy. Ci poniamo allora la domanda

- *ogni successione di Cauchy è convergente?*

Scopo di questo paragrafo è rispondere alle due domande precedenti e definire i punti limite di una successione, in particolare il massimo e il minimo limite.

È evidente che una successione che assume un numero finito di valori, quindi limitata, deve assumere infinite volte uno stesso valore (altrimenti avrebbe un numero finito di indici) e la sottosuccessione con quel valore converge in quanto costante. Per una successione limitata qualunque l'esistenza di sottosuccessioni convergenti è meno banale ed è oggetto del seguente teorema.

Teorema 5.28 - *Per ogni successione limitata (x_n) di numeri reali esistono una sottosuccessione (x_{k_n}) ed un numero reale x tali che $x_{k_n} \to x$.*

Dimostrazione. Siano $a, b \in \mathbf{R}$ una limitazione inferiore e superiore della (x_n) in modo che l'intervallo $[a, b]$ contenga tutta la successione. Il punto medio $m_0 = (a+b)/2$ divide $[a, b]$ nei due intervalli $[a, m_0]$ e $[m_0, b]$ dei quali andiamo a scegliere quello che contiene infiniti elementi della successione; se entrambi ne contengono infiniti ne scegliamo uno a caso o con una regola a piacimento. Poniamo allora

$$\begin{cases} a_1 = a \text{ e } b_1 = m_0 \text{ se si sceglie } [a, m_0] \\ a_1 = m_0 \text{ e } b_1 = b \text{ se si sceglie } [m_0, b] \end{cases}$$

e sia k_1 il primo indice della successione tale che $x_{k_1} \in [a_1, b_1]$. Poi si considera il punto medio m_1 dell'intervallo $[a_1, b_1]$ e dei due intervalli in cui rimane diviso si sceglie quello che contiene infiniti elementi della successione, oppure uno dei due come sopra se entrambi ne contengono infiniti. Si pone

$$\begin{cases} a_2 = a_1 \text{ e } b_2 = m_1 \text{ se si sceglie } [a_1, m_1] \\ a_2 = m_1 \text{ e } b_2 = b_1 \text{ se si sceglie } [m_1, b_1] \end{cases}$$

e sia $k_2 > k_1$ il primo indice della successione, successivo a k_1, tale che $x_{k_2} \in [a_2, b_2]$. Si prosegue sempre in questo modo ottenendo una successione decrescente di intervalli $[a_n, b_n]$ inscatolati ognuno nel precedente e una sottosuccessione (x_{k_n}) della (x_n) selezionata con la regola che k_n sia il primo indice maggiore di k_{n-1} tale che $x_{k_n} \in [a_n, b_n]$. Gli estremi sinistri di questi intervalli formano dunque una successione (a_n) limitata e crescente, quindi convergente ad un certo $\alpha \in \mathbf{R}$, mentre gli estremi destri formano una successione limitata e decrescente e quindi convergente ad un certo $\beta \in \mathbf{R}$ ed essendo $a_n < b_n$ si ha $\alpha \leqslant \beta$. Inoltre $a_n \leqslant x_{k_n} \leqslant b_n$ per ogni $n \in \mathbf{N}$. Ogni intervallo ha lunghezza metà dell'intervallo precedente che lo contiene

$$b_n - a_n = \frac{b_{n-1} - a_{n-1}}{2} = \frac{b_{n-2} - a_{n-2}}{4} = \ldots = \frac{b_1 - a_1}{2^{n-1}} = \frac{b - a}{2^n}$$

e passando al limite per $n \to \infty$ nell'ultimo termine si deduce che $b_n - a_n \to 0$, da cui $\alpha = \beta$. Per il teorema dei carabinieri anche (x_{k_n}) deve tendere a questo valore. Abbiamo così costruito un'estratta convergente dalla successione data.

$\square$

Esercizio 5.24 - *Usare il Teorema 5.28 per dimostrare che anche in* $\mathbf{R}^n$ *ogni successione limitata ammette una sottosuccessione convergente.*

Il seguente esempio mostra che il Teorema 5.28 è vero solo negli spazi euclidei di dimensione finita. Consideriamo in $\mathbf{R}^{\mathbf{N}}$, che ha per elementi le stringhe ordinate di numeri reali del tipo $x = (x_1, x_2, \ldots, x_n, \ldots)$ con un'infinità numerabile di coordinate, la successione

$$\begin{aligned} e_1 &= (1, 0, 0 \ldots, 0, \ldots) \\ e_2 &= (0, 1, 0 \ldots, 0, \ldots) \\ e_3 &= (0, 0, 1 \ldots, 0, \ldots) \\ &\vdots \\ e_n &= (0, 0, 0 \ldots, 0, \overset{posto\, n}{1}, 0 \ldots) \\ &\vdots \end{aligned}$$

Questi vettori hanno tutti lo stesso modulo 1 e formano un insieme limitato, ma ognuno di essi ha distanza $\sqrt{2}$ da tutti gli altri, quindi (e_n) non ha sottosuccessioni convergenti perché non possono essere di Cauchy.

Il Teorema 5.28 risponde affermativamente alla prima questione, vediamo se e come risponde alla seconda.

Teorema 5.29 - *Per ogni successione di Cauchy* $(a_n) \subset \mathbf{R}$ *esiste* $a \in \mathbf{R}$ *tale che* $a_n \to a$.

Dimostrazione. Se (a_n) è di Cauchy per ogni $\varepsilon > 0$ esiste $\nu_1 \in \mathbf{N}$ tale che

$$|a_n - a_m| < \frac{\varepsilon}{2} \qquad \forall m, n > \nu_1 \,.$$

Ma in quanto di Cauchy è anche limitata e per il Teorema 5.28 ammette una sottosuccessione (a_{k_n}) convergente ad un certo $a \in \mathbf{R}$, quindi esiste $\nu_2 \in \mathbf{N}$ tale che

$$|a_{k_n} - a| < \frac{\varepsilon}{2} \qquad \forall n > \nu_2 \,.$$

Allora per ogni $n \geqslant \nu = \max\{\nu_1, \nu_2\}$, tenendo presente che $k_n \geqslant n > \nu$, si ha

$$|a_n - a| \leqslant |a_n - a_{k_n}| + |a_{k_n} - a| < \frac{\varepsilon}{2} + \frac{\varepsilon}{2} = \varepsilon \,,$$

quindi tutta la successione (a_n), non solo l'estratta (a_{k_n}), converge ad a.

$\square$

I vari limiti delle sottosuccessioni convergenti di una successione (a_n) si chiamano *punti limite* della (a_n). Se essa già converge ne ammette uno solo, altrimenti può avere più limiti, addirittura infiniti come la successione $(\operatorname{sen} n)$. Ogni punto limite a di una successione (a_n) è caratterizzato dalla seguente proprietà

$$\forall \varepsilon > 0 \; \forall n \in \mathbf{N} \; \exists k_n \; : \; |a_{k_n} - a| < \varepsilon \,.$$

La proprietà degli elementi a_n di avere distanza da a inferiore a ε non è definitiva, ma *ricorrente*, ossia vera per infiniti elementi, è invece, come sappiamo, definitiva per una sottosuccessione che ammette a come limite. Il Teorema 5.28 stabilisce che ogni successione limitata, anche se non convergente, deve ammettere dei punti limite. Il più piccolo e il più grande dei punti limite si chiamano rispettivamente *minimo limite* e *massimo limite* e si indicano con

$$\liminf_{n\to\infty} a_n \quad \text{e} \quad \limsup_{n\to\infty} a_n\,.$$

Per una successione non limitata inferiormente [superiormente] poniamo

$$\liminf_{n\to\infty} a_n = -\infty \qquad [\limsup_{n\to\infty} a_n = +\infty]$$

che è in accordo con l'esistenza di sottosuccessioni divergenti. Dunque ogni successione ammette sempre massimo e minimo limite, i quali, quando sono finiti, si possono caratterizzare nel seguente modo

$$a' = \liminf_{n\to\infty} a_n \Leftrightarrow \begin{cases} \forall\varepsilon > 0\ \exists\nu \in \mathbf{N}\ :\ a_n > a' - \varepsilon\ \forall n > \nu \\ \forall\varepsilon > 0\ \forall n \in \mathbf{N}\ \exists k_n > n\ :\ a_{k_n} < a' + \varepsilon\,, \end{cases} \tag{5.20}$$

$$a'' = \limsup_{n\to\infty} a_n \Leftrightarrow \begin{cases} \forall\varepsilon > 0\ \exists\nu \in \mathbf{N}\ :\ a_n < a'' + \varepsilon\ \forall n > \nu \\ \forall\varepsilon > 0\ \forall n \in \mathbf{N}\ \exists k_n > n\ :\ a_{k_n} > a'' - \varepsilon\,. \end{cases} \tag{5.21}$$

È sempre vera la relazione

$$\liminf_{n\to\infty} a_n \leqslant \limsup_{n\to\infty} a_n$$

e vale l'uguaglianza se e solo se (a_n) ammette limite, coincidente in questo caso col valore comune del massimo e del minimo limite.

Esempi

5.25 *La successione $a_n = (-1)^n$ ammette come punti limite solo -1, il minimo limite, e 1 che è il massimo limite.*

5.26 *Le successioni $a_n = n^{(-1)^n}$ e $b_n = n(1 + (-1)^n)$ ammettono come punti limite solo 0, il minimo limite, e $+\infty$, il massimo limite.*

5.27 *Le successioni $a_n = 2 + n^{\cos(\pi n/2)}$ e*

$$b_n = \begin{cases} n & \textit{se } n = 3k \\ n/n+1 & \textit{se } n = 3k+1 \\ 1/n & \textit{se } n = 3k+2\,, \quad k \in \mathbf{N}\,, \end{cases}$$

ammettono tre punti limite

$$\liminf_{n\to\infty} a_n = 2\,, \quad \liminf_{n\to\infty} b_n = 0\,, \quad e \quad \limsup_{n\to\infty} a_n = \limsup_{n\to\infty} b_n = +\infty\,.$$

5.28 *L'insieme dei punti limite della successione $(\operatorname{sen} n)$ è l'intervallo $[-1, 1]$, ma non è semplice dimostrarlo.*

Esercizio 5.25 - *Calcolare i punti limite delle successioni, $x \in \mathbf{R}$,*

$$a_n = \frac{n!x^n}{n^n}\operatorname{sen}\frac{n\pi}{2}\,, \qquad b_n = \operatorname{arctg}(-2)^n\,, \qquad c_n = \left(1 + \frac{(-1)^n}{n}\right)^n\,.$$

Data una successione $(a_n) \in \mathbf{R}$, le seguenti due successioni

$$\alpha_n' = \inf_{k\geqslant n} a_k \qquad \text{e} \qquad \alpha_n'' = \sup_{k\geqslant n} a_k$$

sono monotone, la prima crescente e la seconda decrescente, quindi hanno limite e

$$\lim_{n\to\infty} \alpha_n' = \sup_{n\in\mathbf{N}} \inf_{k\geqslant n} a_k \qquad \text{e} \qquad \lim_{n\to\infty} \alpha_n'' = \inf_{n\in\mathbf{N}} \sup_{k\geqslant n} a_k\,.$$

Usando le caratterizzazioni (5.20) e (5.21) non è difficile dimostrare che

$$a' = \liminf_{n\to\infty} a_n = \sup_{n\in\mathbf{N}} \inf_{k\geqslant n} a_k \qquad \text{e} \qquad a'' = \limsup_{n\to\infty} a_n = \inf_{n\in\mathbf{N}} \sup_{k\geqslant n} a_k\,.$$

Esercizio 5.26 - *Data una successione $(a_n) \in \mathbf{R}$, dimostrare che*

$$\inf_{n\in\mathbf{N}} a_n \leqslant \liminf_{n\to\infty} a_n \leqslant \limsup_{n\to\infty} a_n \leqslant \sup_{n\in\mathbf{N}} a_n\,.$$

Esercizio 5.27 - *Ogni successione (a_{k_n}) estratta da (a_n) soddisfa*

$$\liminf_{n\to\infty} a_n \leqslant \liminf_{n\to\infty} a_{k_n} \leqslant \limsup_{n\to\infty} a_{k_n} \leqslant \limsup_{n\to\infty} a_n\,.$$

Esercizio 5.28 - *Siano (a_n) e (b_n) due successioni tali che $a_n \leqslant b_n$ definitivamente. Allora*

$$\liminf_{n\to\infty} a_n \leqslant \liminf_{n\to\infty} b_n \qquad e \qquad \limsup_{n\to\infty} a_n \leqslant \limsup_{n\to\infty} b_n\,.$$

Esercizio 5.29 - *Se $a_n > 0$ per ogni $n \in \mathbf{N}$ si ha*

$$\liminf_{n\to\infty} \frac{1}{a_n} = \frac{1}{\limsup\limits_{n\to\infty} a_n} \qquad e \qquad \liminf_{n\to\infty} \frac{1}{a_n} = \frac{1}{\limsup\limits_{n\to\infty} a_n}$$

con la convenzione $1/\infty = 0$ e $1/0 = \infty$. La seconda relazione è immediata conseguenza della prima.

Esercizio 5.30 - *Per ogni $\lambda \in \mathbf{R}$ si ha*

$$\liminf_{n\to\infty} \lambda a_n = \begin{cases} \lambda \liminf\limits_{n\to\infty} a_n & se\ \lambda \geqslant 0 \\ \lambda \limsup\limits_{n\to\infty} a_n & se\ \lambda < 0\,. \end{cases}$$

Proprietà analoga vale per il massimo limite.

Esercizio 5.31 - *Se (a_n) e (b_n) sono due successioni e la prima è limitata inferiormente*

$$\liminf_{n\to\infty} a_n + \limsup_{n\to\infty} b_n \leqslant \limsup_{n\to\infty}(a_n + b_n) \leqslant \limsup_{n\to\infty} a_n + \limsup_{n\to\infty} b_n$$

e analogamente

$$\liminf_{n\to\infty} a_n + \liminf_{n\to\infty} b_n \leqslant \liminf_{n\to\infty}(a_n + b_n) \leqslant \limsup_{n\to\infty} a_n + \liminf_{n\to\infty} b_n$$

se (a_n) è limitata superiormente.

Esercizio 5.32 - *Se (a_n) e (b_n) sono non negative allora*

$$\liminf_{n\to\infty} a_n \liminf_{n\to\infty} b_n \leqslant \liminf_{n\to\infty} a_n b_n \leqslant \limsup_{n\to\infty} a_n b_n \leqslant \limsup_{n\to\infty} a_n \limsup_{n\to\infty} b_n\,.$$

Ricavare proprietà analoghe nel caso siano non positive o di segno opposto.

Esercizio 5.33 - *Trovare esempi di successioni per le quali le disuguaglianze nei due esercizi precedenti valgono in senso stretto.*

Capitolo 6

Serie numeriche

6.1 Le serie come successioni di somme, esempi

Che cosa significa sommare infiniti numeri? Proviamo a sommare tutte le infinite potenze del numero 2 con esponente naturale e indichiamo con S la loro somma

$$S = 1 + 2 + 2^2 + 2^3 + \ldots + 2^{n-1} + 2^n + \ldots .$$

Se si moltiplicano per 2 tutti i termini si ottiene

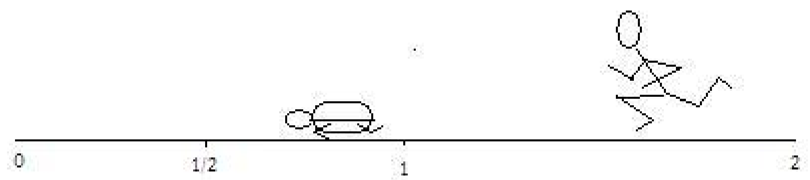
$$2S = 2 + 2^2 + 2^3 + \ldots + 2^n + 2^{n+1} + \ldots = S - 1$$

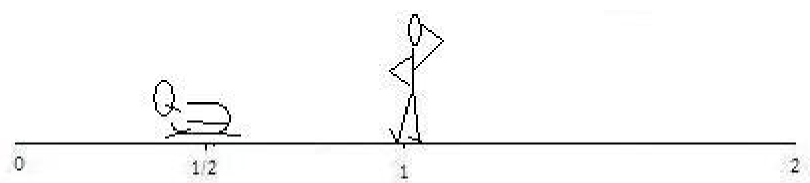
da cui si ricava $S = -1$. Possiamo accettare che una somma di infiniti termini positivi valga -1? Certamente no, ma questo procedimento funziona con la somma

$$S = 1 + \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots + \frac{1}{2^{n-1}} + \frac{1}{2^n} + \ldots$$

che, se divisa per 2, porta alla relazione $S/2 = S - 1$ verificata solo per $S = 2$.

Come si spiega che il valore di una somma di infiniti numeri possa essere finito? Nulla di strano se si ritiene ragionevole accettare la possibilità di suddividere una quantità finita in un'infinità di parti, altrimenti ci si imbatte in paradossi come quello di Zenone: *Achille piè veloce non può raggiungere la tartaruga.*





Un modo sensato di definire questa operazione è quello di passare al limite su una successione di somme finite, allora si noterà che la prima somma considerata diverge, mentre la seconda converge. Precisamente, a partire da una successione (a_n)

di numeri reali (o, come vedremo, eventualmente complessi), si costruisce la nuova successione (s_n) di somme finite

$$s_n = a_0 + a_1 + a_2 + \ldots + a_n = \sum_{k=0}^{n} a_k$$

e poi si calcola, se esiste, il limite

$$S = \lim_{n\to\infty} s_n \,. \tag{6.1}$$

Supponiamo che questo limite esista. Che S sia un numero oppure $+\infty$ o $-\infty$, è naturale che gli venga attribuito il significato di somma di *tutti* gli a_n e si scrive

$$S = \sum_{n=0}^{\infty} a_n \,. \tag{6.2}$$

Ma in ogni caso, che esista o non esista il limite (6.1), il simbolo usato nella (6.2) indica anche la stessa successione (s_n), la successione delle *somme parziali*, e prende il nome di *serie* degli a_n, i quali si chiamano invece *termini della serie.* Il simbolo introdotto nella (6.2) indica due cose diverse, la (s_n) e il suo limite S quando esiste, e questo può generare confusione, ma dal contesto il suo significato sarà chiaro ogni volta. Diciamo che la serie $\sum a_n$ è

convergente con somma S	*se $S \in \mathbf{R}$,*
divergente positivamente [negativamente]	*se $S = +\infty$ $[S = -\infty]$,*
indeterminata	*se S non esiste.*

Non c'è differenza concettuale tra serie e successione, ogni serie è per definizione una successione, la (s_n), e ogni successione (a_n) può essere scritta come serie

$$a_n = a_0 + \sum_{k=1}^{n} (a_k - a_{k-1}) \,. \tag{6.3}$$

Nella gran parte dei casi è molto difficile calcolare la somma di una serie, ma è già interessante di per sé stabilirne il carattere ed eventualmente, nel caso convergente, cercare di ottenerne una stima o un'approssimazione. È possibile calcolarla, invece, quando si riesce a rappresentare in modo esplicito la (s_n) mediante un'espressione finita, trasformando così lo studio della serie nel calcolo di un limite tradizionale. Il calcolo della somma è possibile anche in alcune circostanze fortunate, ma meglio dire miracolose, quando si scoprono sorprendenti collegamenti tra la serie che vogliamo studiare e tecniche o concetti della Matematica talvolta piuttosto lontani, almeno in apparenza.

Come esempio facile di calcolo della somma consideriamo la *serie del Mengoli*

$$s_n = \frac{1}{1\cdot 2} + \frac{1}{2\cdot 3} + \ldots + \frac{1}{n(n+1)} = \sum_{k=1}^{n} \frac{1}{k(k+1)}$$

dove si riesce a scrivere la (s_n) come somma delle "differenze"

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

come nella (6.3). Il calcolo del suo limite S è immediato

$$s_n = 1 - \frac{1}{n+1} \to 1 = S \,.$$

Le serie di questo tipo si chiamano *telescopiche.*

Esercizio 6.1 - *Calcolare la somma della serie telescopica*

$$\sum_{n=2}^{\infty} \frac{n-1}{n!}\,.$$

Un altro esempio di calcolo esplicito è quello della *serie geometrica* di *ragione* x

$$\sum_{n=0}^{\infty} x^n\,, \qquad x \in \mathbf{R}\,, \tag{6.4}$$

di cui i due esempi visti all'inizio rientrano come casi particolari. La successione delle somme parziali è quella dell'Esempio 2.2

$$s_n(x) = 1 + x + x^2 + \ldots + x^n = \begin{cases} n+1 & \text{se } x = 1 \\ \dfrac{1-x^{n+1}}{1-x} & \text{se } x \neq 1\,. \end{cases}$$

Passando al limite per $n \to \infty$ si vede che diverge a $+\infty$ per $x \geq 1$ ed è indeterminata per $x \leqslant -1$; se ad esempio $x = -1$ si ha

$$s_n(-1) = 1 - 1 + 1 - 1 + \ldots + (-1)^n = \frac{1-(-1)^{n+1}}{2} = \begin{cases} 1 & \text{se } n \text{ è pari} \\ 0 & \text{se } n \text{ è dispari} \end{cases} \tag{6.5}$$

che è limitata, non lo è invece $(s_n(x))$ né superiormente, né inferiormente, se $x < -1$. Per $|x| < 1$ la (s_n) converge ed ha per limite

$$S(x) = \frac{1}{1-x} \tag{6.6}$$

che è dunque la somma della serie (6.4).

Esercizio 6.2 - *Usare la* (6.6) *per trasformare in forma frazionaria i numeri decimali periodici* $1,\overline{1}$ *e* $12,3\overline{456}$.

La (6.4) rientra come caso particolare di serie del tipo

$$\sum_{n=0}^{\infty} a_n x^n\,, \qquad x \in \mathbf{R}\,,$$

che studieremo più avanti, ma in relazione all'esercizio precedente osserviamo che per $x = 1/10$, se $a_0 \in \mathbf{Z}$ e i coefficienti a_n, per n da 1 in poi, assumono valori interi da 0 a 9, la sua somma non è altro che il numero reale con allineamento decimale $a_0, a_1 a_2 \ldots a_n \ldots$.

Esercizio 6.3 - *Calcolare il limite della successione*

$$a_n = \frac{1}{11} - \frac{1}{19} + \frac{1}{11^2} - \frac{1}{19^2} + \ldots + \frac{1}{11^n} - \frac{1}{19^n}$$

La versione bidimensionale dell'insieme di Cantor, di cui abbiamo parlato alla fine del § 2.9, forse più facile da capire con l'aiuto di una figura, è anch'essa un'applicazione interessante della serie geometrica.

Problema 6.1 - *Da una regione quadrata di lato 1 si elimina il quadrato centrale di lato* $1/3$, *quindi la parte restante è formata da 8 qudrati di lato* $1/3$. *Poi, se da ciascuno di questi si elimina il quadrato centrale di lato* $1/9$, *rimangono nel quadrato iniziale* $8 \times 8 = 64$ *quadrati di lato* $1/9$, *si prosegue poi sempre nello stesso modo. Ciò che alla fine rimane si chiama Tappeto di Sierpinski. Qual è l'area complessiva dei quadrati eliminati?*

Sommando tutte le aree dei quadrati eliminati si ottiene

$$\begin{aligned} 1\cdot\left(\frac{1}{3}\right)^2+8\cdot\left(\frac{1}{9}\right)^2+8^2\cdot\left(\frac{1}{27}\right)^2+\ldots &= \sum_{n=0}^{\infty} 8^n\cdot\left(\frac{1}{3^2}\right)^{n+1} \\ &= \frac{1}{9}\sum_{n=0}^{\infty}\left(\frac{8}{9}\right)^n = \frac{1}{9(1-8/9)} = 1\,. \end{aligned}$$

Rimane sorprendente che la parte eliminata abbia la stessa area del quadrato iniziale e che ciò che rimane, un tappeto tutto bucherellato, abbia potenza del continuo come il quadrato iniziale! Il ragionamento può essere generalizzato a qualunque dimensione (spugna di Sierpinski in 3 dimensioni ecc.)

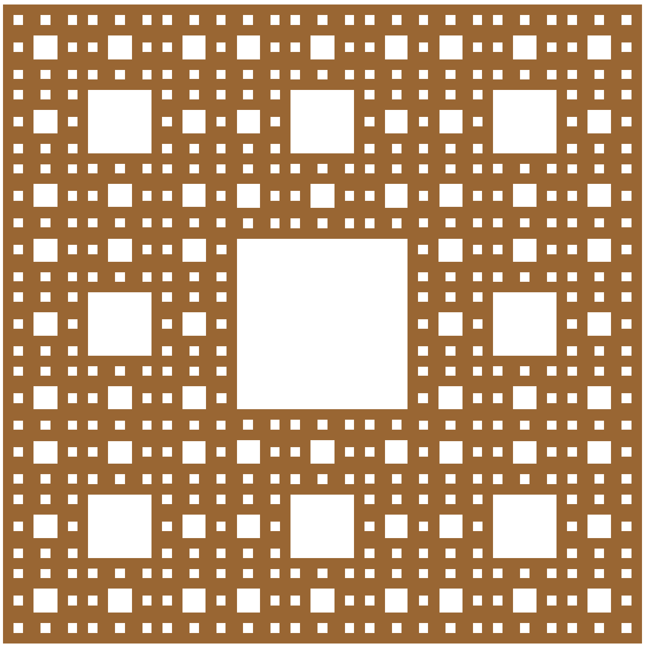


Figura 6.1: Il tappeto di Sierpinski.

Tornando alle serie in generale, dalla ovvia relazione

$$a_n = s_n - s_{n-1} \tag{6.7}$$

segue che se una serie converge la successione dei termini deve essere infinitesima, condizione dunque solo necessaria. Un'altra conseguenza della (6.7) è che nell'ipotesi di termini a segno (definitivamente) costante la serie non può essere indeterminata, se $a_n \geqslant 0$ [$a_n \leqslant 0$] converge o diverge perché (s_n) è crescente [decrescente]. Nel caso positivo allora hanno senso le notazioni

$$\sum_{n=0}^{\infty} a_n < +\infty \quad \text{se convergente} \qquad \text{e} \qquad \sum_{n=0}^{\infty} a_n = +\infty \quad \text{se divergente}\,,$$

analoghe in quello negativo.

La serie *armonica*

$$\sum_{n=1}^{\infty} \frac{1}{n} \tag{6.8}$$

non converge perché la (s_n) non è di Cauchy

$$s_{2n} - s_n = \frac{1}{n+1} + \frac{1}{n+2} + \ldots + \frac{1}{2n} > n \cdot \frac{1}{2n} = \frac{1}{2}\,.$$

Con quale rapidità diverge? Con quale successione nota si può confrontare? Con stime opportune si può non solo dimostrare che diverge, ma anche valutarne il tasso di crescita. Si considerano le ben note disuguaglianze (5.7) che adattiamo qua al caso nostro

$$\left(1+\frac{1}{k}\right)^k < e < \left(1+\frac{1}{k-1}\right)^k\,, \qquad k \geqslant 2\,,$$

poi si passa al logaritmo e si divide per k

$$\log\left(1+\frac{1}{k}\right) < \frac{1}{k} < \log\left(1+\frac{1}{k-1}\right)\,,$$

d'altra parte

$$\log\left(1+\frac{1}{k}\right) = \log(k+1) - \log k \qquad \text{e} \qquad \log\left(1+\frac{1}{k-1}\right) = \log k - \log(k-1)\,,$$

infine, sommando membro a membro da 2 a n e aggiungendo 1, si ottiene

$$1 - \log 2 + \log(n+1) < \sum_{k=1}^{n} \frac{1}{k} < 1 + \log n\,. \tag{6.9}$$

Si vede in questo modo che la (s_n) diverge con un andamento logaritmico.

Strettamente legata alla serie armonica per molte ragioni è la famosa successione

$$\gamma_n = \sum_{k=1}^{n} \frac{1}{k} - \log n \tag{6.10}$$

che per la (6.9) è positiva, parte dal valore $\gamma_1 = 1$ e decresce in quanto

$$\gamma_{n+1} - \gamma_n = \frac{1}{n+1} - \log(n+1) + \log n = \frac{1}{n+1}\left[1 - \log\left(1+\frac{1}{n}\right)^{n+1}\right] < 0\,.$$

La γ_n converge dunque e il suo limite γ è noto come la *costante di Eulero-Mascheroni*. Una sua stima per difetto è 0,5772156649015325 e ancora non si sa se è razionale o meno.

La monotonia della successione delle somme parziali, quando i termini hanno segno costante, può essere sfruttata per stabilire il carattere di una serie mediante il confronto con un'altra serie il cui comportamento è già noto. Da ora in poi supporremo $a_n \geqslant 0$, il caso negativo si tratta allo stesso modo. Vogliamo ad esempio studiare il comportamento della *serie armonica generalizzata* di esponente 2

$$\sum_{n=1}^{\infty} \frac{1}{n^2}\,. \tag{6.11}$$

La somiglianza con la serie del Mengoli

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)}\,,$$

che ha per somma 1, induce nella tentazione di prendere questa come serie di confronto, ma $1/n^2 > 1/n(n+1)$ e a noi servirebbe la stima opposta. Questo non è un problema, basta scrivere i termini della nostra serie in un altro modo

$$\sum_{k=1}^{n} \frac{1}{k^2} = \sum_{k=0}^{n-1} \frac{1}{(k+1)^2} = 1 + \sum_{k=1}^{n-1} \frac{1}{(k+1)^2} < 1 + \sum_{k=1}^{n-1} \frac{1}{k(k+1)} .$$

Le due successioni di somme sono crescenti e quella a destra converge, quindi quella a sinistra è limitata e deve convergere. Eulero per primo è riuscito a calcolare la somma della serie (6.11) che vale $\pi^2/6$ e in seguito sono state trovate altre dimostrazioni, una delle quali verrà trattata in modo un po' intuitivo alla fine di questo capitolo.

Per confronto, sul carattere della *serie armonica generalizzata* di esponente $\alpha \in \mathbf{R}$

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} , \tag{6.12}$$

risulta dunque

$$\begin{aligned} \text{per } \alpha < 1 : \quad & \frac{1}{n^\alpha} > \frac{1}{n} \Rightarrow \text{la (6.12) diverge} \\ \text{per } \alpha > 2 : \quad & \frac{1}{n^\alpha} < \frac{1}{n^2} \Rightarrow \text{la (6.12) converge} . \end{aligned}$$

Come si comporta la (6.12) per $1 < \alpha < 2$? Una stima dei raggruppamenti dei termini, da una potenza di 2 alla successiva, mostra che la nostra serie è convergente per questi valori di α

$$\begin{aligned} & 1 + \left(\frac{1}{2^\alpha} + \frac{1}{3^\alpha}\right) + \left(\frac{1}{4^\alpha} + \frac{1}{5^\alpha} + \frac{1}{6^\alpha} + \frac{1}{7^\alpha}\right) + \ldots \\ & + \left(\frac{1}{(2^n)^\alpha} + \frac{1}{(2^n+1)^\alpha} + \ldots + \frac{1}{(2^{n+1}-1)^\alpha}\right) + \ldots \\ & < 1 + \frac{1}{2^{\alpha-1}} + \frac{1}{(2^{\alpha-1})^2} + \ldots + \frac{1}{(2^{\alpha-1})^n} + \ldots = \frac{2^{\alpha-1}}{2^{\alpha-1}-1} . \end{aligned}$$

Questo metodo non funziona solo per $1 < \alpha < 2$, ma per tutti gli $\alpha > 1$.

Esercizio 6.4 - *In modo analogo, mediante opportuni raggruppamenti, dimostrare che la serie* (6.12) *è divergente per ogni* $\alpha < 1$.

Evidentemente, affinché una serie a termini positivi sia convergente, non basta che (a_n) sia infinitesima, ma deve tendere a 0 in modo abbastanza rapido. Se ad esempio $a_n \geqslant 0$ e decresce si vede facilmente che

$$\sum_{n=1}^{\infty} a_n < +\infty \Rightarrow \lim_{n\to\infty} n a_n \to 0 ,$$

infatti $s_{2n} - s_n \to 0$, d'altra parte

$$s_{2n} - s_n = a_{n+1} + a_{n+2} + \ldots + a_{2n} > n a_{2n} \geqslant 0 ,$$

quindi $2na_{2n} \to 0$ e la conclusione è vera per la sottosuccessione con gli indici pari. Per quella con gli indici dispari

$$(2n+1)a_{2n+1} < (2n+1)a_{2n} = 2na_{2n} + a_{2n} \to 0 .$$

Pertanto tutta la successione (na_n) è infinitesima e questo spiega ancora una volta, se ce ne fosse ancora bisogno, perché la (6.8) diverge.

Molto è stato scritto sulla serie armonica, è stato dimostrato che diverge, con crescita ancora più lenta, anche quella formata dagli inversi dei numeri primi

$$\sum_{p<n} \frac{1}{p} \sim \log \log n \,.$$

Tanto per dare un'idea, la somma degli inversi di tutti i numeri primi che non superano 10^6 è circa $2,887289$ e per calcolare la più piccola somma che supera il valore 4, un computer che impiega 10^{-9} secondi per aggiungere ad ogni passo un nuovo termine impiegherebbe 15 miliardi di anni. Se invece si scelgono solo i numeri "primi gemelli"

$$\left(\frac{1}{3}+\frac{1}{5}\right)+\left(\frac{1}{5}+\frac{1}{7}\right)+\left(\frac{1}{11}+\frac{1}{13}\right)+\left(\frac{1}{17}+\frac{1}{19}\right)+\ldots$$

si ottiene una serie convergente.

Riguardo le serie armoniche generalizzate, Eulero nella sua opera del 1748 *Introductio in analysin infinitorum* elenca i valori della *funzione "zeta" di Riemann*

$$\zeta(x) = \sum_{n=1}^{\infty} \frac{1}{n^x}\,, \quad x > 1\,,$$

il cui grafico è illustrato nella Figura 6.2, per tutti gli interi pari x da 2 a 26 (v. [11]).

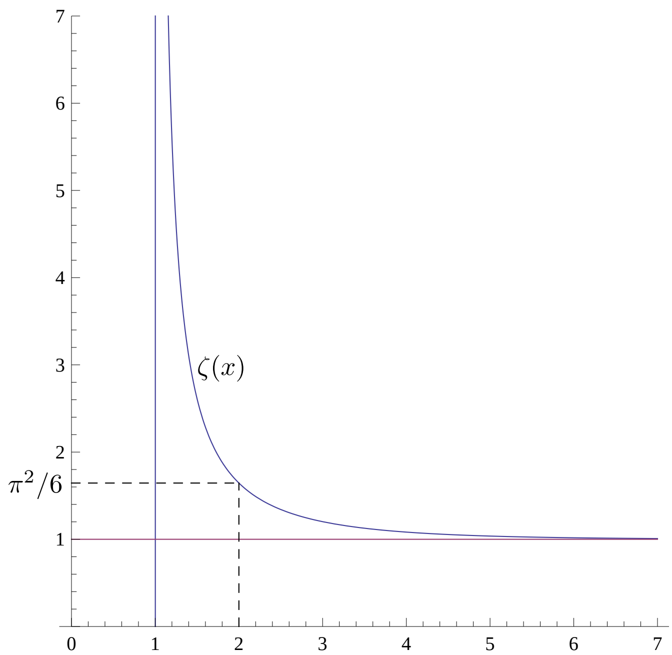


Figura 6.2: La funzione *zeta* di Riemann.

Per esempio

$$\zeta(4) = \frac{\pi^4}{90}\,, \quad \zeta(26) = \frac{2^{24}\cdot 76977927\pi^{26}}{27!} = \frac{1315862\pi^{26}}{11094481976030578125}\,.$$

Nel 1750 scopre la formula generale

$$\zeta(2h) = \sum_{n=1}^{\infty} \frac{1}{n^{2h}} = (-1)^{h-1}\frac{(2\pi)^{2h}}{2(2h)!}B_{2h}$$

dove i B_h sono i *numeri di Bernoulli* (2.14) del Cap. 2. Per x intero dispari non si sa nulla sul valore della somma, il problema è ancora aperto.

6.2 Serie a termini positivi

Questo paragrafo è dedicato alle serie a termini (definitivamente) positivi. Naturalmente ciò che diremo vale con ovvie modifiche anche con i termini negativi, ciò che conta è che abbiano segno costante. Per studiarne il carattere possiamo servirci di certe condizioni sufficienti che sono dette *criteri di convergenza*. Il primo di essi che ora illustriamo l'abbiamo già usato nel § 6.1.

Teorema 6.2 (Criterio del confronto) - *Se definitivamente $b_n \geqslant a_n \geqslant 0$ allora*

$$\sum_{n=0}^{\infty} b_n < +\infty \Rightarrow \sum_{n=0}^{\infty} a_n < +\infty \qquad e \qquad \sum_{n=0}^{\infty} a_n = +\infty \Rightarrow \sum_{n=0}^{\infty} b_n = +\infty\,.$$

Dimostrazione. Se $a_n \leqslant b_n$ per ogni $n \geqslant m$ si ha

$$\sum_{k=m}^{n} a_k \leqslant \sum_{k=m}^{n} b_k \quad \forall n \geqslant m\,.$$

Essendo crescenti rispetto a n, le due somme ammettono limite e la tesi segue dal Teorema 5.12. Rimangono esclusi i termini da 0 a $m-1$, ma sono in numero finito, influiscono solo sul valore della somma della serie, ma non sul carattere.

□

Esempi

6.1 *La serie*

$$\sum_{n=1}^{\infty} \frac{\operatorname{sen}^2 n}{n^2}$$

è convergente perché

$$\frac{\operatorname{sen}^2 n}{n^2} < \frac{1}{n^2}\,.$$

6.2 *La serie*

$$\sum_{n=1}^{\infty} \frac{\log n}{n^2}$$

è convergente. Si può infatti confrontare con una qualunque serie armonica convergente con esponente $2-\varepsilon > 1$. La relazione

$$\frac{\log n}{n^2} < \frac{1}{n^{2-\varepsilon}}$$

è sicuramente vera essendo equivalente a $\log n < n^{\varepsilon}$.

6.3 - *La serie*

$$\sum_{n=2}^{\infty} \frac{1}{(\log n)^{\log n}}$$

è convergente. Dalla

$$\log(\log n)^{\log n} = \log n \log\log n > 2\log n$$

segue $(\log n)^{\log n} > n^2$.

6.4 *Vediamo per quali valori dei parametri $a, \alpha \in \mathbf{R}$, $a > 0$, la serie*

$$\sum_{n=1}^{\infty} \frac{n^{\alpha}}{a^{\log n}}$$

è convergente. Tenendo presente che $a^{\log n} = n^{\log a}$, la serie data è armonica con esponente $\log a - \alpha$ e affinché sia convergente questo numero deve superare 1.

Esercizio 6.5 - *Stabilire il carattere della serie*

$$\sum_{n=2}^{\infty} \frac{1}{\log n!}\,.$$

Esercizio 6.6 - *Discutere il carattere della serie*

$$\sum_{n=1}^{\infty} \frac{\log n!}{n^{\alpha}}$$

al variare di $\alpha \in \mathbf{R}$.

Esercizio 6.7 - *Dimostrare che*

$$\sum_{n=1}^{\infty} \frac{1}{n\left(1+\frac{1}{2}+\frac{1}{3}+\ldots+\frac{1}{n}\right)} = +\infty\,.$$

Esercizio 6.8 - *Dimostrare che se* $a_n \geqslant 0$

$$\sum_{n=0}^{\infty} a_n < +\infty \Rightarrow \sum_{n=0}^{\infty} a_n^2 < +\infty\,.$$

Nel paragrafo precedente abbiamo osservato che se i termini positivi a_n di una serie convergente formano una successione decrescente allora $na_n \to 0$. Senza l'ipotesi sulla monotonia tale proposizione fallisce e l'Esempio 6.2 ci dà l'idea per costruire un controesempio

$$a_n = \begin{cases} \dfrac{\log n}{n} & \text{se } n \text{ è un quadrato perfetto} \\[2ex] \dfrac{1}{n^2} & \text{altrimenti}\,. \end{cases}$$

Ciò che si può dire in generale è che se la serie converge allora

$$\liminf_{n\to\infty} na_n = 0\,.$$

Se il minimo limite fosse un numero $\ell > 0$, preso $\varepsilon > 0$ in modo che $\ell - \varepsilon > 0$, varrebbe definitivamente la disuguaglianza $na_n > \ell - \varepsilon$, da cui $a_n > (\ell - \varepsilon)/n$ e per confronto la serie sarie sarebbe divergente.

Il seguente criterio è molto utile nella pratica.

Teorema 6.3 (Criterio del confronto asintotico) - *Se* $a_n \geqslant 0$ *e* $b_n > 0$ *e*

$$\lim_{n\to\infty} \frac{a_n}{b_n} = L \quad \textit{con } 0 < L < +\infty$$

allora le due serie $\sum a_n$ *e* $\sum b_n$ *hanno lo stesso comportamento.*

Dimostrazione. Per $\varepsilon < L$, ad esempio $\varepsilon = L/2$, si ha definitivamente

$$\frac{L}{2} < \frac{a_n}{b_n} < \frac{3L}{2}\,,$$

quindi per gli stessi n

$$\frac{L}{2} b_n < a_n < \frac{3L}{2} b_n\,.$$

Per il precedente criterio del confronto, Teorema 6.2, il carattere di una serie è identico a quello dell'altra.

□

Esempi

6.5 *La serie*
$$\sum_{n=1}^{\infty} \frac{1/n - \operatorname{sen}(1/n)}{e^{1/n} - \cos(1/n)}$$
è convergente in quanto
$$\frac{1}{n} - \operatorname{sen}\frac{1}{n} \sim \frac{1}{6n^3} \quad e \quad e^{1/n} - \cos\frac{1}{n} \sim \frac{1}{n}\,,$$
quindi il termine generale (a_n) *ha per andamento*
$$a_n \sim \frac{1}{6n^2}$$
e per il confronto asintotico con la serie armonica di esponente 2 la serie converge.

6.6 *Sappiamo che* $a_n = e - (1 + 1/n)^n \searrow 0$. *La serie con questi termini è convergente? Cerchiamo di valutarne l'andamento asintotico scrivendo la successione in un altro modo*
$$a_n = e\left[1 - e^{n\log(1+1/n)-1}\right].$$
D'altra parte
$$\log\left(1 + \frac{1}{n}\right) \sim \frac{1}{n} - \frac{1}{2n^2}\,,$$
quindi
$$a_n \sim e[1 - e^{-1/2n}] \sim \frac{1}{2n}$$
e la serie data diverge. Elevando i termini alla $\alpha \in \mathbf{R}$ *la serie* $\sum a_n^\alpha$ *converge per* $\alpha > 1$ *e diverge per* $\alpha \leqslant 1$.

Esercizio 6.9 - *Se* (a_n) *e* (b_n) *sono successioni positive e definitivamente*
$$\frac{a_{n+1}}{a_n} \leqslant \frac{b_{n+1}}{b_n}$$
allora
$$\sum_{n=0}^{\infty} b_n < +\infty \Rightarrow \sum_{n=0}^{\infty} a_n < +\infty \quad e \quad \sum_{n=0}^{\infty} a_n = +\infty \Rightarrow \sum_{n=0}^{\infty} b_n = +\infty\,.$$

Il metodo di raggruppamento dei termini che abbiamo applicato alle serie armoniche funziona anche in generale e il ragionamento è lo stesso, si associano i termini con gli indici che vanno da 2^n a 2^{n+1} e si perviene al seguente criterio che ci limitiamo ad enunciare senza dimostrarlo (sarebbe grosso modo una ripetizione di quanto detto).

Teorema 6.4 (Criterio di condensazione) - *Se* a_n *è decrescente e infinitesima le due serie*
$$\sum_{n=0}^{\infty} a_n \quad e \quad \sum_{n=0}^{\infty} 2^n a_{2^n}$$
hanno lo stesso comportamento.

Vogliamo ad esempio discutere il carattere della serie
$$\sum_{n=2}^{\infty} \frac{1}{n^\alpha \log^\beta n}$$

al variare di $\alpha, \beta \in \mathbf{R}$. Certamente, per qualunque β, converge per $\alpha > 1$ e diverge per $\alpha < 1$ per confronto diretto con la serie armonica generalizzata (nel secondo caso bisogna prendere quella con un esponente $\alpha + \varepsilon < 1$). Che cosa succede per $\alpha = 1$? La serie da considerare secondo il criterio di condensazione è

$$\sum_{n=0}^{\infty} 2^n \frac{1}{2^n (\log 2^n)^\beta} = \sum_{n=0}^{\infty} \frac{1}{n^\beta \log^\beta 2} \, .$$

Per $\alpha = 1$ la serie data converge per $\beta > 1$ e diverge per $\beta \leqslant 1$.

Teorema 6.5 (Criterio della radice) - *Se $a_n \geqslant 0$ per ogni $n \in \mathbf{N}$*

$$\lim_{n\to\infty} \sqrt[n]{a_n} = L < 1 \quad \Rightarrow \quad \sum_{n=0}^{\infty} a_n < +\infty$$

$$\exists (a_{k_n}) \subset (a_n) \ : \ a_{k_n} \geqslant 1 \quad \Rightarrow \quad \sum_{n=0}^{\infty} a_n = +\infty$$

Dimostrazione. Se $L < 1$ possiamo scegliere $\varepsilon > 0$ abbastanza piccolo in modo che $L + \varepsilon < 1$, così $a_n < (L+\varepsilon)^n$ definitivamente e

$$\sum_{n=0}^{\infty} (L+\varepsilon)^n < +\infty \, ,$$

quindi la serie data converge Se invece per infiniti indici $a_{k_n} \geqslant 1$ la (a_n) non è infinitesima e la serie data diverge.

□

Va osservato che questo criterio non contempla il caso in cui il limite vale 1. Applicato ad esempio alle serie armoniche è inservibile perché il limite della radice è sempre 1, qualunque sia l'esponente.

Facendo attenzione alla dimostrazione della prima parte, dovrebbe risultare evidente che non è servita per intero l'ipotesi $\sqrt[n]{a_n} \to L < 1$ perché è bastata la disuguaglianza $\sqrt[n]{a_n} < L+\varepsilon$. Possiamo dunque sostituirla con la seguente più generale

$$\exists c < 1 \ : \ \sqrt[n]{a_n} \leqslant c$$

(che non significa affatto $\sqrt[n]{a_n} < 1$!) che non richiede l'esistenza del limite, ma è equivalente a

$$\limsup_{n\to\infty} \sqrt[n]{a_n} < 1 \, .$$

Per il caso divergente l'esistenza di una sottosuccessione della (a_n) che si mantiene al di sopra di 1 equivale alla condizione

$$\limsup_{n\to\infty} \sqrt[n]{a_n} \geqslant 1 \, .$$

Esercizio 6.10 - *Verificare che per ogni $\alpha > 0$*

$$\sum_{n=1}^{\infty} \left(1 - \frac{1}{n^\alpha}\right)^{n^{\alpha+1}} < +\infty \, .$$

Teorema 6.6 (Criterio del rapporto) - *Se $a_n > 0$ per ogni $n \in \mathbf{N}$*

$$\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = L < 1 \quad \Rightarrow \quad \sum_{n=0}^{\infty} a_n < +\infty$$

$$\frac{a_{n+1}}{a_n} \geqslant 1 \textit{ definitivamente} \quad \Rightarrow \quad \sum_{n=0}^{\infty} a_n = +\infty$$

Dimostrazione. Se $L < 1$ è il limite di questi rapporti, per $\varepsilon > 0$ abbastanza piccolo in modo che $L + \varepsilon < 1$ si ha

$$\frac{a_{n+1}}{a_n} < L + \varepsilon \quad \forall n > k\,,$$

quindi

$$a_{n+1} < (L+\varepsilon)a_n < (L+\varepsilon)^2 a_{n-1} < \ldots < (L+\varepsilon)^{n-k} a_{k+1} \quad \forall n > k$$

dove la serie

$$\sum_{n=k}^{\infty} (L+\varepsilon)^{n-k}$$

è convergente. Per la seconda parte della tesi, (a_n) è crescente per ipotesi, quindi non può essere infinitesima e la serie diverge.

□

Come quello della radice, anche il criterio del rapporto va modificato se il limite non esiste, ma in modo un poco diverso

$$\frac{a_{n+1}}{a_n} \leqslant c < 1 \quad \text{oppure} \quad \limsup_{n\to\infty} \frac{a_{n+1}}{a_n} < 1 \quad \Rightarrow \quad \text{converge}$$

$$\frac{a_{n+1}}{a_n} \geqslant 1 \text{ definitivamente} \quad \text{oppure} \quad \liminf_{n\to\infty} \frac{a_{n+1}}{a_n} \geqslant 1 \quad \Rightarrow \quad \text{diverge}\,.$$

I due criteri non sono del tutto equivalenti. Consideriamo ad esempio la successione dei termini

$$a_n = \begin{cases} 2^{-n} & \text{se } n \text{ è pari} \\ 3^{-n} & \text{se } n \text{ è dispari.} \end{cases} \tag{6.13}$$

La sottosuccessione della (a_{n+1}/a_n) con gli n pari è infinitesima mentre quella con gli n dispari è addirittura divergente, eppure la serie converge in quanto somma di due serie geometriche convergenti, del resto anche il criterio della radice ce lo conferma. In effetti quest'ultimo è, dei due, il più generale: se quello del rapporto ci dà una risposta sul carattere di una serie anche l'altro ce la dà, ma il viceversa non vale. Questo discende dalla Proposizione 5.27 del Cap. 5.

Esercizio 6.11 - *Studiare il carattere della serie*

$$\sum_{n=1}^{\infty} \frac{n!}{n^n}\,.$$

Esercizio 6.12 - *Dimostrare che se converge la serie $\sum a_n^2$ con $a_n > 0$ e la successione dei rapporti (a_{n+1}/a_n) è decrescente allora converge anche la serie $\sum a_n$.*

Un criterio più generale, che può funzionare anche quando quello della radice fallisce, è il seguente.

Teorema 6.7 (Criterio di Raabe) - *Se $a_n > 0$ per ogni $n \in \mathbf{N}$ ed esiste un numero reale $c > 1$ tale che definitivamente*

$$n\left(\frac{a_n}{a_{n+1}} - 1\right) \geqslant c \tag{6.14}$$

allora $\sum a_n < +\infty$. Se definitivamente

$$n\left(\frac{a_n}{a_{n+1}} - 1\right) \leqslant 1 \tag{6.15}$$

allora $\sum a_n = +\infty$.

Dimostrazione. Scelto $\alpha \in \mathbf{R}$ tale che $1 < \alpha < c$, sappiamo che

$$\lim_{n\to\infty} n\left[\left(1+\frac{1}{n}\right)^{\alpha} - 1\right] = \alpha < c\,,$$

quindi definitivamente

$$\left(1+\frac{1}{n}\right)^{\alpha} = \frac{(n+1)^{\alpha}}{n^{\alpha}} < 1 + \frac{c}{n} < \frac{a_n}{a_{n+1}}$$

per l'ipotesi (6.14). Ne segue

$$\frac{a_{n+1}}{a_n} \leqslant \frac{n^{\alpha}}{(n+1)^{\alpha}}$$

e per l'Esercizio 6.9 la serie data converge come la serie armonica di esponente α.

Con l'altra ipotesi si peviene subito alla condizione

$$\frac{a_{n+1}}{a_n} \geqslant \frac{n}{n+1}$$

e la serie data diverge come la serie armonica.

□

Esercizio 6.13 - *Riprendere tutte le serie viste finora e provare ad applicare il criterio di Raabe.*

6.3 Serie a termini di segno variabile

Negli esempi introduttivi del § 6.1 abbiamo pasticciato un po' con queste somme infinite trattandole come delle somme finite. Abbiamo raggruppato stringhe di termini come se valesse la proprietà associativa, ma basta associare a due a due i termini della serie dell'Esempio 6.5 per trasformarla da indeterminata a convergente con somma 0. Anche riordinare gli stessi termini in un altro modo può cambiarne la somma, ma anche il carattere, se la serie non è a termini positivi. Consideriamo ad esempio la seguente

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} \tag{6.16}$$

che si ottiene dalla serie armonica alternando i segni dei termini. Ricordiamo che la successione delle somme parziali

$$s_n = \sum_{k=1}^{n} \frac{(-1)^{k-1}}{k}$$

è legata all'identità di Catalan (2.11) e che l'Esempio 5.7 ci garantisce che s_{2n} converge. Ma tutta la s_n converge perché $s_{2n+1} - s_{2n} = 1/(2n+1) \to 0$. Per dimostrare che converge a $\log 2$ possiamo usare la successione di Eulero-Mascheroni (6.10) e passare poi al limite per $n \to \infty$ in questa uguaglianza che discende dalla definizione stessa della γ_n

$$\sum_{k=1}^{n} \frac{1}{n+k} = \gamma_{2n} + \log 2n - \gamma_n - \log n = \gamma_{2n} - \gamma_n + \log 2\,.$$

Esercizio 6.14 - *Dimostrare, più in generale, che per ogni $k \in \mathbf{N} - \{0\}$*

$$\lim_{n\to\infty} \sum_{h=1}^{n} \frac{1}{kn+h} = \log\frac{k+1}{k}\,.$$

Riprendiamo adesso la serie (6.16) e sommiamola termine a termine con la stessa divisa per 2

$$\begin{array}{rcccccccccccccccccc} S & = & 1 & - & \frac{1}{2} & + & \frac{1}{3} & - & \frac{1}{4} & + & \frac{1}{5} & - & \frac{1}{6} & + & \frac{1}{7} & - & \frac{1}{8} & + & \dots \\ \frac{1}{2}S & = & & & \frac{1}{2} & & & - & \frac{1}{4} & & & + & \frac{1}{6} & & & - & \frac{1}{8} & + & \dots \end{array}$$

si ottiene

$$\frac{3}{2}S = 1 + \frac{1}{3} - \frac{1}{2} + \frac{1}{5} + \frac{1}{7} - \frac{1}{4} + \frac{1}{9} + \frac{1}{11} - \frac{1}{6} + \dots$$

in cui si riconosce la serie di partenza con i termini in un altro ordine. Evidentemente questa manipolazione altera, in generale, il valore della somma.

È chiaro comunque, dalla teoria delle successioni, che se due serie $\sum a_n$ e $\sum b_n$ sono convergenti con somme S e S' rispettivamente allora anche la serie $\sum(a_n + b_n)$ è convergente con somma $S + S'$, ma la serie somma può convergere senza che lo siano le serie che la compongono. Per esempio quelle coi termini $a_n = 1/n$ e $b_n = -1/(n+1)$ sono divergenti, mentre quella dei termini somma $a_n + b_n$ converge. Altro problema riguarda invece la scelta e l'ordine dei termini di una stessa serie. In generale $\sum a_n$ può convergere anche nel caso siano divergenti certe serie ottenute selezionando da essa solo una parte dei termini, ad esempio quelli positivi o quelli negativi che formano le seguenti

$$\sum_{n=0}^{\infty} a_n^+ \qquad \text{e} \qquad \sum_{n=0}^{\infty} a_n^-$$

dove $a_n^+ = \max\{a_n, 0\}$, $a_n^- = -\min\{a_n, 0\}$ (per cui $a_n = a_n^+ - a_n^-$). La (6.16) è convergente anche se $\sum a_{2n}$ e $\sum a_{2n+1}$ divergono. L'alternanza dei segni ha l'effetto di compensare reciprocamente le parti di segno opposto, ma se una parte prevalesse sull'altra, associando gruppi sempre più numerosi di termini di un certo segno, la serie divergerebbe. Serie di questo tipo si chiamano *condizionatamente convergenti*. Esse soddisfano la proprietà alquanto sorprendente che ci limitiamo ad enunciare senza dimostrazione.

Teorema 6.8 (Riemann) - *Se $\sum a_n$ è condizionatamente convergente allora per ogni $S \in \overline{\mathbf{R}}$ esiste un riordinamento della serie che ha per somma S. Esiste anche un riordinamento che la rende indeterminata.*

Una condizione sufficiente per la convergenza, particolarmente adatta a casi come questi, è espressa dal seguente criterio per le serie a segni alterni.

Teorema 6.9 (Criterio di Leibniz) - *Se (a_n) è decrescente e infinitesima (quindi positiva) la serie*

$$\sum_{n=0}^{\infty} (-1)^n a_n$$

è convergente.

Dimostrazione. Sia (s_n) la successione delle somme parziali della serie data e consideriamo le due sottosuccessioni s_{2n} e s_{2n+1}. Per ogni $n \in \mathbf{N}$ si ha

$$s_{2n} = s_{2n-2} - a_{2n-1} + a_{2n} \leqslant s_{2n-2} \qquad \text{e} \qquad s_{2n+1} = s_{2n-1} + a_{2n} - a_{2n+1} \geqslant s_{2n-1}\,,$$

da cui si vede che (s_{2n}) decresce mentre (s_{2n+1}) cresce. Inoltre la prima è limitata inferiormente e la seconda superiormente dato che

$$s_{2n} = s_{2n-1} + a_{2n} \geqslant s_{2n-1} \geqslant s_1 \qquad \text{e} \qquad s_{2n+1} = s_{2n} - a_{2n+1} \leqslant s_{2n} \leqslant s_0\,,$$

quindi convergono entrambe. Infine hanno lo stesso limite S perché

$$s_{2n} - s_{2n-1} = a_{2n} \to 0\,,$$

quindi $(s_n) \to S$ oscillando in quanto, per la monotonia delle due sottosuccessioni, $s_{2n+1} \leqslant S \leqslant s_{2n}$ per ogni $n \in \mathbf{N}$.

□

Sono dunque convergenti, oltre alla (6.16) che abbiamo già analizzato, anche le serie

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}}\,, \quad \sum_{n=1}^{\infty} \frac{(-1)^n}{\log n}\,, \quad \sum_{n=1}^{\infty} \frac{(-1)^n}{\log\log n}\,,$$

non importa con quale rapidità tendano a 0 gli a_n, basta che lo facciano decrescendo. La successione non monotona

$$a_n = \begin{cases} 1/n & \text{se } n \text{ è pari} \\ 1/n^2 & \text{se } n \text{ è dispari} \end{cases}$$

è infinitesima, ma la serie $\sum(-1)^n a_n$ è positivamente divergente. Nella serie

$$\sum_{n=1}^{\infty} \frac{(-1)^n \operatorname{arctg} n}{n} \tag{6.17}$$

non è così evidente a prima vista, sebbene facile da verificare, che $a_n = (\operatorname{arctg} n)/n$ sia decrescente, ma ricordando la ben nota relazione tra l'arctg n e l'arctg $1/n$ si decompone in due serie convergenti per il criterio di Leibniz, di cui la seconda è anche assolutamente convergente.

Se ciò che serve a far convergere una serie è una sorta di compensazione reciproca tra i termini di segno opposto, non dovrebbe essere strettamente necessario che sia rigorosamente a segni alterni. Possiamo pensare anche ad altre regole e a leggi diverse di distribuzione dei segni, l'importante è che gruppi di termini consecutivi con lo stesso segno siano grosso modo altrettanto numerosi quanto quelli di segno opposto. Possiamo pensare anche a distribuzioni casuali di segni con probabilità del 50%.

Sappiamo bene che la "serie dei segni" $\sum(-1)^n$ non converge, ma la relativa successione delle somme parziali, che assume solo i valori 0 e 1, è comunque limitata. Questo è in realtà il vero motivo per cui funziona il criterio di Leibniz e lo si capisce col seguente criterio nel quale rientra come caso particolare.

Teorema 6.10 (Criterio di Dirichlet) - *Siano (a_n) decrescente e infinitesima e (c_n) tale che la relativa successione delle somme*

$$C_n = \sum_{k=0}^{n} c_k$$

sia limitata. Allora la serie

$$\sum_{n=0}^{\infty} c_n a_n$$

è convergente.

Dimostrazione. Sia $C > 0$ un maggiorante della successione $(|C_n|)$. Dimostriamo che la successione delle somme parziali

$$s_n = \sum_{k=0}^{n} c_k a_k$$

converge in quanto di Cauchy. Per $m \leqslant n$ si ha

$$\begin{aligned} s_n - s_m &= \sum_{k=m+1}^{n} c_k a_k = \sum_{k=m+1}^{n} (C_k - C_{k-1}) a_k = \sum_{k=m+1}^{n} C_k a_k - \sum_{k=m+1}^{n} C_{k-1} a_k \\ &= C_n a_n - C_m a_{m+1} + \sum_{k=m+1}^{n-1} C_k (a_k - a_{k+1}) \end{aligned}$$

e passando ai moduli

$$|s_n - s_m| \leqslant C(a_n + a_{m+1} + a_{m+1} - a_n) = 2C a_{m+1} \,.$$

Fissato allora $\varepsilon > 0$, poiché per un certo $\nu \in \mathbf{N}$ $a_{m+1} < \varepsilon$ per ogni $m > \nu$, per $n \geqslant m > \nu$ si ha $|s_n - s_m| < 2C\varepsilon$.

□

Il criterio di Dirichlet si applica con successo alla serie

$$\sum_{n=1}^{\infty} \frac{\operatorname{sen} nt}{n} \,, \qquad t \in \mathbf{R}$$

o, più in generale, alle *serie trigonometriche*

$$\frac{a_0}{2} + \sum_{n=1}^{\infty} (a_n \cos nt + b_n \operatorname{sen} nt) \tag{6.18}$$

che tratteremo nel corso di Analisi 2. Sembra tagliato su misura proprio per questo tipo di serie, basta che a_n e b_n siano infinitesime e decrescenti ed abbiamo subito una condizione sufficiente per la convergenza della (6.18) (se non è una serie di soli seni dobbiamo porre $t \neq 2h\pi$). Tenendo conto delle formule di Eulero, dobbiamo dimostrare che la successione $(\sum_{k=1}^{n} e^{ikt})$ è limitata per ogni $t \neq 2h\pi$. Infatti

$$\left| \sum_{k=1}^{n} e^{ikt} \right| = \frac{|1 - e^{i(n+1)t}|}{|1 - e^{it}|} \leqslant \frac{2}{|1 - e^{it}|} = \frac{1}{|\operatorname{sen} t/2|} \qquad \forall n \in \mathbf{N} \,.$$

Anche alla serie (6.17) è applicabile il Criterio di Dirichlet perché $(1/n)$ è decrescente e infinitesima mentre la successione

$$C_n = \sum_{k=1}^{n} (-1)^k \operatorname{arctg} k$$

non solo è limitata ma convergente in quanto

$$\operatorname{arctg}(2k) - \operatorname{arctg}(2k-1) = \operatorname{arctg} \frac{1}{1 + 2k(2k-1)} \,.$$

Esercizio 6.15 - *Dimostrare che la serie*

$$\sum_{n=1}^{\infty} \operatorname{sen} t \operatorname{sen} 2t \ldots \operatorname{sen} nt$$

è convergente per ogni $t \in \mathbf{R}$.

6.4 Serie assolutamente convergenti

Abbiamo visto che la serie (6.16) converge, ma quella degli stessi termini presi in valore assoluto, che è poi la serie armonica, diverge. Lo stesso accade per tante altre serie a termini di segno variabile. In questo paragrafo dimostriamo infatti che la convergenza con i termini in valore assoluto è condizione solo sufficiente per la convergenza della serie di partenza. Ad esempio è per questo motivo che converge la serie
$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n^2}\,.$$
Di questa possiamo anche calcolare la somma S supponendo di conoscere il valore $\pi^2/6$ della somma della serie armonica di esponente 2. Dapprima si calcolano le due somme con n solo pari o solo dispari
$$\sum_{n=1}^{\infty} \frac{1}{(2n-1)^2} = \frac{\pi^2}{6} - \sum_{n=1}^{\infty} \frac{1}{(2n)^2} = \frac{\pi^2}{6} - \frac{\pi^2}{24} = \frac{\pi^2}{8}\,,$$
poi si trova il valore di S per differenza
$$S = \frac{\pi^2}{8} - \frac{\pi^2}{24} = \frac{\pi^2}{12}\,.$$
In questo ragionamento abbiamo usato la proprietà di questa serie di essere *incondizionatamente convergente*, convergono cioè separatamente sia la serie formata dai termini positivi che quella dei termini negativi e la somma dell'una più la somma dell'altra coincide con la somma della serie iniziale. Riordinamenti e raggruppamenti dei termini di una serie incondizionatamente convergente sono operazioni che non alterano il valore della somma.

Definizione 6.11 - *Una serie* $\sum a_n$ *è detta* ***assolutamente convergente*** *se*
$$\sum_{n=0}^{\infty} |a_n| < +\infty\,.$$
Tenendo presente che $|a_n| = a_n^+ + a_n^-$, condizione necessaria e sufficiente affinché una serie sia assolutamente convergente è che sia incondizionatamente convergente.

Teorema 6.12 - *Ogni serie assolutamente convergente è convergente e*
$$\left|\sum_{n=0}^{\infty} a_n\right| \leqslant \sum_{n=0}^{\infty} |a_n| \tag{6.19}$$

Dimostrazione. Consideriamo le due successioni di somme
$$s_n = \sum_{k=0}^{n} a_k \qquad \text{e} \qquad \sigma_n = \sum_{k=0}^{n} |a_k|\,.$$
Per ipotesi (σ_n) converge e quindi è di Cauchy. Dimostriamo allora che anche la (s_n) è di Cauchy e di conseguenza convergente. Infatti per ogni $m, n \in \mathbf{N}$, $m < n$, si ha
$$|s_n - s_m| = \left|\sum_{k=m+1}^{n} a_k\right| \leqslant \sum_{k=m+1}^{n} |a_k| = \sigma_n - \sigma_m$$
e se definitivamente $\sigma_n - \sigma_m < \varepsilon$, per gli stessi indici m, n anche $|s_n - s_m| < \varepsilon$. Infine
$$|s_n| \leqslant \sigma_n \leqslant \lim_{n\to\infty} \sigma_n = \sum_{n=0}^{\infty} |a_n|$$

e passando al limite a sinistra si ottiene la (6.19).

□

La Definizione 6.11 e il Teorema 6.12 si applicano anche a termini $a_n \in \mathbf{C}$ senza dover modificare nulla. I criteri di convergenza per le serie a termini positivi, che abbiamo visto nel § 6.2, possono essere usati ora per le serie di segno variabile come criteri di assoluta convergenza.

Nel corso della dimostrazione abbiamo usato la completezza di $\mathbf{R}$, ma in che misura questa proprietà è importante? Vediamo che cosa succede in $\mathbf{Q}$ che non è completo. La serie

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{n(n+1)}$$

è assolutamente convergente. Sappiamo infatti che la serie dei valori assoluti ha per somma $1 \in \mathbf{Q}$, ma la somma della serie data è $1 - 2\log 2 \notin \mathbf{Q}$.

In effetti il teorema può essere dimostrato esattamente nello stesso modo in un contesto molto più ampio, quello degli spazi vettoriali normati e completi. Qua non è necessario dirlo perché abbiamo scelto $\mathbf{R}$ fin dall'inizio, ma addirittura si può dimostrare che se ogni serie assolutamente convergente è convergente allora lo spazio è completo.

Esempi

6.7 *La serie*

$$\sum_{n=1}^{\infty} \frac{\operatorname{sen} nx}{n^2}$$

è assolutamente convergente, quindi convergente, in quanto

$$\left| \frac{\operatorname{sen} nx}{n^2} \right| \leqslant \frac{1}{n^2}\,.$$

Esercizio 6.16 - *Dimostrare che la serie*

$$\sum_{n=1}^{\infty} \frac{\operatorname{sen}(n + 1/n) - \operatorname{sen} n}{n}$$

è assolutamente convergente.

Esercizio 6.17 - *Se $f : [0, +\infty[\to \mathbf{R}$ è una funzione lipschitziana o hölderiana stabilire per quali $\alpha \in \mathbf{R}$ la serie*

$$\sum_{n=1}^{\infty} \frac{f(n + 1/n) - f(n)}{n^{\alpha}}$$

è assolutamente convergente.

Esercizio 6.18 - *Dimostrare che se $\sum a_n^2 < +\infty$ e $\sum b_n^2 < +\infty$ allora $\sum a_n b_n$ è assolutamente convergente. Più in generale, $\sum a_n b_n$ è ancora assolutamente convergente se $\sum |a_n|^p < +\infty$ e $\sum |b_n|^q < +\infty$ dove $p, q > 1$ sono esponenti coniugati, cioè $1/p + 1/q = 1$.*

Sull'operazione di prodotto tra serie, certamente sarebbe sbagliato definirla così

$$\sum a_n \sum b_n = \sum a_n b_n\,,$$

equivarrebbe a confondere $(a+b)(c+d)$ con $ac+bd$. Il modo più ovvio e sensato è quello di ricondursi al prodotto delle relative somme parziali: posto

$$s_n = \sum_{k=0}^{n} a_k \,, \qquad s'_n = \sum_{k=0}^{n} b_k \,,$$

se (s_n) e (s'_n) convergono rispettivamente a S e S' sappiamo bene che

$$\sum_{k=0}^{n} a_k \sum_{k=0}^{n} b_k \to SS' \,.$$

Ma un conto è il prodotto di due serie, altra cosa è la *serie prodotto*. Nelle applicazioni è più utile considerare la serie *prodotto di Cauchy* $\sum c_n$ delle due serie $\sum a_n$ e $\sum b_n$, dove

$$c_n = a_0 b_n + a_1 b_{n-1} + a_2 b_{n-2} + \ldots + a_n b_0 = \sum_{k=0}^{n} a_k b_{n-k} \,, \tag{6.20}$$

però non ne è garantita la convergenza con la sola ipotesi che $\sum a_n$ e $\sum b_n$ siano convergenti. Ad esempio la serie

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{\sqrt{n+1}}$$

converge per il criterio di Leibniz. Moltiplicata per se stessa alla Cauchy, dà luogo alla serie dei termini

$$c_n = \sum_{k=0}^{n} \frac{(-1)^k}{\sqrt{k+1}} \frac{(-1)^{n-k}}{\sqrt{n+1-k}} = (-1)^n \sum_{k=0}^{n} \frac{1}{\sqrt{(k+1)(n+1-k)}} \,,$$

ma $(k+1)(n+1-k) \leqslant (n+1)^2$, quindi

$$|c_n| \geqslant \sum_{k=0}^{n} \frac{1}{n+1} = 1$$

e non può essere infinitesima.

Teorema 6.13 - *Se $\sum a_n$ e $\sum b_n$ sono assolutamente convergenti con somme rispettive S e S' allora anche il loro prodotto di Cauchy è assolutamente convergente ed ha per somma SS'.*

Dimostrazione. Dovendo considerare i vari termini tutti in valore assoluto, possiamo supporre fin dall'inizio $a_n, b_n \geqslant 0$. Indicate con (s_n) e (s'_n) le rispettive somme parziali, basta costruire la tabella dei prodotti $a_i b_j$, con gli a_i in riga e i b_j in colonna, per rendersi conto che

$$s_n s'_n \leqslant \sum_{k=0}^{2n} c_k \leqslant s_{2n} s'_{2n}$$

	a_0	a_1	a_2	a_3	$\ldots$
b_0	a_0b_0	a_1b_0	a_2b_0	a_3b_0	
b_1	a_0b_1	a_1b_1	a_2b_1	$\ldots$	
b_2	a_0b_2	a_1b_2	$\ldots$	$\ldots$	
b_3	a_0b_3	$\ldots$	$\ldots$	$\ldots$	
$\vdots$					

dove ogni c_k è la somma degli elementi della k-esima diagonale $NE - SO$. Passando al limite per $n \to \infty$ si ottiene la tesi per confronto.

□

In realtà si può dimostrare che il Teorema 6.13 vale ancora se una sola delle due serie è assolutamente convergente.

6.5 Serie di potenze

Una categoria importante di serie assolutamente convergenti, a cui si applicano con successo dei metodi generali, è quella delle *serie di potenze*

$$\sum_{n=0}^{\infty} a_n(x-x_0)^n\,, \quad a_n, x_0, x \in \mathbf{R}\,, \qquad \sum_{n=0}^{\infty} c_n(z-z_0)^n\,, \quad c_n, z_0, z \in \mathbf{C}\,.$$

La successione dei *coefficienti*, (a_n) o (c_n), è data, così come il *centro* x_0 o z_0, quindi si tratta di stabilire per quali valori del parametro $x \in \mathbf{R}$ o $z \in \mathbf{C}$ la serie converge e in che senso. Ci mettiamo direttamente nel campo complesso, il caso reale si può ottenere di conseguenza come restrizione della variabile z all'asse reale. Osserviamo che per $z = z_0$ la serie si riduce al solo termine c_0 quindi non si pone il problema della convergenza, anzi si può dire che il *dominio di convergenza*, ossia l'insieme dei numeri $z \in \mathbf{C}$ per i quali la serie converge, non è mai vuoto perché contiene almeno z_0. Neanche per i polinomi, in quanto già per definizione scritti su tutto $\mathbf{C}$ come somme dello stesso tipo, ma finite, si pone il problema della convergenza. In ogni punto del dominio di convergenza rimane ben definita la funzione

$$f(z) = \sum_{n=0}^{\infty} c_n(z-z_0)^n\,, \qquad z \in \mathbf{C}\,. \tag{6.21}$$

Non è restrittivo assumere $z_0 = 0$, dato che a questo caso ci si può sempre ricondurre mediante la traslazione $z' = z - z_0$. Cominciamo dunque con lo stabilire per quali $z \in \mathbf{C}$ la serie

$$\sum_{n=0}^{\infty} c_n z^n \tag{6.22}$$

è assolutamente convergente. Per questo possiamo applicare il criterio della radice nella sua versione più generale, ma funziona anche il criterio del rapporto e porta allo stesso risultato, come si può facilmente verificare.

Teorema 6.14 (di Cauchy-Hadamard) - *Posto*

$$L = \limsup_{n\to\infty} \sqrt[n]{|c_n|}\,, \tag{6.23}$$

se $L = +\infty$ la (6.22) converge solo per $z = 0$ riducendosi al solo termine c_0, se $L = 0$ converge assolutamente per ogni $z \in \mathbf{C}$ e se infine $0 < L < +\infty$ converge assolutamente per ogni $z \in \mathbf{C}$ tale che $|z| < R = 1/L$. Nel caso di R finito la condizione $|z| < R$ definisce il ***cerchio di convergenza*** *il cui raggio è detto* ***raggio di convergenza****. Al di fuori di esso, per $|z| > R$, la (6.22) non converge assolutamente in nessun punto.*

Dimostrazione. Essendo

$$\limsup_{n\to\infty} \sqrt[n]{|c_n z^n|} = \limsup_{n\to\infty} |z| \sqrt[n]{|c_n|} = L|z|\,,$$

per il criterio della radice la serie (6.22) converge assolutamente per ogni $z \in \mathbf{C}$ tale che $L|z| < 1$ e da questa condizione si ottiene

$$\begin{aligned} &L = 0 \Rightarrow z \text{ qualunque in } \mathbf{C}\,,\\ &0 < L < +\infty \Rightarrow |z| < 1/L\,,\\ &L = +\infty \Rightarrow z = 0\,. \end{aligned}$$

Se invece $L|z| > 1$, ovviamente possibile solo se $L > 0$, la serie dei moduli diverge.

□

Esercizio 6.19 - *Enunciare e dimostrare il risultato col criterio del rapporto.*

La scrittura $R = 1/L$ può essere usata anche nei due casi estremi assumendo per convenzione $1/0 = \infty$ e $1/\infty = 0$. Bisogna però fare attenzione al fatto che solo nel caso che esista il limite nella (6.23) possiamo calcolare R così

$$R = \lim_{n\to\infty} \frac{1}{\sqrt[n]{|c_n|}} = \lim_{n\to\infty} \frac{|c_n|}{|c_{n+1}|} \, ,$$

altrimenti dobbiamo passare al minimo limite

$$R = \liminf_{n\to\infty} \frac{1}{\sqrt[n]{|c_n|}} = \liminf_{n\to\infty} \frac{|c_n|}{|c_{n+1}|} \, .$$

Il dominio di assoluta convergenza può dunque essere tutto il piano complesso o il disco aperto di centro 0 e raggio $R = 1/L$, oppure, caso banale, il solo punto 0. Una serie di potenze che converge su tutto $\mathbf{C}$ (per forza in senso assoluto) definisce tramite la (6.21) una funzione $f : \mathbf{C} \to \mathbf{C}$ che viene della *intera*. I polinomi sono esempi di funzioni intere.

Nel caso di R finito non può esistere nessun altro punto a distanza da 0 maggiore di R in cui la serie converga assolutamente. Rimane quindi da capire se nei punti esterni al disco di convergenza la serie possa convergere o meno. Il seguente teorema risponde negativamente, nei punti esterni non solo non vale la convergenza assoluta, ma neanche la semplice convergenza.

Teorema 6.15 - *Se una serie converge in un punto $w \in \mathbf{C} - \{0\}$ allora converge assolutamente in ogni punto $z \in \mathbf{C}$ tale che $|z| < |w|$.*

Dimostrazione. La successione $(c_n w^n)$ è limitata in quanto infinitesima, quindi per un certo $M > 0$ si ha $|c_n w^n| \leqslant M$ per ogni $n \in \mathbf{N}$. Per ogni $z \in \mathbf{C}$ tale che $|z| < |w|$ poniamo $q = |z|/|w| < 1$. Si ha

$$\sum_{n=0}^{\infty} |c_n z^n| = \sum_{n=0}^{\infty} |c_n w^n| q^n \leqslant M \sum_{n=0}^{\infty} q^n = \frac{M}{1-q}$$

e la tesi è provata.

□

Se dunque $R < +\infty$ non si ha nessun tipo di convergenza per $|z| > R$. Nei punti interni invece, con $|z| < R$, la convergenza è sempre assoluta. Che cosa succede sulla circonferenza $|z| = R$? Non si può sapere in generale e bisogna vedere caso per caso, può convergere, talvolta assolutamente, in alcuni di questi punti e non in altri, in nessuno di essi o in tutti, bisogna fare un'indagine specifica caso per caso. Consideriamo ad esempio la serie

$$\sum_{n=0}^{\infty} \frac{z^n}{n} \, . \tag{6.24}$$

Il raggio di convergenza è $R = 1$, quindi la serie converge assolutamente per ogni $z \in \mathbf{C}$ tale che $|z| < 1$. Se $|z| = 1$ allora $z = e^{it}$ con $t \in \mathbf{R}$ e la serie diventa

$$\sum_{n=0}^{\infty} \frac{e^{int}}{n}$$

della quale sappiamo già che diverge per $t = 0$ riducendosi alla serie armonica, ma per ogni $t \neq 0$ converge per il criterio di Dirichlet; al caso particolare $t = \pi$ è applicabile anche il criterio di Leibniz.

Se una serie di potenze con raggio di convergenza $R > 0$ finito converge anche in un punto $\hat{z}$ tale che $|\hat{z}| = R$, che sta quindi sul bordo del disco, in che relazione sta la somma della serie in quel punto con i valori della somma nei punti interni? Vediamo di chiarire la domanda con un esempio. Senza badare troppo al rigore con cui andrebbero maneggiate le somme infinite, mostriamo dapprima che vale lo sviluppo in serie sul disco unitario

$$\log(1+z) = z - \frac{z^2}{2} + \frac{z^3}{3} - \frac{z^4}{4} + \ldots = \sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} z^n \,, \tag{6.25}$$

da cui, scambiando z con $-z$, si ricava tra l'altro la somma $-\log(1-z)$ della (6.24). Possiamo ricavare la (6.25) da quello dell'esponenziale tenendo presente che una funzione è inversa dell'altra

$$e^{\log(1+z)} = 1 + z \,.$$

Al prim'ordine già sappiamo che $\log(1+z) \sim z$, per calcolare i coefficienti c_n dei termini successivi nella rappresentazione

$$\log(1+z) = z + c_2 z^2 + c_3 z^3 + \ldots$$

imponiamo che

$$e^{z + c_2 z^2 + c_3 z^3 + \ldots} - 1 = z + c_2 z^2 + c_3 z^3 + \ldots + \frac{(z + c_2 z^2 + c_3 z^3 + \ldots)^2}{2} + \ldots = z \,.$$

Si ottiene così

$$c_2 = -\frac{1}{2} \,, \; c_3 = \frac{1}{3} \,, \; c_4 = -\frac{1}{4} \ldots \,,$$

dunque la (6.25) è vera. Ora, questa serie converge anche nel punto $z = 1$ che sta sul bordo, ma chi ci assicura che la sua somma sia proprio

$$\lim_{x \to 1} \log(1+x) = \log 2 \,?$$

La risposta viene data nel seguente teorema.

Teorema 6.16 (di Abel) - *Se la serie $\sum c_n z^n$ ha per somma $f(z)$ per $|z| < R$ con R finito e converge con somma $S(\hat{z})$ in un punto $\hat{z}$ tale che $|\hat{z}| = R$ allora esiste il limite di $f(z)$ per $z \to \hat{z}$ lungo il raggio da 0 a $\hat{z}$ e tale limite coincide con $S(\hat{z})$.*

Dimostrazione. A meno di sostituire c_0 con $c_0 - S(\hat{z})$, possiamo supporre $S(\hat{z}) = 0$. Il raggio da 0 a $\hat{z}$ è l'insieme dei $t\hat{z}$ con $0 < t < 1$, quindi per $n \to \infty$

$$s_n(t\hat{z}) = \sum_{k=0}^{n} c_k (t\hat{z})^k = \sum_{k=0}^{n} c_k \hat{z}^k t^k = \sum_{k=0}^{n} \gamma_k t^k \;\to\; f(t\hat{z}) \,,$$

dove $\gamma_k = c_k \hat{z}^k$. Dobbiamo dimostrare che $\lim_{t \to 1} f(t\hat{z}) = 0$. Poiché $\gamma_0 = s_0(\hat{z})$ e $\gamma_k = s_k(\hat{z}) - s_{k-1}(\hat{z})$, si ha

$$\begin{aligned} s_n(t\hat{z}) &= \gamma_0 + \sum_{k=1}^{n} (s_k(\hat{z}) - s_{k-1}(\hat{z})) t^k = \gamma_0 + \sum_{k=1}^{n} s_k(\hat{z}) t^k - \sum_{k=0}^{n-1} s_k(\hat{z}) t^{k+1} \\ &= \gamma_0 - \gamma_0 t + s_n(\hat{z}) t^n + (1-t) \sum_{k=1}^{n-1} s_k(\hat{z}) t^k = s_n(\hat{z}) t^n + (1-t) \sum_{k=0}^{n-1} s_k(\hat{z}) t^k \end{aligned}$$

e passando al limite per $n \to \infty$ si ottiene

$$f(t\hat{z}) = (1-t) \sum_{k=0}^{\infty} s_k(\hat{z}) t^k \,.$$

Ricordando che $s_n(\hat{z}) \to 0$, per ogni $\varepsilon > 0$ si ha per ν abbastanza grande

$$\Big|\sum_{k=\nu+1}^{\infty} s_k(\hat{z})t^k\Big| \leqslant \sum_{k=\nu+1}^{\infty} |s_k(\hat{z})|t^k < \frac{\varepsilon t^{\nu+1}}{1-t} < \frac{\varepsilon}{1-t}\,,$$

mentre per $1-\delta < t < 1$, con $\delta > 0$ opportuno, si ha

$$(1-t)\Big|\sum_{k=0}^{\nu} s_k(\hat{z})t^k\Big| < \varepsilon\,.$$

Pertanto $|f(t\hat{z})| < 2\varepsilon$ per ogni $t \in]1-\delta, 1[$ che è la tesi.

□

Questo risultato ci permette di calcolare come s'è detto la somma della serie

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \log 2$$

che già sapevamo convergente per il criterio di Leibniz. Purtroppo l'approssimazione è pessima, nel senso estremamente lenta, perché la prima somma parziale ad avere le stesse prime due cifre di $\log 2 \sim 0.6931471805599453$ è la 160-esima, $s_{160} = 0,6900319459942253$. La ragione sta nel fatto che siamo in un punto del bordo dove la convergenza non vale nel senso assoluto. Per aumentarne la rapidità facciamo la somma membro a membro delle (6.25) e (6.24) ottenendo la nuova serie di potenze

$$\frac{1}{2}\log\frac{1+z}{1-z} = z + \frac{z^3}{3} + \frac{z^5}{5} + \cdots = \sum_{n=1}^{\infty} \frac{z^{2n-1}}{2n-1} \tag{6.26}$$

che per $z = 3/5$ fornisce un'altra serie numerica con somma $\log 2$

$$\log 2 = \sum_{n=1}^{\infty} \frac{(3/5)^{2n-1}}{2n-1}\,.$$

Stavolta il punto in cui va calcolata si trova all'interno del disco e i termini hanno andamento pressoché esponenziale, già la somma dei primi sei termini soltanto vale 0.6930006452197403.

Guardando ai termini che formano la serie prodotto

$$\begin{aligned}(a_0 + a_1z + a_2z^2 + a_3z^3 + \ldots)&(b_0 + b_1z + b_2z^2 + b_3z^3 + \ldots)\\ &= a_0b_0 + (a_0b_1 + a_1b_0)z + (a_0b_2 + a_1b_1 + a_2b_0)z^2\\ &+ (a_0b_3 + a_1b_2 + a_2b_1 + a_3b_0)z^3 + \ldots\,.\end{aligned}$$

riconosciamo subito proprio quello di Cauchy, il quale appare adesso talmente naturale da indurci a pensare che potrebbe essere stato inventato apposta per le serie di potenze. Poiché esse convergono assolutamente nei punti interni al disco di convergenza, risulta che la serie prodotto ha per somma il prodotto delle somme e converge assolutamente nel più piccolo dei due dischi di convergenza.

Esempi

6.8 *Il prodotto di Cauchy della serie geometrica*

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} z^n$$

per se stessa è dato da

$$(1 + z + z^2 + \ldots + z^n + \ldots)(1 + z + z^2 + \ldots + z^n + \ldots)$$
$$= 1 + 2z + 3z^2 + \ldots + (n+1)z^n + \ldots = \frac{1}{(1-z)^2}$$

nel senso della convergenza assoluta sul disco $|z| < 1$*.*

6.9 *La funzione esponenziale*

$$e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!} \tag{6.27}$$

è ben definita su **C** *come funzione intera. Basta osservare che*

$$\lim_{n\to\infty} \frac{c_{n+1}}{c_n} = \lim_{n\to\infty} \frac{n!}{(n+1)!} = \lim_{n\to\infty} \frac{1}{n+1} = 0\,.$$

Perché la serie (6.27), detta *serie esponenziale*, converge proprio alla funzione esponenziale? Già sappiamo, v. la (5.5), che per $z = 1$ la somma è il numero e. Inoltre si vede facilmente che il prodotto di Cauchy delle serie (assolutamente convergenti) calcolate in z_1 e in z_2 coincide con la serie calcolata in $z_1 + z_2$, infatti

$$\sum_{k=0}^{n} \frac{z_1^k z_2^{n-k}}{k!(n-k)!} = \frac{(z_1+z_2)^n}{n!}$$

per la formula del binomio. Considerate insieme, queste due osservazioni ci permettono di concludere che si tratta della funzione esponenziale.

Vediamo come dalla (6.27) discendano simili rappresentazioni per le funzioni trigonometriche. Per questo basta porre $z = iy$ nella (6.27)

$$e^{iy} = \sum_{n=0}^{\infty} \frac{(iy)^n}{n!}\,.$$

Tenendo presente che $i^{2n} = (-1)^n$ e $i^{2n+1} = (-1)^n i$, separando la parte reale dalla parte immaginaria si ottiene

$$\begin{cases} \cos y = \displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n y^{2n}}{(2n)!} = 1 - \frac{y^2}{2} + \frac{y^4}{4!} - \frac{y^6}{6!} + \ldots \\ \operatorname{sen} y = \displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n y^{2n+1}}{(2n+1)!} = y - \frac{y^3}{3!} + \frac{y^5}{5!} - \frac{y^7}{7!} + \ldots \end{cases}$$

6.10 *Sono funzioni intere anche le estensioni a* **C** *delle precedenti funzioni trigonometriche e delle iperboliche che abbiamo definito nella* (5.14)

$$\begin{cases} \cos z = \displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n z^{2n}}{(2n)!} \\ \operatorname{sen} z = \displaystyle\sum_{n=0}^{\infty} \frac{(-1)^n z^{2n+1}}{(2n+1)!} \end{cases} \quad e \quad \begin{cases} \cosh z = \displaystyle\sum_{n=0}^{\infty} \frac{z^{2n}}{(2n)!} \\ \operatorname{senh} z = \displaystyle\sum_{n=0}^{\infty} \frac{z^{2n+1}}{(2n+1)!} \end{cases}$$

Esercizio 6.20 - *Dimostrare che le serie di potenze centrate in* 0*, che sul disco di convergenza hanno per somma una funzione pari [dispari], presentano solo termini di grado pari [dispari].*

Qualche altro sviluppo in serie di potenze può essere ottenuto algebricamente per confronto diretto dei coefficienti, come nel seguente esercizio, altri ancora ne otterremo più avanti con strumenti del calcolo differenziale.

Esercizio 6.21 - *Ricavare le serie di potenze attorno a* 0 *per le funzioni* $\operatorname{tang} x = ax + bx^3 + \dots$ *e* $\sqrt{1+x} = 1 + ax + \dots$ *sapendo che* $\operatorname{tang} x \cos x = \operatorname{sen} x$ *e* $(\sqrt{1+x})^2 = 1 + x$. *Dedurne quelli per* $\sqrt{1-x}$, $\sqrt{1-x^2}$, $1/\sqrt{1-x^2}$ *e* $\operatorname{arctg} x$.

Possiamo adesso illustrare, rimanendo ad un livello intuitivo, uno dei modi di calcolare la somma della serie armonica generalizzata di esponente $\alpha = 2$. Un polinomio $P(z)$ con radici z_k di molteplicità m_k, $k = 1, \dots, n$, ammette la fattorizzazione

$$P(z) = A \prod_{k=0}^{n} \left(1 - \frac{z}{z_k}\right)^{m_k} .$$

Analogamente, si può dimostrare che anche per le funzioni intere su $\mathbf{C}$ vale una rappresentazione simile, ma in termini di un *prodotto infinito* se gli zeri sono infiniti. Come per le serie, si pone subito la questione della convergenza: il prodotto verrà detto *convergente* se lo è la successione dei prodotti finiti e si scrive

$$\prod_{n=0}^{\infty} \left(1 - \frac{z}{z_n}\right)^{m_n} = \lim_{n\to\infty} \prod_{k=0}^{n} \left(1 - \frac{z}{z_k}\right)^{m_k} .$$

Passando ai logaritmi si riconduce tale definizione a quella di serie convergente e si trova come condizione, solo necessaria, che la successione dei fattori deve tendere a 1.

Anche la funzione $\operatorname{sen} z$, i cui zeri sono i punti $z_n = \pm n\pi$, è rappresentabile in termini di un prodotto infinito

$$\operatorname{sen} z = z \prod_{n=1}^{\infty} \left(1 - \frac{z}{\pi n}\right)\left(1 + \frac{z}{\pi n}\right) = z \prod_{n=1}^{\infty} \left(1 - \frac{z^2}{\pi^2 n^2}\right) . \tag{6.28}$$

Se adesso si confronta la (6.28) con la serie di potenze

$$\operatorname{sen} z = z - \frac{z^3}{3!} + \frac{z^5}{5!} - \frac{z^7}{7!} + \dots = z\left(1 - \frac{z^2}{\pi^2}\right)\left(1 - \frac{z^2}{2^2\pi^2}\right)\left(1 - \frac{z^2}{3^2\pi^2} \cdots\right)$$

e si eguagliano i coefficienti dei termini di terzo grado si ottiene

$$\frac{1}{3!} = \frac{1}{\pi^2} + \frac{1}{2^2\pi^2} + \frac{1}{3^2\pi^2} + \dots$$

e quindi $\pi^2/6$ come somma della serie armonica di esponente 2.

Il lettore accorto avrà certamente notato come l'ampiezza dell'intervallo, o l'estensione del cerchio, di convergenza sia prevedibile in base alla struttura della funzione cui la serie converge. La funzione $f(x) = (1-x)^{-1}$ ha come dominio naturale $\mathbf{R} - \{1\}$, ma il suo sviluppo in serie di potenze attorno a 0 è valido solo su $]-1,1[$. L'intuizione ne suggerisce una spiegazione ovvia: se il dominio di convergenza, che deve essere un intervallo centrato in 0, fosse più grande, comprenderebbe anche il punto 1 dove la f non è definita e intorno al quale non è limitata, per questo il punto 1 è considerato un *punto singolare*, detto anche *polo*. Che cosa succede al di fuori, dove la serie non converge, ma c'è comunque la funzione? La risposta è molto semplice, basta considerare un centro $x_0 \neq 0$ e si ha

$$f(x) = \frac{1}{1-x} = \frac{1}{1 - x_0 - (x - x_0)} = \frac{1}{1-x_0} \cdot \frac{1}{1 - \dfrac{x-x_0}{1-x_0}} = \sum_{n=0}^{\infty} \frac{(x-x_0)^n}{(1-x_0)^{n+1}} .$$

Questa nuova serie ha per coefficienti $(1-x_0)^{-(n+1)}$ e converge per $|x-1| < |x_0-1|$, il raggio di convergenza è ancora la distanza del centro dal punto 1. Per la funzione esponenziale, evidentemente ogni serie di potenze, ovunque centrata, converge ad essa su tutto $\mathbf{R}$, o su tutto $\mathbf{C}$, perché stavolta non ci sono punti singolari. Come si spiega allora che lo sviluppo in serie attorno a 0 della funzione priva di poli

$$f(x) = \frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n} \tag{6.29}$$

converge nell'intervallo limitato $]-1,1[$ e non su $\mathbf{R}$? Perché in realtà si tratta della funzione complessa $f(z) = (1+z^2)^{-1}$ ed ammette $\pm i$ come poli. Il cerchio di convergenza è sempre il massimo cerchio che non li contiene. Ad un altro centro $z_0 \neq 0$ corrisponde un altro cerchio di convergenza su cui la nuova serie di potenze converge alla (6.29), anche in nuovi punti che il cerchio iniziale non comprendeva. Infatti

$$\begin{aligned} f(z) &= \frac{i}{2}\left(\frac{1}{z+i} - \frac{1}{z-i}\right) = \frac{i}{2}\left(\frac{1}{i+z_0+(z-z_0)} + \frac{1}{i-z_0-(z-z_0)}\right) \\ &= \frac{i}{2}\left[\sum_{n=0}^{\infty} \frac{(-1)^n (z-z_0)^n}{(i+z_0)^{n+1}} + \sum_{n=0}^{\infty} \frac{(z-z_0)^n}{(i-z_0)^{n+1}}\right] = \sum_{n=0}^{\infty} c_n (z-z_0)^n \end{aligned}$$

ha per raggio di convergenza la minima tra le distanze di z_0 da i e da $-i$. Man mano che si considerano nuovi cerchi si definiscono altrettante estensioni della funzione.

Definizione 6.17 - *Una funzione $f : A \to \mathbf{C}$, con A aperto di $\mathbf{C}$, viene detta **analitica** in A se per ogni $z_0 \in A$ esiste $r > 0$ tale che f è la somma di una serie di potenze sul cerchio di centro z_0 e raggio r.*

Vediamo, per concludere, un esempio di funzione analitica che non ammette prolungamento analitico al di fuori del cerchio in cui è definita come somma di una serie di potenze. La funzione

$$f(z) = \sum_{n=0}^{\infty} z^{n!}$$

è ben definita sul cerchio unitario dove la serie converge ed è ivi analitica. Però non ammette prolungamento analitico al di fuori perché la serie non converge in un insieme denso di punti del bordo, precisamente quelli individuati dagli angoli che hanno rapporto razionale con π. Fissato uno di questi, $h\pi/k$, chiaramente almeno da k in poi, quindi per ogni $n > k$, si ha $e^{n!h\pi i/k} = 1$, pertanto

$$\sum_{n=0}^{\infty} z^{n!}{}_{|\vartheta = h\pi/k} = \sum_{n=0}^{k} \rho^n e^{n!h\pi i/k} + \sum_{n=k+1}^{\infty} \rho^{n!}\,.$$

Ora, per $\rho \to 1$ il secondo termine diverge. Infatti, fissiamo dapprima un numero naturale $m > k$ e osserviamo che

$$\sum_{n=k+1}^{\infty} \rho^{n!} > \sum_{n=k+1}^{m} \rho^{n!} > (m-k)\rho^{m!}\,.$$

Scelto a piacere $0 < \varepsilon < 1$, esiste $\delta > 0$ tale che se $1-\delta < \rho < 1$ si ha $\rho^{m!} > 1-\varepsilon$. Allora

$$\sum_{n=k+1}^{\infty} \rho^{n!} > (m-k)(1-\varepsilon)$$

per $1-\delta < \rho < 1$.

Capitolo 7

Spazi metrici

7.1 Palle e intorni

Nel § 2.5 abbiamo appena accennato, senza nessuna pretesa di rigore, al fatto che metriche diverse su uno stesso insieme X possono talvolta essere considerate equivalenti rimanendo ad esse associata la stessa nozione di vicinanza tra punti. È possibile formalizzare l'idea di intorno di un punto, inteso come una regione dello spazio formata da punti ad esso vicini, in qualche senso, e pervenire alla definizione di *spazio topologico* utilizzando una distanza. Ma in realtà il concetto di intorno prescinde da quello di distanza; in primo luogo, anche rimanendo nell'ambito degli spazi metrici, distanze diverse possono definire la stessa famiglia di intorni di ogni punto, inoltre è possibile introdurre direttamente, attraverso una lista di assiomi, famiglie di intorni con cui l'insieme X acquisisce una struttura topologica, talvolta non riconducibile ad una distanza. Mentre uno spazio metrico è anche topologico, il viceversa non è vero, vi sono spazi topologici non *metrizzabili.* In ogni caso è conveniente introdurre, per ogni $x \in X$, un *sistema fondamentale di intorni*, cioè una famiglia $\mathscr{B}(x)$ di intorni semplici da descrivere, ma abbastanza "ricca" da permetterci di individuare tutti gli intorni di x. Un insieme $U \subset X$ sarà un intorno di x se contiene almeno un elemento di $\mathscr{B}(x)$. Per gli scopi di questo corso ci limitiamo agli spazi metrici, vediamo quindi come si usa la nozione di distanza per definire gli intorni.

Dato uno spazio metrico (X, d), e definita per ogni punto $x \in X$ la *palla di centro* x e *raggio* $r > 0$

$$B_r(x) = \{y \in X \mid d(x, y) < r\},$$

assumiamo come sistema fondamentale di intorni di x la famiglia

$$\mathscr{B}(x) = \{B_r(x) \mid r > 0\}$$

di tutte le palle di centro x. Gli intorni fondamentali di un punto $x \in \mathbf{R}$ sono gli intervalli aperti $]x-r, x+r[$, in $\mathbf{R}^2$ e in $\mathbf{R}^n$ i cerchi e le sfere, senza il bordo, di centro x.

Definizione 7.1 - *Si chiama **intorno** del punto $x \in X$ ogni insieme $U \subset X$ per cui esiste $r > 0$ tale che $B_r(x) \subset U$. Indicheremo con $\mathscr{I}(x)$ la famiglia di tutti gli intorni di x.*

Ovviamente se U è intorno di x allora $x \in U$ e ogni insieme che contiene un intorno di x è intorno di x. Dalla Definizione 7.1 discendono anche le seguenti che lasciamo per esercizio.

Esercizio 7.1 - *Dimostrare che l'intersezione di due intorni di x è anch'esso un intorno di x.*

Esercizio 7.2 - *Dimostrare che se U è intorno di x esiste un intorno V di x tale che U è intorno di ogni punto $y \in V$.*

Esercizio 7.3 (assioma di Hausdorff) - *Dimostrare che se $x \neq y$ esistono un intorno U di x ed un intorno V di y tali che $U \cap V = \emptyset$.*

Se in $\mathbf{R}$ prendiamo come sistema fondamentale di intorni di un punto x l'insieme di tutte le semirette aperte $]a, +\infty[$ che contengono x la proprietà dell'Esercizio 7.3 non vale più, evidentemente la struttura topologica indotta da intorni del genere non è metrizzabile.

Due metriche d e d' su X si dicono *equivalenti* se inducono la stessa *topologia* su X, cioè se definiscono per ogni punto x la stessa famiglia $\mathscr{I}(x)$ di intorni di x. Una condizione sufficiente è la seguente

$$\exists c, C > 0 \ : \quad cd(x,y) \leqslant d'(x,y) \leqslant Cd(x,y) \qquad \forall x, y \in X \,. \tag{7.1}$$

In tal caso d e d' inducono su X topologie "somiglianti", ogni intorno di x per la prima topologia contiene un intorno di x per la seconda e viceversa. In $\mathbf{R}^2$, ad esempio, le distanze di $x = (x_1, x_2)$ da $y = (y_1, y_2)$

$$d_p(x,y) = \begin{cases} (|x_1 - y_1|^p + |x_2 - y_2|^p)^{1/p} & \text{se } 1 \leqslant p < +\infty \\ \max\{|x_1 - y_1|, |x_2 - y_2|\} & \text{se } p = +\infty \,, \end{cases}$$

tra le quali quella euclidea corrisponde a $p = 2$, sono tutte equivalenti al variare di p perché ogni palla di centro x relativa ad un certo valore di p contiene (ed è contenuta in) una palla relativa ad un altro valore. Si può descrivere la stessa proprietà anche così: per ogni $p, q \in [1, +\infty]$ esiste un'applicazione bigettiva da $(\mathbf{R}^2, d_p)$ in $(\mathbf{R}^2, d_q)$ che trasforma ogni intorno di x per la metrica d_p in un intorno di x per la metrica d_q. Tale applicazione viene detta *omeomorfismo* e gli spazi $(\mathbf{R}^2, d_p)$ e $(\mathbf{R}^2, d_q)$ *omeomorfi*.

Più in generale, anche due spazi metrici distinti (X, d) e (X', d') possono essere muniti di topologie somiglianti e quindi essere omeomorfi. Deve esistere un'applicazione bigettiva da uno all'altro che conserva gli intorni, cioè un *omeomorfismo*. Precisamente, (X, d) e (X', d') saranno omeomorfi, e le loro metriche equivalenti, se esiste un'applicazione bigettiva $\varphi : X \to X'$ tale che per ogni $x \in X$ e per ogni suo intorno U la sua immagine $\varphi(U)$ è un intorno del punto $\varphi(x) \in X'$.

L'analogo della (7.1), più restrittiva anche stavolta, è

$$cd(x,y) \leqslant d'(\varphi(x), \varphi(y)) \leqslant Cd(x,y) \qquad \forall x, y \in X \,. \tag{7.2}$$

Come casi particolari, diciamo che X e X' sono spazi metrici *simili* se $d'(\varphi(x), \varphi(y)) = Cd(x,y)$ e *isometrici* se $C = 1$. Senza pretendere in questa sede di approfondire la questione, possiamo immaginare X come una regione, una curva o una superficie, dello spazio, e un omeomorfismo da X in X' come una deformazione che non causa fratture, tagli, buchi, non ricuce, non incolla e non causa cambiamenti di dimensione. Allora punti tra loro "vicini" (rispetto ad una distanza) rimangono "vicini" (anche rispetto all'altra) durante la trasformazione. La funzione $(x, 0) \to (x, x^2)$ è un esempio di omeomorfismo che deforma la retta in una parabola.

Esercizio 7.4 - *Verificare che l'insieme delle funzioni reali limitate su $[0,1]$ è uno spazio metrico con la distanza*

$$d(f,g) = \sup_{x \in [0,1]} |f(x) - g(x)| \,.$$

Esercizio 7.5 - *Verificare che nell'insieme più ristretto dei polinomi, sempre su $[0,1]$, anche la*

$$d'(P,Q) = \sum_{k=0}^{n} |a_k - b_k| \,,$$

essendo gli a_k i coefficienti di P e i b_k quelli di Q (eventualmente nulli da $m+1$ a n se il grado m di Q fosse minore del grado n di P), è una distanza.

In che relazione stanno d e d' per i polinomi su $[0,1]$? Per ogni $\varepsilon > 0$ esiste $x_\varepsilon \in [0,1]$ tale che

$$\begin{aligned} d(P,Q) &< \varepsilon + \left| \sum_{k=0}^{n} (a_k - b_k) x_\varepsilon^k \right| \leqslant \varepsilon + \sum_{k=0}^{n} |a_k - b_k| x_\varepsilon^k \\ &\leqslant \varepsilon + \sum_{k=0}^{n} |a_k - b_k| = \varepsilon + d'(P,Q)\,, \end{aligned}$$

da cui, essendo ε arbitrario, $d(P,Q) \leqslant d'(P,Q)$. Tuttavia per nessuna costante $C > 0$ può valere la disuguaglianza opposta. Scelto $Q = 0$, riusciamo infatti a costruire una successione di polinomi (P_h), $h \in \mathbf{N}$, tale che

$$\lim_{n\to\infty} \frac{d'(P_h,Q)}{d(P_h,Q)} = +\infty\,,$$

basta definire

$$P_h(x) = \frac{1}{h^2} - \frac{x^2}{h}$$

e si ha

$$d'(P_h,Q) = \frac{1}{h^2} + \frac{1}{h} \qquad \text{e} \qquad d(P_h,Q) = \frac{1}{h^2}\,.$$

7.2 Insiemi chiusi, insiemi aperti e successioni

Diamo adesso delle definizioni sulle relazioni di vicinanza tra punti e insiemi. Sia S un insieme qualunque dello spazio metrico X.

Definizione 7.2 - *Un punto $x \in X$ viene detto* ***aderente*** *ad S se ogni intorno U di x ha intersezione non vuota con S, in simboli*

$$U \cap S \neq \emptyset \quad \forall U \in \mathscr{I}(x)\,. \tag{7.3}$$

L'insieme dei punti aderenti ad S si chiama ***chiusura*** *di S e si indica con $\bar{S}$.*

Ovviamente sono aderenti ad S tutti i suoi punti, quindi $S \subset \bar{S}$, ma possono esistere dei punti non di S ad esso aderenti, anche questi appartengono ad $\bar{S}$. Se ad esempio $S =]0,1[$ i suoi punti aderenti sono tutti e soli quelli dell'intervallo $[0,1]$, che dunque coincide con $\bar{S}$. Oppure, sono aderenti all'insieme $S = \{1/n \mid n \geqslant 1\}$ i punti di S e 0, $\bar{S} = S \cup \{0\}$. Più in generale, la chiusura dell'insieme dei valori di una successione convergente è l'insieme stesso unito il limite. Un sottoinsieme limitato di $\mathbf{R}$ ammette i suoi estremi, inferiore e superiore, come punti aderenti. Se $S =]0,1[\cup \{2\}$ $\bar{S} = [0,1] \cup \{2\}$.

La chiusura di S è anche l'insieme dei punti a distanza nulla da S, essendo la distanza di un punto $x \in X$ da un insieme $S \subset X$ così definita

$$d(x,S) = \inf_{y\in S} d(x,y)\,. \tag{7.4}$$

Questa caratterizzazione della chiusura è una banale applicazione della definizione stessa di estremo inferiore, provare a dimostrarlo.

Definizione 7.3 - *Un insieme $C \subset X$ si dice* ***chiuso*** *se $C \supset \bar{C}$, cioè se $C = \bar{C}$.*

In particolare l'intero spazio X è chiuso. In $\mathbf{R}$ (con la distanza euclidea) gli intervalli chiusi sono dei sottoinsiemi chiusi. Poiché ogni intorno di un numero reale contiene sempre dei numeri razionali, la chiusura di $\mathbf{Q}$ è $\mathbf{R}$, di conseguenza $\mathbf{Q}$ non è chiuso. Per analogia con questo caso diamo la definizione di insieme denso in X.

Definizione 7.4 - *Diciamo che $S \subset X$ è **denso** in X se $\bar{S} = X$, inoltre, dati $S_1 \subset S_2$, S_1 è denso in S_2 se $\bar{S}_1 \supset S_2$.*

Ovviamente se $S_1 \subset S_2 \subset S_3$ con S_1 denso in S_2 e S_2 denso in S_3 allora S_1 è denso in S_3, e se S_1 è denso in S_3 allora anche S_2 è denso in S_3.

Esercizio 7.6 - *Dimostrare che l'insieme dei numeri irrazionali è denso in* $\mathbf{R}$.

Definizione 7.5 - *Si chiama **bordo** o **frontiera** di S, e si indica con ∂S, l'insieme dei punti aderenti sia a S sia a $\complement S$. In simboli*

$$\partial S = \bar{S} \cap \overline{\complement S}\,.$$

Quali nuovi punti dobbiamo aggiungere ad un insieme S non chiuso per ottenerne la chiusura?

Definizione 7.6 - *Un punto $x \in X$ viene detto di **accumulazione** per S se ogni intorno U di x, tolto x stesso, ha intersezione non vuota con S. In simboli*

$$U \cap S - \{x\} \neq \emptyset \quad \forall U \in \mathscr{I}(x)\,. \tag{7.5}$$

*L'insieme dei punti di accumulazione ad S si dice **derivato** di S e si indica con $\mathscr{D}(S)$.*

Tutti i punti di $S =]0,1[$ stanno nella chiusura di S in quanto elementi di S, e sono anche di accumulazione, ma ne fanno parte anche 0 e 1 in quanto di accumulazione anche se non in S, dunque $\bar{S} = \mathscr{D}(S) = [0,1]$. Se $S =]0,1[\cup \{2\}$ il punto 2 è aderente a S solo in quanto elemento di S, ma non è di accumulazione, è *isolato* in S. In altre parole anche ora $\mathscr{D}(S) = [0,1]$, ma $\bar{S} \supsetneqq \mathscr{D}(S)$.

In generale $\bar{S} = S \cup \mathscr{D}(S)$ perché per due motivi x può essere aderente a S: perché già appartiene a S o perché è di accumulazione per S (o per entrambi). In particolare $\mathscr{D}(S) = \emptyset$ implica $S = \bar{S}$.

Definizione 7.7 - *Un punto $x \in S$ viene detto **punto isolato** di S se non è di accumulazione per S.*

I sottoinsiemi finiti di $\mathbf{R}$, ma anche $\mathbf{N}$ e $\mathbf{Z}$, sono composti di soli punti isolati e sono anche chiusi in quanto privi di punti di accumulazione. Anche ogni successione è fatta di punti isolati, ma se non ha sottosuccessioni costanti e ammette dei punti limite questi sono anche tutti e soli i suoi punti di accumulazione. Ne segue che le successioni con dei punti limite sono chiuse solo se assumono per qualche indice n i valori di questi limiti.

Una versione equivalente del Teorema 5.28 afferma che se infiniti punti, comunque distribuiti, occupano una porzione limitata di $\mathbf{R}$ o di $\mathbf{R}^n$ allora devono presentare delle zone di "addensamento". Più precisamente vale il seguente teorema.

Teorema 7.8 (di Bolzano -Weierstraß) - *Ogni insieme S limitato e infinito di numeri reali ammette in $\mathbf{R}$ un punto di accumulazione.*

Dimostrazione. Essendo S infinito, è possibile definire una successione (x_n) di punti di S che non abbia sottosuccessioni costanti. Per il Teorema 5.28 esistono una sottosuccessione $(x_{k_n}) \subset (x_n)$ e un $x \in \mathbf{R}$ tali che $x_{k_n} \to x$. Per la definizione di limite x è di accumulazione per S.

□

Il motivo per cui abbiamo scelto la (x_n) priva di sottosuccessioni costanti è per evitare che qualcuna di esse converga essendo definitivamente costante, in questo caso il suo limite potrebbe essere isolato rispetto alla successione, quindi sì aderente, ma non di accumulazione.

Definizione 7.9 - *Un punto $x \in S$ viene detto* ***interno*** *a S se S è intorno di x, cioè se S contiene, insieme a x, un intorno $U \in \mathscr{I}(x)$ (o una palla di centro x). In simboli*

$$\exists U \in \mathscr{I}(x) \ : \quad U \subset S\,. \tag{7.6}$$

L'insieme dei punti interni a S si chiama ***parte interna*** *di S e si indica con $\overset{\circ}{S}$.*

Ovviamente $\overset{\circ}{S} \subset S$ e $x \in \overset{\circ}{S}$ se e solo se x non è aderente al complementare di S. Ad esempio $\mathbf{Q}$ è privo di punti interni, $\overset{\circ}{\mathbf{Q}} = \emptyset$, perchè $\overline{\complement \mathbf{Q}} = \mathbf{R}$.

Definizione 7.10 - *Un insieme $A \subset X$ si dice* ***aperto*** *se $A \subset \overset{\circ}{A}$, cioè se $A = \overset{\circ}{A}$.*

Dunque è aperto ogni insieme che è intorno di ogni suo punto, formato cioè di soli punti interni. Esempi di aperti sono tutto lo spazio X, come $\mathbf{R}$, gli intervalli aperti, $\complement\mathbf{N}$, $\complement\mathbf{Z}$, ma non $\complement\mathbf{Q}$ che è privo di punti interni. Si verifica banalmente che A è aperto se e solo se $\complement A$ è chiuso. Dunque gli aperti di X sono tutti e soli i complementari dei chiusi e i chiusi sono tutti e soli i complementari degli aperti. In particolare $\emptyset$ e X sono sia aperti che chiusi (non sono gli unici), altri insiemi non sono né aperti, né chiusi.

Esercizio 7.7 - *L'intersezione di due, o di una famiglia finita, di aperti è un aperto e l'unione di due, o di una famiglia finita, di chiusi è un chiuso. L'unione di una famiglia arbitraria di aperti è un aperto e l'intersezione di una famiglia arbitraria di chiusi è un chiuso.*

Le relazioni dell'Esercizio 2.12

$$\bigcap_{n=1}^{\infty} \Big] -\frac{1}{n}, 1+\frac{1}{n}\Big[= [0,1] \qquad \text{e} \qquad \bigcup_{n=1}^{\infty} \Big[\frac{1}{n}, 1-\frac{1}{n}\Big] =]0,1[$$

insieme anche a queste

$$\bigcap_{n=1}^{\infty} \Big] -\frac{1}{n}, 1\Big[= [0,1[\qquad \text{e} \qquad \bigcup_{n=1}^{\infty} \Big[0, 1-\frac{1}{n}\Big] = [0,1[$$

mostrano che un'intersezione infinita di aperti può non essere un aperto e un'unione infinita di chiusi può non essere un chiuso.

Possiamo caratterizzare gli aperti in questo modo.

Esercizio 7.8 - *$A \subset X$ è aperto se e solo se è l'unione di una famiglia di palle.*

Si osservi che se X è uno spazio metrico *separabile*, che ammette cioè un sottoinsieme numerabile e denso, ogni aperto è esprimibile come unione di una famiglia numerabile di palle. In $\mathbf{R}$, ad esempio, nel quale $\mathbf{Q}$ è numerabile e denso, ogni aperto è unione di intervalli di centro razionale.

Esercizio 7.9 - *Dimostrare che per ogni $S, T \subset X$ si ha*

$$\overline{S \cup T} = \bar{S} \cup \bar{T}\,, \quad \overline{S \cap T} \subset \bar{S} \cap \bar{T}\,, \quad \overset{\circ}{\widehat{S \cap T}} = \overset{\circ}{S} \cap \overset{\circ}{T}\,, \quad \overset{\circ}{\widehat{S \cup T}} \supset \overset{\circ}{S} \cup \overset{\circ}{T}\,,$$

dove le inclusioni possono valere in senso stretto.

Esercizio 7.10 - *Dimostrare che*

$$\mathring{S} \subset S \subset \bar{S}\,, \quad \bar{\bar{S}} = \bar{S}\,, \quad \mathring{\mathring{S}} = \mathring{S}\,, \quad \bar{S} = \mathring{S} \cup \partial S\,.$$

Esercizio 7.11 - *Dimostrare che il complementare di un sottoinsieme numerabile di* $\mathbf{R}$ *è denso in* $\mathbf{R}$*, ad esempio* $\mathbf{R} - \mathbf{Z}$ *e* $\mathbf{R} - \mathbf{Q}$ *sono densi in* $\mathbf{R}$*, e che* $\partial\mathbf{Q} = \mathbf{R}$*.*

Esercizio 7.12 - *Dimostrare che la chiusura di un intervallo di qualunque tipo di estremi* a *e* b *è* $[a, b]$*, mentre la sua parte interna è* $]a, b[$*.*

Esercizio 7.13 - *Dimostrare che in ogni spazio metrico*

$$\overline{B}_r(x) = \{y \in X \mid d(x, y) \leqslant r\}\,.$$

Esercizio 7.14 - *Dare un esempio di insieme* $S \subset \mathbf{R}$ *tale che*

$$S\,,\ \mathring{S}\,,\ \bar{S}\,,\ \overline{\mathring{S}}\,,\ \mathring{\bar{S}}\,,\ \overline{\mathring{\bar{S}}}\,,\ \mathring{\overline{\mathring{S}}}$$

siano tutti diversi.

Esercizio 7.15 - *Determinare la chiusura dell'insieme*

$$S = \left\{(-1)^n \frac{n}{n+1} \;\middle|\; n \in \mathbf{N}\right\}.$$

La distanza tra due sottoinsiemi non vuoti S, T dello spazio metrico X è definita da

$$d(S, T) = \inf\{d(x, y) \mid x \in S\,,\ y \in T\}\,.$$

Esercizio 7.16 - *Calcolare la distanza tra due rette sghembe in* $\mathbf{R}^3$*.*

Esercizio 7.17 - *Trovare* $S, T \subset \mathbf{R}$ *tali che* $d(S, T) = 0$ *e* $S \cap T = \emptyset$*.*

Esercizio 7.18 - *Dimostrare che* $\mathbf{N}$ *e l'insieme*

$$S = \left\{n + \frac{1}{n} \;\middle|\; n \geqslant 1\right\}$$

sono a distanza nulla uno dall'altro.

Esercizio 7.19 - *Dimostrare che in* $\mathbf{R}^2$ *l'asse* x *e l'iperbole di equazione* $xy = 1$ *sono a distanza nulla.*

Definizione 7.11 - *Uno spazio metrico* X *è detto* ***limitato*** *se ha diametro finito*

$$\operatorname{diam} X = \sup_{x,y \in X} d(x, y) < +\infty\,.$$

Lo stesso si può dire per un insieme $S \subset X$*, il quale risulta dunque limitato se esiste una palla che lo contiene.*

Evidentemente la limitatezza non è una proprietà intrinseca all'insieme, ma dipende dalla metrica usata. Ogni distanza d è ad esempio equivalente alla

$$d'(x, y) = \frac{d(x, y)}{1 + d(x, y)} \quad \forall x, y \in X$$

che rende X limitato, essendo $\operatorname{diam}_{d'} X \leqslant 1$, in particolare $\operatorname{diam}_{d'} X = 1$ se e solo se (X, d) non è limitato. Per verificare la disuguaglianza triangolare (le altre due sono ovvie) osserviamo che la funzione $t \to t/(1+t)$ è crescente su $[0, +\infty[$. Posto allora $a = d(x, y)$, $b = d(x, z)$ e $c = d(y, z)$, si ha

$$\frac{a}{1+a} \leqslant \frac{b+c}{1+b+c} = \frac{b}{1+b+c} + \frac{c}{1+b+c} \leqslant \frac{b}{1+b} + \frac{c}{1+c} .$$

Per vedere che sono equivalenti, la disuguaglianza

$$d'(x, y) = \frac{d(x, y)}{1 + d(x, y)} \leqslant d(x, y)$$

è banalmente vera e mostra che ogni palla di centro x per la metrica d è contenuta in una palla di centro x per la metrica d'. L'altra disuguaglianza, $d \leqslant Cd'$ per un'opportuna costante C, è vera solo se X è limitato per entrambe le metriche, basta scegliere $C \geqslant 1 + \operatorname{diam}_d X$, altrimenti è falsa. Ma non per questo verrà meno la relativa inclusione tra gli intorni. Infatti, se $y \in B'_r(x)$, cioè se

$$d'(x, y) = \frac{d(x, y)}{1 + d(x, y)} < r ,$$

allora $d(x, y) < r/1 - r$, cioè $y \in B_{r/1-r}(x)$, se $r < 1$. Se $r \geqslant 1$ i punti y tali che $d'(x, y) < r$ sono gli stessi del caso $r = 1$, cioè tutto lo spazio X, il quale è intorno di x sia per una distanza che per l'altra. Rientra in questo esempio la distanza in $\mathbf{R}$ o in $\mathbf{R}^n$

$$d'(x, y) = \frac{|x - y|}{1 + |x - y|} .$$

Un altro esempio di spazio limitato, con diametro 2, è $\mathbf{R}$ con la distanza d_f della (2.6) che induce una metrica equivalente a quella euclidea. Questa distanza e la precedente possono essere estese a $\overline{\mathbf{R}}$ ponendo $d_f(-\infty, +\infty) = 2$ o $d'(-\infty, +\infty) = 1$ che fanno di $\overline{\mathbf{R}}$ uno spazio metrico limitato. Si osservi che i punti $-\infty$ e $+\infty$ sono di accumulazione per $\mathbf{R}$ e per i suoi sottoinsiemi non limitati. Naturalmente $\mathbf{R}$, come sottoinsieme di $\overline{\mathbf{R}}$, è un aperto ma non più un chiuso.

Esercizio 7.20 - *Dimostrare che la*

$$d''(x, y) = \min\{d(x, y), 1\} \quad \forall x, y \in X$$

definisce una metrica di spazio limitato equivalente a quella di d'.

Nel § 7.1 abbiamo detto che ad una distanza resta associato in modo naturale, per ogni punto $x \in X$, un sistema fondamentale di intorni $\mathscr{B}(x)$, quello delle palle di centro x e raggio r. Ma se invece di tutte le palle ne consideriamo solo una parte, purché arbitrariamente piccole, gli intorni non cambiano. In altre parole, se si sceglie una famiglia di palle $\mathscr{B}'(x) \subsetneqq \mathscr{B}(x)$ con la proprietà che per ogni $B \in \mathscr{B}(x)$ esiste $B' \in \mathscr{B}'(x)$ tale che $B' \subset B$ otteniamo un nuovo sistema fondamentale di intorni, più ristretto, che definisce gli stessi intorni di prima. Consideriamo allora il sistema fondamentale

$$\mathscr{B}'(x) = \{B_{1/n}(x) \mid n \in \mathbf{N} - \{0\}\}$$

che essendo una particolare famiglia numerabile di palle ci permette di usare le successioni per caratterizzare gli insiemi chiusi, gli aperti e tutte le proprietà topologiche dello spazio.

Siano $x \in X$ e $S \subset X$. Poiché $x \in \bar{S}$ se e solo se ogni palla $B_{1/n}(x)$ ha intersezione non vuota con S, per ogni $n \geqslant 1$ esiste un punto $x_n \in S \cap B_{1/n}(x)$, cioè tale che $x_n \in S$ e $d(x_n, x) < 1/n$. In questo modo, grazie alla numerabilità del sistema $\mathscr{B}'(x)$, viene a formarsi una successione (x_n) di punti di S che converge al punto $x \in \bar{S}$. In simboli

$$x \in \bar{S} \Leftrightarrow \exists (x_n) \subset S \ : \quad x_n \to x \tag{7.7}$$

dove non si esclude che la successione possa essere costante (definitivamente), $x_n = x$, nel caso che x si trovi già in S, ma se $x \in \mathscr{D}(S)$ la (7.7) va modificata nella forma

$$x \in \mathscr{D}(S) \Leftrightarrow \exists (x_n) \subset S \ : \quad \forall n \geqslant 1 \; x_n \neq x \text{ e } x_n \to x \,.$$

Essendo ad esempio gli estremi inferiore e superiore di un insieme $S \subset \mathbf{R}$ aderenti a S, sono limiti di successioni di punti di S.

Esercizio 7.21 - *Dimostrare che la chiusura in $\mathbf{R}^2$ del grafico della funzione*

$$f(x) = \operatorname{sen} \frac{1}{x} \quad x \neq 0$$

è formata dal grafico di f unito all'insieme $\{(0, y) \mid -1 \leqslant y \leqslant 1\}$.

Esercizio 7.22 - *Dimostrare che S è denso in T se e solo se per ogni $x \in T$ esiste una successione $(x_n) \subset S$ tale che $x_n \to x$.*

Si possono allora caratterizzare gli insiemi chiusi in questo modo:

$$C \subset X \text{ \it è chiuso} \Leftrightarrow \forall (x_n) \subset S \quad x_n \to x \Rightarrow x \in C \,.$$

Poiché A è aperto se nessun punto di A è aderente a $\complement A$, ogni successione $(x_n) \subset X$ convergente ad un punto $x \in A$ deve trovarsi definitivamente dentro A. Questa proprietà caratterizza gli aperti:

$$A \subset X \text{ \it è aperto} \Leftrightarrow \forall (x_n) \subset X \ : \ x_n \to x \in A \; \exists \nu \in \mathbf{N} \ : \ \forall n > \nu \; x_n \in A \,.$$

Ricordando che A, in quanto aperto, è intorno di x, l'affermazione precedente è coerente con la definizione di limite per una successione.

Esercizio 7.23 - *Dimostrare che in X con la distanza atomica* (2.5) *le sole successioni convergenti sono quelle definitivamente costanti. Inoltre ogni insieme formato da un solo punto à aperto, tutti gli insiemi sono aperti e quindi anche chiusi.*

Esercizio 7.24 - *Se in $\mathbf{R}$ si assume la famiglia di intervalli $\{]a, +\infty[\ : \ a < x\}$ come sistema fondamentale di intorni di x una successione convergente a x converge anche ad ogni $y < x$.*

Molto di ciò che è stato detto sulle successioni di numeri reali nel Cap. 5 si può estendere in modo naturale alle successioni a valori in uno spazio metrico qualsiasi. Naturalmente non avranno più significato tutte le proprietà legate alla struttura di $\mathbf{R}$ come corpo ordinato, e anche la distinzione tra divergere e non avere limite non ha più molto senso perché, per un motivo o per un altro, il limite viene comunque a mancare come punto dello spazio. Invece si conservano le proprietà di carattere generale delle successioni convergenti, come l'unicità del limite per l'assioma di separazione dell'Esercizio 7.3, la limitatezza e la condizione Cauchy. Quest'ultima è identica alla Definizione 5.5 che abbiamo dato nel caso di $\mathbf{R}$, pur di sostituire la distanza euclidea del valore assoluto con la distanza relativa a X

$$(x_n) \subset X \text{ \it è di Cauchy se } \forall \varepsilon > 0 \; \exists \nu \in \mathbf{N} \ : \quad d(x_m, x_n) < \varepsilon \quad \forall m, n > \nu \,.$$

Così ogni successione di Cauchy è limitata e, per la disuguaglianza triangolare della distanza, ogni successione convergente è di Cauchy, ne segue che ogni successione convergente è limitata.

I prossimi paragrafi sono strettamente collegati con le due questioni poste alla fine del Cap. 5. A livello di spazi metrici in generale, ma con ricadute interessanti sugli spazi euclidei, il Teorema 5.28 ci porta verso la nozione di *compattezza* mentre la condizione di Cauchy a quella di *completezza*.

7.3 Spazi compatti

Se è vero, stando al Teorema 5.28, che ogni successione limitata di numeri reali ha una sottosuccessione convergente, come deve essere scelto un insieme S affinché ogni successione di punti di S ammetta una sottosuccessione convergente ad un punto di S? Evidentemente deve essere limitato, in modo che lo sia anche ogni successione di punti di S, ma anche chiuso per contenerne i punti limite. Queste due condizioni sono sufficienti? Vediamo tra poco che in $\mathbf{R}$ (e in $\mathbf{R}^n$) la risposta è affermativa per il Teorema 5.28, ma in generale sono solo condizioni necessarie.

Definizione 7.12 - *Un insieme $K \subset X$, o anche X stesso, viene detto* ***compatto*** *se per ogni successione $(x_n) \subset K$ esistono una sottosuccessione $(x_{k_n}) \subset (x_n)$ e un punto $x \in K$ tali che $x_{k_n} \to x$.*

Teorema 7.13 - *Ogni insieme compatto $K \subset X$ è limitato e chiuso.*

Dimostrazione. Supponiamo che K non sia limitato. Fissato un punto $x_0 \in X$, per ogni $n \in \mathbf{N}$ esiste $x_n \in K$ tale che $x_n \notin B_n(x_0)$. Ma la successione (x_n) ammette un'estratta (x_{k_n}) convergente ad un certo $x \in K$ e siccome $x_{k_n} \notin B_{k_n}(x_0)$, si ha

$$k_n < d(x_0, x_{k_n}) \leqslant d(x_0, x) + d(x, x_{k_n})$$

che è assurdo perché a sinistra $k_n \to \infty$, mentre a destra $d(x_0, x) + d(x, x_{k_n}) \to d(x_0, x)$. Per dimostrare che K è anche chiuso consideriamo una successione $(x_n) \subset K$ che converge ad un punto $x \in X$ e dimostriamo che $x \in K$. Per la compattezza la (x_n) ammette una sottosuccessione convergente ad un certo $x' \in K$, ma se già la (x_n) converge a x anche ogni sua sottosuccessione deve convergere a x, quindi $x = x' \in K$.

□

Teorema 7.14 (di Heine-Cantor) - *Ogni insieme limitato e chiuso di numeri reali è compatto.*

Dimostrazione. Sia $K \subset \mathbf{R}$ limitato e chiuso. Ogni successione $(x_n) \subset K$ è dunque limitata, quindi ha una sottosuccessione convergente ad un certo $x \in \mathbf{R}$ e siccome K è chiuso, $x \in K$.

□

Allora in $\mathbf{R}$ e in $\mathbf{R}^n$ essere limitato e chiuso per un insieme equivale ad essere compatto, in dimensione non finita, invece, tutti i compatti sono limitati e chiusi, ma esistono insiemi limitati e chiusi che non sono compatti.

Esercizio 7.25 - *Dimostrare che ogni chiuso di un compatto è compatto.*

Esercizio 7.26 - *Dimostrare che ogni compatto in $\mathbf{R}$ ha massimo e minimo.*

La compattezza è una proprietà particolarmente significativa per l'esistenza di soluzioni nei problemi di massimo e minimo. A titolo di esempio trattiamo il problema della minima distanza da un insieme. Ricordiamo che la distanza di un punto $x \in X$ da un insieme $S \subset X$ è definita da

$$d(x, S) = \inf_{y \in S} d(x, y) .$$

Se esiste in S un punto $\bar{x}$ che realizza il minimo, cioè tale che $d(x, S) = d(x, \bar{x})$, tale punto si chiama *proiezione* di x su S. Una condizione sufficiente per l'esistenza della proiezione è la compattezza.

Problema 7.15 (della proiezione) - *Assegnati un insieme $S \subset X$ e un punto $x \in X$, trovare $\bar{x} \in S$ tale che*

$$d(x, \bar{x}) = d(x, S) .$$

Teorema 7.16 - *Se per S scegliamo un compatto K il Problema 7.15 ha soluzione.*

Dimostrazione. Poniamo
$$l = \inf_{y \in K} d(x, y) .$$
Per note proprietà dell'estremo inferiore, per ogni $n \in \mathbf{N}$ esiste $x_n \in K$ tale che
$$l \leqslant d(x, x_n) < l + \frac{1}{n} \quad \forall n \in \mathbf{N} .$$
La $(x_n) \subset K$, che si chiama *successione minimizzante* in quanto $d(x, x_n) \to l$ per confronto, ammette una sottosuccessione (x_{k_n}) convergente ad un certo $\bar{x} \in K$. Passando al limite nelle disuguaglianze
$$l \leqslant d(x, \bar{x}) \leqslant d(x, \bar{x}_{k_n}) + d(x_{k_n}, \bar{x}) < l + \frac{1}{k_n} + d(x_{k_n}, \bar{x}) ,$$
si ottiene $d(x, \bar{x}) = l$. Dunque l è la minima distanza di x dai punti di K e $\bar{x}$ la sua proiezione su K.

□

È facile costruire esempi in cui non vale l'unicità, in questi casi una successione minimizzante può avere più sottosuccessioni convergenti a proiezioni distinte, ma se la proiezione è unica ogni successione minimizzante deve convergere *tutta* alla soluzione.

Nello stesso modo si dimostra che l'analogo problema, forse meno interessante, della massima distanza ammette soluzione, ma cosa possiamo dire dell'esistenza della proiezione quando l'insieme è semplicemente un chiuso $C \in X$? Certamente non esiste il massimo se C non è limitato. Per il minimo, invece, ci si potrebbe ricondurre al Teorema 7.16 considerando, anziché tutto C, l'insieme $C_R = \overline{B}_R(x) \cap C$ con R abbastanza grande da rendere non vuoto C_R. Se un punto di minimo esiste va cercato in C_R perché i punti di $C - C_R$ sono più lontani da x di quelli di C_R. Questo modo di ragionare è quello giusto, ma richiede che gli insiemi limitati e chiusi siano compatti, quindi è applicabile solo al caso di $\mathbf{R}^n$.

In $\mathbf{R}^n$ una condizione sufficiente per l'unicità della proiezione è la convessità.

Definizione 7.17 - *Un insieme $S \subset \mathbf{R}^n$ viene detto **convesso** se*
$$\forall x, y \in S \quad \lambda x + (1 - \lambda) y \in S \quad \forall \lambda \in [0, 1] .$$
Dimostriamo che la proiezione su un insieme chiuso e convesso $C \in \mathbf{R}^n$ è unica usando l'identità del parallelogramma
$$|x - y|^2 + |x + y|^2 = 2|x|^2 + 2|y|^2 . \tag{7.8}$$
Se $\bar{x}, \bar{y} \in C$ sono due proiezioni di $x \in X$ dalla (7.8) si ottiene
$$|\bar{x} - \bar{y}|^2 = 2|x - \bar{x}|^2 + 2|x - \bar{y}|^2 - 4 \left| x - \frac{\bar{x} + \bar{y}}{2} \right|^2 \leqslant 4l^2 - 4l^2 = 0 .$$

Esercizio 7.27 - *Verificare che il piano π di equazione $ax + by + cz = 1$ è un convesso chiuso di $\mathbf{R}^3$. Calcolare la distanza del punto $O = (0, 0, 0)$ da π, trovare la proiezione P di O su π e mostrare che OP è un vettore ortogonale a π.*

7.4 Spazi completi

Abbiamo già usato il termine *completezza* a proposito della continuità di $\mathbf{R}$. In quel contesto viene coinvolta direttamente la relazione d'ordine totale che è una struttura molto particolare e tipica di $\mathbf{R}$. Tuttavia l'idea di uno "spazio continuo" deve poter essere estesa a insiemi più generali, come $\mathbf{R}^n$ per esempio, in cui un ordinamento naturale non c'è. Si tratta allora di trovare una buona definizione di spazio completo nella quale si possa riconoscere la natura continua. Il modo più semplice di procedere consiste nel cogliere nella completezza di $\mathbf{R}$ quegli aspetti che si possono svincolare dall'ordinamento. In termini di successioni, $\mathbf{Q}$ è denso in $\mathbf{R}$ perché ogni numero reale è il limite di una successione di numeri razionali, la quale quindi soddisfa la condizione di Cauchy. Ora, siccome il limite può essere irrazionale, le successioni in $\mathbf{Q}$ di Cauchy non hanno in generale limite in $\mathbf{Q}$, ed è questo il motivo per cui $\mathbf{Q}$ non è completo secondo la seguente definizione.

Definizione 7.18 - *Uno spazio metrico (X, d) viene detto **completo** se per ogni successione $(x_n) \subset X$ di Cauchy esiste $x \in X$ tale che $x_n \to x$.*

Nel Teorema 5.29 abbiamo dimostrato che $\mathbf{R}$ è completo.

Esercizio 7.28 - *Dimostrare che anche $\mathbf{R}^n$ è completo.*

Esercizio 7.29 - *Dimostrare che un insieme chiuso in uno spazio completo è completo.*

Esercizio 7.30 - *Dimostrare che un insieme completo in uno spazio metrico qualunque è chiuso.*

Esercizio 7.31 - *Dimostrare che ogni spazio metrico compatto è completo, in particolare un compatto di uno spazio qualunque è completo.*

Esercizio 7.32 - *Sia (C_n) una successione di chiusi in uno spazio metrico X completo tale che $C_{n+1} \subset C_n$ e $\operatorname{diam} C_n \to 0$ per $n \to \infty$. Dimostrare che l'intersezione di tutti i C_n non è vuota ed è formata da un unico punto.*

Togliendo una alla volta le ipotesi dell'Esercizio 7.32 ci si rende conto con dei controesempi che esse sono tutte essenziali. Gli intervalli $]0, 1/n[$ non sono chiusi e la loro intersezione è vuota, anche la famiglia di intervalli chiusi $[\sqrt{2}-1/n, \sqrt{2}+1/n] \cap \mathbf{Q}$ dello spazio metrico $\mathbf{Q}$ che non è completo ha intersezione vuota, gli intervalli $[1/n, 2/n]$ non sono una famiglia decrescente e hanno intersezione vuota, infine l'ipotesi $\operatorname{diam} C_n \to 0$ serve ovviamente per affermare che l'intersezione dei C_n è formata da un unico punto.

Come avviene col passaggio da $\mathbf{Q}$ a $\mathbf{R}$, esiste anche negli spazi metrici in generale la possibilità di passare da uno spazio qualsiasi ad uno spazio completo nel quale il primo sia denso. Questa operazione, nota come *completamento*, è enunciato nel seguente teorema che non dimostriamo.

Teorema 7.19 (di completamento) - *Se (X, d) è uno spazio metrico qualunque esiste un insieme $\bar{X}$ che contiene X ed una distanza $\bar{d} : X \times X \to \mathbf{R}$ tali che $\bar{d}_{|X \times X} = d$ e $(\bar{X}, \bar{d})$ è uno spazio metrico completo.*

Introduciamo adesso un'importante classe di applicazioni che hanno un ruolo fondamentale in molti teoremi di esistenza in Analisi.

Definizione 7.20 - *Sia X uno spazio metrico. Una funzione $F : X \to X$ viene detta **contrazione** se esiste una costante $L < 1$ tale che*

$$d(F(x), F(y)) \leqslant L d(x, y) \qquad \forall x, y \in X . \tag{7.9}$$

La (7.9) è la condizione di Lipschitz con una costante minore di 1 e, si faccia attenzione, non equivale a

$$d(F(x), F(y)) < d(x, y) \qquad \forall x, y \in X\,,$$

le due condizioni equivalgono rispettivamente a

$$\sup_{x,y\in X} \frac{d(F(x), F(y))}{d(x, y)} < 1 \qquad \text{e} \qquad \sup_{x,y\in X} \frac{d(F(x), F(y))}{d(x, y)} \leqslant 1\,.$$

Ad esempio la funzione $F(x) = \sqrt{1+x^2}$, $x \in \mathbf{R}$, soddisfa la seconda, ma non è una contrazione.

Teorema 7.21 (di Banach-Caccioppoli) - *Se X è uno spazio metrico completo e $F : X \to X$ è una contrazione allora l'equazione ai punti fissi*

$$F(x) = x\,, \qquad x \in X\,, \tag{7.10}$$

ammette soluzione unica.

Dimostrazione. Il procedimento che seguiremo è di tipo ricorsivo e si chiama *metodo delle approssimazioni successive.* Si parte da un punto iniziale qualunque $x_0 \in X$ e si costruisce la successione

$$x_{n+1} = F(x_n) \qquad \forall n \in \mathbf{N}\,. \tag{7.11}$$

Dimostriamo che la successione così costruita è di Cauchy. Per ogni $k \in \mathbf{N}$ si ha

$$\begin{aligned} d(x_{k+1}, x_k) &= d(F(x_k), F(x_{k-1})) \leqslant L d(x_k, x_{k-1}) = L d(F(x_{k-1}), F(x_{k-2})) \\ &\leqslant L^2 d(x_{k-1}, x_{k-2}) = L^2 d(F(x_{k-2}), F(x_{k-3})) \leqslant \ldots \leqslant L^k d(x_1, x_0)\,, \end{aligned}$$

quindi per ogni $n > m$ si ha

$$d(x_n, x_m) \leqslant \sum_{k=m+1}^{n} d(x_k, x_{k-1}) \leqslant d(x_1, x_0) \sum_{k=m+1}^{n} L^k = d(x_1, x_0) \frac{L^{m+1} - L^{n+1}}{1-L}\,.$$

Essendo infinitesima, la successione $(L^n) \subset \mathbf{R}$ è anche di Cauchy e quindi lo è in X anche la nostra successione (x_n). Per la completezza dello spazio esiste allora $x \in X$ tale che $x_n \to x$. Passando al limite nella (7.11) per $n \to \infty$ si deduce che tale limite x è soluzione della (7.10). Per l'unicità, se anche $x' \neq x$ fosse soluzione si otterrebbe

$$d(x, x') = d(F(x), F(x')) \leqslant L d(x, x') < d(x, x')\,.$$

□

Concludiamo con un esempio. L'algoritmo di Erone (2.8)

$$\begin{cases} a_0 = \alpha > 1 \\ a_{n+1} = \dfrac{1}{2}\left(a_n + \dfrac{\alpha}{a_n}\right) \qquad \forall n \in \mathbf{N} \end{cases}$$

per l'approssimazione di $\sqrt{a}$ rientra proprio nella situazione contemplata nel Teorema 7.21, infatti la funzione $F(x) = (x + \alpha/x)/2$ soddisfa

$$|F(x) - F(y)| \leqslant \frac{1}{2}\left(1 - \frac{\alpha}{xy}\right)|x - y| \leqslant \frac{1}{2}|x - y| \qquad \forall x, y \geqslant \sqrt{\alpha}$$

e quindi è una contrazione come applicazione da $[\sqrt{\alpha}, +\infty[$ in se stesso. Il punto fisso di F è $\sqrt{a}$, che in questo caso è anche il suo punto di minimo, come si vede facilmente risolvendo l'equazione

$$\frac{1}{2}\left(x + \frac{\alpha}{x}\right) = x\,.$$

Capitolo 8

Limiti e funzioni continue

Giunti a questo punto del corso, dopo aver manipolato tanti limiti per successioni e serie, non sarà difficile affrontare la teoria dei limiti anche per le funzioni. Riservando al corso di Analisi 2 il caso delle funzioni di più variabili, nella prima parte di questo capitolo tratteremo la teoria dei limiti per le funzioni reali di una variabile reale. Per agevolare la comprensione di questo argomento mi è sembrato opportuno classificare le varie proprietà con gli stessi criteri usati per le successioni e presentarle nello stesso ordine; tra le due teorie vi sono infatti molte analogie che consigliamo di tenere il più possibile presenti. C'è pure un teorema, diciamo di "collegamento", di portata molto generale, valido in tutti gli spazi metrici, che ci permette di trasferire alle funzioni, senza alcuna fatica, i risultati già ottenuti per le successioni. Uno strumento fondamentale per il calcolo dei limiti è la formula di Taylor che vedremo nel Cap. 9, strettamente correlata all'espansione di una funzione in serie di potenze. Per metterlo a frutto utilmente dobbiamo precisare e approfondire in questo capitolo che cosa s'intende per confronto di infinitesimi e di infiniti, un'idea che abbiamo già applicato, forse in modo non del tutto esplicito, alle successioni e alle serie. Rientra in questo argomento anche l'analisi del comportamento asintotico delle funzioni.

La seconda parte di questo capitolo sarà dedicato alle funzioni continue. Dopo averne dato varie definizioni equivalenti, anche in modo indipendente dal concetto di limite, ne analizzeremo le principali conseguenze, studieremo in particolare le proprietà fondamentali delle funzioni continue quando il dominio è un intervallo o un insieme compatto, infine studieremo in quali circostanze una funzione continua ammette inversa continua. La continuità di una funzione è la traduzione rigorosa del carattere graduale del modo di variare di una grandezza in funzione di un'altra, "piccole variazioni" della variabile indipendente influiscono per variazioni altrettanto piccole sulla variabile dipendente. Si può facilmente immaginare dunque la portata di questo concetto nelle varie Scienze. Tanto per fare un esempio, la posizione di un punto materiale P che si muove nello spazio è una funzione $P(t)$ della variabile temporale $t \in \mathbf{R}$. Assumere che essa sia continua equivale ad escludere che P possa saltare bruscamente da una posizione all'altra, per quanto veloce possa andare, impiegando invece un tempo finito.

8.1 Proprietà generali dei limiti delle funzioni

Ricordiamo che il derivato di un insieme $A \subset \mathbf{R}$ è l'insieme di tutti i suoi punti di accumulazione e si indica con $\mathscr{D}(A)$. Quando diremo che un dato punto appartiene a $\mathscr{D}(A)$ non ci sarà alcuna necessità di ritenere che quel punto appartenga anche ad A. Ricordiamo che se A non è limitato superiormente [inferiormente] il punto $+\infty$ $[-\infty]$ appartiene a $\mathscr{D}(A)$, infatti per le successioni, che sono funzioni definite su $\mathbf{N}$, abbiamo trattato solo il limite per $n \to \infty$. Indicheremo infine con $\mathscr{I}(x)$ la famiglia di tutti gli intorni del punto x.

Sia $f : A \to \mathbf{R}$ con $A \subset \mathbf{R}$. Le condizione essenziale per poter trattare il concetto di limite per $x \to x_0$ è $x_0 \in \mathscr{D}(A)$, altrimenti non ha nessun senso. A seconda che x_0 e il limite l della funzione siano finiti o infiniti, 4 casi in tutto a meno del segno dell'∞, la definizione di limite cambia forma, però ne daremo anche un'altra più generale che comprende tutti i casi.

Definizione 8.1 - *Diciamo che $l \in \overline{\mathbf{R}}$ è il **limite per** x **che tende a** $x_0 \in \overline{\mathbf{R}}$, oppure che f **tende a** l **per** x **che tende a** x_0, e si scrive*

$$\lim_{x \to x_0} f(x) = l \,,$$

se
$x_0\,, l \in \mathbf{R}$:

$$\forall \varepsilon > 0 \; \exists \delta > 0 \; : \; \forall x \in A \,, \; x \neq x_0 \,, \quad |x - x_0| < \delta \Rightarrow |f(x) - l| < \varepsilon \,,$$

$x_0 = +\infty\,, l \in \mathbf{R}$:

$$\forall \varepsilon > 0 \; \exists a \in \mathbf{R} \; : \; \forall x \in A \quad x > a \Rightarrow |f(x) - l| < \varepsilon \,,$$

$x_0 \in \mathbf{R}\,, l = +\infty$:

$$\forall M \in \mathbf{R} \; \exists \delta > 0 \; : \; \forall x \in A \,, \; x \neq x_0 \,, \quad |x - x_0| < \delta \Rightarrow f(x) > M \,,$$

$x_0 = +\infty\,, l = +\infty$:

$$\forall M \in \mathbf{R} \; \exists a \in \mathbf{R} \; : \; \forall x \in A \quad x > a \Rightarrow f(x) > M \,.$$

*Se il limite non esiste f viene detta **indeterminata** in x_0.*

Quelle con $-\infty$ al posto di $+\infty$ sono analoghe e comportano ovvi cambiamenti nelle disugualianze. Se $l \in \mathbf{R}$ f è detta *convergente a* l per $x \to x_0$, altrimenti, se $l = \pm\infty$, *divergente, positivamente* o negativamente.

Commentiamo il I caso. La condizione $|f(x) - l| < \varepsilon$ è una disequazione; se vogliamo che la distanza tra i valori $f(x)$ della funzione e il numero l sia più piccola di ε bisogna scegliere $x \in A$ a distanza inferiore ad una quantità $\delta > 0$ opportuna che dipende da ε. Si tenga ben presente che nessuna relazione intercorre tra $f(x_0)$ e l, tanto meno si richiede che siano uguali. Innanzitutto f potrebbe non essere definita in x_0, non è detto cioè che x_0 appartenga ad A, ma anche se fosse questo il caso sarebbe del tutto irrilevante il valore della funzione in questo punto, valore che non interviene affatto nella definizione. Ribadiamo infine che se x_0 non è di accumulazione per A la definizione non ha più senso perché potrebbe accadere che per un $\delta > 0$ abbastanza piccolo non vi sia nessun punto $x \in A$ a distanza da x_0 minore di δ.

Analogamente a quanto già osservato in relazione ai limiti delle successioni, sempre nel I caso, se, fissato $\varepsilon > 0$, vale la definizione per un certo $\delta > 0$ allora vale con lo stesso ε per ogni $\delta' < \delta$ e continua a valere con lo stesso δ per ogni $\varepsilon' > \varepsilon$, stessi commenti negli altri casi. Per questo motivo si usa dire *per ogni ε piccolo a piacere esiste δ ....* La definizione di limite descrive il comportamento *definitivo* di una funzione. Ciò che nel caso delle successioni avviene da un certo indice in poi, qui avviene in un intorno, nelle vicinanze, di x_0: come una successione conserva lo stesso limite se ne vengono modificati un numero finito di termini, una funzione conserva lo stesso limite se viene modificata solo al di fuori di un intorno di x_0. In modo equivalente, se f e g sono due funzioni tali che $f(x) = g(x)$ per ogni $x \neq x_0$ in un intorno di x_0 se una di esse ha limite $l \in \overline{\mathbf{R}}$ per $x \to x_0$ anche l'altra ha lo stesso limite per $x \to x_0$.

Poiché i parametri ε e δ definiscono gli intorni fondamentali $]l-\varepsilon, l+\varepsilon[$ e $]x_0-\delta, x_0+\delta[$, cui devono appartenere $f(x)$ e x rispettivamente, la definizione può essere data in modo del tutto equivalente usando gli intorni. Ciò significa che la definizione di limite può essere generalizzata anche a funzioni tra spazi metrici, o topologici, qualsiasi. Nelle stese condizioni di prima, dunque, la definizione va espressa in questo modo

$$\lim_{x\to x_0} f(x) = l \Leftrightarrow \forall V \in \mathscr{I}(l)\ \exists U \in \mathscr{I}(x_0)\ :\ \forall x \in A\,,\ x \neq x_0\,, \quad x \in U \Rightarrow f(x) \in V\,,$$

nella parte finale è come dire che $f(A \cap U - \{x_0\}) \subset V$. E la seguente è la versione in termini di distanze

$$\lim_{x\to x_0} f(x) = l \Leftrightarrow \forall \varepsilon > 0\ \exists \delta > 0\ :\ \forall x \in A\,,\ x \neq x_0\,, \quad d(x, x_0) < \delta \Rightarrow d(f(x), l) < \varepsilon\,.$$

Esempi

8.1 *Se f assume su A, o in un intorno di x_0, valore costante c allora*

$$\lim_{x\to x_0} f(x) = c$$

perché $|f(x)-c| = 0 < \varepsilon$ per ogni $\varepsilon > 0$. Ad esempio la funzione $f(x) = 1$ su $A = [0,1[\cup]1,2]$ e il suo prolungamento

$$\bar{f}(x) = \begin{cases} f(x) & \text{se } x \in A \\ 2 & \text{se } x = 1 \end{cases}$$

hanno entrambe limite 1 per $x \to 1$.

Risolviamo il seguente esercizio come esempio di applicazione della definizione di limite.

Esercizio 8.1 - *Applicando la definizione di limite nelle versioni opportune, verificare che*

$$\lim_{x\to 3}(2x^2-x+5) = 20\,, \quad \lim_{x\to+\infty}\frac{3x-2}{4x+1} = \frac{3}{4}\,, \quad \lim_{x\to 0}\frac{1}{x^2} = +\infty\,, \quad \lim_{x\to+\infty}\frac{\sqrt[3]{x}}{\sqrt[4]{2x+1}} = +\infty\,.$$

(I) La disequazione $|(2x^2-x+5)-20| < \varepsilon$ equivale al sistema

$$\begin{cases} 2x^2-x-15-\varepsilon < 0 \\ 2x^2-x-15+\varepsilon > 0 \end{cases}$$

di cui sono soluzioni i numeri reali x tali che

$$3-\delta_1 = \frac{1+\sqrt{121-8\varepsilon}}{4} < x < \frac{1+\sqrt{121+8\varepsilon}}{4} = 3+\delta_2$$

in cui possiamo supporre $\varepsilon < 121/8$. I valori δ_1 e δ_2 sono

$$\delta_1 = \frac{2\varepsilon}{11+\sqrt{121-8\varepsilon}} \quad \text{e} \quad \delta_2 = \frac{2\varepsilon}{11+\sqrt{121+8\varepsilon}}\,,$$

quindi la definizione di limite è soddisfatta con $\delta = \min\{\delta_1, \delta_2\} = \delta_2$.

(II) Vediamo se la disequazione

$$\left|\frac{3x-2}{4x+1} - \frac{3}{4}\right| < \varepsilon$$

è soddisfatta per tutti i numeri reali x che superano un certo a opportuni dipendente da ε. La disequazione equivale a

$$\frac{11}{4|4x+1|} < \varepsilon$$

in cui possiamo supporre $\varepsilon < 11/4$ (ma anche $\varepsilon < 2$), da cui

$$|4x+1| > \frac{11}{4\varepsilon}$$

che è soddisfatta per $x > 3/\varepsilon$.

(III) La disequazione

$$\frac{1}{x^2} > M\,, \quad x \neq 0\,,$$

dove si può supporre $M > 0$, è verificata per $|x| < \delta = 1/\sqrt{M}$.

(IV) La disequazione

$$\frac{\sqrt[3]{x}}{\sqrt[4]{2x+1}} > M\,,$$

è equivalente a

$$x^4 > (2x+1)^3 M^{12}\,,$$

la quale, essendo $x > 0$ e $2x+1 < 3x$, è soddisfatta per $x^4 > 27x^3M^{12}$, cioè per $x > 27M^{12}$.

Le funzioni elementari e quelle che si ottengono combinandole tra loro con operazioni e composizioni hanno in genere limite per $x \to x_0$, con x_0 nel dominio naturale, e tale limite vale proprio $f(x_0)$, sono cioè *continue*, proprietà che vedremo più avanti. Ciò dipende dal fatto che sono localmente lipschitziane. Infatti se $f : I \to \mathbf{R}$, $x_0 \in I$ e

$$|f(x) - f(x_0)| \leqslant k|x - x_0|$$

per qualche $k \geqslant 0$ per ogni $\varepsilon > 0$ basta scegliere $\delta = \varepsilon/k$ e se $|x - x_0| < \delta$ si ha

$$|f(x) - f(x_0)| < k\delta = \varepsilon\,.$$

Stesse conclusioni per le funzioni hölderiane con esponente $\alpha < 1$, basta scegliere $\delta = (\varepsilon/k)^{1/\alpha}$. Si vede in entrambi i casi che δ dipende solo da ε e non da x_0.

Sono due i motivi per cui può non valere la definizione di limite: o il limite non è l oppure non esiste (si provi a scrivere la negazione nei due casi). Come si è detto per le successioni, con le varianti del caso, se una funzione definita su A ammette limite l per $x \to x_0$ allora deve avere lo stesso limite la sua restrizione ad ogni insieme $A' \subset A$ che abbia x_0 come punto di accumulazione. Pertanto se troviamo almeno due sottoinsiemi A_1 e A_2 sui quali i limiti esistono, ma sono distinti, si può concludere che la funzione su A non ha limite. Un caso abbastanza comune è quello in cui differiscono i limiti destro e sinistro. Questi due limiti li possiamo definire così: posto $A^- = \{x \in A \mid x < x_0\}$ e $A^+ = \{x \in A \mid x > x_0\}$,

$$\lim_{x \to x_0^-} f(x) = \lim_{x \to x_0} f_{|A^-}(x) \qquad \text{e} \qquad \lim_{x \to x_0^+} f(x) = \lim_{x \to x_0} f_{|A^+}(x)\,.$$

Se però la restrizione ad ogni sottoinsieme di A ammette sempre lo stesso limite allora la funzione su A ammette quel valore come limite.

8.2 *La funzione*

$$f(x) = \begin{cases} 1 & \textit{se } 0 \leqslant x < 1 \\ 2 & \textit{se } 1 \leqslant x \leqslant 2 \textit{ oppure } 1 < x < 2 \end{cases}$$

non ha limite per $x \to 1$ perché i limiti destro e sinistro esistono, ma sono diversi.

Questa funzione presenta un *salto* in corrispondenza del punto $x_0 = 1$, dove per salto in x_0 si intende la differenza

$$[\![f]\!]_{x_0} = \lim_{x \to x_0^+} f(x) - \lim_{x \to x_0^-} f(x)\,.$$

8.3 *Le funzioni* $1/x$, $\operatorname{arctg} 1/x$, $\mathrm{e}^{1/x}$, $\operatorname{sign} x$, $\operatorname{cotg} x$ *non hanno limite per* $x \to 0$ *perché hanno limite destro e sinistro differenti.*

8.4 *La funzione* $\operatorname{sen} x$ *su* $\mathbf{R}$ *non ha limite per* $x \to +\infty$. *Infatti*

$$\lim_{n \to +\infty} \operatorname{sen}\left(\frac{\pi}{2} + 2n\pi\right) = 1 \qquad e \qquad \lim_{n \to +\infty} \operatorname{sen}\left(\frac{3\pi}{2} + 2n\pi\right) = -1\,.$$

Più in generale una funzione f periodica su $\mathbf{R}$ e non costante non ha limite per $x \to +\infty$. Siano $x_0, y_0 \in \mathbf{R}$ tali che $f(x_0) \neq f(y_0)$ e $T > 0$ il periodo di f. Poiché $f(x_0 + nT) = f(x_0)$ e $f(y_0 + nT) = f(y_0)$ per ogni $n \in \mathbf{N}$, anche i limiti di queste successioni costanti valgono $f(x_0)$ e $f(y_0)$, ma essendo tra loro diversi la f non ha limite.

Esercizio 8.2 - *Dimostrare che la funzione* $f(x) = \operatorname{sen} 1/x$, $x > 0$, *non ha limite per* $x \to 0$.

8.5 *Potenze del tipo* x^α, *con* $x \geqslant 0$ *e* $\alpha > 0$, *divergono positivamente per* $x \to +\infty$ *dato che* $x^\alpha > M$ *per* $x > M^{1/\alpha}$. *Di conseguenza* $|x|^\alpha$ *tende a* $+\infty$ *sia per* $x \to +\infty$ *che per* $x \to -\infty$. *Se l'esponente è un numero naturale* n *a seconda che sia pari o dispari si ha* $x^n \to +\infty$ *e* $x^n \to -\infty$ *per* $x \to -\infty$.

Sono molte le analogie con la teoria dei limiti per le successioni, si può definire una condizione di Cauchy in x_0 anche per le funzioni e dedurne l'equivalenza con l'esistenza del limite finito, vale il teorema di unicità, valgono le stesse proprietà algebriche, di confronto e di monotonia, per cui basta modificare qualche notazione e adattare al caso delle funzioni il linguaggio usato per le successioni e ritroviamo gli stessi risultati con poca fatica. A titolo di esempio possiamo trattare qualche proprietà in maniera indipendente, comunque il seguente teorema ci permette in pratica di convertire automaticamente quasi tutte le proprietà dei limiti per le successioni in analoghe proprietà per le funzioni e di evitare gran parte delle dimostrazioni.

Teorema 8.2 - *Una funzione* $f : A \to \mathbf{R}$ *ammette limite* $l \in \overline{\mathbf{R}}$ *per* $x \to x_0 \in \mathscr{D}(A)$ *se e solo se per ogni successione* $(x_n) \subset A$ *tale che* $x_n \neq x_0$ *per ogni* $n \in \mathbf{N}$ *e* $x_n \to x_0$ *si ha*

$$\lim_{n \to \infty} f(x_n) = l\,.$$

Dimostrazione. Supponiamo che $\lim_{x \to x_0} f(x) = l$ e sia $x_n \to x_0$ come nell'enunciato. Scelti ε e δ come impone la definizione di limite, da un certo indice ν in poi si ha $|x - x_0| < \delta$, ne segue che $|f(x_n) - l| < \varepsilon$ per ogni $n > \nu$.

Viceversa, supponiamo $f(x_n) \to l$ per ogni successione (x_n) che tende a x_0 e, per assurdo, che f non abbia limite l per $x \to x_0$. Neghiamo allora la definizione di limite

$$\exists \varepsilon > 0 \;:\; \forall \delta > 0\; \exists x \in A - \{x_0\} \;:\; |x - x_0| < \delta \quad \text{e} \quad |f(x) - l| \geqslant \varepsilon\,.$$

Se δ è arbitrario, possiamo scegliere $\delta = 1/n$, $n \in \mathbf{N}$, a cui corrisponde un punto $x_n \in A - \{x_0\}$ tale che $|x_n - x_0| < 1/n$ e $|f(x_n) - l| \geqslant \varepsilon$. Passando al limite per $n \to \infty$, $x_n \to x_0$, ma $f(x_n) \nrightarrow l$, in contraddizione con l'ipotesi.

□

Col Teorema 8.2 i limiti notevoli incontrati trattando le successioni diventano adesso limiti notevoli per le funzioni. Ad esempio, avendo già dimostrato che

$$\lim_{n\to\infty} \frac{\operatorname{sen} x_n}{x_n} = 1 \quad \forall (x_n) \subset \mathbf{R} \; : \; x_n \to 0 \,,$$

adesso possiamo affermare che

$$\lim_{x\to 0} \frac{\operatorname{sen} x}{x} = 1$$

con tutti i limiti notevoli trigonometrici legati a questo. Allo stesso modo

$$\lim_{x\to 0} (1+x)^{1/x} = e \,, \quad \lim_{x\to 0} \frac{e^x - 1}{x} = 1 \,, \quad \lim_{x\to 0} \frac{\log(1+x)}{x} = 1 \,.$$

Inoltre se $a > 1$ e $\alpha > 0$ $a^x/x^\alpha \to +\infty$ per $x \to +\infty$ e se $a > 0$, $a \neq 1$ e $\alpha > 0$ $x^\alpha \log_a x \to 0$ per $x \to 0^+$. Vediamo qualche proprietà senza ricorrere alle successioni.

Teorema 8.3 (Unicità del limite) - *Se esiste in $\overline{\mathbf{R}}$ il limite l di una funzione $f : A \to \mathbf{R}$ per $x \to x_0 \in \overline{\mathbf{R}}$, tale limite è unico.*

Dimostrazione. Supponiamo che

$$\lim_{x\to x_0} f(x) = l_1 \quad \text{e} \quad \lim_{x\to x_0} f(x) = l_2 \,, \quad l_1 \neq l_2 \,.$$

Se $l_1 = +\infty$ l_2 non può essere finito altrimenti f sarebbe limitata, tanto meno può essere $l_2 = -\infty$, quindi anche $l_2 = +\infty$. Se sono entrambi finiti, supponiamo $l_1 < l_2$. Scelto $\varepsilon = (l_2 - l_1)/2$ in modo che

$$l_1 + \varepsilon = \frac{l_1 + l_2}{2} = l_2 - \varepsilon \,,$$

esistono $\delta_1, \delta_2 > 0$ tali che per ogni $x \in A$, $x \neq x_0$, si ha

$$|x - x_0| < \delta_1 \Rightarrow |f(x) - l_1| < \varepsilon \qquad \text{e} \qquad |x - x_0| < \delta_2 \Rightarrow |f(x) - l_2| < \varepsilon \,.$$

Ma se $|x - x_0| < \min\{\delta_1, \delta_2\}$ devono valere le due condizioni insieme, quindi

$$f(x) < l_1 + \varepsilon = \frac{l_1 + l_2}{2} = l_2 - \varepsilon < f(x)$$

che è assurdo.

□

L'unicità del limite è una proprietà di carattere generale, nel senso che vale anche per le funzioni tra spazi metrici. In tutti gli spazi metrici vale banalmente, infatti, l'assioma di separazione di Hausdorff (v. l'Esercizio 7.3). Se fosse $l_1 \neq l_2$, scelto $\varepsilon > 0$ in modo che $B_\varepsilon(l_1) \cap B_\varepsilon(l_2) = \emptyset$, ragionando come sopra si otterrebbe la stessa conclusione assurda $f(x) \in B_\varepsilon(l_1) \cap B_\varepsilon(l_2) = \emptyset$ per gli x di un intorno di x_0.

Definizione 8.4 - *Sia $x_0 \in \mathscr{D}(A)$. Diciamo che $f : A \to \mathbf{R}$ soddisfa la **condizione di Cauchy** nel punto x_0 se*

$$\forall \varepsilon > 0 \; \exists U \in \mathscr{I}(x_0) \; : \; \forall x, y \in A \cap U - \{x_0\} \quad |f(x) - f(y)| < \varepsilon \,. \tag{8.1}$$

L'uso di U come intorno generico al posto di $]x_0 - \delta, x_0 + \delta[$ serve a comprendere il caso $x_0 = +\infty$.

Proposizione 8.5 - *Se $f : A \to \mathbf{R}$ converge a l nel punto $x_0 \in \mathscr{D}(A)$ allora è di Cauchy in x_0.*

Dimostrazione. Basta osservare che

$$|f(x) - f(y)| \leqslant |f(x) - l| + |f(y) - l| < 2\varepsilon$$

per ogni $x, y \in A \cap U - \{x_0\}$.

□

Teorema 8.6 - *Se $f : A \to \mathbf{R}$ è di Cauchy nel punto $x_0 \in \mathscr{D}(A)$ allora esiste $U \in \mathscr{I}(x_0)$ dove f è limitata (ciò è vero se, in particolare, f converge in x_0). Se f diverge in x_0 allora non è limitata su nessun intorno di x_0 (quindi non è di Cauchy).*

Dimostrazione. Scegliamo $U \in \mathscr{I}(x_0)$ in modo che $|f(x) - f(y)| < 1$ per ogni $x, y \in A \cap U - \{x_0\}$. Fissando uno dei due punti, ad esempio y, e facendo variare l'altro nell'intorno si ottiene

$$\big||f(x)| - |f(y)|\big| \leqslant |f(x) - f(y)| < 1$$

da cui

$$|f(x)| < 1 + |f(y)| \quad \forall x \in A \cap U - \{x_0\}\,,$$

quindi f è limitata su $A \cap U$. La seconda parte della tesi è banale.

□

Un'altra applicazione del Teorema 8.2 è l'equivalenza tra la condizione di Cauchy e l'esistenza del limite finito. La sua validità dipende dal fatto che il codominio, $\mathbf{R}$ nel nostro caso, è uno spazio metrico completo.

Teorema 8.7 - *Se $f : A \to \mathbf{R}$ è di Cauchy in $x_0 \in \mathscr{D}(A)$ allora esiste $l \in \mathbf{R}$ tale che*

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

Dimostrazione. Siano $\varepsilon > 0$ e $\delta > 0$ come nella (8.1). Se $(x_n) \subset A - \{x_0\}$ e $x_n \to x_0$ esiste $\nu \in \mathbf{N}$ tale che $|x_n - x_0| < \delta$ per ogni $n > \nu$ e per la (8.1)

$$|f(x_n) - f(x_m)| < \varepsilon \qquad \forall n, m > \nu\,.$$

Poiché $\mathbf{R}$ è completo esiste un numero reale l tale che $f(x_n) \to l$. Verifichiamo che il limite l della f non dipende dalla scelta della successione. Se consideriamo un'altra successione (y_n) dello stesso tipo e supponiamo $f(y_n) \to l'$ si ha definitivamente $|f(x_n) - f(y_n)| < \varepsilon$ e passando al limite

$$\lim_{n \to \infty} |f(x_n) - f(y_n)| = |l - l'| \leqslant \varepsilon\,.$$

Quindi $l = l'$ per l'arbitrarietà di ε.

□

Esercizio 8.3 - *Dimostrare che se A è un sottoinsieme infinito di $\mathbf{R}$ e $f : A \to \mathbf{R}$ è limitata allora esiste una successione di punti tutti distinti di A ed un numero reale l tale che $f(x_n) \to l$.*

Talvolta è comodo effettuare un cambio di variabile per semplificare l'espressione che definisce la funzione in esame e ricondurla a un caso di limite più facilmente riconoscibile. Si debba ad esempio calcolare

$$\lim_{x \to +\infty} x^2 (e^{\frac{1}{x}} - e^{\frac{1}{x+1}})\,.$$

Posto $y = 1/x$, il limite diventa

$$\lim_{y \to 0} \frac{e^y - e^{\frac{y}{y+1}}}{y^2} = \lim_{y \to 0} \frac{e^{\frac{y}{y+1}}}{y+1} \cdot \lim_{y \to 0} \frac{e^{\frac{y^2}{y+1}} - 1}{y^2/(y+1)} = 1\,.$$

La procedura si basa sul seguente teorema, anch'esso di carattere generale.

Teorema 8.8 (del cambio di variabile) - *Siano $A, B \subset \mathbf{R}$, $x_0 \in \mathscr{D}(A)$, $y_0 \in \mathscr{D}(B)$ e $f : A \to \mathbf{R}$, $g : B \to \mathbf{R}$ due funzioni tali che $f(A) \subset B$ e $f(x) \neq y_0$ per ogni $x \in A$. Se*

$$\lim_{x\to x_0} f(x) = y_0 \quad e \quad \lim_{y\to y_0} g(y) = l$$

allora esiste il limite di $g \circ f$ per $x \to x_0$ e

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

Dimostrazione. Per ipotesi ad ogni $\varepsilon > 0$ corrisponde un $\delta > 0$ tale che per ogni $y \in B \cap I_\delta(y_0) - \{y_0\}$ si ha $|g(y) - l| < \varepsilon$. D'altra parte a δ corrisponde $\delta' > 0$ tale che per ogni $x \in A \cap I_{\delta'}(x_0) - \{x_0\}$ si ha $|f(x) - y_0| < \varepsilon$. Ora, parte dei punti y che stanno in $B \cap I_\delta(y_0) - \{y_0\}$ sono immagini $f(x)$ dei punti di $x \in A \cap I_{\delta'}(x_0) - \{x_0\}$. In definitiva, tutti i punti $x \in A \cap I_{\delta'}(x_0) - \{x_0\}$ soddisfano la condizione

$$|g(f(x)) - l| < \varepsilon$$

che è la tesi.

□

8.2 Confronto, algebra dei limiti, monotonia

Le proprietà dei limiti di questo paragrafo sono analoghe a quelle trattate in § 5.3 e § 5.5 e dato che dimostrare nei dettagli anche queste sarebbe quasi una ripetizione di cose già dette, ci limitiamo ad enunciarle con al più qualche commento o cenno di dimostrazione. Comunque si può sempre ricorrere al Teorema 8.2.

In alcuni risultati sul confronto tra funzioni e sulla permanenza del segno si deducono informazioni sui limiti a partire da informazioni sulle funzioni, in altri il viceversa. Supponiamo una volta per tutte $x_0 \in \mathscr{D}(A)$, $A \subset \mathbf{R}$, e che le funzioni di cui vogliamo parlare siano definite su A a valori in $\mathbf{R}$.

Teorema 8.9 (confronto I) - *Se $f(x) \leqslant g(x)$ per ogni x in un intorno U di x_0, $f(x) \to l_1$ e $g(x) \to l_2$ per $x \to x_0$ allora $l_1 \leqslant l_2$.*

Dimostrazione. Basta osservare che in un intorno di x_0, eventualmente più piccolo di U, si ha

$$l_1 - \varepsilon < f(x) \leqslant g(x) < l_2 + \varepsilon\,.$$

Se invece si preferisce ragionare con le successioni e utilizzare il Teorema 8.2 bisogna osservare che per ogni successione $(x_n) \subset A - \{x_0\}$ tale che $x_n \to x_0$ si ha $f(x_n) \to l_1$ e $g(x_n) \to l_2$. Dal confronto $f(x_n) \leqslant g(x_n)$ e dallo stesso teorema già dimostrato per le successioni segue la tesi.

□

Un altro modo di usare le successioni è quello di confrontare la funzione della variabile x con la stessa funzione calcolata nella variabile intera $[x]$, così diventa una successione. Ad esempio, dal confronto

$$\frac{a^x}{x^\alpha} \geqslant \frac{a^{[x]}}{(1+[x])^\alpha}$$

con $a > 1$ e $\alpha > 0$ e da quanto sappiamo sulla successione a^n/n^α si deduce che la funzione a sinistra diverge per $x \to +\infty$.

Per un polinomio $P(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$, con $a_n > 0$, ricordiamo la stima (3.10)

$$P(x) \geqslant \frac{a_n}{2}x^n \qquad \forall x \geqslant \max\left\{1, \frac{2Mn}{a_n}\right\}\,,$$

da cui segue che $P(x)$ diverge positivamente per $x \to +\infty$ e, con qualche considerazione in più, che diverge positivamente anche per $x \to -\infty$ se n è pari e negativamente se n è dispari.

Corollario 8.10 (permanenza del segno I) - *Se $f(x) \geqslant 0$ in un intorno di x_0 e $f(x) \to l$ per $x \to x_0$ allora $l \geqslant 0$.*

Teorema 8.11 (permanenza del segno II) - *Se $f(x) \to l$ per $x \to x_0$ e $l > 0$ allora $f(x) > 0$ in un intorno di x_0*

Dimostrazione. Si sceglie $\varepsilon < l$, ad esempio $\varepsilon = l/2$, e si osserva che per la definizione di limite deve essere

$$f(x) > l - \varepsilon = l/2 > 0$$

in un intorno di x_0. □

Teorema 8.12 (confronto II o teorema dei carabinieri) - *Se $f(x) \leqslant g(x) \leqslant h(x)$ per ogni x in un intorno U di x_0, $f(x) \to l$ e $h(x) \to l$ per $x \to x_0$ allora $g(x) \to l$.*

Dimostrazione. Basta osservare che, fissato $\varepsilon > 0$, in un intorno di x_0 si ha

$$l - \varepsilon < f(x) \leqslant g(x) \leqslant h(x) < l + \varepsilon\,.$$

□

Ad esempio dalle stime

$$\frac{x-1}{x+1} \leqslant \frac{x - \operatorname{sen} x}{x + \operatorname{sen} x} \leqslant \frac{x+1}{x-1} \qquad \forall x > 17$$

si deduce che

$$\lim_{x \to +\infty} \frac{x - \operatorname{sen} x}{x + \operatorname{sen} x} = 1\,.$$

Passiamo alle proprietà algebriche.

Teorema 8.13 - *L'insieme delle funzioni convergenti in x_0 è uno spazio vettoriale e un'algebra.*

Dimostrazione. Dal confronto

$$|f(x) + g(x) - (l_1 + l_2)| \leqslant |f(x) - l_1| + |g(x) - l_2|$$

segue che se $f(x) \to l_1$ e $g(x) \to l_2$ per $x \to x_0$ allora $f(x) + g(x) \to l_1 + l_2$ per $x \to x_0$. Inoltre per ogni $\lambda \in \mathbf{R}$, essendo $|f(x) - l| < \varepsilon/|\lambda|$ in un intorno di x_0, si ha $|\lambda f(x) - \lambda l| < \varepsilon$.

Se f è limitata in un intorno di x_0, $|f(x)| \leqslant M$, e h è infinitesima per $x \to x_0$ allora $f(x)h(x) \to 0$ per $x \to x_0$, basta osservare che

$$|f(x)h(x)| \leqslant M|h(x)|\,.$$

Ad esempio $x \operatorname{sen} 1/x \to 0$ per $x \to 0$. Da questa e dalla precedente segue che se $f(x) \to l_1$ e $g(x) \to l_2$ allora $f(x)g(x) \to l_1 l_2$. Infine, se $f(x) \to l \neq 0$ allora $1/f(x) \to 1/l$, infatti $|f(x)| > |l|/2$ definitivamente, quindi

$$\left|\frac{1}{f(x)} - \frac{1}{l}\right| = \frac{|f(x) - l|}{|f(x)l|} < \frac{2|f(x) - l|}{l^2}\,.$$

Da questa e dal prodotto segue che $f(x)/g(x) \to l_1/l_2$ se $f \to l_1$ e $g \to l_2 \neq 0$.

□

Sui casi di indeterminazione non c'è nulla da aggiungere a quanto già detto sulle successioni.

Anche per le funzioni monotone valgono dei risultati di esistenza del limite, solo che rispetto alle successioni si complica un po' la casistica a causa del fatto che il punto di accumulazione x_0 può essere reale, non solo $+\infty$ o $-\infty$, e questo comporta che vi siano punti del dominio della funzione sia a sinistra, sia a destra di x_0, ma una volta chiarita la situazione il ragionamento è identico.

Teorema 8.14 - *Sia $f : A \to \mathbf{R}$ una funzione crescente e $x_0 \in \mathscr{D}(A)$. Se esistono in A e in un intorno di x_0 sia dei punti $x < x_0$, sia dei punti $x > x_0$, allora esistono finiti i limiti di f per $x \to x_0^-$ e per $x \to x_0^+$; se x_0 è il minimo punto di accumulazione per A esiste il limite per $x \to x_0^+$ e può essere finito oppure $-\infty$; se x_0 è il massimo punto di accumulazione per A esiste il limite per $x \to x_0^-$ e può essere finito oppure $+\infty$. In ogni caso*

$$\lim_{x\to x_0^-} f(x) = \sup_{x<x_0} f(x) \qquad e \qquad \lim_{x\to x_0^+} f(x) = \inf_{x>x_0} f(x)\,.$$

Dimostrazione. Supponiamo che A contenga punti di un intorno di x_0 sia a sinistra, sia a destra, di x_0. Allora $f(x) \leqslant f(x')$ per ogni $x < x_0 < x'$, quindi f è limitata superiormente su $A^- = A \cap \{x \in \mathbf{R} \mid x < x_0\}$. Se $l^- = \sup_{A^-} f(x)$ per ogni $\varepsilon > 0$ esiste $\alpha \in A^-$ tale che $f(\alpha) > l^- - \varepsilon$ e siccome f è crescente, posto $\delta = x_0 - \alpha$, per ogni $x \in A \cap]x_0 - \delta, x_0[$ si ha

$$l^- - \varepsilon < f(\alpha) \leqslant f(x) < l^-$$

che è la tesi. Si ragiona nello stesso modo per il limite destro. Se infine siamo su una estremità del dominio, nel caso che f sia limitata non c'è nulla da cambiare, altrimenti basta usare le caratterizzazioni relative a questo caso degli estremi inferiore di f (nell'estremità di sinistra) e superiore (nell'estremità di destra).

□

Ovviamente vale un teorema analogo se f è decrescente.

Esercizio 8.4 *Dimostrare che una funzione crescente $f : I \to I$, con I intervallo, ammette un punto fisso.*

Un'applicazione importante del Teorema 8.14 riguarda le funzioni convesse. Ricordiamo che $f : I \to \mathbf{R}$ è convessa se e solo se per ogni $x_0 \in I$ il rapporto incrementale

$$R(x) = \frac{f(x) - f(x_0)}{x - x_0}$$

è crescente su $I - \{x_0\}$. Se x_0 è interno ad I per il Teorema 8.14 esistono finiti i limiti

$$f'_-(x_0) = \lim_{x\to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \qquad \text{e} \qquad f'_+(x_0) = \lim_{x\to x_0^+} \frac{f(x) - f(x_0)}{x - x_0}\,,$$

detti rispettivamente *derivata sinistra* e *derivata destra* di f nel punto x_0. Inoltre

$$\frac{f(x) - f(x_0)}{x - x_0} \leqslant f'_-(x_0) \leqslant f'_+(x_0) \leqslant \frac{f(x') - f(x_0)}{x' - x_0} \qquad \forall x, x' \in I \,:\, x < x_0 < x'$$

e quindi

$$\frac{f(x) - f(x_0)}{x - x_0} \leqslant m \leqslant \frac{f(x') - f(x_0)}{x' - x_0} \qquad \forall x, x' \in I \,:\, x < x_0 < x'$$

per ogni $m \in \mathbf{R}$ tale che $f'_-(x_0) \leqslant m \leqslant f'_+(x_0)$. Moltiplicando per $x - x_0$ si ottiene un'altra caratterizzazione delle funzioni convesse:

ogni funzione convessa ammette derivata sinistra e derivata destra in ogni punto interno all'intervallo in cui è definita (una sola delle due negli eventuali estremi) e

$$(8.2) \qquad f(x) \geqslant f(x_0) + m(x - x_0) \qquad \forall x, x_0 \in I \text{ e } \forall m \in [f'_-(x_0), f'_+(x_0)]\,.$$

Ogni retta $y = f(x_0)+m(x-x_0)$ con m che soddisfa la (8.2) si chiama *retta d'appoggio*. Il significato geometrico è evidente, f è convessa se e solo se per ogni punto del grafico esiste una stella di rette passanti per quel punto, al di sopra delle quali sta tutto il grafico di f. I loro coefficienti angolari stanno tra quello della tangente sinistra e quello della tangente destra. Se il grafico presenta un punto angoloso, uno spigolo, in x_0, per cui $f'_-(x_0) < f'_+(x_0)$, le rette d'appoggio sono infinite, se invece è abbastanza liscio da avere retta tangente questa diventa l'unica retta d'appoggio e $f'_-(x_0) = f'_+(x_0)$. Il valore comune delle due derivate laterali sarà la *derivata* di f, argomento che tratteremo nel Cap. 9.

Esercizio 8.5 - *Dimostrare che se $f : I \to \mathbf{R}$ è convessa allora $x_0 \in I$ è di minimo per f se e solo se $f'_-(x_0) \leqslant 0 \leqslant f'_+(x_0)$ se x_0 è interno, $f'_+(x_0) \geqslant 0$ se x_0 è l'estremità di sinistra e $f'_-(x_0) \leqslant 0$ se x_0 è l'estremità di destra.*

Esercizio 8.6 - *Dimostrare che se $f : \mathbf{R} \to \mathbf{R}$ è convessa non costante allora soddisfa almeno una delle condizioni*

$$\lim_{x\to-\infty} f(x) = +\infty \qquad \textit{oppure} \qquad \lim_{x\to+\infty} f(x) = +\infty$$

e valgono entrambe se e solo se f ha minimo e l'intervallo formato dai punti di minimo è limitato.

Esercizio 8.7 - *Trovare le rette d'appoggio agli estremi per le funzioni convesse*

$$f(x) = -\sqrt{x}\,,\ x \in [0,1] \qquad e \qquad f(x) = \begin{cases} x^2 & se\ |x| < 1 \\ 2 & se\ |x| = 1\,. \end{cases}$$

8.3 Punti limite, minimo e massimo limite

Come abbiamo visto per le successioni nel § 5.8, anche per le funzioni è possibile definire l'insieme dei punti limite e, in particolare, il minimo limite e il massimo limite.

Siano, al solito, $f : A \to \mathbf{R}$ e $x_0 \in \mathscr{D}(A)$.

Definizione 8.15 - *Un numero reale l è un **punto limite** per f per $x \to x_0$ se*

$$\forall \varepsilon > 0 \text{ e } \forall \delta > 0\ \exists x \in A \cap I_\delta(x_0) - \{x_0\} \ :\ |f(x) - l| < \varepsilon\,,$$

$l = +\infty$ $[-\infty]$ è un punto limite per f per $x \to x_0$ se

$$\forall M \in \mathbf{R} \text{ e } \forall \delta > 0\ \exists x \in A \cap I_\delta(x_0) - \{x_0\} \ :\ f(x) > M\ [f(x) < M]$$

che è come dire che f è non limitata superiormente [inferiormente] intorno a x_0.

Esercizio 8.8 - *Dimostrare che f ammette $l \in \overline{\mathbf{R}}$ come punto limite se e solo se esiste un insieme $A' \subset A$ che ammette x_0 come punto di accumulazione, ad esempio una successione che tende a x_0, tale che*

$$\lim_{x\to x_0} f_{|A'}(x) = l\,.$$

Si verifica facilmente che l'insieme $L \subset \overline{\mathbf{R}}$ dei punti limite di una funzione f per $x \to x_0$ è sempre chiuso, ma in $\overline{\mathbf{R}}$ i chiusi sono anche compatti quindi L ammette massimo e minimo (eventualmente $+\infty$ o $-\infty$).

Definizione 8.16 - *Se f è limitata in un intorno di x_0 poniamo*

$$\liminf_{x\to x_0} f(x) = \min L \qquad e \qquad \limsup_{x\to x_0} f(x) = \max L\,.$$

Se f non è limitata inferiormente o superiormente in nessun intorno di x_0 poniamo

$$\liminf_{x\to x_0} f(x) = -\infty \qquad o \qquad \limsup_{x\to x_0} f(x) = +\infty\,.$$

*Questi due elementi di $\overline{\mathbf{R}}$ si chiamano **minimo limite** e **massimo limite** di f per $x \to x_0$*

Ad esempio

$$\liminf_{x\to 0} \operatorname{sen}\frac{1}{x} = -1 \qquad \text{e} \qquad \limsup_{x\to 0} \operatorname{sen}\frac{1}{x} = 1\,.$$

La disuguaglianza

$$\liminf_{x\to x_0} f(x) \leqslant \limsup_{x\to x_0} f(x)$$

è sempre vera e si riduce ad un'uguaglianza se e solo se f ammette limite, pari al loro comune valore, caso in cui L è formato da un unico punto.

Esercizio 8.9 - *Dimostrare che $l = \liminf_{x\to x_0} f(x) \in \mathbf{R}$ se e solo se sono soddisfatte le due condizioni*

(a) $\forall\varepsilon > 0\ \exists\delta > 0\ :\ \forall x \in A \cap I_\delta(x_0) - \{x_0\}$ *si ha* $f(x) > l - \varepsilon$,

(b) $\forall\varepsilon > 0\ \forall\delta > 0\ :\ \exists x \in A \cap I_\delta(x_0) - \{x_0\}\ :\ f(x) < l + \varepsilon$.

Esercizio 8.10 - *Dimostrare che*

$$\liminf_{x\to x_0} f(x) = \sup_{\rho>0}\, \inf_{I_\rho(x_0)} f(x) = \lim_{\rho\to 0}\, \inf_{I_\rho(x_0)} f(x)\,,$$
$$\limsup_{x\to x_0} f(x) = \inf_{\rho>0}\, \sup_{I_\rho(x_0)} f(x) = \lim_{\rho\to 0}\, \sup_{I_\rho(x_0)} f(x)\,.$$

8.4 Confronto di infinitesimi e di infiniti

Vogliamo attribuire un significato rigoroso all'idea intuitiva di "quanto rapidamente" una funzione convergente "si avvicini" al suo limite o, se diverge, "si allontani" verso $+\infty$ o $-\infty$. Naturalmente non sarà possibile stabilire in *assoluto* che cosa vuol dire tendere rapidamente ad un limite, bensì va intesa in un senso *relativo*, rispetto ad un'altra funzione con lo stesso limite, utilmente scelta talvolta come *funzione campione*. In altre parole, sarà possibile soltanto stabilire quale, tra due funzioni, tenda più rapidamente dell'altra ad un certo limite comune.

Per quanto riguarda le funzioni convergenti, non è restrittivo limitarsi a quelle infinitesime. Infatti se $x_0 \in \overline{\mathbf{R}}$, $U \in \mathscr{I}(x_0)$ e $f : U - \{x_0\} \to \mathbf{R}$ allora f tende a $l \in \mathbf{R}$ per $x \to x_0$ se e solo se f differisce da l per una funzione infinitesima che è esattamente la funzione $\sigma(x) = f(x) - l$. In termini simbolici

$$\lim_{x\to x_0} f(x) = l \in \mathbf{R} \Leftrightarrow \exists\sigma : U - \{x_0\} \to \mathbf{R}\ :\ f(x) = l + \sigma(x) \quad \text{e} \quad \lim_{x\to x_0} \sigma(x) = 0\,,$$

così la rapidità di convergenza di f verso l può essere misurata da quella con cui σ tende a 0.

Siano dunque $x_0 \in \overline{\mathbf{R}}$, $U \in \mathscr{I}(x_0)$ e $f, g : U - \{x_0\} \to \mathbf{R}$ tali che

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

Definizione 8.17 - *Diciamo che f e g sono infinitesimi dello **stesso ordine**, si scrive $f = O(g)$, se esistono $k \in \mathbf{R} - \{0\}$ e un infinitesimo $\sigma : U - \{x_0\} \to \mathbf{R}$ tali che*

$$f(x) = g(x)(k + \sigma(x)) \quad \forall x \in U - \{x_0\}\,. \tag{8.3}$$

Se g non ha zeri in $U - \{x_0\}$ la (8.3) equivale a

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = k\,. \tag{8.4}$$

Ad esempio $\operatorname{sen} x = O(x)$, $1 - \cos x = O(x^2) = O(\operatorname{sen}^2 x)$. La Definizione 8.17 stabilisce una relazione di equivalenza tra gli infinitesimi dello stesso ordine. La proprietà simmetrica, immediata da verificare usando la (8.4)

$$\lim_{x \to x_0} \frac{g(x)}{f(x)} = \frac{1}{k}\,,$$

è riconoscibile anche dalla forma (8.3) che implica

$$g(x) = f(x)\Big(\frac{1}{k} + \tau(x)\Big)\,, \quad \text{con} \quad \tau(x) = -\frac{\sigma(x)}{k(k + \sigma(x))}\,.$$

Va osservato che ci sono infinitesimi tra loro non confrontabili, ad esempio il rapporto tra gli infinitesimi $x \operatorname{sen} 1/x$ e x non ha limite per $x \to 0$.

Definizione 8.18 - *Diciamo che f è un infinitesimo di **ordine superiore** a g, si scrive $f = o(g)$, se esiste un infinitesimo $\sigma : U - \{x_0\} \to \mathbf{R}$ tale che*

$$f(x) = g(x)\sigma(x) \quad \forall x \in U - \{x_0\}\,. \tag{8.5}$$

Il prodotto tra due infinitesimi è un infinitesimo di ordine superiore ad ognuno dei due fattori. Si tratta della stessa relazione di prima, ma con $k = 0$ che, sempre che g non abbia zeri in $U - \{x_0\}$, porta alla condizione equivalente

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = 0\,. \tag{8.6}$$

Ad esempio $\operatorname{sen} x = o(\sqrt{x})$, $1 - \cos x = o(x)$ e $\log(1 + x^2) = o(x)$.

Definizione 8.19 - *Diciamo che f è un infinitesimo di ordine $\alpha > 0$ rispetto a g se $f = O(|g|^\alpha)$. Nella relazione*

$$f(x) = |g(x)|^\alpha (k + \sigma(x))$$

*il termine $k|g(x)|^\alpha$ si chiama **parte principale** dell'infinitesimo f rispetto all'infinitesimo g.*

Le due funzioni $f(x) = e^{-x}$ e $g(x) = 1/x$ soddisfano $f = o(g)$ per $x \to +\infty$, ma per nessun valore di $\alpha > 0$ si ha $f = O(g^\alpha)$.

Teorema 8.20 (Principio di sostituzione degli infinitesimi) - *Siano $f, g, F, G : U - \{x_0\}$ quattro funzioni infinitesime per $x \to x_0$ tali che $F = o(f)$, $G = o(g)$ e che f, g, $G + g$ siano prive di zeri in $U - \{x_0\}$. Se esiste in $\overline{\mathbf{R}}$ il limite di f/g per $x \to x_0$ allora esiste anche il limite per $x \to x_0$ di $(f + F)/(g + G)$ e*

$$\lim_{x \to x_0} \frac{f(x) + F(x)}{g(x) + G(x)} = \lim_{x \to x_0} \frac{f(x)}{g(x)}\,.$$

Dimostrazione. Basta osservare che

$$\frac{f(x)+F(x)}{g(x)+G(x)} = \frac{f(x)}{g(x)} \cdot \frac{1+\dfrac{F(x)}{f(x)}}{1+\dfrac{G(x)}{g(x)}} \tag{8.7}$$

e applicare i teoremi algebrici.

□

Ad esempio

$$\lim_{x\to 0} \frac{x-1+\cos x}{\operatorname{sen}^2 x^2 + \log(1+x)} = \lim_{x\to 0} \frac{x}{\log(1+x)} = 1$$

perché $1-\cos x = o(x)$ e $\operatorname{sen}^2 x^2 = O(x^4) = o(\log(1+x))$. Oppure

$$\begin{aligned}\lim_{x\to+\infty} \frac{e^{-x}+(\pi/2-\operatorname{arctg} x)^{3/2}}{\sqrt{1+\sqrt{x}\operatorname{sen}^2 1/x}-\cos 1/x} &= \lim_{x\to+\infty} \frac{e^{-x}+\operatorname{arctg}^{3/2} 1/x}{1+1/2\sqrt{x}\operatorname{sen}^2 1/x - \cos 1/x} \\ &= \lim_{y\to 0^+} \frac{2\sqrt{y}\operatorname{arctg}^{3/2} y}{\operatorname{sen}^2 y} = 2\end{aligned}$$

essendo $e^{-x} = o(1/x^\alpha)$ per ogni $\alpha > 0$, $\sqrt{1+t} = 1+t/2+o(t)$ e $1-\cos t = o(t^{3/2})$ per $t \to 0$.

Le due funzioni $f(x) = \log(\cos x + \operatorname{sen} x)$ e $g(x) = \log(e^x + \operatorname{sen} x)$ sono entrambe infinitesime per $x \to 0$ e dello stesso ordine di x, lo si vede dal fatto che sia $\cos x$, sia e^x tendono a 1 e $\operatorname{sen} x = O(x)$. Ma sarebbe sbagliato dedurne che $f(x)/x$ e $g(x)/x$ hanno lo stesso limite. Infatti non è vero perché

$$\lim_{x\to 0} \frac{\log(\cos x + \operatorname{sen} x)}{x} = 1 \qquad \text{e} \qquad \lim_{x\to 0} \frac{\log(e^x + \operatorname{sen} x)}{x} = 2\,.$$

Una funzione $f : U-\{x_0\} \to \mathbf{R}$ viene detta un *infinito* per $x \to x_0$ se f è positivamente o negativamente divergente per $x \to x_0$. Certamente un infinito non può avere zeri in un intorno di x_0 che possiamo supporre ancora U.

Definizione 8.21 - *Diciamo che f e g sono infiniti dello* ***stesso ordine****, e si scrive $f = O(g)$, se*

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = k \in \mathbf{R} - \{0\}\,. \tag{8.8}$$

Ad esempio $\sqrt{1+x^2} = O(x)$ per $x \to +\infty$. Possiamo anche ricondurre la Definizione 8.21 alla Definizione 8.17 passando dalle funzioni f e g alle reciproche $1/f$ e $1/g$, quindi valgono le stesse osservazioni di prima.

Definizione 8.22 - *Diciamo che f è un infinito di* ***ordine superiore*** *a g, e si scrive $f = o(g)$, se*

$$\lim_{x\to x_0} \frac{f(x)}{g(x)} = \pm\infty\,. \tag{8.9}$$

Osserviamo che $f = o(g)$ se e solo se $1/f$ è un infinitesimo di ordine superiore a $1/g$.

Definizione 8.23 - *Diciamo che f è un infinito di ordine $\alpha > 0$ rispetto a g se $f = O(|g|^\alpha)$, cioè se*

$$\lim_{x\to x_0} \frac{f(x)}{|g(x)|^\alpha} = k \in \mathbf{R} - \{0\}\,,$$

o, in modo equivalente, $1/f = O(1/|g|^\alpha)$ nel senso degli infinitesimi. Il termine $k|g(x)|^\alpha$ si chiama ***parte principale*** *dell'infinito f rispetto all'infinito g.*

Teorema 8.24 (Principio di sostituzione degli infiniti) - *Siano $f, g, F, G : U - \{x_0\}$ quattro infiniti per $x \to x_0$ tali che $f = o(F)$ e $g = o(G)$. Se esiste in $\overline{\mathbf{R}}$ il limite di f/g per $x \to x_0$ allora esiste anche il limite per $x \to x_0$ di $(f+F)/(g+G)$ e*

$$\lim_{x \to x_0} \frac{f(x) + F(x)}{g(x) + G(x)} = \lim_{x \to x_0} \frac{f(x)}{g(x)} .$$

Ad esempio

$$\lim_{x \to 0} \frac{\operatorname{senh}(1/x^2) + \log x^2}{e^{1/x^2} + \operatorname{cotg}^2 x} = \lim_{x \to 0} \frac{\operatorname{senh}(1/x^2)}{e^{1/x^2}} = \frac{e^{1/x^2}}{2e^{1/x^2}} = \frac{1}{2} .$$

8.5 Asintoti

Sia $f : A \to \mathbf{R}$ divergente, positivamente o negativamente, per $x \to x_0^+$ o per $x \to x_0^-$ con $x_0 \in \mathscr{D}(A)$. Nelle vicinanze di x_0 il grafico di f si adagia, da una parte o dall'altra del punto x_0, sulla retta di equazione $x = x_0$. Questa viene detta *asintoto verticale* di f nel punto x_0. Se f presenta questo comportamento da una sola parte rispetto a x_0 allora di parla di asintoto *destro* o *sinistro*. La funzione $e^{1/x}$ su $\mathbf{R} - \{0\}$, ad esempio, ammette come asintoto destro la retta $x = 0$ e tende a 0 per $x \to 0^-$.
Se invece $x_0 = +\infty$ (il caso $x_0 = -\infty$ è analogo) con A quindi non limitato superiormente, diciamo che f ammette la retta di equazione $y = \alpha x + \beta$ come *asintoto a* $+\infty$ se esiste una funzione $\sigma : A \to \mathbf{R}$ infinitesima per $x \to +\infty$ tale che

$$f(x) = \alpha x + \beta + \sigma(x) , \tag{8.10}$$

proprietà che si descrive anche dicendo che f è *lineare all'*∞. Seguendo il grafico di f, lo vediamo avvicinarsi a quella retta. Un caso particolare è quello di f convergente, con $\alpha = 0$ e $\beta \in \mathbf{R}$, per cui $f(x) = \beta + \sigma(x)$ e l'asintoto è la retta $y = \beta$.
Se f diverge all'∞, affinché valga la (8.10) per qualche $\alpha, \beta \in \mathbf{R}$ dobbiamo verificare che esistano e siano finiti i due limiti

$$\alpha = \lim_{x \to +\infty} \frac{f(x)}{x} \neq 0 \qquad \text{e} \qquad \beta = \lim_{x \to +\infty} (f(x) - \alpha x) .$$

Il primo ci dice che f ha andamento lineare all'∞, poi dal secondo si deduce che $f(x) \sim \alpha x + \beta$. Ad esempio $\sqrt{x}$, $\log x$, x^2 e e^x non hanno asintoto perché divise per x tendono o a 0 o a $+\infty$.

Esempi

8.6 *La funzione $f(x) = \sqrt{1 + x^2}$, $x \in \mathbf{R}$, ammette gli asintoti $y = x$ a $+\infty$ e $y = -x$ a $-\infty$, ai quali il grafico si avvicina da sopra. Infatti $f(x) > |x|$ per ogni $x \in \mathbf{R}$ e*

$$\lim_{x \to +\infty} \sqrt{1 + x^2} - |x| = 0 \qquad \forall x \in \mathbf{R} .$$

8.7 *La funzione $f(x) = x + \sqrt{x}$, $x > 0$, diverge senza asintoto perché $f(x)/x \to 1$, ma $f(x) - x \to +\infty$.*

8.8 *Una funzione razionale $f(x) = P(x)/Q(x)$ ha lo stesso asintoto orizzontale, oppure obliquo, a $+\infty$ e a $-\infty$ se e solo se P e Q sono polinomi, rispettivamente, con lo stesso grado o con il grado di P superiore di 1 a quello di Q.*

8.9 *La funzione*

$$f(x) = \frac{x^2 + \operatorname{sen} e^x}{x + 1} , \qquad x \neq -1 ,$$

ammette la retta $y = x - 1$ come asintoto a $+\infty$ e a $-\infty$, per vederlo basta scrivere la f così

$$f(x) = x - 1 + \frac{1 + \operatorname{sen} e^x}{x+1}\,,$$

come somma di una funzione lineare con una funzione infinitesima.

8.10 *La funzione*

$$f(x) = x(1 + e^{-x})^{e^x}\,, \qquad x \in \mathbf{R}\,,$$

che diverge sia a $+\infty$ che a $-\infty$, soddisfa

$$\lim_{x\to+\infty} \frac{f(x)}{x} = e \qquad e \qquad \lim_{x\to-\infty} \frac{f(x)}{x} = 1\,.$$

Sul comportamento asintotico a $+\infty$ si ha

$$\begin{aligned} x(1+e^{-x})^{e^x} - ex &= ex(e^{e^x \log(1+e^{-x})-1} - 1) \sim ex(e^x \log(1+e^{-x}) - 1) \\ &= ex\frac{\log(1+e^{-x}) - e^{-x}}{e^{-x}} \sim -ex\frac{e^{-2x}/2}{e^{-x}} = -\frac{xe^{1-x}}{2} \to 0\,, \end{aligned}$$

l'asintoto c'è ed è la retta $y = ex$. A $-\infty$, posto $y = e^{-x} \to +\infty$, si ha

$$x(1+e^{-x})^{e^x} - x = -\left[\exp\left(\frac{\log(1+y)}{y}\right) - 1\right]\log y \sim -\frac{\log(1+y)\log y}{y} \to 0$$

e l'asintoto è la retta $y = x$.

8.11 *La funzione*

$$f(x) = \frac{1}{x}\operatorname{sen}\frac{1}{x}\,, \qquad x > 0\,,$$

non diverge per $x \to 0^+$ perché non ha limite, ma non è limitata né superiormente, né inferiormente e ammette nell'intorno infinite oscillazioni sempre più ampie. La retta $x = 0$ può comunque essere considerata asintoto verticale.

8.12 *La funzione*

$$f(x) = \frac{1}{\operatorname{sen}(1/x)}\,, \qquad x > 0\,,\ x \neq \frac{1}{k\pi}\,,\ k \in \mathbf{N}\,,$$

soddisfa la relazione $|f(x)| \geqslant 1$, in particolare

$$f(x) \geqslant 1 \quad se\ \frac{1}{(2k+1)\pi} < x < \frac{1}{2k\pi} \qquad e \qquad f(x) \leqslant -1 \quad se\ \frac{1}{2k\pi} < x < \frac{1}{(2k-1)\pi}\,.$$

Quindi la f non ammette limite per $x \to 0^+$, inoltre $\sup f = +\infty$ e $\inf f = -\infty$ e in ogni intorno di 0 presenta un'infinità di asintoti verticali, ma non si può sostenere che l'asse y sia un asintoto!

8.13 *La funzione*

$$f(x) = \frac{\operatorname{sen} x^2}{x}\,, \qquad x > 0\,,$$

ammette come asintoto orizzontale la retta $y = 0$, alla quale il grafico di f si avvicina per $x \to +\infty$ presentando oscillazioni sempre più fitte.

8.6 Proprietà generali delle funzioni continue

In questa definizione non si richiede che un punto x_0 sia di accumulazione per il dominio A di una funzione, ma che vi appartenga.

Definizione 8.25 - *Una funzione $f : A \to \mathbf{R}$, con $A \subset \mathbf{R}$, si dice **continua** nel punto $x_0 \in A$ se*

$$\forall \varepsilon > 0\, \exists \delta > 0 \ : \ \forall x \in A \quad |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon\,. \tag{8.11}$$

Anche per questa definizione vi sono versioni più generali in termini di distanze e intorni

$$\forall \varepsilon > 0\, \exists \delta > 0 \ : \ \forall x \in A \quad d(x, x_0) < \delta \Rightarrow d(f(x), f(x_0) < \varepsilon\,,$$
$$\forall \varepsilon > 0\, \exists \delta > 0 \ : \ f(B_\delta(x_0)) \subset B_\varepsilon(f(x_0))\,,$$
$$\forall V \in \mathscr{I}(f(x_0))\, \exists U \in \mathscr{I}(x_0) \ : \ f(U) \subset V\,.$$

Definizione 8.26 - *Se f è continua in ogni punto di A allora diciamo che è **continua in** A. L'insieme di tutte le funzioni continue in A si indica con $C^0(A)$.*

Ovviamente la restrizione di una funzione continua ad un sottoinsieme qualunque del dominio rimane continua.

La somiglianza con la definizione di limite suggerisce che possa sussistere una relazione tra l'essere continua in x_0 e l'essere convergente per $x \to x_0$, ma il limite può essere considerato solo se x_0 è anche di accumulazione per A. Con questa ipotesi aggiuntiva la (8.11) diventa

$$\forall \varepsilon > 0\, \exists \delta > 0 \ : \ \forall x \in A\,,\ x \neq x_0\,, \quad |x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon\,,$$

cioè

$$\lim_{x \to x_0} f(x) = f(x_0)\,.$$

Dunque per due motivi f può essere *discontinua*, non continua, in un punto $x_0 \in A \cap \mathscr{D}(A)$:

1. esiste finito $l = \lim_{x \to x_0} f(x)$, ma $l \neq f(x_0)$,
2. esiste, ma non è finito, o non esiste $\lim_{x \to x_0} f(x)$.

Se invece x_0 è un punto isolato di A si può sempre scegliere $\delta > 0$ così piccolo in modo che $A \cap I_\delta(x_0)$ contenga solo x_0, ma allora la condizione $|f(x) - f(x_0)| < \varepsilon$ si riduce alla $|f(x_0) - f(x_0)| < \varepsilon$ che è banalmente vera. Quindi ogni funzione è continua nei punti isolati di A, per esempio le successioni sono funzioni continue da $\mathbf{N} \to \mathbf{R}$. Siccome il caso dei punti isolati è poco significativo, da ora in poi supporremo che i punti di A siano tutti di accumulazione per A. Ne segue che possiamo servirci di gran parte dei risultati, praticamente tutti, sulle funzioni convergenti in un punto $x_0 \in \mathbf{R}$ per dedurne altrettanti sulle funzioni continue.

Esempi

8.14 *Ogni funzione costante su A è continua. Infatti se $f(x) = c$ per ogni $x \in A$ il suo limite per $x \to x_0$ vale c che è anche $f(x_0)$.*

8.15 *Le funzioni lipschitziane, le hölderiane, anche localmente, e quindi tutte le funzioni elementari sono continue nei loro domini naturali. Se f soddisfa*

$$|f(x_1) - f(x_2)| \leqslant |x_1 - x_2|^\alpha \qquad \forall x_1, x_2 \in A$$

con $k \geqslant 0$ e $\alpha \in]0,1]$ basta scegliere $\delta = (\varepsilon/k)^{1/\alpha}$ per ogni $\varepsilon > 0$. Si noti che in questi casi δ dipende solo da ε, proprietà su cui torneremo nel § 8.9.

8.16 *La funzione*

$$f(x) = \frac{x^2-1}{x-1}, \quad x \in \mathbf{R} - \{1\},$$

è continua. Essa coincide infatti con la funzione $x+1$ che in $\mathbf{R}-\{1\}$ è continua. Si osservi che il punto 1 è di accumulazione per questo dominio e che esiste e vale 2 il limite della f per $x \to 1$, quindi la funzione $x+1$ in $\mathbf{R}$ è un prolungamento continuo della f a tutto $\mathbf{R}$.

8.17 *La funzione $|x|$ è continua su $\mathbf{R}$, le funzioni $x/|x|$, $1/x$, $\operatorname{arctg} 1/x$, $e^{1/x}$, $\operatorname{sen} 1/x$, $\operatorname{tang} x$ e $(\operatorname{sen} x)/x$ sono tutte continue su $\mathbf{R}-\{0\}$, ma solo l'ultima, l'unica che ammette limite finito per $x \to 0$, ammette prolungamento continuo in 0 ed è la funzione $\bar f : \mathbf{R} \to \mathbf{R}$ definita da*

$$\bar f(x) = \begin{cases} \dfrac{\operatorname{sen} x}{x} & \text{se } x \neq 0 \\ 1 & \text{se } x = 0 . \end{cases}$$

Più in generale, se $f : A \to \mathbf{R}$ è continua e $x_0 \notin A$ è di accumulazione per A allora esiste un'unica funzione continua $\bar f : A \cup \{x_0\} \to \mathbf{R}$ che prolunga la f al punto x_0 se e solo se esiste finito il limite finito di f per $x \to x_0$. Essa è definita da

$$\bar f(x) = \begin{cases} f(x) & \text{se } x \in A \\ \lim\limits_{x \to x_0} f(x) & \text{se } x = x_0 . \end{cases}$$

L'unicità del prolungamento continuo è un'immediata conseguenza dell'unicità del limite. Naturalmente se x_0 non è di accumulazione ogni prolungamento è continuo.

La continuità di una funzione è una proprietà *puntuale*, può essere verificata in un punto, in certi punti o in tutti i punti del dominio, ma la continuità in un punto non implica la continuità in un suo intorno, non è in altre parole una proprietà *locale*. Per vederlo costruiamo una funzione continua in 0 che in ogni intorno di 0 ammette punti di discontinuità . A partire dalla $g : [1, +\infty[\to \mathbf{R}$ definita da

$$g(x) = \begin{cases} 1 & \text{se } x \in [2n-1, 2n[\\ -1 & \text{se } x \in [2n, 2n+1[, \end{cases} \qquad n \in \mathbf{N} - \{0\},$$

la quale è discontinua nei numeri naturali, si può costruire la funzione $f : [0,1] \to \mathbf{R}$

$$f(x) = \begin{cases} x g(1/x) & \text{se } x \in]0,1] \\ 0 & \text{se } x = 0 \end{cases}$$

che è continua in 0, ma presenta salti in ogni suo intorno.

8.18 *La funzione*

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

è discontinua in ogni punto. Infatti ogni $x_0 \in \mathbf{R}$ è di accumulazione sia per $\mathbf{Q}$, sia per $\mathbf{R} - \mathbf{Q}$ e non esiste il limite di f per $x \to x_0$ perché

$$\lim_{x \to x_0} f_{|\mathbf{Q}}(x) = 1 \qquad \text{e} \qquad \lim_{x \to x_0} f_{|\mathbf{R}-\mathbf{Q}}(x) = 0\,.$$

Esercizio 8.11 - *Definire, se esiste, il prolungamento continuo in* 0 *delle funzioni* $x \log |x|$, $e^{1/x}$ *e* $x \operatorname{sen} 1/x$, $x \neq 0$.

Esercizio 8.12 - *Dimostrare che una funzione* $f : A \to \mathbf{R}$ *è continua se e solo se per ogni aperto* $B \subset \mathbf{R}$ $f^{-1}(B)$ *è aperto in* A.

Le funzioni continue ereditano in modo ovvio proprietà importanti dalla teoria dei limiti tra le quali segnaliamo le seguenti (si parla di un punto ma possono valere in ogni punto).

(C^0)1. Una funzione continua in un punto è limitata in un intorno del punto, ne segue che una funzione continua è localmente limitata,

(C^0)2. se f è continua in $x_0 \in A$ e $f \geqslant 0$ in $A - \{x_0\}$ allora $f(x_0) \geqslant 0$ e, viceversa, se $f(x_0) > 0$ esiste un intorno di x_0 nei punti x del quale $f(x) > 0$,

(C^0)3. se $f : A \to \mathbf{R}$ è continua in x_0, $f(A) \subset B$ e $g : B \to \mathbf{R}$ è continua in $y_0 = f(x_0)$ allora $f \circ g$ è continua in x_0,

(C^0)4. f è continua in x_0 se e solo se per ogni successione $(x_n) \subset A$ tale che $x_n \to x_0$ si ha $f(x_n) \to f(x_0)$,

(C^0)5. somme, prodotti, rapporti di due funzioni continue f, g in x_0 e anche $f \vee g$ e $f \wedge g$ sono funzioni continue in x_0, dunque $C^0(A)$ è uno spazio vettoriale, un'algebra e un reticolo.

Esercizio 8.13 - *Dimostrare che l'insieme degli zeri di una funzione continua su* A *è un sottoinsieme chiuso di* A.

Definizione 8.27 - *Una funzione* $f : A \to \mathbf{R}$ *si dice* ***semicontinua inferiormente*** *(s.c.i.)* ***[superiormente]*** *(s.c.s.) in* $x_0 \in A$ *se*

$$\forall \varepsilon > 0\ \exists \delta > 0\ :\ \forall x \in A \quad |x - x_0| < \delta \Rightarrow f(x) > f(x_0) - \varepsilon \quad [f(x) < f(x_0) + \varepsilon]\,.$$

Se ciò avviene in ogni punto diciamo che f *è semicontinua inferiormente [superiormente] in* A.

In altre parole f è semicontinua se nella Definizione 8.25 vale una soltanto delle due disuguaglianze

$$f(x_0) - \varepsilon < f(x) < f(x_0) + \varepsilon$$

e f è continua se e solo se è semicontinua sia inferiormente che superiormente.

Se x_0 è isolato in A non c'è nulla da dire, se è di accumulazione si ha immediata la seguente caratterizzazione

$$f \text{ è s.c.i. se e solo se } f(x_0) \leqslant \liminf_{x \to x_0} f(x)\,,$$

$$f \text{ è s.c.s. se e solo se } f(x_0) \geqslant \limsup_{x \to x_0} f(x)\,.$$

Anche da questa osservazione si capisce che f è continua in x_0 se e solo se valgono insieme le due condizioni di semicontinuità. Infatti dal momento che la relazione

$$\liminf_{x \to x_0} f(x) \leqslant \limsup_{x \to x_0} f(x)$$

è sempre vera, aggiungere ad essa la condizione

$$\limsup_{x\to x_0} f(x) \leqslant f(x_0) \leqslant \liminf_{x\to x_0} f(x)$$

equivale ad affermare che il limite esiste e coincide con $f(x_0)$.

Naturalmente c'è un'altra caratterizzazione che si basa sulle successioni

$$f \text{ è s.c.i. se e solo se } \forall (x_n) \subset A \; x_n \to x_0 \Rightarrow f(x_0) \leqslant \liminf_{n\to\infty} f(x_n)\,,$$
$$f \text{ è s.c.s. se e solo se } \forall (x_n) \subset A \; x_n \to x_0 \Rightarrow f(x_0) \geqslant \limsup_{n\to\infty} f(x_n)\,.$$

8.19 *La funzione*

$$f(x) = \begin{cases} \operatorname{arctg} 1/x & \text{se } x \neq 0 \\ c & \text{se } x = 0 \end{cases}$$

è semicontinua inferiormente se $c \leqslant -\pi/2$ *e superiormente se* $c \geqslant \pi/2$.

Esercizio 8.14 - *Determinare il massimo prolungamento a* $\mathbf{R}$ *semicontinuo inferiormente della funzione* $e^{1/x}$ *su* $\mathbf{R} - \{0\}$.

8.7 Funzioni continue su un intervallo

La definizione di insieme connesso è di natura topologica e la evitiamo volentieri, ma a livello intuitivo un insieme connesso lo possiamo intendere come fatto di un unico pezzo, che non sia formato da più parti separate. In $\mathbf{R}$ si ha una facile caratterizzazione, gli unici insiemi connessi sono gli intervalli, i quali, tra l'altro, sono anche gli unici insiemi convessi. Poiché essi non presentano interruzioni, possiamo ritenere che in qualche modo le informazioni possano correre da una zona all'altra trasformando proprietà puntuali in proprietà globali, ciò che succede in un punto ha influenza sugli altri. Le conseguenze sono notevoli, ad esempio si può dimostrare che se una funzione è crescente in ogni punto di un intervallo allora è crescente sull'intervallo, ma ve ne sono tante altre. Qui siamo interessati alle proprietà speciali che una funzione continua soddisfa per il solo fatto di essere definita su un intervallo, proprietà che discendono tutte dal teorema degli zeri, a noi già noto (v. il Teorema 3.7), che adesso ridimostriamo col metodo della *bisezione*, un tipo di ragionamento che ricorda, mutatis mutandis, il modo in cui abbiamo dimostrato il Teorema di Bolzano-Weierstraß.

Teorema 8.28 (degli zeri) - *Se* $f \in C^0[a,b]$ *e* $f(a)f(b) < 0$ *allora esiste* $\xi \in [a,b]$ *tale che* $f(\xi) = 0$.

Dimostrazione. Supponiamo $f(a) < 0 < f(b)$. Se il punto medio $m_0 = (a+b)/2$ è uno zero di f l'abbiamo trovato e non c'è nulla da aggiungere, altrimenti poniamo

$$a_1 = m_0 \text{ e } b_1 = b \text{ se } f(m_0) < 0 \quad \text{oppure} \quad a_1 = a \text{ e } b_1 = m_0 \text{ se } f(m_0) > 0$$

e scegliamo $m_1 = (a_1 + b_1)/2$. Se m_1 è uno zero di f l'abbiamo trovato e non c'è nulla da aggiungere, altrimenti poniamo

$$a_2 = m_1 \text{ e } b_2 = b_1 \text{ se } f(m_1) < 0 \quad \text{oppure} \quad a_2 = a_1 \text{ e } b_2 = m_1 \text{ se } f(m_1) > 0$$

e scegliamo $m_2 = (a_2 + b_2)/2$. Se m_2 è uno zero di f l'abbiamo trovato e non c'è nulla da aggiungere, altrimenti andiamo avanti procedendo sempre nello stesso modo. Se dopo un certo numero di passi, mettiamo al k-esimo, il punto medio dell'intervallo $I_k =$

$[a_k, b_k]$ è uno zero di f ci si ferma, altrimenti si ottiene una successione decrescente di intervalli I_n, $I_{n+1} \subset I_n$, con gli estremi sinistri a_n che formano una successione crescente e quelli destri, i b_n, che formano una successione decrescente. Per costruzione si ha $f(a_n) < 0$ e $f(b_n) > 0$ per ogni $n \in \mathbf{N}$. In quanto monotone e limitate le due successioni (a_n) e (b_n) convergono ai rispettivi estremi superiore e inferiore

$$a_n \to \alpha = \sup_{n \in \mathbf{N}} a_n \leqslant \inf_{n \in \mathbf{N}} b_n = \beta \leftarrow b_n \,,$$

quindi $b_n - a_n \to \beta - \alpha$. Ma $\alpha = \beta$ perché le misure $b_n - a_n$ degli I_n soddisfano

$$\begin{aligned} b_n - a_n &= \frac{1}{2}(b_{n-1} - a_{n-1}) = \frac{1}{2^2}(b_{n-2} - a_{n-2}) = \frac{1}{2^3}(b_{n-3} - a_{n-3}) = \dots \\ &= \frac{1}{2^k}(b_{n-k} - a_{n-k}) = \dots = \frac{1}{2^{n-1}}(b_1 - a_1) = \frac{1}{2^n}(b - a) \end{aligned}$$

e passando al limite a destra per $n \to \infty$ si ottiene

$$b_n - a_n \to 0\,.$$

Non resta adesso che dimostrare che il punto $\xi = \alpha = \beta$ è uno zero di f. Usando la continuità per successioni e la permanenza del segno, si ha

$$f(a_n) \to f(\xi) \leqslant 0 \quad \text{e} \quad f(b_n) \to f(\xi) \geqslant 0$$

pertanto $f(\xi) = 0$.

□

Questo risultato, importante di per sé come teorema di esistenza (dove si usa la completezza di $\mathbf{R}$?), ha notevoli conseguenze che ora vediamo. Bisogna precisare però che l'assumere valori di segno opposto agli estremi di un intervallo è soltanto una condizione sufficiente, anche funzioni come la x^2, $x \in [-1, 1]$, ammettono zeri.

Esercizio 8.15 - *Costruire controesempi eliminando una alla volta le ipotesi del Teorema 8.28.*

Si chiama *punto fisso* di una funzione $f : X \to X$ un punto $\bar{x} \in X$ che la f lascia in se stesso, cioè tale che $f(\bar{x}) = \bar{x}$. Vi sono molti *teoremi del punto fisso* che sotto ipotesi varie ne garantiscono l'esistenza. Uno di essi è il seguente.

Corollario 8.29 - *Ogni funzione continua $f : [a, b] \to [a, b]$ ha un punto fisso.*

Dimostrazione. Se $f(a) = a$ oppure $f(b) = b$ questi sono i punti fissi. Altrimenti sarà $f(a) > a$ e $f(b) < b$. La funzione continua $\varphi(x) = f(x) - x$, $x \in [a, b]$ soddisfa dunque $\varphi(a) > 0$ e $\varphi(b) < 0$, quindi esiste $\bar{x} \in [a, b]$ tale che $\varphi(\bar{x}) = 0$ per il Teorema 8.28.

□

Immediata conseguenza del Teorema degli zeri, ma anche ad esso equivalente, è la proprietà di una funzione continua di trasformare intervalli in intervalli. Nel seguito indicheremo con I un intervallo.

Corollario 8.30 (Teorema dei valori intermedi) - *Se $f \in C^0(I)$ allora $f(I)$ è un intervallo.*

Dimostrazione. Dobbiamo dimostrare che, scelti ad arbitrio $y_1, y_2 \in I$ con $y_1 < y_2$, se $y_1 < \eta < y_2$ esiste $\xi \in I$ tale che $f(\xi) = \eta$. Siano $x_1, x_2 \in I$ due punti tali che $f(x_1) = y_1$ e $f(x_2) = y_2$. La funzione $g(x) = f(x) - \eta$, $x \in I$, è continua e soddisfa

$$g(x_1) = f(x_1) - \eta < 0 \qquad \text{e} \qquad g(x_2) = f(x_2) - \eta > 0\,.$$

Per il Teorema 8.28 esiste ξ nell'intervallo di estremi x_1 e x_2 (non ne abbiamo stabilito l'ordine) tale che $g(\xi) = f(\xi) - \eta = 0$.

□

Esercizio 8.16 - *Dimostrare che una funzione continua su I a valori in un insieme finito, o in $\mathbf{N}$, in $\mathbf{Q}$, in $\mathbf{R}-\mathbf{Q}$, è necessariamente costante.*

Nel Corollario 8.30 si afferma che l'immagine continua di un intervallo è un intervallo, ma che relazione c'è tra il carattere di I e quello della sua immagine $f(I)$? Vengono conservate anche proprietà come la limitatezza e il fatto di essere aperto o chiuso? La risposta è negativa ed è facile costruire le situazioni più disparate: uno dei due può essere limitato e l'altro no, si pensi alle funzioni $\operatorname{tang} x$ su $]-\pi/2, \pi/2[$ e $\operatorname{arctg} x$ su $\mathbf{R}$; I può essere aperto e $f(I)$ chiuso o semiaperto, a meno che non sia monotona, basta considerare la funzione $\operatorname{sen} x$ che trasforma l'aperto $]0, 2\pi[$ nel chiuso $[-1,1]$. Solo se I è limitato e chiuso, quindi compatto, anche la sua immagine è un intervallo limitato e chiuso. Deve essere un intervallo per il Corollario 8.30, ma anche compatto per un risultato di cui parleremo nel prossimo paragrafo.

Rimanendo ancora nell'ambito delle funzioni continue su un intervallo, vogliamo stabilire la continuità della funzione inversa quando esiste. Certamente la stretta monotonia garantisce l'esistenza dell'inversa, opportunamente definita come funzione surgettiva, ma in generale non la sua continuità. La funzione $f : [0,1] \cup]2,3] \to [0,2]$ definita da

$$f(x) = \begin{cases} x & \text{se } 0 \leqslant x \leqslant 1 \\ x-1 & \text{se } 2 < x \leqslant 3 \end{cases} \tag{8.12}$$

è continua e invertibile, ma l'inversa $f^{-1} : [0,2] \to [0,1] \cup]2,3]$ definita da

$$f^{-1}(x) = \begin{cases} x & \text{se } 0 \leqslant x \leqslant 1 \\ x+1 & \text{se } 1 < x \leqslant 2 \end{cases}$$

non è continua (anche perché se lo fosse l'immagine sarebbe un intervallo). Se però il dominio di f è un intervallo la monotonia diventa una condizione necessaria e sufficiente sia per l'esistenza dell'inversa che per la sua continuità. Per vederlo dobbiamo invertire il Corollario 8.30 con l'ipotesi aggiuntiva che f sia monotona.

Lemma 8.31 - *Se $f : I \to \mathbf{R}$ è monotona e $f(I)$ è un intervallo allora f è continua.*

Dimostrazione. Supponiamo che f sia crescente e discontinua in un punto x_0 interno ad I. Per il Teorema 8.14 esistono finiti i limiti sinistro e destro, l^- e l^+, di f in x_0 e si ha

$$f(x) \leqslant l^- < l^+ \leqslant f(x')$$

per ogni $x, x' \in I$ tali che $x < x_0 < x'$. Basta considerare allora un punto $y \in [l^-, l^+]$ che sia diverso da $f(x_0)$ e non esiste nessun $x \in I$ tale che $f(x) = y$, in contraddizione con l'ipotesi che $f(I)$ sia un intervallo. Se il punto di salto x_0 è un estremo dell'intervallo ci sarà uno solo dei due limiti, ma la dimostrazione non cambia. □

Ad esempio la funzione $f : [0,2[\to [0,2[$ definita da

$$f(x) = \begin{cases} x & \text{se } 0 \leqslant x \leqslant 1 \\ 3-x & \text{se } 1 < x < 2 \end{cases}$$

non è continua in quanto non monotona.

Teorema 8.32 - *Se $f \in C^0(I)$ è iniettiva allora è strettamente monotona e l'inversa $f^{-1} : f(I) \to I$ è continua.*

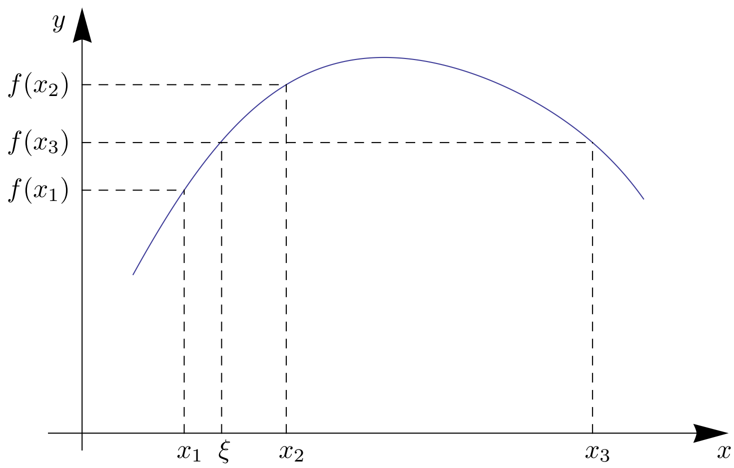
Dimostrazione. È chiaro che se è monotona, in quanto iniettiva, deve esserlo strettamente. Ricordiamo inoltre che f è monotona (strettamente) se e solo se per ogni terna di punti $x_1, x_2, x_3 \in I$ tali che $x_1 < x_2 < x_3$ si ha

$$(f(x_1) - f(x_2))(f(x_2) - f(x_3)) > 0 .$$

Supponiamo per assurdo che questa condizione non sia soddisfatta. Esistono allora tre punti come sopra tali che

$$f(x_1) < f(x_2) \qquad \text{e} \qquad f(x_2) > f(x_3) \tag{8.13}$$

oppure con le disuguaglianze scambiate. Se vale la (8.13) e ad esempio $f(x_1) < f(x_3)$ (l'altro caso è analogo), essendo $f(x_3)$ un valore intermedio tra $f(x_1)$ e $f(x_2)$, esiste $\xi \in [x_1, x_2]$ tale che $f(\xi) = f(x_3)$ e quindi f non sarebbe iniettiva (v. Figura 8.7).



La f^{-1} è strettamente monotona nello stesso senso della f e trasforma l'intervallo $f(I)$ nell'intervallo I, quindi per il Lemma 8.31 deve essere è continua.

□

Risultano dunque continue le funzioni $\operatorname{arcsen} x$ e $\arccos x$ su $[-1, 1]$, $\operatorname{arctg} x$ e $\operatorname{settsenh} x$ su $\mathbf{R}$, $\log x$ su $]0, +\infty[$ se viste come le funzioni inverse di $\operatorname{sen} x$, $\cos x$ ecc. sugli intervalli opportunamente scelti.

8.8 Funzioni continue su un compatto

Un'altra proprietà che si conserva nel passaggio dal dominio all'immagine di una funzione continua è, come accennato, la compattezza. Si tratta di un risultato di carattere generale con conseguenze di notevole portata nell'Analisi Matematica in fatto di massimi e minimi.

Teorema 8.33 - *Siano X e Y spazi metrici e $K \subset X$ compatto. Se $f : K \to Y$ è una funzione continua allora $f(K)$ è compatto in Y.*

Dimostrazione. Dobbiamo dimostrare che ogni successione $(y_n) \subset f(K)$ ammette un'estratta convergente ad un punto $y \in f(K)$. Per ogni $n \in \mathbf{N}$ sia $x_n \in K$ tale che $f(x_n) = y_n$. Poiché K è compatto esiste una sottosuccessione $(x_{k_n}) \subset (x_n)$ convergente ad un certo $x \in K$ ed essendo f continua $f(x_{k_n}) \to f(x)$. Ma la successione $y_{k_n} = f(x_{k_n})$ è un'estratta della (y_n) ed è convergente al punto $y = f(x) \in K$.

□

Conseguenza immediata del Teorema 8.33 è il seguente risultato, fondamentale per lo studio dei massimi e dei minimi delle funzioni a valori reali.

Corollario 8.34 (Teorema di Weierstraß) - *Se $K \subset X$ è un insieme compatto ogni funzione continua $f : K \to \mathbf{R}$ ammette massimo e minimo.*

Dimostrazione. Per il Teorema 8.33 $f(K)$ è un compatto di $\mathbf{R}$, quindi limitato e chiuso. In quanto limitato ammette estremi superiore e inferiore finiti e in quanto chiuso li contiene come punti aderenti.

□

Il Teorema di Weierstraß ci permette di inquadrare in un contesto generale il Problema 7.15 della proiezione di un punto su un insieme compatto di uno spazio metrico e ci dà una spiegazione della validità del Teorema 7.16 di esistenza di tale proiezione: fissato il punto $x \in X$ la funzione distanza $y \to d(x, y)$ è continua, per questo ha minimo su K, infatti

$$|d(x, y_1) - d(x, y_2)| \leqslant d(y_1, y_2)$$

per la disuguaglianza triangolare.

Se invece della proiezione su un compatto si volesse affrontare lo stesso problema su un chiuso $C \subset X$, il Teorema di Weierstraß non si può applicare altro che nei casi in cui i chiusi e limitati sono tutti compatti. In $\mathbf{R}^n$ ciò è vero ed è per questo che nel § 7.3 abbiamo potuto osservare che la proiezione esiste anche in questo caso. È bastato considerare l'intersezione, compatta, tra C e la chiusura di una palla di centro x e osservare poi che il punto di minima distanza da tale intersezione non può che essere anche di minima distanza da C. Infine l'ulteriore condizione di natura vettoriale che C sia convesso ci ha garantito l'unicità della proiezione.

Vediamo adesso un'interessante variante al Teorema di Weierstraß. Se siamo interessati solo al minimo di una funzione $f : K \to \mathbf{R}$, sempre con $K \subset X$ compatto, non c'è bisogno di supporre che f sia continua, basta che sia semicontinua inferiormente. Analogamente per il massimo basta che lo sia superiormente. È chiaro che se f è semicontinua, sia superiormente che inferiormente, allora è continua e ritroviamo il Teorema di Weierstraß.

Definiamo *successione minimizzante* per una funzione $f : A \to \mathbf{R}$ ogni successione $(x_n) \subset A$ tale che

$$\lim_{n\to\infty} f(x_n) = \inf_{x\in A} f(x)\,.$$

Indichiamo con J l'estremo inferiore e osserviamo che una successione minimizzante esiste sempre per le solite proprietà dell'estremo inferiore

$$J = -\infty \qquad \forall n \in \mathbf{N}\ \exists x_n \in A\ :\ f(x_n) < -n \Rightarrow \lim_{n\to\infty} f(x_n) = -\infty,$$

$$J > -\infty \qquad \forall n \in \mathbf{N}\ \exists x_n \in A\ :\ f(x_n) < \frac{1}{n} + J\,,\ f(x_n) \geqslant J \Rightarrow \lim_{n\to\infty} f(x_n) = J.$$

Teorema 8.35 - *Se $K \subset X$ è compatto e $f : K \to \mathbf{R}$ è semicontinua inferiormente allora esiste il minimo di f.*

Dimostrazione. Dimostriamo che $J = \inf_K f$ è il minimo di f. Sia $(x_n) \subset K$ una successione minimizzante per f e (x_{k_n}) un'estratta convergente e sia $x_0 \in K$ il suo limite. Ora, $f(x_{k_n}) \to J$, d'altra parte, essendo f semicontinua inferiormente, si ha

$$J \leqslant f(x_0) \leqslant \liminf_{n\to\infty} f(x_{k_n}) = \lim_{n\to\infty} f(x_{k_n}) = J\,,$$

quindi $f(x_0) = J$ e x_0 è un punto di minimo per f su K.

□

Quanto detto riguarda l'esistenza del punto di minimo, ma talvolta è di qualche interesse stabilirne anche l'unicità, che per la compattezza del dominio corrisponde al caso in cui tutte le successioni minimizzanti convergono allo stesso punto. Quest'affermazione è falsa se il dominio non è compatto: la funzione x^2e^{-x}, $x \in \mathbf{R}$, ha un solo punto di minimo, $x_0 = 0$, ma sono minimizzanti sia le successioni infinitesime, sia quelle positivamente divergenti. Di condizioni generali che garantiscono l'unicità ve ne sono diverse, ma una di esse, la più naturale e la più frequente, è la condizione di stretta convessità della funzione, per questo le funzioni convesse sono delle buone funzioni per i problemi di minimo.

Compattezza e semicontinuità sono gli strumenti fondamentali di un grande settore dell'Analisi Matematica che si chiama Calcolo delle Variazioni. In esso rientrano problemi dove la f rappresenta un costo da minimizzare, la distanza stessa ad esempio, oppure l'energia, la dispersione di calore, il tempo ecc.

Torniamo adesso al caso di $\mathbf{R}$ che ci interessa per lo studio dei grafici e vediamo altre varianti al Teorema di Weierstraß. Se il dominio è un intervallo e si rinuncia alla compattezza dobbiamo fare ulteriori ipotesi sul comportamento della funzione.

Corollario 8.36 - *Se $f \in C^0]a,b[$, $a,b \in \overline{\mathbf{R}}$, e $\lim_{x\to a} f(x) = \lim_{x\to b} f(x) = +\infty$ allora f ha minimo su $]a,b[$.*

Dimostrazione. Siano $\bar{x} \in]a,b[$, un punto qualunque e $M > f(\bar{x})$. Per l'ipotesi che f diverge agli estremi esistono $\alpha, \beta \in]a,b[$, $\alpha < \beta$, tali che $f(x) > M$ per ogni $x \in A_{\alpha,\beta} =]a,\alpha[\cup]\beta, b[$, ne segue che $\bar{x} \in [\alpha,\beta]$. Sul compatto $[\alpha,\beta]$ la f raggiunge il minimo in un punto x_0, ma questo punto è di minimo anche su tutto $]a,b[$, infatti per ogni $x \in]a,b[$ si ha

$$x \in A_{\alpha,\beta} \Rightarrow f(x) > M > f(\bar{x}) \geqslant f(x_0)\,.$$

□

Corollario 8.37 - *Se $f \in C^0[a,+\infty[$ e $\lim_{x\to+\infty} f(x) = l \geqslant f(a)$ allora f ha minimo su $[a,+\infty[$.*

Dimostrazione. Non c'è niente da dimostrare se $f(x) \geqslant f(a)$ per ogni $x \in [a,+\infty[$, situazione che comprende il caso di f crescente (o costante). Altrimenti scegliamo un $\bar{x} \in [a,+\infty[$ tale che $f(\bar{x}) < f(a)$. Fissato $\varepsilon > 0$ abbastanza piccolo in modo che $l - \varepsilon > f(\bar{x})$, esiste $b > a$ tale che

$$f(x) > l - \varepsilon \qquad \forall x > b\,,$$

quindi $\bar{x} \in [a,b]$. Sul compatto $[a,b]$ la f raggiunge il minimo in un punto x_0, ma questo punto è di minimo anche su tutto $[a,+\infty[$ dal momento che per ogni $x \in [a,+\infty[$ si ha

$$x > b \Rightarrow f(x) > l - \varepsilon > f(\bar{x}) \geqslant f(x_0)\,.$$

□

Esercizio 8.17 - *Dimostrare che la funzione $f(x) = x^2 - \log x$ ammette nell'intervallo $]0,+\infty[$ un unico punto di minimo.*

Concludiamo il paragrafo con un altro risultato sulla continuità della funzione inversa, si raccomanda di rivedere il Teorema 8.32 e di confrontarlo e combinarlo con esso.

Teorema 8.38 - *Se $K \subset X$ è un insieme compatto e $f : K \to Y$ è continua e iniettiva allora $f^{-1} : f(K) \to K$ è continua.*

Dimostrazione. Scelta una successione qualunque $(y_n) \subset f(K)$ che converge ad un punto $y_0 \in f(K)$, dobbiamo dimostrare che $f^{-1}(y_n) \to f^{-1}(y_0)$. Posto $x_n = f^{-1}(y_n)$, supponiamo per assurdo che la successione $(x_n) \subset K$ non converga a $x_0 = f^{-1}(y_0)$. In questo caso esiste un'estratta (x_{k_n}) convergente ad un certo $\hat{x} \in K$ con $\hat{x} \neq x_0$, ma per la continuità di f si ha $f(x_{k_n}) \to f(\hat{x}) \neq f(x_0) = y_0$ essendo anche iniettiva. Questa conclusione contraddice il fatto che $f(x_{k_n})$ è una sottosuccessione della (y_n) convergente a y_0.

□

La funzione (8.12) può essere usata come controesempio anche in questo contesto, si tratta di una funzione continua e invertibile, ma definita su un insieme non compatto, la sua inversa è discontinua.

8.9 Funzioni uniformemente continue

Abbiamo osservato che la continuità di una funzione è una proprietà puntuale: una funzione $f : A \to \mathbf{R}$ può essere continua in un punto, in alcuni sì e in altri no o in tutti i punti. L'uniforme continuità è invece una proprietà di natura globale ed è più restrittiva della semplice continuità su A. Ricordiamo che f è continua su A se e solo se

$$(8.14) \quad \forall x \in A \text{ e } \forall \varepsilon > 0 \; \exists \delta(x,\varepsilon) > 0 \; : \; \forall y \in A \quad |x-y| < \delta(x,\varepsilon) \Rightarrow |f(x)-f(y)| < \varepsilon\,.$$

La funzione $\varepsilon \to \delta(x,\varepsilon)$ può essere interpretata come una misura del *tasso di continuità* nel punto x, ma non è univocamente determinata perché la (8.14) continua a valere con lo stesso ε se $\delta'(x,\varepsilon) < \delta(x,\varepsilon)$. Naturalmente come può essere scelto δ dipende dalla funzione: è arbitrario se f è costante, ma, a parità di ε, quanto più rapidamente varia la f vicino a x, tanto più piccolo va scelto $\delta(x,\varepsilon)$ (v. Figura 8.1).

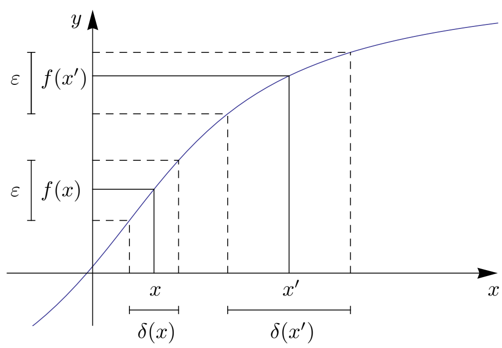


Figura 8.1: Come varia δ da punto a punto.

Se nella (8.14) si può usare un $\delta = \delta(\varepsilon) > 0$ indipendente da x, lo stesso in tutti i punti, essendo verificata la condizione

$$\delta(\varepsilon) = \inf_{x \in A} \delta(x,\varepsilon) > 0\,,$$

allora diciamo che f è uniformemente continua.

Definizione 8.39 - *Una funzione $f : A \to \mathbf{R}$ viene detta* ***uniformemente continua*** *se*

$$(8.15) \qquad \forall \varepsilon > 0 \; \exists \delta(\varepsilon) > 0 \; : \; \forall x, y \in A \quad |x-y| < \delta(\varepsilon) \Rightarrow |f(x)-f(y)| < \varepsilon\,.$$

Come al solito, questa definizione ha perfettamente senso anche per una funzione tra spazi metrici, basta sostituire la distanza del valore assoluto con le distanze proprie del dominio e del codomino.

Una funzione uniformemente continua è ovviamente continua. Vediamo invece alcuni esempi di funzioni continue che non sono uniformemente continue.

Esempi

8.20 *La funzione*

$$f(x) = \begin{cases} 1 & se\ 0 \leqslant x < 1 \\ 2 & se\ 1 < x \leqslant 2 \end{cases}$$

è continua su $[0,1[\cup]1,2]$*, ma non uniformemente continua perché se, per esempio, si sceglie* $\varepsilon = 1/2$*, o un numero qualsiasi minore di* 1*, per nessun* $\delta > 0$ *esistono* $x < 1 < y$ *con* $|x-y| < \delta$ *tali che* $|f(x) - f(y)| < 1/2$ *dato che* $|f(x) - f(y)| = 1$.

8.21 *La funzione* $f(x) = x^2$*,* $x \geqslant 0$*, è continua, ma non uniformemente continua. Infatti, poiché*

$$|x^2 - x_0^2| < \varepsilon \Leftrightarrow |x - x_0| < \frac{\varepsilon}{x + x_0}\,, \tag{8.16}$$

se $0 \leqslant x \leqslant x_0$ *e* $|x - x_0| < \delta(x_0, \varepsilon) = \varepsilon/2x_0$ *allora* $|x^2 - x_0^2| < \varepsilon$*, ma*

$$\inf_{x_0>0} \frac{\varepsilon}{2x_0} = 0\,.$$

Invece la restrizione di f *ad ogni intervallo* $[0,a]$ *è uniformemente continua perché*

$$\delta(\varepsilon) = \inf_{0\leqslant x_0\leqslant a} \frac{\varepsilon}{2x_0} = \frac{\varepsilon}{2a} > 0\,.$$

8.22 *La funzione* $f(x) = 1/x$*,* $0 < x \leqslant 1$*, è continua ma non uniformemente continua. Infatti, poiché*

$$\left|\frac{1}{x} - \frac{1}{x_0}\right| < \varepsilon \Leftrightarrow |x - x_0| < \varepsilon x_0 x\,, \tag{8.17}$$

se $0 < x_0 \leqslant x$ *e* $|x - x_0| < \delta(x_0, \varepsilon) = \varepsilon x_0^2$ *allora* $|1/x - 1/x_0| < \varepsilon$*, ma*

$$\inf_{0<x_0\leqslant 1} \varepsilon x_0^2 = 0\,.$$

Queste tre funzioni non sono uniformemente continue per motivi diversi, riconoscibili con opportune condizioni necessarie. Ma prima di parlare di queste, e poi anche di condizioni sufficienti, vediamo un'utile condizione necessaria e sufficiente.

Teorema 8.40 - *Una funzione* $f : A \to \mathbf{R}$ *è uniformemente continua se e solo se per ogni coppia di successioni* $(x_n), (y_n) \subset A$ *si ha*

$$\lim_{n\to\infty} |x_n - y_n| = 0 \Rightarrow \lim_{n\to\infty} |f(x_n) - f(y_n)| = 0\,.$$

Dimostrazione. Se f è uniformemente continua rimane ad essa associata una funzione $\delta(\varepsilon)$ che soddisfa la (8.15). Prese allora due successioni $(x_n), (y_n) \subset A$ tali che $|x_n - y_n| \to 0$, si avrà definitivamente $|x_n - y_n| < \delta(\varepsilon)$, quindi per gli stessi indici si avrà anche $|f(x_n) - f(y_n)| < \varepsilon$ per la Definizione 8.39.

Dimostriamo per assurdo l'implicazione contraria. Se f non fosse uniformemente continua, verificherebbe la negazione della Definizione 8.39

$$\exists \varepsilon > 0 \ : \ \forall \delta > 0\ \exists x, y \in A \ : \quad |x - y| < \delta \text{ e } |f(x) - f(y)| \geqslant \varepsilon\,.$$

Ma dal momento che δ è arbitrario, possiamo prendere $\delta = 1/n$, $n \in \mathbf{N} - \{0\}$, al quale corrispondono punti $x_n, y_n \in A$ tali che

$$|x_n - y_n| < \frac{1}{n} \quad \text{e} \quad |f(x_n) - f(y_n)| \geqslant \varepsilon\,.$$

Passando al limite per $n \to \infty$, $|x_n - y_n| \to 0$, ma le immagini $f(x_n)$ e $f(y_n)$ si mantengono a distanza finita, non inferiore a ε, in contraddizione con l'ipotesi.

□

Ad esempio la funzione continua $\operatorname{sen} x^2$ (e anche $\cos x^2$, v. Figura 8.2), $x \in \mathbf{R}$, non è uniformemente continua, la differenza tra le due successioni

$$x_n = \sqrt{\frac{\pi}{2} + 2n\pi} \qquad \text{e} \qquad y_n = \sqrt{\frac{3\pi}{2} + 2n\pi}\,, \qquad\qquad n \in \mathbf{N}\,,$$

è infinitesima, mentre $f(x_n) - f(y_n) = 2$.

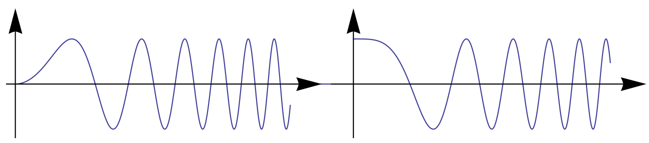


Figura 8.2: i grafici di $\operatorname{sen} x^2$ e $\cos x^2$.

Condizioni necessarie

Teorema 8.41 - *Una funzione $f : A \to \mathbf{R}$ uniformemente continua trasforma successioni di Cauchy in successioni di Cauchy.*

Dimostrazione. Dobbiamo dimostrare che se $(x_n) \subset A$ è di Cauchy allora anche $(f(x_n))$ è di Cauchy in $\mathbf{R}$. Scelto arbitrariamente $\varepsilon > 0$, esiste un indice $\nu \in \mathbf{N}$ tale che $|x_n - x_m| < \delta(\varepsilon)$ per ogni $m, n > \nu$. Allora $|f(x_n) - f(x_m)| < \varepsilon$ per gli stessi indici n m, quindi $(f(x_n))$ è di Cauchy.

□

Teorema 8.42 - *Se $f : A \to \mathbf{R}$ è uniformemente continua e $x_0 \in \mathbf{R}$ è di accumulazione per A allora esiste ed è finito il limite di f per $x \to x_0$.*

Dimostrazione. Se il limite non esistesse o non fosse finito, f non sarebbe di Cauchy in x_0, esisterebbe dunque un $\varepsilon > 0$ tale che per ogni $n \in \mathbf{N}$ esistono x_n e y_n in A tali che $|x_n - y_n| < 1/n$ e $|f(x_n) - f(y_n)| \geqslant \varepsilon$. Ma allora si contraddice la condizione necessaria e sufficiente per l'uniforme continuità del Teorema 8.40.

□

Entrambi i teoremi forniscono una spiegazione della mancata continuità uniforme degli Esempi 8.20 e 8.22. Si osservi inoltre che il Teorema 8.42 continua a valere, con la stessa dimostrazione, quando il codominio è uno spazio metrico completo e lo stesso può essere detto per il seguente importante corollario.

Corollario 8.43 - *Ogni funzione $f : A \to \mathbf{R}$ uniformemente continua ammette un unico prolungamento $\bar{f} : \bar{A} \to \mathbf{R}$ che è uniformemente continuo con lo stesso $\delta(\varepsilon)$.*

Dimostrazione. Ricordando che $\bar{A} = A \cup \mathscr{D}(A)$, il Teorema 8.42 garantisce che il prolungamento di f

$$\bar{f}(x) = \lim_{\substack{t \to x \\ t \in A}} f(t) \qquad \forall x \in \bar{A}\,,$$

unico come sappiamo per l'unicità del limite, sia ben definito. Rimane da dimostrare che $\bar{f}$ è uniformemente continua . Scelti $x, x' \in \bar{A}$ tali che $|x - x'| < \delta(\varepsilon)$, due successioni $(x_n), (x'_n) \subset A$ convergenti a x e a x' soddisfano definitivamente

$$|x_n - x'_n| < \delta(\varepsilon)\,,$$

da cui segue

$$|f(x_n) - f(x'_n)| < \varepsilon$$

per gli stessi indici. La tesi si ottiene passando al limite su n

$$|\bar{f}(x) - \bar{f}(x')| = \lim_{n\to\infty} |f(x_n) - f(x'_n)| \leqslant \varepsilon\,.$$

□

Teorema 8.44 - *Una funzione uniformemente continua $f : A \to \mathbf{R}$ trasforma ogni insieme limitato contenuto in A in un insieme limitato di $\mathbf{R}$.*

Dimostrazione. Sia $B \subset A$ limitato e, per assurdo, non lo sia $f(B)$, ad esempio superiormente. Per ogni $n \in \mathbf{N}$ esiste un punto $x_n \in B$ tale che $f(x_n) > n$. In quanto limitata, la (x_n) ammette una sottosuccessione (x_{k_n}) convergente in $\mathbf{R}$, però $(f(x_{k_n}))$ non converge perché $f(x_{k_n}) > k_n \to +\infty$ e questo contraddice il Teorema 8.42.

□

La funzione continua $\operatorname{sen} 1/x$, $x > 0$, trasforma limitati in limitati in quanto limitata, ma non è uniformemente continua perché non ha limite per $x \to 0$.

Citiamo senza dimostrarla un'ultima condizione necessaria che spiega l'Esempio 8.21.

Teorema 8.45 - *Se $f : [0, +\infty[\to \mathbf{R}$ è uniformemente continua esistono $a, b \in \mathbf{R}$ tali che*

$$|f(x)| \leqslant ax + b \qquad \forall x \in [0, +\infty[\,. \tag{8.18}$$

In altre parole, una funzione uniformemente continua ha un andamento asintotico *sublineare* all'∞. La funzione $x \operatorname{sen} x$, $x \geqslant 0$, soddisfa la (8.18), ma non è uniformemente continua, infatti la differenza tra le due successioni

$$x_n = 2n\pi \qquad \text{e} \qquad x'_n = 2n\pi + \arcsin\frac{1}{n}\,, \qquad\qquad n \in \mathbf{N}\,,$$

è infinitesima, mentre

$$|f(x_n) - f(x'_n)| = \left(2n\pi + \arcsin\frac{1}{n}\right)\frac{1}{n} \to 2\pi\,.$$

Condizioni sufficienti

Le condizioni di Lipschitz e di Hölder

$$|f(x_1) - f(x_2)| \leqslant k|x_1 - x_2| \qquad \text{e} \qquad |f(x_1) - f(x_2)| \leqslant k|x_1 - x_2|^\alpha\,, \quad 0 < \alpha < 1\,,$$

garantiscono la continuità uniforme, basta scegliere $\delta(\varepsilon) = \varepsilon/k$ e $\delta(\varepsilon) = (\varepsilon/k)^{1/\alpha}$ rispettivamente.

Vediamo qui un'altra condizione sufficiente in cui di nuovo si combina la compattezza del dominio con la continuità della funzione.

Teorema 8.46 (di Heine) - *Ogni funzione $f \in C^0(K)$ con K compatto è uniformemente continua.*

Dimostrazione. Se f non fosse uniformemente continua esisterebbero un $\varepsilon > 0$ e due successioni $(x_n), (x'_n) \subset K$ tali che

$$|x_n - x'_n| \to 0 \qquad \text{e} \qquad |f(x_n) - f(x'_n)| > \varepsilon \tag{8.19}$$

definitivamente. Una delle due, per esempio la (x_n), ammette una sottosuccessione (x_{k_n}) convergente ad un punto $x_0 \in K$, ma anche la (x'_{k_n}) converge a x_0 per la disuguaglianza triangolare. Dalla continuità di f segue

$$|f(x_{k_n}) - f(x'_{k_n})| \to |f(x_0) - f(x_0)| = 0$$

che contraddice la (8.19).

□

Vediamo un esempio di funzione non lipschitziana, né hölderiana, ma uniformemente continua per il Teorema di Heine. La funzione

$$f(x) = \begin{cases} \dfrac{1}{\log x} & \text{se } 0 < x \leqslant \dfrac{1}{2} \\ 0 & \text{se } x = 0 \end{cases}$$

è continua sul compatto $[0, 1/2]$ e quindi uniformemente continua. Però non soddisfa la relazione

$$|f(x)| \leqslant k|x|^\alpha \qquad \forall x \in [0, 1/2]$$

per nessun $k \geqslant 0$ e per nessun $\alpha \in]0, 1]$. Se fosse vera risulterebbe

$$-kx^\alpha \log x \geqslant 1 \qquad \forall x \in [0, 1/2]\,,$$

ma la funzione a primo membro è infinitesima per $x \to 0$.

Una variante del Teorema 8.46 consiste nel sostituire la compattezza del dominio con un'ipotesi sul comportamento della funzione all'∞.

Teorema 8.47 - *Una funzione $f \in C^0[a, +\infty[$ che ammette asintoto all'∞ è uniformemente continua.*

Dimostrazione. Supponiamo dapprima che l'asintoto di f sia l'asse x. Prese due successioni $(x_n), (x'_n) \subset [a, +\infty[$ tali che $|x_n - x'_n| \to 0$, dobbiamo dimostrare che $|f(x_n) - f(x'_n)| \to 0$. Questo è vero se una delle due, e quindi anche l'altra, è limitata perché su un compatto che le contiene vale il Teorema di Heine. Se non sono limitate non è restrittivo supporre che divergano, dato che lo stesso Teorema vale per ogni eventuale sottosuccessione limitata. Ma se divergono si ha per ipotesi $f(x_n) \to 0$ e $f(x'_n) \to 0$, quindi $|f(x_n) - f(x'_n)| \to 0$.

Se adesso f ammette come asintoto all'∞ una retta diversa, il grafico di una certa $\varphi(x) = ax + b$ uniformemente continua in quanto lineare, per il ragionamento precedente la funzione $g(x) = f(x) - \varphi(x)$, continua e infinitesima a $+\infty$, è uniformemente continua, quindi anche $f(x) = g(x) + \varphi(x)$ lo è.

□

In particolare ogni funzione continua su $\mathbf{R}$ che si annulla identicamente al di fuori di un insieme limitato è uniformemente continua. Queste funzioni sono dette a *supporto compatto*, essendo il supporto di f l'insieme

$$\operatorname{supp} f = \overline{\{x \in \mathbf{R} \mid f(x) \neq 0\}}\,.$$

Lo spazio delle funzioni a supporto compatto su $\mathbf{R}$ viene indicato con $C^0_0(\mathbf{R})$.

Per almeno due motivi funzioni come $e^{-x} \operatorname{sen} x$, $\dfrac{\operatorname{sen} x}{x}$ e $x^2 \operatorname{sen} \dfrac{1}{x}$ sono uniformemente continue su $[0, +\infty[$, perché sono continue con asintoto e perché sono lipschitziane. C'è una relazione tra queste due proprietà? Ora, così come $\operatorname{sen} x$ è lipschitziana senza asintoto, sarebbe errato ritenere che una funzione continua con asintoto debba essere lipschitziana. Per esempio $\dfrac{\operatorname{sen} x^2}{\sqrt{x}}$, $x > 0$, ha asintoto ma non è lipschitziana, quindi è uniformemente continua per il solo motivo di avere asintoto.

Capitolo 9

Calcolo differenziale

9.1 La derivata

Consideriamo una funzione $f : A \to \mathbf{R}$, con $A \subset \mathbf{R}$, e un punto $x_0 \in A$. Il rapporto incrementale di f in x_0

$$R(x) = \frac{f(x) - f(x_0)}{x - x_0} \qquad \forall x \in A - \{x_0\}$$

è, come ben noto, il valore del coefficiente angolare della retta passante per i punti $(x_0, f(x_0))$ e $(x, f(x))$ del grafico di f. Se per $x \to x_0$ esiste finito il limite di R significa che la retta considerata tende a disporsi verso un "retta limite" che è ragionevole interpretare come la *retta tangente al grafico* di f nel punto $(x_0, f(x_0))$ come illustrato nella Figura 9.1. L'operazione di passaggio al limite, ricordiamo, richiede che x_0 sia di accumulazione per A, dunque dobbiamo supporre $x_0 \in A \cap \mathscr{D}(A)$. In questo capitolo assumiamo allora che A sia privo di punti isolati, tutti i punti di A saranno di accumulazione per A.

Definizione 9.1 - *Diciamo che $f : A \to \mathbf{R}$ è* ***derivabile*** *in $x_0 \in A$ se esiste finito il limite*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}\,. \tag{9.1}$$

Questo numero si chiama ***derivata*** *di f nel punto x_0 e viene indicato con $f'(x_0)$.*

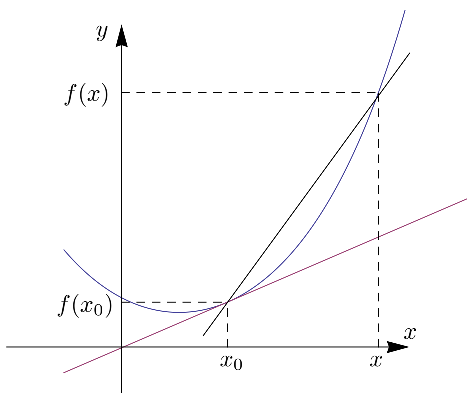


Figura 9.1: La retta tangente come posizione limite di rette secanti.

Diremo ***retta tangente*** *al grafico di f nel punto x_0 la retta di equazione*

$$y = f(x_0) + f'(x_0)(x - x_0)\,.$$

Altre notazioni di uso comune per la derivata in x_0 sono

$$Df(x_0) \quad \text{e} \quad \frac{df(x_0)}{dx}\,.$$

Se f è derivabile in x_0 fra tutte le rette passanti per il punto $(x_0, f(x_0))$, di equazione quindi $y = f(x_0) + m(x - x_0)$, ne esiste una ed una sola che ha un "contatto" col grafico di f superiore alle altre. Esiste cioè un unico valore di m tale che

$$f(x) - f(x_0) - m(x - x_0) = o(x - x_0)\,,$$

o, in modo equivalente, tale che

$$\lim_{x \to x_0} \frac{f(x) - f(x_0) - m(x - x_0)}{x - x_0} = 0\,. \tag{9.2}$$

Il valore di m che soddisfa questa condizione è proprio $f'(x_0)$, la retta è quella tangente. Infatti, come ogni funzione differisce dal suo limite per una quantità infinitesima, così avviene anche per il rapporto incrementale

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) + \sigma(x) \qquad \forall x \in A\,, \tag{9.3}$$

dove $\sigma(x) \to 0$ per $x \to x_0$. Moltiplicando per $x - x_0$ si ottiene

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(x - x_0) \tag{9.4}$$

Viceversa, se per un certo $m \in \mathbf{R}$ vale la (9.2), da essa si ricava subito il suo valore

$$m = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0} = f'(x_0)\,.$$

Una funzione che soddisfa la (9.2) viene detta *differenziabile* e l'applicazione lineare $h \to df_{x_0}(h) = f'(x_0)h$ si chiama *differenziale* di f nel punto x_0. Il ragionamento appena illustrato dimostra che la derivabilità e la differenziabilià sono proprietà equivalenti.

Conviene talvolta usare la nuova variabile $h = x - x_0$ con la quale le (9.1) e (9.4) diventano rispettivamente

$$f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} \qquad \text{e} \qquad f(x_0 + h) = f(x_0) + f'(x_0)h + o(h)\,.$$

Ovvia conseguenza della (9.4) è la seguente proprietà.

Teorema 9.2 - *Una funzione $f : A \to \mathbf{R}$ differenziabile in un punto $x_0 \in A$ è necessariamente continua in tal punto.*

Dimostrazione. Basta passare al limite per $x \to x_0$ nella (9.4).

□

Definizione 9.3 - *La funzione $f : A \to \mathbf{R}$ viene detta* ***derivabile in*** *A se lo è in ogni punto di A.*

Se f è derivabile in A rimane ben definita la *funzione derivata*

$$x \to f'(x) = \lim_{h\to 0} \frac{f(x+h) - f(x)}{h} = \lim_{t\to x} \frac{f(t) - f(x)}{t - x} \qquad \forall x \in A\,.$$

L'esistenza di f' su A implica la continuità di f, ma non è detto che f' sia continua, come vedremo più avanti con un esempio. Se lo è allora diciamo che f è di classe C^1 . Indicheremo con $C^1(A)$ l'insieme delle funzioni derivabili definite su A con derivata continua su A.

Esercizio 9.1 - *Qual è l'equazione della retta normale al grafico di f in un suo punto $(x_0, f(x_0))$?*

Esempi

9.1 *Ogni funzione costante, $f(x) = c$, è derivabile e ha derivata identicamente nulla. Infatti è già nullo il rapporto incrementale in ogni punto x_0*

$$\frac{f(x) - f(x_0)}{x - x_0} = \frac{c - c}{x - x_0} = 0 \qquad \forall x \in A - \{x_0\}\,,$$

quindi lo è anche il suo limite per $x \to x_0$.

9.2 *Le potenze $p_n(x) = x^n$, $n \in \mathbf{N}$, sono derivabili in $\mathbf{R}$ e $p_n'(x) = nx^{n-1}$, basta fermarsi al termine lineare in h nella formula del binomio*

$$(x+h)^n = \sum_{k=0}^{n} \binom{n}{k} h^k x^{n-k} = x^n + nx^{n-1}h + o(h)\,.$$

9.3 *La funzione esponenziale e^x ha per derivata se stessa*

$$\lim_{h\to 0} \frac{e^{x+h} - e^x}{h} = e^x \lim_{h\to 0} \frac{e^h - 1}{h} = e^x\,.$$

9.4 *Per ogni $x > 0$, $D \log x = 1/x$. Infatti*

$$\lim_{h\to 0} \frac{\log(x+h) - \log x}{h} = \lim_{h\to 0} \frac{\log(1 + h/x)}{h} = \frac{1}{x}\,.$$

9.5 *Riguardo le funzioni trigonometriche si ha*

$$D \operatorname{sen} x = \cos x \qquad e \qquad D \cos x = -\operatorname{sen} x\,.$$

Infatti

$$\lim_{h\to 0} \frac{\operatorname{sen}(x+h) - \operatorname{sen} x}{h} = \lim_{h\to 0} \frac{2}{h} \cos\left(\frac{2x+h}{2}\right) \operatorname{sen} \frac{h}{2} = \cos x\,.$$

Il calcolo della derivata di $\cos x$ viene lasciata per esercizio.

Esercizio 9.2 - *Scrivere l'equazione della retta tangente e della retta normale al grafico della funzione $\operatorname{sen} x$ nel punto $x = \pi/4$.*

Esercizio 9.3 - *Trovare la più semplice funzione $f \in C^1(\mathbf{R})$ che sia costante pari a 0 sulla semiretta $x \leqslant 0$ e costante pari a 1 sulla semiretta $x \geqslant 1$.*

Una funzione può non essere derivabile perché non esiste il limite del rapporto incrementale oppure perché esiste ma non è finito. La funzione $|x|$, $x \in \mathbf{R}$, non è derivabile in 0 perché

$$\lim_{x\to 0^+} \frac{|x|}{x} = 1 \qquad \text{e} \qquad \lim_{x\to 0^-} \frac{|x|}{x} = -1\,,$$

infatti il suo grafico presenta in $(0,0)$ un punto angoloso, privo di retta tangente.

Il rapporto incrementale in 0 della funzione continua

$$f(x) = \begin{cases} x \operatorname{sen} \dfrac{1}{x} & \text{se } x \neq 0 \\ 0 & \text{se } \ x = 0 \end{cases}$$

è la funzione

$$R(x,0) = \operatorname{sen} \frac{1}{x}\,, \qquad x \neq 0\,,$$

che notoriamente non ha limite per $x \to 0$.

La funzione $\sqrt{x}$, $x \geqslant 0$, non è derivabile in 0 perché

$$R(x,0) = \frac{\sqrt{x}}{x} \to +\infty\,.$$

Applichiamo la (9.4) al calcolo del limite nel seguente caso di indeterminazione.

Teorema 9.4 (I regola di L'Hôpital) - *Siano f e g due funzioni continue nell'intorno U di x_0 e derivabili nel punto x_0 tali che $f(x_0) = g(x_0) = 0$, $g(x) \neq 0$ per ogni $x \in U - \{x_0\}$ e $g'(x_0) \neq 0$. Allora*

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \frac{f'(x_0)}{g'(x_0)}\,.$$

Dimostrazione. Annullandosi le due funzioni in x_0, per la (9.4) il loro rapporto in un punto x può essere scritto nella forma

$$\frac{f(x)}{g(x)} = \frac{f'(x_0)(x-x_0) + o(x-x_0)}{g'(x_0)(x-x_0) + o(x-x_0)}$$

e poi si passa al limite trascurando gli infinitesimi di ordine superiore.

□

Ad esempio,

$$\lim_{x \to 0} \frac{\operatorname{sen} x}{x} = \frac{\cos 0}{1} = 1\,.$$

Il Teorema 9.4 non è applicabile se $g'(x_0) = 0$, sappiamo che $(1 - \cos x)/x^2 \to 1/2$ per $x \to 0$, ma non rientra nel Teorema 9.4 perché $g'(0) = 0$.

9.2 Regole di derivazione

Proposizione 9.5 - *Se le funzioni $f, g : A \to \mathbf{R}$ sono derivabili e $\lambda \in \mathbf{R}$ allora sono derivabili anche λf, $f+g$, fg e, dove hanno senso, anche $1/f$ e f/g e le rispettive derivate sono*

$$(\lambda f)' = \lambda f'\,, \qquad (f+g)' = f' + g' \quad e \quad (fg)' = f'g + fg'\,,$$

$$\left(\frac{1}{f}\right)' = -\frac{f'}{f^2} \qquad e \qquad \left(\frac{f}{g}\right)' = \frac{f'g - fg'}{g^2}\,.$$

Dimostrazione. Per le prime due basta osservare che

$$\frac{\lambda f(x) - \lambda f(x_0)}{x - x_0} = \lambda \frac{f(x) - f(x_0)}{x - x_0}\,,$$

$$\frac{(f(x)+g(x)) - (f(x_0)+g(x_0))}{x-x_0} = \frac{f(x)-f(x_0)}{x-x_0} + \frac{g(x)-g(x_0)}{x-x_0}\,.$$

Per il prodotto

$$\frac{f(x)g(x) - f(x_0)g(x_0)}{x - x_0} = f(x)\frac{g(x) - g(x_0)}{x - x_0} + g(x_0)\frac{f(x) - f(x_0)}{x - x_0}$$

e poi si passa al limite ricordando che f è continua in x_0 (v. Teorema 9.2). Infine

$$\frac{\dfrac{1}{f(x)} - \dfrac{1}{f(x_0)}}{x - x_0} = \frac{f(x_0) - f(x)}{f(x)f(x_0)(x - x_0)}$$

e anche qua si usa la continuità di f nel passaggio al limite. La derivata del rapporto si lascia per esercizio.

□

Esempi

9.6 *I polinomi, in quanto combinazioni lineari di potenze ad esponente intero positivo, sono derivabili e*

$$P(x) = \sum_{k=0}^{n} a_k x^k \quad \Rightarrow \quad P'(x) = \sum_{k=1}^{n} k a_k x^{k-1} .$$

9.7 *Per la regola di derivazione dell'inverso,* $1/x^n$ *è derivabile per ogni* $x \neq 0$ *e*

$$D\frac{1}{x^n} = -\frac{nx^{n-1}}{x^{2n}} = -\frac{n}{x^{n+1}} .$$

9.8 *Applichiamo la regola di derivazione del rapporto a* $\operatorname{tang} x$ *per* $x \neq \pi/2 + k\pi$

$$D \operatorname{tang} x = \frac{D \operatorname{sen} x \cdot \cos x - \operatorname{sen} x \cdot D \cos x}{\cos^2 x} = \frac{1}{\cos^2 x} = 1 + \operatorname{tang}^2 x .$$

9.9 *Se* $a > 0$ *e* $a \neq 1$*, essendo* $\log_a x = \log x / \log a$*, si ha*

$$D \log_a x = \frac{1}{x \log a} .$$

9.10 *Se* f *è derivabile in* $x \in \mathbf{R}$ *anche* $e^x f(x)$ *lo è e*

$$D(e^x f(x)) = e^x (f(x) + f'(x)) .$$

Esercizio 9.4 - *Calcolare la derivata del prodotto di* n *funzioni derivabili* $f_1, \ldots, f_n$.

Esercizio 9.5 - *Calcolare la derivata di* $\operatorname{cotg} x$*,* $x \neq k\pi$*,* $\cosh x$*,* $\operatorname{senh} x$ *e* $\operatorname{tangh} x$*.*

Teorema 9.6 - *Siano* f *e* g *due funzioni derivabili tali che il dominio di* f *contenga l'immagine di* g*. Allora* $f \circ g$ *è derivabile e*

$$Df(g(x)) = f'(g(x))g'(x) . \tag{9.5}$$

Dimostrazione. Applichiamo la (9.4) alla f e poi anche alla g

$$\begin{aligned} f(g(x+h)) - f(g(x)) &= f'(g(x))(g(x+h) - g(x)) + o(g(x+h) - g(x)) \\ &= f'(g(x))(g'(x)h + o(h)) + o(g'(x)h + o(h)) \\ &= f'(g(x))g'(x)h + o(h) . \end{aligned}$$

□

Esempi

9.11 *Per ogni $\alpha \in \mathbf{R}$ si ha*

$$Dx^\alpha = \alpha x^{\alpha-1}$$

come per le potenze ad esponente in $\mathbf{N}$ ($x \neq 0$ se $\alpha < 0$). Infatti

$$Dx^\alpha = De^{\alpha \log x} = \frac{\alpha}{x} e^{\alpha \log x} = \frac{\alpha}{x} x^\alpha = \alpha x^{\alpha-1} \, .$$

Se $\alpha = 1/2$, per esempio, $D\sqrt{x} = \dfrac{1}{2\sqrt{x}}$.

9.12 *Se $a > 0$*

$$Da^x = a^x \log a \, ,$$

basta cambiar base $a^x = e^{x \log a}$ e applicare il Teorema 9.6.

Esercizio 9.6 - *Verificare che $D \log |x| = 1/x$ per ogni $x \neq 0$ e, più in generale, che $D \log |f(x)| = f'(x)/f(x)$ per ogni x in cui f è definita, derivabile e $f(x) \neq 0$.*

Esercizio 9.7 - *Determinare in quali punti è derivabile la funzione $f(x) = |x|^\alpha$ al variare di $\alpha \in \mathbf{R}$ e, dove la derivata esiste, calcolarla.*

Esercizio 9.8 - *Calcolare la derivata della funzione $f(x) = x^x$ sul suo dominio naturale e trovare una formula per la derivata di $\varphi(x)^{\psi(x)}$.*

Teorema 9.7 - *Siano $x_0 \in A$ e $f : A \to \mathbf{R}$ una funzione invertibile in un intorno U di x_0 e derivabile in x_0 tale che $f'(x_0) \neq 0$. Allora la funzione inversa $f^{-1} : f(U) \to U$ è derivabile in $y_0 = f(x_0)$ e*

$$\frac{d}{dy} f^{-1}(y_0) = \frac{1}{f'(f^{-1}(y_0))} \, .$$

Dimostrazione. Per calcolare il limite per $y \to y_0$ del rapporto incrementale

$$R(y, y_0) = \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0}$$

è lecito effettuare la sostituzione $y = f(x)$ perché $f(x) \neq f(x_0)$ in U. Dunque

$$\lim_{y \to y_0} \frac{f^{-1}(y) - f^{-1}(y_0)}{y - y_0} = \lim_{x \to x_0} \frac{x - x_0}{f(x) - f(x_0)} = \frac{1}{f'(x_0)} \, ,$$

dove $x_0 = f^{-1}(y_0)$.

$\square$

9.13 *Applichiamo il Teorema 9.7 alle funzioni inverse di quelle trigonometriche*

$$D \operatorname{arctg} y = \frac{1}{D \operatorname{tang} x_{\,|x = \operatorname{arctg} y}} = \frac{1}{1 + \operatorname{tang}^2 \operatorname{arctg} y} = \frac{1}{1 + y^2} \qquad \forall y \in \mathbf{R} \, ,$$

$$\begin{aligned} D \operatorname{arcsen} y &= \frac{1}{D \operatorname{sen} x_{\,|x = \operatorname{arcsen} y}} = \frac{1}{\cos \operatorname{arcsen} y} \\ &= \frac{1}{\sqrt{1 - \operatorname{sen}^2 \operatorname{arcsen} y}} = \frac{1}{\sqrt{1 - y^2}} \qquad \forall y \in]-1, 1[\, , \end{aligned}$$

essendo $-\pi/2 \leqslant \operatorname{arcsen} y \leqslant \pi/2$. In modo simile si dimostra che

$$D \arccos y = -\frac{1}{\sqrt{1 - y^2}} \quad \forall y \in]-1, 1[\qquad e \qquad D \operatorname{arcotg} y = -\frac{1}{1 + y^2} \quad \forall y \in \mathbf{R} \, .$$

Esercizio 9.9 - *Si dia una spiegazione del fatto che nei domini di queste funzioni*

$$D(\operatorname{arctg} x + \operatorname{arcotg} x) = 0 \qquad e \qquad D(\operatorname{arcsen} x + \arccos x) = 0\,.$$

Esercizio 9.10 - *Calcolare la derivata di* $\log x$*, di* $\sqrt{x}$ *e di* $\operatorname{settsenh} x$ *usando il Teorema 9.7.*

Esercizio 9.11 - *Verificare che la funzione differenziabile* $f(x) = x^5 + x$ *ammette inversa* $f^{-1} : \mathbf{R} \to \mathbf{R}$ *differenziabile e di questa calcolare la derivata in* 2*.*

Esercizio 9.12 - *Si chiama funzione di Lambert* $W(x)$ *la funzione multivoca inversa della* $f(x) = xe^x$ *su* $\mathbf{R}$*. Indichiamo con* w *l'inversa della* f *su* $[-1, +\infty[$*. Calcolare* $w'(0)$ *e* $w'(e)$ *e mostrare che* w *non è derivabile per* $x = -e^{-1}$*.*

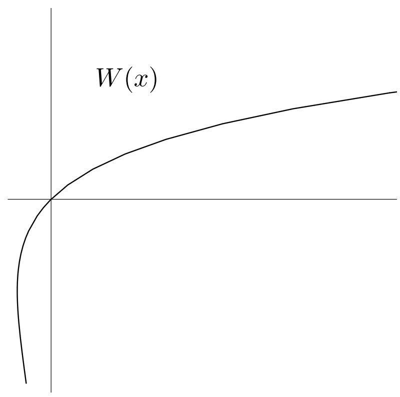


Figura 9.2: La funzione multivoca di Lambert.

Vediamo adesso un esempio di funzione ovunque derivabile ma con derivata discontinua.

9.14 *La funzione* $f : \mathbf{R} \to \mathbf{R}$ *definita da*

$$f(x) = \begin{cases} x^2 \operatorname{sen} \dfrac{1}{x} & \text{se } x \neq 0 \\ 0 & \text{se } x = 0 \end{cases}$$

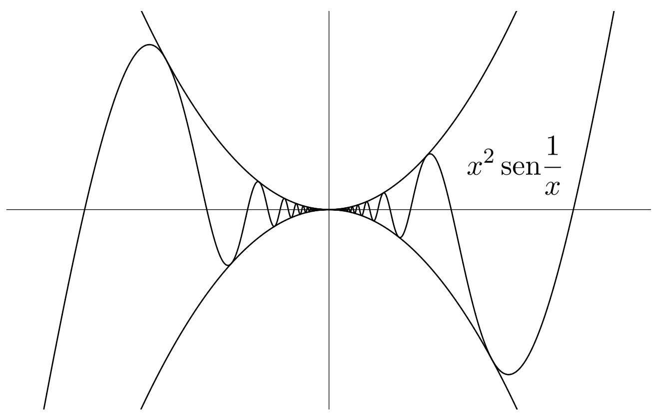


Figura 9.3: Esempio di funzione con derivata discontinua, $f(x) = x^2 \operatorname{sen} 1/x$.

è derivabile per ogni $x \in \mathbf{R}$. *Infatti, usando le regole di derivazione per* $x \neq 0$ *e la definizione di derivata per* $x = 0$, *si ottiene*

$$f'(x) = \begin{cases} 2x \operatorname{sen} \dfrac{1}{x} - \cos \dfrac{1}{x} & se\ x \neq 0 \\ \\ \lim\limits_{x \to 0} x \operatorname{sen} \dfrac{1}{x} = 0 & se\ x = 0\,. \end{cases}$$

Però f' *non ha limite per* $x \to 0$ *e quindi è discontinua.*

Definizione 9.8 - *Se la derivata* f' *di una funzione* $f : A \to \mathbf{R}$ *è a sua volta derivabile in un punto di* A, *o su tutto* A, *allora diciamo che* f *è derivabile due volte, nel punto o su* A, *e la derivata della derivata*

$$f''(x) = \lim_{t \to x} \frac{f'(t) - f'(x)}{t - x}$$

si chiama derivata seconda di f. *Per induzione, definiamo la derivata* k*-esima di* f

$$\begin{cases} f^{(0)}(x) = f(x) \\ f^{(k)}(x) = \lim\limits_{t \to x} \dfrac{f^{(k-1)}(t) - f^{(k-1)}(x)}{t - x} = Df^{(k-1)}(x) \quad \forall k \geqslant 1\,. \end{cases}$$

che si indica anche con le notazioni $D^k f(x)$ *e* $\dfrac{d^k}{dx^k} f(x)$.

Esercizio 9.13 - *Qual è la formula della derivata seconda della funzione inversa?*

Naturalmente tutto quello che abbiamo detto su f' a partire da f vale anche per $f^{(k)}$ a partire da $f^{(k-1)}$. L'esistenza della derivata di un certo ordine in un punto presuppone l'esistenza nell'intorno di quella di ordine precedente e ne implica la continuità nel punto. Se dunque f ammette derivata fino all'ordine k su A non è detto che $f^{(k)}$ sia continua, ma dovranno essere certamente continue su A tutte le derivate fino all'ordine $k-1$. Se poi è continua anche la $f^{(k)}$ allora diciamo che f è di classe C^k, si tratta di una misura di quanto f sia *regolare*. Indichiamo con $C^k(A)$ l'insieme delle funzioni derivabili su A fino all'ordine k con derivata k-esima continua. Ovviamente $C^k(A) \subset C^{k-1}(A)$. Poniamo infine

$$C^\infty(A) = \bigcap_{k \in \mathbf{N}} C^k(A)$$

per indicare l'insieme delle funzioni su A che ammettono derivata di ogni ordine, tutte necessariamente continue.

I polinomi stanno in $C^\infty(\mathbf{R})$ perché ogni polinomio è di classe C^1 e la derivata di un polinomio è ancora un polinomio. Riprendiamo l'Esempio 9.6 e deriviamo ancora

$$P''(x) = \sum_{k=2}^{n} k(k-1)a_k x^{k-2}\,, \quad P'''(x) = \sum_{k=3}^{n} k(k-1)(k-2)a_k x^{k-3}\,,$$

fino alla generica derivata di ordine $h \leqslant n$

$$P^{(h)}(x) = \sum_{k=h}^{n} k(k-1)(k-2)\ldots(k-h+1)a_k x^{k-h}\,. \tag{9.6}$$

Da una all'altra il grado diminuisce finché non si ottiene, per $h = n$, la costante

$$P^{(n)}(x) = n! a_n$$

e poi sempre 0 se si deriva ancora. Nella (9.6) per $x = 0$ rimane solo il termine noto $P^{(h)}(0) = h!a_h$ da cui segue la rappresentazione

$$P(x) = \sum_{k=0}^{n} \frac{P^{(k)}(0)}{k!} x^k \quad \text{oppure} \quad P(x) = \sum_{k=0}^{n} \frac{P^{(k)}(x_0)}{k!} (x - x_0)^k \tag{9.7}$$

a seconda della scelta del centro x_0.

Esercizio 9.14 - *Dimostrare che $x_0 \in \mathbf{R}$ è una radice di molteplicità k per il polinomio $P(x)$ se e solo se*

$$P(x_0) = P'(x_0) = P''(x_0) = \ldots = P^{(k-1)}(x_0) = 0 \quad e \quad P^{(k)}(x_0) \neq 0\,.$$

Anche le altre funzioni elementari sono di classe C^∞. Dopo aver visto infatti che stanno tutte in C^1 (fanno eccezione, ma solo in certi punti particolari, le potenze ad esponente $\alpha \in]0,1[$, l'arcsen x e l'arcos x), la loro derivata prima, in quanto funzione elementare, è anch'essa di classe C^1. Per induzione, dunque, la funzione di partenza ammette derivata di ogni ordine, necessariamente continua.

Esercizio 9.15 - *Dimostrare per induzione la formula di Leibniz*

$$D^k(fg) = \sum_{h=0}^{k} \binom{k}{h} f^{(h)} g^{(k-h)}$$

per il prodotto di due funzioni derivabili fino all'ordine k.

Esercizio 9.16 - *Stabilire la regolarità delle funzioni $x|x|$ e $x^2|x|$ su $\mathbf{R}$.*

Esercizio 9.17 - *Dimostrare che per $k \leqslant n$*

$$\frac{d^k}{dx^k}(x^n e^x) = e^x \sum_{h=0}^{k} h! \binom{k}{h} \binom{n}{h} x^{n-h}\,.$$

Che cosa si ottiene per $k > n$?

Esercizio 9.18 - *Un filo elastico con gli estremi nei punti $(-2, 0)$ e $(2, 0)$ viene teso e appoggiato al di sopra della parabola $y = 1 - x^2$, pensata come un ostacolo rigido. Assumendo che la configurazione di equilibrio del filo debba essere grafico di una funzione di $C^1]-2, 2[$, qual è questa funzione? Ve ne sono altre? Che cosa succede se l'ostacolo ha equazione $y = 1 - |x|$?*

9.3 Estremi relativi e monotonia in un punto

In questo paragrafo analizziamo il comportamento della derivata in punti particolari, dove la funzione presenta degli estremi relativi e dove è monotona. Per le relative definizioni rimandiamo al Capitolo 3.

Teorema 9.9 (di Fermat) - *Se $x_0 \in A$ è un punto di minimo o di massimo relativo interno ad A e $f : A \to \mathbf{R}$ è derivabile in x_0, allora $f'(x_0) = 0$.*

Dimostrazione. Supponiamo che x_0 sia di minimo. Esiste un intorno U di x_0, per esempio un intervallo $I_\delta(x_0) =]x_0 - \delta, x_0 + \delta[$ che possiamo supporre tutto contenuto in A, tale che

$$f(x) \geqslant f(x_0) \qquad \forall x \in I_\delta(x_0)\,. \tag{9.8}$$

Tenendo conto della (9.3) e della (9.8) e passando al limite separatamente da destra e da sinistra si ottiene

$$f'(x_0) = \lim_{x \to x_0^+} \frac{f(x) - f(x_0)}{x - x_0} \geqslant 0 \quad \text{e} \quad f'(x_0) = \lim_{x \to x_0^-} \frac{f(x) - f(x_0)}{x - x_0} \leqslant 0\,,$$

da cui $f'(x_0) = 0$.

□

Definizione 9.10 - *Diciamo che $x_0 \in A$ è stazionario per f, o che f è stazionaria in x_0, se $f'(x_0) = 0$.*

La condizione che x_0 sia un estremo locale è solo sufficiente, ma non necessaria per la stazionarietà. Il punto $x = 0$ è stazionario per la funzione $f(x) = x^3$, ma non è né di minimo, né di massimo relativo per f. L'utilità del Teorema 9.9 sta nel fatto che gli estremi relativi interni a un dominio vanno ricercati tra i punti stazionari, oltreché tra quelli eventuali in cui f non è derivabile. Che cosa si può dire riguardo ai punti di frontiera? La funzione $f(x) = x$ su $[0, 1]$ ha minimo in 0 e massimo in 1 ma non è stazionaria in questi punti.

Esercizio 9.19 - *Facendo attenzione alla dimostrazione del Teorema 9.9, studiare il segno di $f'(x_0)$ se x_0, di minimo o di massimo relativo, è un estremo del dominio.*

Teorema 9.11 - *Se $f : A \to \mathbf{R}$ è crescente nel punto $x_0 \in A$ e derivabile in x_0 allora $f'(x_0) \geqslant 0$.*

Dimostrazione. Usiamo la definizione di funzione crescente in un punto insieme alla (9.3)

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(x_0) + \sigma(x) \geqslant 0\,, \qquad \lim_{x \to x_0} \sigma(x) = 0\,,$$

poi si passa al limite ottenendo $f'(x_0) \geqslant 0$.

□

Con ovvie modifiche si dimostra che $f'(x_0) \leqslant 0$ se f è decrescente in x_0.

Fin qui abbiamo usato informazioni puntuali sulla f per dedurre informazioni sulla f', però sarebbe più utile procedere nel senso opposto: studiare le proprietà di f' per ottenere informazioni sul comportamento di f. Ora, come il II teorema della permanenza del segno non è esattamente l'implicazione inversa del I, allo stesso modo per invertire l'enunciato del Teorema 9.11 non si può scambiare l'ipotesi con la tesi così come sono scritte, ma dobbiamo assumere la condizione più restrittiva che la derivata nel punto non si annulli e di conseguenza la monotonia diventa stretta.

Teorema 9.12 - *Se $f'(x_0) > 0$ allora f è strettamente crescente nel punto x_0.*

Dimostrazione. Essendo $f'(x_0)$ il limite del rapporto incrementale di f per $x \to x_0$, per il II teorema della permanenza del segno esiste $\delta > 0$ tale che

$$\frac{f(x) - f(x_0)}{x - x_0} > 0 \qquad \forall x \in]x_0 - \delta, x_0 + \delta[\,,$$

ma questa è esattamente la definizione di funzione crescente nel punto x_0.

□

Tuttavia, se anche fosse $f'(x) > 0$ per ogni $x \in A$, l'unica proprietà che possiamo dedurre per la f è che si tratti di una funzione crescente in ogni punto di A, ma non che sia crescente su A. In altre parole, senza ulteriori ipotesi, la monotonia rimane una proprietà puntuale. La funzione

$$f(x) = \begin{cases} x & \text{se } x \in [0, 1] \\ x - 2 & \text{se } x \in [2, 3] \end{cases}$$

è crescente in ogni punto ma non lo è su $[0,1] \cup [2,3]$. In quali circostanze allora la monotonia su tutto il dominio può essere dedotta dal segno della derivata in ogni punto, cioè dalla monotonia in ogni punto?

9.4 Funzioni derivabili su un intervallo

Se il dominio di f è un intervallo molte proprietà puntuali si trasformano facilmente in proprietà globali, come se le informazioni potessero "viaggiare" da una parte all'altra per l'assenza di vuoti. Gli intervalli sono infatti gli unici insiemi connessi di $\mathbf{R}$ e abbiamo già visto con le funzioni continue quanto questa situazione possa avere conseguenze tanto importanti.

Vediamo adesso tre teoremi fondamentali nei quali è essenziale l'ipotesi che il dominio sia un intervallo. Per il fatto di essere tra loro equivalenti si possono dimostrare in un ordine qualsiasi, a patto naturalmente che il primo venga dimostrato in maniera indipendente per non cadere in una tautologia.

Teorema 9.13 (di Rolle) - *Sia $f \in C^0[a,b]$, derivabile in $]a,b[$ e tale che $f(a) = f(b)$. Allora esiste $\xi \in]a,b[$ tale che $f'(\xi) = 0$.*

Dimostrazione. Per il Teorema di Weierstraß esistono $x_1, x_2 \in [a,b]$ tali che

$$f(x_1) \leqslant f(x) \leqslant f(x_2) \qquad \forall x \in [a,b]\,.$$

Se $f(x_1) = f(x_2)$ f è costante su $[a,b]$ con valore $k = f(a) = f(b)$, quindi $f'(x) = 0$ per ogni $x \in [a,b]$ e la tesi è verificata. Se f non è costante esiste un $\bar{x} \in]a,b[$ tale che

$$f(\bar{x}) < k \qquad \text{oppure} \qquad f(\bar{x}) > k\,.$$

Nel primo caso il punto di minimo x_1 è interno all'intervallo essendo

$$f(x_1) \leqslant f(\bar{x}) < k\,,$$

quindi $f'(x_1) = 0$ per il Teorema 9.9. Analogamente, nel secondo, sarà interno il punto di massimo x_2, per cui $f'(x_2) = 0$. Naturalmente potranno essere interni e stazionari entrambi. In ogni caso abbiamo trovato almeno un punto $\xi \in]a,b[$ stazionario per f. □

Nel Teorema di Rolle l'ipotesi che f sia continua fino agli estremi è essenziale. Come controesempio possiamo considerare la funzione $f : [-1,1] \to \mathbf{R}$ definita da

$$f(x) = \begin{cases} x & \text{se } x \in]-1,1[\\ 0 & \text{se } x = \pm 1\,. \end{cases}$$

Esercizio 9.20 - *Con l'aiuto del Corollario 8.36 dimostrare la seguente variante al Teorema di Rolle. Sia $f \in C^0]a,b[$, con $a,b \in \bar{\mathbf{R}}$, derivabile su $]a,b[$ e tale che*

$$\lim_{x\to a} f(x) = \lim_{x\to b} f(x) = L \in \bar{\mathbf{R}}\,.$$

Allora esiste un punto $\xi \in]a,b[$ tale che $f'(\xi) = 0$.

Possiamo adesso estendere il teorema degli zeri ad una classe più ampia delle funzioni continue, quelle che sono la derivata di una funzione derivabile.

Teorema 9.14 - *Siano $F : [a,b] \to \mathbf{R}$ una funzione derivabile e $f(x) = F'(x)$ per ogni $x \in [a,b]$. Se $f(a) < 0 < f(b)$ allora esiste $\xi \in [a,b]$ tale che $f(\xi) = 0$.*

Dimostrazione. Poiché $f(a)$ è il limite del rapporto incrementale di F in a e $f(a) < 0$, per il II teorema della permamenza del segno esiste $\delta > 0$ tale che

$$\frac{F(x) - F(a)}{x - a} < 0 \qquad \forall x \in]a, a + \delta[\,,$$

quindi $F(x) < F(a)$ per $a < x < a + \delta$. Ragionando nello stesso modo in b si ottiene $F(x) < F(b)$ per $b - \delta < x < b$. Allora il minimo di F (che esiste per il Teorema di Weierstraß) viene atteso in un punto interno $\xi \in]a, b[$ dove

$$f(\xi) = F'(\xi) = 0 \,.$$

□

Naturalmente per queste funzioni vale anche il teorema dei valori intermedi e tutte le proprietà che discendono dal teorema degli zeri.

Teorema 9.15 - *Siano $f, g \in C^0[a, b]$ due funzioni derivabili in $]a, b[$ tali che $g'(x) \neq 0$ per ogni $x \in]a, b[$. Allora esiste un punto $\xi \in]a, b[$ tale che*

$$\frac{f(b) - f(a)}{g(b) - g(a)} = \frac{f'(\xi)}{g'(\xi)} \,.$$

Dimostrazione. L'ipotesi $g'(x) \neq 0$ in ogni punto garantisce, per il Teorema di Rolle, che il denominatore $g(b) - g(a)$ non si annulli. La funzione

$$F(x) = f(x) - \lambda g(x)$$

soddisfa in $[a, b]$ le ipotesi del Teorema di Rolle per

$$\lambda = \frac{f(b) - f(a)}{g(b) - g(a)} \,.$$

Pertanto esiste $\xi \in]a, b[$ tale che

$$F'(\xi) = f'(\xi) - \frac{f(b) - f(a)}{g(b) - g(a)} g'(\xi) = 0 \,.$$

□

Teorema 9.16 (di Lagrange) - *Sia $f \in C^0[a, b]$ e derivabile in $]a, b[$. Allora esiste un punto $\xi \in]a, b[$ tale che*

$$\frac{f(b) - f(a)}{b - a} = f'(\xi) \,.$$

Dimostrazione. Basta scegliere $g(x) = x$ nel Teorema 9.15.

□

In termini geometrici, il Teorema 9.16, detto anche *Teorema del valor medio*, afferma che esiste un punto del grafico in cui la retta tangente è parallela alla corda passante per gli estremi.

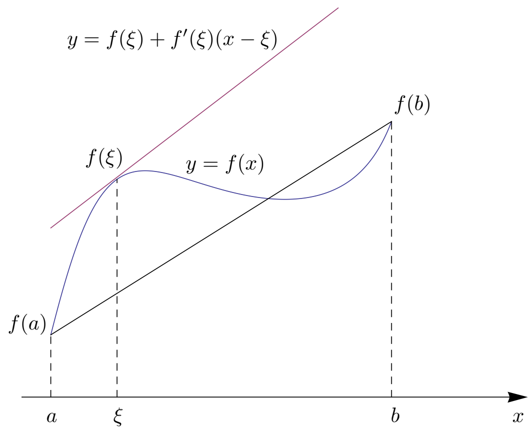


Figura 9.4: Il significato geometrico del valor medio.

Il Teorema di Lagrange fornisce anche la seguente versione alternativa della (9.4)

$$f(x) = f(x_0) + f'(\xi)(x - x_0)\,, \tag{9.9}$$

dove l'assenza dell'infinitesimo è in qualche modo compensata dalla valutazione della derivata in un punto ξ, opportunamente scelto nell'intervallo di estremi x_0 e x, che in generale sarà diverso da x_0.

Vediamo alcune importanti applicazioni del Teorema del valor medio.

Teorema 9.17 - *Sia I un intervallo di natura qualsiasi. Se $f : I \to \mathbf{R}$ è derivabile allora*

($\to$) $\quad f'(x) = 0 \quad \forall x \in I \quad \Rightarrow \quad f$ *è costante su I,*

($\nearrow$) $\quad f'(x) \geqslant 0 \quad \forall x \in I \quad \Rightarrow \quad f$ *è crescente su I,*

($\angle$) $\quad |f'(x)| \leqslant k \quad \forall x \in I \quad \Rightarrow \quad f$ *è lipschitziana con costante k su I.*

Dimostrazione. Scegliamo due punti qualsiasi $x_1, x_2 \in I$, con $x_1 < x_2$, e applichiamo il Teorema del valor medio all'intervallo $[x_1, x_2]$

$$\exists\, \xi \in [x_1, x_2] \ : \quad \frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(\xi)\,.$$

Allora

($\to$) $\quad f'(\xi) = 0 \quad \Rightarrow \quad f(x_1) - f(x_2)$,

($\nearrow$) $\quad f'(\xi) \geqslant 0 \quad \Rightarrow \quad f(x_1) \leqslant f(x_2)$,

($\angle$) $\quad |f'(\xi)| \leqslant k \quad \Rightarrow \quad |f(x_1) - f(x_2)| \leqslant k|x_1 - x_2|$.

Nella ($\nearrow$) si ottiene la monotonia stretta se $f' > 0$.

□

La funzione

$$f(x) = \operatorname{arctg} x + \operatorname{arctg}\frac{1}{x}\,, \qquad x \in \mathbf{R} - \{0\}\,,$$

ha derivata identicamente nulla su $\mathbf{R} - \{0\}$, ma non è costante, lo è invece su ciascuna delle due semirette $x > 0$ e $x < 0$ separatamente dove i valori che prende si calcolano subito:

$$f(1) = \frac{\pi}{2} \Rightarrow f(x) = \frac{\pi}{2} \quad \forall x > 0 \quad \text{e} \quad f(-1) = -\frac{\pi}{2} \Rightarrow f(x) = -\frac{\pi}{2} \quad \forall x < 0\,.$$

Definizione 9.18 - *Si chiama* ***primitiva*** *di una funzione* $f : I \to \mathbf{R}$ *ogni funzione* $F : I \to \mathbf{R}$ *tale che* $F'(x) = f(x)$ *per ogni* $x \in I$.

Il fatto che la derivata di una funzione costante sia nulla implica che se F è una primitiva di f allora anche $F + c$ lo è, infatti

$$(F + c)' = F' + c' = f\,,$$

ma il Teorema 9.17 ($\to$) garantisce, tenendo presente che il dominio è un intervallo, che due primitive della stessa funzione differiscono per una costante. Infatti se $F' = G' = f$ su I allora

$$D(F - G) = F' - G' = f - f = 0$$

quindi esiste $c \in \mathbf{R}$ tale che

$$F(x) - G(x) = c \qquad \forall x \in I\,.$$

La questione dell'esistenza di primitive per le funzioni continue verrà trattata nel Cap. 10, comunque adesso sappiamo che basta conoscerne una per conoscerle tutte.

Torniamo alla monotonia con un'importante osservazione. Abbiamo detto che se f è crescente in un punto x, o in ogni punto x, allora $f'(x) \geqslant 0$, disuguaglianza questa che non è detto valga in senso stretto neanche se f fosse strettamente crescente, si pensi al caso della funzione x^3 su $\mathbf{R}$, strettamente crescente in ogni punto, ma con derivata nulla in 0. Viceversa, per garantire la stretta monotonia, l'ipotesi $f'(x) > 0$ discussa nel Teorema 9.17 è fin troppo restrittiva, lo stesso esempio della funzione x^3 lo dimostra. Come possiamo indebolire questa ipotesi?

Corollario 9.19 - *Se* $f'(x) \geqslant 0$ *su un intervallo* I *e se* f' *si annulla al più su un insieme di punti isolati (quindi su un insieme finito se* I *è limitato) allora* f *è strettamente crescente.*

Dimostrazione. Supponiamo che f sia crescente su I, ma non strettamente. Allora esistono due punti $x_1, x_2 \in I$, con $x_1 < x_2$, tali che $f(x_1) = f(x_2)$, ma dal momento che $f(x_1) \leqslant f(x) \leqslant f(x_2)$ per $x \in [x_1, x_2]$, f sarebbe costante tra i due punti e la sua derivata sarebbe identicamente nulla sull'intero intervallo $[x_1, x_2]$. Questo contraddice l'ipotesi che sia nulla su un insieme di punti isolati.

□

Rivediamo adesso il Teorema 9.7, sulla derivata della funzione inversa, alla luce di quanto abbiamo detto sulla relazione tra la monotonia di f e il segno di f'. L'ipotesi $f'(x_0) \neq 0$ ci è servita in quel teorema per evitare che il rapporto incrementale della f^{-1} fosse divergente. Ma adesso ci possiamo chiedere se supporre ad esempio $f'(x_0) > 0$ non garantisca l'esistenza stessa di f^{-1} almeno in un intorno di x_0, dato che in questo caso f è strettamente crescente in x_0. La risposta è negativa perché la monotonia in un punto non implica la monotonia in un intorno. Per esempio la funzione

$$f(x) = \begin{cases} \dfrac{1}{2}x + x^2 \operatorname{sen} \dfrac{1}{x} & \text{se } x \neq 0 \\ 0 & \text{se } x = 0 \end{cases}$$

è strettamente crescente in 0 perché $f'(0) = 1/2 > 0$, ma non è crescente in nessun intorno di 0 perché $f'(1/2n\pi) = -1/2 < 0$ per ogni $n \in \mathbf{N} - \{0\}$.

Se invece si suppone che f sia derivabile in tutto un intorno di x_0 con derivata continua in x_0, se la derivata è positiva in x_0 rimane positiva su un intervallo $I_\delta(x_0) =]x_0 - \delta, x_0 + \delta[$, di conseguenza è crescente in senso stretto e quindi invertibile su I_δ. Vale quindi la seguente versione del Teorema 9.7.

Teorema 9.20 (di invertibilità locale) - *Sia $f : A \to \mathbf{R}$ una funzione derivabile con derivata continua nel punto $x_0 \in A$ e $f'(x_0) \neq 0$. Allora esiste un intorno U di x_0 sul quale f è invertibile con inversa $f^{-1} : f(U) \to U$ di classe C^1 e naturalmente*

$$\frac{d}{dy} f^{-1}(y) = \frac{1}{f'(f^{-1}(y))} \qquad \forall y \in f(U)$$

per il Teorema 9.7.

Osservazione 9.21 - *Se poi il dominio è un intervallo I e la derivata di f non si annulla mai, mantenendo quindi lo stesso segno, il Teorema 9.20 acquista carattere globale perché f è strettamente monotona su tutto I. In altre parole, una funzione $f : I \to \mathbf{R}$ derivabile con derivata non nulla in ogni punto, quindi localmente invertibile, è in realtà invertibile anche su I, cioè globalmente, con inversa differenziabile.*

Una funzione bigettiva e differenziabile con inversa differenziabile si chiama *diffeomorfismo*. Volendo esprimere l'Osservazione 9.21 con questa terminologia, ogni diffeomorfismo locale $f : I \to \mathbf{R}$ è anche un diffeomorfismo su I (globale). Vedremo che questo passaggio, dal carattere locale a quello globale, non si verifica in generale per le funzioni di più variabili, ma nemmeno per quelle di una sola variabile a valori vettoriali, cioè in $\mathbf{R}^n$ con $n > 1$. Consideriamo ad esempio la funzione $\varphi : \mathbf{R} \to \mathbf{R}^2$

$$t \to \varphi(t) = (\cos t, \operatorname{sen} t)$$

che "avvolge" tutta la retta reale infinite volte sulla circonferenza unitaria. Ovviamente non essendo iniettiva non ammette inversa, però è una bigezione differenziabile con derivata $\varphi'(t) = (-\operatorname{sen} t, \cos t)$, che non si annulla mai, dell'intorno $I_\delta(t_0)$ di ogni punto $t_0 \in \mathbf{R}$, con $\delta < \pi$, in un arco della circonferenza. Inoltre la sua inversa, che "appiattisce" quell'arco su $I_\delta(t_0)$, è differenziabile. Si tratta quindi di un diffeomorfismo locale, ma non globale.

Nel seguente teorema diamo alcune interessanti caratterizzazioni delle funzioni convesse con l'ipotesi che siano derivabili. Una di esse richiede l'uso del Teorema di Lagrange, ovviamente lecito dal momento che il dominio di una funzione convessa è per definizione un intervallo.

Teorema 9.22 - *Per una funzione $f : I \to \mathbf{R}$ derivabile le seguenti proprietà sono equivalenti:*

($\smile$)**1.** *f è convessa,*

($\smile$)**2.** $f(x) \geqslant f(x_0) + f'(x_0)(x - x_0) \quad \forall x, x_0 \in I$,

($\smile$)**3.** *f' è crescente,*

($\smile$)**4.** $f''(x) \geqslant 0 \quad \forall x \in I$ *se f ammette anche la derivata seconda in I.*

Dimostrazione. Una condizione necessaria e sufficiente per la convessità di f l'abbiamo già ottenuta nel Cap. 8 quando abbiamo ricavato la (8.2)

$$f(x) \geqslant f(x_0) + m(x - x_0)$$

che ci dice che f è convessa se e solo se il suo grafico sta al di sopra di ogni retta d'appoggio. Questa relazione vale per ogni $x, x_0 \in I$ e per ogni $m \in [f'_-(x_0), f'_+(x_0)]$. Ma se f è derivabile $f'_-(x_0) = f'_+(x_0) = f'(x_0)$, quindi l'equivalenza tra ($\smile$)1 e ($\smile$)2 è ovvia. Le infinite rette d'appoggio in un punto vengono a coincidere con la retta tangente, quindi le funzioni convesse e derivabili sono tutte e sole quelle funzioni il cui grafico rimane tutto al di sopra di ogni sua retta tangente.

Dimostriamo che $(\smile)2 \Rightarrow (\smile)3$. Scriviamo la $(\smile)2$ in due punti $x_1, x_2 \in I$ nelle due forme possibili

$$f(x_1) \geqslant f(x_2) + f'(x_2)(x_1 - x_2)\,,$$
$$f(x_2) \geqslant f(x_1) + f'(x_1)(x_2 - x_1)\,.$$

Sommando membro a membro si ottiene

$$0 \geqslant f'(x_2)(x_1 - x_2) + f'(x_1)(x_2 - x_1)$$

che equivale a

$$(f'(x_2) - f'(x_1))(x_2 - x_1) \geqslant 0\,.$$

Viceversa $(\smile)3 \Rightarrow (\smile)2$ per il Teorema 9.16. Infatti, presi due punti qualsiasi $x, x_0 \in I$, se $x > x_0$ esiste $\xi \in [x_0, x]$ tale che

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(\xi) \geqslant f'(x_0)\,,$$

mentre se $x < x_0$ esiste $\xi \in [x, x_0]$ tale che

$$\frac{f(x) - f(x_0)}{x - x_0} = f'(\xi) \leqslant f'(x_0)\,.$$

Moltiplicando a sinistra e a destra per $x - x_0$ si ottiene in entrambi i casi la $(\smile)2$.

L'equivalenza con la $(\smile)4$ è, a questo punto, ovvia per il Teorema 9.17.

$\square$

Esercizio 9.21 - *Per quali $a \in \mathbf{R}$ si ha $e^x \geqslant ax$ per ogni $x \in \mathbf{R}$?*

Corollario 9.23 - *Se $f : I \to \mathbf{R}$ è convessa e derivabile allora $f \in C^1(I)$.*

Dimostrazione. Per il Teorema 9.14 l'immagine di I tramite f', cioè $f'(I)$, è un intervallo, inoltre f' è crescente quindi continua per il Lemma 8.31 del Cap. 8.

$\square$

Il Teorema 9.22 suggerisce la seguente definizione di convessità puntuale.

Definizione 9.24 - *Una funzione $f : A \to \mathbf{R}$ viene detta convessa nel punto $x_0 \in A$ se esiste un intorno U di x_0 tale che*

$$f(x) \geqslant f(x_0) + f'(x_0)(x - x_0) \qquad \forall x \in U\,. \tag{9.10}$$

Chiaramente nel Teorema 9.22 si afferma, tra le altre cose, che sugli intervalli la convessità in ogni punto implica la convessità.

I punti del dominio in corrispondenza dei quali il grafico della funzione viene attraversato dalla retta tangente si chiamano punti di flesso. In altre parole, si tratta di punti in cui la funzione cambia la sua concavità, da concava diventa convessa, *flesso ascendente*, o, viceversa, da convessa diventa concava, *flesso discendente*. Tenendo presente la (9.10), la differenza, in questo caso, tra la funzione e la sua parte lineare nell'intorno di x_0 sarà, nei due rispettivi casi, concorde o discorde con $x - x_0$, come nella seguente definizione.

Definizione 9.25 - *Se $f : A \to \mathbf{R}$, diciamo che un punto $x_0 \in A$ è di flesso ascendente [discendente] per f se esiste un intorno U di x_0 tale che*

$$\frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} \geqslant 0 \quad [\leqslant 0] \qquad \forall x \in U\,. \tag{9.11}$$

Se f è due volte derivabile in x_0, come condizione necessaria per la (9.11) deve essere $f''(x_0) = 0$. La situazione è del tutto analoga a quella del Teorema di Fermat per f'. Infatti, poiché la (9.11) equivale ad affermare che l'espressione

$$\frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{(x - x_0)^2}$$

assume in U valori di segno opposto a sinistra e a destra di x_0, possiamo concludere che il suo limite per $x \to x_0$ deve essere nullo, ma questo limite vale $f''(x_0)/2$ come si può vedere applicando la seconda regola di l'Hôpital che vediamo tra poco.

Trattandosi di condizioni solo necessarie, come si può stabilire la natura di un punto stazionario per f, se è di massimo, di minimo ecc., o di un punto stazionario per f', se f è convessa, concava o ha flesso? Una possibilità è studiare il comportamento della funzione nell'intorno del punto: se f decresce a sinistra e cresce a destra significa che il punto è di minimo e, analogamente, se f è concava a sinistra e convessa a destra quel punto è di flesso ascendente. Nel prossimo paragrafo vedremo, come applicazione della formula di Taylor, che questo modo di procedere non è l'unico. Una risposta altrettanto esauriente, ma spesso operativamente più semplice, consiste in uno studio delle derivate successive calcolate soltanto nel punto in questione.

9.5 La regola di L'Hôpital e la formula di Taylor

Il seguente risultato vale con ipotesi opposte a quelle della Proposizione 9.4.

Teorema 9.26 (II regola di L'Hôpital) - *Sia I_δ l'intervallo $[x_0, x_0 + \delta[$, con $\delta > 0$, e siano $f, g : I_\delta \to \mathbf{R}$ due funzioni continue in x_0 e derivabili in $I_\delta - \{x_0\}$ tali che $f(x_0) = g(x_0) = 0$, $g(x) \neq 0$ e $g'(x) \neq 0$ per ogni $x \in I_\delta - \{x_0\}$. Se esiste il limite*

$$\lim_{x \to x_0} \frac{f'(x)}{g'(x)}$$

allora esiste anche il limite

$$\lim_{x \to x_0} \frac{f(x)}{g(x)}$$

e i due limiti sono uguali.

Dimostrazione. Indichiamo con $L \in \bar{\mathbf{R}}$ il limite del rapporto delle derivate e dimostriamo che per ogni successione $(x_n) \subset I_\delta$ tale che $x_n \to x_0$ si ha

$$\lim_{n \to \infty} \frac{f(x_n)}{g(x_n)} \to L\,.$$

Per il Teorema di Cauchy 9.15, per ogni $n \in \mathbf{N}$ esiste $\xi_n \in]0, x_n[$ tale che

$$\frac{f(x_n)}{g(x_n)} = \frac{f(x_n) - f(x_0)}{g(x_n) - g(x_0)} = \frac{f'(\xi_n)}{g'(\xi_n)}\,.$$

Per ipotesi

$$\lim_{n \to \infty} \frac{f'(\xi_n)}{g'(\xi_n)} = L$$

perché $\xi_n \to x_0$, dunque anche

$$\lim_{n \to \infty} \frac{f(x_n)}{g(x_n)} = L\,.$$

□

Esercizio 9.22 - *Con un opportuno cambio di variabile dimostrare che il Teorema 9.26 vale anche nel caso del limite per $x \to +\infty$.*

Il Teorema 9.26 continua a valere anche quando f e g divergono invece di tendere a 0, ma ne omettiamo la dimostrazione.

Osservazione 9.27 - *Non sempre è applicabile la regola di L'Hôpital, per esempio sappiamo che*

$$\lim_{x\to 0} \frac{x^2 \operatorname{sen}(1/x)}{\operatorname{sen} x} = 0\,, \qquad \lim_{x\to +\infty} \frac{x + \operatorname{sen} x}{x + \cos x} = 1\,,$$

mentre i rapporti fra la derivata del numeratore e del denominatore non hanno limite.

Poiché il Teorema 9.26 vale ovviamente anche su $]x_0-\delta, x_0]$, da ora in poi, salvo avviso contrario, sceglieremo come intorno U del punto x_0 un intervallo aperto di centro x_0.

Corollario 9.28 - *Sia $f : U \to \mathbf{R}$ continua in x_0 e derivabile in $U - \{x_0\}$. Se esiste finito il limite*

$$\lim_{x\to x_0} f'(x) = L$$

allora f è derivabile anche in x_0 e $f'(x_0) = L$.

Dimostrazione. Basta applicare il Teorema 9.26 al rapporto incrementale

$$R(x, x_0) = \frac{f(x) - f(x_0)}{x - x_0}$$

osservando che il rapporto tra le derivate delle funzioni $f(x) - f(x_0)$ e $x - x_0$ coincide proprio con $f'(x)$.

□

In altre parole, il corollario afferma che se f è continua in x_0 basta che il limite per $x \to x_0$ della derivata esista e sia finito affinché questa sia anche continua in x_0 ed esclude la possibilità che il limite esista ma sia diverso dal valore della derivata nel punto.

Con il Teorema 9.26 possiamo spingere fino al II ordine l'espansione asintotica

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + o((x - x_0)^2) \tag{9.12}$$

se f naturalmente è sufficientemente regolare, basta applicare la II regola di L'Hôpital al calcolo del limite

$$\lim_{x\to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{(x - x_0)^2}\,.$$

La (9.12) è del tutto analoga alla (9.4), la parabola che meglio approssima il grafico di f nell'intorno di x_0 è il grafico della funzione quadratica che compare a II membro della (9.12). Il passo successivo consiste nel trovare la cubica ottimale: si porta a I membro della (9.12) la parte quadratica, si divide per $(x - x_0)^3$ e si passa al limite per $x \to x_0$. Per il Teorema 9.26 tale limite esiste se supponiamo che f sia derivabile tre volte in x_0. Infatti

$$\begin{aligned} &\lim_{x\to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0) - \frac{1}{2} f''(x_0)(x - x_0)^2}{(x - x_0)^3} \\ &= \lim_{x\to x_0} \frac{f'(x) - f'(x_0) - f''(x_0)(x - x_0)}{3(x - x_0)^2} = \lim_{x\to x_0} \frac{f''(x) - f''(x_0)}{6(x - x_0)} = \frac{f'''(x_0)}{6}\,. \end{aligned}$$

Ciò equivale a

$$\frac{f(x) - f(x_0) - f'(x_0)(x - x_0) - \dfrac{f''(x_0)}{2}(x - x_0)^2}{(x - x_0)^3} = \frac{f'''(x_0)}{6} + \sigma(x)$$

dove σ è un infinitesimo, e così si ottiene infine lo sviluppo di f fino al III ordine

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{f''(x_0)}{2}(x - x_0)^2 + \frac{f'''(x_0)}{6}(x - x_0)^3 + o((x - x_0)^3)\,.$$

Applicando lo stesso ragionamento ad ogni passo fino alla derivata n-esima, si perviene ad una rappresentazione approssimata della f con un polinomio P_n di grado n, il *polinomio di Taylor*, che differisce da f per un infinitesimo r_n di ordine superiore a $(x - x_0)^n$, il *resto di Peano.*

Teorema 9.29 (Formula di Taylor con il resto di Peano) - *Se $f \in C^{n-1}(U)$ ed esiste $f^n(x_0)$ allora*

$$f(x) = \sum_{k=1}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + r_n(x)$$

dove $r_n(x) = o((x - x_0)^n)$.

Dimostrazione. Bisogna dividere $r_n = f - P_n$ per $(x - x_0)^n$ e vedere se tende a 0, ma questo equivale a mostrare che

$$\lim_{x \to x_0} \frac{f(x) - P_{n-1}(x)}{(x - x_0)^n} = \frac{f^{(n)}(x_0)}{n!}\,.$$

Applichiamo la II regola di L'Hôpital

$$\begin{aligned}
\lim_{x \to x_0} \frac{f(x) - P_{n-1}(x)}{(x - x_0)^n} &= \lim_{x \to x_0} \frac{f'(x) - P'_{n-1}(x)}{n(x - x_0)^{n-1}} = \lim_{x \to x_0} \frac{f''(x) - P''_{n-1}(x)}{n(n-1)(x - x_0)^{n-2}} \\
&= \ldots = \lim_{x \to x_0} \frac{f^{(k)}(x) - P^{(k)}_{n-1}(x)}{n(n-1)\ldots(n-k+1)(x - x_0)^{n-k}} \\
&= \ldots = \lim_{x \to x_0} \frac{f^{(n-1)}(x) - P^{(n-1)}_{n-1}(x)}{n!(x - x_0)}\,.
\end{aligned}$$

Poiché $P^{(n-1)}_{n-1}(x)$ è costante e vale $f^{(n-1)}(x_0)$, essendo f derivabile n volte in x_0, l'ultimo limite esiste e vale $f^{(n)}(x_0)/n!$. Allora, per il Teorema 9.26, esistono ed hanno lo stesso valore tutti i limiti precedenti fino a quello iniziale. Esiste dunque un infinitesimo σ per $x \to x_0$ tale che

$$\begin{aligned}
f(x) &= P_{n-1}(x) + \left[\frac{f^{(n)}(x_0)}{n!} + \sigma(x)\right](x - x_0)^n \\
&= \sum_{k=1}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + r_n(x)\,,
\end{aligned}$$

dove il resto n-esimo r_n è un infinitesimo di ordine superiore a $(x - x_0)^n$ per $x \to x_0$. □

Il risultato ottenuto non è altro che l'analogo della (9.4) all'n-esimo ordine. Vediamo come si applica questo risultato allo studio della convessità puntuale con le relative conseguenze sulla natura dei punti stazionari.

Corollario 9.30 - *Sia $f \in C^{n-1}(U)$ una funzione derivabile n volte nel punto x_0 tale che*

$$f''(x_0) = f'''(x_0) = \ldots = f^{(n-1)}(x_0) = 0 \quad e \quad f^{(n)}(x_0) \neq 0\,.$$

Allora

$$\text{se } n \text{ è pari} \qquad \begin{cases} f^{(n)}(x_0) > 0 \Rightarrow f \text{ convessa in } x_0 \\ f^{(n)}(x_0) < 0 \Rightarrow f \text{ concava in } x_0 \end{cases}$$

$$\text{e se } n \text{ è dispari} \qquad \begin{cases} f^{(n)}(x_0) > 0 \Rightarrow x_0 \text{ di flesso ascendente per } f \\ f^{(n)}(x_0) < 0 \Rightarrow x_0 \text{ di flesso discendente per } f\,. \end{cases}$$

Dimostrazione. Per le ipotesi fatte, la formula di Taylor si riduce a

$$f(x) - f(x_0) - f'(x_0)(x - x_0) = \left[\frac{f^{(n)}(x_0)}{n!} + \sigma(x)\right](x - x_0)^n$$

in U. Poiché σ è infinitesima per $x \to x_0$, se U è abbastanza piccolo il segno della parentesi quadra coincide col segno di $f^{(n)}(x_0)$. Se n è pari $(x - x_0)^n > 0$ in $U - \{x_0\}$ e quindi

$$f^{(n)}(x_0) > 0 \Rightarrow f(x) - f(x_0) - f'(x_0)(x - x_0) > 0 \;\; \forall x \in U - \{x_0\}\,,$$

$$f^{(n)}(x_0) < 0 \Rightarrow f(x) - f(x_0) - f'(x_0)(x - x_0) < 0 \;\; \forall x \in U - \{x_0\}\,.$$

Nel primo caso f è strettamente convessa e nel secondo strettamente concava. Se n è dispari dividiamo per $x - x_0$ ottenendo in $U - \{x_0\}$

$$\frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} = \left[\frac{f^{(n)}(x_0)}{n!} + \sigma(x)\right](x - x_0)^{n-1}\,.$$

Adesso $(x - x_0)^{n-1} > 0$, quindi

$$f^{(n)}(x_0) > 0 \Rightarrow \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} > 0 \;\; \forall x \in U - \{x_0\}\,,$$

$$f^{(n)}(x_0) < 0 \Rightarrow \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} < 0 \;\; \forall x \in U - \{x_0\}\,.$$

Nel primo caso x_0 è di flesso ascendente per f e nel secondo discendente.

□

In particolare, supponiamo in più $f'(x_0) = 0$. Combinando questa condizione aggiuntiva con i primi due casi si ottiene rispettivamente che x_0 è di minimo o di massimo relativo isolato, con gli altri due si ottiene invece che f è strettamente crescente o decrescente in x_0.

Il caso della funzione di classe C^∞

$$f(x) = \begin{cases} e^{-1/x^2} & \text{se } x \neq 0 \\ 0 & \text{se } x = 0\,, \end{cases}$$

che ammette $x = 0$ come punto di minimo relativo isolato nel quale le derivate tutte nulle, mostra che le condizioni enunciate nel Corollario 9.30 sono solo sufficienti.

Osservazione 9.31 - *Usando il Teorema 9.29 possiamo generalizzare la I regola di L'Hôpital. Se $f^{(k)}(x_0) = 0$ per $k < m$ e $f^{(m)}(x_0) \neq 0$ e se $g^{(k)}(x_0) = 0$ per $k < n$ e $g^{(n)}(x_0) \neq 0$ allora*

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \lim_{x \to x_0} \frac{\dfrac{f^{(m)}(x_0)}{m!}(x - x_0)^m + o((x - x_0)^m)}{\dfrac{g^{(n)}(x_0)}{n!}(x - x_0)^n + o((x - x_0)^n)}$$

$$= \frac{n!\,f^{(m)}(x_0)}{m!\,g^{(n)}(x_0)} \lim_{x \to x_0} (x - x_0)^{m-n}$$

che si riduce a

$$\lim_{x\to x_0}\frac{f(x)}{g(x)}=\frac{f^{(n)}(x_0)}{g^{(n)}(x_0)}$$

per $n=m$, quando le due funzioni sono infinitesime dello stesso ordine. Se $n\neq m$ poco interessa la presenza di $n!$ e $m!$ dal momento che il limite è 0 o ∞. Ad esempio

$$\lim_{x\to 0}\frac{\operatorname{tang}x}{x-\operatorname{sen}x}=\lim_{x\to 0}\frac{1}{\cos^2 x(1-\cos x)}=+\infty$$

anche se

$$\frac{\operatorname{tang}x}{x-\operatorname{sen}x}\sim\frac{6x}{x^3}\quad e\quad \frac{1}{\cos^2 x(1-\cos x)}\sim\frac{2}{x^2}\,.$$

In questa forma l'Osservazione 9.31 ricade anche all'interno del Teorema 9.26 del quale possiamo considerarla un miglioramento per il fatto di fornire esplicitamente il valore del limite. Non ne è tuttavia del tutto equivalente, ma i soli casi che di esso non contempla sono quelli in cui una almeno delle due funzioni ha le derivate tutte nulle in x_0. Si tratta d'altra parte di casi ai quali il Teorema 9.26, pur potendosi teoricamente applicare, è di fatto inservibile.

Vediamo adesso un'altra versione della formula di Taylor che generalizza la (9.9) all'n-esimo ordine. Al posto del resto di Peano, nell'ultimo termine dello sviluppo si calcola la derivata in un punto opportuno dell'intorno di x_0. Quel termine si chiama *resto di Lagrange*.

Teorema 9.32 (Formula di Taylor col resto di Lagrange) - *Se $f\in C^n(U)$ allora per ogni $x\in U$ esiste $\xi\in U$ tale che*

$$f(x)=\sum_{k=0}^{n-1}\frac{f^{(k)}(x_0)}{k!}(x-x_0)^k+\frac{f^{(n)}(\xi)}{n!}(x-x_0)^n\,.$$

Dimostrazione. Ricordiamo che $P_n^{(k)}(x_0)=f^{(k)}(x_0)$, quindi

$$\frac{f(x)-P_{n-1}(x)}{(x-x_0)^n}=\frac{f(x)-P_{n-1}(x)-(f(x_0)-P_{n-1}(x_0))}{(x-x_0)^n-(x_0-x_0)^n}\,.$$

A questa espressione è applicabile il Teorema di Cauchy: supponendo ad esempio $x>x_0$, esiste $\xi_1\in[x_0,x[$ tale che

$$\frac{f(x)-P_{n-1}(x)}{(x-x_0)^n}=\frac{f'(\xi_1)-P'_{n-1}(\xi_1)}{n(\xi_1-x_0)^{n-1}}\,.$$

Per lo stesso motivo, aggiungendo come prima i termini nulli $f'(x_0)-P'_{n-1}(x_0)$ al numeratore e $n(x_0-x_0)^{n-1}$ al denominatore, esiste $\xi_2\in[x_0,\xi_1[$ tale che

$$\frac{f'(\xi_1)-P'_{n-1}(\xi_1)}{n(\xi_1-x_0)^{n-1}}=\frac{f''(\xi_2)-P''_{n-1}(\xi_2)}{n(n-1)(\xi_2-x_0)^{n-2}}\,.$$

Ripetendo lo stesso ragionamento ad ogni passo, troviamo $n-1$ punti $\xi_k\in[x_0,x[$, $\xi_k>\xi_{k+1}$, tali che

$$\begin{aligned}\frac{f(x)-P_{n-1}(x)}{(x-x_0)^n}&=\ldots=\frac{f^{(k)}(\xi_k)-P_{n-1}^{(k)}(\xi_k)}{n(n-1)\ldots(n-k+1)(\xi_k-x_0)^{n-k}}\\&=\ldots=\frac{f^{(n-1)}(\xi_{n-1})-P_{n-1}^{(n-1)}(\xi_{n-1})}{n!(\xi_{n-1}-x_0)}=\frac{f^{(n-1)}(\xi_{n-1})-f^{(n-1)}(x_0)}{n!(\xi_{n-1}-x_0)}\end{aligned}$$

essendo $P_{n-1}^{(n-1)}(x) = f^{(n-1)}(x_0)$ per ogni $x \in \mathbf{R}$. Non rimane che l'ultimo passo

$$\frac{f^{(n-1)}(\xi_{n-1}) - f^{(n-1)}(x_0)}{n!(\xi_{n-1} - x_0)} = \frac{f^{(n)}(\xi)}{n!}$$

per un certo $\xi \in [x_0, \xi_{n-1}[$. In conclusione abbiamo ottenuto

$$\frac{f(x) - P_{n-1}(x)}{(x - x_0)^n} = \frac{f^{(n)}(\xi)}{n!}$$

che è la tesi.

□

Il resto di Lagrange

$$r_{n-1}(x) = \frac{f^{(n)}(\xi)}{n!}(x - x_0)^n\,, \qquad x \in I_\delta\,,$$

si presta in modo particolarmente efficace alla valutazione numerica delle funzioni, tanto più precisa quanto più è piccolo l'intervallo a cui appartiene il punto ξ. Una stima della derivata n-esima su tale intervallo ci permette di controllare l'errore, cioè il resto, con una maggiorazione del tipo

$$|r_{n-1}(x)| \leqslant \frac{M}{n!}|x - x_0|^n\,,$$

basta scegliere δ e M in modo che $|f^{(n)}(x)| \leqslant M$ per ogni $x \in I_\delta$.

Se per esempio dobbiamo calcolare le prime tre cifre decimali di $\cos 1$, sviluppiamo $\cos x$ intorno a 0 fino al generico ordine n

$$\cos x = 1 - \frac{x^2}{2} + \ldots \pm \frac{D^n \cos\xi}{n!}x^n\,,$$

poi osserviamo che l'ultimo termine scritto, $r_{n-1}(x)$, soddisfa

$$|r_{n-1}(x)| \leqslant \frac{|x|^n}{n!} \qquad \text{e} \qquad |r_{n-1}(1)| \leqslant \frac{1}{n!}\,.$$

Il più piccolo $n \in \mathbf{N}$ tale che $1/n! < 10^{-3}$ è $n = 7$ e quindi possiamo trascurare il resto $r_6(1) \leqslant 1/7! = 1/5.040$. Si ottiene pertanto

$$\cos 1 = 1 - \frac{1}{2} + \frac{1}{4!} - \frac{1}{6!} = \frac{389}{720} = 0.540$$

e altre cifre decimali che non ci interessano.

Dimostriamo che

$$1 - \frac{x^2}{6} < \frac{\operatorname{sen} x}{x} < 1 - \frac{x^2}{6} + \frac{x^4}{120} \qquad \forall x \in \left]-\frac{\pi}{2}, \frac{\pi}{2}\right[\,.$$

Dato che si tratta di una funzione pari ragioniamo solo per $x > 0$. Per ogni $x \in]0, \pi/2[$ esiste $\xi \in]0, x[$ tale che

$$\operatorname{sen} x = x - \frac{x^3}{6} + \frac{\cos\xi}{120}x^5$$

ed essendo $0 < \cos\xi < 1$ si ottiene

$$x - \frac{x^3}{6} < \operatorname{sen} x < x - \frac{x^3}{6} + \frac{x^5}{120}\,.$$

9.6 Il metodo delle corde e il metodo delle tangenti

Fra i tanti metodi di approssimazione degli zeri di una funzione continua, oltre a quello della bisezione che abbiamo usato nel Teorema 8.28 degli zeri, ne abbiamo scelti un paio (i più noti tra l'altro) per il loro significato geometrico particolarmente chiaro ed efficace. Per approfondire l'argomento si veda ad esempio questo documento di Dario Bini, Università di Pisa, dove viene trattata anche la questione dell'ordine di convergenza.

Il metodo delle corde - Si sceglie un intervallo $[a, b]$ tale che $f(a)f(b) < 0$ che contenga un solo zero della f, il che è sicuramente possibile se f è strettamente monotona in una zona intorno allo zero, e su cui f sia strettamente convessa o concava. Supponiamo, tanto per fissare un caso, che f sia crescente e convessa. Per $a < x < b$ si ha

$$f(x) < \varphi_1(x) = f(a) + \frac{f(b) - f(a)}{b - a}(x - a)$$

dove φ_1 si annulla nel punto

$$x_1 = \frac{af(b) - bf(a)}{f(b) - f(a)}\,,$$

quindi per la relazione precedente, di convessità, $f(x_1) < 0$. Allora ripetiamo il procedimento per $x_1 < x < b$

$$f(x) < \varphi_2(x) = f(x_1) + \frac{f(b) - f(x_1)}{b - x_1}(x - x_2)$$

dove φ_2 si annulla nel punto

$$x_2 = \frac{x_1 f(b) - bf(x_1)}{f(b) - f(x_1)}\,,$$

dove $f(x_2) < 0$ per lo stesso motivo di prima. Si genera in questo modo la successione ricorsiva

$$x_{n+1} = \frac{x_n f(b) - bf(x_n)}{f(b) - f(x_n)} \tag{9.13}$$

dove per induzione $f(x_n) < 0$ per ogni $n \in \mathbf{N}$ e (x_n) cresce, quindi converge, in quanto

$$x_{n+1} - x_n = \frac{f(x_n)(x_n - b)}{f(b) - f(x_n)} > 0\,.$$

Se x_0, necessariamente minore di b, è il limite della (x_n), per continuità $f(x_n) \to f(x_0)$ e passando al limite nella (9.13) si ottiene $f(x_0) = 0$.

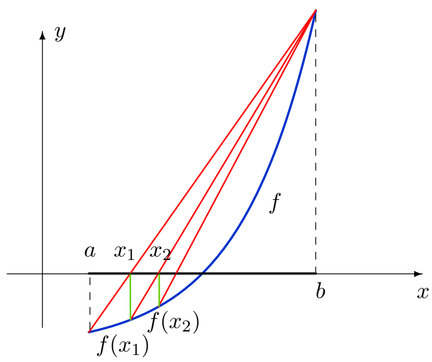


Figura 9.5: Il metodo delle corde.

Il metodo delle tangenti - È noto anche come metodo di Newton-Raphson. Sia $f \in C^1[a,b]$ strettamente convessa tale che $f'(x) > 0$ per ogni $x \in [a,b]$. Per $a < x < b$ si ha

$$f(x) > \psi_1(x) = f(b) + f'(b)(x-b)$$

dove ψ_1 si annulla nel punto

$$x_1 = b - \frac{f(b)}{f'(b)}$$

e $f(x_1) > 0$. Ripetiamo il procedimento per $a < x < x_1$

$$f(x) > \psi_2(x) = f(x_1) + f'(x_1)(x - x_1)$$

dove ψ_2 si annulla nel punto

$$x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}$$

dove $f(x_2) > 0$. Si genera in questo modo la successione ricorsiva

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{9.14}$$

dove $f(x_n) > 0$ per ogni $n \in \mathbf{N}$ e (x_n) decresce quindi converge. Se $x_0 > a$ è il limite della (x_n), per continuità $f(x_n) \to f(x_0)$ e passando al limite nella (9.14), dove per le ipotesi fatte la successione $(f'(x_n))$ non può tendere a 0, si ottiene $f(x_0) = 0$.

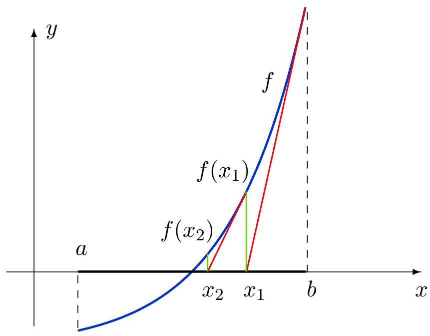


Figura 9.6: Il metodo delle tangenti.

9.7 Studio qualitativo del grafico

Allo scopo di mettere in evidenza le proprietà essenziali e il comportamento qualitativo di una funzione reale di una variabile reale, daremo nei punti che seguono delle indicazioni generali sulle fasi più importanti attraverso cui tale studio deve procedere. Si tenga presente, tuttavia, che queste indicazioni non vanno prese come uno schema rigido, ma devono essere adattate di volta in volta alle peculiarità della funzione in esame. Tutte le informazioni che questa analisi fornisce verranno infine raccolte e riassunte in un disegno (chiaro e senza pasticci, su un foglio a quadretti) che ne rappresenta il grafico. Ricordiamo che il grafico di una funzione $f : A \to \mathbf{R}$, dove $A \subset \mathbf{R}$, è il sottoinsieme di $\mathbf{R}^2$ definito da

$$G(f) = \{(x,y) \in \mathbf{R}^2 \mid x \in A,\ y = f(x)\}.$$

Esso sarà generalmente formato da una curva o da un numero finito di curve da rappresentare rispetto ad una coppia di assi cartesiani ortogonali, dei quali uno, l'asse x, contiene il dominio A e l'altro, l'asse y, l'immagine $f(A)$. L'unità di misura verrà scelta in base all'estensione della funzione, e dovrà essere abbastanza grande, compatibilmente con lo spazio disponibile, in modo che tutte le proprietà siano leggibili con chiarezza.

1. Data l'espressione che definisce la funzione f, verrà subito scelto come dominio A il più grande sottoinsieme di $\mathbf{R}$ su cui tale espressione è ben definita, cioè il *dominio naturale* di f. Di solito si tratta di un intervallo o di una unione di più intervalli separati, limitati o non limitati. È bene specificare il grado di regolarità di f, cioè se f è continua e a quale classe $C^k(A)$ appartiene, se è derivabile infinite volte su tutto A o se lo è in A eccetto qualche punto dove è discontinua oppure derivabile fino ad un certo ordine. Eventuali simmetrie permettono di limitare lo studio ad una sola parte di A, ad esempio ad $A^+ = \{x \in A \mid x \geqslant 0\}$ se f è una funzione pari o dispari con A simmetrico rispetto a 0, o ad un intervallo di ampiezza T se f è periodica di periodo T; il prolungamento pari, dispari o periodico di questa restrizione a tutto l'insieme A sarà la f assegnata.

2. Si cerca di calcolare o di dare una stima del valore di $f(0)$ che è l'ordinata dell'intersezione con l'asse y, e degli zeri di f, cioè dei punti $x_i \in A$ tali che $f(x_i) = 0$, in corrispondenza dei quali il grafico attraversa o tocca l'asse x. Se ciò è possibile si potrà anche stabilire il segno di f. Ad esempio, se f è continua sull'intervallo I e $x_1, x_2 \in I$ sono due zeri consecutivi, se scegliendo a caso un punto $x_0 \in]x_1, x_2[$ si trova che $f(x_0) > 0$, allora si può concludere che $f(x) > 0$ per ogni $x \in]x_1, x_2[$.

3. Sia I uno dei massimi intervalli, di estremi $a, b \in \bar{\mathbf{R}}$ con $a < b$, contenuti in A. Se $a \in I$ e/o $b \in I$ e $f \in C^0(I)$ basta calcolare $f(a)$ e $f(b)$, altrimenti bisogna calcolare anche i limiti di f per $x \to a^+$ e/o per $x \to b^-$. In un punto in cui f presenta un salto deve essere calcolato il limite destro e il limite sinistro.

4. Se x_0 è un numero reale, punto di accumulazione per A, il dominio di f, e uno dei due limiti, da una parte o dall'altra o da entrambe, non è finito, allora la retta di equazione $x = x_0$ è *asintoto verticale* per f. La conoscenza del segno della funzione può essere utile per stabilire se questo il limite è $+\infty$ o $-\infty$. Se A non è limitato, per esempio superiormente, e il limite

$$a = \lim_{x \to +\infty} f(x)$$

esiste finito, la retta di equazione $y = a$ è *asintoto orizzontale* per f a $+\infty$. Se questo limite esiste ma non è finito, f potrebbe avere un *asintoto obliquo*, situazione che si verifica se e solo se esistono finiti i limiti

$$\alpha = \lim_{x \to +\infty} \frac{f(x)}{x} \; (= \lim_{x \to +\infty} f'(x) \text{ quando esiste}) \qquad \text{e} \qquad \beta = \lim_{x \to +\infty} (f(x) - \alpha x)$$

e l'asintoto ha equazione $y = \alpha x + \beta$. L'asintoto orizzontale rientra in questo come caso particolare con $\alpha = 0$. Si avrà cura di disegnare, magari con un tratto più sottile, anche gli asintoti, mettendone in evidenza gli eventuali punti di intersezione col grafico. Verso di essi la curva verrà fatta avvicinare dalla parte giusta e con la giusta rapidità.

5. Dove f è derivabile si calcola f' e i suoi zeri e se f' è continua lo studio del suo segno può essere evitato perché, come si è detto per f, tra due zeri consecutivi il segno non cambia. L'analisi della derivata e dei suoi zeri ci permette di individuare gli estremi locali interni al dominio. Per stabilire la natura di un punto stazionario x_0, talvolta può essere più conveniente calcolare le derivate successive di f in x_0 finché $f^{(n)}(x_0) \neq 0$ e concludere sulla base di noti risultati. Inoltre f può presentare dei massimi o dei minimi locali in corrispondenza dei quali non è derivabile; anche gli stessi estremi a e b di I possono essere punti di massimo o di minimo relativo o assoluto senza che la derivata esista, oppure con derivata non nulla. Basterà allora studiare il segno delle derivate destre e/o sinistre. Infine dovranno essere calcolati i valori di f in tutti questi punti in modo da stabilire gli estremi assoluti per confronto.

Per avere un'idea precisa sulla direzione che prende il grafico, conviene calcolare la derivata in qualche punto significativo, ad esempio nei punti di attraversamento degli assi o degli asintoti o alle estremità del dominio. Dove la derivata destra ha valore diverso dalla derivata sinistra verranno messe in evidenza le due tangenti distinte, e nei punti del dominio in cui il limite della derivata o il limite del rapporto incrementale non è finito verrà disegnata una retta tangente verticale.

6. Dove f' è derivabile si calcola f'' che va studiata in modo analogo a f' per stabilire dove f è convessa, dove è concava e dove presenta dei flessi. Anche nei punti di flesso può essere utile calcolare il valore di f per sapere dove disegnarli, e di f' per sapere come sono orientati. Il segno di f'' a sinistra e a destra, oppure il segno di f''' nel punto di flesso, indicano se si tratta di un flesso ascendente o discendente.

Raccolte tutte queste informazioni, si procede al disegno del grafico. Saranno molto apprezzate tutte quelle osservazioni da cui si possano dedurre delle previsioni sul comportamento di f. Esse sono di grande utilità perché ci permettono di evitare noiosi calcoli, inoltre, anche quando questi sono inevitabili, ci chiariscono che cosa stiamo cercando. Situazioni tipiche sono quelle riconducibili al teorema degli zeri per le funzioni continue o ai suoi corollari, al teorema di Weierstraß, e varianti, che risulta particolarmente efficace per lo studio dei massimi e minimi se combinato con considerazioni sulla convessità, oppure riconducibili a risultati generali sulla monotonia e la convessità che non coinvolgano il calcolo, talvolta troppo complicato, della derivata seconda. In ogni caso è sempre bene tenere presenti tutti quei risultati, che discendono dai teoremi del tipo "valor medio" (Rolle, Lagrange, Cauchy), sulle proprietà delle funzioni derivabili su un intervallo. Il caso in esame suggerisce volta per volta gli aspetti più interessanti e degni di nota.

9.8 Funzioni analitiche

Questo argomento è stato introdotto nel § 6.5 del Cap. 6, si riveda la Definizione 6.17 e si tengano ben presenti gli esempi presentati e commentati in quella sede. Ci limitiamo qui a ricordare, per chiarezza, solo alcuni fatti importanti di quanto si è già detto per poi dedurre ulterriori proprietà con gli strumenti del calcolo differenziale.

Una serie di potenze di centro $z_0 \in \mathbf{C}$ può convergere assolutamente in ogni punto di un cerchio aperto $B_R(z_0) = \{z \in \mathbf{C} \mid |z - z_0| < R\}$ oppure assolutamente in ogni punto $z \in \mathbf{C}$, non consideriamo il caso banale e poco interessante della convergenza nel solo punto z_0. La somma della serie in ogni punto z del dominio di convergenza è una funzione analitica $f(z)$. Abbiamo osservato che anche se la f si estende al di fuori di $B_R(z_0)$, su un dominio A più grande, questo rimane comunque il massimo cerchio di convergenza della serie perché altrimenti verrebbe a contenere dei punti singolari di f. Però possiamo sempre scegliere un altro punto $z_0' \in A$ come centro di un'altra serie di potenze con raggio di convergenza R' che converge a f, nello stesso senso di prima, su $B_{R'}(z_0')$. Naturalmente sull'eventuale intersezione dei due cerchi la somma f sarà la stessa, sebbene con una diversa rappresentazione. Procedendo sempre in questo modo si arriva a ricoprire tutto il dominio A con l'unione dei cerchi così trovati. È sufficiente conoscere la f su un primo cerchio di convergenza perché sia univocamente determinata anche al di fuori di esso, fino a un dominio "massimo", sul quale viene chiamata *prolungamento analitico*. Abbiamo visto infatti vari esempi di sviluppi in serie, tra i quali

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

sul cerchio unitario di centro 0, ma accanto a questi vi sono altri sviluppi su altri cerchi per le stesse funzioni, come

$$\frac{1}{1-z} = \sum_{n=0}^{\infty} \frac{(z-z_0)^n}{(1-z_0)^{n+1}} \qquad \text{e} \qquad \sum_{n=0}^{\infty} \frac{(-1)^n[(z_0+i)^{n+1}+(z_0-i)^{n+1}]}{2i(z_0^2+1)^{n+1}}(z-z_0)^n\,.$$

Tornando alla serie geometrica, abbiamo visto nell'Esempio 6.8 che

$$\frac{1}{(1-z)^2} = \sum_{n=1}^{\infty} nz^{n-1}\,, \qquad |z|<1\,, \tag{9.15}$$

come applicazione del prodotto di Cauchy (6.20), ma adesso che conosciamo le derivate, la (9.15) trova una spiegazione alternativa

$$\frac{1}{(1-z)^2} = \frac{d}{dz}\frac{1}{1-z} = \frac{d}{dz}\sum_{n=0}^{\infty} z^n = \sum_{n=0}^{\infty} \frac{dz^n}{dz} = \sum_{n=1}^{\infty} nz^{n-1}\,.$$

Si noti che la definizione di derivata che abbiamo dato per le funzioni di variabile reale vale anche per quelle di variabile complessa

$$f'(z) = \lim_{h\to 0} \frac{f(z+h)-f(z)}{h}$$

con le stesse conseguenze sulle regole di derivazione e sulle derivate delle funzioni elementari. Ricordiamo che anche i risultati sulla convergenza delle serie di potenze sono gli stessi, differisce solo l'interpretazione che si deve dare al dominio di convergenza che in $\mathbf{R}$ è un intervallo e in $\mathbf{C}$ è un cerchio.

La regola di derivazione di una somma, che da quanto ci risulta si applica soltanto alle somme finite, evidentemente sembra funzionare anche con le serie, o almeno con tutte le serie di potenze relative alle funzioni elementari, per esempio

$$\frac{de^z}{dz} = \frac{d}{dz}\sum_{n=0}^{\infty}\frac{z^n}{n!} = \sum_{n=1}^{\infty}\frac{nz^{n-1}}{n!} = \sum_{n=1}^{\infty}\frac{z^{n-1}}{(n-1)!} = \sum_{n=0}^{\infty}\frac{z^n}{n!} = e^z\,,$$

oppure

$$\begin{aligned}\frac{d}{dz}\operatorname{sen} z &= \frac{d}{dz}\sum_{n=0}^{\infty}(-1)^n\frac{z^{2n+1}}{(2n+1)!}\\ &= \sum_{n=0}^{\infty}(-1)^n\frac{(2n+1)z^{2n}}{(2n+1)!} = \sum_{n=0}^{\infty}(-1)^n\frac{z^{2n}}{(2n)!} = \cos z\,.\end{aligned}$$

Sebbene poter scambiare il segno di derivata con quello di somma possa sembrare ovvio, in realtà non è così scontato perché c'è anche un problema di convergenza. Non si tratta di una somma, ma di un limite di somme, lo scambio quindi è tra la derivata e un passaggio al limite. Ammesso poi che tale operazione sia possibile, rimane da precisare in quali punti la derivata della serie e la serie delle derivate coincidono.

Teorema 9.33 - *Se $R>0$ è il raggio di convergenza della serie $\sum c_n z^n$ allora anche la serie $\sum nc_n z^{n-1}$ ammette R come raggio di convergenza. Inoltre la somma della serie, la funzione*

$$f(z) = \sum_{n=0}^{\infty} c_n z^n\,, \qquad z\in B_R(0)\,,$$

è derivabile in $B_R(0)$ e

$$f'(z) = \sum_{n=1}^{\infty} nc_n z^{n-1} \qquad \forall z\in B_R(0)\,.$$

Dimostrazione. Calcoliamo il raggio di convergenza della serie delle derivate

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

Veniamo adesso alla derivabilità di f mostrando che, se $|z| < R$, allora

$$f(z+h) - f(z) = h \sum_{n=1}^{\infty} n c_n z^{n-1} + h\sigma(h)$$

dove $\sigma(h)$ è un infinitesimo per $h \to 0$. Dalla formula del binomio segue che

$$\begin{aligned} f(z+h) - f(z) &= \sum_{n=0}^{\infty} c_n[(z+h)^n - z^n] = \sum_{n=1}^{\infty} c_n \sum_{k=1}^{n} \binom{n}{k} h^k z^{n-k} \\ &= \sum_{n=1}^{\infty} c_n \left[\binom{n}{1} h z^{n-1} + \binom{n}{2} h^2 z^{n-2} + \cdots + \binom{n}{n} h^n \right] \\ &= h \sum_{n=1}^{\infty} n c_n z^{n-1} + h \sum_{n=2}^{\infty} c_n \left[\binom{n}{2} h z^{n-2} + \cdots + \binom{n}{n} h^{n-1} \right]. \end{aligned}$$

Adesso mostriamo che

$$\sigma(h) = \sum_{n=2}^{\infty} c_n \left[\binom{n}{2} h z^{n-2} + \cdots + \binom{n}{n} h^{n-1} \right]$$

è infinitesima per $h \to 0$. Si ha infatti

$$\begin{aligned} \sigma(h) &= \sum_{n=2}^{\infty} c_n \sum_{k=2}^{n} \binom{n}{k} h^{k-1} z^{n-k} = \sum_{n=2}^{\infty} c_n \sum_{i=0}^{n-2} \binom{n}{i+2} h^{i+1} z^{n-2-i} \\ &= h \sum_{n=2}^{\infty} c_n \sum_{i=0}^{n-2} \binom{n}{i+2} h^i z^{n-2-i}. \end{aligned}$$

D'altra parte

$$\binom{n}{i+2} = \frac{n(n-1)(n-2)!}{(i+2)(i+1)i!(n-2-i)!} \leqslant \frac{n(n-1)}{2} \binom{n-2}{i},$$

quindi

$$\begin{aligned} |\sigma(h)| &\leqslant |h| \sum_{n=2}^{\infty} |c_n| \sum_{i=0}^{n-2} \frac{n(n-1)}{2} \binom{n-2}{i} |h|^i |z|^{n-2-i} \\ &= |h| \sum_{n=2}^{\infty} |c_n| \frac{n(n-1)}{2} (|h| + |z|)^{n-2}. \end{aligned}$$

Se adesso scegliamo $|h|$ abbastanza piccolo in modo che $|z| + |h| < R$, si ha

$$\sum_{n=2}^{\infty} |c_n| \frac{n(n-1)}{2} (|h| + |z|)^{n-2} < +\infty$$

e quindi $\sigma(h) \to 0$.

□

Anche alla derivata, in quanto somma di una serie di potenze, è applicabile lo stesso teorema, per cui f ammette derivata seconda e

$$f''(z) = \sum_{n=2}^{\infty} n(n-1)z^{n-2}\,, \qquad z \in B_R(z_0)\,,$$

nel senso dell'assoluta convergenza nei punti interni allo stesso cerchio. Immediata conseguenza del Teorema 9.33 è il seguente risultato.

Corollario 9.34 - *Se $f : B_R(z_0) \to \mathbf{R}$ è la somma di una serie di potenze*

$$f(z) = \sum_{n=0}^{\infty} c_n(z-z_0)^n \qquad \forall z \in B_R(x_0)$$

allora $f \in C^\infty(B_R(z_0))$ e ogni derivata è a sua volta la somma di una serie di potenze sullo stesso cerchio

$$f^{(k)}(z) = \sum_{n=k}^{\infty} n(n-1)\dots(n-k+1)c_n(z-z_0)^{n-k} \qquad \forall z \in B_R(z_0)$$

In particolare $f^{(k)}(z_0) = k!c_k$, relazione che ci permette di ricavare i *coefficienti di Taylor* $c_k = f^{(k)}(z_0)/k!$ conoscendo la funzione. Così, nota la f, possiamo costruirne la rappresentazione in *serie di Taylor*

$$f(z) = \sum_{n=0}^{\infty} \frac{f^{(n)}(z_0)}{n!}(z-z_0)^n\,. \tag{9.16}$$

Osservazione 9.35 - *Riguardo agli estremi dell'intervallo di convergenza, o nei punti del bordo del cerchio di convergenza, non c'è nessuna relazione tra il comportamento della serie associata ad f e quello della serie associata ad f', lo possiamo verificare negli esempi che seguono.*

Il Teorema 9.33 è estremamente utile per ottenere altri sviluppi in serie di potenze. Ad esempio, se nella rappresentazione

$$\frac{1}{1+x} = \sum_{n=0}^{\infty}(-1)^n x^n\,, \qquad |x| < 1\,, \tag{9.17}$$

consideriamo la funzione a I membro come la derivata di qualche funzione, per il Teorema 9.33 si ottiene

$$\log(1+x) = c + \sum_{n=0}^{\infty}(-1)^n \frac{x^{n+1}}{n+1}\,, \qquad |x| < 1\,,$$

ma siccome $\log 1 = 0$, deve essere $c = 0$, da cui

$$\log(x+1) = x - \frac{x^2}{2} + \frac{x^3}{3} - \ldots = \sum_{n=0}^{\infty}(-1)^n \frac{x^{n+1}}{n+1}\,, \qquad |x| < 1\,. \tag{9.18}$$

Esercizio 9.23 - *Tenendo presente l'Esercizio 6.21, ricavare lo sviluppo in serie di potenze per la funzione* $\operatorname{arcsen} x$.

Per $x = 1$ la serie nella (9.18) è convergente, ma quella delle derivate nella (9.17) è indeterminata. Similmente, ancora per $|x| < 1$ si ottiene lo sviluppo per l'arctg x

$$\frac{1}{1+x^2} = \sum_{n=0}^{\infty} (-1)^n x^{2n} \quad \Rightarrow \quad \operatorname{arctg} x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}\,.$$

Anche in questo caso, la seconda converge per $x = 1$ mentre la prima è indeterminata. Ricordiamo che per il Teorema di Abel 6.16 è lecito in questi due esempi porre $x = 1$, che non sta dentro il cerchio, per ricavare i valori delle somme

$$\sum_{n=1}^{\infty} \frac{(-1)^{n-1}}{n} = \log 2 \qquad \text{e} \qquad \sum_{n=1}^{\infty} \frac{(-1)^n}{2n+1} = \frac{\pi}{2}\,.$$

Serie binomiale - Vediamo adesso il caso più generale della potenza $(1+x)^\alpha$, $\alpha \in \mathbf{R}$ (che abbiamo già considerato per $\alpha = 1/2$). Per $\alpha \in \mathbf{N}$ si riduce allo sviluppo finito del binomio di Newton. Per α qualunque in $\mathbf{R}$, introdotti i coefficienti binomiali

$$\binom{\alpha}{0} = 1 \qquad \text{e} \qquad \binom{\alpha}{n} = \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!} \qquad \forall n \geqslant 1\,,$$

dimostriamo che

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \qquad \forall x \in]-1,1[\,.$$

Prima di tutto calcoliamo il raggio di convergenza

$$\begin{aligned} \lim_{n\to\infty} \sqrt[n]{\left|\binom{\alpha}{n}\right|} &= \lim_{n\to\infty} \left|\binom{\alpha}{n+1}\right| \left|\binom{\alpha}{n}\right|^{-1} \\ &= \lim_{n\to\infty} \frac{|\alpha(\alpha-1)\ldots(\alpha-n)|n!}{|\alpha(\alpha-1)\ldots(\alpha-n+1)|(n+1)!} = \lim_{n\to\infty} \frac{|\alpha-n|}{n+1} = 1\,, \end{aligned}$$

dunque $R = 1$. Poniamo allora

$$f(x) = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \qquad \forall x \in]-1,1[$$

e cerchiamo una relazione tra f e la sua derivata

$$f'(x) = \sum_{n=1}^{\infty} n \binom{\alpha}{n} x^{n-1} = \sum_{n=0}^{\infty} (n+1) \binom{\alpha}{n+1} x^n$$

ottenuta derivando termine a termine. Osserviamo che

$$x f'(x) = \sum_{n=1}^{\infty} n \binom{\alpha}{n} x^n = \sum_{n=0}^{\infty} n \binom{\alpha}{n} x^n\,.$$

Sommando membro a membro le due relazioni si ottiene

$$(1+x)f'(x) = \sum_{n=0}^{\infty} \left[(n+1)\binom{\alpha}{n+1} + n\binom{\alpha}{n}\right] x^n = \alpha \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n = \alpha f(x)\,.$$

Infatti

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

Dunque f soddisfa l'equazione differenziale

$$(1+x)f'(x) = \alpha f(x) \qquad \forall x \in]-1,1[\,.$$

Per continuità, essendo $f(0) = 1$, f dovrà rimanere positiva in un intorno U di 0 e l'equazione può essere scritta nella forma

$$\frac{f'(x)}{f(x)} = \frac{\alpha}{1+x} \qquad \forall x \in U\,.$$

D'altra parte

$$\frac{f'(x)}{f(x)} = \frac{d}{dx}\log f(x) \qquad \text{e} \qquad \frac{\alpha}{1+x} = \alpha\frac{d}{dx}\log(1+x)\,,$$

pertanto

$$\log f(x) = c + \log(1+x)^\alpha$$

nella quale possiamo supporre $U =]-1,1[$. Passando all'esponenziale si ottiene

$$f(x) = C(1+x)^\alpha\,,$$

ma, ancora, $f(0) = 1$, quindi $C = 1$ e $f(x) = (1+x)^\alpha$. In particolare, per $\alpha = -1$ si riconosce la serie geometrica.

Per $\alpha = 1/2$, tanto per fare un esempio, si ottiene

$$\sqrt{1+x} = \sum_{n=0}^{\infty} \binom{1/2}{n} x^n = 1 + \frac{x}{2} - \frac{x^2}{8} + \frac{3x^3}{48} + \dots\,.$$

Qui si possono reperire molti altri interessanti sviluppi in serie.

Per il Corollario 9.34 una funzione analitica su un intervallo aperto $I \subset \mathbf{R}$ o su un aperto $\Omega \subset \mathbf{C}$ è necessariamente di classe C^∞. Sull'implicazione contraria il caso reale e il caso complesso differiscono:

• *in* $\mathbf{C}$*: una funzione di variabile complessa che ammette derivata prima in* Ω *è necessariamente analitica in* Ω *(e quindi ammette derivata di ogni ordine).*

• *in* $\mathbf{R}$*: una funzione definita su un intervallo aperto, anche se ammette derivata di ogni ordine, non è necessariamente analitica.*

La dimostrazione della prima affermazione esula dagli scopi di questo corso, per la seconda basta considerare il seguente esempio

$$f(x) = \begin{cases} e^{-1/x^2} & \text{se } x \neq 0 \\ 0 & \text{se } x = 0\,. \end{cases}$$

Questa funzione, di cui abbiamo già parlato a proposito della formula di Talyor, è positiva e nulla solo in 0 che dunque è di minimo isolato. Inoltre appartiene a $C^\infty(\mathbf{R})$ perché la derivata di ogni prdine è il prodotto di e^{-1/x^2} per una funzione razionale, quindi in 0 ha prolungamento nullo. Ora se f fosse analitica, essendo i suoi valori nell'intorno di 0 determinati dalle derivate in 0 che sono tutte nulle, sarebbe identicamente nulla, in contraddizione col fatto che è positiva.

Come si spiega questo fenomeno un po' sorprendente, che rende il caso reale così diverso da quello complesso? Il vero motivo per cui questa funzione non è analitica va ricercato ancora una volta in $\mathbf{C}$: la funzione e^{-1/x^2} non è altro che la restrizione all'asse reale della funzione analitica e^{-1/z^2}, $z \in \mathbf{C} - \{0\}$, che non ammette nemmeno prolungamento continuo in 0. Lungo l'asse immaginario infatti, per $z = iy$, diventa e^{1/y^2} che tende a $+\infty$ per $y \to 0$.

È naturale allora chiedersi: quando una funzione infinitamente derivabile è analitica? Come si passa dalla formula di Taylor finita, ma valida fino a qualunque ordine, alla serie di Taylor?

Proposizione 9.36 - *Condizione sufficiente affinché $f \in C^\infty]-R,R[$ sia sviluppabile in serie di Taylor in $]-R,R[$ è che esista $M \geqslant 0$ tale che*

$$\sup_{|x|<R} |f^{(n)}(x)| \leqslant M \frac{n!}{R^n} \qquad \forall n \in \mathbf{N}\,. \tag{9.19}$$

Dimostrazione. Usiamo la formula di Taylor col resto di Lagrange. Per ogni $x \in]-R,R[$ esiste $\xi \in]-R,R[$, $|\xi| \leqslant |x|$, tale che

$$r_n(x) = \frac{f^{(n+1)}(\xi)}{(n+1)!} x^{n+1}\,,$$

quindi

$$|r_n(x)| \leqslant M \frac{(n+1)!}{R^{n+1}} \frac{|x|^{n+1}}{(n+1)!} = M \left(\frac{|x|}{R}\right)^{n+1}\,.$$

Ne segue che

$$\lim_{n\to\infty} |r_n(x)| = 0$$

e il polinomio di Taylor diventa la successione convergente delle somme parziali della serie di potenze associata a f.

□

Osservazione 9.37 *Essendo $R^n \leqslant n!$ definitivamente, una situazione che implica la (9.19) è la condizione che esista una costante $M \geqslant 0$ indipendente da n tale che*

$$|f(x)| \leqslant M \qquad \forall x \in]-R,R[\,.$$

Il teorema di derivazione termine a termine di una serie non è vero in generale. Vediamo un interessante controesempio con gli strumenti che abbiamo acquisito sulle serie di potenze. Consideriamo la serie

$$\sum_{n=0}^{\infty} \frac{\operatorname{sen}(2n+1)x}{2n+1} \tag{9.20}$$

convergente per il criterio di Dirichlet. La serie delle derivate $\sum_{n=0}^\infty \cos(2n+1)x$ di certo non converge. Per calcolare la somma della (9.20), conviene vederla come la parte immaginaria della serie complessa

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

che è la restrizione alla circonferenza unitaria della serie di potenze

$$\sum_{n=0}^{\infty} \frac{z^{2n+1}}{2n+1}\,.$$

La somma di questa all'interno del disco è stata calcolata nella (6.26)

$$\sum_{n=0}^{\infty} \frac{z^{2n+1}}{2n+1} = \frac{1}{2} \log\left(\frac{1+z}{1-z}\right) = \frac{1}{2}\left[\log\left|\frac{1+z}{1-z}\right| + i \operatorname{Arg}\left(\frac{1+z}{1-z}\right)\right]\,,$$

dove la scelta dell'argomento principale è dovuta al valore che si ottiene al I membro per $z=0$. Passando sul bordo, segue dal Teorema di Abel

$$\sum_{n=0}^{\infty} \frac{\operatorname{sen}(2n+1)x}{2n+1} = \frac{1}{2} \operatorname{Arg}\left(\frac{1+e^{ix}}{1-e^{ix}}\right) = \begin{cases} -\dfrac{\pi}{4} & \text{se } (2k-1)\pi < x < 2k\pi \\ \dfrac{\pi}{4} & \text{se } 2k\pi < x < (2k+1)\pi\,, \end{cases}$$

essendo $(1+e^{ix})/(1-e^{ix})$ immaginario. Pertanto la somma della serie è costante a tratti ed ha derivata ovunque nulla eccetto nei punti di salto.

Capitolo 10

Calcolo integrale

10.1 Dall'area all'integrale

L'area di una regione piana è il numero reale che esprime il rapporto tra la sua *estensione* e quella del quadrato unitario, analogamente a come si è definito nel § 2.10 la lunghezza di un segmento. Facciamo alcune ipotesi preliminari sull'area, peraltro del tutto ragionevoli, per precisare con un certo rigore di che cosa si sta parlando e per poterci lavorare.

In una certa classe di regioni piane, che preciseremo tra poco, supponiamo sia definita una funzione *area* che associa ad ogni regione E un numero reale $\mathscr{A}(E)$, l'area di E, tale che

($\mathscr{A}$)1. $\mathscr{A}(E) \geqslant 0$;

($\mathscr{A}$)2. $\mathscr{A}(E)$ è invariante per isometrie, cioè rotazioni e traslazioni di E non alterano la sua area;

($\mathscr{A}$)3. $\mathscr{A}(E)$ è additiva, cioè $\mathscr{A}(E_1 \cup E_2) = \mathscr{A}(E_1) + \mathscr{A}(E_2)$ se $E_1 \cap E_2 = \emptyset$ (per induzione vale anche per una famiglia finita di regioni E_i);

($\mathscr{A}$)4. nel caso di un rettangolo $R = I \times J$, con I, J intervalli di lunghezze l_1 e l_2, $\mathscr{A}(R) = l_1 l_2$, in particolare un segmento, in quanto rettangolo degenere, ha area nulla.

Si noti che le ($\mathscr{A}$)1 e ($\mathscr{A}$)3 insieme implicano che l'area $\mathscr{A}$ è una funzione crescente di insieme rispetto all'inclusione, nel senso che $E_1 \subset E_2 \Rightarrow \mathscr{A}(E_1) \leqslant \mathscr{A}(E_2)$. Infatti

$$\mathscr{A}(E_1) \leqslant \mathscr{A}(E_1) + \mathscr{A}(E_2 - E_1) = \mathscr{A}(E_1 \cup (E_2 - E_1)) = \mathscr{A}(E_2)\,.$$

Inoltre combinando tra loro le proprietà ($\mathscr{A}$)1 − ($\mathscr{A}$)4 possiamo calcolare le aree di tutti i poligoni*. Ogni poligono può essere infatti decomposto in una unione finita di triangoli disgiunti (o con dei lati in comune che però non hanno nessuna influenza sull'area), ognuno dei quali ha area pari alla metà di quella del rettangolo costruito su uno dei suoi lati e di pari altezza. Di fronte però a insiemi più generali le cose si fanno piuttosto complicate. Anche se ci limitiamo alle sole regioni delimitate da unioni finite di curve, perfino che cosa si debba intendere per curva non è del tutto chiaro e vi sono molti livelli di generalità, certi insiemi hanno una frontiera molto complicata e lontana dall'idea di curva che abbiamo in mente e senza ulteriori precisazioni non sarebbe chiaro neanche che cosa significhi misurarne l'area. A partire dalla classe più ristretta dei poligoni, conviene allora costruire direttamente una classe

*In questo contesto conviene intendere per *poligono* anche un'eventuale unione finita di poligoni, intesi nel senso usuale, senza punti in comune.

ragionevolmente ampia di insiemi del piano a cui sia possibile estendere la nozione di area. Il procedimento, legittimato dalla completezza di $\mathbf{R}$, consiste nell'approssimare un insieme dall'*interno* e dall'*esterno* con opportuni poligoni. Precisamente, dato un insieme $E \subset \mathbf{R}^2$, consideriamo le due famiglie $\mathscr{P}^-$ e $\mathscr{P}^+$ di tutti i poligoni contenuti in E e contenenti E rispettivamente. Per la monotonia dell'area, che abbiamo già verificato, le aree dei poligoni di $\mathscr{P}^-$ e quelle dei poligoni di $\mathscr{P}^+$ costituiscono due insiemi separati di numeri reali.

Definizione 10.1 - *Diciamo che E è **misurabile secondo Jordan** se*

$$\sup_{P\in\mathscr{P}^-} \mathscr{A}(P) = \inf_{P\in\mathscr{P}^+} \mathscr{A}(P)\,.$$

*Il valore comune dei due estremi viene detto **area** di E, o **misura di Jordan** di E, e si indica con $\mathscr{A}(E)$.*

In altre parole, E è misurabile se questi due insiemi separati di aree sono anche contigui e l'unico elemento separatore è l'area di E. La stessa definizione può essere data in modo analogo in più dimensioni, per il volume di E in $\mathbf{R}^3$ o per la misura n-dimensionale in $\mathbf{R}^n$, a partire da prodotti cartesiani di tre o n intervalli che sono i parallelepipedi, l'analogo dei rettangoli in $\mathbf{R}^2$ come si può facilmente immaginare.

Dalla Definizione 2.3 del Cap. 2 segue che $E \subset \mathbf{R}^2$ è misurabile se e solo se

$$\forall\varepsilon > 0\ \exists P^- \in \mathscr{P}^-\,,\ P^+ \in \mathscr{P}^+\ :\ \mathscr{A}(P^+) - \mathscr{A}(P^-) < \varepsilon\,. \tag{10.1}$$

I poligoni P^- e P^+ dipendono da ε che può essere scelto arbitrariamente piccolo. La particolare scelta $\varepsilon = 1/n$, $n \in \mathbf{N} - \{0\}$, ci dà un'ulteriore condizione necessaria e sufficiente in termini di successioni di poligoni: $E \subset \mathbf{R}^2$ è misurabile se e solo se

$$\exists (P_n^-) \in \mathscr{P}^-\,,\ (P_n^+) \in \mathscr{P}^+\ :\ \lim_{n\to\infty} [\mathscr{A}(P_n^+) - \mathscr{A}(P_n^-)] = 0\,.$$

Da queste considerazioni segue immediatamente la conferma che l'area di ogni segmento del piano, anche se non l'avessimo considerato come un rettangolo, è nulla, basta infatti, per ogni $\varepsilon > 0$, ricoprirlo con un rettangolo di area ε. Inoltre un'attenta rilettura della (10.1), dove $\mathscr{A}(P^+) - \mathscr{A}(P^-) = \mathscr{A}(P^+ - P^-)$, ci fornisce una caratterizzazione della misurabilità di un insieme: è misurabile ogni insieme $E \subset \mathbf{R}^2$ la cui frontiera ha misura nulla.

In questo capitolo siamo interessati al calcolo dell'area di regioni piane delimitate da un intervallo I dell'asse x, dal grafico di una funzione a valori reali definita su I e, nel caso che gli estremi a e b di I siano finiti, dalle rette $x = a$ e $x = b$. Per una funzione f non negativa su I si tratta dell'area del *sottografico*

$$\Gamma(f) = \{(x, y) \in \mathbf{R}^2 \mid x \in I\,,\ 0 \leqslant y \leqslant f(x)\}\,.$$

Per gli insiemi di questo tipo possiamo scegliere come poligoni approssimanti dall'interno e dall'esterno delle unioni finite di rettangoli, con i lati paralleli agli assi che per comodità possiamo supporre disgiunti (oppure, al solito, con al più dei lati in comune). Per meglio chiarire il motivo di questa scelta, vediamo come si può applicare la definizione che abbiamo dato poc'anzi all'esempio semplice della regione che sta tra l'asse x e un arco di parabola. Assegnata la funzione $f(x) = x^2$ con $0 \leqslant x \leqslant a$, per calcolare l'area di $\Gamma(f)$ scegliamo gli intervalli $I_h = [x_{h-1}, x_h]$, $h = 1, \ldots, n$, di uguale lunghezza a/n con $x_h = ha/n$ e su ogni I_h costruiamo il rettangolo di altezza $f(x_{h-1}) = (h-1)^2a^2/n^2$. La loro unione è un poligono P_n^- contenuto in $\Gamma(f)$ la cui area è

$$\mathscr{A}(P_n^-) = \sum_{h=1}^{n} \frac{(h-1)^2a^3}{n^3} = \frac{n(n-1)(2n-1)a^3}{6n^3}\,.$$

Sugli stessi intervalli I_h costruiamo adesso i rettangoli di altezze $f(x_h) = h^2a^2/n^2$. La loro unione è un poligono P_n^+ contenente $\Gamma(f)$ la cui area è

$$\mathscr{A}(P_n^+) = \sum_{h=1}^{n} \frac{h^2a^3}{n^3} = \frac{n(n+1)(2n+1)a^3}{6n^3}\,.$$

Le due successioni di aree $(\mathscr{A}(P_n^-))$ e $(\mathscr{A}(P_n^+))$, la prima crescente e la seconda decresente, hanno lo stesso limite $a^3/3$ che è anche il loro unico elemento separatore. Dunque $\Gamma(f)$ è misurabile e la sua area è $a^3/3$.

Per una funzione qualsiasi $f : [a,b] \to \mathbf{R}$ non negativa si può procedere nello stesso modo: per ogni $n \in \mathbf{N}$ si considera un'opportuna suddivisione di $[a,b]$ mediante certi punti $x_0 = a < x_1 < x_2 < \ldots < x_n = b$, su ogni intervallo $I_h = [x_{h-1}, x_h]$ si costruiscono due rettangoli di altezze α_h e β_h tali che $\alpha_h \leqslant f(x) \leqslant \beta_h$ per ogni $x \in I_h$. Se le due successioni

$$\sum_{h=1}^{n} \alpha_h(x_h - x_{h-1}) \quad \text{e} \quad \sum_{h=1}^{n} \beta_h(x_h - x_{h-1})\,, \quad n \in \mathbf{N}\,,$$

che sono anche due classi separate di numeri reali, sono contigue allora $\Gamma(f)$ è misurabile ed ha per area l'unico elemento separatore. È facile rendersi conto a questo punto che se $f(x) \geqslant g(x) \geqslant 0$ per ogni $x \in I$ e $\Gamma(f)$ e $\Gamma(g)$ sono misurabili, allora lo è anche l'insieme delimitato dai loro grafici

$$\Gamma(g,f) = \{(x,y) \in \mathbf{R}^2 \mid x \in I\,,\ g(x) \leqslant y \leqslant f(x)\}$$

ed ha per area $\mathscr{A}(\Gamma(f)) - \mathscr{A}(\Gamma(g))$, la quale coincide con $\mathscr{A}(\Gamma(f-g))$.

Nella teoria dell'integrazione che esporremo in questo capitolo non faremo alcun riferimento all'area di $\Gamma(f)$, l'integrale riguarda la funzione direttamente e va visto come un'applicazione che associa ad essa un numero reale. Il primo passo sarà definire l'integrale per una funzione $\varphi : \mathbf{R} \to \mathbf{R}$ del tipo

$$\varphi(x) = \begin{cases} \alpha_h & \text{se } x \in I_h \\ 0 & \text{se } x \notin [a,b]\,, \end{cases}$$

che diremo costante a tratti. È evidente che il sottografico di una φ come questa è l'unione finita dei rettangoli che hanno per base gli intervalli I_h e altezze i numeri reali α_h. L'integrale di φ sarà dato dalla somma algebrica

$$\sum_{h=1}^{n} \alpha_h(x_h - x_{h-1})$$

che formalmente coincide con l'espressione dell'area che abbiamo già visto. Mentre le differenze $m(I_h) = x_h - x_{h-1}$ sono sempre positive, e rappresentano le misure (unidimensionali) degli intervalli I_h, gli α_h possono avere segno qualunque, in particolare quelli eventualmente negativi danno un contributo negativo al valore complessivo della somma. Il secondo passo consiste nel definire l'integrale per una generica funzione limitata f definita su $[a,b]$ e prolungata a 0 sul complementare $\mathbf{R} - [a,b]$. Se gli integrali delle funzioni costanti a tratti minoranti la f e quelli delle maggioranti formano due classi contigue di numeri reali allora f verrà detta integrabile e il suo integrale sarà l'unico elemento separatore.

10.2 L'integrale per le funzioni costanti a tratti

Dato un insieme $A \subset \mathbf{R}$, si chiama *funzione caratteristica* di A la funzione $\chi_A : \mathbf{R} \to \mathbf{R}$

$$\chi_A(x) = \begin{cases} 1 & \text{se } x \in A \\ 0 & \text{se } x \in \complement A \,. \end{cases}$$

A noi interessa il caso in cui A è un intervallo. Consideriamo dunque la funzione caratteristica χ_I di un intervallo I. Se questa viene moltiplicata per una costante $\alpha \in \mathbf{R}$ si ottiene la funzione $\alpha\chi_I$ che vale α si I e 0 sul complementare. Il prodotto di due funzioni caratteristiche χ_I e χ_J è $\chi_{I\cap J}$ (la funzione caratteristica del vuoto è nulla ovunque), invece una loro combinazione lineare, $\varphi = \alpha\chi_I + \beta\chi_J$, vale $\alpha + \beta$ su $I \cap J$, α su $I - J$, β su $J - I$ e 0 su tutto il resto di $\mathbf{R}$. È evidente che non c'è un unico modo per rappresentare φ, $I \cup J$ può essere spezzato ad esempio nei tre intervalli disgiunti $I_1 = I - J$, $I_2 = I \cap J$ e $I_3 = J - I$ e, posto $\gamma = \alpha + \beta$, la stessa funzione può essere riscritta nella forma

$$\varphi = \alpha\chi_{I_1} + \gamma\chi_{I_2} + \beta\chi_{I_3} \tag{10.2}$$

e in infiniti altri modi spezzando ulteriormente gli intervalli I_1, I_2 e I_3. La (10.2) è stata costruita scegliendo come estremi degli intervalli i punti di salto della funzione.

Definizione 10.2 *Una combinazione lineare finita di funzioni caratteristiche di intervalli limitati si chiama* ***funzione costante a tratti*** *o anche* ***funzione step****.*

Si tratta quindi di una funzione $\varphi : \mathbf{R} \to \mathbf{R}$ del tipo

$$\varphi(x) = \sum_{h=1}^{k} \alpha_h \chi_{I_h}(x) \quad \forall x \in \mathbf{R}$$

con $\alpha_h \in \mathbf{R}$ e nella quale possiamo supporre $I_h = [x_{h-1}, x_h]$, $h = 1, \dots, k$, perché l'unico punto che I_h e I_j possono avere in comune, se adiacenti, ha misura nulla. In definitiva una funzione step è nulla al di fuori di un intervallo limitato e assume un numero finito di valori. Se si ha a che fare con due funzioni step da trattare insieme è sempre possibile ridefinirle su una stessa suddivisione di intervalli separati, basta spezzare opportunamente gli intervalli associati all'una e all'altra in modo da ottenere come estremi i punti di salto di ognuna delle due funzioni.

Definiamo *misura* dell'intervallo I di estremi $a, b \in \mathbf{R}$ la sua lunghezza $b - a$ che indicheremo con $m(I)$. Osserviamo che $m(I)$ non risente della presenza in I dei suoi estremi: che I sia chiuso, aperto o semiaperto, la sua misura è sempre la stessa.

Definizione 10.3 - *Si chiama* ***integrale*** *della funzione step*

$$\varphi = \sum_{h=1}^{k} \alpha_h \chi_{I_h}$$

il numero reale

$$\int_{\mathbf{R}} \varphi(x)\, dx = \sum_{h=1}^{k} \alpha_h m(I_h)\,. \tag{10.3}$$

Si osservi che se adesso ridefiniamo i vari I_h senza preoccuparci della presenza o meno degli estremi, che sono in numero finito e quindi di misura complessiva nulla, per quanto possano risentirne i valori della φ, l'integrale rimane inalterato.

Indichiamo con $\mathscr{S}(\mathbf{R})$ lo spazio vettoriale di tutte le funzioni step su $\mathbf{R}$ e lasciamo per esercizio le seguenti proprietà dell'integrale.

($\int$)1. $\int_{\mathbf{R}}(\varphi+\psi)=\int_{\mathbf{R}}\varphi+\int_{\mathbf{R}}\psi$ e $\int_{\mathbf{R}}\lambda\varphi=\lambda\int_{\mathbf{R}}\varphi$ $\quad\forall\varphi,\psi\in\mathscr{S}(\mathbf{R}),\forall\lambda\in\mathbf{R}$;

($\int$)2. $\forall\varphi\in\mathscr{S}(\mathbf{R})\quad\varphi\geqslant 0\Rightarrow\int_{\mathbf{R}}\varphi\geqslant 0$,

($\int$)3. $\forall\varphi\in\mathscr{S}(\mathbf{R})\quad\left|\int_{\mathbf{R}}\varphi\right|\leqslant\int_{\mathbf{R}}|\varphi|$.

Nella ($\int$)1. si afferma che l'applicazione $\int_{\mathbf{R}}:\mathscr{S}(\mathbf{R})\to\mathbf{R}$ è lineare dallo spazio vettoriale $\mathscr{S}(\mathbf{R})$ nel campo $\mathbf{R}$ degli scalari, per questo si chiama *funzionale lineare*. Nella ($\int$)2. si afferma che l'integrale è *positivo*, proprietà che, insieme alla precedente, permette di dimostrare che è crescente

$$\forall\varphi,\psi\in\mathscr{S}(\mathbf{R})\quad\varphi\leqslant\psi\Rightarrow\int_{\mathbf{R}}\varphi\leqslant\int_{\mathbf{R}}\psi\,. \tag{10.4}$$

10.3 L'integrale di Riemann

Indichiamo con $\mathscr{L}_0(\mathbf{R})$ l'insieme delle funzioni $f:\mathbf{R}\to\mathbf{R}$ limitate e nulle al di fuori di un intervallo limitato. Data una funzione $f\in\mathscr{L}_0(\mathbf{R})$, siano $\mathscr{S}^-(f)$ e $\mathscr{S}^+(f)$ le due classi di funzioni step

$$\begin{aligned}\mathscr{S}^-(f)&=\{\varphi\in\mathscr{S}\mid\varphi(x)\leqslant f(x)\ \forall x\in\mathbf{R}\}\\ \mathscr{S}^+(f)&=\{\psi\in\mathscr{S}\mid\psi(x)\geqslant f(x)\ \forall x\in\mathbf{R}\}\,.\end{aligned}$$

Per costruzione si ha

$$\varphi\leqslant\psi\qquad\forall\varphi\in\mathscr{S}^-(f)\,,\forall\psi\in\mathscr{S}^+(f)$$

e dalla (10.4) segue che le due classi di numeri reali

$$S^-=\left\{\int_{\mathbf{R}}\varphi\mid\varphi\in\mathscr{S}^-(f)\right\}\quad\text{e}\quad S^+=\left\{\int_{\mathbf{R}}\psi\mid\psi\in\mathscr{S}^+(f)\right\}$$

sono separate.

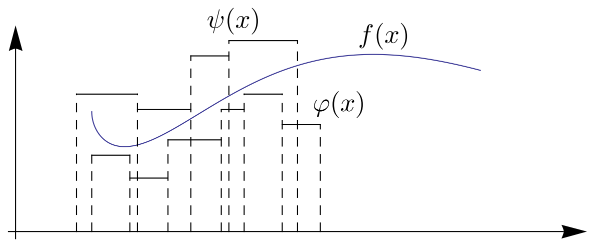


Figura 10.1: Le step approssimanti dal basso e dall'alto.

Definizione 10.4 - *Si chiamano* ***integrale inferiore*** *e* ***integrale superiore*** *di* $f\in\mathscr{L}_0(\mathbf{R})$ *i due numeri*

$$\int_{\mathbf{R}}^{-}f(x)\,dx=\sup S^-\qquad e\qquad\int_{\mathbf{R}}^{+}f(x)\,dx=\inf S^+\,.$$

Se

$$\int_{\mathbf{R}}^{-}f=\int_{\mathbf{R}}^{+}f$$

allora f è detta ***integrabile secondo Riemann*** *e questo valore comune è l'****integrale di Riemann*** *di f, indicato al solito con la notazione*

$$\int_{\mathbf{R}}f(x)\,dx\,.$$

In altre parole, gli integrali delle funzioni semplici minoranti e quelli delle maggioranti la f sono due classi separate di numeri reali, ma se sono anche contigue f è integrabile e l'unico elemento separatore è il suo integrale. Osserviamo che se f, in particolare, è una funzione step allora è integrabile secondo Riemann e il suo integrale coincide con quello della Definizione 10.3. In questo caso infatti $f \in \mathscr{S}^-(f) \cap \mathscr{S}^+(f)$, quindi le due classi numeriche S^- e S^+ hanno un elemento in comune che è allo stesso tempo il massimo della prima e il minimo della seconda. Ciò mostra che l'integrale di Riemann è un'estensione naturale di quello definito all'inizio per le funzioni step e possiamo senz'altro interpretare l'integrale di f come l'area, con segno, della regione piana compresa tra il suo grafico e l'asse x.

Ricordando proprietà ben note sulle classi contigue e sugli estremi inferiore e superiore, le seguenti utili condizioni necessarie e sufficienti per l'integrabilità sono immediate

$$\forall \varepsilon > 0 \; \exists \varphi \in \mathscr{S}^-(f), \psi \in \mathscr{S}^+(f) \;:\; \int_{\mathbf{R}} (\psi - \varphi) < \varepsilon \,; \tag{10.5}$$

$$\exists (\varphi_n) \subset \mathscr{S}^-(f), (\psi_n) \subset \mathscr{S}^+(f) \;:\; \lim_{n\to\infty} \int_{\mathbf{R}} (\psi_n - \varphi_n) = 0 \,. \tag{10.6}$$

La (10.6) discende dalla (10.5), basta scegliere $\varepsilon = 1/n$. Se, viceversa, vale la (10.6) si fissa arbitrariamente $\varepsilon > 0$ e una coppia di funzioni $\varphi_n \in \mathscr{S}^-(f)$, $\psi_n \in \mathscr{S}^+(f)$ che soddisfa definitivamente

$$\int_{\mathbf{R}} (\psi_n - \varphi_n) < \varepsilon \,.$$

Esercizio 10.1 - *Dimostrare che f è integrabile se e solo se per ogni $\varepsilon > 0$ esistono due funzioni integrabili f_1 e f_2 tali che $f_1 \leqslant f \leqslant f_2$ e $\int (f_2 - f_1) < \varepsilon$.*

Indichiamo con $\mathscr{R}(\mathbf{R})$ il sottoinsieme di $\mathscr{L}_0(\mathbf{R})$ delle funzioni integrabili secondo Riemann e vediamo con l'aiuto della (10.5) come si possono estendere a $\mathscr{R}(\mathbf{R})$ le proprietà $(\int)$1., 2., 3. del § 10.2.

Teorema 10.5 - *Valgono le seguenti proprietà:*

$(\int)$1. *se $f, g \in \mathscr{R}(\mathbf{R})$ e $\lambda \in \mathbf{R}$ allora $f + g, \lambda f \in \mathscr{R}(\mathbf{R})$ e*

$$\int_{\mathbf{R}} (f + g) = \int_{\mathbf{R}} f + \int_{\mathbf{R}} g \quad e \quad \int_{\mathbf{R}} \lambda f = \lambda \int_{\mathbf{R}} f \,,$$

$(\int)$2. *se $f \in \mathscr{R}(\mathbf{R})$ e $f \geqslant 0$ allora $\int_{\mathbf{R}} f \geqslant 0$,*

$(\int)$3. *se $f \in \mathscr{R}(\mathbf{R})$ allora $|f| \in \mathscr{R}(\mathbf{R})$ e $\left|\int_{\mathbf{R}} f\right| \leqslant \int_{\mathbf{R}} |f|$,*

$(\int)$4. *se $f \in \mathscr{R}(\mathbf{R})$ allora $f^2 \in \mathscr{R}(\mathbf{R})$,*

$(\int)$5. *se $f, g \in \mathscr{R}(\mathbf{R})$ allora $fg \in \mathscr{R}(\mathbf{R})$.*

Dimostrazione. $(\int)$1. In questa proprietà si afferma che $\mathscr{R}(\mathbf{R})$ è uno spazio vettoriale e che l'integrale è un funzionale lineare su $\mathscr{R}(\mathbf{R})$. Per ogni $\varepsilon > 0$, scegliamo $\varphi_1, \psi_1, \varphi_2, \psi_2 \in \mathscr{S}$ tali che

$$\varphi_1 \leqslant f \leqslant \psi_1 \quad \text{e} \quad \varphi_2 \leqslant g \leqslant \psi_2 \tag{10.7}$$

e

$$\int_{\mathbf{R}} (\psi_1 - \varphi_1) < \frac{\varepsilon}{2} \quad \text{e} \quad \int_{\mathbf{R}} (\psi_2 - \varphi_2) < \frac{\varepsilon}{2} \,. \tag{10.8}$$

Sommando nella (10.7) si ottiene

$$\varphi_1 + \varphi_2 \leqslant f + g \leqslant \psi_1 + \psi_2 \tag{10.9}$$

dove, per la (10.8), si ha

$$\int_{\mathbf{R}} [(\psi_1 + \psi_2) - (\varphi_1 + \varphi_2)] = \int_{\mathbf{R}} (\psi_1 - \varphi_1) + \int_{\mathbf{R}} (\psi_2 - \varphi_2) < \varepsilon$$

e quindi $f + g$ è integrabile. Integrando membro a membro la (10.9) si ottiene

$$\int_{\mathbf{R}} (\varphi_1 + \varphi_2) \leqslant \int_{\mathbf{R}} (f + g) \leqslant \int_{\mathbf{R}} (\psi_1 + \psi_2)\,,$$

mentre integrando nella (10.7) e poi sommando si ottiene

$$\int_{\mathbf{R}} \varphi_1 + \int_{\mathbf{R}} \varphi_2 \leqslant \int_{\mathbf{R}} f + \int_{\mathbf{R}} g \leqslant \int_{\mathbf{R}} \psi_1 + \int_{\mathbf{R}} \psi_2\,.$$

Ne segue dalla (10.8)

$$\left| \int_{\mathbf{R}} (f + g) - \int_{\mathbf{R}} f - \int_{\mathbf{R}} g \right| < \varepsilon$$

ed essendo ε arbitrario si ottiene

$$\int_{\mathbf{R}} (f + g) = \int_{\mathbf{R}} f + \int_{\mathbf{R}} g\,.$$

La seconda parte si dimostra in modo simile e viene lasciata per esercizio.

($\int$)2. Consideriamo la funzione identicamente nulla come una funzione step il cui integrale è ovviamente nullo. Allora

$$\int_{\mathbf{R}} f \geqslant \int_{\mathbf{R}} 0 = 0\,.$$

($\int$)3. In termini delle parti positiva e negativa di f si ha $f = f^+ - f^-$ e $|f| = f^+ + f^-$. Ora, se f è integrabile e φ, ψ sono due funzioni step tali che $\varphi \leqslant f \leqslant \psi$ e $\int(\psi - \varphi) < \varepsilon$, si ha

$$\varphi^+ \leqslant f^+ \leqslant \psi^+ \qquad \text{e} \qquad \int (\psi^+ - \varphi^+) \leqslant \int (\psi - \varphi) < \varepsilon\,,$$

quindi f^+ è integrabile e nello stesso modo si dimostra che lo è anche f^-. Ne segue che $|f| = f^+ + f^-$ è integrabile. A questo punto basta integrare membro a membro la relazione

$$-|f| \leqslant f \leqslant |f|$$

ottenendo

$$-\int_{\mathbf{R}} |f| \leqslant \int_{\mathbf{R}} f \leqslant \int_{\mathbf{R}} |f|$$

che è equivalente alla disuguaglianza richiesta.

($\int$)4. Supponiamo dapprima $f \geqslant 0$. Fissato $\varepsilon > 0$, scegliamo $\varphi, \psi \in \mathscr{S}(\mathbf{R})$ tali che $0 \leqslant \varphi \leqslant f \leqslant \psi$, $\int(\psi - \varphi) < \varepsilon$ e sia inoltre $M \in \mathbf{R}$ un maggiorante comune a queste due funzioni step. In queste condizioni possiamo elevare al quadrato ottenendo

$$\varphi^2 \leqslant f^2 \leqslant \psi^2 \qquad \text{e} \qquad \int_{\mathbf{R}} (\psi^2 - \varphi^2) = \int_{\mathbf{R}} (\psi + \varphi)(\psi - \varphi) \leqslant 2M\varepsilon\,,$$

quindi f^2 è integrabile. Se adesso f è di segno qualsiasi $|f|$ è comunque integrabile e non negativa e poiché $f^2 = |f|^2$, anche in questo caso f^2 è integrabile.

($\int$)5. Basta osservare che

$$fg = \frac{(f + g)^2 - f^2 - g^2}{2}\,.$$

$\square$

Esercizio 10.2 - *Dimostrare che se $f \in \mathscr{R}(\mathbf{R})$ è non negativa allora esistono $\alpha, \beta \in \mathbf{R}$ tali che $f(x) > 0$ per ogni $x \in]\alpha, \beta[$ se e solo se l'integrale di f è strettamente positivo.*

Sulla composizione tra funzioni si può dimostrare che se f è integrabile e $h : \mathbf{R} \to \mathbf{R}$ è continua allora $h \circ f$ è integrabile, ma per i nostri scopi ci basta ampiamente il seguente teorema la cui dimostrazione è notevolmente più semplice.

Teorema 10.6 - *Siano $f \in \mathscr{R}(\mathbf{R})$ nulla al di fuori di $[a, b]$ e $h : \mathbf{R} \to \mathbf{R}$ localmente lipschitziana. Allora la funzione $h(f(x))$ per $x \in [a, b]$ e nulla al di fuori di $[a, b]$ è integrabile.*

Dimostrazione. Essendo integrabile, f è limitata, quindi a valori in un intervallo J limitato dove la h è proprio lipschitziana con una costante $L \geqslant 0$.

Per ogni $\varepsilon > 0$ consideriamo una suddivisione di $[a, b]$ in intervalli $I_h = [x_{h-1}, x_h]$, $1 \leqslant h \leqslant k$, tale che, posto

$$\alpha_h = \inf_{I_h} f \qquad \text{e} \qquad \beta_h = \sup_{I_h} f\,,$$

le due funzioni step

$$\varphi(x) = \sum_{h=1}^{k} \alpha_h \chi_{I_h}(x) \qquad \text{e} \qquad \psi(x) = \sum_{h=1}^{k} \beta_h \chi_{I_h}(x)$$

soddisfino la condizione

$$\int_{\mathbf{R}} (\psi - \varphi) < \frac{\varepsilon}{L}\,.$$

Per come sono definite si ha $\varphi \leqslant f \leqslant \psi$ e se adesso poniamo

$$\alpha'_h = \inf_{x \in I_h} h(f(x)) \qquad \text{e} \qquad \beta'_h = \sup_{x \in I_h} h(f(x))\,,$$

ricordando l'Esercizio 3.20 si ha

$$\begin{aligned}\sum_{h=1}^{k} (\beta'_h - \alpha'_h) m(I_h) &= \sum_{h=1}^{k} \sup_{x,y \in I_h} |h(f(x)) - h(f(y))| m(I_h) \\ &\leqslant L \sum_{h=1}^{k} \sup_{x,y \in I_h} |f(x) - f(y)| m(I_h) = L \sum_{h=1}^{k} (\beta_h - \alpha_h) m(I_h) < \varepsilon\,.\end{aligned}$$

□

10.4 Classi di funzioni integrabili

In questo paragrafo ci chiediamo: quali sono le funzioni integrabili secondo Riemann? A questo livello non è semplice dare una risposta completa, di caratterizzazioni se ne possono dare nell'ambito di teorie più generali come quella di Lebesgue, ma per ora ci dobbiamo accontentare di condizioni solo sufficienti.

Teorema 10.7 - *Ogni funzione $f : \mathbf{R} \to \mathbf{R}$ limitata, nulla al di fuori di un intervallo $[a, b]$ e monotona su $[a, b]$ è integrabile.*

Dimostrazione. Supponiamo che f sia crescente su $[a, b]$, l'altro caso è analogo. Per ogni $\epsilon > 0$ consideriamo dei punti

$$x_0 = a < x_1 < x_2 < \ldots < x_k = b$$

scelti in modo tale che $x_h - x_{h-1} < \varepsilon/(f(b) - f(a))$. Posto $I_h = [x_{h-1}, x_h]$ per $1 \leqslant h \leqslant k$, le funzioni step

$$\varphi = \sum_{h=1}^{k} f(x_{h-1})\chi_{I_h} \quad \text{e} \quad \psi = \sum_{h=1}^{k} f(x_h)\chi_{I_h}$$

soddisfano la solita relazione $\varphi \leqslant f \leqslant \psi$ essendo $f(x_{h-1}) \leqslant f(x_h)$. Inoltre

$$\int_{\mathbf{R}} (\psi - \varphi) = \sum_{h=1}^{k} (f(x_h) - f(x_{h-1}))m(I_h) < \frac{\varepsilon}{f(b) - f(a)} \sum_{h=1}^{k} (f(x_h) - f(x_{h-1})) = \varepsilon\,,$$

quindi f è integrabile.

□

Osserviamo che se f non è monotona su $[a, b]$, ma lo è su ogni intervallo $[c_{i-1}, c_i]$ di una partizione finita di $[a, b]$, le funzioni $f_i(x) = f(x)\chi_{[c_{i-1}, c_i]}(x)$, $x \in \mathbf{R}$, sono integrabili perché soddisfano le ipotesi del Teorema 10.7 e la loro somma è proprio la f che pertanto risulta integrabile.

Teorema 10.8 - *Ogni funzione $f : \mathbf{R} \to \mathbf{R}$ limitata, nulla al di fuori di un intervallo $[a, b]$ e continua eccetto al più un numero finito di punti è integrabile.*

Dimostrazione. Siano $x_i \in [a, b]$, $1 \leqslant i \leqslant m$, i punti di discontinuità di f. Scelto $\varepsilon > 0$, consideriamo gli intervalli aperti $J_\varepsilon(x_i)$ di centro x_i e ampiezza ε. Il complementare della loro unione, $K = [a, b] - \cup_i J_\varepsilon(x_i)$, è un insieme compatto sul quale f è continua e quindi uniformemente continua per il Teorema 8.46 con un suo modulo di continuità $\delta(\varepsilon) > 0$. Se adesso suddividiamo K in intervalli I_h, con $1 \leqslant h \leqslant k$, tali che $m(I_h) < \delta(\varepsilon)$, posto

$$\alpha = \inf_{\mathbf{R}} f\,,\ \beta = \sup_{\mathbf{R}} f\,, \quad \alpha_h = \inf_{I_h} f\,, \beta_h = \sup_{I_h} f\,,$$

le due funzioni step

$$\varphi = \alpha \sum_{i=1}^{m} \chi_{J_\varepsilon(x_i)} + \sum_{h=1}^{k} \alpha_h \chi_{I_h} \quad \text{e} \quad \psi = \beta \sum_{i=1}^{m} \chi_{J_\varepsilon(x_i)} + \sum_{h=1}^{k} \beta_h \chi_{I_h}$$

soddisfano la relazione $\varphi \leqslant f \leqslant \psi$ e poiché $\beta_h - \alpha_h < \varepsilon$, si ha

$$\begin{aligned}\int_{\mathbf{R}} (\psi - \varphi) &= (\beta - \alpha) \sum_{i=1}^{m} m(J_\varepsilon(x_i)) + \sum_{h=1}^{k} (\beta_h - \alpha_h)m(I_h) \\ &\leqslant (\beta - \alpha)m\varepsilon + \varepsilon(b - a) = C\varepsilon\,,\end{aligned}$$

quindi f è integrabile.

□

In particolare ogni funzione $f \in C_0^0(\mathbf{R})$, cioè continua su $\mathbf{R}$ e nulla al di fuori di un insieme limitato, è integrabile.

Fin qui, a partire dalla definizione di integrale, abbiamo sempre dovuto considerare due classi insieme, o due successioni, di funzioni step, minoranti e maggioranti la f. Ai fini pratici, per il calcolo di un integrale è preferibile la seconda opzione, costruire cioè opportune successioni approssimanti che non applicare la definizione, come abbiamo fatto del resto per l'area del segmento parabolico del § 10.1, un procedimento che trova piena giustificazione nella (10.6). È naturale allora chiedersi se, per maggiore semplicità, non sia possibile invece costruire in qualche modo una sola successione di funzioni step, anziché due, i cui integrali convergano direttamente all'integrale richiesto. Se già sappiamo che f, nulla al di fuori di $[a, b]$, è integrabile, la risposta è affermativa e ben illustrata dai numerosi *applet*, come questo, reperibili in rete. Precisamente, si considera per ogni $n \in \mathbf{N}$ una suddivisione di $[a, b]$ in n intervalli $I_h = [x_{h-1}, x_h]$, $x_0 = a < x_1 < \ldots < x_n = b$, $h = 1, \ldots, n$, di lunghezza $x_h - x_{h-1}$ infinitesima per $n \to \infty$, ad esempio $x_h = a + h(b - a)/n$. Sappiamo inoltre che esistono due successioni, (φ_n) di step minoranti e (ψ_n) di step maggioranti, costanti su ogni I_h, tali che

$$\lim_{n\to\infty} \int_{\mathbf{R}} (\psi_n - \varphi_n) = 0\,, \qquad \lim_{n\to\infty} \int_{\mathbf{R}} \varphi_n = \lim_{n\to\infty} \int_{\mathbf{R}} \psi_n = \int_{\mathbf{R}} f\,.$$

È evidente allora che se si sceglie un punto qualunque $\xi_h \in I_h$, $h = 1, \ldots, n$, a caso o con una certa regola, la successione di funzioni step definita da

$$\eta_n(x) = \sum_{h=1}^{n} f(\xi_h)\chi_{I_h}(x)\,, \qquad x \in \mathbf{R}\,,$$

è compresa tra φ_n e ψ_n, quindi

$$\lim_{n\to\infty} \int_{\mathbf{R}} \eta_n(x) = \int_{\mathbf{R}} f$$

per il Teorema dei Carabinieri. Gli integrali così costruiti

$$\int_{\mathbf{R}} \eta_n(x)\, dx = \sum_{h=0}^{n} f(\xi_h) m(I_h)$$

si chiamano *somme di Riemann.*

Esercizio 10.3 - *Dimostrare che la funzione caratteristica di un solo punto è integrabile ed ha integrale nullo. Dedurne che una funzione del tipo*

$$\varphi = \sum_{h=1}^{k} \alpha_h \chi_{\{x_h\}}\,, \tag{10.10}$$

nulla ovunque eccetto un insieme finito di punti, è integrabile ed ha integrale nullo.

Gli insiemi finiti di punti, come quello considerato nell'esercizio, ed altri anche non finiti, ma con un numero finito di punti di accumulazione, sono esempi di insiemi di misura nulla. Se una funzione integrabile viene modificata su un insieme di misura nulla rimane integrabile e il valore dell'integrale non cambia. Il motivo sta nel fatto che in realtà questa operazione equivale a passare da una f integrabile ad una $f + g$ con g avente integrale nullo in quanto, come nella (10.10), identicamente nulla al di fuori un insieme di misura nulla.

Un esempio di funzione non integrabile è la funzione di Dirichlet $\chi_{\mathbf{Q}\cap[0,1]}$, la funzione caratteristica dell'insieme dei numeri razionali dell'intervallo $[0, 1]$. Essa differisce dalla funzione integrabile identicamente nulla solo su un insieme numerabile di punti, che sono però distribuiti in modo ovunque denso in $[0, 1]$. Ogni intervallo che non si riduca ad un solo punto contiene, per la densità di $\mathbf{Q}$ in $\mathbf{R}$, sia numeri razionali che irrazionali, quindi ogni funzione step φ minorante è non positiva su $\mathbf{R}$ e ogni funzione step ψ maggiorante assume valori non inferiori a 1 su $[0, 1]$. In altre parole l'integrale inferiore di f vale 0 mentre l'integrale superiore vale 1. Ne segue che $\psi - \varphi \geqslant \chi_{[0,1]}$ e

$$\int_{\mathbf{R}} (\psi - \varphi) \geqslant 1$$

per ogni $\varphi \in \mathscr{S}^-$ e per ogni $\psi \in \mathscr{S}^+$, quindi la funzione di Dirichlet non è integrabile secondo Riemann (ma lo è secondo Lebesgue, per il quale tutti gli insiemi numerabili hanno misura nulla, ed ha integrale nullo).

10.5 Funzioni integrabili su un intervallo

Definizione 10.9 - *Se I è un intervallo limitato una funzione limitata $f : I \to \mathbf{R}$ è detta **integrabile su** I secondo Riemann se il suo prolungamento nullo al di fuori di I*

$$f^*(x) = \begin{cases} f(x) & \text{se } x \in I \\ 0 & \text{se } x \in \complement I \end{cases}$$

è integrabile secondo Riemann. In questo caso poniamo

$$\int_I f(x)\,dx = \int_{\mathbf{R}} f^*(x)\,dx\,.$$

Esercizio 10.4 - *Ogni funzione costante su I, $f(x) = c$, è integrabile e*

$$\int_I f\,dx = cm(I)\,.$$

Dai risultati fin qui ottenuti possiamo dedurre che su un intervallo limitato I ogni funzione limitata e monotona è integrabile e che ogni funzione limitata e continua, o continua eccetto i punti di un insieme di misura nulla, è integrabile. Ovviamente per il teorema di Weierstraß la limitatezza è conseguenza della continuità se I è anche chiuso. Infine, se i valori di una funzione integrabile su I vengono modificati nei punti di un insieme di misura nulla l'integrabilità si conserva e l'integrale non cambia.

Indichiamo con $\mathscr{R}(I)$ lo spazio vettoriale delle funzioni integrabili secondo Riemann sull'intervallo limitato I.

Definizione 10.10 - *Si chiama **media integrale** di $f \in \mathscr{R}(I)$ il numero*

$$\mu_I(f) = \frac{1}{m(I)} \int_I f(x)\,dx\,.$$

Spesso viene detta *a media nulla* una funzione su I con integrale nullo.

Esercizio 10.5 - *Per quale valore di $c \in \mathbf{R}$ la funzione $f(x) + c$ è a media nulla?*

Teorema 10.11 (della media) - *Per ogni $f \in \mathscr{R}(I)$ si ha*

$$\inf_I f \leqslant \mu_I(f) \leqslant \sup_I f \tag{10.11}$$

e se f è continua esiste $\xi \in I$ tale che

$$\mu_I(f) = f(\xi)\,.$$

Dimostrazione. Siano $\alpha = \inf_I f$ e $\beta = \sup_I f$. Il prolungamento nullo f^* di f al di fuori di I e le due funzioni step

$$\varphi(x) = \alpha\chi_I(x) \quad \text{e} \quad \psi(x) = \beta\chi_I(x)$$

soddisfano la relazione $\varphi \leqslant f^* \leqslant \psi$ dalla quale, per integrazione su $\mathbf{R}$, si ottiene

$$\alpha m(I) \leqslant \int_{\mathbf{R}} f^* = \int_I f \leqslant \beta m(I)\,.$$

Ciò mostra che $\mu_I(f)$ è un numero compreso tra l'estremo inferiore e l'estremo superiore e se f è continua la sua media è anche uno dei valori che f assume su I per il teorema dei valori intermedi.

□

Il significato geometrico della media integrale è evidente: poiché

$$\int_I f = \mu_I(f)m(I)\,,$$

$\mu_I(f)$ rappresenta l'altezza, con segno, che deve avere un rettangolo di base I affinché abbia la stessa area, con segno, della regione compresa tra il grafico di f e l'asse x. Ad esempio la media integrale della funzione $f(x) = x^2$ su $[0, a]$, di cui abbiamo calcolato l'integrale nel § 10.1, vale $a^2/3$ e siccome f è continua esiste un punto $\xi \in [0, a]$ tale che $f(\xi) = \xi^2 = a^2/3$. Il punto che realizza la media integrale è in questo caso $\xi = a/\sqrt{3}$. Il fatto che cada all'interno dell'intervallo non è un caso, certamente può anche essere uno degli estremi, ma si può sempre trovare un punto all'interno nel quale il valore di f coincida con la media. Infatti se in un intervallo I fosse

$$f(\xi) \neq \mu_I(f) \qquad \forall \xi \,:\, a < \xi < b\,,$$

essendo f continua deve essere vera una soltanto delle seguenti

$$f(\xi) < \frac{1}{m(I)}\int_I f\,dx \quad \text{ o } \quad f(\xi) > \frac{1}{m(I)}\int_I f\,dx \qquad \forall \xi \in (a, b)\,.$$

Se fosse vera la prima, per integrazione membro a membro e tenendo presente l'Esercizio 10.4 si otterrebbe l'assurdo

$$\int_I f\,d\xi < \left(\frac{1}{m(I)}\int_I f\,dx\right) m(I) = \int_I f\,dx\,,$$

l'altra si tratta allo stesso modo.

La media integrale è la versione continua della media aritmetica di un insieme finito di numeri a_h con $h = 1, \ldots, k$, basta considerare una f costante a tratti che sugli intervalli $[h-1, h]$ assume i valori a_h

$$\frac{1}{m[0,k]}\int_{[0,k]} f\,dx = \frac{a_1 + a_2 + \ldots + a_k}{k}\,.$$

Abbiamo visto nell'Esercizio 3.33 che ogni funzione $\varphi : \mathbf{R} \to \mathbf{R}$ convessa "alza la media"

$$\varphi\left(\frac{a_1 + a_2 + \ldots + a_n}{n}\right) \leqslant \frac{\varphi(a_1) + \varphi(a_2) + \ldots + \varphi(a_n)}{n}\,.$$

Se gli a_i sono i valori di una funzione step f, come l'abbiamo definita poco fa, possiamo riscriverla nella forma

$$\varphi\left(\frac{1}{m(I)}\int_I f\,dx\right) \leqslant \frac{1}{m(I)}\int_I \varphi(f(x))\,dx \tag{10.12}$$

che però rimane valida per ogni funzione integrabile f, non solo per una step, e si chiama *disuguaglianza di* Jensen. Ad esempio

$$\exp\left(\frac{1}{b-a}\int_{[a,b]} f(x)\,dx\right) \leqslant \frac{1}{b-a}\int_{[a,b]} e^{f(x)}\,dx\,.$$

Dal Teorema 10.6 discende prima di tutto che $\varphi \circ f \in \mathscr{R}(I)$, inoltre ci possiamo servire della ben nota condizione necessaria e sufficiente per la convessità

$$\forall y, y_0 \in \mathbf{R}\ \exists k \in \mathbf{R} \,:\, \varphi(y) \geqslant \varphi(y_0) + k(y - y_0)$$

scritta con $y_0 = \mu(f)$ e $y = f(x)$

$$\varphi(f(x)) \geqslant \varphi(\mu(f)) + k(f(x) - \mu(f))$$

da cui, integrando membro a membro, si ottiene

$$\int_I \varphi(f(x))\,dx \geqslant m(I)\varphi(\mu(f)) + k\left(\int_I f(x)\,dx - \mu(f)m(I)\right) = m(I)\varphi(\mu(f))$$

che è la (10.12).

Per una funzione f integrabile e strettamente positiva su I possiamo definire la *media armonica* e la *media geometrica* nel seguente modo

$$\nu(f) = \mu\left(\frac{1}{f}\right)^{-1} \qquad \text{e} \qquad \gamma(f) = e^{\mu(\log f)}$$

rispettivamente, espressioni che generalizzano quelle già viste nel caso discreto.

Esercizio 10.6 - *Dimostrare che*

$$\nu(f) \leqslant \gamma(f) \leqslant \mu(f)\,.$$

Esercizio 10.7 - *Il momento d'inerzia rispetto all'asse x del sottografico di una funzione positiva $f : [a,b] \to \mathbf{R}$ è definito da*

$$I_x = \frac{1}{3}\int_{[a,b]} f(x)^3\,dx\,.$$

Dimostrare che fra tutte le funzioni integrabili su $[a,b]$, il cui sottografico ha area assegnata A, quella che corrisponde al momento d'inerzia minimo è la funzione costante $f(x) = A/(b-a)$ (la soluzione è quindi il rettangolo) ed è unica.

Esercizio 10.8 - *Siano $f \in \mathscr{R}(I)$ e $\rho(x)$ una funzione integrabile non negativa (si pensi ad esempio alla densità di massa). Allora vale la seguente versione generalizzata, con la media* ***pesata****, del Teorema 10.11*

$$\inf_{x\in I} f(x)\int_I \rho(x)\,dx \leqslant \int_I f(x)\rho(x)\,dx \leqslant \sup_{x\in I} f(x)\int_I \rho(x)\,dx$$

e se $f \in C^0(I)$ esiste $\xi \in I$ tale che

$$\int_I f(x)\rho(x)\,dx = f(\xi)\int_I \rho(x)\,dx\,.$$

Proposizione 10.12 - *Se $f \in \mathscr{R}(I)$ e $J \subset I$ allora $f_{|J} \in \mathscr{R}(J)$.*

Dimostrazione. Basta osservare che $f^*_{|J}$, il prolungamento nullo di $f_{|J}$, è prodotto di due funzioni integrabili

$$f^*_{|J} = f^*\chi_J\,.$$

□

Se I ha estremi $a < b$ l'integrale viene di solito indicato anche con la notazione

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

È però conveniente dare significato a questo simbolo anche quando $a \geqslant b$ adottando la convenzione

$$\int_a^b f = -\int_b^a f \quad \text{se } a > b \qquad \text{e} \qquad \int_a^b f = 0 \quad \text{se } a = b\,,$$

coerente tra l'altro col fatto che l'integrale su un singolo punto ha valore nullo.

Teorema 10.13 - *Se $f \in \mathscr{R}(I)$ e $a, b, c \in I$ allora*

$$\int_a^b f + \int_b^c f = \int_a^c f\,.$$

Dimostrazione. Sia f^* il prolungamento nullo di f al di fuori di I e supponiamo dapprima $a < b < c$. Allora

$$\int_a^b f + \int_b^c f = \int_{\mathbf{R}} f^* \chi_{[a,b]} + \int_{\mathbf{R}} f^* \chi_{[b,c]} = \int_{\mathbf{R}} f^* (\chi_{[a,b]} + \chi_{[b,c]}) = \int_{\mathbf{R}} f^* \chi_{[a,c]} = \int_a^c f\,.$$

Se ad esempio $a < c < b$ si ha

$$\int_a^c f + \int_c^b f = \int_a^b f$$

da cui

$$\int_a^c f = \int_a^b f - \int_c^b f = \int_a^b f + \int_b^c f$$

e analogamente per gli altri casi.

□

Si noti che il teorema della media continua a valere in qualunque ordine siano presi a e b, purché si scriva la (10.11) nella forma

$$\inf_{[a,b]} f \leqslant \frac{1}{b-a} \int_a^b f(x)\,dx \leqslant \sup_{[a,b]} f\,.$$

Come applicazione del Teorema 10.13 dimostriamo che se f è una funzione localmente integrabile e T-periodica su $\mathbf{R}$ si ha

$$\int_a^{a+T} f(x)\,dx = \int_0^T f(x)\,dx \qquad \forall a \in \mathbf{R}\,.$$

Infatti

$$\int_a^{a+T} f(x)\,dx = \int_a^0 f(x)\,dx + \int_0^T f(x)\,dx + \int_T^{a+T} f(x)\,dx\,,$$

ma per la periodicità di f e l'invarianza per traslazioni della misura si ha

$$\int_T^{a+T} f(x)\,dx = \int_T^{a+T} f(x-T)\,dx = \int_0^a f(x)\,dx$$

che, inserita nella relazione precedente, dimostra l'affermazione.

Esercizio 10.9 - *Usando l'invarianza della misura per isometrie, dimostrare che se f è integrabile su $[-a, a]$ si ha*

$$f \text{ pari} \Rightarrow \int_{-a}^a f = 2\int_0^a f \qquad e \qquad f \text{ dispari} \Rightarrow \int_{-a}^a f = 0\,.$$

Esercizio 10.10 - *Dimostrare che se $f \in C^0[a, b]$, $f \geqslant 0$ e $\int_a^b f = 0$ allora $f = 0$ su $[a, b]$. Inoltre se $f \in C^0[a, b]$ ed ha integrale nullo su ogni intervallo $[\alpha, \beta] \subset [a, b]$ allora $f = 0$ su $[a, b]$.*

Esercizio 10.11 - *Dimostrare che se se $f \in C^0[a, b]$ e $\int_a^b fg \geqslant 0$ per ogni $g \in C^0_0[a, b]$ allora $f = 0$ su $[a, b]$.*

10.6 La funzione integrale

La Proposizione 10.12 ci assicura che la restrizione di una $f \in \mathscr{R}(I)$ ad ogni intervallo $[x_1, x_2] \subset I$ è integrabile e possiamo quindi considerare l'integrale

$$F(x_1, x_2) = \int_{x_1}^{x_2} f(t)\, dt \tag{10.13}$$

come una funzione degli estremi x_1 e x_2 definita su $I \times I$ che per il Teorema 10.13 soddisfa

$$F(x_1, x_2) = -F(x_2, x_1)$$

e si annulla in tutti i punti della diagonale $x_1 = x_2$ del quadrato $I \times I$. Tuttavia, una funzione $f : I \to \mathbf{R}$ può essere integrabile su ogni intervallo $[a, b] \subset I$ senza esserlo su I, basta considerare ad esempio il caso di I aperto su cui f è continua ma non limitata agli estremi. In questo paragrafo supponiamo che I sia un intervallo qualsiasi, aperto, chiuso o semiaperto, eventualmente non limitato come $\mathbf{R}$ stesso o una semiretta.

Definizione 10.14 - *La funzione $f : I \to \mathbf{R}$ viene detta* ***localmente integrabile****, e si scrive $f \in \mathscr{R}_{loc}(I)$, se $f \in \mathscr{R}[a, b]$ per ogni intervallo $[a, b] \subset I$ i cui estremi siano distinti da quelli di I.*

Ad esempio ogni funzione continua su I è localmente integrabile. La funzione (10.13) è ben definita anche se f è solo localmente integrabile. Però adesso, per i nostri scopi, ci interessa scegliere il primo estremo, diciamo un certo $a \in I$, e considerare la F come funzione solo del secondo. Definiamo quindi la *funzione integrale*

$$F_a(x) = \int_a^x f(t)\, dt \qquad \forall x \in I$$

con *punto iniziale* $a \in I$. Per prima cosa osserviamo che sostituendo a con un altro punto iniziale $\alpha \in I$ la relativa funzione integrale F_α differisce dalla F_a per una costante essendo

$$F_\alpha(x) = F_a(x) + \int_\alpha^a f(t)\, dt\,,$$

ma siccome siamo interessati a studiare le proprietà di regolarità di queste funzioni e a stabilire in che relazione stanno con la funzione integranda f, ce ne basta una soltanto, ad esempio la F_a, da cui possiamo a questo punto eliminare il pedice a. Fissato $a \in I$, poniamo dunque

$$F(x) = \int_a^x f(t)\, dt \qquad \forall x \in I\,.$$

Per esempio la funzione

$$f(x) = \begin{cases} 1 & \text{se } 0 \leqslant x \leqslant 1 \\ 2 & \text{se } 1 < x \leqslant 2 \end{cases}$$

ammette come funzione integrale su $[0, 2]$

$$F(x) = \int_0^x f(t)\, dt = \begin{cases} x & \text{se } 0 \leqslant x \leqslant 1 \\ 2x - 1 & \text{se } 1 < x \leqslant 2\,. \end{cases} \tag{10.14}$$

Si noti che f è discontinua, ma F è lipschitziana (quindi continua).

Esercizio 10.12 - *Dimostrare che se $f \geqslant 0$ allora F è crescente.*

Esercizio 10.13 - *Dimostrare che se f è costante, $f(x) = k$ per ogni $x \in I$, allora $F(x) = k(x - a)$ per ogni $x \in I$.*

L'esempio trattato nel § 10.1 ci mostra che la funzione integrale, ad esempio con $a = 0$, della $f(x) = x^2$ per $x \geqslant 0$ è la funzione $F(x) = x^3/3$ per $x \geqslant 0$.

Esercizio 10.14 - *Dimostrare che se $f : [-a, a] \to \mathbf{R}$ è localmente integrabile la funzione integrale*

$$F(x) = \int_0^x f(t)\, dt$$

è dispari se f è pari e pari se f è dispari.

Il seguente risultato spiega la continuità della (10.14).

Teorema 10.15 - *Se $f \in \mathscr{R}_{loc}(I)$ allora F è localmente lipschitziana.*

Dimostrazione. Sia $[a, b] \subset I$. Posto $L = \sup_{[a,b]} |f|$ e scelti $x_1 < x_2$ in $[a, b]$, si ha

$$|F(x_2) - F(x_1)| = \left| \int_{x_1}^{x_2} f(t)\, dt \right| \leqslant \int_{x_1}^{x_2} |f(t)|\, dt \leqslant L(x_2 - x_1)\,.$$

□

Dunque la funzione integrale F di una funzione localmente integrabile f è continua su I, ed è uniformemente continua in quanto lipschitziana se f è integrabile su I.

Vediamo adesso come aumentando la regolarità di f aumenta anche quella di F conservando sempre un po' di regolarità in più. La funzione $f(x) = 2|x|$ è continua su $\mathbf{R}$, quindi localmente intergabile, ma non è derivabile in 0, mentre la sua funzione integrale $F(x) = x|x|$ è derivabile anche in 0, sta in $C^1(\mathbf{R})$.

Teorema 10.16 (Fondamentale del Calcolo Integrale) - *Se $f \in \mathscr{R}_{loc}(I)$ è continua in un punto $x \in I$ allora la sua funzione integrale con punto iniziale $a \in I$*

$$F(x) = \int_a^x f(t)\, dt\,, \qquad x \in I\,,$$

è derivabile in x e $F'(x) = f(x)$.

Dimostrazione. Per la continuità di f nel punto x esiste una funzione $\sigma(t)$, $t \in I$, infinitesima per $t \to x$ tale che

$$f(t) = f(x) + \sigma(t) \qquad \forall t \in I\,,$$

quindi

$$F(x + h) - F(x) = \int_x^{x+h} f(t)\, dt = \int_x^{x+h} (f(x) + \sigma(t))\, dt = f(x)h + \int_x^{x+h} \sigma(t)\, dt\,.$$

Se facciamo vedere che l'ultimo termine è un $o(h)$ possiamo concludere che F è differenziabile e $F'(x) = f(x)$. Ma questo è vero perché se $h > 0$ (il caso opposto è analogo), posto

$$s(h) = \sup_{x \leqslant t \leqslant x+h} |\sigma(t)|$$

e ricordando la definizione di massimo limite, si ha

$$\lim_{h\to 0} s(h) = \lim_{h\to 0} \sup_{x\leqslant t\leqslant x+h} |\sigma(t)| = \limsup_{t\to 0} |\sigma(t)| = \lim_{t\to 0} |\sigma(t)| = 0\,,$$

quindi

$$\left| \int_x^{x+h} \sigma(t)\, dt \right| \leqslant \int_x^{x+h} |\sigma(t)|\, dt \leqslant s(h)h = o(h)\,.$$

□

Osservazione 10.17 - *Se $f \in C^0(I)$ allora F è derivabile su I e $F'(x) = f(x)$ per ogni $x \in I$, ma allora in questo caso possiamo affermare che $F \in C^1(I)$.*

Esercizio 10.15 - *Dimostrare che se $f \in C^0(I)$ e $\alpha, \beta : J \to I$ sono funzioni derivabili definite su un intervallo J allora la funzione*

$$F(x) = \int_{\alpha(x)}^{\beta(x)} f(t)\, dt \qquad \forall x \in J$$

appartiene a $C^1(J)$ e $F'(x) = f(\beta(x))\beta'(x) - f(\alpha(x))\alpha'(x)$.

Definizione 10.18 *Si chiama **primitiva** di f su I ogni funzione $G : I \to \mathbf{R}$ derivabile tale che $G'(x) = f(x)$ per ogni $x \in I$.*

Dal Teorema 10.16, insieme all'Osservazione 10.17, segue che ogni funzione continua su un intervallo ammette una primitiva perché almeno la funzione integrale lo è. In realtà le primitive sono infinite e basta conoscerne una, come la funzione integrale, per determinarle tutte. Valgono infatti le seguenti proprietà:

($\int$)1. se G è una primitiva di f lo è anche $G + c$ per ogni $c \in \mathbf{R}$, basta osservare che

$$(G(x) + c)' = G'(x) = f(x) \qquad \forall x \in \mathbf{I}\,;$$

($\int$)2. se G_1 e G_2 sono primitive di f su I esiste $c \in \mathbf{R}$ tale che $G_1 = G_2 + c$ su I, infatti

$$(G_1(x) - G_2(x))' = f(x) - f(x) = 0 \qquad \forall x \in I$$

e quindi $G_1 - G_2$ è costante.

L'insieme di tutte le primitive di f si chiama *integrale indefinito* di f, si indica col simbolo

$$\int f(x)\, dx\,, \qquad x \in I\,,$$

e può essere caratterizzato in termini della funzione integrale nel seguente modo

$$\int f(x)\, dx = \left\{ c + \int_a^x f(t)\, dt \mid c \in \mathbf{R} \right\} \qquad \forall x \in I\,.$$

Va sottolineato che l'integrale indefinito non è una funzione, bensì una famiglia a un parametro di funzioni, quindi è con un certo abuso di notazione che si usa comunemente la scrittura

$$\int f(x)\, dx = G(x) + c\,, \qquad c \in \mathbf{R}\,, \tag{10.15}$$

avendo scelto una primitiva particolare G, ma è certamente meno pesante ed altrettanto efficace della notazione insiemistica. Il seguente risultato è estremamente utile perché ci permette di ricondurre il problema del calcolo dell'integrale su un intervallo a quello della ricerca di primitive.

Teorema 10.19 (di Torricelli) - *Se $f \in C^0(I)$ e G è una primitiva di f allora*

$$\int_a^b f(x)\, dx = G(b) - G(a) \tag{10.16}$$

per ogni $a, b \in I$. Per il II membro si usa in genere la comoda notazione $[G(x)]_a^b$.

Dimostrazione. Se G è una primitiva di f esiste $c \in \mathbf{R}$ tale che

$$G(x) = c + \int_a^x f(t)\,dt \qquad \forall x \in I\,.$$

Per $x = a$ si ottiene il valore di c, che è $G(a)$, allora

$$G(x) = G(a) + \int_a^x f(t)\,dt \qquad \forall x \in I\,,$$

da cui segue la tesi scegliendo $x = b$.

$\square$

Esempi

10.1 *Delle tre regioni piane formate dal grafico della cubica $f(x) = x(x-1)^2$, $x \in \mathbf{R}$, e dall'asse x, l'area di quella limitata è*

$$\int_0^1 f(x)\,dx = \int_0^1 (x^3 - 2x^2 + x)\,dx = \left[\frac{x^4}{4} - \frac{2}{3}x^3 + \frac{x^2}{2}\right]_0^1 = \frac{1}{4} - \frac{2}{3} + \frac{1}{2} = \frac{1}{12}\,.$$

10.2 *Le funzioni* $\cos x$ *e* $\operatorname{sen} x$ *sono a media nulla, infatti*

$$\int_0^{2\pi} \cos x\,dx = \operatorname{sen} 2\pi - \operatorname{sen} 0 = 0 \qquad e \qquad \int_0^{2\pi} \operatorname{sen} x\,dx = -\cos 2\pi + \cos 0 = 0\,.$$

Più in generale, se f è T-periodica e derivabile in $\mathbf{R}$ la sua derivata, ovviamente T-periodica, è a media nulla

$$\int_0^T f'(x)\,dx = f(T) - f(0) = 0\,.$$

Le primitive di una funzione T-periodica sono T-periodiche? La funzione 2π-periodica $f(x) = 1 + \cos x$ ha per primitive le funzioni non periodiche $G(x) = x + \operatorname{sen} x + c$ che infatti sono strettamente crescenti. Se però f è a media nulla ogni sua primitiva è T-periodica e una di esse è a media nulla. Infatti la funzione integrale con punto iniziale $a \in \mathbf{R}$ soddisfa

$$F(x+T) = \int_a^{x+T} f(t)\,dt = \int_a^x f(t)\,dt + \int_x^{x+T} f(t)\,dt = \int_a^x f(t)\,dt = F(x)\,,$$

quindi anche ogni primitiva di f, $G(x) = F(x) + c$, è T-periodica e il valore di c che ci fornisce quella a media nulla è

$$c = \frac{1}{T}\int_a^{a+T} F(t)\,dt\,.$$

10.3 *Vediamo l'integrale di qualche funzione con primitive immediate*

$$\int_0^{\pi/3} \operatorname{tang}^2 x\,dx = \int_0^{\pi/3} \left(\frac{1}{\cos^2 x} - 1\right)dx = [\operatorname{tang} x - x]_0^{\pi/3} = \sqrt{3} - \frac{\pi}{3}\,,$$

$$\int_{\pi/6}^{\pi/2} \frac{1}{\operatorname{sen} x}\,dx = \int_{\pi/6}^{\pi/2} \frac{\operatorname{sen}^2 x/2 + \cos^2 x/2}{2\operatorname{sen} x/2 \cos x/2}\,dx = \ldots = \log(2+\sqrt{3})\,,$$

$$\int_1^4 \frac{e^{\sqrt{x}}}{\sqrt{x}}\,dx = 2[e^{\sqrt{x}}]_1^4 = 2e(e-1)\,,$$

$$\int_0^a \frac{\operatorname{senh} x}{\cosh^2 x}\,dx = \left[-\frac{1}{\cosh x}\right]_0^a = 1 - \frac{1}{\cosh a}\,.$$

Si faccia bene attenzione al fatto che mentre la derivazione ha carattere puntuale, l'integrazione va sempre intesa su un intervallo, altrimenti la (10.15), così com'è scritta, può risultare sbagliata. Il motivo per cui due primitive differiscono per una costante discende dalla nota conseguenza del Teorema di Lagrange, valido solo su un intervallo, secondo cui una funzione con derivata nulla è necessariamente costante. Invece le primitive di una funzione definita su due intervalli disgiunti, la cui unione non forma un intervallo, non differiscono per una costante dal momento che le primitive ristrette ad uno di essi e all'altro sono completamente indipendenti, è costante la differenza su ogni intervallo, ma non sull'unione. Si tratta in realtà di integrali indefiniti di due funzioni distinte e quindi di due famiglie a un parametro di primitive. Se è vero che per ogni $x \in \mathbf{R} - \{0\}$ si ha

$$\frac{d}{dx} \log|x| = \frac{1}{x}$$

sarebbe invece errato scrivere

$$\int \frac{1}{x}\, dx = \log|x| + c\,, \qquad c \in \mathbf{R}\,,$$

sebbene molto diffuso. Le due funzioni

$$G_1(x) = \begin{cases} \log(-x) & \text{se } x < 0 \\ \log x & \text{se } x > 0 \end{cases} \qquad \text{e} \qquad G_2(x) = \begin{cases} \log(-x) & \text{se } x < 0 \\ 1 + \log x & \text{se } x > 0 \end{cases}$$

sono entrambe primitive di $1/x$, ma la loro differenza non è costante, $G_2 - G_1$ vale 0 per $x < 0$ e 1 per $x > 0$. Il modo concettualmente corretto di scrivere è il seguente

$$\int \frac{1}{x}\, dx = c_1 + \log x\,,\ x \in]0, +\infty[\,, \qquad \int \frac{1}{x}\, dx = c_2 + \log(-x)\,,\ x \in]-\infty, 0[\,,$$

al variare di $c_1, c_2 \in \mathbf{R}$, dove è ovviamente lecito, ma inutile, scrivere $\log|x|$.

Un po' diversa è la situazione per la funzione $f(x) = 1/\sqrt{|x|}$, la quale, pur essendo definita su $\mathbf{R} - \{0\}$ come la precedente, ammette primitive per $x < 0$ e $x > 0$ che hanno limite finito per $x \to 0$. Ogni primitiva a sinistra può essere prolungata in 0 e raccordata con continuità ad una particolare primitiva a destra mettendo in relazione opportuna le due relative costanti. Precisamente, pur in un senso generalizzato, ammette come integrale indefinito

$$\int \frac{1}{\sqrt{|x|}}\, dx = 2\sqrt{|x|}\,\mathrm{sign}\, x + c\,, \qquad c \in \mathbf{R}\,.$$

Evidentemente la singolarità in 0 della f è la causa della non derivabilità in 0 delle sue primitive, punto nel quale comunque esse rimangono continue, anzi hölderiane.

Il Teorema 10.16 ci dice che per una funzione $F \in C^1(I)$ che si annulla in a vale la relazione

$$F(x) = \int_a^x F'(t)\, dt \qquad \forall x \in I$$

e analogamente anche la (10.16) può essere scritta nella forma

$$G(b) - G(a) = \int_a^b G'(t)\, dt\,. \tag{10.17}$$

Viene allora spontaneo chiedersi se la (10.17) non possa valere anche quando venisse a mancare la derivata di G in un punto o in qualche punto, dato che il valore dell'integrale non dipende da quello che succede su un insieme di misura nulla. Ad esempio si vede subito che la (10.17) è vera per la funzione F definita nella (10.14) anche se non è derivabile per $x = 1$. Se però consideriamo la funzione $H : [-1, 1] \to \mathbf{R}$

$$H(x) = \begin{cases} 0 & \text{se} \quad -1 \leqslant x < 0 \\ 1 & \text{se} \quad 0 \leqslant x \leqslant 1 \end{cases}$$

si ha $H(1) - H(-1) = 1$, mentre

$$\int_{-1}^{1} H'(t)\, dt = 0$$

essendo $H'(x) = 0$ ovunque, eccetto per $x = 0$ dove non è derivabile. In che cosa differiscono le due situazioni? Eppure sia F che H sono derivabili ovunque, ad eccezione di un solo punto. Certamente la funzione integrale può non essere di classe C^1 e non avere la derivata in qualche punto, ma affinché valga la (10.17) deve essere almeno continua e la H non lo è. Se viene a mancare la continuità di G la (10.17) non è più vera. Come va corretta? Se ad esempio $G : [a, b] \to \mathbf{R}$ è discontinua in $x_0 \in]a, b[$ e derivabile in tutti gli altri punti, applicando correttamente il Teorema 10.19 si ottiene

$$\int_a^b G' = \int_a^{x_0} G' + \int_{x_0}^b G' = G(x_0^-) - G(a) + G(b) - G(x_0^+) = G(b) - G(a) - [\![G]\!]_{x_0}$$

se si ammette, per semplicità, che esistano finiti i limiti destro e sinistro in x_0. Se dunque gli x_i sono i punti di salto di G la (10.17) va corretta nella forma

$$G(b) - G(a) = \int_a^b G'(x)\, dx + \sum_i [\![G]\!]_{x_i} .$$

10.7 Integrali impropri

Vogliamo adesso estendere la nozione di integrale a funzioni non limitate, ma definite su un intervallo limitato, e a funzioni definite su un intervallo non limitato. Evidentemente anche l'area di una regione piana illimitata può essere finita.

Definizione 10.20 - *Una funzione $f : [a, b[\to \mathbf{R}$, $b < +\infty$, non necessariamente limitata in un intorno di b viene detta **integrabile in senso improprio**, o anche **in senso generalizzato**, se è localmente integrabile ed esiste finito il limite*

$$\lim_{\beta \to b} \int_a^{\beta} f(x)\, dx .$$

In tal caso assumiamo questo limite come il valore dell'integrale di f su $[a, b[$. In modo analogo si definisce l'integrale improprio nel caso di f non limitata nell'intorno di a.

Che si tratti di un'estensione dell'integrale di Riemann è evidente. Se $f \in \mathscr{R}[a, b]$ per il Teorema 10.15 ogni primitiva G è lipschitziana quindi uniformemente continua, pertanto esiste finito il limite

$$\lim_{\beta \to b} [G(\beta) - G(a)] = \lim_{\beta \to b} \int_a^{\beta} f(x)\, dx$$

e siccome G è continua fino a b questo limite coincide con $G(b) - G(a)$ che è l'integrale di Riemann di f su $[a, b]$. È allora naturale indicare con $\mathscr{R}(I)$ anche il più ampio insieme delle funzioni integrabili in senso generalizzato su I.

Definizione 10.21 - *Se $f : [a,c[\cup]c,b] \to \mathbf{R}$ non è limitata nell'intorno di c diciamo che è* ***integrabile in senso improprio*** *su $[a,b]$ se lo sono le sue restrizioni ai due intervalli $[a,c[$ e $]c,b]$ separatamente e si pone*

$$\int_a^b f\,dx = \int_a^c f\,dx + \int_c^b f\,dx\,.$$

Se vi sono più punti nell'intorno dei quali f non è limitata quanto detto si estende in modo ovvio.

10.4 *La funzione $f(x) = 1/x^\alpha$ non è limitata su $]0,1]$ nell'intorno di 0 se $\alpha > 0$ (il problema non si pone se $\alpha \leqslant 0$). Ovviamente, in quanto continua, è limitata e integrabile su ogni intervallo $[h,1]$. Vediamo per quali $\alpha \in \mathbf{R}$ è integrabile in senso improprio. L'integrale*

$$\int_h^1 \frac{1}{x^\alpha}\,dx = \begin{cases} \left[\dfrac{x^{1-\alpha}}{1-\alpha}\right]_h^1 = \dfrac{1-h^{1-\alpha}}{1-\alpha} & \textit{se } \alpha \neq 1 \\ [\log x]_h^1 = -\log h & \textit{se } \alpha = 1 \end{cases} \tag{10.18}$$

è convergente per $h \to 0$ se e solo se $0 < \alpha < 1$ e il suo valore è $(1-\alpha)^{-1}$.

Caccia all'errore: La funzione $1/x$ è integrabile in senso improprio su $[-1,1]$

$$\int_{-1}^1 \frac{1}{x}\,dx = [\log|x|]_{-1}^1 = \log 1 - \log 1 = 0\,. \tag{10.19}$$

L'errore sta nel fatto che la funzione in questione non soddisfa la Definizione 10.21. Poiché non è limitata nell'intorno di 0, bisogna studiare separatamente il suo comportamento sui due intervalli $[-1,0[$ e $]0,1]$ e su nessuno dei due è integrabile in senso improprio. Il fatto che l'integrale da una parte possa compensare l'integrale dall'altra, nel senso che

$$\lim_{\varepsilon\to 0}\left[\int_{-1}^{-\varepsilon} \frac{1}{x}\,dx + \int_\varepsilon^1 \frac{1}{x}\,dx\right] = 0\,, \tag{10.20}$$

non è rilevante, basta infatti scegliere come estremi dipendenti da ε due espressioni tra loro un po' diverse, come ε da una parte e 2ε o ε^2 dall'altra, e il limite della somma può assumere qualsiasi valore in $\bar{\mathbf{R}}$. Affinché la funzione sia integrabile, è necessario e sufficiente che questo limite non dipenda da come vengono scelti gli estremi in funzione di ε. Nella (10.20) esiste ed è nullo il limite della somma, ma non è definita la somma dei limiti. C'è anche un altro errore nella (10.19), ma ne abbiamo parlato nel § 10.6.

Neanche la funzione $1/\sqrt{|x|}$ è limitata in un intorno di 0, ma è integrabile in senso improprio su $[-1,0[\cup]0,1]$, infatti

$$\lim_{\varepsilon\to 0}\int_\varepsilon^1 \frac{1}{\sqrt{x}}\,dx = 2\lim_{\varepsilon\to 0}[\sqrt{x}]_\varepsilon^1 = 2$$

e trattandosi di una funzione pari il suo integrale su $[-1,1]$ vale 4.

Definizione 10.22 - *Una funzione $f : [a,+\infty[\to \mathbf{R}$ viene detta* ***integrabile in senso improprio****, o anche* ***in senso generalizzato****, se è localmente integrabile in $[a,+\infty[$ ed esiste finito il limite*

$$\lim_{b\to+\infty}\int_a^b f(x)\,dx\,.$$

In tal caso assumiamo questo limite come il valore dell'integrale

$$\int_a^{+\infty} f(x)\,dx\,.$$

In modo analogo si definisce l'integrale su $]-\infty, 0]$. *Infine* $f : \mathbf{R} \to \mathbf{R}$ *è integrabile in senso improprio se è localmente integrabile e sono integrabili in senso improprio, separatamente, le restrizioni di* f *ai due intervalli* $]-\infty, 0]$ *e* $[0, +\infty[$. *In tal caso si pone*

$$\int_{\mathbf{R}} f(x)\,dx = \int_{-\infty}^{+\infty} f(x)\,dx = \int_{-\infty}^{0} f(x)\,dx + \int_0^{+\infty} f(x)\,dx\,.$$

Anche qui bisogna stare attenti coi passaggi al limite, la funzione $f(x) = x$, $x \in \mathbf{R}$, non è integrabile su $\mathbf{R}$ sebbene

$$\lim_{a\to+\infty} \int_{-a}^{a} x\,dx = 0\,,$$

ma si tratta di una funzione non integrabile né su $[0, +\infty[$, né su $]-\infty, 0]$.

La situazione più generale è quella di una funzione definita su un certo numero di intervalli $]-\infty, a_1[$, $]a_1, a_2[$, ..., $]a_{k-1}, a_k[$, $]a_k, +\infty[$ e non limitata vicino agli a_i. Per essere integrabile in senso improprio sulla loro unione deve esserlo su ogni intervallo, il valore dell'integrale è la somma degli integrali sui vari intervalli.

Esempi

10.5 *La funzione* e^{-x} *è integrabile su* $[0, +\infty[$, *infatti*

$$\int_0^h e^{-x}\,dx = [-e^{-x}]_0^h = 1 - e^{-h} \to 1 \quad \textit{per } h \to \infty\,.$$

10.6 *Vediamo per quali* $\alpha \in \mathbf{R}$ *la funzione* $f(x) = 1/x^\alpha$ *è integrabile in senso improprio su* $[1, +\infty[$. *Per* $h > 1$ *l'integrale su* $[1, h]$

$$\int_1^h \frac{1}{x^\alpha}\,dx = \begin{cases} \left[\dfrac{x^{1-\alpha}}{1-\alpha}\right]_1^h = \dfrac{h^{1-\alpha}-1}{1-\alpha} & \textit{se } \alpha \neq 1 \\ [\log x]_1^h = \log h & \textit{se } \alpha = 1 \end{cases}$$

è convergente per $h \to +\infty$ *se e solo se* $\alpha > 1$ *e il suo valore è* $(\alpha - 1)^{-1}$.

10.7 *La funzione* $\operatorname{sen} x$ *su* $[0, +\infty[$ *non è integrabile perché* $[-\cos t]_0^x = 1 - \cos x$ *non ha limite per* $x \to +\infty$.

Esercizio 10.16 - *Dimostrare che la funzione* Γ *di Eulero*

$$\Gamma(x) = \int_0^{+\infty} t^{x-1} e^{-t}\,dt \qquad \forall x > 0$$

è ben definita per $x > 0$ *e soddisfa la relazione* $\Gamma(x+1) = x\Gamma(x)$. *Dedurne per induzione che* $\Gamma(n+1) = n!$ *per ogni* $n \in \mathbf{N}$, *quindi si tratta di un'estensione ai reali positivi del fattoriale.*

Esercizio 10.17 - *Stabilire per quali* $p, q, r \in \mathbf{R}$ *è convergente l'integrale*

$$\int_e^{+\infty} \frac{dx}{x^p (\log x)^q (\log\log x)^r}\,.$$

Illustriamo adesso alcuni criteri di convergenza degli integrali impropri in una teoria unificata, senza dover distinguere il caso contemplato nella Definizione 10.20 da quello della Definizione 10.22. Salvo precisazioni esplicite, indicheremo con I un intervallo $[a, b[$ con $a \in \mathbf{R}$, mentre b può essere un numero reale oppure $+\infty$.

Una prima osservazione ovvia è la seguente

$$\exists \text{ finito } \lim_{x\to b} \int_a^x f(t)\,dt = \int_a^b f(t)\,dt \Leftrightarrow \lim_{x\to b} \int_x^b f(t)\,dt = 0\,.$$

Ricordando invece la Definizione 8.4 e la Proposizione 8.5, si può dedurre il seguente utile criterio di convergenza.

Teorema 10.23 (Criterio di Cauchy) - *Una funzione $f \in \mathscr{R}_{loc}(I)$ è integrabile in senso improprio se e solo se per ogni $\varepsilon > 0$ esiste $c \in I$ tale che*

$$\left| \int_{x_1}^{x_2} f(t)\,dt \right| < \varepsilon$$

per ogni $x_1, x_2 \in]c, b[$.

Dimostrazione. Sia F la funzione integrale di f con punto iniziale a. Esiste finito il limite

$$\lim_{x\to b} F(x) = \int_a^b f(t)\,dt$$

se e solo se per ogni $\varepsilon > 0$ esiste $c \in]a, b[$ tale che

$$|F(x_1) - F(x_2)| = \left| \int_{x_1}^{x_2} f(t)\,dt \right| < \varepsilon \tag{10.21}$$

per ogni $x_1, x_2 \in]c, b[$.

□

In termini di successioni, $f \in \mathscr{R}(I)$ se e solo se per ogni coppia di successioni $x_n, y_n \in I$

$$\lim_{n\to\infty} |x_n - y_n| = 0 \Rightarrow \lim_{n\to\infty} \left| \int_{x_n}^{y_n} f(t)\,dt \right| = 0\,.$$

Si noti che la funzione integrale è già uniformemente continua, in quanto lipschitziana, su ogni intervallo compatto contenuto in I dove, per ipotesi, la f è integrabile secondo Riemann. La condizione aggiuntiva (10.21) ci permette di affermare che $f \in \mathscr{R}_{loc}(I)$ è integrabile su I se e solo se F è uniformemente continua su I (senza che però sia su I necessariamente lipschitziana!) come mostra l'esempio della $f(x) = 1/\sqrt{x}$ su $]0, 1]$. Si potrebbe obiettare che in questo caso la funzione $F(x) = 2\sqrt{x}$ è uniformemente continua in quanto hölderiana e non come primitiva di una funzione integrabile, ma se si considera la $f(x) = 1/(x \log^2 x)$, $0 < x < 1/2$, questa è integrabile perché ammette per primitiva $F(x) = -1/\log x$ che è uniformemente continua ma non hölderiana.

La non integrabilità della funzione $\operatorname{sen} x$ su $[0, +\infty[$ può essere verificata anche così

$$2n\pi, (2n+1)\pi \to +\infty\,, \quad \text{ma} \quad \int_{2n\pi}^{(2n+1)\pi} \operatorname{sen} x\,dx = 2$$

che di certo non tende a 0.

Teorema 10.24 (Criterio di confronto) - *Siano $f \in \mathscr{R}_{loc}(I)$, $h, g \in \mathscr{R}(I)$ e $g \leqslant f \leqslant h$. Allora anche $f \in \mathscr{R}(I)$.*

Dimostrazione. Si può usare la condizione di Cauchy (con $x_1 < x_2$)

$$-\varepsilon < \int_{x_1}^{x_2} g(t)\,dt \leqslant \int_{x_1}^{x_2} f(t)\,dt \leqslant \int_{x_1}^{x_2} h(t)\,dt < \varepsilon$$

da cui

$$\left|\int_{x_1}^{x_2} f(t)\,dt\right| < \varepsilon\,.$$

□

Si noti che se $f \geqslant 0$ su I basta la stima superiore ad assicurarne l'integrabilità e in questo caso si può applicare un ragionamento alternativo basato sulla monotonia: la funzione integrale di f

$$F(x) = \int_a^x f(t)\,dt$$

è crescente e quindi ammette limite per $x \to b$, ma questo limite è finito perché

$$F(x) \leqslant H(x) = \int_a^x h(t)\,dt$$

e H converge crescendo all'integrale di h su I per $x \to b$.

10.8 *Per ogni $\alpha \in \mathbf{R}$ la funzione $x^\alpha e^{-x}$ è integrabile in senso improprio su $[1, +\infty[$ (è integrabile anche su $]0, +\infty[$, ma solo per $\alpha > -1$ da quanto visto nella (10.18)).*

Infatti ciò è vero per $\alpha = 0$ come abbiamo già visto. Se $\alpha < 0$ $x^\alpha e^{-x} \leqslant e^{-x}$. Per $\alpha > 0$ basta ricordare che per x abbastanza grande $x^\alpha < e^{x/2}$, quindi moltiplicando per e^{-x} si ottiene $x^\alpha e^{-x} < e^{-x/2}$. Per confronto è integrabile.

10.9 *La funzione $f(x) = e^{-x^2}$ è integrabile su $[0, +\infty[$ (e quindi su $\mathbf{R}$ essendo pari).*

Le primitive di questa funzione non sono esprimibili in forma chiusa mediante funzioni elementari, ma con un passaggio all'integrale doppio, che verrà illustrato nel corso di Analisi 2, è possibile dimostrare che l'integrale su $[0, +\infty[$ vale $\sqrt{\pi}/2 \sim 0.886227$. Data l'importanza che la funzione integrale normalizzata riveste nell'ambito delle equazioni differenziali alle derivate parziali, in Statistica e nel Calcolo delle Probabilità, è stata introdotta direttamente col nome di *funzione degli errori di Gauss*

$$\operatorname{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-t^2}\,dt \qquad \forall x \in \mathbf{R}\,,$$

Per ora ci accontentiamo di dimostrare che l'integrale converge. Un metodo consiste

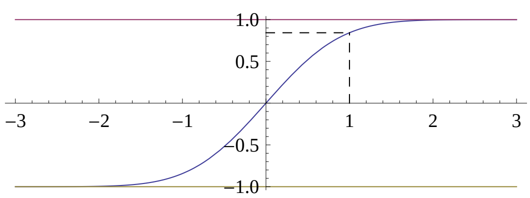


Figura 10.2: La funzione degli errori di Gauss.

nel maggiorare la funzione su $[1, +\infty[$ con e^{-x}, ma si ottiene una stima più precisa per confronto con la funzione xe^{-x^2} che si integra facilmente

$$\int_1^{+\infty} e^{-x^2}\,dx < \int_1^{+\infty} xe^{-x^2}\,dx = -\frac{1}{2}\int_1^{+\infty} -2xe^{-x^2}\,dx = -\frac{1}{2}\lim_{a\to+\infty}[e^{-x^2}]_1^a = \frac{1}{2e}\,.$$

Possiamo anche fornire una stima dell'integrale su $[0,1]$ usando la disuguaglianza $e^{x^2} \geqslant 1+x^2$

$$\int_0^1 e^{-x^2}\,dx < \int_0^1 \frac{1}{1+x^2}\,dx = \operatorname{arctg} 1 = \frac{\pi}{4}\,.$$

Sommando i due integrali si ottiene

$$\int_0^{+\infty} e^{-x^2}\,dx < \frac{\pi}{4} + \frac{1}{2e} \sim 0.969338\,.$$

Il Teorema 10.24 è utile anche come *criterio integrale* di convergenza per alcune serie. Ogni termine a_n è il valore costante che la funzione $f : [0,+\infty[\to \mathbf{R}$ definita da

$$f(x) = \sum_{n=1}^{\infty} a_n \chi_{[n-1,n]}(x) \qquad \forall x \in [0,+\infty[$$

assume su $[n-1,n]$. Dal momento che la somma della serie coincide col limite per $n\to\infty$ dell'integrale di f su $[0,n]$, per verificare che la serie converge basta trovare due funzioni g,h, $g \leqslant f \leqslant h$, di facile integrazione con integrale convergente. Oppure, nei casi divergenti, trovare $g \leqslant f$ con integrale divergente positivamente o $h \geqslant f$ con integrale divergente negativamente.

Possiamo ad esempio rivedere con questa tecnica il comportamento della serie

$$\sum_{n=2}^{\infty} \frac{1}{n\log^{\alpha} n}$$

al variare di $\alpha > 0$. Per ogni $x \in [n-1,n]$, $n \geqslant 2$, si ha

$$\frac{1}{n\log^{\alpha} n} \leqslant \frac{1}{x\log^{\alpha} x}$$

e per $\alpha > 1$

$$\int_2^{+\infty} \frac{1}{x\log^{\alpha} x}\,dx = \frac{-1}{\alpha-1}\lim_{h\to+\infty}\left[\frac{1}{\log^{\alpha-1} x}\right]_2^h = \frac{1}{(\alpha-1)\log^{\alpha-1} 2}\,,$$

pertanto la serie converge per $\alpha > 1$. Si vede facilmente in modo simile che diverge per $\alpha < 1$.

Esercizio 10.18 *Dimostrare il seguente criterio di confronto asintotico. Siano $f,g \geqslant 0$ due funzioni localmente integrabili su I tali che*

$$\lim_{x\to b}\frac{f(x)}{g(x)} = L \in \mathbf{R}\,,\ L > 0\,.$$

Se una delle due funzioni è integrabile su I lo è anche l'altra e se una non lo è neanche l'altra lo è.

10.10 *Per ogni $\alpha > 0$ la funzione $f(x) = (\operatorname{sen} x)/x^{\alpha}$ è integrabile su $[1,+\infty[$ (ovviamente non lo è per $\alpha \leqslant 0$).*

La situazione è simile a quella delle serie a segni alterni, l'integrale converge perché i termini che danno contributi di segni opposti si compensano a vicenda. Per ogni $x_1, x_2 \geqslant 1$ si ha

$$\int_{x_1}^{x_2} \frac{\operatorname{sen} x}{x^{\alpha}}\,dx = \frac{\cos x_1}{x_1^{\alpha}} - \frac{\cos x_2}{x_2^{\alpha}} - \alpha\int_{x_1}^{x_2} \frac{\cos x}{x^{\alpha+1}}\,dx\,.$$

I primi due termini vanno a 0 per $x_1, x_2 \to +\infty$ e riguardo al terzo

$$\left| \int_{x_1}^{x_2} \frac{\cos x}{x^{\alpha+1}}\, dx \right| \leqslant \int_{x_1}^{x_2} \left| \frac{\cos x}{x^{\alpha+1}} \right| dx \leqslant \int_{x_1}^{x_2} \frac{1}{x^{\alpha+1}}\, dx = \frac{1}{\alpha}\left(\frac{1}{x_1^\alpha} - \frac{1}{x_2^\alpha} \right) \to 0\,,$$

come c'era da aspettarsi dato che $1/x^{\alpha+1}$ è integrabile.

Esercizio 10.19 - *Verificare che sull'intervallo* $]0,1]$ *la* f *di quest'ultimo esempio è integrabile per* $\alpha < 2$. *Per* $\alpha \leqslant 1$ *è addirittura continua e limitata e quindi integrabile nel senso di Riemann.*

Il seguente criterio di integrabilità ci ricorda da vicino quello di Dirichlet per le serie, v. Teorema 6.10.

Teorema 10.25 (Criterio di Abel-Dirichlet) - *Siano* f *derivabile e decrescente in* $[a, b[$ *e infinitesima per* $x \to b$ *e* $\psi \in C^0[a, b[$ *la cui funzione integrale*

$$\Psi(x) = \int_a^x \psi(t)\, dt$$

è limitata. Allora ψf *è integrabile in* $[a, b[$.

Dimostrazione. Per verificare la condizione di Cauchy scegliamo $x_1, x_2 \in [a, b[$ e integriamo per parti

$$\int_{x_1}^{x_2} \psi(x) f(x)\, dx = \Psi(x_2) f(x_2) - \Psi(x_1) f(x_1) - \int_{x_1}^{x_2} \Psi(x) f'(x)\, dx\,. \tag{10.22}$$

Poiché $-f'(x) \geqslant 0$ per ogni $x \in [a, b[$, per il teorema della media generalizzata, v. l'Esercizio 10.8, e per la continuità di Ψ esiste un punto $\xi \in [x_1, x_2]$ tale che

$$\int_{x_1}^{x_2} \Psi(x)(-f'(x))\, dx = \Psi(\xi) \int_{x_1}^{x_2} -f'(x)\, dx = \Psi(\xi)(f(x_1) - f(x_2))\,. \tag{10.23}$$

Ora $|\Psi| \leqslant M$ su $[a, b[$ per un certo $M > 0$, quindi si ottiene dalle (10.22) e (10.23)

$$\left| \int_{x_1}^{x_2} \psi(x) f(x)\, dx \right| \leqslant M(f(x_2) + f(x_1)) + M(f(x_1) - f(x_2)) = 2M f(x_1) \to 0$$

per $x_1 \to b$ (e per $x_2 > x_1$ che per forza! tende a b).

□

Così ancora una volta viene confermata la convergenza dell'integrale dell'Esempio 10.10. A questo tipo di integrale si riducono gli *integrali di Fresnel*

$$\int_0^{+\infty} \cos x^2\, dx \qquad \text{e} \qquad \int_0^{+\infty} \operatorname{sen} x^2\, dx$$

col semplice cambio di variabile $y = x^2$.

Esercizio 10.20 - *Stabilire per quali valori di* $\alpha \in \mathbf{R}$ *sono convergenti gli integrali*

$$\int_0^{+\infty} \cos x^\alpha\, dx \qquad e \qquad \int_0^{+\infty} \operatorname{sen} x^\alpha\, dx\,.$$

Gli esempi fatti finora, sugli integrali impropri relativi ad un intervallo non limitato, ci inducono a credere che una funzione $f : [a, +\infty[\to \mathbf{R}$ debba essere infinitesima per $x \to +\infty$ affinché sia integrabile, come del resto avviene per la successione dei termini di una serie numerica, ma non è così. Certamente se f ammette limite $L > 0$ $[L < 0]$ per $x \to +\infty$ l'integrale diverge positivamente [negativamente] perché per x

abbastanza grande $f(x) > L-\varepsilon$ $[f(x) < L+\varepsilon]$ e la funzione costante $L-\varepsilon$ $[L+\varepsilon]$ non ha integrale finito, però si costruiscono facilmente esempi in cui la f non ha limite, e magari non è neanche limitata, con integrale convergente. Ad esempio la funzione continua

$$f(x) = \sum_{n=2}^{\infty}(n - n^4|x-n|)\chi_{[1-\frac{1}{n^3},1+\frac{1}{n^3}]}(x) \qquad \forall x \geqslant 0\,,$$

il cui grafico presenta dei picchi triangolari sempre più alti, ma più stretti, in corrispondenza dei punti $x = n$, è integrabile perché

$$\lim_{a\to+\infty}\int_0^a f = \sum_{n=2}^{\infty}\frac{1}{n^2} < +\infty\,.$$

Si capisce che questo fenomeno viene impedito se si impone una limitazione alla pendenza del grafico.

Proposizione 10.26 - *Se $f : [0,+\infty[\to \mathbf{R}$ è lipschitziana e integrabile allora f è infinitesima per $x \to +\infty$.*

Dimostrazione. Ragioniamo per assurdo assumendo che $f(x)$ non tenda a 0. Non potendo avere limite, altrimenti l'integrale sarebbe divergente, deve ammettere più punti limite per $x \to +\infty$ che formano un intervallo L perché f è continua. Precisamente, $L = [0,l]$ o $L = [l',0]$ (devono contenere lo 0 altrimenti l'integrale diverge) oppure $L = [l',l]$. In ogni caso (a parte il secondo che è analogo al primo) fissiamo un $a \in \mathbf{R}$ tale che $0 < a < l$ e consideriamo il *sopralivello* $A = \{x \in [0,+\infty[\mid f(x) \geqslant a\}$. Sempre per la continuità di f, tale insieme è una unione numerabile di intervalli disgiunti I_n. Scegliamo adesso per ogni $n \in \mathbf{N}$ due punti $x_n, y_n \in I_n$, dunque $x_n, y_n \to +\infty$, tali che $x_n < y_n$, $f(x_n) \to \alpha$ e $f(y_n) \to \beta$ con $a \leqslant \alpha < \beta \leqslant l$. Per il Teorema della media integrale esiste $\xi_n \in [x_n, y_n]$ tale che

$$\int_{x_n}^{y_n} f(t)\,dt = f(\xi_n)(y_n - x_n)\,.$$

L'integrale a primo membro tende a 0 per $n \to \infty$ perché per ipotesi l'integrale di f su $[0,+\infty[$ è convergente e $\liminf_{n\to\infty} f(\xi_n) \geqslant a > 0$, quindi $y_n - x_n \to 0$. Ne segue che

$$\lim_{n\to\infty}\frac{f(y_n)-f(x_n)}{y_n - x_n} = +\infty$$

contro l'ipotesi che f sia lipschitziana.

□

A differenza di quanto avviene per le funzioni integrabili secondo Riemann, una funzione può essere integrabile in senso improprio senza che lo sia il suo valore assoluto.

Definizione 10.27 - *Una funzione $f \in \mathscr{R}_{loc}(I)$ viene detta* ***assolutamente integrabile*** *se*

$$\int_I |f(x)|\,dx < +\infty\,.$$

La situazione è analoga a quella delle serie assolutamente convergenti: ci si aspetta che le funzioni assolutamente integrabili sono integrabili. Le funzioni dell'Esempio 10.10 $f(x) = (\operatorname{sen} x)/x^\alpha$ sono tutte integrabili per $\alpha > 0$ su $[1,+\infty[$ (e per $0 < \alpha < 2$ anche su $[0,+\infty[$), ma solo per $\alpha > 1$ lo sono assolutamente. Per $\alpha = 1$, ma si ragiona allo stesso modo per $\alpha < 1$, si ha per ogni $n \geqslant 2$

$$\int_\pi^{n\pi}\left|\frac{\operatorname{sen} x}{x}\right|dx = \sum_{k=2}^{n}\int_{(k-1)\pi}^{k\pi}\frac{|\operatorname{sen} x|}{x}\,dx > \sum_{k=2}^{n}\frac{1}{k\pi}\int_{(k-1)\pi}^{k\pi}|\operatorname{sen} x|\,dx = \frac{2}{\pi}\sum_{k=2}^{n}\frac{1}{k}$$

dove l'ultimo termine diverge positivamente per $n \to \infty$.

Teorema 10.28 - *Ogni funzione $f \in \mathscr{R}_{loc}(I)$ assolutamente integrabile su I è integrabile.*

Dimostrazione. Possiamo usare, come del resto abbiamo fatto per le serie, la condizione di Cauchy. Supponiamo di dover studiare la convergenza dei nostri integrali su $I = [a, b[$ dove eventualmente $b = +\infty$. Per ipotesi la funzione

$$\Phi(x) = \int_a^x |f(t)|\, dt\,, \qquad a \leqslant x < b\,,$$

ammette limite finito per $x \to b$. Allora soddisfa la condizione di Cauchy in b

$$\forall \varepsilon > 0\ \exists U \in \mathscr{I}(b)\ :\ \forall x_1, x_2 \in I \cap U \quad |\Phi(x_1) - \Phi(x_2)| < \varepsilon\,.$$

Posto

$$F(x) = \int_a^x f(t)\, dt\,,$$

per gli stessi $x_1, x_2 \in I$ (supponiamo $x_1 < x_2$) si ha

$$|F(x_2) - F(x_1)| = \left| \int_{x_1}^{x_2} f(t)\, dt \right| \leqslant \int_{x_1}^{x_2} |f(t)|\, dt = \Phi(x_2) - \Phi(x_1) < \varepsilon\,.$$

Pertanto anche F è di Cauchy in b e quindi converge per $x \to b$.

□

Per concludere vediamo il carattere asintotico della media per una funzione definita su $[0, +\infty[$, cioè il comportamento per $x \to +\infty$ della funzione

$$\mu(x) = \frac{1}{x} \int_0^x f(t)\, dt$$

con f localmente integrabile. È ovvio che nel caso f sia integrabile su $[0, +\infty[$ si ha $\lim_{x\to+\infty} \mu(x) = 0$, ma questo è l'analogo di ciò che avviene con una serie convergente: se a_n è una successione di numeri reali le cui somme parziali s_n convergono si ha $\mu_n = s_n/n \to 0$. Il fatto però che la successione μ_n delle medie sia infinitesima non è dovuto al carattere convergente della s_n, ma al carattere infinitesimo della a_n che è condizione solo necessaria. Questo perché sappiamo che la μ_n ha lo stesso limite della a_n quando questo esiste. Non c'è da sorprendersi se anche con gli integrali avviene la stessa cosa. Precisamente, dimostriamo che se f ha limite $l \in \bar{\mathbf{R}}$ per $x \to +\infty$ allora anche μ ha lo stesso limite.

Supponiamo dapprima che l sia finito. Fissato un $\varepsilon > 0$ arbitrario, esiste $c > 0$ tale che $l - \varepsilon < f(t) < l + \varepsilon$ per ogni $t > c$. Per $x > c$ si ha

$$\mu(x) = \frac{1}{x} \int_0^c f(t)\, dt + \frac{1}{x} \int_c^x f(t)\, dt$$

dove il primo termine tende a 0 per $x \to +\infty$, mentre il secondo soddisfa

$$\frac{1}{x}(l - \varepsilon)(x - c) < \frac{1}{x} \int_c^x f(t)\, dt < \frac{1}{x}(l + \varepsilon)(x - c) \qquad .$$

Passando al limite a sinistra e a destra per $x \to +\infty$, non avendo informazioni sull'esistenza del limite nel termine di mezzo, si ha

$$l - \varepsilon \leqslant \liminf_{x\to+\infty} \frac{1}{x} \int_c^x f(t)\, dt \leqslant \limsup_{x\to+\infty} \frac{1}{x} \int_c^x f(t)\, dt \leqslant l + \varepsilon\,,$$

ma per l'arbitrarietà di ε $\liminf = \limsup = l$, quindi il limite esiste e vale l.

Nell'ipotesi che sia $f(x) > 0$ per ogni $x \geqslant 0$ possiamo considerare anche le medie armonica e geometrica e verificare che anch'esse tendono a l. Infatti se $0 < l < +\infty$

$$\lim_{x\to+\infty} \frac{1}{f(x)} = \frac{1}{l} \Rightarrow \nu(f)^{-1} = \mu\left(\frac{1}{f}\right) \to \frac{1}{l} \Rightarrow \nu(f) \to l$$

e quindi anche $\gamma(f) \to l$.

Se $l = 0$ $\mu(f) \to 0$ e quindi $\nu(f), \mu(f) \to 0$ essendo queste due medie più piccole e positive, e se $l = +\infty$ $\nu(f)^{-1} = \mu(1/f) \to 0$, quindi $\nu(f) \to +\infty$ e siccome $\nu(f) \leqslant \gamma(f) \leqslant \mu(f)$ anche $\gamma(f), \mu(f) \to +\infty$.

10.8 Metodi d'integrazione

Riguardo le funzioni elementari, basta leggere al contrario la tabella delle derivate alla rovescia e diventa subito la tabella delle primitive, per cui ad esempio

$$\int x^n\,dx = \frac{x^{n+1}}{n+1} + c\,, \quad \int e^x\,dx = e^x + c\,, \quad \int \cos x\,dx = \operatorname{sen} x + c\,, \quad c \in \mathbf{R}\,,$$

e così via, dove si intende $x \in \mathbf{R}$ o in $x \in I$ a seconda della funzione integranda. Tuttavia, combinate tra loro con prodotti, rapporti e composizioni, definiscono nuove funzioni le cui primitive non sono così immediate da riconoscere, addirittura in molti casi non è proprio possibile esprimerle in forma chiusa, in termini cioè di espressioni esplicite, finite, che coinvolgano le sole funzioni elementari. Non si conoscono le primitive di e^{x^2}, di $\operatorname{sen} x/x$, di e^x/x e di tante altre. Illustriamo allora alcuni metodi che ci permettono di determinare le primitive quando è possibile in molti casi non banali.

Integrazione delle funzioni razionali - L'integrando si presenta nella forma di un rapporto tra due polinomi primi tra loro, $f(x) = P(x)/Q(x)$, di cui possiamo supporre che il grado di P sia strettamente inferiore al grado di Q, dato che a questo caso ci si può sempre ridurre facendo direttamente la divisione. La prima cosa da fare è fattorizzare Q in un prodotto di polinomi irriducibili del tipo $(x - a_i)^{m_i}$, dove a_i, $i = 1, \ldots, h$, sono le radici reali con molteplicità m_i, e del tipo $((x - b_j)^2 + c_j^2)^{\mu_j}$, $j = 1, \ldots, k$, dove $b_j \pm ic_j$ sono le radici complesse con molteplicità μ_j, per cui

$$Q(x) = \prod_{i=1}^{h} (x - a_i)^{m_i} \prod_{j=1}^{k} ((x - b_j)^2 + c_j^2)^{\mu_j}\,.$$

Ad ogni radice reale a_i corrispondono m_i coefficienti reali A_{ir}, $r = 1, \ldots, m_i$, e ad ogni coppia di radici complesse $b_j + ic_j$ corrispondono μ_j coppie di coefficienti reali B_{js}, C_{js}, $s = 1, \ldots, \mu_j$ tali che

$$\frac{P(x)}{Q(x)} = \sum_{i=1}^{h} \sum_{r=1}^{m_i} \frac{A_{ir}}{(x - a_i)^r} + \sum_{j=1}^{k} \sum_{s=1}^{\mu_j} \frac{B_{js}x + C_{js}}{((x - b_j)^2 + c_j^2)^s}$$

che possono essere determinati in virtù del principio d'identità dei polinomi. Vediamo come si integrano i vari termini.

- $r = 1$:

$$\int \frac{1}{x - a_i}\,dx = \begin{cases} \log(x - a_i) + c & \text{se } x > a_i \\ -\log(a_i - x) + c & \text{se } x < a_i\,. \end{cases}$$

- $r \geqslant 2$:

$$\int \frac{1}{(x - a_i)^r}\,dx = \frac{1}{(1 - r)(x - a_i)^{r-1}} + c\,.$$

• $s = 1$: manipolando opportunamente i coefficienti b_j, c_j, B_{j1} e C_{j1}, possiamo facilmente esprimere la frazione come combinazione lineare di funzioni del tipo

$$\frac{\alpha_j}{(\alpha_j x - \beta_j)^2 + 1} \qquad \text{e} \qquad \frac{2(x - b_j)}{(x - b_j)^2 + c_j}$$

i cui integrali sono

$$\operatorname{arctg}(\alpha_j x - \beta_j) + c \qquad \text{e} \qquad \log((x - b_j)^2 + c_j) + c\,.$$

• $s \geqslant 2$: con simili manipolazioni la frazione si esprime come combinazione lineare di funzioni del tipo

$$\frac{\alpha_j}{((\alpha_j x - \beta_j)^2 + 1)^s} \qquad \text{e} \qquad \frac{2(x - b_j)}{((x - b_j)^2 + c_j)^s}$$

delle quali quella di destra ha per integrale

$$\frac{1}{(1-s)((x - b_j)^2 + c_j)^{s-1}} + c\,,$$

mentre l'integrazione di quella di sinistra si riconduce al calcolo dell'integrale

$$I_n(x) = \int \frac{1}{(1 + x^2)^n}\, dx\,, \qquad n \geqslant 2\,, \tag{10.24}$$

che vediamo tra poco.

Esercizio 10.21 - *Calcolare l'integrale*

$$\int_{-\infty}^{+\infty} \frac{1}{1 + x^4}\, dx\,.$$

Per $m, n \in \mathbf{N}$ con $n > m \geqslant 0$ la funzione continua e limitata

$$f(x) = \frac{x^{2m}}{1 + x^{2n}}\,, \qquad x \in \mathbf{R}\,,$$

è integrabile su $\mathbf{R}$ (provarlo!), proviamo a calcolarne l'integrale. Gli zeri del denominatore sono i $2n$ numeri complessi unitari

$$e^{i(2k+1)\pi/2n} = a_k + ib_k\,, \qquad k = 0, \ldots, 2n-1\,.$$

Scomponiamo la f in fratti semplici

$$f(x) = \sum_{k=0}^{2n-1} A_k \left(\frac{x - a_k}{(x - a_k)^2 + b_k^2} + i \frac{b_k}{(x - a_k)^2 + b_k^2} \right)\,, \quad A_k = -\frac{(a_k + ib_k)^{2m+1}}{2n}\,,$$

e integriamo su $[-h, h]$

$$\int_{-h}^{h} f(x)\, dx = \sum_{k=0}^{2n-1} A_k \left[\log \frac{(h - a_k)^2 + b_k^2}{(h + a_k)^2 + b_k^2} + i \left(\operatorname{arctg} \frac{h - a_k}{b_k} + \operatorname{arctg} \frac{h + a_k}{b_k} \right) \right]\,.$$

Passando al limite per $h \to +\infty$ si ottiene

$$\int_{-\infty}^{+\infty} \frac{x^{2m}}{1 + x^{2n}}\, dx = \pi i \sum_{k=0}^{2n-1} A_k \operatorname{sign} b_k = \pi \left(n \operatorname{sen} \frac{(2m+1)\pi}{2n} \right)^{-1}\,.$$

Un altro metodo di integrazione delle funzioni razionali, conveniente in presenza di radici multiple al denominatore, si basa sull'uso della *formula di Hermite*

$$\frac{P(x)}{Q(x)} = \sum_{i=1}^{h} \frac{A_i}{x-a_i} + \sum_{j=1}^{k} \frac{B_j x + C_j}{(x-b_j)^2 + c_j^2} + \frac{d}{dx}\frac{S(x)}{T(x)},$$

dove

$$T(x) = \prod_{i=1}^{h}(x-a_i)^{m_i-1} \prod_{j=1}^{k}((x-b_j)^2 + c_j^2)^{\mu_j-1}$$

e $S(x)$ è un polinomio a coefficienti da determinarsi, di grado pari a quello di $T(x)$ diminuito di 1. Il vantaggio è evidente, una volta calcolati i coefficienti di S, i termini nelle sommatorie si integrano senza difficoltà, l'altro compare già col segno di derivata e S/T ne è una primitiva.

Integrazione per parti - Questo metodo è particolarmente indicato quando ci troviamo di fronte ad un prodotto di funzioni di natura diversa e si basa sulla formula di derivazione del prodotto. Se $f, g : I \to \mathbf{R}$ sono la prima continua e la seconda derivabile, ovviamente sono anche localmente integrabili e si ha

$$\int f(x)g(x)\,dx = F(x)g(x) - \int F(x)g'(x)\,dx$$

dove F è una primitiva di f. Ai fini pratici conviene scegliere la funzione da integrare e quella da derivare in modo tale che l'integrale che resta da calcolare a secondo membro sia più semplice di quello dato. La dimostrazione è immediata

$$(10.25)\quad \frac{d}{dx}(F(x)g(x)) = F'(x)g(x) + F(x)g'(x) = f(x)g(x) + F(x)g'(x) \qquad \forall x \in I.$$

Ad esempio

$$\int x\cos x\,dx = x \operatorname{sen} x - \int \operatorname{sen} x\,dx = x\operatorname{sen} x + \cos x + c,$$

per cui si ha

$$\int_0^\pi x\cos x\,dx = [x\operatorname{sen} x]_0^\pi - \int_0^\pi \operatorname{sen} x\,dx = [\cos x]_0^\pi = -2.$$

Oppure

$$\int e^{-x}\operatorname{sen} x = -e^{-x}\operatorname{sen} x + \int e^{-x}\cos x = -e^{-x}\operatorname{sen} x - e^{-x}\cos x - \int e^{-x}\operatorname{sen} x$$

e portando a I membro l'ultimo termine

$$\int e^{-x}\operatorname{sen} x\,dx = -\frac{1}{2}e^{-x}(\operatorname{sen} x + \cos x) + c.$$

E ancora, pensando $f(x)$ come $1 \cdot f(x)$

$$\int \log x = x\log x - \int x\cdot\frac{1}{x} = x\log x - x,$$

$$\int \operatorname{arctg} x = x\operatorname{arctg} x - \int x\cdot\frac{1}{1+x^2} = x\operatorname{arctg} x - \frac{1}{2}\log(1+x^2).$$

La (10.25) è applicabile al calcolo dell'integrale di Riemann

$$\int_a^b f(x)g(x)\,dx = [F(x)g(x)]_a^b - \int_a^b F(x)g'(x)\,dx\,, \tag{10.26}$$

ne segue ad esempio che

$$\int_0^{+\infty} e^{-x}\operatorname{sen} x\,dx = -\frac{1}{2}\lim_{a\to+\infty}[e^{-x}(\operatorname{sen} x+\cos x)]_0^a = \frac{1}{2}\,.$$

Per un integrale improprio bisogna fare attenzione col passaggio al limite. Non ha senso scrivere

$$\int_a^b f(x)g(x)\,dx = \lim_{\beta\to b}[F(x)g(x)]_a^\beta - \lim_{\beta\to b}\int_a^\beta F(x)g'(x)\,dx$$

se i due limiti a destra sono infiniti con lo stesso segno, la (10.26) significa

$$\int_a^b f(x)g(x)\,dx = \lim_{\beta\to b}\left\{[F(x)g(x)]_a^\beta - \int_a^\beta F(x)g'(x)\,dx\right\}.$$

Esercizio 10.22 - *Calcolare l'integrale*

$$\int_0^1 \operatorname{arctg} x\,dx\,.$$

Esercizio 10.23 - *Verificare che*

$$\int_0^1 \left(\frac{\log(1+x)}{x}+\frac{\log x}{1+x}\right)dx = 0\,,$$

ma non si tenti di cercare le primitive delle due funzioni, non si conoscono, come non si conoscono le primitive di $x/\log x$, e^x/x, $(\operatorname{arctg} x)/x$ *e di tante altre, alcune già viste.*

Esercizio 10.24 - *Verificare che*

$$\int_0^{+\infty}\frac{\operatorname{sen}^2 x}{x^2}\,dx = \int_0^{+\infty}\frac{\operatorname{sen} x}{x}\,dx\,.$$

Esercizio 10.25 - *Dimostrare che*

$$\sum_{n=1}^{\infty}\frac{x^n}{n^2} = -\int_0^x \frac{\log(1-t)}{t}\,dt \qquad \forall x\in[-1,1]\,,$$

in particolare

$$-\int_0^1 \frac{\log(1-t)}{t}\,dt = \frac{\pi^2}{6} \quad e \quad \int_{-1}^0 \frac{\log(1-t)}{t}\,dt = -\frac{\pi^2}{12}\,.$$

Esercizio 10.26 - *Verificare che per ogni* $p\in\mathbf{R}$ *e per ogni* $\varepsilon>0$ *si ha*

$$\int_\varepsilon^1 (-\log x)^p\,dx = -\varepsilon(-\log\varepsilon)^p + p\int_\varepsilon^1 (-\log\varepsilon)^{p-1}\,dx\,,$$

e che quindi per ogni $n\in\mathbf{N}$ *la funzione* $(-\log x)^n$ *è integrabile in senso improprio su* $[0,1]$ *e*

$$\int_0^1 (-\log x)^n\,dx = n!\,.$$

Dedurne, per confronto, che anche la funzione $(-\log x)^p$ *è integrabile in senso improprio su* $]0,1]$ *per ogni* $p\in]0,+\infty[$ *(per* $p\leqslant 0$ *lo è nel senso di Riemann in quanto continua e limitata).*

Esercizio 10.27 - *Ricavare formule ricorsive per*

$$\int x^n e^x \, dx \qquad e \qquad \int \operatorname{sen}^n x \, dx \,.$$

Caccia all'errore: Dimostriamo che $\operatorname{senh} x = \cosh x$

$$\int e^x \operatorname{senh} x \, dx = e^x \cosh x - \int e^x \cosh x \, dx = e^x \cosh x - e^x \operatorname{senh} x + \int e^x \operatorname{senh} x \, dx$$

e semplificando si ottiene $e^x(\cosh x - \operatorname{senh} x) = 0$, è possibile?

In vari esempi del § 9.8 abbiamo usato il Teorema 9.33 di derivazione di una serie di potenze per passare da una funzione analitica ad una sua primitiva. Che ciò sia consentito è evidente, la serie $\sum c_n z^{n+1}/(n+1)$ ha lo stesso raggio di convergenza della serie iniziale $\sum c_n z^n$ e la somma della prima ha per derivata la somma della seconda sullo stesso cerchio di convergenza, bordo a parte. In altre parole anche l'integrazione passa al limite sotto il segno di somma, una proprietà che può essere utilmente applicata per rappresentare, e quindi approssimare con arbitraria precisione, il valore di un integrale definito come somma di una serie numerica. Ad esempio non si conoscono le primitive della funzione continua $f(x) = x^x$, $x > 0$, ma in $[0, 1]$ è limitata e quindi integrabile e, tenuto conto che la serie esponenziale converge su tutto $\mathbf{R}$, si ha

$$\int_0^1 x^x \, dx = \int_0^1 e^{x \log x} \, dx = \int_0^1 \sum_{n=0}^{\infty} \frac{x^n \log^n x}{n!} \, dx = \sum_{n=0}^{\infty} \frac{1}{n!} \int_0^1 x^n \log^n x \, dx \,.$$

D'altra parte integrando ripetutamente per parti si vede facilmente che

$$\int_0^1 x^n \log^n x \, dx = \frac{(-1)^n n!}{(n+1)^{n+1}} \,,$$

da cui

$$\int_0^1 x^x \, dx = \sum_{n=0}^{\infty} \frac{(-1)^n}{(n+1)^{n+1}} \,.$$

La somma dei primi 10 termini vale $0,783430510708738$ dove le prime 10 cifre (dopo la virgola) sono definitive.

Esercizio 10.28 - *Stabilire se è convergente l'integrale*

$$\int_0^1 \frac{1}{1 - x^x} \, dx \,.$$

Occupiamoci adesso della (10.24). Integrando per parti si ottiene

$$\begin{aligned} I_n(x) &= \frac{x}{(1+x^2)^n} + 2n \int \frac{x^2 \, dx}{(1+x^2)^{n+1}} \\ &= \frac{x}{(1+x^2)^n} + 2n \int \left(\frac{1}{(1+x^2)^n} - \frac{1}{(1+x^2)^{n+1}} \right) dx \,, \end{aligned}$$

da cui segue la formula ricorsiva

$$I_{n+1}(x) = \frac{x}{2n(1+x^2)^n} + \left(1 - \frac{1}{2n}\right) I_n(x) \,.$$

Ad esempio

$$I_3(x) = \int \frac{dx}{(1+x^2)^3} = \frac{x}{6(1+x^2)^3} + \frac{5}{6} I_2(x) = \frac{x}{6(1+x^2)^3} + \frac{5}{6} \operatorname{arctg} x + c \,,$$

$$\begin{aligned}I_4(x) = \int \frac{dx}{(1+x^2)^4} &= \frac{x}{8(1+x^2)^4} + \frac{7}{8} I_3(x) \\ &= \frac{x}{8(1+x^2)^4} + \frac{7}{8}\left(\frac{x}{6(1+x^2)^3} + \frac{5}{6}\operatorname{arctg} x\right) + c\end{aligned}$$

e così via.

Esercizio 10.29 - *Dimostrare per induzione che se $f \in C^{n+1}(I)$ e $x_0 \in I$ vale la formula di Taylor col resto integrale*

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x-x_0)^k + \frac{1}{n!}\int_{x_0}^{x} (x-t)^n f^{(n+1)}(t)\, dt \quad \forall x \in I\,.$$

Cambiamento di variabile - Talvolta, esprimendo la variabile d'integrazione come funzione opportuna di una nuova variabile, l'integrale dato si trasforma in un altro integrale, sperabilmente più semplice da calcolare o riconducibile a casi noti come le funzioni razionali. Indovinare la sostituzione giusta è un'arte che può essere acquisita con una buona palestra di esercizi.

Siano $f \in C^0(I)$ e $F \in C^1(I)$ una sua primitiva, legate dalla solita relazione

$$\int f(x)\, dx = F(x) + c\,.$$

Se adesso esprimiamo la variabile $x \in I$ in funzione della nuova variabile $t \in J$, ponendo $x = \varphi(t)$ con $\varphi : J \to I$ di classe $C^1(J)$, si ha

$$F(\varphi(t)) + c = \int \frac{d}{dt} F(\varphi(t))\, dt = \int f(\varphi(t))\varphi'(t)\, dt$$

e dal confronto con la predente si ottiene la formula del cambiamento di variabile

$$\int f(x)\, dx_{|x=\varphi(t)} = \int f(\varphi(t))\varphi'(t)\, dt\,.$$

Senza la precisazione a I membro sulla dipendenza di x da t tramite la φ tale formula non sarebbe corretta perché un'uguaglianza tra una funzione di x e una funzione di t non può essere identicamente soddisfatta.

Per un integrale definito, la formula del cambio di variabile con una $\varphi : [\alpha, \beta] \to [a, b]$ tale che $\varphi(\alpha) = a$ e $\varphi(\beta) = b$ si può dimostrare direttamente in questo modo

$$\begin{aligned}\int_a^b f(x)\, dx &= F(b) - F(a) = F(\varphi(\beta)) - F(\varphi(\alpha)) \\ &= \int_\alpha^\beta \frac{d}{dt} F(\varphi(t))\, dt = \int_\alpha^\beta f(\varphi(t))\varphi'(t)\, dt\,.\end{aligned}$$

Se $\varphi(\alpha) = b$ e $\varphi(\beta) = a$ si ottiene

$$\int_a^b f(x)\, dx = \int_\beta^\alpha f(\varphi(t))\varphi'(t)\, dt = -\int_\alpha^\beta f(\varphi(t))\varphi'(t)\, dt\,.$$

Ne segue che se φ è monotona (φ' di segno costante) vale in ogni caso la seguente formula

$$\int_a^b f(x)\, dx = \int_\alpha^\beta f(\varphi(t))|\varphi'(t)|\, dt\,.$$

Esempi:

10.11 *Calcoliamo l'area della regione E delimitata dell'ellisse di semiassi a e b*

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1\,.$$

Per simmetria basta moltiplicare per 4 l'area della sola parte contenuta nel I quadrante, cioè del sottografico della funzione

$$f(x) = \frac{b}{a}\sqrt{a^2 - x^2}\,, \qquad 0 \leqslant x \leqslant a\,.$$

All'integrale

$$\int \sqrt{a^2 - x^2}\, dx$$

si addice la sostituzione $x = a \operatorname{sen} t$

$$\int \sqrt{a^2 - x^2}\, dx_{|x = a \operatorname{sen} t} = a \int |\cos t| \frac{d}{dt} a \operatorname{sen} t\, dt = a^2 \int |\cos t| \cos t\, dt\,,$$

da cui sappiamo come togliere il valore assoluto a seconda dell'intervallo che si vuole considerare. Ad esempio

$$\int_0^a \sqrt{a^2 - x^2}\, dx = a^2 \int_0^{\pi/2} \cos^2 t\, dt = a^2 \int_0^{\pi/2} \frac{1 + \cos 2t}{2}\, dt = \frac{a^2 \pi}{4}\,.$$

Pertanto $\mathscr{A}(E) = \pi ab$.

10.12 *Si perviene all'integrale*

$$\frac{b}{a} \int_0^a \sqrt{a^2 + x^2}\, dx$$

se si vuole calcolare l'area della regione, detta settore iperbolico, compresa tra un ramo dell'iperbole

$$\frac{y^2}{b^2} - \frac{x^2}{a^2} = 1\,,$$

l'intervallo $[0, a]$ dell'asse x e le rette $x = 0$ e $x = a$.

Conviene usare in questo caso la sostituzione $x = a \operatorname{senh} t$ ($x = a \cosh t$ nel caso dell'integrale di $\sqrt{x^2 - a^2}$), il calcolo si lascia per esercizio.

Sostituzioni razionalizzanti - Con opportuni cambi di variabile standard è possibile ricondurre l'integrale assegnato a quello di una funzione razionale.

- $f(e^x)$ *con* $f(y)$ *funzione razionale.* Si pone $e^x = t$ e l'integrale diventa

$$\int f(e^x)\, dx_{|x = \log t} = \int \frac{f(t)}{t}\, dt\,.$$

Al posto dell'esponenziale possono trovarsi le funzioni iperboliche, ma si tratta dello stesso caso. Rientrano in questo espressioni razionali dipendenti da $\sqrt{a^2 + x^2}$ o da $\sqrt{x^2 - a^2}$, una volta trattate con le stesse sostituzioni dell'Esempio 10.12.

10.13 *Calcoliamo l'integrale*

$$I = \int_1^{+\infty} \frac{e^x - 1}{e^{2x} + e^x}\, dx$$

ponendo $y = e^x$.

Si ottiene

$$I = \int_e^{+\infty} \frac{y-1}{y^2+1} \cdot \frac{1}{y}\, dy = \int_e^{+\infty} \left(\frac{1}{y^2+1} + \frac{y}{y^2+1} - \frac{1}{y} \right) dy$$
$$= \lim_{a\to+\infty} [\operatorname{arctg} y + \log\sqrt{y^2+1} - \log y]_e^a = \frac{\pi}{2} - \log\sqrt{e^2+1}\,.$$

•$f(\operatorname{sen} x, \cos x)$ *con* $f(y_1, y_2)$ *funzione razionale.* Si pone $\operatorname{tang} x/2 = t$ e l'integrale diventa

$$\int f(\operatorname{sen} x, \cos x)\, dx_{|x=2\operatorname{arctg} t} = \int f\left(\frac{2t}{1+t^2}, \frac{1-t^2}{1+t^2} \right) \frac{2}{1+t^2}\, dt\,.$$

Rientrano in questo caso espressioni razionali dipendenti da $\sqrt{a^2-x^2}$, una volta trattate con le stesse sostituzioni dell'Esempio 10.11.

•$f(x, y(x)^{m_1/n_1}, \ldots, y(x)^{m_k/n_k})$ *con* $f(x, y_1, \ldots, y_k)$ *funzione razionale,*

$$y(x) = \frac{ax+b}{cx+d}\,, \qquad ad - bc \neq 0\,, \qquad m_i, n_i \in \mathbf{Z}\,.$$

Se $n = mcm\{n_1, \ldots, n_k\}$ con la sostituzione

$$\frac{ax+b}{cx+d} = t^n\,, \quad \text{da cui } x(t) = \ldots\,,$$

l'integrale diventa

$$\int f(x, y(x)^{m_1/n_1}, \ldots, y(x)^{m_k/n_k})\, dx = \int f(x(t), t^{nm_1/n_1}, \ldots, t^{nm_k/n_k}) x'(t)\, dt\,.$$

Rientra in questo il caso di una funzione razionale $f(x,y)$ con $y(x) = \sqrt{ax^2+bx+c}$ dove $\Delta = b^2 - 4ac > 0$. Se $a > 0$, dette $x_1 < x_2$ le radici del polinomio, si ha

$$\sqrt{ax^2+bx+c} = \sqrt{a(x-x_1)(x-x_2)} = (x-x_2)\sqrt{\frac{a(x-x_1)}{x-x_2}}$$

se stiamo integrando sull'intervallo $[x_2, +\infty[$ e analogamente su $]-\infty, x_1]$. Se $a < 0$ siamo necessariamente sull'intervallo $[x_1, x_2]$, si agisce nello stesso modo, ma tenendo presente che $\sqrt{ax^2+bx+c} = \sqrt{-a(x-x_1)(x_2-x)}$. In ogni caso, anche per $\Delta < 0$, possiamo sempre operare con la sostituzione

$$\sqrt{ax^2+bx+c} = \sqrt{|a|}x \pm t\,.$$

•$f(x) = x^p(ax^q+b)^r\,,\ p, q, r \in \mathbf{Q}$. Gli integrali di funzioni di questo tipo sono detti *integrali binomi* e sono riconducibili a integrali di funzioni razionali se e solo se è intero almeno uno dei numeri r, $(p+1)/q$ o $r + (p+1)/q$. Conviene usare le seguenti sostituzioni:

(a) $x = t^n$ se $r \in \mathbf{Z}$, essendo n il minimo comune multiplo tra i denominatori delle frazioni di p e q;

(b) $ax^q + b = t^n$ se $(p+1)/q \in \mathbf{Z}$ e $r = m/n$ con $n > 0$;

(c) $a + bx^{-p} = t^n$ se $r + (p+1)/q \in \mathbf{Z}$ e e $r = m/n$ con $n > 0$.

Bibliografia

[1] E. Acerbi, G. Buttazzo, *Primo corso di Analisi Matematica*, Pitagora, Bologna 1997.

[2] E. Acerbi, G. Buttazzo, *Analisi Matematica ABC, 1. Funzioni di una variabile*, Pitagora, Bologna 2003.

[3] F. Conti, *Calcolo, Teoria e applicazioni*, McGraw-Hill, 1993.

[4] F. Conti, P. Acquistapace, A. Savojni, *Analisi Matematica, Teoria e Applicazioni*, McGRaw-Hill, 2001.

[5] C.D. Pagani, S. Salsa, *Analisi Matematica, vol. 1*, Masson, 1994.

[6] G.H. Hardy, *A Course of Pure Mathematics*, Cambridge University Press, 1908.

[7] E. Giusti, *Analisi Matematica 1*, Bollati Boringhieri, Torino, 1991.

[8] P. Marcellini, C. Sbordone, *Esercitazioni di Matematica, vol. 1*, Liguori Editore, Napoli, 1995.

[9] E. Giusti, *Esercizi e Complementi di Analisi Matematica, vol.1*, Bollati Boringhieri, Torino, 1991.

[10] E. Acerbi, L. Modica, S. Spagnolo, *Problemi scelti di Analisi Matematica I*, Liguori Editore, Napoli, 1985.

[11] J. Havil, *Gamma: exploring Euler's constant*, Princeton University Press, Princeton, New Jersey, 2003.