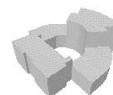




Sapienza, Università di Roma
Dipartimento di Matematica "G.Castelnuovo"



Note di base di

Analisi Matematica

Parte prima

versione 1.2 (4 ottobre 2011)

Lamberto LAMBERTI
Corrado MASCIA



Licenza © 2008 Lamberto Lamberti & Corrado Mascia

Distribuzione Creative Commons

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

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

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

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

Indice

Capitolo 1. I numeri reali	1
1. Numeri naturali, interi, razionali e irrazionali	1
2. Ordine e disequaglianze	14
3. La struttura metrica: il modulo	16
4. La verità sui reali	19
5. Estremo superiore ed estremo inferiore	23
Capitolo 2. Funzioni: anno zero	29
1. Ingredienti di base	29
2. Operazioni elementari su grafici	34
3. Funzioni invertibili e funzioni monotòne	38
4. Classi di funzioni più o meno comuni	42
5. Problemi di massimo e minimo	47
Capitolo 3. Incontri ravvicinati con i limiti: le successioni	49
1. Limite di successioni	51
2. Il limite entra in società	58
3. Calcolo di alcuni limiti	64
4. Successioni monotòne	66
5. Serie numeriche	67

CAPITOLO 1

I numeri reali

Il primo obiettivo che ci proponiamo in queste Note è di introdurre i numeri reali, indispensabili per tutto quello che vedremo nel seguito. Lo scopo non è di proporre una costruzione rigorosa dei reali (che va ricercata in qualche altro testo), quanto di presentare il concetto di numero reale in maniera “intuitivamente convincente”, in modo da potere dedurne le proprietà fondamentali. Per completezza, comunque, verranno dati cenni relativi ad alcune versioni degli assiomi necessari per la definizione precisa dell’insieme dei numeri reali.

1. Numeri naturali, interi, razionali e irrazionali

Il percorso classico che porta alla definizione dell’insieme \mathbb{R} , pietra miliare di qualsiasi corso di Analisi Matematica, comincia dai numeri *naturali*, passa per gli *interi*, si sofferma sui *razionali*, ed arriva, infine, ai *reali*. Più in là presenteremo un ulteriore passo in avanti che conduce ai *numeri complessi*.

Numeri naturali. I numeri naturali nascono da uno dei problemi primordiali dell’uomo (specie se insonne): il “conteggio”. Essi nascono dall’esigenza di rispondere alla domanda: “quanti?": $0, 1, 2, 3, \dots$. L’insieme di questi numeri, detti **numeri naturali**, si indica con il simbolo \mathbb{N} :

$$\text{numeri naturali:} \quad \mathbb{N} = \{0, 1, 2, 3, \dots\}.$$

Nell’insieme \mathbb{N} sono definite le operazioni di addizione e moltiplicazione, per cui valgono le ben note *leggi commutative, associative, distributiva*.

I numeri naturali hanno due difetti fondamentali, che costringono, in qualche modo, a guardare più in là, cercando nuovi insiemi di numeri.

1. *Difetto “algebrico”.* Le operazioni inverse, sottrazione e divisione, non sono sempre possibili nell’insieme dei numeri naturali: non è possibile sottrarre 2 da 1 o dividere 1 con 2 e ottenere un numero naturale. La soluzione “alla Salomone” (cioè dividere in due metà) nei naturali non è sempre praticabile: nel caso di un numero dispari, il concetto di “metà” non è rappresentato da nessun numero naturale.

Il desiderio di risolvere il difetto “algebrico” ci porterà, tra poche righe, ad introdurre prima i numeri interi relativi e poi quelli razionali.

2. *Difetto “metrico”.* Il secondo problema è collegato ad una questione fondamentale: la *misura*. Il concetto di misura si basa sulla procedura seguente:

- si sceglie un’unità di misura (ad esempio, il metro campione);
- si contano il numero di copie dell’oggetto unitario che ricoprono l’oggetto da misurare.

Nella maggior parte dei casi, la lunghezza del segmento da misurare non è pari ad multiplo intero del segmento unitario. Come fare? Semplice: si divide in sotto-segmenti il segmento unitario e si contano quante “parti” della sotto-unità servono per misurare il nuovo oggetto. Questo è il concetto di frazione, cioè di numero razionale. Come vedremo, però, bisognerà prendere atto della realtà delle cose: l’essere razionali non basta...

Numeri interi relativi. Per rendere possibili senza restrizioni le operazioni di sottrazione, si estende il concetto degli interi “negativi”, ottenendo l’insieme dei **numeri interi relativi** (o semplicemente **interi**):

$$\text{numeri interi:} \quad \mathbb{Z} = \{0, \pm 1, \pm 2, \pm 3, \dots\}$$

L’insieme \mathbb{Z} è una estensione insiemistica di \mathbb{N} , nel senso che contiene l’insieme dei numeri naturali, ed inoltre è possibile definire l’operazione di somma in \mathbb{Z} in modo che valgano le stesse proprietà che valevano in \mathbb{N} . C’è anche qualcosa in più: grazie all’introduzione dei numeri negativi, oltre a sommare, è sempre possibile sottrarre. In maniera più precisa, possiamo affermare che ogni elemento di \mathbb{Z} ha un *elemento inverso* rispetto all’operazione di somma:

$$\forall k \in \mathbb{Z} \quad \exists h \in \mathbb{Z} \quad \text{tale che} \quad k + h = 0.$$

Il numero h è quello che si ottiene cambiando di segno k . La presenza in \mathbb{Z} di un’operazione di somma per cui vale la proprietà commutativa, l’esistenza di 0 e l’esistenza dell’elemento inverso, fanno di \mathbb{Z} un **gruppo commutativo** (o **gruppo abeliano**).

Numeri razionali. Non abbiamo ancora risolto completamente il difetto “algebrico” di \mathbb{N} , dato che anche nell’insieme \mathbb{Z} non è sempre possibile dividere. Per questo motivo, introduciamo un terzo insieme di numeri: l’insieme dei **numeri razionali**, cioè numeri che sono rapporto di numeri interi

$$\text{numeri razionali:} \quad \mathbb{Q} = \left\{ \frac{p}{q} : p \in \mathbb{Z}, q \in \mathbb{N}, q \neq 0 \right\}.$$

Attenzione! I numeri razionali possono essere scritti in molti modi diversi: ad esempio,

$$\frac{2}{1} = \frac{4}{2} = \frac{142}{71} = \dots$$

In genere si preferisce avere un'unica rappresentazione del numero e, per questo, si richiede che q sia positivo e che p e q abbiano massimo comune divisore pari a 1. Ci si riconduce sempre ad un'espressione di questo genere attraverso la "semplificazione" dei fattori comuni.

Per costruzione valgono le inclusioni insiemistiche

$$\mathbb{N} \subset \mathbb{Z} \subset \mathbb{Q}.$$

Inoltre, nell'insieme dei numeri razionali \mathbb{Q} sono ben definite tutte le *operazioni razionali*, cioè addizione, moltiplicazione, sottrazione e divisione, eccetto la divisione per zero. La maniera precisa di formulare il fatto che sia possibile "dividere per un qualsiasi numero razionale non nullo" consiste nell'affermare che ogni elemento di \mathbb{Q} , tranne 0, ha un *elemento inverso* rispetto all'operazione di prodotto:

$$\forall x \in \mathbb{Q}, x \neq 0 \quad \exists y \in \mathbb{Q} \quad \text{tale che} \quad xy = 1.$$

L'elemento y si indica con $\frac{1}{x}$ oppure con x^{-1} .

La presenza di somma e prodotto (e relative proprietà) fa di \mathbb{Q} un **campo**.

ESERCIZIO 1.1. *Dimostrare che in \mathbb{Q} vale la legge di annullamento del prodotto: se $ab = 0$, allora almeno uno degli elementi a, b è nullo.*

Soluzione. Infatti, se $a \neq 0$, allora moltiplicando $ab = 0$ per $\frac{1}{a}$ a destra e a sinistra dell'uguale, si ottiene $\frac{1}{a}ab = \frac{1}{a}0$, da cui segue $b = 0$.

Risolto il problema delle operazioni inverse, rimane il difetto "metrico". È ragionevole sospettare che la definizione di \mathbb{Q} possa risolvere in un colpo solo sia il problema della divisione che quello della misurazione di lunghezze... ma è così?

Rappresentazione grafica dei razionali. Disegniamo una retta R (o *asse numerico*) e procediamo secondo questa scaletta.

(i) Scegliamo un punto di R come rappresentante di 0, che chiameremo **origine**, e un altro punto arbitrario come rappresentante del numero 1. Definiamo la direzione da 0 a 1 come **direzione positiva**, in questo modo, la retta R diviene una retta *orientata* (d'ora in poi, penseremo di aver scelto 1 alla destra di 0).

(ii) Replicando copie del segmento unitario nella direzione positiva di R (cioè alla destra di 1), otteniamo una famiglia di punti che indichiamo con $2, 3, \dots$. Detto in altre parole, rappresentiamo tutti i numeri naturali come punti sulla retta R .

(iii) Ripetendo lo stesso procedimento nella direzione negativa (alla sinistra di 0), otteniamo, allo stesso modo, una rappresentazione per gli interi negativi. Con questo procedimento, insieme a quanto fatto in (ii), arriviamo ad una rappresentazione i numeri interi sulla retta R .

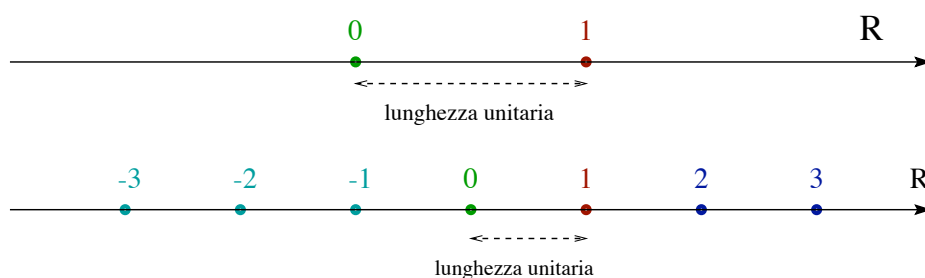


FIGURA 1. Dall'alto verso il basso: (a) i punti 0 e 1 sulla retta R determinano il segmento di lunghezza unitaria; (b) la rappresentazione dei numeri interi su R .

(iv) Rappresentiamo ora i numeri razionali $\frac{p}{q}$ su R per cui $p, q \in \mathbb{N}$ con p più piccolo di q . Dato che il numero $\frac{p}{q}$ è, moralmente, “ p volte $1/q$ ”, si divide l'intervallo unitario in q parti di uguale lunghezza e si prende come rappresentante di $\frac{p}{q}$ l'estremo destro del p -esimo intervallo.

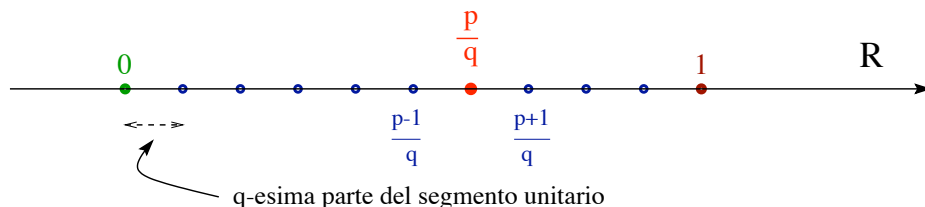


FIGURA 2. La rappresentazione su R di un numero razionale $\frac{p}{q}$ per cui $p, q \in \mathbb{N}$ e $p < q$.

(v) Nel caso di un qualsiasi numero razionale positivo p/q , si determinano p' e r positivi, con r più piccolo di q tali che

$$\frac{p}{q} = p' + \frac{r}{q},$$

e si ripete l'operazione spiegata nel passo precedente nel segmento di estremi p' e $p' + 1$. Analogamente per i numeri razionali negativi.

Il significato geometrico della somma di numeri razionali è facile: se $x, y \in \mathbb{Q}$, il punto in R che corrisponde a $x + y$ corrisponde al punto che si ottiene applicando una copia del segmento di estremi 0 e y sul punto x in modo da far coincidere la copia del punto 0 con x .

Ordine, modulo e distanza nei numeri razionali. Una volta rappresentati i numeri razionali su una retta R che è orientata, è possibile mettere ordine in \mathbb{Q} .

DEFINIZIONE 1.2. Ordinamento in \mathbb{Q} . Se $x, y \in \mathbb{Q}$, allora x è minore di y (o, equivalentemente, y è maggiore di x), se, nella rappresentazione su R , x si trova alla

sinistra di y . In tal caso si scrive

$$x < y.$$

Se x è minore di y o uguale ad y , si scrive $x \leq y$

$$x \leq y \iff x < y \text{ oppure } x = y.$$

Un numero $x \in \mathbb{Q}$ è **positivo**¹ se $x > 0$ ed è **negativo** se $x < 0$. Se x è positivo o è zero, si dice che è **non negativo** e si scrive $x \geq 0$ e, analogamente, se x è negativo o è zero, è **non positivo** e si scrive $x \leq 0$.

A questo punto, si introduce un oggetto di fondamentale importanza: il *modulo*.

DEFINIZIONE 1.3. Modulo e distanza in \mathbb{Q} . Dato $x \in \mathbb{Q}$, il **modulo** di x (detto anche **valore assoluto** o **norma**) si indica con $|x|$ ed è definito da

$$|x| := \begin{cases} x & x \geq 0, \\ -x & x < 0 \end{cases}$$

Dati due numeri razionali $x, y \in \mathbb{Q}$, si chiama **distanza** di x da y il numero $|x - y|$.

Geometricamente, il numero (non negativo) $|x|$ rappresenta la lunghezza del segmento in R che ha per estremi il punto x ed il punto 0; analogamente $|x - y|$ è la lunghezza del segmento che ha per estremi il punto x ed il punto y .

Non tutte le lunghezze sono razionali. Passiamo al collaudo di \mathbb{Q} (rappresentato sulla retta R), per la misurazione di lunghezze. Armiamoci di una fettuccia

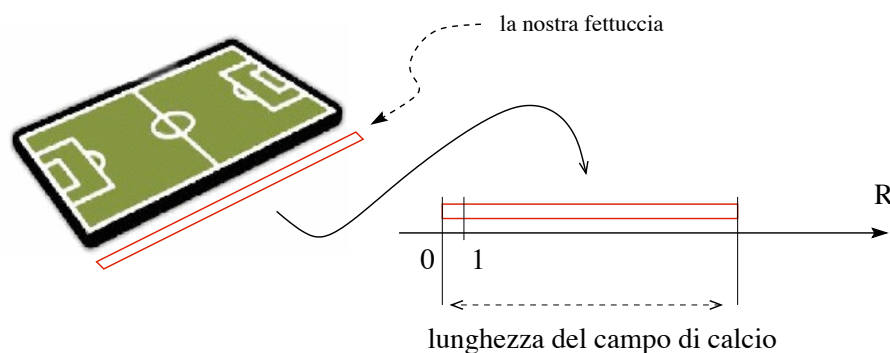


FIGURA 3. Il procedimento di misurazione: un campo di calcio.

di stoffa (o di fantasia) e procediamo nel modo più semplice possibile: se vogliamo

¹Qualcuno usa una terminologia leggermente diversa: con il termine “positivo” indica un numero alla destra di 0, eventualmente anche 0 stesso, mentre un numero positivo, ma non zero, viene detto “strettamente positivo”. Analogo discorso per i numeri negativi e per la relazione d’ordine. Se y è alla destra di x ed eventualmente è x dice che “ y è maggiore di x ”; se y è alla destra di x , ma non coincide con x , dice che “ y è strettamente maggiore di x ”. Basta mettersi d’accordo.

misurare la lunghezza di un certo oggetto (un tavolo, un campo di calcio, quel che sia...), fissiamo un estremo della fettuccia ad una estremità dell'oggetto da misurare ed estendiamola fino all'altro estremo, tagliamo la fettuccia in concomitanza con il secondo estremo e riportiamo la fettuccia lungo la retta R . Collochiamo il primo estremo in corrispondenza del punto 0, stendiamo la fettuccia in tutta la sua lunghezza nella direzione positiva e vediamo il secondo estremo dove va a finire. **Se** quest'ultimo finisce in corrispondenza di un numero razionale, quel numero è la lunghezza desiderata... Non è un errore di stampa il fatto che il “Se” sia scritto in neretto: questo procedimento non sempre funziona: alcune lunghezze non corrispondono a nessun numero razionale!

Già i matematici greci scoprirono che esistono segmenti la cui lunghezza non è un numero razionale, cioè esistono punti della retta R che non corrispondono a nessun numero razionale: in simboli, $R \setminus \mathbb{Q} \neq \emptyset$. L'esempio più elementare di lunghezza non razionale è la lunghezza ℓ della diagonale di un quadrato di lato unitario. Infatti, per

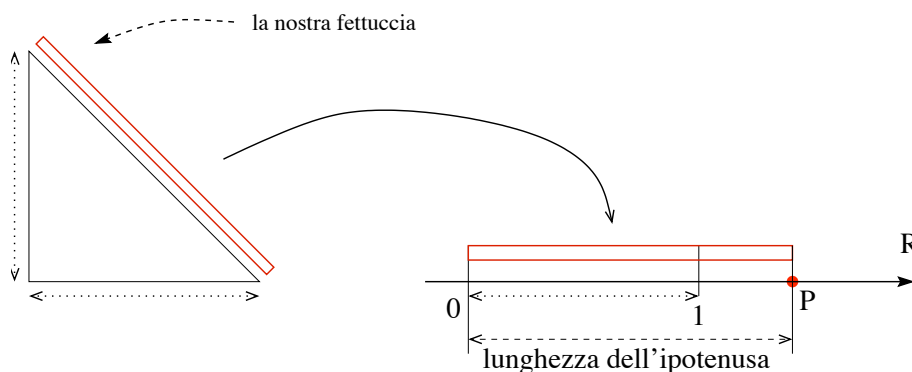


FIGURA 4. Il procedimento di misurazione dell'ipotenusa di un triangolo rettangolo di lato unitario: il punto P non corrisponde a nessun numero razionale.

il teorema di Pitagora, $\ell^2 = 2$, ma *nessun numero razionale elevato al quadrato dà per risultato il valore 2*.²

ESERCIZIO 1.4. *Dimostrare che \sqrt{p} non è razionale per ogni numero p primo. Lo stesso per $\sqrt[n]{p}$ con $n \in \{2, 3, 4, \dots\}$. Quali altre classi di numeri irrazionali sai immaginare a partire da questo esempio?*

²Per dimostrare questa affermazione, supponiamo, al contrario, che $\ell = p/q$ per opportuni p, q interi positivi. Senza restrizione, possiamo supporre che p e q non abbiano fattori comuni (altrimenti si possono semplificare). Allora si deve avere $p^2 = 2q^2$, quindi p^2 deve essere un numero pari. Dato che il quadrato di un numero dispari è dispari, anche p deve essere pari, cioè della forma $p = 2r$ con r intero positivo. Sostituendo, si ottiene $4r^2 = 2q^2$, e, semplificando il fattore 2, $2r^2 = q^2$. Ne segue che q^2 è pari e, di conseguenza, lo è anche q . Quindi p e q avrebbero un fattore comune in contraddizione con la nostra ipotesi. Pertanto ℓ non può essere razionale.

L'introduzione dei numeri razionali, che sembrava così promettente, non ha risolto il difetto “metrico” che avevamo già trovato in \mathbb{N} . Infatti è possibile costruire oggetti la cui lunghezza non è misurabile con un elemento di \mathbb{Q} . Prossima tappa: estendere \mathbb{Q} in modo da ottenere un insieme (con le stesse proprietà algebriche e metriche di \mathbb{Q}) in cui sia possibile misurare tutte le lunghezze possibili. Questa estensione è l'oggetto che chiamiamo *insieme dei numeri reali*.

Descrizione (naïf) dei numeri reali. Dato che i numeri razionali non sono sufficienti per le misurazioni, è necessario “inventare” nuovi numeri che permettano di misurare tutti i possibili segmenti. Prendiamo il toro per le corna e dichiariamo che: “ogni punto della retta R è un numero”, che chiameremo **numero reale**:

insieme dei numeri reali: $\mathbb{R} := \mathbb{Q} \cup \{\text{punti di } R \text{ che non sono in } \mathbb{Q}\}.$

Un elemento di \mathbb{R} che non sia in \mathbb{Q} si dice **numero irrazionale**. I numeri reali quindi, per definizione, coincidono con quelli della retta reale R . Esattamente come detto e fatto in precedenza, pensiamo la retta \mathbb{R} con orientamento da sinistra verso destra. La scelta del simbolo \mathbb{R} (che sostituisce R) sta a ricordare che stiamo pensando i punti della retta come oggetti per cui sono definite operazioni di somma e prodotto.

Questa definizione di numero reale grida vendetta: è intuitiva e andrebbe precisata rigorosamente. A questo livello, però, ci accontentiamo di questa versione naïf.³

Si tratta ora di capire quale rappresentazione possiamo dare ad un qualsiasi numero reale, cosa significano le operazioni di somma e prodotto in \mathbb{R} e i concetti di ordine e distanza?

Per la somma (e quindi la differenza) basta ricordare il significato della somma di razionali come punti sulla retta. Se x, y sono due numeri reali, per determinare dove si trovi sulla retta \mathbb{R} il punto $x + y$, basta procedere come segue. Rappresentiamo y come una freccia che parte da 0 e arriva nel punto corrispondente y . Per ottenere $x + y$ basta fare un “cut'n'paste” della freccia da 0 a y : se ne fa una copia e si trasla in modo da far coincidere il punto di partenza della freccia con x . Il nuovo punto di arrivo della freccia determina la posizione di $x + y$. Nel caso della differenza $x - y$, bisogna invertire la freccia che rappresenta y .

Le operazioni di prodotto e divisione in \mathbb{Q} possono essere estese, per approssimazione, ad \mathbb{R} , ma non ci soffermeremo qui sulla questione. Ci limitiamo a comunicare che valgono le stesse proprietà elencate per i numeri razionali, che, per completezza, riportiamo qui di seguito:

³L'idea intuitiva di numero reale come punto dell'asse numerico è stata alla base della matematica per lunghissimo tempo. Solo più tardi, nel XIX secolo, tale ipotesi è stata giustificata in modo rigoroso.

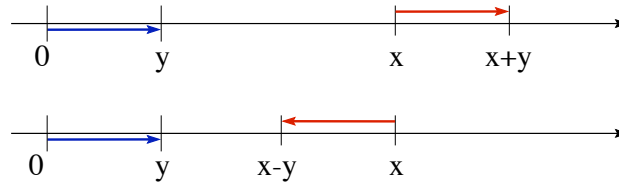


FIGURA 5. Somma e differenza di numeri reali.

leggi associative: $a + (b + c) = (a + b) + c$ e $a(bc) = (ab)c$ per ogni $a, b, c \in \mathbb{R}$;

leggi commutative: $a + b = b + a$ e $ab = ba$ per ogni $a, b \in \mathbb{R}$;

esistenza dell'elemento neutro per la somma: $a + 0 = a$ per ogni $a \in \mathbb{R}$;

esistenza dell'elemento neutro per il prodotto: $a \cdot 1 = a$ per ogni $a \in \mathbb{R}$;

esistenza dell'opposto per la somma: $a + (-a) = 0$ per ogni $a \in \mathbb{R}$;

esistenza dell'inverso per il prodotto: $a \cdot \frac{1}{a} = 1$ per ogni $a \in \mathbb{R}$, $a \neq 0$;

legge distributiva: $a(b + c) = ab + ac$;

annullamento del prodotto: se $ab = 0$, allora almeno uno tra a , b è 0.

Anche ordinamento e distanza si possono estendere da \mathbb{Q} ad \mathbb{R} .

DEFINIZIONE 1.5. Ordinamento in \mathbb{R} . Se $x, y \in \mathbb{R}$, allora x è minore di y (o y è maggiore di x), se x si trova alla sinistra di y . In tal caso si scrive $x < y$.

Per i simboli \leq e \geq , e per i termini positivo/negativo/non negativo/non positivo si utilizza lo stesso significato già visto per i numeri razionali.

DEFINIZIONE 1.6. Modulo e distanza in \mathbb{R} . Dato $x \in \mathbb{R}$, il modulo di x (detto anche valore assoluto o norma) si indica con $|x|$ ed è definito da

$$|x| := \begin{cases} x & x \geq 0, \\ -x & x < 0 \end{cases}$$

Dati due numeri reali $x, y \in \mathbb{R}$, si chiama distanza di x da y il numero $|x - y|$.

Tutte le proprietà che abbiamo descritto fanno di \mathbb{R} (così come lo era \mathbb{Q}) un **campo totalmente ordinato dotato di metrica**... più qualcosa... Mentre \mathbb{Q} è, in un certo senso, “bucato”, \mathbb{R} non lo è... Cosa vuol dire rigorosamente che “ \mathbb{R} non ha buchi”? Ci dedicheremo tra una manciata di pagine a spiegare in maniera più precisa come trasformare questa idea intuitiva in un oggetto matematicamente chiaro.

M'approssimo in un denso. Sebbene i numeri razionali non coprano tutta la retta dei numeri reali \mathbb{R} per via della presenza di numeri irrazionali, è vero che ci vanno molto vicini... Comunque si fissa una soglia di errore ammissibile, è possibile approssimare un numero reale con un numero razionale commettendo un errore più piccolo della soglia consentita. Infatti, supponiamo (per semplicità) che la soglia sia

della forma $1/q_0$ con $q_0 \in \mathbb{N}$ (ad esempio, per $q_0 = 1000$, l'errore ammesso è $1/1000$). Dividiamo la retta reale \mathbb{R} in segmenti di lunghezza $1/q_0$ e consideriamo i punti della forma p/q_0 con $p \in \mathbb{Z}$. Dato che \mathbb{R} è l'unione dei segmenti con estremi p/q_0 e $(p+1)/q_0$ per $p \in \mathbb{Z}$, per ogni punto x di \mathbb{R} esiste p_0 tale che

$$\frac{p_0}{q_0} \leq x \leq \frac{p_0 + 1}{q_0}.$$

Quindi è possibile approssimare il punto x con un razionale p_0/q_0 commettendo un errore minore di $1/q_0$. Il fatto che i punti razionali siano arbitrariamente vicini ad ogni punto x di \mathbb{R} si esprime in simboli come segue

$$\forall x \in \mathbb{R}, \quad \forall \varepsilon > 0 \quad \exists y \in \mathbb{Q} \quad \text{tale che} \quad |x - y| < \varepsilon,$$

e si esprime dicendo che *l'insieme dei razionali è denso nell'insieme dei numeri reali*.

Dal punto di vista della misurazione concreta, la densità è una proprietà notevole! In linea di principio, tenendo conto che c'è sempre un errore di misurazione, ogni misura può essere compiuta in \mathbb{Q} : un qualsiasi elemento di \mathbb{R} può essere individuato con un arbitrario grado di precisione usando punti razionali. Ad esempio, anche un qualsiasi computer ragiona sempre e solo con (un sottoinsieme dei) numeri razionali.

Rappresentazione decimale dei numeri reali. I numeri reali possono essere rappresentati tramite una “successione di decimali”. Vediamo in che maniera.

Fissato $x \in \mathbb{R}$, c'è (almeno) un intero $c_0 \in \mathbb{Z}$ per cui $c_0 \leq x \leq c_0 + 1$. Indichiamo con I_0 il segmento che ha come estremi i punti c_0 e $c_0 + 1$, cioè

$$I_0 = \{y \in \mathbb{R} : c_0 \leq y \leq c_0 + 1\}.$$

L'insieme I_0 si chiama *intervallo*. Quindi possiamo dire che

$$\forall x \in \mathbb{R} \quad \exists c_0 \in \mathbb{Z} \quad \text{tale che} \quad x \in I_0$$

Se dividiamo I_0 in dieci parti uguali tramite i punti $c_0 + \frac{1}{10}, c_0 + \frac{2}{10}, \dots, c_0 + \frac{9}{10}$, il numero x è in almeno uno dei sottointervalli di I_0 . In altre parole, c'è un intero c_1 , compreso tra $0, 1, 2, \dots, 9$ (nel seguito, diciamo che intero c compreso in $\{0, 1, \dots, 9\}$ è una *cifra*) tale che x appartiene all'intervallo I_1 dove

$$I_1 = \left\{ y \in \mathbb{R} : c_0 + \frac{1}{10} c_1 \leq y \leq c_0 + \frac{1}{10} c_1 + \frac{1}{10} \right\}.$$

Quindi vale la seguente affermazione

$$\forall x \in \mathbb{R} \quad \exists c_0 \in \mathbb{Z}, \quad c_1 \in \{0, 1, \dots, 9\} \quad \text{tali che} \quad x \in I_1.$$

Di nuovo, dividiamo I_1 in dieci parti uguali e troviamo una cifra c_2 tale che x sia nell'intervallo I_2 definito da:

$$I_2 = \left\{ y \in \mathbb{R} : c_0 + \frac{1}{10}c_1 + \frac{1}{100}c_2 \leq y \leq c_0 + \frac{1}{10}c_1 + \frac{1}{100}c_2 + \frac{1}{100} \right\}.$$

Ripetendo il procedimento, dopo n passi, troviamo $c_0 \in \mathbb{Z}$, $c_1, c_2, \dots, c_n \in \{0, 1, \dots, 9\}$. tali che $x \in I_n$ con I_n dato da

$$I_n = \left\{ y \in \mathbb{R} : c_0 + \frac{1}{10}c_1 + \dots + \frac{1}{10^n}c_n \leq y \leq c_0 + \frac{1}{10}c_1 + \dots + \frac{1}{10^n}c_n + \frac{1}{10^n} \right\}.$$

In modo equivalente possiamo dire

$$\forall x \in \mathbb{R} \quad \forall n \in \mathbb{N} \quad \exists c_0 \in \mathbb{Z}, c_1, \dots, c_n \in \{0, 1, \dots, 9\} \quad \text{tali che } x \in I_n.$$

L'errore che si commette approssimando il numero x con la sequenza c_0, c_1, \dots, c_n è minore o uguale a 10^{-n} . Dato che, per n che cresce, l'errore tende a zero, una volta assegnati tutti i numeri c_0, c_1, c_2, \dots , (che sono una successione infinita!) è possibile determinare il punto x . In altre parole, un qualsiasi numero reale può essere descritto completamente con una successione infinita di interi. Nella notazione decimale ordinaria, la corrispondenza tra x e la successione c_0, c_1, c_2, \dots è indicata scrivendo

$$x = c_0, c_1 c_2 c_3 \dots$$

dove anche l'intero c_0 viene espresso in notazione decimale.⁴ La sequenza $c_0, c_1, \dots, c_n, \dots$ corrisponde (in qualche senso) al “numero civico” dell’abitante x di “via dei Reali”: ad ogni decimale in più che considera, un alacre postino determina la collocazione di x con un errore più piccolo di un fattore dieci. Per determinare esattamente la collocazione del numero x , il poveretto dovrà considerare un numero infinito di cifre!

Nella rappresentazione decimale dei numeri reali, il ruolo speciale giocato dal numero dieci è puramente accidentale ed è legato ad un problema di dita (dieci nel caso degli esseri umani). L'unica ragione dell'ampio uso del sistema decimale è proprio dovuto alla facilità di contare fino a dieci con le nostre mani. Non per niente, in inglese “cifra” si dice “digit”, che manifesta in modo più immediato l'origine “manuale” del conteggio.... Avremmo potuto utilizzare un qualsiasi altro intero $p > 1$ e procedere in modo analogo, ottenendo una rappresentazione diversa, ma altrettanto valida. Ad

⁴E' possibile che ci siano due diverse rappresentazioni decimali dello stesso numero. Per esempio,

$$1 = 0,99999\dots = 1,00000\dots$$

Nella costruzione, l'intero c_0 è determinato in maniera univoca da x solo se x non è esso stesso un intero. Se $x \in \mathbb{Z}$, è possibile scegliere sia $c_0 = x$ che $c_0 = x - 1$. Una volta fatta questa scelta, la cifra c_1 è unica a meno che x non sia uno dei nuovi punti che suddividono I_0 in dieci parti uguali. Continuando così, troviamo che c_0 e tutti i c_k sono determinati unicamente da x , a meno che x non sia uno dei punti della suddivisione ad un certo livello.

esempio, scegliendo $p = 3$, quale rappresentazione decimale si ottiene per $\frac{1}{2}$? Per la risposta, provate ad immaginarvi postino che è in grado di percorrere passi di lunghezza $1/3^n$. Segnate sulla retta il numero $1/2$. Partite da 0: il numero di passi di lunghezza $1 = 1/3^0$ che potete percorrere senza superare il numero $1/2$ è c_0 . Partendo ora da c_0 , quanti passi di lunghezza $1/3$ potete percorrere? Procedete in questo modo con passi da $1/3^2$, $1/3^3$ e così via dicendo...

Intervalli limitati ed illimitati. Dati $a < b$, il segmento in \mathbb{R} di estremi a, b si chiama **intervallo**. Se gli estremi a, b sono inclusi nell'intervallo, l'intervallo si dice **chiuso**, se invece vengono esclusi si dice **aperto**:

$$\begin{aligned}\text{intervallo aperto:} \quad & (a, b) := \{x \in \mathbb{R} : a < x < b\}, \\ \text{intervallo chiuso:} \quad & [a, b] := \{x \in \mathbb{R} : a \leq x \leq b\}.\end{aligned}$$

In entrambi i casi il valore $b - a$ è la **lunghezza** dell'intervallo (o **misura** dell'intervallo).

Si possono considerare anche intervalli semiaperti (o semichiusi) includendo uno solo dei due estremi: $(a, b]$ oppure $[a, b)$. Anche le semirette sono usualmente considerate “intervalli” e si indicano con

$$(a, +\infty) := \{x \in \mathbb{R} : x > a\} \quad [a, +\infty) := \{x \in \mathbb{R} : x \geq a\},$$

e varianti. Anche l'insieme \mathbb{R} può essere pensato come intervallo e, in tal caso, viene indicato con $(-\infty, +\infty)$.

Se necessario, per distinguere il caso degli intervalli ad estremi in \mathbb{R} da quello delle semirette, si parla, nel primo caso, di *intervalli limitati*, nel secondo di *intervalli illimitati*.

OSSERVAZIONE 1.7. Gli intervalli (limitati e illimitati) sono tutti e soli i sottoinsiemi I di \mathbb{R} che godono della proprietà seguente:

$$(x_1, x_2) \subset I \quad \forall x_1, x_2 \in I, x_1 < x_2.$$

Questa proprietà si esprime dicendo che *gli intervalli sono insiemi connessi*.

Il piano ed altre realtà. Dati due insiemi A e B si indica con $A \times B$ l'insieme

$$\text{prodotto cartesiano di } A \text{ e } B: \quad A \times B = \{(x, y) : x \in A, y \in B\},$$

Ad esempio, a quale insieme appartiene l'oggetto “giorno dell'anno”? Due numeri lo individuano: il giorno del mese (che appartiene all'insieme $I := \{1, 2, 3, \dots, 31\}$) e il numero del mese (che appartiene all'insieme $J := \{1, 2, 3, \dots, 12\}$), quindi è un elemento del prodotto cartesiano $I \times J$. Ad esempio, la data 3 aprile corrisponde all'elemento $(3, 4)$ dell'insieme $I \times J$.

Con il simbolo $\mathbb{R} \times \mathbb{R}$ o con \mathbb{R}^2 si indica il prodotto cartesiano dell'insieme \mathbb{R} con sé stesso, ossia l'insieme costituito dalle coppie ordinate (x, y) dove $x, y \in \mathbb{R}$:

$$\mathbb{R}^2 = \mathbb{R} \times \mathbb{R} = \{(x, y) : x, y \in \mathbb{R}\}.$$

La sottolineatura della parola “ordinate” sta a segnalare che, in generale, l'elemento (x, y) è diverso da (y, x) . Ad esempio, $(1, 2) \neq (2, 1)$ (il primo febbraio è diverso dal 2 gennaio!). Per rappresentare l'insieme \mathbb{R}^2 si utilizza in genere un piano. Scegliete due rette orientate di riferimento, ortogonali fra loro e battezzatele, rispettivamente “asse x ” e “asse y ”. Per disegnare l'elemento (x_0, y_0) , si segna sull'asse x il punto H corrispondente al numero reale x_0 e sull'asse y quello K corrispondente al numero reale y_0 . Dopo di che si tracciano la retta per H e parallela all'asse y (quindi ortogonale all'asse x) e la retta per K e parallela all'asse x . Il punto intersezione rappresenta il punto (x, y) . I numeri x e y sono detti **coordinate** di (x, y) . Pensando $\mathbb{R} \times \mathbb{R}$ come

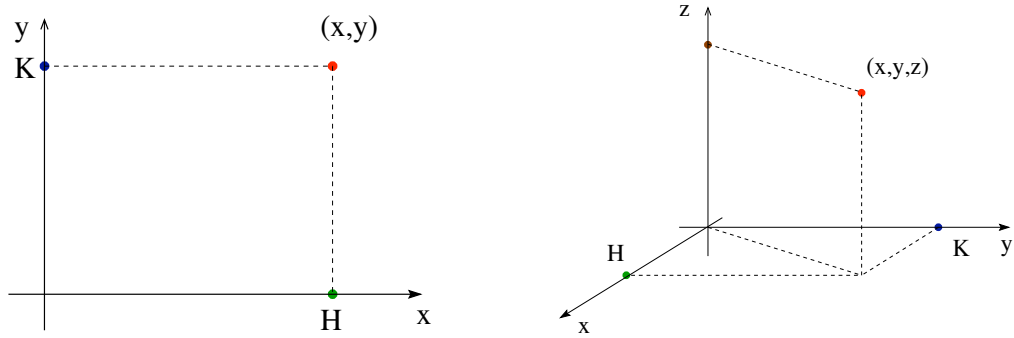


FIGURA 6. Il piano cartesiano \mathbb{R}^2 e lo spazio cartesiano \mathbb{R}^3 .

un piano, come si rappresentano gli insiemi $\{1\} \times \{2\}$, $[0, 1] \times \{2\}$ e $[0, 1] \times [0, 2]$? Rispettivamente, un punto, un segmento ed un rettangolo (Fig. 7).

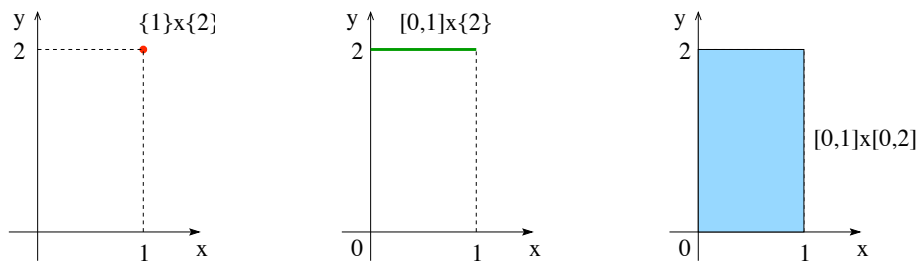


FIGURA 7. Gli insiemi $\{1\} \times \{2\}$, $[0, 1] \times \{2\}$ e $[0, 1] \times [0, 2]$.

Il piano reale \mathbb{R}^2 è dotato in maniera naturale dell'operazione di somma: si tratta della somma definita componente per componente:

$$(1) \quad (x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2).$$

Non è evidente, al contrario, come dare una nozione di prodotto tra due punti⁵ del piano \mathbb{R}^2 . E' possibile, invece, introdurre la nozione di **prodotto per uno scalare**: dato $\lambda \in \mathbb{R}$, si definisce

$$(2) \quad \lambda(x, y) := (\lambda x, \lambda y).$$

Con le definizioni (1) e (2), il piano \mathbb{R}^2 viene dotato della struttura di **spazio vettoriale** e i suoi elementi vengono detti **vettori**.

Analogamente $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$, o \mathbb{R}^3 , indica l'insieme di terne ordinate di numeri reali (x, y, z) . L'insieme \mathbb{R}^3 si rappresenta con lo spazio, tramite una scelta di tre assi coordinati ortogonali tra loro. In generale, \mathbb{R}^n (dove $n \in \mathbb{N}$) indica l'insieme delle n -ple del tipo (x_1, \dots, x_n) dove $x_1, \dots, x_n \in \mathbb{R}$. In sintesi

$$\mathbb{R}^n := \{x = (x_1, \dots, x_n) : x_1, \dots, x_n \in \mathbb{R}\}.$$

I casi $n = 2$ e $n = 3$ possono essere rappresentati geometricamente come un piano e tutto lo spazio, rispettivamente. In questa rappresentazione, i valori della coppia/terna corrispondono alle coordinate cartesiane di un punto. Nel caso di \mathbb{R}^n , i valori della n -pla possono essere interpretati come coordinate cartesiane n -dimensionali, ma poco si può visualizzare a meno di non possedere dei superpoteri.

Nell'insieme \mathbb{R}^n sono definite le operazioni di somma e di prodotto per uno scalare, dati $x = (x_1, \dots, x_n), \xi = (\xi_1, \dots, \xi_n) \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$

$$(3) \quad x + \xi = (x_1, \dots, x_n) + (\xi_1, \dots, \xi_n) := (x_1 + \xi_1, \dots, x_n + \xi_n).$$

$$(4) \quad \lambda(x_1, \dots, x_n) := (\lambda x_1, \dots, \lambda x_n).$$

In questo modo, \mathbb{R}^n acquisisce la struttura di **spazio vettoriale** e, anche in questo caso, i suoi elementi vengono detti **vettori**.

E' possibile definire anche in \mathbb{R}^n i concetti di modulo e distanza.

DEFINIZIONE 1.8. Modulo e distanza in \mathbb{R}^n . Dato $x = (x_1, \dots, x_n) \in \mathbb{R}^n$, il **modulo** di x si indica con $|x|_n$ ed è definito da

$$(5) \quad |x|_n := \sqrt{x_1^2 + \dots + x_n^2} = (x_1^2 + \dots + x_n^2)^{1/2}$$

Dati due vettori $x, y \in \mathbb{R}^n$, si chiama **distanza** di x da y il numero $|x - y|_n$.

E' possibile, in realtà, introdurre anche altre definizioni di modulo, ma, per ora, ci atterremo a quella data in (5).

⁵Definire il prodotto componente per componente, non è una buona idea (perché?). Una possibilità per introdurre il concetto di prodotto nel piano è alla base della definizione di numero complesso, che verrà ripresa più avanti.

2. Ordine e disequaglianze

Nella pratica spesso è difficile determinare con precisione una quantità x . Ben più facile è ottenere una *stima di x* , cioè mostrare che x è compreso tra una certa quantità a e un'altra quantità b . e, in molte situazioni, una buona stima di x è un'informazione sufficiente per la soluzione del problema. Le disequaglianze sono, perciò, un oggetto fondamentale nell'uso dei numeri reali.

Regole di base per le disequazioni

- (i) La somma di numeri positivi è positiva: $a > 0, b > 0 \Rightarrow a + b > 0$.
- (ii) Il prodotto ab è positivo se e solo se a e b sono di segno concorde ed è negativo se e solo se sono di segno discorde
- $$ab > 0 \iff a > 0, b > 0 \text{ oppure } a < 0, b < 0$$
- $$ab < 0 \iff a > 0, b < 0 \text{ oppure } a < 0, b > 0.$$
- (iii) Le disequaglianze possono essere sommate⁶: $a > b, c > d \Rightarrow a + c > b + d$.
- (iv) Se si moltiplica una disequaglianza per un numero positivo il segno $>$ rimane invariato:

$$a > b, c > 0 \Rightarrow ac > bc.$$

- (v) Vale la *proprietà transitiva*: $a < b, b < c \Rightarrow a < c$.

Le regole precedenti valgono anche sostituendo al simbolo $<$ il simbolo \leq .

ESERCIZIO 2.1. *Dimostrare la seguente disequazione*

$$(1 + h)^n \geq 1 + nh \quad \forall h > 0, n \in \mathbb{N}.$$

Prima soluzione (euristica). Partiamo da un paio di situazioni semplici:

$$n = 2 : \quad (1 + h)^2 = 1 + 2h + h^2 \geq 1 + 2h,$$

$$n = 3 : \quad (1 + h)^3 = 1 + 3h + 3h^2 + h^3 \geq 1 + 3h.$$

In generale, dato che $(1 + h)^n = (1 + h) \dots (1 + h)$, si ha

$$(1 + h)^n = 1 + h + \dots + h + \dots$$

In quanti modi possibili compare il termine h ? Uno per ciascuno dei termini $(1 + h)$, cioè n . In definitiva

$$(1 + h)^n = 1 + nh + \dots \text{altri termini} \dots$$

Gli "...altri termini..." sono positivi perché prodotto di positivi, quindi $(1 + h)^n \geq 1 + nh$.

⁶Sottrarre le disuguaglianze ottenendo $a - c > b - d$, invece, non è legittimo. Ad esempio, $2 > 1$ e $3 > 1$, ma $2 - 3 = -1 < 0 = 1 - 1$.

Seconda soluzione (per induzione). Risolviamo di nuovo lo stesso problema usando il metodo di induzione. L'affermazione da dimostrare consiste in una famiglia di disequazioni, una per ogni numero naturale n . Il metodo di induzione si divide in due passi: dimostrare l'affermazione per il primo naturale ($n = 1$ in questo caso); supporre l'affermazione valida nel caso n e dimostrarla per $n + 1$.

Il primo passo è semplicissimo: una verifica diretta mostra che la disuguaglianza è in realtà una uguaglianza quando si sceglie $n = 1$.

Supponiamo ora che la proprietà valga con n . Di conseguenza, si hanno

$$\begin{aligned}(1+h)^{n+1} &= (1+h)^n(1+h) \geq (1+nh)(1+h) \\ &= 1+h+nh+nh^2 = 1+(n+1)h+nh^2 \\ &\geq 1+(n+1)h,\end{aligned}$$

dove l'ultima disuguaglianza discende dal fatto che nh^2 è non negativo.

Oltre a sommare/moltiplicare disequazioni, è chiaro che si può ambire ad applicare altre operazioni. Vediamo il caso dell'elevazione a potenza e delle radici n -esime.

Per ogni $a, b \in \mathbb{R}$, vale $a^2 - b^2 = (a+b)(a-b)$. Se $a, b > 0$, allora $a+b > 0$ e quindi

$$a, b > 0, \quad a > b \quad \Rightarrow \quad a^2 > b^2.$$

Quindi l'operazione "elevare al quadrato" preserva l'ordine dei numeri positivi⁷: se i numeri sono ordinati, anche i loro quadrati lo sono. Dalla relazione

$$a - b = \frac{1}{a+b}(a^2 - b^2)$$

valida per ogni $a, b > 0$, si deduce che vale anche il viceversa: se $a, b > 0$ e $a^2 > b^2$, allora $a > b$. La stessa proprietà vale per qualsiasi potenza $n \in \mathbb{N}$.

PROPOSIZIONE 2.2. *Siano $a, b > 0$ e $n \in \mathbb{N}$. Allora*

$$a^n > b^n \quad \Longleftrightarrow \quad a > b.$$

DIMOSTRAZIONE. Dato che, nella scomposizione

$$a^n - b^n = (a-b)(a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1}),$$

il secondo fattore a destra è positivo, $a^n - b^n$ ha lo stesso segno di $a - b$. □

Dato $x \geq 0$, con il simbolo \sqrt{x} si indica la **radice quadrata** di x , cioè il numero *non negativo* il cui quadrato è x . Con questa convenzione, dato che $|c| \geq 0$ e $|c|^2 = c^2$,

$$\sqrt{c^2} = |c| \quad \forall c \in \mathbb{R},$$

(un errore comune è scrivere $\sqrt{c^2} = c$, che, invece, è vera solo per $c \geq 0$).

⁷Attenzione, lo stesso NON è vero per numeri reali qualsiasi! Ad esempio, $-2 < -1$, ma $(-2)^2 = 4 > 1 = 1^2$.

Applicando il risultato precedente ad $a = \sqrt{x}$ e $b = \sqrt{y}$ con $x, y > 0$, troviamo che

$$\text{se } x, y > 0, \quad \text{allora } x > y \iff \sqrt{x} > \sqrt{y}.$$

Più in generale, se $x \geq y \geq 0$, allora $\sqrt{x} \geq \sqrt{y}$. Quindi è legittimo prendere le radici quadrate di entrambi i membri di una disequazione *quando i termini sono non negativi*. Lo stesso si può dire per le radici n -esime:

$$\text{se } x, y > 0, \quad \text{allora } x > y \iff \sqrt[n]{x} > \sqrt[n]{y}.$$

Le proprietà che abbiamo descritto rientrano in una problematica più generale. Siano $x, y \in \mathbb{R}$ con $x < y$. Che succede se trasformiamo i due valori secondo una certa regola (ad esempio, l'elevazione alla potenza n o la radice n -esima)? Proprietà di questo genere sono dette *proprietà di monotonia* e verranno riprese più in là.

Se n è dispari, la conclusione della Proposizione 2.2 vale per ogni $a, b \in \mathbb{R}$. Infatti, dalle regole di base per le disuguaglianze, se n è dispari, allora $a^n > 0$ se e solo se $a > 0$. Quindi, nel caso $b < 0 < a$ la proprietà è evidente. Se, invece, a e b sono di segno concorde, basta notare che il termine $a^{n-1} + a^{n-2}b + \dots + ab^{n-2} + b^{n-1}$ è sempre positivo, e ragionare come nella dimostrazione della Proposizione 2.2. Riassumendo:

$$\text{se } n \text{ è dispari, allora } a^n > b^n \iff a > b.$$

3. La struttura metrica: il modulo

Ricordiamo che, dato $x \in \mathbb{R}$, il **modulo** di x è definito da

$$|x| := \begin{cases} x, & x \geq 0, \\ -x, & x < 0. \end{cases}$$

L'innocuo simbolo $|\cdot|$ gode di tre proprietà che gli conferiscono poteri strabilianti.

PROPOSIZIONE 3.1 (Proprietà del modulo). *Valgono le seguenti proprietà:*

- (i) $|a| \geq 0$ per ogni $a \in \mathbb{R}$ e $|a| = 0$ se e solo se $a = 0$;
- (ii) il prodotto dei moduli è il modulo del prodotto: $|ab| = |a||b|$ per ogni $a, b \in \mathbb{R}$;
- (iii) vale la disuguaglianza triangolare:

$$|a + b| \leq |a| + |b| \quad \forall a, b \in \mathbb{R}.$$

DIMOSTRAZIONE. La prima proprietà è banale. Per la seconda, basta tenere conto della regola dei segni per il prodotto di numeri reali: ad esempio, consideriamo il caso $a < 0 < b$, allora $ab < 0$ e quindi $|ab| = -ab = a(-b) = |a||b|$. Analogamente per gli altri casi. Resta da dimostrare la disuguaglianza triangolare (iii). Distinguiamo i casi $a + b \geq 0$ e $a + b < 0$. Nel primo caso, la disuguaglianza afferma $a + b \leq |a| + |b|$, che discende direttamente da $a \leq |a|$ e $b \leq |b|$ e dalla somma di queste due disequazioni.

Nell'altro caso, la disuguaglianza diviene $-(a+b) \leq |a| + |b|$, che discende dalla somma delle disequazioni $-a \leq |a|$ e $-b \leq |b|$. \square

ESERCIZIO 3.2. Verificare la validità dell'uguaglianza: $|x| = \max\{x, -x\}$.

Dato che il modulo è indipendente dal segno, cioè $|-x| = |x|$, la disuguaglianza triangolare vale anche con il segno $-$ al posto di $+$, cioè vale anche $|a - b| \leq |a| + |b|$. E' facile trovare esempi che mostrano che per certe scelte di a e b la disuguaglianza $|a - b| \leq |a| - |b|$ è falsa.

Applicando due volte la disuguaglianza triangolare, si ottiene

$$|a + b + c| = |(a + b) + c| \leq |a + b| + |c| \leq |a| + |b| + |c|.$$

Allo stesso modo, si ottiene la disuguaglianza più generale

$$|a_1 + a_2 + \cdots + a_n| \leq |a_1| + |a_2| + \cdots + |a_n| \quad \forall a_1, \dots, a_n \in \mathbb{R},$$

o, equivalentemente (tramite il simbolo di sommatoria⁸)

$$(6) \quad \left| \sum_{k=1}^n a_k \right| \leq \sum_{k=1}^n |a_k| \quad \forall a_1, \dots, a_n \in \mathbb{R},$$

Alcune conseguenze dirette della disuguaglianza triangolare sono molto utili.

COROLLARIO 3.3. Per ogni $a, b \in \mathbb{R}$, valgono

$$(i) \quad |a| \leq |a \pm b| + |b|, \quad (ii) \quad ||a| - |b|| \leq |a \pm b|,$$

dove il simbolo \pm indica che la disuguaglianza è vera sia con $+$ che con $-$.

DIMOSTRAZIONE. Dimostriamo le versioni di (i) e (ii) con il segno $+$. Dato che

$$|a| = |(a + b) + (-b)| \leq |a + b| + |-b| = |a + b| + |b|,$$

la (i) è vera. Per la (ii), basta riutilizzare la (i) ottenendo $|a| - |b| \leq |a + b|$. Scambiando il ruolo di a e di b , si deduce anche $|b| - |a| \leq |a + b|$. La disequazione (ii) segue da $\pm(|a| - |b|) \leq |a + b|$ e da $|x| = \max\{x, -x\}$. \square

ESERCIZIO 3.4. Dimostrare che, per ogni $a > 0$,

$$|x| \leq a \quad \Longleftrightarrow \quad -a \leq x \leq a.$$

⁸La somma di n numeri x_1, \dots, x_n viene indicata con $\sum_{k=1}^n x_k := x_1 + x_2 + \cdots + x_n$. Ad esempio $\sum_{k=1}^5 k = 1 + 2 + 3 + 4 + 5 = 15$, $\sum_{k=1}^4 k^2 = 1^2 + 2^2 + 3^2 + 4^2 = 30$.

DEFINIZIONE 3.5. *Insiemi limitati.* Un sottoinsieme $E \subset \mathbb{R}$ è **limitato** se esiste $R > 0$ tale che E è contenuto in un intervallo del tipo $[-R, R]$:

$$E \subset \mathbb{R} \quad \text{limitato} \quad \Longleftrightarrow \quad \exists R > 0 \quad \text{tale che} \quad |x| \leq R \quad \forall x \in E.$$

Ad esempio, gli intervalli $[a, b]$, (a, b) , $(a, b]$ e $[a, b)$ sono limitati, mentre tutte le semirette sono illimitate.

ESERCIZIO 3.6. Dire quali tra i seguenti insiemi sono limitati e quali non lo sono:

$$\{0\} \cup \{1\}, \quad (-\infty, 0) \cup [1, 2], \quad [-1, 1] \cap [0, +\infty), \quad \mathbb{N}.$$

La distanza nella retta reale. Come si è detto, il numero $|x|$ rappresenta la distanza del punto x dall'origine 0 della retta reale \mathbb{R} . Quindi, tramite il modulo, si può dare senso alla “distanza tra numeri reali”: dati $a, b \in \mathbb{R}$,

$$\text{distanza di } a \text{ da } b: \quad |a - b|.$$

Come impareremo ad apprezzare pian piano, il fatto che sia possibile introdurre in \mathbb{R} una nozione di distanza permette di sviluppare una teoria estremamente ricca ed interessante, definendo in modo rigoroso il concetto di *limite*. Ma non corriamo troppo...

Dalle proprietà del modulo discendono direttamente alcune proprietà (anche queste più che utili) per la distanza:

- (a) *positività*: la distanza di a da b è non negativa ed è nulla se e solo se $a = b$;
- (b) *simmetria*: $|a - b| = |b - a|$;
- (c) *disuguaglianza triangolare*: vale

$$|a - b| \leq |a - c| + |c - b| \quad \forall a, b, c \in \mathbb{R}.$$

ESERCIZIO 3.7. Dimostrare le proprietà (a)-(b)-(c) della distanza utilizzando le proprietà (i)-(ii)-(iii) del modulo.

DEFINIZIONE 3.8. Dato $x_0 \in \mathbb{R}$ e $r > 0$, si chiama **intorno di x_0 di raggio r** e si indica con $I(x_0; r)$ (o con $I_r(x_0)$), l'intervallo dei numeri $x \in \mathbb{R}$ tali che $x_0 - r < x < x_0 + r$:

$$I(x_0; r) = (x_0 - r, x_0 + r).$$

In maniera equivalente, si può dire che $I(x_0; r)$ è l'insieme dei numeri reali che distano da x_0 meno del raggio r . Questo insieme, sulla retta reale, si rappresenta come un segmento di lunghezza $2r$, centrato nel punto x_0 . Dalle proprietà del modulo di un numero reale discende che la condizione “ x appartiene all'intorno di x_0 di raggio r ” si può esprimere scrivendo la condizione $|x - x_0| < r$:

$$|x - x_0| < r \quad \Longleftrightarrow \quad x \in I(x_0; r).$$

Dato che valgono (perché?)

$$|x - x_0| < r \iff -r < x - x_0 < r \iff x_0 - r < x < x_0 + r.$$

un intorno può essere scritto in maniere diverse, tutte equivalenti,

$$\begin{aligned} I(x_0; r) &= (x_0 - r, x_0 + r) = \{x : x_0 - r < x < x_0 + r\} \\ &= \{x : -r < x - x_0 < r\} = \{x : |x - x_0| < r\}. \end{aligned}$$

4. La verità sui reali

Prologo. È giunto il momento di riprendere la questione della *completezza* dei numeri reali. Partiamo da un esempio illustrativo che spiega la situazione: disegniamo due curve nel piano così come in Figura 8. La domanda è: *queste due curve si interse-*

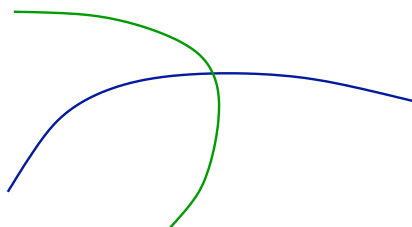


FIGURA 8. Due curve nel piano.

cano oppure no? Un sondaggio del 2003 dà queste percentuali di risposta: il 78% degli intervistati dice “SI, SEMPRE”, il 5% dice “QUALCHE VOLTA”, il 3% dice “QUASI MAI”, il 10% risponde “NON SO” e il 4% fugge scappando per timore di fare brutta figura. Evidentemente la risposta è racchiusa in quello che succede vicino al punto di incrocio delle due curve. Proponiamo tre maniere diverse di ragionare.

Versione “atomistica” (o “alla Democrito”). In questa versione, si immagina che le curve siano costituite da punti equidistanti (a distanza tanto piccola che l’occhio non è in grado di distinguerli, e vede solo un linea apparentemente continua). Con un ingrandimento di scala si vede bene che (a meno di casi particolarmente fortunati) le due curve, discretizzate in tanti atomi, non si incontrano. La risposta in questo caso è “(QUASI) MAI”.

Versione “razionale”. Questa volta, immaginiamo le curve come rette, formate dall’unione di soli punti razionali, deformate. Qui il disegno non è facile: dato che in ogni intervallo cadono infiniti punti di \mathbb{Q} , nessun ingrandimento permette di riconoscere “a occhio” se ci sia intersezione oppure no. Però sappiamo già che in alcuni casi non c’è intersezione: ad esempio, non esistono numeri razionali x tali che $x^2 = 2$, cioè le

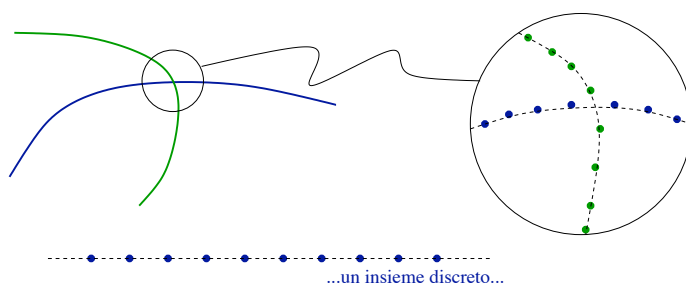


FIGURA 9. Versione “quantizzata”: una retta è un’unione di punti a distanza fissata.

curve nel piano (x, y) (con x, y razionali) definite da $y = x^2$ e $y = 2$ non si intersecano. Il seguace di questa corrente di pensiero risponde “QUALCHE VOLTA”.

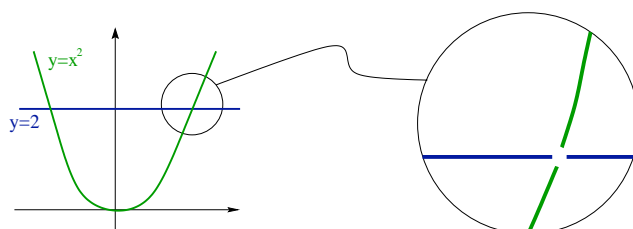


FIGURA 10. Versione “razionale”: la parabola $y = x^2$ e la retta $y = 2$ nel piano (x, y) con $x, y \in \mathbb{Q}$ non si intersecano mai.

Versione “reale”. Infine c’è la versione reale: rette e curve costituiscono un “continuo” di punti senza interruzione. Dunque le due curve si intersecano sempre. Ogni ingrandimento del punto di incontro delle curve dà sempre e comunque lo stesso tipo di figura.

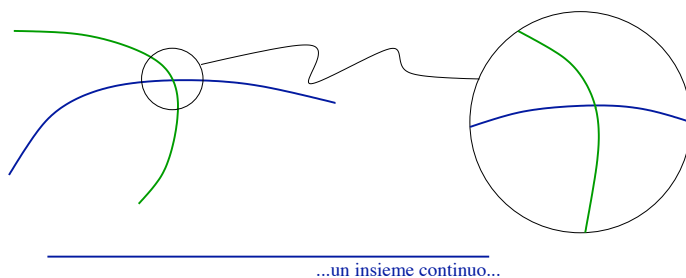


FIGURA 11. Versione “reale”: l’idea intuitiva di “continuo”.

Qual’è la risposta “esatta”? Tutte allo stesso tempo... Quello che conta, infatti, è decidere fin dall’inizio qual’è il tipo di visione che vogliamo prediligere e seguirla con coerenza e chiarezza. Scegliere una strada significa decidere qual’è l’ambiente base

(discreto, razionale, continuo) con cui lavoriamo chiarendo bene quali siano le proprietà di cui gode. Tali proprietà vengono tradotte in *assiomi* (o *postulati*) che si dichiarano veri al principio. *La nostra scelta è di lavorare con l'insieme dei numeri reali*. I motivi sono tanti, primo fra tutti il fatto che la percezione del “continuo”, cioè di un universo “senza buchi” (seppure sbagliata a livello microscopico!), è estremamente naturale.

Occorre ora rendere chiaro cosa voglia dire con precisione la frase “l'insieme dei numeri reali non ha buchi”. Ci sono vari modi per esprimere attraverso un assioma la *completezza dei numeri reali*. Qui scegliamo di utilizzare quali assiomi di completezza di \mathbb{R} la validità di due proprietà: il *postulato degli intervalli incapsulati* e il *principio di Archimede*. Per versioni equivalenti e maggiori chiarimenti sull'assioma di continuità bisogna ricorrere ad altri libri⁹.

Gli intervalli incapsulati. Un intervallo chiuso e limitato in \mathbb{R} è un insieme del tipo $[a, b] = \{x \in \mathbb{R} : a \leq x \leq b\}$ con $a, b \in \mathbb{R}$.

Postulato degli intervalli incapsulati. *Per ogni successione di intervalli $I_0, I_1, \dots, I_n, \dots$ chiusi e limitati che siano incapsulati, cioè tali che $I_{n+1} \subset I_n$ per ogni $n \in \mathbb{N}$, esiste sempre almeno un punto $x_0 \in \mathbb{R}$ tale che $x_0 \in I_n$ per ogni n .*

Il postulato degli intervalli incapsulati (detto anche *principio di Cantor*) esprime il fatto che, se si prende una sequenza di intervalli chiusi e limitati in \mathbb{R} , ciascuno dei quali sia contenuto nel precedente (e quindi contenga il successivo), allora c'è sempre almeno un numero reale contenuto in tutti quanti gli intervalli.

L'assioma si può anche esprimere affermando che:

$$I_0, \dots, I_n, \dots \text{ intervalli chiusi e limitati, } I_{n+1} \subset I_n \quad \forall n \in \mathbb{N} \quad \Rightarrow \quad \bigcap_{n \in \mathbb{N}} I_n \neq \emptyset.$$

In \mathbb{Q} , il postulato degli intervalli incapsulati è falso! Infatti basta rappresentare \mathbb{Q} con punti di una retta orientata e “circondare” il punto $\ell = \sqrt{2}$ con una sequenza di intervalli incapsulati con lunghezza che diventa sempre più piccola in modo che l'intersezione degli intervalli sia data dal solo punto ℓ . Dato che $\ell \notin \mathbb{Q}$, non c'è nessun punto di \mathbb{Q} che sia nell'intersezione di questa sequenza di intervalli incapsulati.

In effetti, il postulato degli intervalli incapsulati è stato già usato nella rappresentazione dei numeri reali tramite decimali: in quell'occasione abbiamo rappresentato un numero reale attraverso una sequenza di intervalli incapsulati di lunghezza via via più piccola (l'intervallo I_n è di lunghezza $1/10^n$). Lì l'assioma degli intervalli incapsulati

⁹Si veda, ad esempio, E.Giusti, “Esercizi e Complementi di Analisi Matematica, vol. I”, Bollati Boringhieri, cap.2.

garantisce all'ipotetico postino che deve consegnare un plico all'indirizzo $c_0, c_1 c_2 c_3 \dots$ di trovare certamente un destinatario.

L'assioma di Archimede. C'è un'altra proprietà che abbiamo usato senza grande preoccupazione perché estremamente intuitiva: esistono intervalli di lunghezza arbitrariamente piccola. In altri termini, si assume di poter effettuare ingrandimenti arbitrariamente forti di una piccola zona della retta reale a cui si è interessati. Equivalente a questa proprietà c'è quella che afferma che se riproduciamo copie del segmento unitario della retta reale \mathbb{R} ... “arriviamo all'infinito”, cioè copriamo tutta la retta. In altre parole ogni elemento di \mathbb{R} è sempre contenuto in un intervallo che ha per estremi due numeri interi. Nonostante sembri una proprietà evidente, abbiamo bisogno di un assioma *ad hoc*, cioè una regola che stabiliamo valida una volta per tutte e che, quindi, potremo utilizzare tutte le volte che ci verrà utile.

Assioma di Archimede. *Per ogni numero reale a , esiste un numero naturale n più grande di a : in simboli,*

$$\forall a \in \mathbb{R}, \quad \exists n \in \mathbb{N}, \quad \text{tale che } a < n.$$

Grazie all'Assioma di Archimede (o Proprietà di Archimede) è possibile dimostrare la densità dell'insieme dei numeri razionali in quello dei numeri reali.

PROPOSIZIONE 4.1. *Per ogni $a, b \in \mathbb{R}$ con $a < b$, esiste $r \in \mathbb{Q}$ tale che $a < r < b$.*

DIMOSTRAZIONE. Per semplicità, consideriamo il caso $0 < a < b$. Grazie alla proprietà di Archimede, dato che $1/(b-a) \in \mathbb{R}$, esiste $N \in \mathbb{N}$ tale che $N \geq 1/(b-a)$. Consideriamo il sottoinsieme di \mathbb{N} definito da

$$K := \{k \in \mathbb{N} : k \leq Na\},$$

che è composto di un numero finito di elementi e indichiamo con h il più grande degli elementi di K .

Affermazioni. Sia $r := (h+1)/N$, allora (i) $r \in \mathbb{Q}$; (ii) $r > a$ e (iii) $r < b$.

(i) Il fatto che r sia razionale è evidente.

(ii) Dato che h è il più grande degli elementi di K , $h+1 \notin K$, cioè $h+1 > Na$. Ne segue che $a < \frac{h+1}{N} = r$.

(iii) Resta da verificare che $r < b$:

$$b - r = b - \frac{h}{N} - \frac{1}{N} \geq b - \frac{Na}{N} - \frac{1}{N} = b - a - \frac{1}{N} \geq (b-a) - (b-a) = 0,$$

dato che $N \geq 1/(b-a)$. □

Una conseguenza dell'assioma di Archimede verrà utilizzata migliaia di volte nel seguito. Eccola.

PROPOSIZIONE 4.2. *Vale la seguente implicazione*

$$x \leq \varepsilon \quad \forall \varepsilon > 0 \quad \Rightarrow \quad x \leq 0.$$

DIMOSTRAZIONE. Supponiamo per assurdo $x > 0$. Per la proprietà di Archimede, esiste $n \in \mathbb{N}$ tale che $\frac{1}{x} < n$. Dato che $x > 0$, anche $1/x > 0$, quindi n è diverso da zero. Moltiplicando $\frac{1}{x} < n$ per x/n (che è positivo), si deduce che

$$\exists n \in \mathbb{N} \quad \text{tale che} \quad 0 < \frac{1}{n} < x$$

che contraddice l'ipotesi. □

5. Estremo superiore ed estremo inferiore

Tra le conseguenze notevoli della completezza dei numeri reali, una delle più significative è la possibilità di definire due oggetti fondamentali dell'analisi matematica: l'*estremo superiore* e l'*estremo inferiore* di un sottoinsieme dell'asse reale.

DEFINIZIONE 5.1. *Un valore $\Lambda \in \mathbb{R}$ è un maggiorante di E se tutti gli elementi $y \in E$ verificano $y \leq \Lambda$; un valore $\lambda \in \mathbb{R}$ è un minorante di E se tutti gli elementi $y \in E$ verificano $\lambda \leq y$. Se esiste almeno un maggiorante per l'insieme $E \subset \mathbb{R}$, E si dice limitato superiormente; se esiste almeno un minorante per l'insieme $E \subset \mathbb{R}$, E si dice limitato inferiormente.*

In sostanza, un maggiorante Λ è una stima per eccesso di tutti gli elementi dell'insieme E e un minorante λ ne è una stima per difetto. Se, in qualche modo, siamo in grado di procurarci un minorante λ ed un maggiorante Λ , sappiamo già che l'insieme E è un sottoinsieme dell'intervallo chiuso $[\lambda, \Lambda]$.

Non è difficile convincersi che un insieme è limitato (secondo la Definizione 3.5) se e solo se è limitato superiormente e limitato inferiormente. Fatelo.

Il problema è che per uno stesso insieme E è possibile fornire stime diverse! Ad esempio, se

$$E := \left\{ \frac{1}{n} : n = 1, 2, 3, \dots \right\} = \left\{ 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \dots \right\},$$

è vero sia che $E \subset [0, 1]$ che $E \subset [-100, 100]$, ma, evidentemente la prima delle due stime è preferibile all'altra perché più "precisa". La domanda successiva è più che ovvia: è possibile determinare una stima che sia "ottimale"? Quando un maggiorante Λ (o un minorante λ) è elemento dell'insieme E è chiaro che Λ (o λ) dà la migliore stima per eccesso (per difetto). In questo caso si parla di *massimo* (o di *minimo*).

DEFINIZIONE 5.2. Massimo e minimo. Il valore $M \in \mathbb{R}$ è il **massimo** di E , e si scrive $M = \max E$, se

(i) M è un maggiorante di E , (ii) M è un elemento di E .

Analogamente, il valore $m \in \mathbb{R}$ è il **minimo** di E , e si scrive $m = \min E$, se

(i) m è un minorante di E , (ii) m è un elemento di E .

Quindi, se siamo informati (da qualche agenzia investigativa segreta) del fatto che $M = \max E$, sappiamo che tutto l'insieme E è limitato dall'alto da M e che questa è la stima migliore possibile.

ESERCIZIO 5.3. Dimostrare che, se M è il massimo dell'insieme E , allora:

se M' è un altro maggiorante di E , vale $M \leq M'$.

Il problema è che ci sono sottoinsiemi di \mathbb{R} che non hanno massimo o minimo, o nessuno dei due! Ad esempio, \mathbb{R} stesso, per il principio di Archimede, non ha né massimo né minimo. Se $E = \{1, 1/2, 1/3, \dots\}$, è facile convincersi che $1 = \max E$ e che E non ha elemento minimo. L'intervallo aperto $(0, 1)$ non ha invece né massimo né minimo. Quindi dato un qualsiasi sottoinsieme di \mathbb{R} non è detto che abbia senso scrivere $\max E$ e/o $\min E$. Bisogna, perciò, introdurre dei nuovi oggetti che siano ben definiti anche quando il massimo e/o il minimo non esistono. L'idea è semplice: sostituire la condizione (ii) della Definizione 5.2 (troppo restrittiva!), con la condizione dell'Esercizio 5.3.

DEFINIZIONE 5.4. Estremi superiore e inferiore. Il valore $\Lambda \in \mathbb{R}$ è l'**estremo superiore** di E , e si scrive $\Lambda = \sup E$, se

(i) Λ è un maggiorante di E , (ii) ogni maggiorante L di E verifica $\Lambda \leq L$.

Analogamente, il valore $\lambda \in \mathbb{R}$ è l'**estremo inferiore** di E , e si scrive $\lambda = \inf E$, se

(i) λ è un minorante di E , (ii) ogni minorante ℓ di E verifica $\ell \leq \lambda$.

La proprietà (ii) dell'estremo superiore Λ garantisce che *non esiste una stima per eccesso migliore di Λ* . Ogni altro possibile maggiorante dell'insieme, necessariamente è maggiore (o uguale) a Λ . Similmente, la proprietà (ii) dell'estremo inferiore λ garantisce che *non esiste una stima per difetto migliore di λ* . In altre parole, l'estremo superiore è il più piccolo dei maggioranti e l'estremo inferiore è il più piccolo dei minoranti.

Dalle definizioni segue che, se E ammette massimo M (o minimo m), questo valore è anche l'estremo superiore (o estremo inferiore) di E . Infatti se $M = \max E$, dato che $M \in E$, si ha $M \leq L$ per ogni L maggiorante di E e quindi la condizione (ii) dell'estremo superiore è soddisfatta.

A prima vista passare dalla definizione di massimo/minimo a quella di estremo superiore/inferiore sembra una vera e propria truffa. Infatti, la definizione di estremo superiore discende da questa strategia:

- dato l'insieme E , costruirne l'insieme $F = \{\Lambda \in \mathbb{R}; x \leq \Lambda \forall x \in E\}$ dei maggioranti,
- dichiarare che l'estremo superiore di E è il minimo dell'insieme F : $\sup E = \min F$.

Ma (attenzione!) ci sono due tranelli in questa definizione: il primo è che l'insieme F potrebbe essere anche vuoto (a questo problema penseremo tra poco), il secondo, più grave, è che abbiamo appena dichiarato che non abbiamo nessuna garanzia che un sottoinsieme di \mathbb{R} abbia minimo! Dunque, chi ci assicura che $\min F$, e quindi $\sup E$, esista? Una proprietà fondamentale, che discende dall'assioma di continuità dei numeri reali, è che se esiste almeno un maggiorante (o minorante) dell'insieme E , allora *esiste sempre l'estremo superiore (o inferiore)*.

TEOREMA 5.5. Esistenza degli estremi superiore e inferiore. *Sia $E \subset \mathbb{R}$ un insieme non vuoto. Allora*

- (i) *se E è limitato superiormente, esiste $\Lambda = \sup E \in \mathbb{R}$;*
- (ii) *se E è limitato inferiormente, esiste $\lambda = \inf E \in \mathbb{R}$.*

Vedremo la dimostrazione di questo risultato più avanti, quando considereremo la strategia del *divide et impera* in dettaglio.

ESEMPIO 5.6. Gli intervalli $(-1, 1)$, $[-1, 1)$, $(-1, 1]$ e $[-1, 1]$ hanno tutti estremo superiore uguale a 1 e estremo inferiore uguale a -1 , ma solo due di loro ammettono massimo e due ammettono minimo (quali?).

Rimane da decidere il da farsi nel caso in cui l'insieme dei maggioranti e/o quello dei minoranti siano vuoti:

- se non esistono maggioranti, E è **illimitato superiormente**, si scrive $\sup E = +\infty$;
- se non esistono minoranti E è **illimitato inferiormente**, si scrive $\inf E = -\infty$.

I simboli $+\infty$ e $-\infty$ non corrispondono a nessun numero reale: non si tratta di punti sulla retta reale! Nel seguito, se parlerà di “ x finito” nel caso in cui sia utile sottolineare che si tratta di un numero reale e non di uno dei simboli $\pm\infty$.

ESERCIZIO 5.7. *Dimostrare che $\inf \left\{ \frac{1}{n} : n = 1, 2, 3, \dots \right\} = 0$.*

Le definizioni di maggiorante, minorante, massimo, minimo, estremo superiore/inferiore, limitato superiormente/inferiormente, illimitato superiormente/inferiormente si basano sull'ordinamento di \mathbb{R} , cioè sul simbolo \leq (e varianti). Quindi non hanno estensioni ai sottoinsiemi del piano \mathbb{R}^2 , dello spazio \mathbb{R}^3 o di qualsivoglia altro oggetto privo di ordine!

Forse non tutti sanno che... il fatto che l'estremo inferiore dei numeri $1/n$ con n naturale (non zero) sia 0 non è così banale come un disegno potrebbe far credere...

Torniamo all'esempio dell'insieme E dei razionali della forma $\frac{1}{n}$ con $n = 1, 2, \dots$

$$E = \left\{ 1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \frac{1}{5}, \frac{1}{6}, \dots \right\}.$$

Questo insieme è limitato inferiormente da 0 e limitato superiormente da 1. Dato che 1 è elemento dell'insieme, si ha $\sup E = \max E = 1$. Quanto vale $\inf E$? Con una

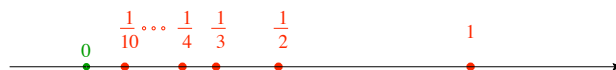


FIGURA 12.

figura ci si convince facilmente che $\inf E = 0$, ma... come si dimostra? Le figure sono sempre estremamente interessanti, ma, alle volte, possono trarre in inganno. Il fatto (forse sorprendente) è che $\inf E = 0$ *discende dall'assioma di Archimede*. Vediamone il perché.

PROPOSIZIONE 5.8. *Se vale la proprietà di Archimede, allora*

$$\inf \left\{ \frac{1}{n} : n = 1, 2, 3, \dots \right\} = 0.$$

Dimostrazione. Sia $E = \left\{ \frac{1}{n} : n = 1, 2, 3, \dots \right\}$, vogliamo dimostrare che l'insieme D dei minoranti di E , $D := \{x \in \mathbb{R} : x \leq y, y \in E\}$, coincide con la semiretta $(-\infty, 0]$.

È chiaro che $(-\infty, 0] \subset D$ dato che tutti gli elementi di E sono nonnegativi. Inoltre, dalla proprietà di Archimede, discende che per ogni $x > 0$ esiste $\bar{n} \in \mathbb{N}$ tale che $\bar{n} > \frac{1}{x}$ o, equivalentemente, che per ogni $x > 0$ esiste $\bar{n} \in \mathbb{N}$ tale che $x > \frac{1}{\bar{n}}$. Questo significa che nessun numero reale x positivo è minorante dell'insieme E . In altre parole, se $x \in (0, +\infty)$ allora $x \notin D$. Incollando quest'informazione con l'inclusione precedente, si deduce che $D = (-\infty, 0]$.

Altre partenze (ma con lo stesso punto di arrivo). Assumendo il Postulato degli intervalli incapsulati e l'Assioma di Archimede, si dimostra la validità del Teorema 5.5, relativo all'esistenza dell'estremo superiore ed inferiore.

Si può anche scegliere una strada diversa, non supporre valide *a priori* le affermazioni relative agli intervalli incapsulati e la proprietà archimedea e considerare direttamente la tesi del Teorema 5.5 come un assioma

Postulato dell'estremo superiore. *Ogni insieme superiormente limitato, ammette estremo superiore.*

In questo caso, è possibile dedurre il *Postulato degli intervalli incapsulati* e l'*Assioma di Archimede* come conseguenza dell'assunzione dell'esistenza dell'estremo superiore. Infatti, se non valesse la proprietà di Archimede, l'insieme \mathbb{N} sarebbe superiormente limitato ed ammetterebbe quindi un estremo superiore Λ , che dovrebbe necessariamente essere naturale. Quindi, esisterebbe $\Lambda \in \mathbb{N}$ tale che $n \leq \Lambda$ per ogni $n \in \mathbb{N}$, in contraddizione con il fatto che $\Lambda + 1$ è naturale ed è più grande di Λ .

Non dimostreremo in questa sede come è possibile dedurre la proprietà relativa agli intervalli incapsulati a partire dal *Postulato dell'estremo superiore*.

In quello che segue, quindi, possiamo fare riferimento al *Postulato degli intervalli incapsulati* e l'*Assioma di Archimede* o al *Postulato dell'estremo superiore* in maniera del tutto equivalente. A seconda del problema considerato, potremo utilizzare uno o l'altro come punto di partenza per dedurre nuove proprietà relative all'insieme dei numeri reali \mathbb{R} . Senza offesa per nessuno.

CAPITOLO 2

Funzioni: anno zero

1. Ingredienti di base

In tutti i campi della scienza compaiono, in modo del tutto naturale, oggetti chiamati *funzioni*: la pressione di un gas ideale è funzione della densità e della temperatura, la posizione di una particella in movimento è funzione del tempo, il volume e la superficie di un cilindro sono funzioni del raggio e dell'altezza, etc. etc.. In generale, quando certe quantità a, b, c, \dots , dette *variabili dipendenti*, sono determinate da altre quantità x, y, z, \dots , dette *variabili indipendenti*, si dice che a, b, c, \dots “sono funzioni di” x, y, z, \dots o, in modo equivalente che a, b, c, \dots “dipendono da” x, y, z, \dots . L'idea è semplice: cambiando il valore delle variabili indipendenti, cambia il valore delle variabili dipendenti.

Ecco alcuni esempi tanto per cominciare.

- i. L'area A di un quadrato di lato ℓ è data da $A = \ell^2$, quindi la variabile dipendente *area* A è funzione della variabile indipendente *lato* ℓ .
- ii. Il Teorema di Pitagora afferma: *la lunghezza ℓ dell'ipotenusa è pari alla radice quadrata della somma dei quadrati delle lunghezze a e b dei cateti*, o, in formule, $\ell = \sqrt{a^2 + b^2}$. In questo caso, la lunghezza dell'ipotesa è una funzione delle lunghezze dei cateti: la variabile dipendente è ℓ , mentre le variabili indipendenti sono a e b .
- iii. Esistono anche oggetti che associano ad una sola variabile indipendente t , due variabili dipendenti x e y . Ad esempio, la funzione

$$x = t + 1, \quad y = 2 - t^3.$$

Interpretando x e y come coordinate di un punto P nel piano e t come il tempo, queste equazioni descrivono la posizione di P al tempo t , cioè il *moto* del punto P .

In generale, una **funzione** è una legge che associa ad ogni dato valore di una variabile (indipendente) un unico valore di un'altra variabile (dipendente). In termini più informatici, si può pensare alla variabile indipendente come **Input** della funzione e alla variabile dipendente come **Output**.

$$\text{Input} \longrightarrow \boxed{\text{Funzione}} \longrightarrow \text{Output}$$

Nella prima parte di queste Note, approfondiremo il caso delle *funzioni che associano ad un numero reale un altro numero reale* (vedi esempio **i.**). In questa situazione, si usa la notazione

$$f : I \subset \mathbb{R} \rightarrow \mathbb{R},$$

che esprime che la funzione f è definita per valori $x \in I$ dove I è un assegnato sottoinsieme della retta reale \mathbb{R} e che la f trasforma x nel valore $y = f(x)$ di \mathbb{R} . Quindi, per definire una funzione occorre conoscere:

- i valori della variabile indipendente per cui la funzione f è considerata (l'insieme I);
- in quale insieme “vive” la variabile dipendente (qui l'insieme dei numeri reali);
- la regola definita dalla funzione f .

Volete qualche altro esempio? Eccovene un paio:

$$f(x) = x^2 \quad x \in \mathbb{R} \quad \text{oppure} \quad f(x) = \sqrt{1 - x^2} \quad -1 \leq x \leq 1$$

Frequentemente useremo i seguenti vocaboli, con cui ci si familiarizza col tempo.

Piccolo glossario per le funzioni

- x : variabile indipendente
- y : variabile dipendente
- I : dominio di definizione (o campo di esistenza)
- \mathbb{R} (di arrivo): codominio
- $f(x)$: immagine di x (o trasformato di x)
- x : (una) pre-immagine, o contro-immagine, di $f(x)$
- $f(I) = \{y \in \mathbb{R} : y = f(x) \text{ per qualche } x \in I\}$:
insieme immagine o immagine (di I tramite f)¹

OSSERVAZIONE 1.1. L'assegnazione di una funzione include anche la definizione del dominio della funzione. Funzioni con la stessa espressione analitica, ma differente dominio di definizione sono da considerarsi funzioni diverse! Ad esempio, la funzione $f(x) = x^2$ per $0 < x < 2$ non coincide con la funzione $g(x) = x^2$ per $x \in \mathbb{R}$, dato che il loro dominio di definizione è diverso.

Nell'esempio appena descritto, però, le funzioni f e g coincidono in $(0, 2)$, cioè nell'insieme in cui è definita la funzione f . In questo caso, si utilizza la definizione seguente.

DEFINIZIONE 1.2. Restrizione ed estensione. La funzione $f : I_f \subset \mathbb{R} \rightarrow \mathbb{R}$ è una restrizione della funzione $g : I_g \subset \mathbb{R} \rightarrow \mathbb{R}$ (e g è una estensione di f) se l'insieme di

¹In inglese, si parla di *range* della funzione f .

definizione di g contiene quello di f e le due funzioni coincidono dove sono definite entrambe, cioè se

$$I_f \subset I_g \quad e \quad f(x) = g(x) \quad \forall x \in I_f.$$

Usualmente, se una funzione viene assegnata dandone l'espressione analitica, ma senza specificarne l'insieme di definizione, si intende che la funzione è considerata nell'insieme più grande in cui le operazioni richieste sono lecite. Ad esempio, la funzione $f(x) = x^3 + 1$ si considera definita in $I = \mathbb{R}$, mentre la funzione $f(x) = \frac{1}{x}$ è definita in $I = \mathbb{R} \setminus \{0\}$.

Grafico di funzioni. Per individuare proprietà delle funzioni è utile realizzarne una rappresentazione grafica: il **grafico della funzione** f è il sottoinsieme del piano

$$\Gamma := \{(x, y) \in \mathbb{R}^2 : x \in I, y = f(x)\}.$$

Per iniziare, vediamo alcuni esempi.

(i) y è una “funzione affine” di x , cioè la funzione f è un polinomio di grado 1:

$$f(x) = ax + b \quad \text{per qualche } a, b \in \mathbb{R}, \quad a \neq 0.$$

Come è noto dalla geometria elementare, il grafico è una retta nel piano.

(ii) y è inversamente proporzionale a x ,

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

Questa funzione è definita per $x \neq 0$ dato che la divisione per zero non ha senso. Il grafico rappresenta una *iperbole (rettangolare)*.

(iii) y è il quadrato di x ,

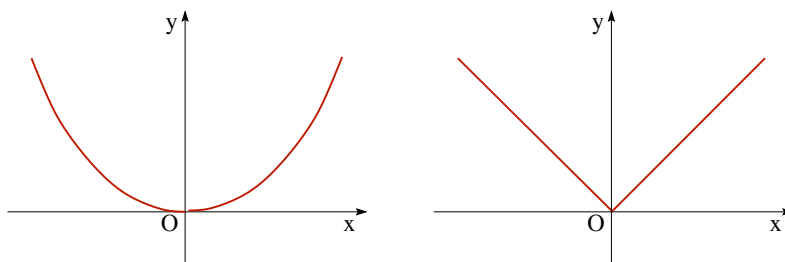
$$f(x) = x^2$$

come è ben noto questa funzione ha per grafico una parabola (vedi Fig.1(a)). Lo stesso vale per le funzioni del tipo $f(x) = ax^2 + bx + c$, con $a, b, c \in \mathbb{R}$, $a \neq 0$.

(iv) y è uguale a $|x|$. Dato che, per definizione

$$|x| := \begin{cases} x, & x \geq 0, \\ -x, & x < 0 \end{cases}$$

il grafico della funzione è composto da due semirette (vedi Fig.1(b)).

FIGURA 1. (a) La parabola $y = x^2$; (b) Il modulo: $y = |x|$.

Polinomi. Il tipo più semplice di funzione si ottiene utilizzando le sole operazioni di somma e moltiplicazione: un **polinomio** (di grado n) è una funzione della forma

$$y = a_0 + a_1x + \cdots + a_nx^n \quad a_n \neq 0.$$

dove a_0, a_1, \dots, a_n (con $a_n \neq 0$) sono $n+1$ numeri reali assegnati. Quindi $y = 3x+1$, $y = x^2 - 2x + 5$, $5x^{47} + 47x^5$ sono esempi di polinomi.

ESERCIZIO 1.3. *Disegnare i grafici delle seguenti funzioni*

$$f(x) = 1, \quad f(x) = x, \quad f(x) = 2x + 1, \quad f(x) = 2x^2 + x + 1.$$

E' una buona idea quella di sperimentare al calcolatore come siano fatti i grafici di polinomi. In particolare è interessante contare il “numero di oscillazioni” delle funzioni al variare del grado, dove per “numero di oscillazioni” si intende il numero delle zone in cui il grafico “sale” e di quelle in cui “scende”. Qual è la regola generale?

Funzioni razionali. I rapporti di polinomi sono dette **funzioni razionali**

$$y = \frac{a_0 + a_1x + \cdots + a_nx^n}{b_0 + b_1x + \cdots + b_mx^m} \quad \text{con } a_i, b_j \in \mathbb{R}, \quad (b_j \text{ non tutti nulli})$$

e sono definite per tutti i valori di x per cui il denominatore è diverso da zero. Per quanto riguarda gli zeri della funzione, questi sono tutti e soli gli zeri del polinomio a numeratore (l'unico modo per ottenere zero da un rapporto è che il numeratore sia zero). Lo studio del segno si traduce invece in un sistema di disequazioni.

Una buona classe per iniziare lo studio delle funzioni razionali è

$$f(x) = \frac{ax + b}{cx + d} \quad a, b, c, d \in \mathbb{R}.$$

Ad esempio, consideriamo la funzione

$$f(x) = \frac{2x + 3}{x + 1}.$$

L'insieme di definizione è $I = \{x : x \neq -1\}$, inoltre la funzione è positiva per $x > 1$ e per $x \leq -3/2$ e negativa nel resto dell'insieme. Il grafico è in Figura 2(a).

Un altro esempio di funzione razionale facile è $f(x) = 1/x^2$ (vedi Figura 2(b)).

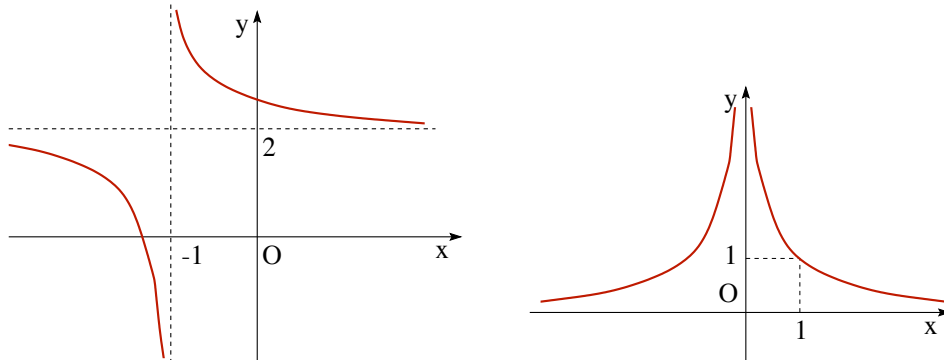


FIGURA 2. (a) La funzione $y = (2x + 3)/(x + 1)$; (b) la funzione $y = 1/x^2$.

Funzioni trigonometriche. Non è possibile in poche righe ricordare tutto il necessario sulle funzioni trigonometriche. Qui ci limitiamo alle proprietà principali. Le funzioni trigonometriche di base sono $\sin x$ e $\cos x$ le cui proprietà fondamentali sono:

- entrambe sono definite per ogni valore reale x ;
- $\cos 0 = 1$ e $\sin 0 = 0$;
- per ogni $x \in \mathbb{R}$, si hanno $\cos(x + 2\pi) = \cos x$ e $\sin(x + 2\pi) = \sin x$;
- per ogni $x \in \mathbb{R}$, vale la relazione $\cos^2 x + \sin^2 x = 1$;
- per ogni $\alpha, \beta \in \mathbb{R}$, valgono le formule di somma e sottrazione

$$\cos(\alpha \pm \beta) = \cos \alpha \cos \beta \mp \sin \alpha \sin \beta \quad \sin(\alpha \pm \beta) = \sin \alpha \cos \beta \pm \cos \alpha \sin \beta$$

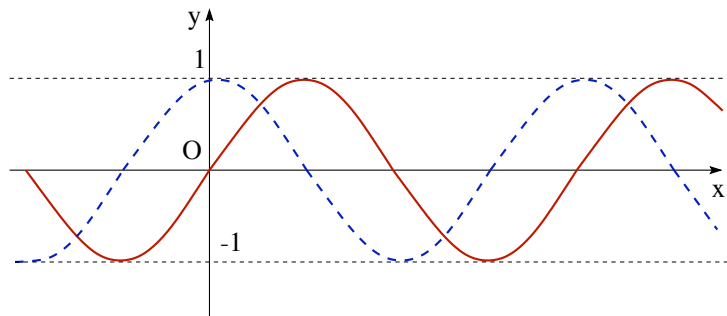


FIGURA 3. Il grafico della funzione $\sin x$ (linea continua) e della funzione $\cos x$ (linea tratteggiata).

Dato che $\sin^2 x + \cos^2 x = 1$, è sempre vero che $|\sin^2 x|, |\cos^2 x| \leq 1$ e quindi

$$|\sin x| \leq 1, \quad |\cos x| \leq 1 \quad \forall x \in \mathbb{R}.$$

ESERCIZIO 1.4. Dimostrare che $|\sin x| \leq |x|$ per ogni $x \in \mathbb{R}$.

ESERCIZIO 1.5. Dedurre, dalle formule di somma e sottrazione, la formula

$$\sin x - \sin y = 2 \cos \left(\frac{x+y}{2} \right) \sin \left(\frac{x-y}{2} \right).$$

Soluzione. Poniamo $\xi = \frac{x+y}{2}$ e $\eta = \frac{x-y}{2}$. Allora $x = \xi + \eta$ e $y = \xi - \eta$. Dunque $\sin x - \sin y = \sin(\xi + \eta) - \sin(\xi - \eta) = \sin \xi \cos \eta + \cos \xi \sin \eta - \sin \xi \cos \eta + \cos \xi \sin \eta = 2 \cos \xi \sin \eta$, e ricordando la definizione di ξ ed η si giunge alla conclusione.

Tramite le funzioni $\sin x$ e $\cos x$ si definiscono le funzioni **tangente** e **cotangente**:

$$\tan x := \frac{\sin x}{\cos x} \quad \text{e} \quad \cot x := \frac{\cos x}{\sin x}.$$

Dalla definizione e dalle proprietà di seno e coseno, discende che $\tan x$ è definita per $x \neq \frac{\pi}{2} + k\pi$ per $k \in \mathbb{Z}$ e $\cot x$ è definita per $x \neq k\pi$ per $k \in \mathbb{Z}$.

2. Operazioni elementari su grafici

Una volta noto il grafico di una funzione f è possibile, a partire da questo, ricostruire il grafico di altre funzioni g che si ottengano dalla prima per via elementare. Vediamo alcuni esempi significativi, tenendo conto che, qui, l'unica maniera per capire è *sperimentare* (anche usando un computer o una calcolatrice grafica, se possibile).

(i) *Traslazioni.* Il grafico di $g(x) = f(x) + c$ dove $c \in \mathbb{R}$ è dato da una traslazione in verticale del grafico di f della quantità c (la traslazione sarà quindi verso l'alto se $c > 0$ e verso il basso se $c < 0$).

Il grafico di $g(x) = f(x + c)$ dove $c \in \mathbb{R}$ è dato da una traslazione in orizzontale di $-c$ del grafico di f . Nota bene! La traslazione è di $-c$, quindi è verso sinistra se $c > 0$ e verso destra se $c < 0$.

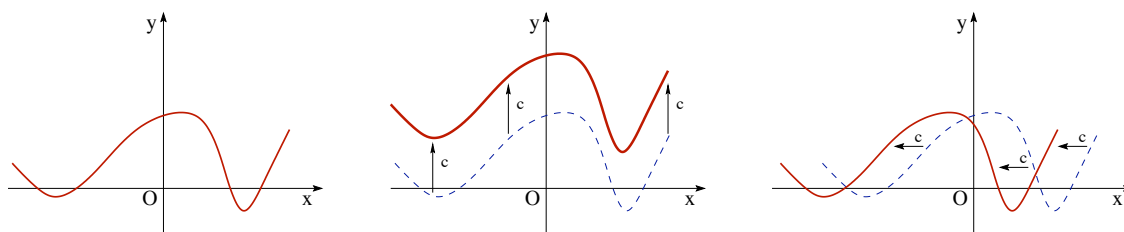


FIGURA 4. I grafici di (a) $y = f(x)$, (b) $y = f(x) + c$; e (c) $y = f(x + c)$.

(ii) *Dilatazioni/Compressioni.* Il grafico di $g(x) = kf(x)$ è ottenuto dilatando la variabile dipendente di un fattore k , il grafico è pertanto dilatato nella direzione verticale. Il grafico di $g(x) = f(kx)$ è ottenuto dilatando la variabile indipendente di un fattore

$1/k$, quello che per la funzione f accadeva in x ora per la funzione g accade in x/k . Questo vuol dire che se $k > 1$ il grafico risulta compresso in orizzontale verso l'asse y , mentre se $k < 1$ il grafico risulta dilatato. Un esempio? Fate il grafico di

$$f(x) = |x| - 1, \quad g(x) = |2x| - 1, \quad h(x) = \left| \frac{x}{2} \right| - 1.$$

Visto che ci siete, fate anche $l(x) = 2||x| - 1|$ e $m(x) = \frac{1}{2}||x| - 1|$.

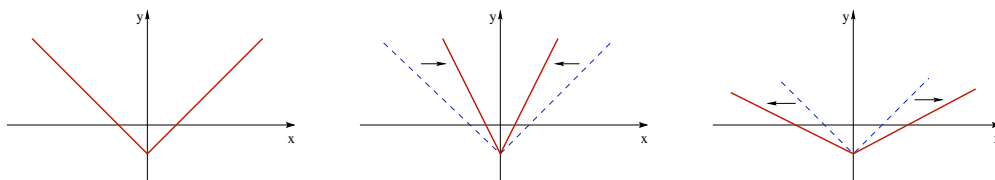


FIGURA 5. I grafici di (a) $y = f(x) = |x| - 1$, (b) $g(x) = |2x| - 1$, (c) $h(x) = \left| \frac{x}{2} \right| - 1$.

ESERCIZIO 2.1. Disegnare i grafici delle funzioni

$$f(x) = ||x| - 1|, \quad g(x) = ||3x| - 1|, \quad h(x) = \frac{1}{3}||x| - 1|.$$

(iii) *Somma/Sottrazione*. Dati i grafici di f e g è possibile stabilire un andamento qualitativo anche delle funzioni $h = f + g$ e $l = f - g$. Basta disegnare i due grafici di f e g sullo stesso piano (x, y) e poi calcolare punto per punto la somma e la differenza. Nel caso della differenza, il significato del grafico è di “distanza” con segno (cioè l è positiva se f è sopra g e negativa se f è sotto g) tra i punti, aventi stessa ascissa, dei grafici delle due funzioni. Quindi la “distanza” qui è calcolata in verticale (non è la distanza nel piano...).

(iv) *Passaggio al reciproco*. Dato il grafico della funzione f è possibile anche determinare i grafici delle funzioni $g(x) = \frac{1}{f(x)}$ e $h(x) = f(1/x)$. Come si dovrebbe essere capito dai casi precedenti, nel primo caso si ottiene una trasformazione “in verticale” (nel senso della variabile dipendente y), mentre nel secondo “in orizzontale” (nel senso della variabile dipendente x).

Il grafico di g si ottiene dalla f notando che i valori che vengono mandati da f vicino a zero sono trasformati per g in valori grandi, mentre i valori che la f trasforma in valori grandi, sono mandati da g in zero. I valori che vanno in ± 1 rimangono gli stessi. Un grafico di quel che fa la trasformazione $t \rightarrow s = 1/t$ dall'asse t all'asse s aiuta a capire cosa sta succedendo.

Per quanto riguarda il grafico della funzione h , questa volta l'inversione è compiuta sulla variabile indipendente x , quindi l'inversione è in orizzontale.

(v) *Modulo di una funzione.* Una classe significativa è quella delle funzioni della forma $g(x) = |f(x)|$ dove si suppone noto il grafico della funzione f . Dato che il modulo $|\cdot|$ trasforma un numero in sé stesso se è positivo, e nel suo opposto se è negativo, per fare il grafico di g basta lasciare invariata la parte del grafico di f che corrisponde a valori positivi della variabile dipendente, cioè la parte che è sopra l'asse delle x , e ribaltare attorno all'asse x la parte del grafico che si trova al di sotto.

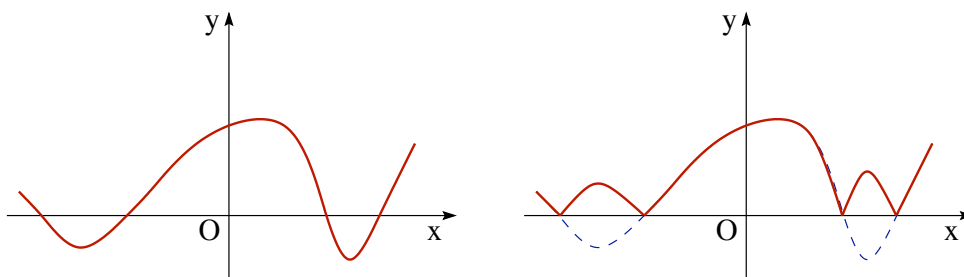


FIGURA 6. I grafici di (a) $y = f(x)$, (b) $y = |f(x)|$.

(vi) *Parte positiva e parte negativa di una funzione.* Data una funzione f , la funzione $\max\{f(x), 0\}$ è detta **parte positiva** di f e la funzione $\max\{-f(x), 0\}$ è detta **parte negativa** di f . Il grafico della prima delle due si ottiene molto facilmente a partire da quello della f : coincide con quest'ultimo dove $f(x) \geq 0$ e, in quel che resta, coincide con l'asse delle x . Il grafico della parte negativa richiede un piccolo sforzo in più: prima si ribalta il grafico della funzione f attorno all'asse x (cioè si disegna il grafico della funzione $-f$) e poi si procede come per la parte positiva. Si noti, in particolare, che sia il grafico della parte positiva che quello della parte negativa giacciono nel semipiano $y \geq 0$.

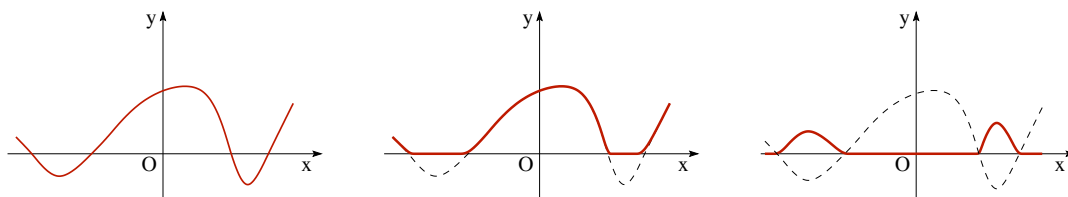


FIGURA 7. I grafici di (a) $y = f(x)$, (b) $y = \max\{f(x), 0\}$, (c) $y = \max\{-f(x), 0\}$.

Simmetrie di grafici. Capita spesso che le funzioni che si studiano abbiano della *simmetrie*, cioè abbiano la proprietà che il loro grafico rimane immutato quando vengono compiute certe trasformazioni. Ad esempio, il grafico di una funzione costante, dato che è una retta orizzontale, rimane invariato se viene traslato in orizzontale. Se si riconosce una simmetria di una funzione, lo studio è in genere semplificato, perché il

grafico può essere determinato studiandone semplicemente una parte e poi applicando una trasformazione opportuna. Vediamo rapidamente i principali tipi di simmetria.

Se il grafico di una funzione f è simmetrico rispetto all'asse y si dice che la funzione è **pari**. Analiticamente, questa proprietà corrisponde a

$$\text{funzione pari:} \quad f(-x) = f(x) \quad \forall x \in I.$$

Ad esempio le funzioni $y = x^2$, $y = |x|$ sono funzioni pari.

Se il grafico è simmetrico rispetto all'origine, la funzione è **dispari**

$$\text{funzione dispari:} \quad f(-x) = -f(x) \quad \forall x \in I.$$

Ad esempio, le funzioni $y = x^3$ e $y = 1/x$ sono dispari.

Le funzioni pari più semplici sono i polinomi che includano solo potenze pari di x . Le funzioni dispari più semplici sono i polinomi che includano solo potenze dispari di x . La funzione $\cos x$ è pari: sono pari quindi somme, differenze e prodotti di $\cos x$. La funzione $\sin x$ è dispari. E' vero che somme/prodotti di $\sin x$ sono dispari?

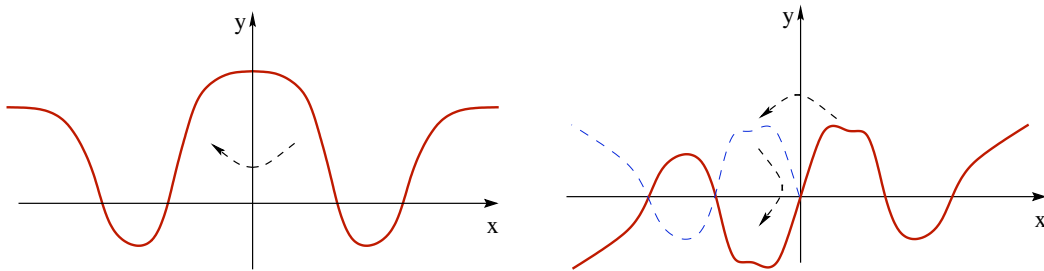


FIGURA 8. (a) una funzione pari, (b) una funzione dispari.

Una funzione $y = f(x)$ si dice **periodica** se

$$\text{funzione periodica:} \quad \exists T > 0 \text{ tale che } \forall x \quad f(x + T) = f(x).$$

Qualora esista, il più piccolo valore T per cui vale questa proprietà è il **periodo** della funzione f . Graficamente questa proprietà corrisponde al fatto che il grafico può essere ricostruito con “copia/incolla”: si determina il grafico della funzione in un intervallo di lunghezza T e poi lo si riproduce a destra e a sinistra dell'intervallo. Esempi di funzioni periodiche sono le funzioni trigonometriche: $\sin x$, $\cos x$, $\tan x$, $\cot x$,

$$\sin(x + 2\pi) = \sin x, \quad \cos(x + 2\pi) = \cos x,$$

$$\tan(x + \pi) = \tan x, \quad \cot(x + \pi) = \cot x.$$

Tanto per provare, verifichiamo la periodicità di $\tan x$:

$$\tan(x + \pi) = \frac{\sin(x + \pi)}{\cos(x + \pi)} = \frac{\sin x \cos \pi + \cos x \sin \pi}{\cos x \cos \pi - \sin x \sin \pi} = \frac{-\sin x}{-\cos x} = \frac{\sin x}{\cos x} = \tan x$$

ESERCIZIO 2.2. Qualcuna delle funzioni seguenti è pari/dispari/periodica?

$$\cos(x^2), \quad |\cos x|, \quad x|x|.$$

Soluzione. La prima funzione è pari, infatti $f(-x) = \cos((-x)^2) = \cos(x^2) = f(x)$. Anche la seconda è pari (verificare!). E' anche periodica di periodo π , infatti

$$|\cos(x + \pi)| = |\cos(x)\cos(\pi) - \sin(x)\sin(\pi)| = |-\cos(x)| = |\cos x|.$$

La terza funzione è dispari: $f(-x) = -x|-x| = -x|x| = -f(x)$.

Un altro esempio (meno frequente) di funzione periodica è la *parte frazionaria*. Sia

$$\text{parte intera di } x: \quad [x] := z \in \mathbb{Z},$$

dove z è l'unico intero per cui $z \leq x < z + 1$. Allora la funzione

$$\text{parte frazionaria (o mantissa) di } x: \quad \{x\} := x - [x]$$

è una funzione periodica, con periodo 1 (verificare!).

3. Funzioni invertibili e funzioni monotone

Data $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ e dato $y \in \mathbb{R}$, un problema tipico è determinare se ci sono (e quante) soluzioni di $f(x) = y$. In altri termini, data y si vuole sapere quante sono le sue pre-immagini tramite f . Per definizione di insieme immagine (vd. il "Piccolo Glossario per Funzioni"), il problema ammette almeno una soluzione se e solo se $y \in f(I)$.

Problema: sia $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ una funzione assegnata,

dato $y \in \mathbb{R}$, quante soluzioni esistono dell'equazione $f(x) = y$?

DEFINIZIONE 3.1. Una funzione f è *iniettiva* se manda valori di x diversi in valori y diversi, ossia

$$f \text{ è iniettiva se } f(x_1) = f(x_2) \Rightarrow x_1 = x_2.$$

In parole povere, se l'equazione $f(x) = y$ ammette sempre non più di una soluzione la funzione f si dice *iniettiva* (o *uno a uno*). Potrebbero però esserci dei valori y per cui il problema non ha soluzione.

Graficamente, l'iniettività corrisponde al fatto che *rette parallele all'asse delle x intersecano il grafico della funzione f al più una volta*. Ad esempio, la funzione x è iniettiva, mentre le funzioni x^2 e $|x|$ non lo sono.

Vediamo, in alcuni esempi concreti, come verificare se una funzione è iniettiva. La strategia pratica è supporre che valga l'uguaglianza $f(x_1) = f(x_2)$ per x_1, x_2 generici, e domandarsi se ne segue $x_1 = x_2$:

$$f(x_1) = f(x_2) \stackrel{?}{\Rightarrow} x_1 = x_2$$

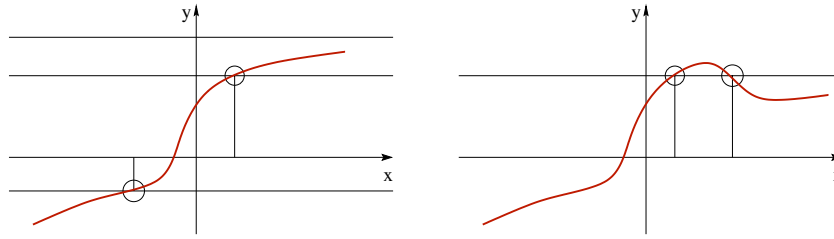


FIGURA 9. (a) una funzione iniettiva, (b) una funzione non iniettiva.

ESEMPIO 3.2. Consideriamo la funzione $f(x) = 3x - 2$ e supponiamo che esistano x_1, x_2 tale che $f(x_1) = f(x_2)$. Allora

$$3x_1 - 2 = 3x_2 - 2 \iff 3x_1 = 3x_2 \iff x_1 = x_2,$$

quindi la funzione è iniettiva. Alla stessa conclusione si giunge disegnando il grafico della retta $y = 3x - 2$ e osservando che, dato che la retta è obliqua, la proprietà geometrica dell'iniettività è soddisfatta.

ESEMPIO 3.3. Consideriamo

$$f(x) = \frac{x+1}{x+2} \quad I = \{x \neq -2\}.$$

Studiamone l'iniettività: siano $x_1, x_2 \in I$ tali che

$$\begin{aligned} \frac{x_1+1}{x_1+2} = \frac{x_2+1}{x_2+2} &\iff (x_2+2)(x_1+1) = (x_2+1)(x_1+2) \\ &\iff 2x_1 + x_2 = 2x_2 + x_1 \iff x_1 = x_2. \end{aligned}$$

Quindi la funzione è iniettiva.

ESEMPIO 3.4. Consideriamo la funzione $f(x) = x^2 + x$. Pensando al suo grafico, è evidente che non si tratta di una funzione iniettiva, ma come si può riconoscere questo fatto direttamente dai conti algebrici? Procediamo come in precedenza:

$$x_1^2 + x_1 = x_2^2 + x_2 \iff (x_1 - x_2)(x_1 + x_2 + 1) = 0.$$

Se supponiamo $x_1 \neq x_2$, allora otteniamo $x_1 + x_2 + 1 = 0$, da leggersi come una condizione su x_1 e x_2 che garantisce $f(x_1) = f(x_2)$. Ad esempio, scegliendo $x_1 = -1$ e $x_2 = 0$, otteniamo

$$x_1 = -1, \quad x_2 = 0 \implies f(x_1) = f(x_2) = 0,$$

quindi la funzione non è iniettiva.

Se una funzione è iniettiva, per tutti i valori y per cui l'equazione $y = f(x)$ ha soluzione è possibile associare un unico valore x dato proprio dalla soluzione dell'equazione $f(x) = y$. In altre parole, se una funzione è iniettiva è possibile definire

una nuova funzione che permette di “tornare indietro”, cioè che associa ad ogni y dell’insieme immagine, l’unico valore x di cui è immagine.

DEFINIZIONE 3.5. *Data una funzione f iniettiva, la funzione $f^{-1} : f(I) \subset \mathbb{R} \rightarrow \mathbb{R}$ che gode della proprietà*

$$f(f^{-1}(y)) = y \quad f^{-1}(f(x)) = x \quad \forall y \in f(I), x \in I,$$

si dice funzione inversa di f . La funzione f si dice invertibile.

Per determinarne l’inversa f^{-1} dobbiamo, sostanzialmente, esplicitare la funzione in termini della variabile y , cioè, a partire dall’espressione $y = f(x)$, arrivare ad una espressione *equivalente* della forma $x = f^{-1}(y)$.

ESEMPIO 3.6. Nel caso della funzione $f(x) = 3x - 2$, si ha

$$y = 3x - 2 \quad \Longleftrightarrow \quad 3x = y + 2 \quad \Longleftrightarrow \quad x = \frac{y + 2}{3}.$$

La funzione inversa è $f^{-1}(y) = \frac{y + 2}{3}$.

ESEMPIO 3.7. Per la funzione

$$f(x) = \frac{x + 1}{x + 2} \quad I = \{x \neq -2\}.$$

si ha

$$y = \frac{x + 1}{x + 2} \quad \Longleftrightarrow \quad xy + 2y = x + 1 \quad \Longleftrightarrow \quad x(y - 1) = 1 - 2y \quad \Longleftrightarrow \quad x = \frac{1 - 2y}{y - 1}.$$

Tale espressione ha senso solo per $y \neq 1$. In effetti, nel caso $y = 1$, si trova la relazione $x + 1 = x + 2$ cioè $1 = 2$ che è falsa, quindi $1 \notin f(I)$. La funzione inversa è definita in $\{y \neq 1\}$ ed è data da

$$f^{-1}(y) = \frac{1 - 2y}{y - 1}.$$

Conoscendo il grafico di una funzione f , si può sempre ottenere il grafico dell’inversa f^{-1} . Infatti, dato che $y = f(x)$ se e solo se $x = f^{-1}(y)$, si ha $(x, f(x)) = (f^{-1}(y), y) \in \Gamma$; quindi il grafico della funzione inversa si ottiene scambiando il ruolo dell’asse x e dell’asse y , ossia ribaltando il grafico attorno alla retta $y = x$ (vedere Figura 10).

Funzioni monotone. Una funzione $y = f(x)$ il cui valore immagine cresce se cresce la variabile indipendente, cioè tale che, per ogni $x, x' \in I$,

$$x < x' \quad \Longleftrightarrow \quad f(x) < f(x')$$

si dice **monotona strettamente crescente** in I o, più semplicemente, **strettamente crescente** in I . Analogamente, se, per ogni $x, x' \in I$,

$$x < x' \quad \Longleftrightarrow \quad f(x) > f(x'),$$

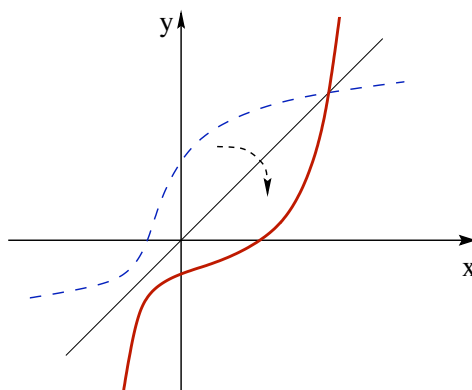


FIGURA 10. Il grafico della funzione inversa di una funzione assegnata.

la funzione è **monotona strettamente decrescente** o semplicemente **strettamente decrescente**. Equivalentemente, si può scrivere

$$f \text{ strettamente crescente} \iff [f(x) - f(x')](x - x') > 0 \quad \forall x \neq x'$$

$$f \text{ strettamente decrescente} \iff [f(x) - f(x')](x - x') < 0 \quad \forall x \neq x'$$

Per ogni $n \in \mathbb{N}$, la funzione $y = x^n$ ristretta a $x \geq 0$ è una funzione monotona strettamente crescente (come si è visto nella sezione delle disequazioni). Più precisamente, per n dispari, la funzione x^n è strettamente crescente in \mathbb{R} (quindi anche per i negativi), mentre per n pari la funzione x^n non è monotona in \mathbb{R} .

ESERCIZIO 3.8. *Dimostrare che se f e g sono funzioni strettamente crescenti allora anche la funzione $f + g$ è strettamente crescente. E' vero che anche la funzione fg è strettamente crescente?*

Soluzione. Per ipotesi, se $x < y$, allora $f(x) < f(y)$ e $g(x) < g(y)$. Quindi, sommando i termini di destra e i termini di sinistra, si ottiene

$$f(x) + g(x) < f(y) + g(y) \quad \forall x, y \quad x < y.$$

La risposta alla domanda finale è “NO”. Basta infatti considerare $f(x) = g(x) = x$, che è strettamente crescente, mentre $f(x)g(x) = x^2$ non lo è. Quale ipotesi aggiuntiva occorre per dedurre che il prodotto di funzioni strettamente crescenti è crescente?

Chiaramente valgono le implicazioni

$$f \text{ strettamente monotona} \implies f \text{ iniettiva} \iff f \text{ invertibile}$$

Ad esempio, le funzioni x^n per $x \geq 0$ e n pari e x^n per $x \in \mathbb{R}$ e n dispari sono funzioni invertibili.

OSSERVAZIONE 3.9. La monotonia di una funzione f ne garantisce l'invertibilità senza passare per la determinazione della formula per la funzione inversa: la funzione

f^{-1} c'è, ma non si vede! Ad esempio, per ogni n dispari, la funzione $f(x) = x + x^n$ è una funzione strettamente crescente, dato che è somma di funzioni strettamente crescenti. Quindi è anche una funzione invertibile.

Se la funzione f è strettamente crescente/decrescente, anche la sua inversa lo è. Infatti, siano $y = f(x)$ e $y' = f(x')$, allora $x = f^{-1}(y)$ e $x' = f^{-1}(y')$ e quindi

$$[f(x) - f(x')](x - x') = (y - y')[f^{-1}(y) - f^{-1}(y')].$$

Quindi il segno del secondo membro è lo stesso del primo, ossia le funzioni f e f^{-1} hanno lo stesso tipo di monotonia.

Nel caso in cui, nella definizione di monotonia, il simbolo di disequazione $<$ venga sostituito con la versione indebolita \leq e $>$ con \leq si parla di funzioni **non decrescenti** o **non crescenti**: Attenzione però al fatto che le funzioni non crescenti e quelle non decrescenti possono essere non iniettive, e quindi non invertibili (ad esempio, le funzioni costanti!).

4. Classi di funzioni più o meno comuni

Radici n -esime. Sia n un numero naturale $n \geq 2$. Dato che $y = x^n$ è strettamente crescente per $x \geq 0$, essa è iniettiva e quindi invertibile. La sua inversa si indica con

$$y = \sqrt[n]{x} = x^{1/n}.$$

Per definizione questa radice è sempre non negativa.

Per n dispari, però, la funzione x^n è strettamente crescente per tutti i valori $x \in \mathbb{R}$ (quindi anche per i negativi) e, di conseguenza, *per n dispari, $\sqrt[n]{x}$ è definita per tutti i valori della x* ; in questo caso $\sqrt[n]{x}$ è negativa per x negativa.

Più in generale, data una funzione g , possiamo considerare funzioni della forma

$$f(x) = \sqrt[n]{g(x)}.$$

Nel caso in cui n sia pari, una funzione di questo genere è definita solo per i valori della x per cui $g(x) \geq 0$. Ad esempio, dove è definita la funzione $f(x) = \sqrt{1 - x^2}$? Facile. Basta imporre la condizione $1 - x^2 \geq 0$, quindi per $x \in [-1, 1]$. Invece la funzione $h(x) = \sqrt[3]{1 - x^2}$ è definita per ogni valore x .

Inverse delle funzioni trigonometriche. Le funzioni trigonometriche sono periodiche, quindi non iniettive e non invertibili se considerate in tutto l'insieme in cui sono definite. Opportune restrizioni di queste funzioni sono però monotone e quindi invertibili. Vediamole in dettaglio.

Funzione arcoseno. La funzione $f(x) = \sin x$ è una funzione periodica su \mathbb{R} , quindi ad ogni elemento della sua immagine corrispondono infinite pre-immagini. Ad esempio,

$f^{-1}(1) = \{\frac{\pi}{2} + 2k\pi : k \in \mathbb{Z}\}$. Perciò la funzione $\sin x$ non è invertibile. Invece, la restrizione di $\sin x$ all'intervallo $[-\frac{\pi}{2}, \frac{\pi}{2}]$ è una funzione crescente e quindi anche invertibile. Questa inversa si chiama **(funzione) arcoseno** e si indica con $\arcsin x$. Il dominio dell'arcoseno è $[-1, 1]$ e l'insieme immagine è $[-\frac{\pi}{2}, \frac{\pi}{2}]$:

$$\arcsin : [-1, 1] \longrightarrow \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

Infine, dato che la funzione $\sin x$ è crescente nell'intervallo in cui la stiamo considerando, anche $\arcsin x$ è crescente.

Allo stesso modo si sarebbe potuto decidere di invertire la restrizione della funzione $\sin x$ ad un altro intervallo, ad esempio in $[\frac{3\pi}{2}, \frac{5\pi}{2}]$. La scelta dell'intervallo $[-\frac{\pi}{2}, \frac{\pi}{2}]$ è puramente convenzionale, per questo, qualche volta, si dice che $\arcsin x$ è il *valore principale dell'arcoseno*.

Funzione arcocoseno. In modo analogo, considerando la restrizione della funzione $\cos x$ all'intervallo $[0, \pi]$ e osservando che tale funzione è decrescente, è possibile definire la sua funzione inversa

$$\arccos : [-1, 1] \longrightarrow [0, \pi]$$

detta **(funzione) arcocoseno** (o anche *valore principale dell'arcocoseno*). Dalla monotonia di $\cos x$ in $[0, \pi]$ discende che la funzione $\arccos x$ è decrescente nel suo insieme di definizione.

Funzione arcotangente. La funzione $\tan x$ ristretta all'intervallo $(-\frac{\pi}{2}, \frac{\pi}{2})$ è crescente e quindi invertibile, con inversa crescente. La sua inversa si indica con $\arctan x$ ed è detta **(funzione) arcotangente**:

$$\arctan : \mathbb{R} \longrightarrow \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Anche per l'arcotangente (come per arcoseno e arcocoseno) si sarebbe potuto decidere di invertire $\tan x$ in un altro intervallo.

Esponenziali e logaritmi (costruzione naïf). Oltre alle funzioni elementari sono importanti le funzioni esponenziali con base $a > 0$ e le loro inverse, i logaritmi in base $a > 0$:

$$y = a^x \quad \text{e} \quad y = \log_a x.$$

Diamo qui solo una definizione “leggera” (non rigorosa) delle funzioni esponenziali. Fissiamo $a > 0$, se $x = p/q \in \mathbb{Q}$ è possibile definire a^x

$$\text{fissato } a > 0, \quad a^x := \sqrt[q]{a^p} \quad \forall x = \frac{p}{q} \in \mathbb{Q},$$

dove la radice (come sempre) è scelta positiva. Per definire il valore a^x anche nel caso in cui x sia un numero irrazionale, è naturale approssimare $x \in \mathbb{R} \setminus \mathbb{Q}$ con numeri razionali sempre più vicini.

Per le funzioni esponenziali valgono le proprietà

- (i) $a^x > 0 \quad \forall x \in \mathbb{R}$,
- (ii) $a^0 = 1$,
- (iii) $a^x a^y = a^{x+y}$,
- (iv) $(a^x)^y = a^{xy}$
- (v) a^x è $\begin{cases} \text{decrescente} & \text{se } 0 < a < 1 \\ \text{crescente} & \text{se } a > 1. \end{cases}$

Dato che a^x è monotona per $a \neq 1$, essa è invertibile. La funzione inversa è $y = \log_a x$: è definita sull'insieme immagine dell'esponenziale, quindi solo per $x > 0$ ed associa ad x l'unico valore y che verifica $a^y = x$. Per i logaritmi valgono le proprietà

- (i) $\log_a x$ è definito per $x > 0$,
- (ii) $\log_a 1 = 0$,
- (iii) $\log_a x + \log_a y = \log_a(xy)$,
- (iv) $\log_a(x^\alpha) = \alpha \log_a x$
- (v) $\log_a x$ è $\begin{cases} \text{decrescente} & \text{se } 0 < a < 1 \\ \text{crescente} & \text{se } a > 1 \end{cases}$

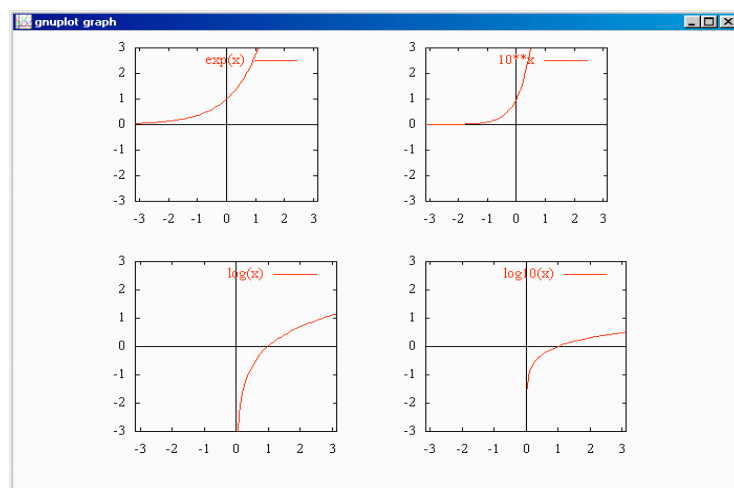


FIGURA 11. (a) $y = e^x$, (b) $y = 10^x$, (c) $y = \ln x$, (d) $y = \log_{10} x$

OSSERVAZIONE 4.1. Perché non si può definire “ a elevato ad x ” per a negativi? In effetti, se $a < 0$, l’espressione a^x ha senso per x del tipo p/q con p e q interi e q dispari. Quindi si potrebbe sperare di estendere questa strana funzione in modo da definire a^x anche per tutti gli altri valori di $x \in \mathbb{R}$. Ma non esiste una estensione che preservi le belle proprietà che abbiamo elencato per l’esponenziale. Ecco un esempio: supponiamo per un attimo che a^x sia ben definita per ogni a e per ogni x , allora $-1 = (-1)^1 = (-1)^{2 \cdot \frac{1}{2}} = [(-1)^2]^{\frac{1}{2}} = 1^{\frac{1}{2}} = 1$. Ops!

Funzioni composte. Si possono generare funzioni anche con la *composizione di funzioni*: se $\phi : \mathbb{R} \rightarrow \mathbb{R}$ e $g : \mathbb{R} \rightarrow \mathbb{R}$, allora la formula

$$f(x) := g(\phi(x))$$

definisce una funzione $f : \mathbb{R} \rightarrow \mathbb{R}$. Ad esempio,

$$f(x) = \sin(1 + x^2) = g(\phi(x)) \quad \text{dove} \quad \begin{cases} \phi(x) = 1 + x^2, \\ g(u) = \sin u. \end{cases}$$

Analogamente

$$f(x) = 2^{\cos x} = g(\phi(x)) \quad \text{dove} \quad \begin{cases} \phi(x) = \cos x, \\ g(u) = 2^u. \end{cases}$$

Spesso la funzione composta $f(x) = g(\phi(x))$ si indica con $f = g \circ \phi$ (che si legge “ g composto ϕ ” o anche “ g dopo ϕ ”)².

La composizione di funzioni non è un’operazione commutativa: in generale, $g \circ \phi$ e $\phi \circ g$ non sono la stessa funzione. L’ordine con cui si fanno le operazioni è importante! Se, per esempio, l’operazione ϕ sta per “sommare 1” e g per “moltiplicare per 2”, allora

$$g(\phi(x)) = g(x + 1) = 2x + 2, \quad \phi(g(x)) = \phi(2x) = 2x + 1.$$

Nel contesto delle funzioni composte, la nozione di funzione inversa diviene ancora più chiara. La funzione $Id(x) = x$, si chiama **identità** (è una funzione affine che ha per grafico la bisettrice del primo e terzo quadrante). Se la funzione ϕ è iniettiva su \mathbb{R} e la sua funzione inversa ϕ^{-1} è anch’essa definita su tutto \mathbb{R} , ϕ^{-1} è l’unica funzione che gode delle proprietà $\phi^{-1} \circ \phi = Id$ e $\phi \circ \phi^{-1} = Id$, cioè

$$(\phi^{-1} \circ \phi)(x) = x \quad (\phi \circ \phi^{-1})(x) = x \quad \forall x \in \mathbb{R},$$

La composizione ha senso anche per funzioni non definite in tutto \mathbb{R} . Sia $\phi : I \subset \mathbb{R} \rightarrow \mathbb{R}$ e $g : J \subset \mathbb{R} \rightarrow \mathbb{R}$, allora $f(x) := g(\phi(x))$ è ben definita per ogni $x \in I$ per cui $\phi(x) \in J$. Ad esempio, la funzione $f(x) = \ln(1 + x)$ è la composizione della funzione $\phi(x) = 1 + x$, definita in $I = \mathbb{R}$, e della funzione $g(u) = \ln u$, definita in $J = (0, +\infty)$.

²Alle volte si omette il simbolo \circ e si scrive semplicemente $f = g\phi$, ma bisogna stare attenti a non fare confusione con la funzione prodotto.

La composizione f è definita per tutte i valori $x \in \mathbb{R}$ tali che $1 + x \in J$, cioè per $x \in (-1, +\infty)$.

Per poter comporre due funzioni ϕ e g e definire una nuova funzione $g \circ \phi$, il dominio di g deve includere almeno una parte dell'immagine di ϕ . Ad esempio, non possiamo formare la funzione $g \circ \phi$ quando $g(u) = \sqrt{u}$ e $\phi(x) = -1 - x^2$, dato che il dominio di g è $[0, +\infty)$ e l'immagine di ϕ è $(-\infty, -1)$.

Chiaramente è possibile comporre più di due funzioni. Ad esempio,

$$f(x) = \frac{1}{1 + \tan(x^2)}$$

può essere ottenuta componendo (nell'ordine) $\phi(x) = x^2$, $\psi(\phi) = 1 + \tan \phi$, $g(\psi) = \frac{1}{\psi}$. Quindi $f = g \circ \psi \circ \phi$.

Funzioni meno comuni. Esistono infiniti modi per definire funzioni. Una possibilità (vagamente esotica) è di decomporre l'insieme di definizione in un certo numero di sottoinsiemi disgiunti ed associare una opportuna regola di calcolo per ciascuno di tali sottoinsiemi.

(i) Sia $I \subset \mathbb{R}$, allora si definisce

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

La funzione $\chi_{\mathbb{R}}$ vale identicamente 1, mentre χ_{\emptyset} vale sempre 0, la funzione $\chi_{[0,1]}$ vale 1 nell'intervallo $[0, 1]$ e 0 nel complementare.

(ii) E' possibile anche fare scelte più originali: ad esempio,

$$f(x) = \begin{cases} 3x + 1 & \text{se } x \leq 0, \\ x^2 & \text{se } x > 0. \end{cases}$$

(iii) Un altro modo per generare nuove funzioni è tramite i "comandi" max e min. Ad esempio, si è già visto che

$$\max\{-x, x\} = |x|.$$

Più in generale, date f e g ,

$$\begin{aligned} \max\{f, g\}(x) &= \begin{cases} f(x) & \text{se } g(x) \leq f(x), \\ g(x) & \text{se } f(x) \leq g(x) \end{cases} \\ \min\{f, g\}(x) &= \begin{cases} f(x) & \text{se } f(x) \leq g(x), \\ g(x) & \text{se } g(x) \leq f(x) \end{cases} \end{aligned}$$

5. Problemi di massimo e minimo

Molti problemi pratici conducono a problemi di massimo e di minimo di funzioni: qual è il carico massimo sopportato da una trave? Qual è l'energia minima che occorre perchè un satellite sfugga dall'attrazione gravitazionale di un pianeta? Qual è il minimo sforzo che bisogna compiere per passare l'esame? Diamo perciò una definizione rigorosa di cosa si intenda per massimo e minimo di una funzione.

DEFINIZIONE 5.1. *Sia $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$. Se esiste $x_0 \in I$ tale che $f(x) \geq f(x_0)$ per ogni $x \in I$, la funzione ammette (valore) minimo in I e x_0 è un punto di minimo. Si scrive*

$$f(x_0) = \min_{x \in I} f(x) \quad (\text{valore) minimo di } f.$$

Analogamente se esiste un punto $x_1 \in I$ tale che $f(x) \leq f(x_1)$ per ogni $x \in I$, si dice che la funzione ammette (valore) massimo in I e x_1 è un punto di massimo. Si scrive

$$f(x_1) = \max_{x \in I} f(x) \quad (\text{valore) massimo di } f.$$

Il massimo ed il minimo dipendono dall'insieme di definizione I . In generale una restrizione di una funzione, se ammette massimo, ha un massimo minore o uguale a quello della funzione di partenza e, se ammette minimo, questo è maggiore o uguale di quello della funzione di partenza.

In effetti il massimo ed il minimo della funzione non sono altro che il massimo ed il minimo dell'insieme immagine $f(I)$:

$$\begin{aligned} \min_{x \in I} f(x) &= \min f(I) = \min\{f(x) : x \in I\}, \\ \max_{x \in I} f(x) &= \max f(I) = \max\{f(x) : x \in I\}. \end{aligned}$$

Come si fa a determinare il massimo o il minimo di una funzione di una variabile reale a partire dal grafico? Il significato geometrico di un punto di massimo è chiaro: il grafico della funzione f è al di sotto della retta di equazione $y = f(x_0) = \text{costante}$. Quindi per determinare il massimo a partire dal grafico, basta stabilire se esista una retta con tale proprietà.

Come si vede a partire da alcuni esempi, *non tutte le funzioni ammettono massimo e/o minimo nel loro insieme di definizione!* Per superare questo ostacolo si introducono i concetti di *estremo superiore* e di *estremo inferiore*. L'estremo superiore ed inferiore di una funzione f sono l'estremo superiore ed inferiore dell'insieme immagine $f(I)$

$$\begin{aligned} \inf_{x \in I} f(x) &= \inf f(I) = \inf\{f(x) : x \in I\}, \\ \sup_{x \in I} f(x) &= \sup f(I) = \sup\{f(x) : x \in I\}. \end{aligned}$$

Il risultato sull'esistenza di estremo superiore ed inferiore garantisce che se esiste almeno un maggiorante allora esiste l'estremo superiore. Quindi, dato il grafico della funzione f ci sono solo due possibilità: o esiste almeno una retta orizzontale di equazione $y = c \in \mathbb{R}$ che sia completamente sopra il grafico di f o non ne esiste nessuna. Nel primo caso, l'estremo superiore di f è il valore minimo che si può dare al valore c facendo sempre in modo che la retta $y = c$ sia sopra il grafico di f (tale retta può anche non intersecare il grafico). Nel secondo caso, la funzione f è **illimitata superiormente** e $\sup_I f = +\infty$. Analogamente per l'estremo inferiore.

DEFINIZIONE 5.2. Una funzione $f : I \subset \mathbb{R} \rightarrow \mathbb{R}$ tale che $\inf_{x \in I} f(x) \in \mathbb{R}$ e $\sup_{x \in I} f(x) \in \mathbb{R}$ (quindi non sono $\pm\infty$) si dice **limitata**.

Il significato geometrico della limitatezza di una funzione è immediato: una funzione è limitata se e solo se il suo grafico è interamente contenuto in una striscia orizzontale $\{(x, y) \in \mathbb{R}^2 : c \leq y \leq d\}$ per qualche $c, d \in \mathbb{R}$.

La lattina più conveniente. Supponiamo che si voglia progettare una scatola di latta di forma cilindrica (altezza h e raggio di base r). Il problema è: fissato il volume della scatola, esiste una scelta di r e h che minimizzi il quantitativo di latta da utilizzare (cioè la superficie totale del cilindro)? Sia V il volume della scatola e S la sua superficie

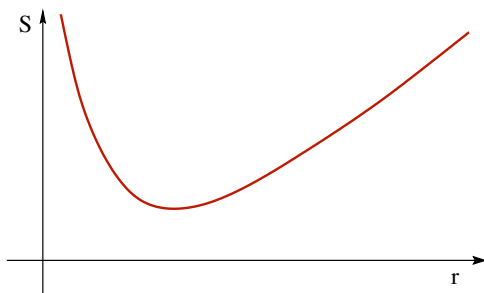


FIGURA 12. Il grafico di $S(r) = 2\pi \left(r^2 + \frac{V}{\pi r} \right)$.

totale. Allora $S = 2\pi r^2 + 2\pi hr = 2\pi(r^2 + hr)$ e $V = \pi r^2 h$. Fissare il volume V e minimizzare S , vuol dire scegliere $V = \text{costante}$. Quindi $h = V/\pi r^2$, e, sostituendo in S ,

$$h = \frac{V}{\pi r^2} \quad \Rightarrow \quad S = 2\pi \left(r^2 + \frac{V}{\pi r} \right).$$

Dal grafico della funzione S (che si può ottenere sommando i grafici $s = r^2$ e $s = V/\pi r$) si vede che tale minimo esiste. Il calcolo esplicito di quanto valga non è possibile per via elementare (ma con un minimo di cognizione di derivate, si può fare!).

CAPITOLO 3

Incontri ravvicinati con i limiti: le successioni

Una funzione $a : \mathbb{N} \rightarrow \mathbb{R}$ che associa ad ogni numero naturale n un valore $a(n)$ è una **successione (numerica)**. In genere, l' n -esimo elemento della successione si indica con a_n (invece di $a(n)$), questione di tradizione. Gli elementi della successione a_n possono essere pensati come una sequenza di valori ordinati in base al loro indice n

$$a_0, a_1, a_2, a_3, \dots$$

Un primo esempio è la successione dei numeri pari: $2, 4, 6, \dots$. In questo caso $a_n = 2n$. Un altro esempio semplice di funzione di n è l'espressione **n -fattoriale**, definita dal prodotto dei primi n numeri interi

$$a_n = n! = 1 \cdot 2 \cdot 3 \cdots n,$$

che dà luogo alla successione $1, 1, 2, 6, 24, 120, \dots$ (per definizione $0! := 1$).

Una successione può essere rappresentata disegnando nel piano cartesiano sopra ogni valore $n \in \mathbb{N}$ (dell'asse x) il valore definito da a_n , proprio come nel caso delle funzioni. Questo primo metodo è molto pratico nel caso di successioni definite da $a_n = f(n)$, dove si conosca il grafico della funzione f : basta prendere sul grafico di f solamente i punti con coordinata $x \in \mathbb{N}$. In alternativa, assegnata la successione a_n si può considerare come sua rappresentazione il grafico della funzione g definita da

$$g(x) = a_n \quad \text{per } x \in [n, n+1).$$

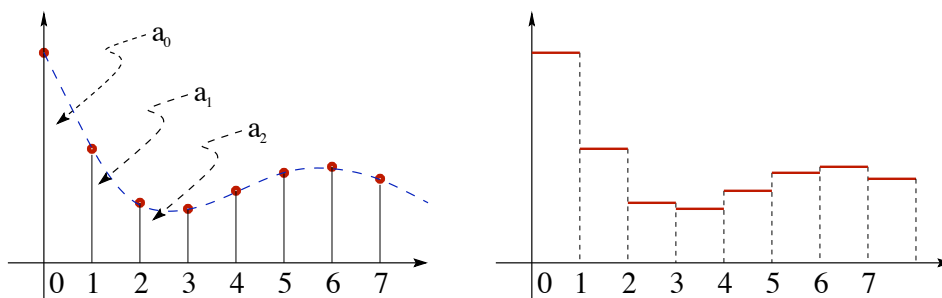


FIGURA 1.

Successioni definite per ricorrenza. In quel che segue utilizzeremo le successioni numeriche come una “cavia da laboratorio” per imparare, in una situazione particolarmente semplice, il concetto di limite e le procedure di base di calcolo. In realtà le successioni numeriche possono emergere anche da semplici modelli applicati. Supponiamo di voler studiare una popolazione di individui e di indicare con a_n il numero di abitanti all’anno n . Per controllare l’evoluzione della popolazione occorre conoscere il *tasso di incremento* R , definito da

$$R = \frac{a_{n+1} - a_n}{a_n},$$

che descrive quanto valga l’aumento di popolazione $a_{n+1} - a_n$ rispetto alla popolazione a_n all’anno n . Se si suppone che il tasso di crescita R sia costante ed uguale ad $r \in \mathbb{R}$, si ottiene, $a_{n+1} - a_n = ra_n$, cioè, esplicitando rispetto ad a_{n+1} ,

$$a_{n+1} = (1 + r)a_n.$$

Quindi, se si assegna la popolazione a_0 all’anno iniziale, si deduce che $a_1 = (1 + r)a_0$, $a_2 = (1 + r)a_1 = (1 + r)^2 a_0$. In generale questo semplice modello dà luogo ad una popolazione che cresce esponenzialmente in n , infatti

$$a_{n+1} = (1 + r)a_n = (1 + r)^2 a_{n-1} = \cdots = (1 + r)^{n+1} a_0.$$

Ad esempio, se si parte da una popolazione di 100 abitanti e si suppone che il tasso di crescita annuale sia del 10%, cioè $r = 0,1$, dopo dieci anni la popolazione sarà di $a_{10} = 1,1^{10} \times 100$ abitanti (circa 259). Se si sceglie il tasso di incremento della forma $R = r(N - a_n)$ (questo vuol dire che c’è una popolazione critica, in questo caso pari a N , tale che se $a_n > N$ la popolazione decresce, mentre se $a_n < N$ la popolazione aumenta), si ottiene

$$a_{n+1} = (1 + Nr)a_n - ra_n^2 \quad (\text{equazione logistica}).$$

Già un oggetto così semplice e apparentemente innocuo è in grado di generare (per scelte opportune del parametro r) dinamiche particolarmente “stravaganti” e molto interessanti.

L’esempio precedente rientra nella classe delle successioni *definite per ricorrenza*: il termine $(n + 1)$ -esimo si ottiene in funzione dei termini precedenti. Nella forma più semplice, il termine $(n + 1)$ -esimo è determinato dal solo termine n -esimo: assegnata la funzione f (dipende dal modello), si pone $a_{n+1} = f(a_n)$. Come nell’esempio precedente occorre anche assegnare una *condizione iniziale*, cioè deve essere dato il valore iniziale a_0 .

1. Limite di successioni

Il concetto fondamentale su cui si basa l'analisi matematica è quello di **limite**. L'idea che esprime il concetto di limite di una successione è semplice: assegnata la successione a_n , siamo in grado di “prevedere” quello che succederà per valori di n molto grandi? Più precisamente: è vero che la successione a_n “si stabilizza” per $n \rightarrow +\infty$, ovvero tende ad avvicinarsi ad un valore ℓ fissato? In caso affermativo, si dice che la successione ammette limite ℓ per $n \rightarrow +\infty$, altrimenti si dice che la successione non ha limite. Molte parole che abbiamo scritto nelle righe precedenti vanno precisate: che vuol dire “avvicinarsi”? E mandare n a $+\infty$ è da considerarsi un terribile insulto?

Partiamo da alcuni esempi. Sia

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

cioè consideriamo la successione $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \dots$. Nessun numero di questa successione è zero, ma, per n che cresce, a_n si avvicina a zero. La frase “si avvicina a zero” va interpretata in questo senso: se decidiamo che l'essere vicino vuol dire che la distanza tra a_n e 0 deve essere minore di $1/10$, allora basta considerare gli elementi a_n della successione con indice $n \geq 10$; se rendiamo la condizione più stringente, ad esempio richiedendo che la distanza sia minore di $1/100$, basta considerare $n \geq 100$, e così via. In generale, comunque fissiamo una distanza $\varepsilon > 0$, da un certo indice n_ε in poi (n_ε dipende da ε) la distanza di a_n da 0 (che è data da $|a_n - 0|$) è minore di ε , cioè

$$(7) \quad \forall \varepsilon > 0 \quad \exists n_\varepsilon \in \mathbb{N} \quad \text{t.c.} \quad |a_n - 0| < \varepsilon \quad \forall n > n_\varepsilon.$$

In questo caso si dice che a_n **tende a 0** per $n \rightarrow +\infty$ (che si legge “ n tende a $+\infty$ ”). Per la successione $b_n = \frac{(-1)^n}{n+1}$ la situazione è esattamente la stessa, dato che

$$|b_n - 0| = \left| \frac{(-1)^n}{n+1} - 0 \right| = \left| \frac{1}{n+1} - 0 \right| = |a_n - 0|.$$

L'unica differenza è che i numeri b_n sono alternativamente più grandi e più piccoli di zero, cioè la successione *oscilla* attorno al valore limite 0, ma anche in questo caso vale la proprietà (7).

Consideriamo $a_n = \frac{n}{n+1}$. Scrivendo la successione nella forma

$$a_n = \frac{n}{n+1} = 1 - \frac{1}{n+1} \quad \Rightarrow \quad |a_n - 1| = \frac{1}{n+1},$$

vediamo che, per $n \rightarrow +\infty$, la distanza di a_n da 1 tende a zero, cioè il valore a_n si avvicina ad 1. Anche la successione $a_n = \frac{n^2 - 1}{n^2 + n + 1}$ si comporta in modo analogo,

infatti:

$$a_n = 1 - \frac{n+2}{n^2+n+1} \Rightarrow |a_n - 1| = \frac{n+2}{n^2+n+1} \leq \frac{2n}{n^2} = \frac{2}{n} \quad \forall n \geq 2,$$

da cui si deduce che la distanza di a_n da 1 tende a zero per $n \rightarrow \infty$.

DEFINIZIONE 1.1. Limite di successione. Si dice che la successione a_n converge ad $\ell \in \mathbb{R}$ per $n \rightarrow +\infty$ e si scrive $a_n \rightarrow \ell$ per $n \rightarrow +\infty$ o $\lim_{n \rightarrow +\infty} a_n = \ell$, se

$$(8) \quad \forall \varepsilon > 0 \quad \exists n_\varepsilon \in \mathbb{N} \quad t.c. \quad |a_n - \ell| < \varepsilon \quad \forall n > n_\varepsilon.$$

Il valore ℓ è il limite della successione a_n .

Se $a_n \rightarrow 0$ per $n \rightarrow +\infty$ (cioè se vale (7)), si dice che a_n è infinitesima.

La definizione esprime che, comunque si fissi una soglia di errore $\varepsilon > 0$, tutti gli elementi a_n della successione distano dal limite ℓ meno di ε , tranne al più un numero finito (quelli con indice da 1 ad n_ε). Quindi una maniera equivalente di dire che ℓ è il limite di a_n è affermare che *ogni intorno di ℓ contiene tutti i valori della successione a_n tranne al più un numero finito*.

Si faccia bene attenzione al fatto che la soglia ε vive sull'asse delle ordinate (e non su quello delle ascisse). In generale, scegliendo valori più piccoli per il margine di errore ε occorre scegliere valori più grandi di n_ε ; in altre parole, in generale, n_ε cresce quando ε tende a zero.

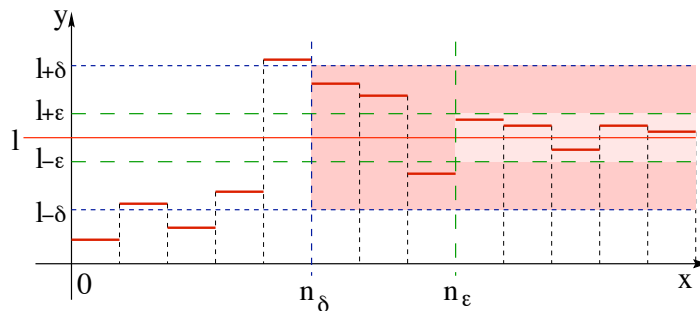


FIGURA 2.

OSSERVAZIONE 1.2. Per quale motivo si richiede che ε possa essere scelto arbitrariamente? Non basterebbe scegliere un fissato ε sufficientemente piccolo, ad esempio ε pari a un miliardesimo o a un miliardesimo di miliardesimo? Il problema è che i concetti di “grande” e “piccolo” sono soggettivi, mentre quello che si vuole definire qui è un criterio assoluto che vada bene sia per l’astronomo che usa la distanza Terra-Luna come parametro di “vicinanza”, sia per il fisico atomico per cui un millimetro è già una distanza abissale. La richiesta di una proprietà che valga *per ogni scelta di ε* rende la definizione “universale”, cioè indipendente dalla personale idea di piccolo o grande.

ESERCIZIO 1.3. Sia a_n una successione convergente ad ℓ per $n \rightarrow +\infty$ e sia b_n un'altra successione tale che $b_n = a_n$ per ogni $n > N_A$ dove N_A è il numero di Avogadro¹. Dimostrare che anche b_n converge ad ℓ per $n \rightarrow +\infty$.

Soluzione. Niente di più facile dato che la definizione di limite non dipende dal comportamento di un numero finito di elementi della successione. Per ipotesi,

$$\forall \varepsilon > 0 \quad \exists n_\varepsilon \in \mathbb{N} \quad \text{t.c.} \quad |a_n - \ell| < \varepsilon \quad \forall n > n_\varepsilon.$$

Ora vogliamo far vedere che vale una frase analoga anche per la successione b_n . Fissato $\varepsilon > 0$, scegliamo $n'_\varepsilon := \max\{n_\varepsilon, N_A\}$. Allora, se $n > n'_\varepsilon$, dato che $n > N_A$ si ha $b_n = a_n$ e dato che $n > n_\varepsilon$ vale anche $|a_n - \ell| < \varepsilon$. Dunque

$$|b_n - \ell| = |a_n - \ell| < \varepsilon \quad \forall n > n'_\varepsilon := \max\{n_\varepsilon, N_A\},$$

cioè la conclusione. E' essenziale che N_A sia proprio il numero di Avogadro o lo stesso ragionamento vale per N_A qualsiasi?

Calcolo diretto di un limite. Abbiamo una perfetta definizione di limite: logicamente ineccepibile. Ma come fare per verificarne la validità in un caso concreto? Proviamo a vedere un esempio. Tenete però conto che, nella pratica, non è questo il modo con cui si calcolano la maggior parte dei limiti! L'esempio che segue serve solo per acquisire maggiore familiarità con la definizione.

Consideriamo la successione

$$a_n = \frac{n^2}{n^2 + 1}.$$

Ammette limite? Ecco subito il primo problema: nella definizione di limite compare il valore ℓ del limite stesso, ma in generale ci si trova ad avere un'espressione per la successione, non per il suo (eventuale) limite. Questo va ottenuto per un'altra strada. Proviamo a ragionare in maniera casereccia. La domanda di fondo è: cosa succede dei valori a_n per n molto grande? Ad esempio, se $n = 1000$,

$$a_{1000} = \frac{1000^2}{1000^2 + 1} = \frac{1000000}{1000001}.$$

Bene... e se $n = 100000$? Allora

$$a_{100000} = \frac{100000^2}{100000^2 + 1} = \frac{10000000000}{10000000001}.$$

Come si vede, per valori di n molto grandi il termine $+1$ a denominatore diventa sempre più ridicolo perché va a sommarsi ad una quantità enormemente più grande. Allora è sensato aspettarsi che per $n \rightarrow +\infty$ valga un'approssimazione del tipo $n^2 + 1 \approx n^2$ e quindi $a_n = \frac{n^2}{n^2 + 1} \approx \frac{n^2}{n^2} = 1$. Questo per ora non dimostra un bel nulla, ma fa sospettare che la successione abbia limite e che il suo limite sia $\ell = 1$. Rimbocchiamoci le maniche

¹Per chi non lo ricorda il numero di Avogadro è $N_A = 6,02214199 \times 10^{23}$.

e proviamo a dimostrarlo. Qual'è l'affermazione racchiusa nella definizione di limite? la distanza di a_n da ℓ è piccola se n è grande. Prima di tutto, quindi, scriviamo $|a_n - \ell|$:

$$|a_n - \ell| = \left| \frac{n^2}{n^2 + 1} - 1 \right| = \left| \frac{n^2 - n^2 - 1}{n^2 + 1} \right| = \frac{1}{n^2 + 1}.$$

Ora si tratta di far vedere che, fissato $\varepsilon > 0$, questa distanza è minore di ε se scegliamo $n > n_\varepsilon$ dove *abbiamo completa libertà di scelta per n_ε* . Imponiamo la disequazione a cui vogliamo arrivare e riscriviamola come condizione su n :

$$\frac{1}{n^2 + 1} < \varepsilon \quad \Longleftrightarrow \quad n^2 > \frac{1}{\varepsilon} - 1 \quad \Longleftrightarrow^{\varepsilon < 1} \quad n > \sqrt{\frac{1}{\varepsilon} - 1}.$$

Il gioco è fatto, basta scegliere $n_\varepsilon \geq \sqrt{\frac{1}{\varepsilon} - 1}$, ad esempio,

$$n_\varepsilon := \left[\sqrt{\frac{1}{\varepsilon} - 1} \right] + 1$$

dove $[\cdot]$ indica la funzione parte intera.

ESERCIZIO 1.4. *Dimostrare a partire dalla definizione la validità di*

$$\lim_{n \rightarrow +\infty} \frac{n}{2n + 1} = \frac{1}{2}, \quad \lim_{n \rightarrow +\infty} \frac{1}{n^k} = 0 \quad k \in \mathbb{N}, k \neq 0.$$

ESERCIZIO 1.5. *Dimostrare che, se $\lim_{n \rightarrow +\infty} a_n = a$, allora $\lim_{n \rightarrow +\infty} |a_n| = |a|$.*

Come dimostrare che una successione non ha limite? Data una successione a_n , una **sottosuccessione** a_{n_k} di a_n si ottiene scegliendo un sottoinsieme infinito degli elementi di a_n , scelti in modo che ogni elemento abbia indice strettamente maggiore di quello del precedente. Ad esempio, gli elementi di indice dispari $a_1, a_3, a_5, a_7, \dots$ costituiscono una sottosuccessione di $a_1, a_2, a_3, a_4, \dots$, così come gli elementi di indice pari $a_0, a_2, a_4, a_6, \dots$. Invece $a_5, a_3, a_{12}, a_{101}, \dots$ non è una sottosuccessione, perché il primo elemento ha indice maggiore del secondo.

Come individuare una sottosuccessione? Bisogna indicare il primo elemento, poi il secondo, quindi il terzo e così via. In definitiva bisogna scegliere un'applicazione da \mathbb{N} ad \mathbb{N} che ci dica al k -esimo posto quale l'elemento n_k della successione scelto: al numero $k \in \mathbb{N}$ viene quindi associato un numero naturale n_k . Per fare in modo che l'ordine degli elementi venga preservato occorre che n_k sia crescente, cioè se $k_1 < k_2$ allora $n_{k_1} < n_{k_2}$. La sottosuccessione dei termini di indice dispari è espressa da $n_k = 2k + 1$: per $k = 0$ si prende $n_k = 1$, per $k = 1$ si prende $n_k = 3$ e così via...

ESERCIZIO 1.6. *Dire quale delle seguenti espressioni possono essere i primi elementi di una sottosuccessione di $a_n = n$*

$$\{1, 1, 3, 5, 7, \dots\}, \quad \{1, 4, 9, 16, 25, \dots\}, \quad \{1, 3, 5, 7, 6, 9, \dots\}.$$

ESERCIZIO 1.7. *Data la successione $a_n = n$, quali sono i primi termini della sottosuccessione a_{k^2} ? E se $k_n = 2n$? Ripetere l'esercizio nel caso in cui $a_n = n^2 + 1$.*

PROPOSIZIONE 1.8. *Sia a_n una successione convergente ad ℓ per $n \rightarrow +\infty$. Allora ogni sua sottosuccessione a_{n_k} converge allo stesso limite ℓ per $k \rightarrow +\infty$.*

La dimostrazione è lasciata per esercizio.

La Proposizione 1.8 può essere utilizzata “in negativo” per dimostrare che una assegnata successione non ammette limite. Consideriamo ad esempio la successione $a_n = (-1)^n$. La sottosuccessione dei termini di indice pari è $a_{2k} = (-1)^{2k} = 1$ per ogni k , quindi, essendo costantemente uguale ad 1, converge ad 1 per $k \rightarrow +\infty$. Invece, la sottosuccessione dei termini di indice dispari è $a_{2k+1} = (-1)^{2k+1} = -1$ per ogni k , quindi converge a -1 per $k \rightarrow +\infty$. Dato che due sottosuccessioni diverse convergono a limiti diversi, la conclusione della Proposizione 1.8 e quindi l'ipotesi non può essere vera: la successione $(-1)^n$ non è convergente. In generale, *se da una successione possono essere estratte due sottosuccessioni convergenti a limiti diversi, la successione non è convergente.*

Prime proprietà delle successioni convergenti. Prima di enunciare alcuni risultati che permettono di calcolare limiti in maniera più semplice di come si è fatto finora, dimostriamo alcune proprietà generali delle successioni convergenti.

PROPOSIZIONE 1.9. *Se una successione è convergente, allora il suo limite è unico.*

DIMOSTRAZIONE. Mostriamo che se la successione a_n tende sia ad ℓ che ad ℓ' , allora deve essere $\ell = \ell'$. Per definizione di limite, è vero che, per ogni $\varepsilon > 0$,

$$\exists n_\varepsilon \in \mathbb{N} \quad \text{t.c.} \quad |a_n - \ell| < \varepsilon \quad \forall n > n_\varepsilon \quad \text{e} \quad \exists n'_\varepsilon \in \mathbb{N} \quad \text{t.c.} \quad |a_n - \ell'| < \varepsilon \quad \forall n > n'_\varepsilon.$$

Allora, per $n > \max\{n_\varepsilon, n'_\varepsilon\}$, sono vere entrambe le affermazioni e quindi

$$0 \leq |\ell - \ell'| = |\ell - a_n + a_n - \ell'| \leq |\ell - a_n| + |a_n - \ell'| < 2\varepsilon.$$

In definitiva, abbiamo dimostrato che, per ogni $\varepsilon > 0$, vale $0 < |\ell - \ell'| < 2\varepsilon$. Quindi $|\ell - \ell'| = 0$, cioè $\ell = \ell'$. \square

DEFINIZIONE 1.10. *Una successione a_n tale che esista $M > 0$ per cui $|a_n| \leq M$ per ogni n , (cioè tutti gli a_n appartengono all'intervallo $[-M, M]$) si dice **limitata**.*

Se si ricorda che una successione a_n è una funzione da \mathbb{N} a \mathbb{R} , la condizione espressa nella Definizione 1.10 è equivalente alla frase “ $a(\mathbb{N})$ è un sottoinsieme limitato di \mathbb{R} ”, che è proprio la definizione di limitatezza data per funzioni di una variabile.

PROPOSIZIONE 1.11. *Se a_n è una successione convergente allora è anche limitata.*

DIMOSTRAZIONE. Fissato $\varepsilon = 1$, dato che a_n è convergente, esiste $N \in \mathbb{N}$ tale che $\ell - 1 < a_n < \ell + 1$ per ogni $n > N$. Quindi la “coda” della successione a_{N+1}, a_{N+2}, \dots vive in un insieme limitato. Dato che i termini mancanti sono in numero finito, tutta la successione “vive” in un insieme limitato (se necessario più grande del precedente). \square

Il viceversa non è vero: *esistono successioni limitate, non convergenti*. Ad esempio, la successione $a_n = (-1)^n$ non è convergente, ma è limitata dato che $|(-1)^n| = 1$ per ogni n (quindi si può scegliere $M = 1$ nella Definizione 1.10).

Successioni divergenti. Oltre alle successioni che tendono ad un limite, ci sono anche quelle il cui valore a_n diventa arbitrariamente grande: ad esempio la successione dei numeri pari $2n$, o la successione del fattoriale $n!$. La prossima definizione esprime cosa significhi in modo preciso l’affermazione “una successione tende a $+\infty$ ”.

DEFINIZIONE 1.12. *La successione a_n diverge a $+\infty$ (rispettivamente a $-\infty$) per $n \rightarrow +\infty$ se, comunque si scelga un valore M tutti i valori a_n sono più grandi di M (risp. più piccoli) tranne al più un numero finito. In tal caso si scrive*

$$\lim_{n \rightarrow +\infty} a_n = +\infty \quad (\text{risp. } -\infty).$$

In modo equivalente si può scrivere

$$(9) \quad \forall M \quad \exists n_M \in \mathbb{N} \quad \text{t.c.} \quad a_n \geq M \quad (\text{risp. } a_n \leq M) \quad \forall n > n_M.$$

Come nel caso delle successioni divergenti è utile vedere almeno un esempio di verifica diretta del fatto che una successione è divergente. Ad esempio, dimostriamo

$$\lim_{n \rightarrow +\infty} \frac{n^2}{n+1} = +\infty.$$

Fissato $M > 0$, dobbiamo far vedere che è possibile determinare n_M per cui valga $a_n > M$ per ogni $n > n_M$. Dato che

$$\frac{n^2}{n+1} > M \quad \Longleftrightarrow \quad n^2 - Mn - M > 0,$$

la domanda da porsi è: per quali n è vera la disequazione finale? Le radici del polinomio di secondo grado $x^2 - Mx - M$ sono $x_{\pm} := (M \pm \sqrt{M^2 + 4M})/2$, quindi, se $n > x_+$, è vero che $n^2 - Mn - M > 0$. Ottimo, allora scegliamo

$$N_M := \left\lceil \frac{M + \sqrt{M^2 + 4M}}{2} \right\rceil + 1.$$

Strade diverse portano alla stessa conclusione. Lo stesso problema può essere risolto in un modo diverso, meno contoso. Prima di tutto si osserva che (verificate i

passaggi)

$$a_n = \frac{n^2}{n+1} = n - \frac{1}{n+1} \geq n-1.$$

Quindi, se $n-1 > M$, vale anche $a_n > M$. Perciò si può scegliere $N_M = [M+1] + 1$ per dedurre la stessa conclusione. In effetti, il valore n_M della Definizione 1.12 (così come n_ε della Definizione 1.1), non è definito in maniera univoca, tutt'altro!

ESERCIZIO 1.13. *Dimostrare le implicazioni*

$$\begin{aligned} \lim_{n \rightarrow +\infty} a_n = 0 & \quad \Rightarrow \quad \lim_{n \rightarrow +\infty} \frac{1}{|a_n|} = +\infty. \\ \lim_{n \rightarrow +\infty} |a_n| = +\infty & \quad \Rightarrow \quad \lim_{n \rightarrow +\infty} \frac{1}{a_n} = 0. \end{aligned}$$

Criterio di convergenza di Cauchy. Ogni successione convergente definisce un numero ℓ , il suo limite, ma l'unico test di convergenza che emerge dalla definizione consiste nel dimostrare che la differenza $|a_n - \ell|$ è infinitesima, quindi è applicabile solo se il numero ℓ è già noto. Invece, è essenziale avere un test “intrinseco” di convergenza che non richieda la conoscenza *a priori* del valore del limite, ma che coinvolga solamente i termini stessi della successione.

TEOREMA 1.14. Criterio di convergenza di Cauchy. *Una successione a_n è convergente se e solo se è una successione di Cauchy, cioè se vale la condizione*

$$\forall \varepsilon > 0, \quad \exists n_\varepsilon \in \mathbb{N} \quad t.c. \quad |a_m - a_n| < \varepsilon \quad \forall m, n > n_\varepsilon.$$

La condizione di Cauchy descrive il fatto che gli elementi della successione distano tra loro meno di una soglia arbitraria $\varepsilon > 0$ a patto di considerare termini con un indice sufficientemente grande. Il fatto che i termini si avvicinino l'un l'altro per $n \rightarrow +\infty$ quando una successione è convergente è del tutto naturale: i termini si avvicinano al limite ℓ e quindi si avvicinano tra loro. La proprietà notevole è che vale anche il viceversa: se gli elementi della successione si avvicinano, allora la successione converge. Non è questa la sede per approfondire ulteriormente questo criterio, ma non si può mancare di dire che si tratta di una pietra miliare nella costruzione dei numeri reali a partire dai numeri razionali.

Piccolo glossario per le successioni

Se una successione $a_n \dots$

- \dots è convergente o divergente (a $\pm\infty$), allora è **regolare**;
- \dots non è convergente né divergente, allora è **non regolare**;
- \dots tende a 0 per $n \rightarrow +\infty$, allora è **infinitesima**;

... è tale che $|a_n| \leq M$ per qualche M e per ogni n , (cioè se i suoi elementi sono in un intervallo limitato) allora è **limitata**;

... è tale che $\lim_{n \rightarrow +\infty} |a_n| = +\infty$, allora **diverge in modulo**.

$$\begin{array}{ccccc}
 a_n \text{ è infinitesima} & \Rightarrow & a_n \text{ converge} & \Rightarrow & a_n \text{ è limitata} \\
 & & \Downarrow & & \\
 & & a_n \text{ è regolare} & & \\
 & & \Uparrow & & \\
 & & a_n \text{ diverge} & \Rightarrow & a_n \text{ diverge in modulo}
 \end{array}$$

2. Il limite entra in società

Fin qui abbiamo definito il senso della parola *limite* per successioni di numeri reali. Una successione in \mathbb{R} può avere una struttura complicata: ad esempio, può essere somma/prodotto di vari termini. Come si comporta l'operazione di "passaggio al limite" rispetto alle operazioni $+$ e \cdot definite in \mathbb{R} ? E rispetto ai segni \leq e $<$?

Operazioni razionali con i limiti. Per il calcolo dei limiti è possibile usare le operazioni elementari di somma, moltiplicazione, sottrazione e divisione.

(i) *Somma e sottrazione.* Il limite della somma/sottrazione di successioni convergenti è la somma/sottrazione dei limiti:

$$\lim_{n \rightarrow +\infty} a_n = a, \quad \lim_{n \rightarrow +\infty} b_n = b \quad \Rightarrow \quad \lim_{n \rightarrow +\infty} a_n \pm b_n = a \pm b.$$

(ii) *Prodotto.* Il limite del prodotto di successioni convergenti è il prodotto dei limiti:

$$\lim_{n \rightarrow +\infty} a_n = a, \quad \lim_{n \rightarrow +\infty} b_n = b \quad \Rightarrow \quad \lim_{n \rightarrow +\infty} a_n b_n = ab.$$

(iii) *Rapporto.* Il limite del rapporto di successioni convergenti è il rapporto dei limiti *a patto che la successione a denominatore non tenda a zero*:

$$\lim_{n \rightarrow +\infty} a_n = a, \quad \lim_{n \rightarrow +\infty} b_n = b \neq 0 \quad \Rightarrow \quad \lim_{n \rightarrow +\infty} \frac{a_n}{b_n} = \frac{a}{b}.$$

In altre parole: possiamo *invertire l'ordine di applicazione* delle operazioni razionali e del procedimento di limite ottenendo lo stesso risultato, o anche *operazioni razionali e limiti commutano*.

Dimostriamo la proprietà del prodotto. Supponiamo $a_n \rightarrow a$ e $b_n \rightarrow b$ per $n \rightarrow +\infty$. Allora²

$$\forall \varepsilon > 0 \quad \exists n_\varepsilon \quad : \quad |a - a_n| < \varepsilon, \quad |b - b_n| < \varepsilon \quad \forall n > n_\varepsilon.$$

²Il valore di n_ε per la successione $\{a_n\}$ e quello per la successione $\{b_n\}$ potrebbero essere diversi, ma, in questo caso, potremmo scegliere il più grande dei due, per cui sono soddisfatte le relazioni sia per a_n che per b_n .

Scrivendo $ab - a_nb_n = b(a - a_n) + a_n(b - b_n)$, e ricordando che una successione convergente è sempre limitata (per cui esiste $M > 0$ tale che $|a_n| \leq M$ per ogni n), si ha che

$$|ab - a_nb_n| \leq |b||a - a_n| + |a_n||b - b_n| < (|b| + M)\varepsilon \quad \forall n > n_\varepsilon.$$

Dato che la quantità $(|b| + M)\varepsilon$ può essere resa arbitrariamente piccola scegliendo ε sufficientemente piccolo, la distanza tra ab e a_nb_n diviene arbitrariamente piccola scegliendo valori di n sufficientemente grandi e quindi vale la conclusione.

Tramite queste regole è possibile calcolare molti limiti. Ad esempio:

$$\lim_{n \rightarrow +\infty} \frac{2n^2 - 1}{3n^2 + n + 1} = \lim_{n \rightarrow +\infty} \frac{2 - \frac{1}{n^2}}{3 + \frac{1}{n} + \frac{1}{n^2}} = \frac{\lim_{n \rightarrow +\infty} (2 - \frac{1}{n^2})}{\lim_{n \rightarrow +\infty} (3 + \frac{1}{n} + \frac{1}{n^2})} = \frac{2}{3},$$

passando al limite sia nel numeratore che nel denominatore.

Per ora, non è chiaro cosa dire sul comportamento al limite della successione $\frac{a_n}{b_n}$ nel caso in cui la successione b_n sia infinitesima. Torneremo tra breve sulla questione.

ESERCIZIO 2.1. *Calcolare il limite per $n \rightarrow +\infty$ delle seguenti successioni*

$$\frac{n^4 + 1}{n^4 + n^2}, \quad \frac{3n^2 + 1}{n(2n^2 + 1)}, \quad \frac{(n+1)(2n+2)(3n+3)}{n^3}, \quad \frac{(n+1)^2}{(n^2+1)^2}.$$

Limiti e disequazioni. Un'altra questione importante è come si comporti l'operazione di limite rispetto all'ordinamento dei numeri reali.

TEOREMA 2.2. Monotonia del limite. *Sia $\lim_{n \rightarrow +\infty} a_n = a$ e $\lim_{n \rightarrow +\infty} b_n = b$. Se, per qualche $N \in \mathbb{N}$, vale $a_n < b_n$ (oppure $a_n \leq b_n$) per ogni $n \geq N$, allora $a \leq b$.*

DIMOSTRAZIONE. Dato che a_n tende a a e b_n tende a b , si ha che per ogni $\varepsilon > 0$, $a - \varepsilon < a_n \leq b_n < b + \varepsilon$ per ogni $n > N$ per un opportuno N . Guardando il primo e l'ultimo termine nella catena di disequaglianze, si deduce che $b - a > -2\varepsilon$ per ogni $\varepsilon > 0$. Quindi, necessariamente, $b - a \geq 0$. \square

Il Teorema 2.2 afferma che l'operazione di limite è “monotona non decrescente”, nel senso che vale l'implicazione

$$a_n < b_n \quad \Rightarrow \quad a := \lim_{n \rightarrow +\infty} a_n \leq \lim_{n \rightarrow +\infty} b_n =: b.$$

Si noti che, nel passaggio al limite, la disuguaglianza stretta diviene una disuguaglianza non stretta. Ad esempio, scegliendo $a_n = 1/2n$ e $b_n = 1/n$, si ha $a_n < b_n$ per ogni n , ma $\lim_{n \rightarrow +\infty} a_n = \lim_{n \rightarrow +\infty} b_n = 0$.

In maniera analoga, si dimostra il seguente (utilissimo) risultato.

TEOREMA 2.3. *Teorema dei carabinieri.* Siano a_n, b_n, c_n successioni tali che

(i) esiste $N \in \mathbb{N}$ tale che $a_n \leq b_n \leq c_n$ per ogni $n > N$,

(ii) $\lim_{n \rightarrow +\infty} a_n = \lim_{n \rightarrow +\infty} c_n = \ell \in \mathbb{R}$.

Allora la successione b_n è convergente e $\lim_{n \rightarrow +\infty} b_n = \ell$.

DIMOSTRAZIONE. L'obiettivo è dimostrare che per ogni $\varepsilon > 0$ c'è una scelta di $n_\varepsilon \in \mathbb{N}$ tale che $\ell - \varepsilon < b_n < \ell + \varepsilon$ per ogni $n > n_\varepsilon$. Quindi occorre stimare dall'alto e dal basso i termini b_n . Dalle ipotesi segue che comunque si fissi $\varepsilon > 0$ esiste $\tilde{n}_\varepsilon \in \mathbb{N}$ tale che, per ogni $n > \tilde{n}_\varepsilon$,

$$\ell - \varepsilon < a_n < \ell + \varepsilon \quad \text{e} \quad \ell - \varepsilon < c_n < \ell + \varepsilon.$$

Quindi, dato che per ipotesi esiste N tale che $a_n \leq b_n \leq c_n$ per ogni $n > N$, si ottiene

$$\ell - \varepsilon < a_n \leq b_n \leq c_n < \ell + \varepsilon \quad \forall n > N := \max\{N, \tilde{n}_\varepsilon\},$$

ossia la conclusione. □

ESEMPIO 2.4. Dato $x \in (0, 1)$, applichiamo il Teorema 2.3, alla successione

$$b_n = x^n.$$

Chiaramente esiste $h > 0$ tale che $x = 1/(1+h)$. Poiché $(1+h)^n \geq 1+nh$ per ogni $n \in \mathbb{N}$ (dimostrarlo!),

$$0 < b_n = \frac{1}{(1+h)^n} \leq \frac{1}{1+nh} \quad \forall n \in \mathbb{N}.$$

Dato che $\lim_{n \rightarrow +\infty} \frac{1}{1+nh} = 0$, scegliendo $a_n = 0$ e $c_n = \frac{1}{1+nh}$ e applicando il Teorema 2.3

$$\lim_{n \rightarrow +\infty} x^n = 0 \quad \text{per } x \in (0, 1).$$

ESERCIZIO 2.5. Dati $\lambda \in [0, 1)$ e $a_0 \in \mathbb{R}$, sia a_n la successione definita per ricorrenza da $a_{n+1} = \lambda a_n$. Dimostrare che la successione a_n è infinitesima.

Conseguenza del Teorema 2.3 è questo piccolo criterio, che è una versione più generale dell'esercizio precedente: sostanzialmente si suppone che l'uguaglianza $a_{n+1} = \lambda a_n$ con $\lambda \in [0, 1)$ valga "all'infinito"...

COROLLARIO 2.6. Sia $a_n > 0$ per ogni n una successione tale che

$$(10) \quad \lim_{n \rightarrow +\infty} \frac{a_{n+1}}{a_n} = \lambda$$

Allora, se $\lambda \in [0, 1)$, la successione a_n è infinitesima.

DIMOSTRAZIONE. *Passo 1.* Dimostriamo che esistono $\sigma \in [0, 1)$ e $N \in \mathbb{N}$ tale che

$$(11) \quad a_{n+1} \leq \sigma a_n \quad \forall n \geq N.$$

Infatti, scegliamo $\sigma \in (\lambda, 1)$ e sia $\varepsilon := \sigma - \lambda$. Utilizzando la definizione di limite per la successione a_{n+1}/a_n , si deduce che esiste $n_\varepsilon \in \mathbb{N}$ tale che

$$\lambda - \varepsilon < \frac{a_{n+1}}{a_n} < \lambda + \varepsilon = \lambda + (\sigma - \lambda) = \sigma \quad \forall n > n_\varepsilon.$$

che porta alla (11) scegliendo $N = n_\varepsilon$.

Passo 2. Dimostriamo che a_n è infinitesima. Per semplicità, supponiamo $N = 0$ (cioè che (11) valga per ogni n), allora

$$0 < a_{n+1} \leq \sigma a_n \leq \sigma^2 a_{n-1} \leq \dots \leq \sigma^{n+1} a_0.$$

Dato che $\sigma \in [0, 1)$, per $n \rightarrow +\infty$ la quantità $\sigma^n a_1$ tende a zero e quindi si ha la conclusione. Nel caso generale, dato che

$$0 < a_{n+1} \leq \sigma^n \sigma^{1-N} a_N \quad \forall n \geq N,$$

la dimostrazione è analoga. □

È possibile dare criteri analoghi al Teorema 2.3 per dimostrare la divergenza di una successione. Ad esempio, una successione a_n che sia più grande di una successione b_n divergente a $+\infty$ è divergente a $+\infty$.

ESEMPIO 2.7. Consideriamo di nuovo la successione $b_n = x^n$, questa volta con $x > 1$. Allora $x = 1 + h$ con $h > 0$. Dalla disuguaglianza $(1 + h)^n \geq 1 + nh$, segue

$$b_n = (1 + h)^n \geq 1 + nh$$

Dato che $1 + nh \rightarrow +\infty$ per $n \rightarrow \infty$, anche la successione b_n diverge a $+\infty$.

Zeri a denominatore ed uso degli infiniti. Dall'analisi che abbiamo presentato fin qui restano fuori alcuni casi significativi:

- che succede della successione $\frac{a_n}{b_n}$ nel caso in cui $\lim_{n \rightarrow +\infty} b_n = 0$?
- che succede di somma/sottrazione/prodotto/rapporto quando qualcuno dei termini in gioco tende a $+\infty$ o a $-\infty$?

Partiamo dalla prima delle due questioni. Un numero molto piccolo a denominatore rende tutta la frazione molto grande. Quindi è ragionevole aspettarsi che, qualora il denominatore sia infinitesimo, il rapporto tenda a $+\infty$. Il banalissimo esempio:

$$a_n = 1, \quad b_n = \frac{1}{n} \quad \Rightarrow \quad \frac{a_n}{b_n} = \frac{1}{1/n} = n \rightarrow +\infty \quad \text{per } n \rightarrow +\infty,$$

dà conforto a questa prima ipotesi di lavoro. Arrischiamoci in una congettura:

$$\text{Congettura 1:} \quad \lim_{n \rightarrow +\infty} b_n = 0 \quad \Rightarrow \quad \lim_{n \rightarrow +\infty} \frac{a_n}{b_n} = +\infty$$

Ma se consideriamo il caso

$$a_n = \frac{1}{n}, \quad b_n = \frac{1}{n} \quad \Rightarrow \quad \frac{a_n}{b_n} = \frac{1/n}{1/n} = 1 \rightarrow 1 \quad \text{per } n \rightarrow +\infty.$$

Cosa succede? Molto semplice, il denominatore è infinitesimo, ma lo è anche il numeratore. Quindi, può capitare che il tendere a zero del denominatore sia (in qualche modo) compensato dal tendere a zero del numeratore! Proviamo una nuova versione:

$$\text{Congettura 2:} \quad \lim_{n \rightarrow +\infty} a_n = a \neq 0, \quad \lim_{n \rightarrow +\infty} b_n = 0 \quad \Rightarrow \quad \lim_{n \rightarrow +\infty} \frac{a_n}{b_n} = +\infty$$

Qui non siamo troppo lontani dal vero, ma ancora siamo stati un po' troppo leggeri nella questione dei segni: nel caso

$$a_n = 1, \quad b_n = -\frac{1}{n} \quad \Rightarrow \quad \frac{a_n}{b_n} = \frac{1}{-1/n} = -n \rightarrow -\infty \quad \text{per } n \rightarrow +\infty,$$

Proponiamo allora una versione (si spera finale) più precisa:

$$\text{Congettura 3:} \quad \lim_{n \rightarrow +\infty} a_n = a \neq 0, \quad \lim_{n \rightarrow +\infty} b_n = 0 \quad \Rightarrow \quad \lim_{n \rightarrow +\infty} \left| \frac{a_n}{b_n} \right| = +\infty$$

La Congettura 3 è vera. Per dimostrarla, bisogna mostrare che per ogni $M > 0$, vale la disuguaglianza $|a_n/b_n| \geq M$ per n sufficientemente grande. Dato che la disequazione precedente è equivalente a $|a_n| - M|b_n| \geq 0$ bisogna mostrare che

$$\forall M \quad \exists n_M \in \mathbb{N} \quad \text{t.c.} \quad |a_n| - M|b_n| \geq 0 \quad \forall n > n_M.$$

Niente di più facile, dato che

$$\lim_{n \rightarrow +\infty} (|a_n| - M|b_n|) = |a| > 0.$$

ESERCIZIO 2.8. *Siamo tranquilli e soddisfatti del nostro risultato, quando, d'improvviso, giunge un tipo, sicuro del fatto suo, che afferma*

$$\text{Congettura 4:} \quad \left. \begin{array}{l} \exists \nu > 0 \quad \text{t.c.} \quad a_n \geq \nu \quad \forall n, \\ \lim_{n \rightarrow +\infty} b_n = 0 \end{array} \right\} \quad \Rightarrow \quad \lim_{n \rightarrow +\infty} \left| \frac{a_n}{b_n} \right| = +\infty$$

Credergli o non credergli?

Resta fuori dalla nostra analisi il caso in cui sia il numeratore che il denominatore tendano a 0 per $n \rightarrow +\infty$. In questo caso può succedere di tutto! Ci sono casi in cui il rapporto tende a zero, casi in cui tende ad ∞ , casi in cui il limite del rapporto non esiste! Non esiste una regola generale e per questo si dice che si tratta di una forma

indeterminata (nel senso che non si può determinare subito se esista e quanto valga il limite ed occorre un'analisi più raffinata):

$$\text{forma indeterminata } \frac{0}{0} : \quad \lim_{n \rightarrow +\infty} a_n = \lim_{n \rightarrow +\infty} b_n = 0 \quad \Rightarrow \quad \lim_{n \rightarrow +\infty} \left| \frac{a_n}{b_n} \right| = ?$$

ESERCIZIO 2.9. *Trovare due successioni a_n e b_n infinitesime tali che la successione dei rapporti a_n/b_n non abbia limite.*

Consideriamo il caso in cui uno dei termini sia divergente (per fissare le idee, divergente a $+\infty$) e l'altro convergente, cioè supponiamo

$$\lim_{n \rightarrow +\infty} a_n = +\infty \quad \text{e} \quad \lim_{n \rightarrow +\infty} b_n = b.$$

Cosa succede di somma e prodotto? La regola, detta in maniera molto poco ortodossa, è che “finché i termini in gioco non si contrastano tra loro tutto va bene...”. Ad esempio:

$$\lim_{n \rightarrow +\infty} a_n + b_n = +\infty, \quad \text{se } b \neq 0, \quad \lim_{n \rightarrow +\infty} |a_n b_n| = +\infty,$$

(dimostrate queste proprietà!). L'unica situazione di “contrasto” è quella in cui la successione da studiare sia prodotto di una successione divergente e di una infinitesima

$$\text{forma indeterminata } \infty \cdot 0 : \quad \left. \begin{array}{l} \lim_{n \rightarrow +\infty} a_n = \pm\infty, \\ \lim_{n \rightarrow +\infty} b_n = 0 \end{array} \right\} \Rightarrow \lim_{n \rightarrow +\infty} a_n \cdot b_n = ?$$

Infine, resta la situazione più drammatica di tutte: entrambe le successioni a_n e b_n sono divergenti. Nel caso in cui si sommino due successioni divergenti a $+\infty$, la conclusione è evidente: la successione somma è anch'essa divergente a $+\infty$:

$$\lim_{n \rightarrow +\infty} a_n = \lim_{n \rightarrow +\infty} b_n = +\infty \quad \Rightarrow \quad \lim_{n \rightarrow +\infty} a_n + b_n = +\infty, \quad \lim_{n \rightarrow +\infty} a_n \cdot b_n = +\infty.$$

Nel caso in cui le due successioni divergano una a $+\infty$ e l'altra a $-\infty$, non si può dedurre nessuna conclusione generale:

$$\text{forma indeterminata } +\infty - \infty : \quad \left. \begin{array}{l} \lim_{n \rightarrow +\infty} a_n = +\infty, \\ \lim_{n \rightarrow +\infty} b_n = -\infty \end{array} \right\} \Rightarrow \lim_{n \rightarrow +\infty} a_n + b_n = ?$$

Il prodotto di successioni divergenti non crea nessun problema particolare (bisogna solo stare attenti al segno di ∞ , che si deduce con la buona vecchia regola del segno di un prodotto). Ad esempio,

$$\lim_{n \rightarrow +\infty} a_n = \lim_{n \rightarrow +\infty} b_n = +\infty \quad \Rightarrow \quad \lim_{n \rightarrow +\infty} a_n \cdot b_n = +\infty.$$

Per quanto riguarda il rapporto, pochi minuti di riflessione portano alla conclusione che l'unica situazione problematica è la seguente:

$$\text{forma indeterminata } \frac{\infty}{\infty} : \left. \begin{array}{l} \lim_{n \rightarrow +\infty} a_n = \infty, \\ \lim_{n \rightarrow +\infty} b_n = \infty \end{array} \right\} \Rightarrow \lim_{n \rightarrow +\infty} \frac{a_n}{b_n} = ?$$

3. Calcolo di alcuni limiti

Utilizziamo ora le proprietà dei limiti per studiare alcune successioni specifiche.

ESEMPIO 3.1.

$$\lim_{n \rightarrow +\infty} \sqrt[n]{x} = 1 \quad \text{per ogni } x > 0.$$

Poniamo $a_n = \sqrt[n]{x}$ e usiamo la disequazione $(1+h)^n \geq 1+nh$. Nel caso $x > 1$, allora anche $\sqrt[n]{x} > 1$ e quindi $h_n := \sqrt[n]{x} - 1 > 0$ e $a_n = 1 + h_n$. Esplicitando la disequazione rispetto a h_n ,

$$x = (a_n)^n = (1 + h_n)^n \geq 1 + nh_n \Rightarrow 0 < h_n \leq \frac{x-1}{n},$$

da cui segue $\lim_{n \rightarrow +\infty} h_n = 0$ e quindi otteniamo la conclusione nel caso $x > 1$.

Se $x = 1$ la successione è costante ed il risultato banale. Se $x < 1$, allora $1/x > 1$ e quindi $\sqrt[n]{1/x}$ converge ad 1 per quanto già visto. Dato che $\sqrt[n]{x} = \frac{1}{\sqrt[n]{1/x}}$, segue la conclusione.

ESEMPIO 3.2.

$$\lim_{n \rightarrow +\infty} \sqrt{n+1} - \sqrt{n} = 0.$$

In questo caso $a_n = \sqrt{n+1} - \sqrt{n}$ è la differenza di termini che tendono a $+\infty$. Passare al limite separatamente sui due termini dà l'espressione $+\infty - \infty$, che è *senza senso*.³ Ci troviamo di fronte ad una forma indeterminata e quindi per determinare l'esistenza o meno del limite bisogna lavorare un po' d'astuzia. Qui possiamo riscrivere a_n come

$$a_n = \frac{(\sqrt{n+1} - \sqrt{n})(\sqrt{n+1} + \sqrt{n})}{\sqrt{n+1} + \sqrt{n}} = \frac{(n+1) - n}{\sqrt{n+1} + \sqrt{n}} = \frac{1}{\sqrt{n+1} + \sqrt{n}},$$

che tende a 0 per $n \rightarrow +\infty$.

ESEMPIO 3.3. (Molto importante!)

$$\lim_{n \rightarrow +\infty} x^n = \begin{cases} 0 & x \in (-1, 1), \\ 1 & x = 1, \\ \text{non esiste} & x = -1, \\ +\infty & x > 1, \\ \text{diverge in modulo} & x < -1. \end{cases}$$

³In generale si pone la questione: si possono dare delle "regole algebriche" per l'uso dei simboli $\pm\infty$? Occorre ricordarsi che questi simboli sono definiti dall'operazione di limite e quindi, per definire tali regole, bisogna ricondursi alle proprietà dei limiti.

Abbiamo già visto che $\lim_{n \rightarrow +\infty} x^n = 0$ per $x \in (0, 1)$ e che $\lim_{n \rightarrow +\infty} x^n = +\infty$ per $x > 1$. Anche per $x \in (-1, 0]$, la successione x^n è infinitesima, dato che $\lim_{n \rightarrow +\infty} |x|^n = 0$.

Se $x = 1$ la successione è costantemente 1 (che quindi è il suo limite). Se $x = -1$, la successione oscilla tra i valori ± 1 ed è non regolare. Nel caso $x < -1$, la successione oscilla tra valori positivi e negativi e non è regolare, ma in valore assoluto diverge.

ESEMPIO 3.4.

$$\lim_{n \rightarrow +\infty} \frac{n^p}{x^n} = 0. \quad \forall x > 1, p \in \mathbb{N}.$$

Per dimostrare questo limite, applichiamo il Corollario 2.6 ad $a_n = \frac{n^p}{x^n}$:

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)^p}{x^{n+1}} \frac{x^n}{n^p} = \frac{(n+1)^p}{x n^p} = \frac{1}{x} \left(1 + \frac{1}{n}\right)^p \longrightarrow \frac{1}{x} \in (0, 1) \quad \text{per } n \rightarrow +\infty.$$

Dato che il rapporto $\frac{a_{n+1}}{a_n}$ tende ad un numero in $(0, 1)$, sono verificate le ipotesi del Corollario e quindi vale la conclusione.

ESEMPIO 3.5.

$$(12) \quad \lim_{n \rightarrow +\infty} \frac{x^n}{n!} = 0 \quad \forall x > 1.$$

Applichiamo di nuovo il Corollario 2.6

$$\frac{a_{n+1}}{a_n} = \frac{x^{n+1}}{(n+1)!} \frac{n!}{x^n} = \frac{x}{n+1} \longrightarrow 0 \quad \text{per } n \rightarrow +\infty,$$

da cui segue la conclusione.

ESEMPIO 3.6.

$$\lim_{n \rightarrow +\infty} \frac{n!}{n^n} = 0.$$

Infatti

$$0 \leq a_n = \frac{n(n-1) \cdots 2 \cdot 1}{n \cdot n \cdots n} = \frac{n}{n} \frac{n-1}{n} \cdots \frac{1}{n} = 1 \left(1 - \frac{1}{n}\right) \cdots \frac{1}{n} \leq \frac{1}{n},$$

quindi, per il Teorema 2.3, vale la conclusione.

ESEMPIO 3.7. **La serie geometrica.** Fissato $q \in \mathbb{R}$, la successione

$$S_n := \sum_{k=0}^n q^k,$$

è convergente se e solo se $q \in (-1, 1)$. Inoltre

$$\lim_{n \rightarrow +\infty} S_n = \frac{1}{1-q} \quad \forall q \in (-1, 1).$$

Infatti, la successione S_n si può riscrivere nella forma

$$S_n = 1 + q + q^2 + \cdots + q^n = \begin{cases} \frac{1 - q^{n+1}}{1 - q} & q \neq 1, \\ n + 1 & q = 1 \end{cases}$$

Passando al limite per $n \rightarrow +\infty$ è utilizzando i risultati dell'Esempio 3.3 si ottiene la conclusione. Nel caso $q \geq 1$ la successione diverge, mentre per $q \leq -1$ la serie è non regolare. Per esprimere la convergenza di S_n a $1/(1 - q)$ per $|q| < 1$, si usa la notazione

$$\text{serie geometrica :} \quad \sum_{k=0}^{+\infty} q^k := \lim_{n \rightarrow +\infty} \sum_{k=0}^n q^k = \frac{1}{1 - q} \quad \forall q \in (-1, 1).$$

Sul significato della parola “serie” in generale torneremo tra poco.

4. Successioni monotone

DEFINIZIONE 4.1. *Successioni monotone.* Una successione a_n è **strettamente crescente** se ogni termine è maggiore del precedente, cioè se $a_n < a_{n+1}$ per ogni $n \in \mathbb{N}$. Analogamente, è **strettamente decrescente** se $a_n > a_{n+1}$ per ogni $n \in \mathbb{N}$. Una successione a_n è **strettamente monotona** se è o strettamente crescente o strettamente decrescente.

Nel caso in cui valgano le disuguaglianze non strette valgono delle definizioni analoghe: a_n è **non decrescente** se $a_n \leq a_{n+1}$ per ogni n ed è **non crescente** se $a_n \geq a_{n+1}$. Una successione a_n è **monotona** se è o non decrescente o non crescente.

Per verificare se una successione è monotona occorre risolvere delle disequazioni. Ad esempio, la successione $a_n = \frac{1}{1 + n^2}$ è strettamente decrescente, infatti

$$a_{n+1} < a_n \quad \Longleftrightarrow \quad \frac{1}{1 + (n+1)^2} < \frac{1}{1 + n^2} \quad \Longleftrightarrow \quad 1 + n^2 < 1 + (n+1)^2$$

che è verificata per ogni $n \in \mathbb{N}$.

ESERCIZIO 4.2. Dire se la successione $a_n = \frac{n}{1+n^2}$ è strettamente decrescente.

ESERCIZIO 4.3. Siano a_n e b_n due successioni positive e non decrescenti. Dimostrare che $a_n + b_n$ e $a_n b_n$ sono anch'esse successioni non decrescenti. E' ancora vera la conclusione se si rimuove l'ipotesi di positività?

Se una successione è monotona, allora è preclusa la possibilità che abbia delle oscillazioni: i termini o salgono sempre o scendono sempre, non possono fare “un po' su e un po' giù”. In termini di esistenza/non esistenza del limite questa proprietà semplifica molto la casistica.

TEOREMA 4.4. Regolarità delle successioni monotone. Una successione a_n monotona è sempre regolare (cioè o è convergente o è divergente) e

$$\begin{aligned} a_n \text{ non decrescente} &\Rightarrow \lim_{n \rightarrow +\infty} a_n = \sup_{n \in \mathbb{N}} a_n, \\ a_n \text{ non crescente} &\Rightarrow \lim_{n \rightarrow +\infty} a_n = \inf_{n \in \mathbb{N}} a_n. \end{aligned}$$

Se a_n è anche limitata, allora è convergente.

DIMOSTRAZIONE. Supponiamo a_n non decrescente e superiormente limitata,

$$a_n \leq a_m \leq \ell := \sup_{n \in \mathbb{N}} a_n < +\infty, \quad n \leq m.$$

Per dimostrare che tale successione converge ad ℓ bisogna mostrare che comunque scelto $\varepsilon > 0$ è possibile scegliere $N \in \mathbb{N}$ per cui si ha $\ell - \varepsilon < a_n < \ell + \varepsilon$ per ogni $n > N$. La seconda delle due disequazioni è sempre verificata per definizione di ℓ (è un maggiorante). Inoltre, dato che ℓ è il più piccolo dei maggioranti, per ogni $\varepsilon > 0$ esiste certamente un indice N tale che $\ell - \varepsilon < a_N$ (altrimenti $\ell - \varepsilon$ sarebbe un maggiorante e, quindi, ℓ non sarebbe il più piccolo!). Usando la monotonia si ha

$$\ell - \varepsilon < a_N \leq a_n < \ell \quad \forall n \geq N,$$

cioè la conclusione.

I casi rimanenti si dimostrano in maniera simile. □

ESERCIZIO 4.5. Sia a_n una successione non decrescente. Dimostrare che:

- (i) se da a_n si può estrarre una sottosuccessione a_{n_k} convergente, allora a_n è convergente;
- (ii) se da a_n si può estrarre una sottosuccessione a_{n_k} divergente, allora a_n diverge.

5. Serie numeriche

In generale, una **serie numerica** è definita da una successione a_n , con la richiesta di sommare i termini nell'ordine dato dall'indice n . In parole povere, si tratta di dare senso alla *somma di un numero infinito di termini*. Il procedimento più naturale (utilizzato nell'Esempio 3.7) è di considerare la successione S_n delle **somme parziali**

$$S_0 := a_0, \quad S_1 := a_0 + a_1, \quad \dots, \quad S_n := \sum_{k=0}^n a_k,$$

cioè la successione il cui termine n -esimo è la somma dei primi $n+1$ termini a_0, \dots, a_n . Se la successione delle somme parziali S_n è convergente, la serie si dice **semplicemente convergente**. La **somma della serie**, che si indica con $\sum_{n=0}^{\infty} a_n$, è definita da

$$\sum_{n=0}^{\infty} a_n := \lim_{n \rightarrow +\infty} S_n,$$

nel caso in cui tale limite esista. Si usa la stessa terminologia delle successioni: una serie è **convergente/divergente/regolare/non regolare** se la successione delle somme parziali è convergente/divergente/regolare/non regolare.

Grazie alla linearità del limite di successioni, anche le serie godono della proprietà di linearità: combinazioni lineari di serie convergenti danno luogo a serie convergenti.

TEOREMA 5.1. Linearità delle serie. *Siano a_k e b_k i termini generici di due serie convergenti. Allora, per ogni $\lambda, \mu \in \mathbb{R}$, la serie di termine generico $\lambda a_k + \mu b_k$ è convergente e vale la formula*

$$\sum_{k=0}^{\infty} (\lambda a_k + \mu b_k) = \lambda \sum_{k=0}^{\infty} a_k + \mu \sum_{k=0}^{\infty} b_k.$$

In generale non è facile stabilire se una serie sia o non sia convergente. Un primo ingrediente utile per questo studio è il seguente: *se una serie è convergente, allora il suo termine generico deve essere infinitesimo:*

$$\text{condizione necessaria} : \quad \sum_{n=0}^{\infty} a_n \text{ convergente} \Rightarrow \lim_{n \rightarrow +\infty} a_n = 0.$$

Infatti, sia $\sum_{n=0}^{\infty} a_n$ convergente, cioè esista finito $\lim_{n \rightarrow +\infty} S_n = S$. Dato che $a_n = S_n - S_{n-1}$ e, per ipotesi, la successione S_n converge a S per $n \rightarrow +\infty$, allora

$$\lim_{n \rightarrow +\infty} a_n = \lim_{n \rightarrow +\infty} (S_n - S_{n-1}) = \lim_{n \rightarrow +\infty} S_n - \lim_{n \rightarrow +\infty} S_{n-1} = S - S = 0.$$

Questa condizione necessaria rispecchia il fatto, suggerito dall'intuizione, secondo cui occorre sommare termini che tendono a zero per sperare che la somma di un numero infinito di termini dia un valore finito. E' importante notare che esistono serie il cui termine generico a_n è infinitesimo, ma che non sono convergenti. Ad esempio, possiamo considerare la serie definita dalla successione seguente: il primo termine vale 1, i due termini successivi valgono $\frac{1}{2}$, i tre termini successivi valgono $\frac{1}{3}$ e così via:

$$1, \frac{1}{2}, \frac{1}{2}, \frac{1}{3}, \frac{1}{3}, \frac{1}{3}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{4}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \frac{1}{5}, \dots$$

E' chiaro che la successione è infinitesima, ma le corrispondenti somme ridotte verificano:

$$S_0 = 1, S_3 = 2, S_6 = 3, S_{10} = 4, S_{15} = 5, \dots$$

e quindi danno luogo ad una successione divergente.

Criteri di confronto per serie a termini positivi. Stabilire la convergenza o meno di una serie è più facile se la successione a_k è costituita da termini non negativi. Infatti, dato che $a_k \geq 0$, si ha $S_{n+1} - S_n = a_{n+1} \geq 0$, cioè $S_n \leq S_{n+1}$ per ogni n . Quindi la successione S_n è non decrescente e pertanto regolare: o converge o diverge,

$$a_k \geq 0 \quad \Rightarrow \quad \sum_{k=0}^{+\infty} a_k < +\infty \quad \text{oppure} \quad \sum_{k=0}^{+\infty} a_k = +\infty.$$

Dato che una serie a termini positivi può solo convergere o divergere a $+\infty$, condizione sufficiente di convergenza è che esista una serie maggiorante b_k convergente.

TEOREMA 5.2. Criterio di confronto per serie. *Sia $a_k \geq 0$ per ogni k .*

(i) *Se esiste b_k tali che $0 \leq a_k \leq b_k$ per ogni k , allora*

$$\sum_{k=0}^{\infty} b_k < +\infty \quad \Rightarrow \quad \sum_{k=0}^{\infty} a_k < +\infty$$

(ii) *se esiste $b_k \geq 0$ tali che $a_k \geq b_k$ per ogni k , allora*

$$\sum_{k=0}^{\infty} b_k = +\infty \quad \Rightarrow \quad \sum_{k=0}^{\infty} a_k = +\infty.$$

OSSERVAZIONE 5.3. Se in (i) e in (ii) la richiesta di stima con i termini b_k è soddisfatta definitivamente, cioè per ogni $k \geq \bar{N}$ per un opportuno \bar{N} , anziché per ogni $k \in \mathbb{N}$, la conclusione vale lo stesso, dato che *le proprietà di convergenza non cambiano quando si cambiano un numero finito di termini.*

DIMOSTRAZIONE. Dimostriamo solo la parte (i) lasciando la (ii) come esercizio. La successione delle somme parziali di a_n è non decrescente, quindi per dimostrarne la convergenza, è sufficiente mostrarne la limitatezza (grazie al Teorema 4.4). Dalle maggiorazioni

$$0 \leq \sum_{k=0}^n a_k \leq \sum_{k=0}^n b_k \leq \sum_{k=0}^{\infty} b_k < +\infty,$$

segue la conclusione. □

ESEMPIO 5.4. Consideriamo la serie (a termini non negativi)

$$\sum_{n=1}^{\infty} \frac{2^n - 1}{3^n + n}.$$

Euristicamente, è ragionevole aspettarsi che, per valori di n grandi il termine generico $\frac{2^n - 1}{3^n + n}$ si possa approssimare come segue

$$\frac{2^n - 1}{3^n + n} \approx \frac{2^n}{3^n} = \left(\frac{2}{3}\right)^n,$$

quindi il sospetto è che sia convergente, vista la somiglianza con la serie geometrica. Poniamoci quindi l'obiettivo di stimarla dall'alto

$$\frac{2^n - 1}{3^n + n} \leq \frac{2^n}{3^n} = \left(\frac{2}{3}\right)^n.$$

Dato che il termine generico della serie è minore di quello della serie geometrica di ragione $\frac{2}{3}$, la serie è convergente.

OSSERVAZIONE 5.5. Una maniera facile per determinare una stima del tipo $a_k \leq Cb_k$ per k sufficientemente grande è tramite il limite del rapporto a_k/b_k . Supponiamo, ad esempio, di sapere che

$$\lim_{k \rightarrow +\infty} \frac{a_k}{b_k} = \ell \geq 0 \quad \text{e} \quad \sum_{k=1}^{\infty} b_k < +\infty,$$

Per definizione di limite, scelto $\varepsilon > 0$, esiste $N \in \mathbb{N}$ opportuno per cui, per ogni $n \geq N$,

$$\ell - \varepsilon < \frac{a_k}{b_k} < \ell + \varepsilon \quad \Rightarrow \quad a_k < (\ell + \varepsilon)b_k.$$

Quindi, applicando il precedente risultato, si dimostra che la serie $\sum_{k=0}^{\infty} a_k$ è convergente.

ESEMPIO 5.6. Fissato $x \geq 0$, studiamo la **serie esponenziale**

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

Sappiamo già che il termine generico di questa serie è infinitesimo. Purtroppo questo non basta a stabilire la convergenza della serie. L'impressione è che, dato che il fattoriale cresce più rapidamente di qualsiasi esponenziale (vedi Esempio 3.5), il termine generico vada a zero molto rapidamente tanto da far convergere la serie. Scegliamo, allora, come serie b_k di confronto quella di termine generico $b_k = y^k$ con $y \in (0, 1)$. Dato che, per (12),

$$\lim_{k \rightarrow +\infty} \frac{x^k/k!}{y^k} = \lim_{k \rightarrow +\infty} \frac{(x/y)^k}{k!} = 0,$$

effettivamente, per k sufficientemente grande vale la stima $x^k/k! \leq y^k$ che mostra che la serie esponenziale è convergente per ogni scelta di $x > 0$.

Serie numeriche a segno qualsiasi. Come si è detto, verificare la convergenza semplice di una serie può essere molto complicato. Ben diverso è se la serie è a termini positivi, dato che in questo caso la successione delle somme parziali è non decrescente. Per questo motivo, per serie numeriche con termini a segno qualsiasi si introduce un nuovo tipo di convergenza che è più restrittivo di quella semplice, ma che è più facile da verificare, dato che richiede la verifica della convergenza di una serie a termini positivi.

DEFINIZIONE 5.7. La serie $\sum_{n=1}^{\infty} a_n$ converge assolutamente se è convergente la serie dei valori assoluti, cioè se $\sum_{n=1}^{\infty} |a_n|$ è convergente.

Nel caso in cui la serie sia a termini non negativi, cioè $a_n \geq 0$ per ogni n , la convergenza assoluta è equivalente alla convergenza semplice. In generale vale l'implicazione (mentre l'implicazione opposta è falsa!)

$$\text{convergenza assoluta} \quad \Rightarrow \quad \text{convergenza semplice}$$

Per dimostrarla, introduciamo due nuovi oggetti. Dato $a \in \mathbb{R}$, la parte positiva a^+ e la parte negativa a^- di a sono definite da

$$a^+ = \max(a, 0) \quad , \quad a^- = \max(-a, 0).$$

Quindi, se $a = 2$, $a^+ = 2$ e $a^- = 0$, mentre se $a = -3$, $a^+ = 0$ e $a^- = 3$. Dalla definizione discendono le proprietà seguenti:

$$0 \leq a^+ \leq |a|, \quad 0 \leq a^- \leq |a|, \quad a = a^+ - a^-, \quad |a| = a^+ + a^-.$$

ESERCIZIO 5.8. Disegnare il grafico delle funzioni parte positiva $f(x) = x^+$ e parte negativa $g(x) = x^-$.

Se a_k è il termine generico di una serie, consideriamo le altre due serie di termine generico a_k^+ e a_k^- . Ad esempio, se $a_k = (-1)^k / (k+1)$,

$$\begin{aligned} a_k &: 1, -\frac{1}{2}, \frac{1}{3}, -\frac{1}{4}, \frac{1}{5}, -\frac{1}{6}, \dots \\ a_k^+ &: 1, 0, \frac{1}{3}, 0, \frac{1}{5}, 0, \dots \\ a_k^- &: 0, \frac{1}{2}, 0, \frac{1}{4}, 0, \frac{1}{6}, \dots \end{aligned}$$

Che collegamento c'è tra la convergenza assoluta e la convergenza della serie delle parti positive e delle parti negative? Semplice... si tratta della stessa cosa!

PROPOSIZIONE 5.9. Una serie di termine generico a_k converge assolutamente se e solo se convergono (semplicemente) le serie di termine generico a_k^+ e a_k^- .

DIMOSTRAZIONE. Supponiamo che la serie $\sum a_k$ sia assolutamente convergente. Allora è possibile applicare il criterio di confronto per serie a termini positivi (Teorema 5.2): dato che $0 \leq a_k^\pm \leq |a_k|$ e la serie $\sum |a_k|$ è convergente, anche le serie $\sum a_k^+$ e $\sum a_k^-$ sono convergenti.

Viceversa, supponiamo che le serie $\sum a_k^+$ e $\sum a_k^-$ siano convergenti, allora per la linearità delle serie (Teorema 5.1) anche la loro somma è convergente. Dato che la loro somma è proprio la serie dei valori assoluti. \square

Da questa caratterizzazione della convergenza assoluta, discende rapidamente il risultato seguente.

TEOREMA 5.10. *Sia a_k il termine generico di un serie. Se $\sum_{k=0}^{\infty} a_k$ converge assolutamente, allora converge anche semplicemente.*

DIMOSTRAZIONE. Se la serie converge assolutamente, allora le serie $\sum a_k^+$ e $\sum a_k^-$ sono convergenti, quindi per il Teorema 5.1 anche la loro differenza è convergente: ma $a_k = a_k^+ - a_k^-$ e quindi la conclusione è a portata di mano. Allungatevi. \square

Se una serie converge assolutamente, allora vale la seguente versione (per somme infinite) della disuguaglianza triangolare:

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

La dimostrazione (facile) è lasciata per esercizio.

ESEMPIO 5.11. Fissato $x \in \mathbb{R}$, consideriamo la serie

$$\sum_{k=1}^{\infty} \frac{\sin(kx)}{2^k}.$$

Questa serie è assolutamente convergente per ogni scelta di $x \in \mathbb{R}$. Infatti

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

e la serie $\sum \frac{1}{2^k}$ è convergente perché è la serie geometrica con ragione $\frac{1}{2}$. Quindi la serie $\sum \frac{\sin(kx)}{2^k}$ converge per ogni $x \in \mathbb{R}$.

La serie esponenziale. Dato $x \in \mathbb{R}$, consideriamo la serie

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} \quad x \in \mathbb{R}.$$

Abbiamo già visto che la serie è convergente per $x \geq 0$. Cosa succede per i valori di $x < 0$? Facile, basta ricorrere alla convergenza assoluta! Dato che

$$\sum_{k=0}^{\infty} \left| \frac{x^k}{k!} \right| = \sum_{k=0}^{\infty} \frac{|x|^k}{k!}$$

è convergente perché $|x| \geq 0$, la serie assegnata converge assolutamente e quindi converge anche semplicemente. Si definisce

$$(13) \quad \text{funzione esponenziale:} \quad e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!} \quad x \in \mathbb{R}.$$

Questa posizione è del tutto rigorosa dato che la serie è sempre convergente. A rigore, bisognerebbe verificare che questa definizione sia compatibile con la costruzione naïf e con le proprietà viste in precedenza per le funzioni esponenziali, ma non ci soffermeremo su tale aspetto.

Approssimazione del numero di Nepero. A partire dalla formula

$$e = \sum_{n=0}^{\infty} \frac{1}{n!},$$

è possibile ottenere approssimazioni del valore di e . Prima di tutto stimiamo l'errore che si commette sostituendo il numero e con la somma parziale $s_n = \sum_{k=0}^n \frac{1}{k!}$, cioè la differenza $e - s_n$

$$\begin{aligned} 0 < e - s_n &= \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \cdots = \frac{1}{(n+1)!} \left(1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \cdots \right) \\ &< \frac{1}{(n+1)!} \left(1 + \frac{1}{n+1} + \frac{1}{(n+1)^2} + \cdots \right) = \frac{1}{(n+1)!} \cdot \frac{1}{1 - \frac{1}{n+1}} = \frac{1}{n!n}, \end{aligned}$$

quindi vale

$$\left| e - \sum_{k=0}^n \frac{1}{k!} \right| \leq \frac{1}{n!n}.$$

Da questa stima si può determinare il numero di elementi che occorre sommare per ottenere una approssimazione di e con errore prescritto. Ad esempio, supponiamo di ammettere un errore massimo di 10^{-4} . Dato che

n	1	2	3	4	5	6	7
$n!n$	1	4	18	96	600	4320	35280
$1/(n!n)$	1	0,25	0,05	0,010412	0,0016	0,000231	0,000028

va bene scegliere $n = 7$. Quindi l'approssimazione richiesta è

$$e \approx 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} + \frac{1}{720} + \frac{1}{5040} = 2,718254.$$

La serie armonica e la serie armonica generalizzata. Una delle serie numeriche più famose è quella di termine generico $\frac{1}{n}$, che è detta **serie armonica**. Questa serie è divergente

$$(14) \quad \sum_{n=1}^{\infty} \frac{1}{n} = +\infty.$$

Infatti, sia $S_n := \sum_{k=1}^n \frac{1}{k}$ e consideriamo la sottosuccessione S_{2^k} di S_n . Allora si ha

$$S_1 = 1, \quad S_2 = 1 + \frac{1}{2}, \quad S_4 = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right), \dots$$

e così via. Manteniamo i termini S_1 e S_2 così come sono e preoccupiamoci dei successivi.

Il termine S_4 può essere stimato dal basso:

$$S_4 = 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) \geq 1 + \frac{1}{2} + \left(\frac{1}{4} + \frac{1}{4}\right) = 1 + \frac{1}{2} + \frac{2}{4} = 1 + \frac{1}{2} + \frac{1}{2} = 1 + \frac{2}{2}.$$

Analogamente, vale

$$\begin{aligned} S_8 &= 1 + \frac{1}{2} + \left(\frac{1}{3} + \frac{1}{4}\right) + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) = S_4 + \left(\frac{1}{5} + \frac{1}{6} + \frac{1}{7} + \frac{1}{8}\right) \\ &\geq S_4 + \left(\frac{1}{8} + \frac{1}{8} + \frac{1}{8} + \frac{1}{8}\right) = S_4 + \frac{1}{2} \geq 1 + \frac{1}{2} + \frac{1}{2} + \frac{1}{2} = 1 + \frac{3}{2}. \end{aligned}$$

In generale, consideriamo il termine $S_{2^{k+1}}$:

$$\begin{aligned} S_{2^{k+1}} &= S_{2^k} + \left(\frac{1}{2^k+1} + \frac{1}{2^k+2} + \dots + \frac{1}{2^{k+1}}\right) \\ &\geq S_{2^k} + \left(\frac{1}{2^{k+1}} + \frac{1}{2^{k+1}} + \dots + \frac{1}{2^{k+1}}\right) = S_{2^k} + \frac{2^k}{2^{k+1}} = S_{2^k} + \frac{1}{2} \end{aligned}$$

dato che i termini in parentesi sono $2^{k+1} - 2^k = 2^k$. Quindi, ogni volta che l'indice k della sottosuccessione S_{2^k} aumenta di 1, la stima dal basso aumenta di $1/2$. Dato che $S_{2^0} = 1$, se ne deduce che

$$S_{2^{k+1}} \geq 1 + \frac{k+1}{2},$$

da cui segue che la sottosuccessione S_{2^k} è divergente. Dato che S_n è crescente ed ammette una sottosuccessione divergente, essa stessa è divergente.

In generale, la **serie armonica generalizzata**

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} \quad \alpha > 0,$$

è convergente se e solo se $\alpha > 1$.

La divergenza nel caso $\alpha \in (0, 1)$ segue da $n^\alpha < n$ per ogni n (dimostrare!), che implica:

$$\sum_{k=1}^n \frac{1}{k^\alpha} > \sum_{k=1}^n \frac{1}{k}.$$

Nel caso $\alpha > 1$ tale serie è convergente. Dato che si tratta di una “somma” di termini positivi, la successione delle somme parziali è monotona crescente. Per la proprietà delle successioni monotone (Teorema 4.4), per dimostrarne la convergenza, basta verificare che la serie non diverga a $+\infty$. Perciò, con astuzia, consideriamo gli insiemi $I_k = \{n \in \mathbb{N} : 2^k \leq n < 2^{k+1}\}$,

notando che $\mathbb{N} = \bigcup_{k=0}^{\infty} I_k$. L'insieme I_k ha $2^{k+1} - 2^k = 2^k(2 - 1) = 2^k$ termini e, se $n \in I_k$, vale la maggiorazione $1/n^\alpha \leq 1/2^{\alpha k}$. Stimiamo la somma dei termini $1/n^\alpha$ per $n \in I_k$:

$$\sum_{n \in I_k} \frac{1}{n^\alpha} \leq \sum_{n \in I_k} \frac{1}{2^{\alpha k}} = 2^k \cdot \frac{1}{2^{\alpha k}} = \left(\frac{1}{2^{\alpha-1}} \right)^k.$$

Quindi, notando che $\frac{1}{2^{\alpha-1}} \in (-1, 1)$ per $\alpha > 1$,

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} = \sum_{k=0}^{\infty} \sum_{n \in I_k} \frac{1}{n^\alpha} \leq \sum_{k=0}^{\infty} \left(\frac{1}{2^{\alpha-1}} \right)^k < +\infty$$

grazie alla convergenza della serie geometrica.