Sapienza, Università di Roma

a.a. 2008/09

Note per il corso di

Analisi

Parte prima

versione 1.0 (September 29, 2008)

L.Lamberti e C.Mascia

Corsi di laurea in Fisica & Fisica ed Astrofisica

Indice

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.

1. Numeri naturali, interi, razionali e irrazionali

Per arrivare ai numeri reali, seguiremo un percorso classico: dai numeri *naturali* ai *razionali*, passando per gli *interi*, e giungendo, alla fine, 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 $\mathbb{N}$:

$$\textsf{numeri naturali:} \qquad \mathbb{N} = \{0, 1, 2, 3, \cdots\}.$$

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 ci 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):

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

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:

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

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:

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

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.

(iv) Rappresentiamo ora i numeri razionali $\dfrac{p}{q}$ 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

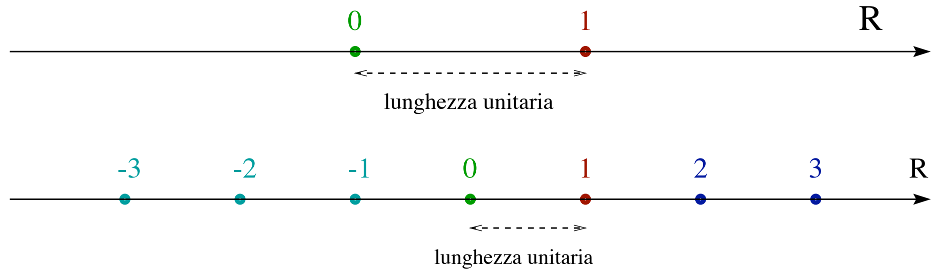

FIGURA 1. Dall'alto verso il basso: *(a)* i punti 0 e 1 sulla retta R determinano il segmento di lunghezza unitaria; *(b)* la rappresentazione dei numeri interi su R.

in q parti di uguale lunghezza e si prende come rappresentante di $\frac{p}{q}$ l'estremo destro del $p-$esimo intervallo.

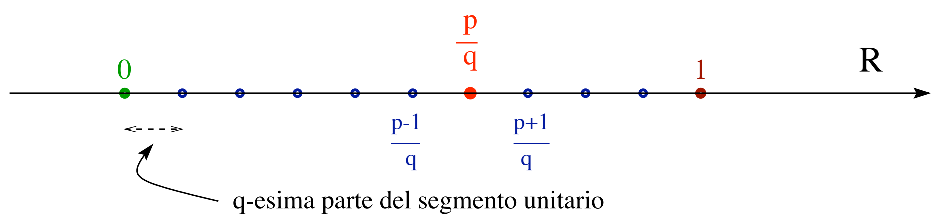

FIGURA 2. La rappresentazione su R di un numero razionale $\frac{p}{q}$ per cui $p, q \in \mathbb{N}$ e $p < q$.

(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 $x \leq y$

$$x \leq y \qquad \Longleftrightarrow \qquad x < y \quad \textit{oppure} \quad x = y.$$

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 $x \geq 0$.*

A questo punto, è possibile introdurre 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 $|x - y|$.*

Dalla definizione, segue che

$$|x| \geq 0 \qquad \forall\, x \in \mathbb{Q} \qquad \text{e} \qquad |x| = 0 \qquad \Longleftrightarrow \qquad x = 0.$$

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

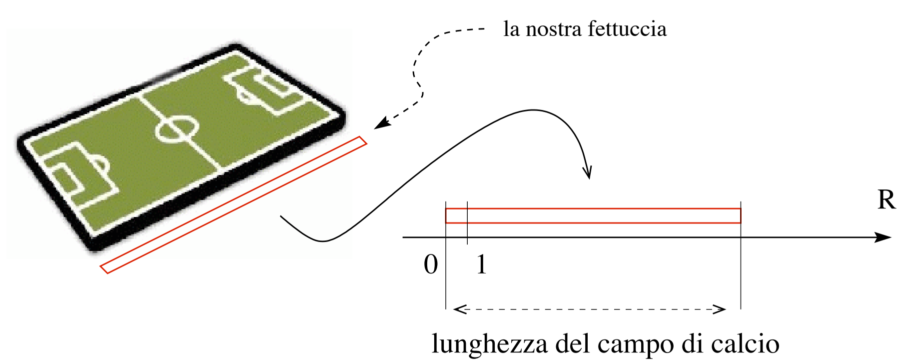

Figura 3. Il procedimento di misurazione: un campo di calcio.

[1]Qualcuno 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.

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

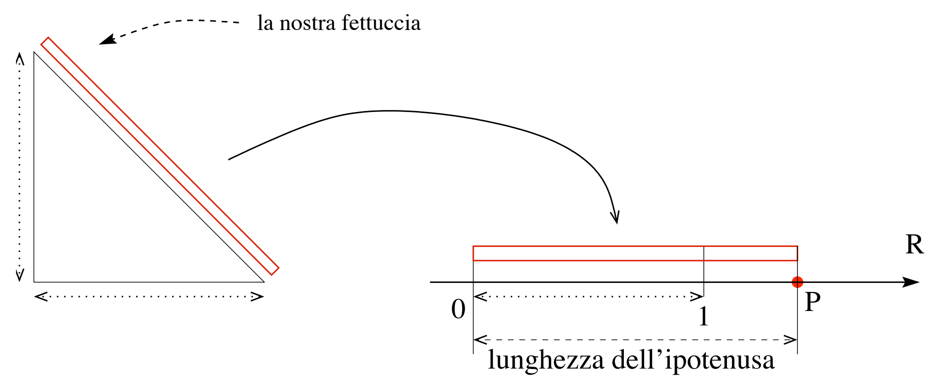

FIGURA 4. Il procedimento di misurazione dell'ipotenusa di un triangolo rettangolo di lato unitario: il punto P non corrisponde a nessun numero razionale!

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?*

[2]Per 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:

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

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 sostuisce 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:

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

[3] L'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.

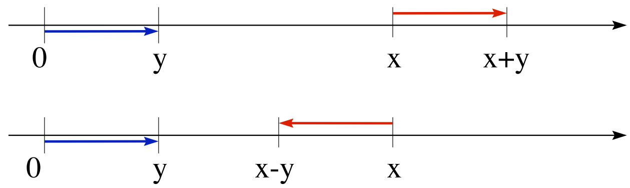

Figura 5. Somma e differenza di numeri reali.

leggi 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: $a \cdot \dfrac{1}{a} = 1$ 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.

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:

$$\text{intervallo aperto:} \qquad (a, b) := \{x \in \mathbb{R} \,:\, a < x < b\},$$

$$\text{intervallo chiuso:} \qquad [a, b] := \{x \in \mathbb{R} \,:\, a \leq x \leq b\}.$$

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

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

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:

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

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

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

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 <u>ordinate</u> (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

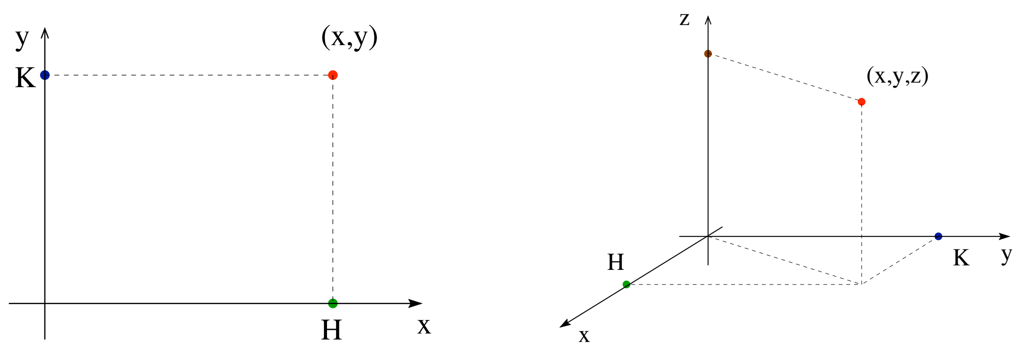

FIGURA 6. Il piano cartesiano $\mathbb{R}^2$ e lo spazio cartesiano $\mathbb{R}^3$.

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).

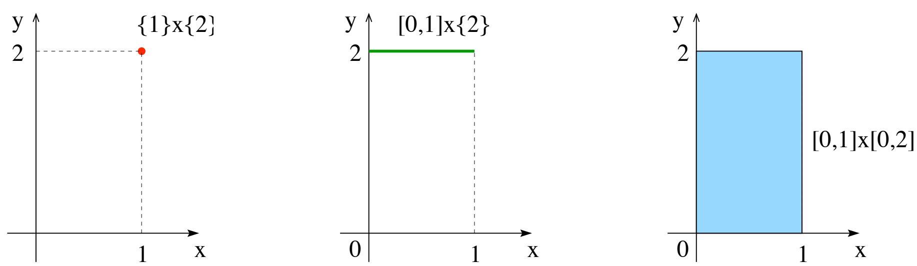

FIGURA 7. Gli insiemi $\{1\} \times \{2\}$, $[0,1] \times \{2\}$ e $[0,1] \times [0,2]$.

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

$$(x_1, y_1) + (x_2, y_2) := (x_1 + x_2, y_1 + y_2). \tag{1}$$

Non è evidente, al contrario, come dare una nozione di prodotto tra due punti[4] del piano $\mathbb{R}^2$. E' possibile, invece, introdurre la nozione di prodotto per uno scalare: dato $\lambda \in \mathbb{R}$, si definisce

$$\lambda\,(x, y) := (\lambda\, x, \lambda\, y). \tag{2}$$

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

[4] Definire 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.

coordinati ortogonali tra loro. In generale, $\mathbb{R}^n$ (dove $n \in \mathbb{N}$) indica l'insieme delle $n-$ple del tipo $(x_1, \ldots, x_n)$ dove $x_1, \ldots, x_n \in \mathbb{R}$. In sintesi

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

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 $x = (x_1, \ldots, x_n), \xi = (\xi_1, \ldots, \xi_n) \in \mathbb{R}^n$ e $\lambda \in \mathbb{R}$

$$x + \xi = (x_1, \ldots, x_n) + (\xi_1, \ldots, \xi_n) := (x_1 + \xi_1, \ldots, x_n + \xi_n). \tag{3}$$

$$\lambda\,(x_1, \ldots, x_n) := (\lambda\, x_1, \ldots, \lambda\, x_n). \tag{4}$$

In questo modo, $\mathbb{R}^n$ acquisice 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 $x = (x_1, \ldots, x_n) \in \mathbb{R}^n$, il* modulo di *$x$ si indica con $|x|_n$ ed è definito da*

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

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).

2. Ordine e diseguaglianze

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. Non solo! In molte situazioni, la stima di x è un'informazione sufficiente per la soluzione del problema. Le diseguaglianze sono, perciò, un oggetto fondamentale nell'uso dei numeri reali. In questa Sezione richiamiamo alcune regole elementari a riguardo.

Regole di base per le disequazioni

(i) La somma di numeri positivi è positiva: $a > 0,\, b > 0 \quad \Rightarrow \quad a + b > 0.$

(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

$$ab > 0 \iff a > 0,\, b > 0 \quad \text{oppure} \quad a < 0,\, b < 0$$

$$ab < 0 \iff a > 0,\, b < 0 \quad \text{oppure} \quad a < 0,\, b > 0.$$

(iii) Le diseguaglianze possono essere sommate[5]: $a > b,\, c > d \quad \Rightarrow \quad a + c > b + d.$

(iv) Se si moltiplica una diseguaglianza per un numero positivo il segno $>$ rimane invariato:

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

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

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

Esercizio 2.1. *Dimostrare le regole di base delle disequazioni.*

Soluzione. Ecco le soluzioni di alcune delle proprietà.
(iii) $(a + c) - (b + d) = (a - b) + (c - d)$ è positivo perché somma di positivi.
(iv) Infatti $ac - bc = (a - b)c$ è positivo perché prodotto di numeri positivi. Moltiplicare per $c < 0$, a quale diseguaglianza porta?
(v) La transitività segue subito da $a - c = (a - b) + (b - c)$ e dalla proprietà (i).

Esercizio 2.2. *Dimostrare la seguente disequazione*

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

Soluzione. Partiamo da un paio di situazioni semplici:

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

Cosa succede in generale? Dato che $(1 + h)^n = (1 + h) \dots (1 + h)$ (il termine $(1 + h)$ compare n volte), quando si fa il prodotto di 1 per n volte si ottiene 1, quando si moltiplica il termine 1 in tutti i termini tranne uno, si ottiene h, quindi $(1 + h)^n = 1 + h + \dots + h + \dots$. In quanti modi possibili compare il termine h? Uno per ciascuno dei termini $(1 + h)$, cioè n. In definitiva

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

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

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

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

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

Quindi l'operazione "elevare al quadrato" preserva l'ordine dei numeri positivi[6]: 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 2.3. *Siano $a, b > 0$ e $n \in \mathbb{N}$. Allora*

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

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 $|c|^2 = c^2$,

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

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

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

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

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:

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

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 monotonía* e verranno riprese più in là.

Se n è dispari, la conclusione della Proposizione 2.3 vale per ogni $a, b \in \mathbb{R}$. Infatti, dalle regole di base per le diseguaglianze, 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 $a^{n-1} + a^{n-2}b + \cdots + ab^{n-2} + b^{n-1}$ è sempre positivo, e ragionare come nella dimostrazione della Proposizione 2.3. Riassumendo:

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

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

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 simboletto $|\cdot|$ 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*:*

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

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 diseguglianza triangolare (iii). Distinguiamo i casi $a + b \geq 0$ e $a + b < 0$. Nel primo caso, la diseguaglianza 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 diseguaglianza 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:* $|x| = \max\{x, -x\}$.

Soluzione. Se $x \geq 0$, allora $-x \leq 0 \leq x$, quindi $\max\{x, -x\} = x = |x|$; se $x < 0$, allora $x < 0 < -x$, quindi $\max\{x, -x\} = -x = |x|$.

Dato che il modulo è indipendente dal segno, cioè $|-x| = |x|$, la diseguaglianza 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 diseguaglianza $|a - b| \leq |a| - |b|$ è falsa.

Applicando due volte la diseguaglianza triangolare, si ottiene

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

Allo stesso modo, si ottiene la diseguaglianza più generale

$$|a_1 + a_2 + \cdots + a_n| \leq |a_1| + |a_2| + \cdots + |a_n| \qquad \forall a_1, \ldots, a_n \in \mathbb{R}, .$$

o, equivalentemente (tramite il simbolo di sommatoria[7])

$$\left|\sum_{k=1}^{n} a_k\right| \leq \sum_{k=1}^{n} |a_k| \qquad \forall a_1, \ldots, a_n \in \mathbb{R}, \tag{6}$$

Alcune conseguenze dirette della diseguaglianza triangolare sono molto utili.

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

$$\textit{(i)} \qquad |a| \leq |a \pm b| + |b|, \qquad\qquad \textit{(ii)} \qquad \big|\,|a| - |b|\,\big| \leq |a \pm b|,$$

dove il simbolo $\pm$ indica che la diseguaglianza 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 $a > 0$,*

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

Soluzione. Dato che $|x| = \max\{-x, x\}$, la diseguaglianza $|x| \leq a$ è equivalente a dire che sia x che $-x$ sono minori od uguali di a, cioè $x \leq a$ e $-x \leq a$. Quindi $|x| \leq a$ è equivalente a $-a \leq x \leq a$. D'altronde, se si pensa ad $|x|$ come distanza di x da 0....

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 $[-R, R]$:*

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

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:*

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

[7]La somma di n numeri $x_1, \ldots, x_n$ viene indicata con $\sum\limits_{k=1}^{n} x_k := x_1 + x_2 + \cdots + x_n$. Ad esempio $\sum\limits_{k=1}^{5} k = 1 + 2 + 3 + 4 + 5 = 15$, $\sum\limits_{k=1}^{4} k^2 = 1^2 + 2^2 + 3^2 + 4^2 = 30$.

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 $a, b \in \mathbb{R}$,

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

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:* $|a - b| = |b - a|$, cioè la distanza di a da b è uguale a quella di b da a;

(c) *diseguaglianza triangolare*: vale

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

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?)

$$|x - x_0| < r \quad \Longleftrightarrow \quad -r < x - x_0 < r \quad \Longleftrightarrow \quad x_0 - r < x < x_0 + r.$$

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 intersecano 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.

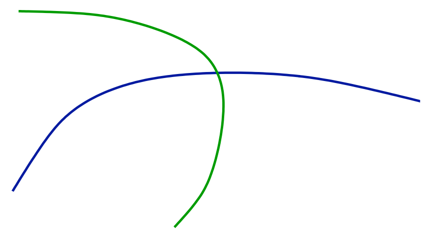

FIGURA 8. Due curve nel piano.

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".

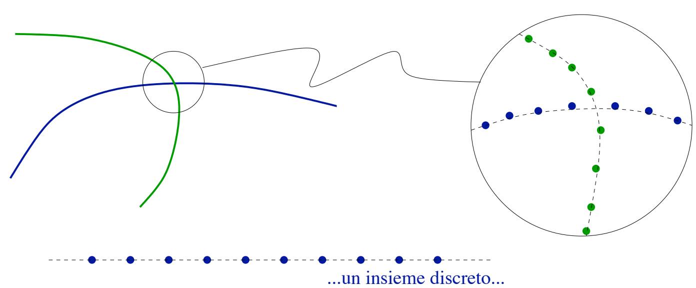

FIGURA 9. Versione "quantizzata": una retta è un'unione di punti a distanza fissata.

Versione "razionale". Questa volta, immaginaniamo 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 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".

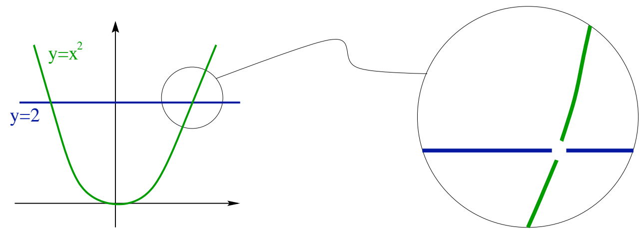

FIGURA 10. Versione "razionale": la parabola $y = x^2$ e la retta $y = 2$ nel piano (x, y) con $x, y \in \mathbb{R}$ non si intersecano mai.

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.

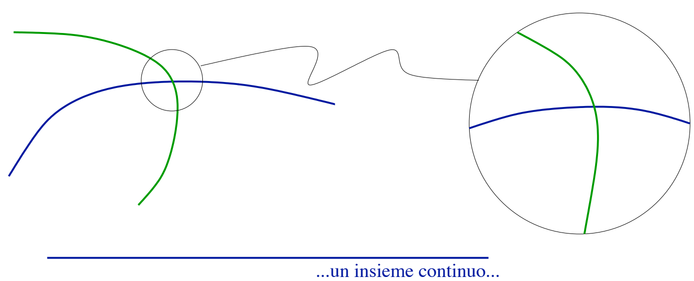

FIGURA 11. Versione "reale": l'idea intuitiva di "continuo".

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 libri[8].

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 $I_0, I_1, \ldots, I_n, \ldots$ <u>chiusi e limitati</u> 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:

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

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.*

L'assioma di Archimede. C'è un'altra proprietà che abbiamo usato senza grande preoccupazione perché estremamente intuitiva: 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. Evidente, vero? Si, ma tanto quanto la proprietà di "non avere buchi", quindi, anche per questo fatto 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,*
>
> $$\forall\, a \in \mathbb{R}, \qquad \exists\, n \in \mathbb{N}, \qquad \textit{tale che} \quad a < n.$$

[8] Si veda, ad esempio, E.Giusti, "Esercizi e Complementi di Analisi Matematica, vol. I", Bollati Boringhieri, cap.2.

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

PROPOSIZIONE 4.1. *Vale la seguente implicazione*

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

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

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

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. <u>*Massimo e minimo.*</u> *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. <u>*Estremi superiore e inferiore.*</u> *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 superiore 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.

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!

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`.

$$\texttt{Input} \longrightarrow \boxed{\textsf{Funzione}} \longrightarrow \texttt{Output}$$

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

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

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:

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

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* $f : I_f \subset \mathbb{R} \to \mathbb{R}$ *è una* restrizione *della funzione* $g : I_g \subset \mathbb{R} \to \mathbb{R}$ *(e* g *è una* estensione *di* f*) se l'insieme di*

[1]In 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

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

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:

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

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)).

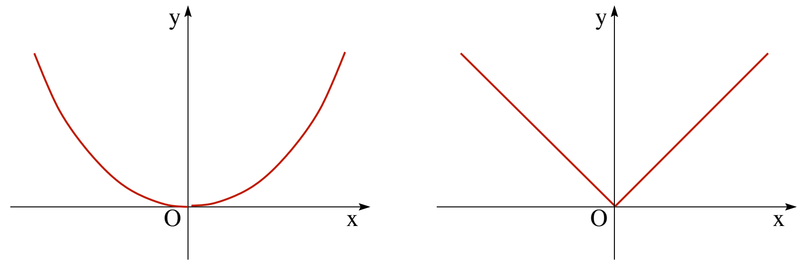

Figura 1. (a) La parabola $y = x^2$; (b) Il modulo: $y = |x|$.

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

$$y = a_0 + a_1 x + \cdots + a_n x^n \qquad a_n \neq 0.$$

dove $a_0, a_1, \ldots, a_n$ (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*

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

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

$$y = \frac{a_0 + a_1 x + \cdots + a_n x^n}{b_0 + b_1 x + \cdots + b_m x^m} \quad \text{con } a_i, b_j \in \mathbb{R}, \quad (b_j \text{ non tutti nulli})$$

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

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

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)).

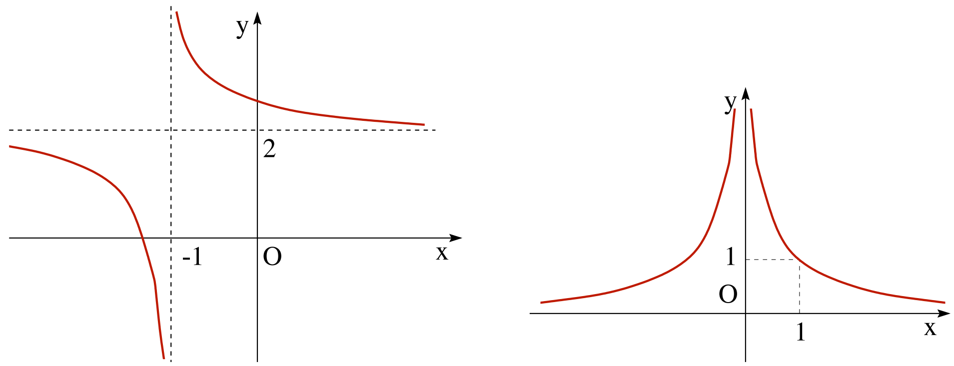

FIGURA 2. (a) La funzione $y = (2x+3)/(x+1)$; (b) la funzione $y = 1/x^2$.

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

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

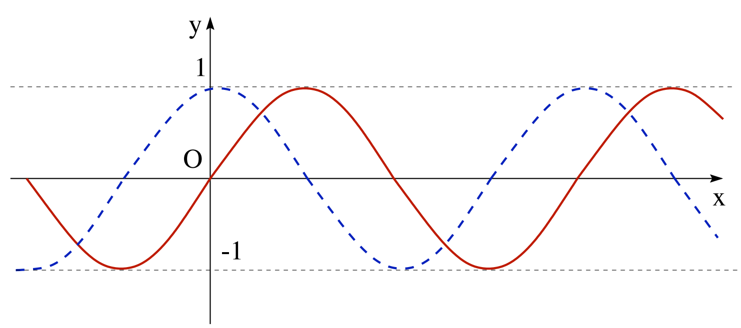

FIGURA 3. Il grafico della funzione $\sin x$ (linea continua) e della funzione $\cos x$ (linea tratteggiata).

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

$$|\sin x| \le 1, \qquad |\cos x| \le 1 \qquad\qquad \forall\, x \in \mathbb{R}.$$

Esercizio 1.4. *Dimostrare che* $|\sin x| \le |x|$ *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:

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

Dalla definizione e dalle proprietà di seno e coseno, discende che $\tan x$ è definita per $x \ne \frac{\pi}{2} + k\pi$ per $k \in \mathbb{Z}$ e $\cot x$ è definita per $x \ne k\pi$ 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 $c < 0$.

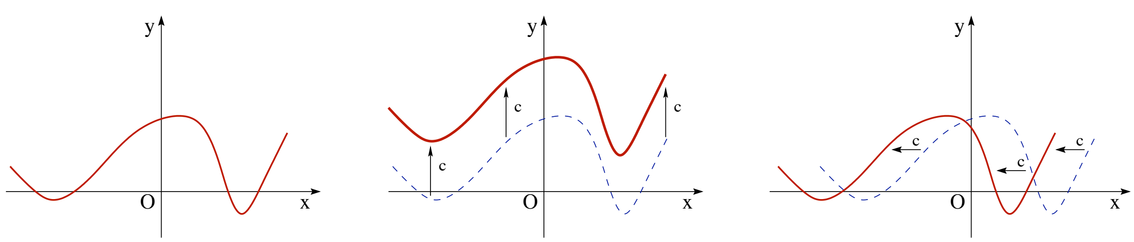

Figura 4. I grafici di (a) $y = f(x)$, (b) $y = f(x) + c$; e (c) $y = f(x + c)$.

(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

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

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

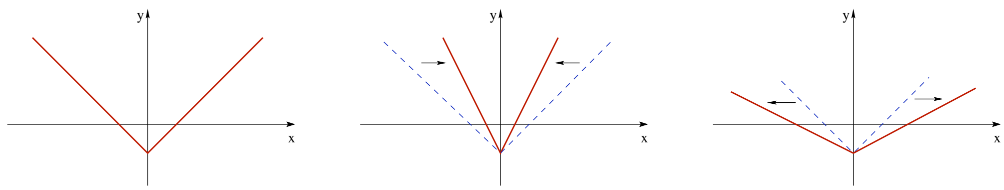

FIGURA 5. I grafici di (a) $y = f(x) = |x| - 1$, (b) $g(x) = |2x| - 1$, (c) $h(x) = \left|\frac{x}{2}\right| - 1$.

ESERCIZIO 2.1. *Disegnare i grafici delle funzioni*

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

(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 $t \to s = 1/t$ 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.

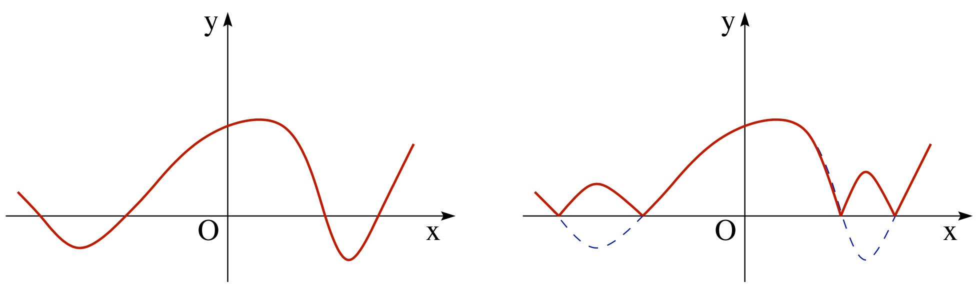

FIGURA 6. I grafici di (a) $y = f(x)$, (b) $y = |f(x)|$.

(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 negativo 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 $y \geq 0$.

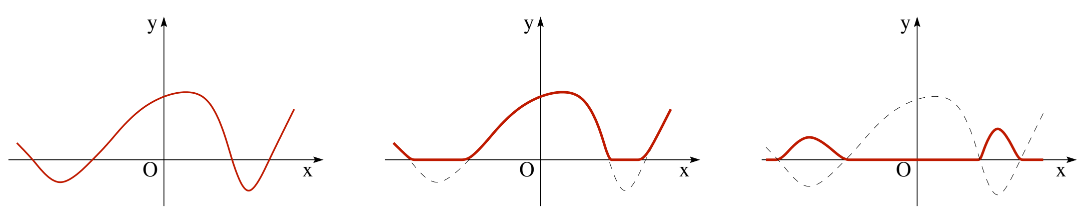

FIGURA 7. I grafici di (a) $y = f(x)$, (b) $y = \max\{f(x), 0\}$, (c) $y = \max\{-f(x), 0\}$,.

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

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

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

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

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

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?

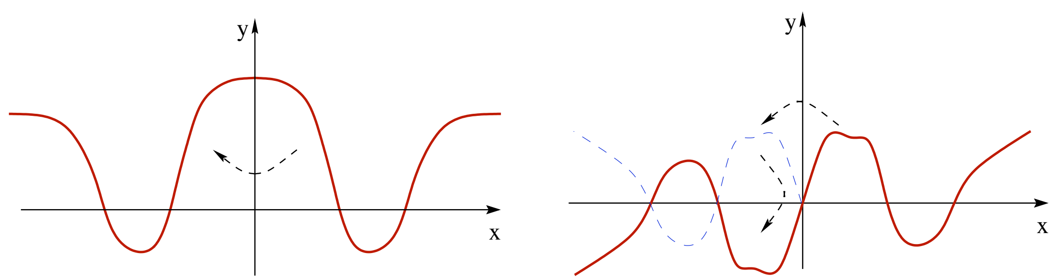

FIGURA 8. (a) una funzione pari, (b) una funzione dispari.

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

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

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: $\sin x$, $\cos x$, $\tan x$, $\cot x$,

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

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

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?*

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

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

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

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

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

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

3. Funzioni invertibili e funzioni monotòne

Data $f : I \subset \mathbb{R} \to \mathbb{R}$ 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 $f : I \subset \mathbb{R} \to \mathbb{R}$ 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*

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

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 $x_1 = x_2$:

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

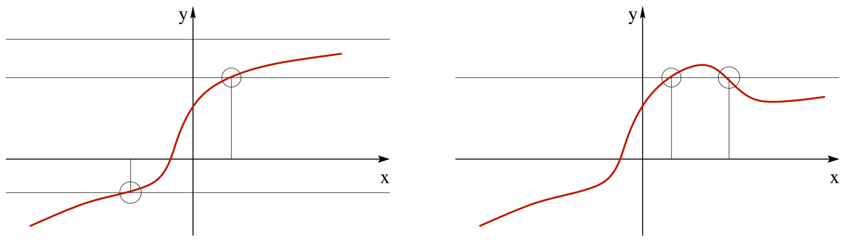

FIGURA 9. (a) una funzione iniettiva, (b) una funzione non iniettiva.

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

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

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

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

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

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

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:

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

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

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

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 <u>unico</u> 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 $f^{-1} : f(I) \subset \mathbb{R} \to \mathbb{R}$ che gode della proprietà*

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

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

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

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

Esempio 3.7. Per la funzione

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

si ha

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

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

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

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

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

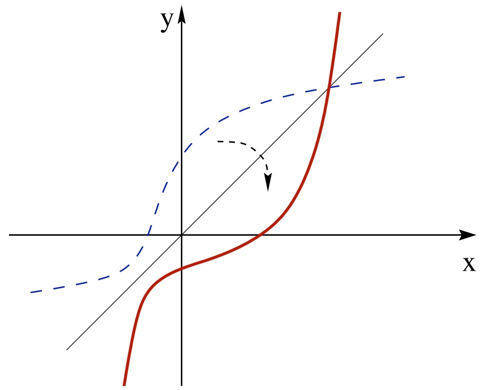

FIGURA 10. Il grafico della funzione inversa di una funzione assegnata.

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

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

Per ogni $n \in \mathbb{N}$, la funzione $y = x^n$ ristretta a $x \geq 0$ è una funzione monotòna 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 monotòna 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

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

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

$$f \textit{ strettamente monotòna} \quad \Rightarrow \quad f \textit{ iniettiva} \quad \iff \quad f \textit{ invertibile}$$

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

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

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 monotonía, 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 $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ 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 $a > 0$:

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

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

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

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 proprietá

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

Dato che a^x è monotòna 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 proprietá

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

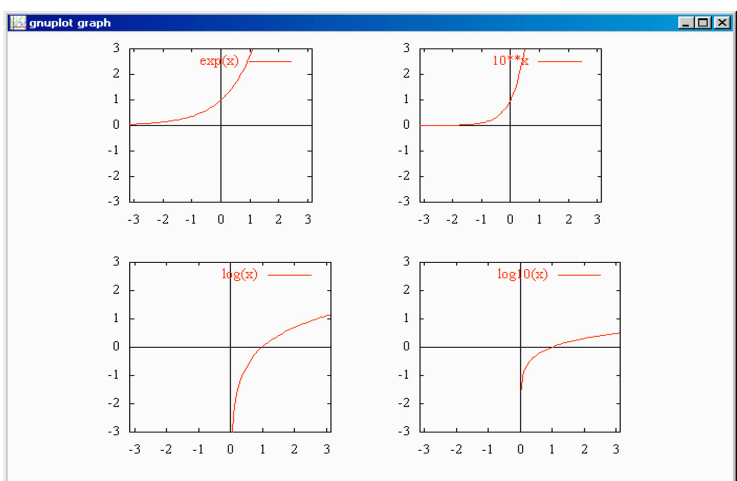

FIGURA 11. (a) $y = e^x$, (b) $y = 10^x$, (c) $y = \ln x$, (d) $y = \log_{10} x$

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 $-1 = (-1)^1 = (-1)^{2\frac{1}{2}} = [(-1)^2]^{\frac{1}{2}} = 1^{\frac{1}{2}} = 1$. Ops!

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

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

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

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

Analogamente

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

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

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

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

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

La composizione ha senso anche per funzioni non definite in tutto $\mathbb{R}$. Sia $\phi : I \subset \mathbb{R} \to \mathbb{R}$ e $g : J \subset \mathbb{R} \to \mathbb{R}$, 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 $J = (0, +\infty)$.

[2] Alle 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 un 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

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

La funzione $\chi_{\mathbb{R}}$ vale identicamente 1, mentre $\chi_{\emptyset}$ 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,

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

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

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 $f : I \subset \mathbb{R} \to \mathbb{R}$. 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*

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

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*

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

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 $f(I)$:

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

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 $y = f(x_0) =$ 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 $f(I)$

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

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* $f : I \subset \mathbb{R} \to \mathbb{R}$ *tale che* $\inf\limits_{x\in I} f(x) \in \mathbb{R}$ *e* $\sup\limits_{x\in I} f(x) \in \mathbb{R}$ *(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 $\{(x, y) \in \mathbb{R}^2 : c \le y \le d\}$ 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

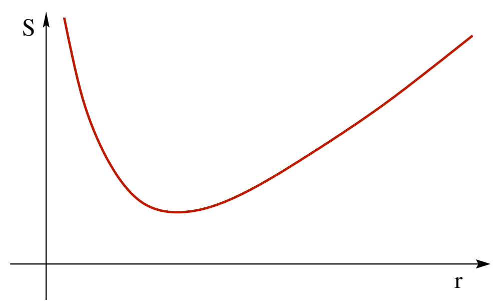

FIGURA 12. Il grafico di $S(r) = 2\pi\left(r^2 + \frac{V}{\pi r}\right)$.

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 $V =$ costante. Quindi $h = V/\pi r^2$, e, sostituendo in S,

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

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 $a : \mathbb{N} \to \mathbb{R}$ 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

$$a_0,\ a_1,\ a_2,\ a_3,\ \ldots$$

Un primo esempio è la successione dei numeri pari: $2, 4, 6, \ldots$. 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 $1, 1, 2, 6, 24, 120, \ldots$ (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

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

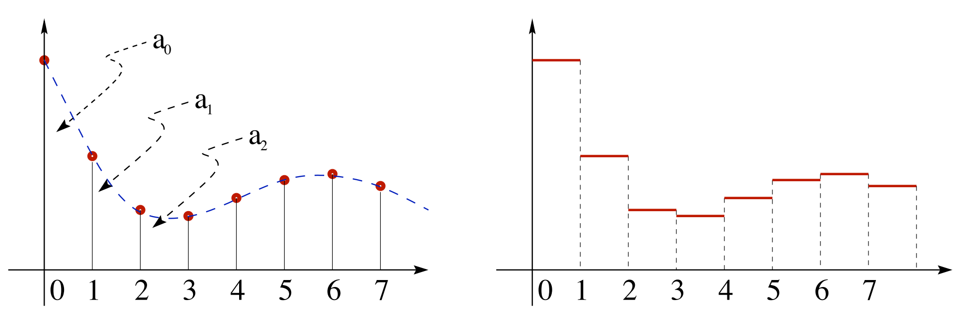

FIGURA 1.

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

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

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 $n \to +\infty$, ovvero tende ad avvicinarsi ad un valore ℓ fissato? In caso affermativo, si dice che la successione ammette limite ℓ per $n \to +\infty$, 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

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

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

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

In questo caso si dice che a_n tende a 0 per $n \to +\infty$ (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 $a_n = \dfrac{n}{n+1}$. 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 $n \to +\infty$, la distanza di a_n da 1 tende a zero, cioè il valore a_n si avvicina ad 1. Anche la successione $a_n = \dfrac{n^2 - 1}{n^2 + n + 1}$ si comporta in modo analogo, infatti:

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

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

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

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

Il valore ℓ è il limite della successione *a_n.*

Se $a_n \to 0$ per $n \to +\infty$ (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.

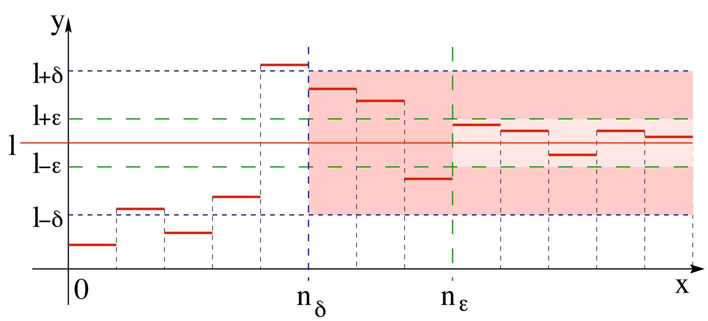

Figura 2.

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 $n \to +\infty$ 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 Avogadro*[1]*. Dimostrare che anche b_n converge ad ℓ per $n \to +\infty$.*

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,

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

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

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

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 $n \to +\infty$ 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

[1] Per 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:

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

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 validitá di*

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

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

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, ... costituiscono una sottosuccessione di a_1, a_2, a_3, a_4, ..., così come gli elementi di indice pari a_0, a_2, a_4, a_6, ... 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 qualè 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 $k = 2$ 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* $a_n = n$

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

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 $n \to +\infty$. Allora ogni sua sottosuccessione a_{n_k} converge allo stesso limite ℓ per $k \to +\infty$.*

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 $k \to +\infty$. Invece, la sottosuccessione dei termini di indice dispari è $a_{2k+1} = (-1)^{2k+1} = -1$ per ogni k, quindi converge a -1 per $k \to +\infty$. 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 $\varepsilon > 0$,

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

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

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

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

Definizione 1.10. *Una successione a_n tale che esista $M > 0$ per cui $|a_n| \le M$ 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 $n \to +\infty$ 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*

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

In modo equivalente si può scrivere

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

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

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

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

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

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

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

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*

$$\lim_{n\to+\infty} a_n = 0 \qquad \Rightarrow \qquad \lim_{n\to+\infty} \frac{1}{|a_n|} = +\infty.$$

$$\lim_{n\to+\infty} |a_n| = +\infty \qquad \Rightarrow \qquad \lim_{n\to+\infty} \frac{1}{a_n} = 0.$$

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 $n \to +\infty$ 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 ...

... è convergente o divergente (a $\pm\infty$), allora è regolare;

... non è convergente né divergente, allora è non regolare;

... tende a 0 per $n \to +\infty$, 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 $\lim_{n\to+\infty} |a_n| = +\infty$, 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:

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

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

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

(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*:

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

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 $a_n \to a$ e $b_n \to b$ per $n \to +\infty$. Allora[2]

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

[2]Il 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 $ab - a_n b_n = b(a - a_n) + a_n(b - b_n)$, 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

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

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

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

$$\lim_{n\to+\infty} \frac{2n^2 - 1}{3n^2 + n + 1} = \lim_{n\to+\infty} \frac{2 - \frac{1}{n^2}}{3 + \frac{1}{n} + \frac{1}{n^2}} = \frac{\lim\limits_{n\to+\infty} \left(2 - \frac{1}{n^2}\right)}{\lim\limits_{n\to+\infty} \left(3 + \frac{1}{n} + \frac{1}{n^2}\right)} = \frac{2}{3},$$

passando al limite sia nel numeratore che nel denominatore.

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

Esercizio 2.1. *Calcolare il limite per $n \to +\infty$ delle seguenti successioni*

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

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

Teorema 2.2. Monotonìa del limite. *Sia $\lim\limits_{n\to+\infty} a_n = a$ e $\lim\limits_{n\to+\infty} b_n = b$. 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 diseguaglianze, 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 "monotòna non decrescente", nel senso che vale l'implicazione

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

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 $\lim\limits_{n\to+\infty} a_n = \lim\limits_{n\to+\infty} b_n = 0$.

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

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

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 $n > \tilde{n}_\varepsilon$,

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

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

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

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!),

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

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

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

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*

$$\lim_{n\to+\infty} \frac{a_{n+1}}{a_n} = \lambda \tag{10}$$

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

$$a_{n+1} \leq \sigma a_n \qquad \forall\, n \geq N. \tag{11}$$

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

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

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 $n \to +\infty$ la quantità $\sigma^n a_1$ tende a zero e quindi si ha la conclusione. Nel caso generale, dato che

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

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 $1 + nh \to +\infty$ per $n \to \infty$, 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 $\dfrac{a_n}{b_n}$ nel caso in cui $\lim\limits_{n\to+\infty} b_n = 0$?

– 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:

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

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

$$\textit{Congettura 1:} \qquad \lim_{n\to+\infty} b_n = 0 \qquad \Rightarrow \qquad \lim_{n\to+\infty} \frac{a_n}{b_n} = +\infty$$

Ma se consideriamo il caso

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

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:

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

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

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

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

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

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

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

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

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

Credergli o non credergli?

Resta fuori dalla nostra analisi il caso in cui sia il numeratore che il denominatore tendano a 0 per $n \to +\infty$. 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):

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

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

$$\lim_{n\to+\infty} a_n = +\infty \qquad \text{e} \qquad \lim_{n\to+\infty} b_n = b.$$

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:

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

(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

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

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 $+\infty$:

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

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

$$\textbf{forma indeterminata} \quad +\infty - \infty\ : \qquad \left.\begin{array}{l} \lim\limits_{n\to+\infty} a_n = +\infty, \\ \lim\limits_{n\to+\infty} b_n = -\infty \end{array}\right\} \quad \Rightarrow \quad \lim_{n\to+\infty} a_n + b_n = ?$$

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,

$$\lim_{n\to+\infty} a_n = \lim_{n\to+\infty} b_n = +\infty \qquad \Rightarrow \qquad \lim_{n\to+\infty} a_n \cdot b_n = +\infty.$$

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

$$\textbf{forma indeterminata}\quad \frac{\infty}{\infty}\ : \qquad \left.\begin{array}{l}\lim\limits_{n\to+\infty} a_n = \infty,\\ \lim\limits_{n\to+\infty} b_n = \infty\end{array}\right\} \quad\Rightarrow\quad \lim_{n\to+\infty}\frac{a_n}{b_n} = ?$$

3. Calcolo di alcuni limiti

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

Esempio 3.1.

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

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,

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

da cui segue $\lim\limits_{n\to+\infty} h_n = 0$ 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 $\sqrt[n]{x} = \dfrac{1}{\sqrt[n]{1/x}}$, segue la conclusione.

Esempio 3.2.

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

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 $n \to +\infty$.

Esempio 3.3. (Molto importante!)

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

[3] In 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 $\lim_{n\to+\infty} x^n = 0$ per $x \in (0,1)$ e che $\lim_{n\to+\infty} x^n = +\infty$ per $x > 1$ Anche per $x \in (-1,0]$, la successione x^n è infinitesima, dato che $\lim_{n\to+\infty} |x|^n = 0$.

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.

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

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

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

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.

$$\lim_{n\to+\infty} \frac{x^n}{n!} = 0 \qquad \forall\, x > 1. \tag{12}$$

Applichiamo di nuovo il Corollario 2.6

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

da cui segue la conclusione.

Esempio 3.6.

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

Infatti

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

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

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

Infatti, la successione S_n si può riscrivere nella forma

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

Passando al limite per $n \to +\infty$ 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

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

Sul significato della parola "serie" in generale torneremo tra poco.

4. Successioni monotòne

DEFINIZIONE 4.1. <u>Successioni monotòne.</u> *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 monotòna *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 monotòna se è o non decrescente o non crescente.

Per verificare se una successione è monotona occorre risolvere delle disequazioni. Ad esempio, la successione $a_n = \dfrac{1}{1+n^2}$ 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 postive 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 monotóna, 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. <u>Regolarità delle successioni monotòne.</u> *Una successione a_n monotòna è sempre regolare (cioè o è convergente o è divergente) e*

$$a_n \text{ non decrescente} \quad \Rightarrow \quad \lim_{n\to+\infty} a_n = \sup_{n\in\mathbb{N}} a_n,$$
$$a_n \text{ non crescente} \quad \Rightarrow \quad \lim_{n\to+\infty} a_n = \inf_{n\in\mathbb{N}} a_n.$$

Se a_n è anche limitata, allora è convergente.

DIMOSTRAZIONE. Supponiamo a_n non decrescente e superiormente limitata,

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

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 piú piccolo!). Usando la monotonia si ha

$$\ell - \varepsilon < a_N \le a_n < \ell \qquad \forall n \ge N,$$

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

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

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*:

condizione necessaria di convergenza : $\displaystyle\sum_{n=0}^{\infty} a_n$ convergente $\Rightarrow$ $\displaystyle\lim_{n\to+\infty} a_n = 0.$

Infatti, sia $\sum_{n=0}^{\infty} a_n$ convergente, cioè esista finito $\lim_{n\to+\infty} S_n = S$. Dato che $a_n = S_n - S_{n-1}$ e, per ipotesi, la successione S_n converge a S per $n \to +\infty$, allora

$$\lim_{n\to+\infty} a_n = \lim_{n\to+\infty} (S_n - S_{n-1}) = \lim_{n\to+\infty} S_n - \lim_{n\to+\infty} S_{n-1} = S - S = 0.$$

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,

$$a_k \geq 0 \qquad \Rightarrow \qquad \sum_{k=0}^{+\infty} a_k < +\infty \qquad \text{oppure} \qquad \sum_{k=0}^{+\infty} a_k = +\infty.$$

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

$$\lim_{k\to+\infty} \frac{a_k}{b_k} = \ell \geq 0 \qquad \text{e} \qquad \sum_{k=1}^{\infty} b_k < +\infty,$$

Per definizione di limite, scelto $\varepsilon > 0$, esiste $N \in \mathbb{N}$ opportuno per cui, per ogni $n \geq N$,

$$\ell - \varepsilon < \frac{a_k}{b_k} < \ell + \varepsilon \quad \Rightarrow \qquad a_k < (\ell + \varepsilon) b_k.$$

Quindi, applicando il precedente risultato, si dimostra che la serie $\sum\limits_{k=0}^{\infty} a_k$ 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),

$$\lim_{k\to+\infty} \frac{x^k/k!}{y^k} = \lim_{k\to+\infty} \frac{(x/y)^k}{k!} = 0,$$

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!)

$$\text{convergenza assoluta} \qquad \Rightarrow \qquad \text{convergenza semplice}$$

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:

$$0 \leq a^+ \leq |a|, \qquad 0 \leq a^- \leq |a|, \qquad a = a^+ - a^-, \qquad |a| = a^+ + a^-.$$

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 $a_k = (-1)^k/(k+1)$,

$$\begin{array}{llllllll}
a_k : & 1, & -\frac{1}{2}, & \frac{1}{3}, & -\frac{1}{4}, & \frac{1}{5}, & -\frac{1}{6}, \dots \\
a_k^+ : & 1, & 0, & \frac{1}{3}, & 0, & \frac{1}{5}, & 0, \dots \\
a_k^- : & 0, & \frac{1}{2}, & 0, & \frac{1}{4}, & 0, & \frac{1}{6} \dots .
\end{array}$$

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 applicale 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. □

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 $x \in \mathbb{R}$, considerariamo la serie

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad x \in \mathbb{R}.$$

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

$$\text{funzione esponenziale:} \qquad e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad x \in \mathbb{R}. \tag{13}$$

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

$$\left| e - \sum_{k=0}^{n} \frac{1}{k!} \right| \le \frac{1}{n!\,n}.$$

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	$0,25$	$0,0\bar{5}$	$0,010412$	$0,001\bar{6}$	$0,000231$	$0,000028$

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 $\dfrac{1}{n}$, che è detta serie armonica. Questa serie è divergente

$$\sum_{n=1}^{\infty} \frac{1}{n} = +\infty. \tag{14}$$

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

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha} \qquad \alpha > 0,$$

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.

CAPITOLO 4

Le funzioni continue

Dopo l'*excursus* del Capitolo precedente sulle successioni numeriche, torniamo a parlare di funzioni reali di variabile reale in generale. Per fissare le idee, supponiamo di voler studiare funzioni f, definite in $I \subset \mathbb{R}$, dove I è un intervallo (limitato o illimato) di $\mathbb{R}$. L'obiettivo principale del Capitolo è definire il significato della parola *continuità*.

1. Limite di funzioni

Tutto nasce dalla definizione di "limite". Come abbiamo visto per le successioni, il limite formalizza l'idea di "previsione" del comportamento di un oggetto sotto osservazione per opportuni valori della variabile.

Limiti all'infinito. Partiamo prima di tutto dal concetto di funzione infinitesima per $x \to +\infty$. Una funzione $d : \mathbb{R} \to \mathbb{R}$ è infinitesima per $x \to +\infty$, se

$$\forall \varepsilon > 0 \quad \exists M \quad \text{tale che} \quad |d(x)| < \varepsilon \quad \forall x > M. \tag{15}$$

Il numero $M \in \mathbb{R}$ dipende dalla scelta di ε (come n_ε per le successioni): $M = M(\varepsilon)$.

La proprietà $|d(x)| < \varepsilon$ (equivalente a $-\varepsilon < d(x) < \varepsilon$) indica che il grafico della funzione d vive nella striscia infinita delimitata dalle retta $y = -\varepsilon$ e $y = \varepsilon$ per x sufficientemente grandi (Fig.1(a)), quindi la condizione (15) significa che il grafico della funzione d "tende a confondersi" con l'asse x per $x \to +\infty$.

Data una funzione $f : \mathbb{R} \to \mathbb{R}$, questa tende ad un limite ℓ per $x \to +\infty$ se è la funzione $f(x) - \ell$ ad essere infinitesima.

Definizione 1.1. *Data $f : \mathbb{R} \to \mathbb{R}$, si dice che f* converge ad $\ell \in \mathbb{R}$ per $x \to +\infty$

$$\lim_{x \to +\infty} f(x) = \ell, \tag{16}$$

se $|f(x) - \ell|$ è infinitesima per $x \to +\infty$, cioè se

$$\forall \varepsilon > 0 \quad \exists M \quad \textit{tale che} \quad |f(x) - \ell| < \varepsilon \quad \forall x > M.$$

La funzione $d(x) := |f(x) - \ell|$ rappresenta la lunghezza del segmento verticale di estremi $(x, f(x))$ e (x, ℓ) e "tendere ad ℓ" indica che tale lunghezza tende a zero.

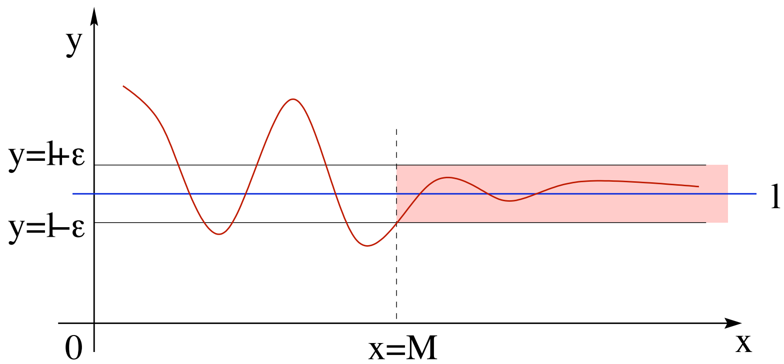

FIGURA 1. Una funzione che tende ad un limite per $x \to +\infty$.

Buona parte di quanto visto per le successioni si può ripetere. Ad esempio,

$$\lim_{x \to +\infty} \frac{x^2}{1+x^2} = 1.$$

Infatti, per ogni $\varepsilon \in (0,1)$, si ha

$$\left| \frac{x^2}{1+x^2} - 1 \right| = \frac{1}{1+x^2} < \varepsilon \qquad \forall x > M := \sqrt{\frac{1}{\varepsilon} - 1}.$$

OSSERVAZIONE 1.2. Nella definizione di limite di funzione per $x \to +\infty$, siamo partiti da una funzione f definita in tutto $\mathbb{R}$. Per definire il limite per $x \to +\infty$, basta anche di meno: l'unica cosa indispensabile è che l'insieme di definizione sia non limitato superiormente. Pensate al caso delle successioni: sono funzioni definite su $\mathbb{N}$ (e quindi non su una semiretta) e il limite per $n \to +\infty$ ha perfettamente senso.

Analogamente si possono definire anche:

– limiti divergenti: $\lim\limits_{x \to +\infty} f(x) = +\infty$ o $= -\infty$;

– limiti per x che tende a $-\infty$: $\lim\limits_{x \to -\infty} f(x)$

Limiti in un punto. Per funzioni definite in intervalli, è possibile parlare di limite in un punto. Procediamo come in precedenza chiarendo prima il concetto di "funzione infinitesima in un punto" e poi il concetto di "limite di funzione in un punto".

Sia $x_0 \in [a,b]$. Una funzione $d : (a,b) \to \mathbb{R}$ è infinitesima per $x \to x_0$ se

$$\forall \varepsilon > 0 \quad \exists \delta = \delta(\varepsilon) > 0 \quad \text{t.c.} \quad |d(x)| < \varepsilon \quad \forall x \in (a,b),\ 0 < |x - x_0| < \delta. \tag{17}$$

Rispetto alla definizione di funzione infinitesima per $x \to +\infty$, l'unica differenza sta nelle scelte di x per cui è soddisfatta la condizione $-\varepsilon < d(x) < \varepsilon$. In questo caso si tratta di tutti i valori x, diversi da x_0, che distano da x_0 meno di $\delta > 0$.

Definizione 1.3. *Sia $f : (a,b) \to \mathbb{R}$ e sia $x_0 \in [a,b]$. La funzione f* tende ad $\ell \in \mathbb{R}$ per *$x \to x_0$ se $f(x) - \ell$ è infinitesima per $x \to x_0$, cioè se*

$$\forall\, \varepsilon > 0 \quad \exists\, \delta = \delta(\varepsilon) > 0 \qquad t.c. \qquad |f(x) - \ell| < \varepsilon \quad \forall\, x \in (a,b),\ 0 < |x - x_0| < \delta.$$

In questo caso, si scrive

$$\lim_{x \to x_0} f(x) = \ell. \tag{18}$$

Il punto fondamentale è nella definizione di funzione infinitesima: per dimostrare che il limite della funzione è ℓ bisogna verificare che la distanza tra $f(x)$ e ℓ, cioè la quantità $d(x) := |f(x) - \ell|$, diventa piccola quando x è sufficientemente vicino a x_0.

Esempio 1.4. Proviamo a dimostrare il limite

$$\lim_{x \to 1} (3x - 5) = -2.$$

In questo caso $f(x) = 3x - 5$, $x_0 = 1$ e $\ell = -2$. Per definizione, basta mostrare che la quantità $|f(x) - \ell|$ è infinitesima per x che tende ad 1. Poniamoci quindi l'obiettivo di stimarla in termini di una funzione in cui compaia la distanza $|x - 1|$:

$$|f(x) - \ell| = |(3x - 5) - (-2)| = |3x - 3| = 3|x - 1|.$$

Perfetto! Da queste uguaglianze segue la conclusione. Se vogliamo conoscere esplicitamente il valore di δ in funzione di ε, così come richiesto dalla definizione, basta osservare che se $|x - 1| < \delta$, allora $|f(x) - \ell| < 3\delta$, quindi dato $\varepsilon > 0$, basta scegliere δ in modo che $3\delta = \varepsilon$, cioè $\delta = \varepsilon/3$.

Esempio 1.5. Fissato $x_0 \in \mathbb{R}$, calcoliamo $\lim\limits_{x \to x_0} \sin x$. E' ragionevole aspettarsi che tale limite esista e che valga $\ell = \sin x_0$, quindi proviamo a stimare $|\sin x - \sin x_0|$. Usando una delle (diaboliche) formule di prostaferesi,

$$|\sin x - \sin x_0| = \left| 2 \sin\left(\frac{x - x_0}{2}\right) \cos\left(\frac{x + x_0}{2}\right) \right| \leq 2 \left| \sin\left(\frac{x - x_0}{2}\right) \right|.$$

Dato che $|\sin x| \leq |x|$ per ogni $x \in \mathbb{R}$ (Cap. 2, Es. 1.4), si ottiene

$$|\sin x - \sin x_0| \leq 2 \left| \frac{x - x_0}{2} \right| = |x - x_0|,$$

da cui segue

$$\lim_{x \to x_0} \sin x = \sin x_0.$$

Volendo determinare esplicitamente δ, dato $\varepsilon > 0$, basta scegliere $\delta := \varepsilon$ per fare in modo che, se $|x - x_0| < \delta$, allora $|\sin x - \sin x_0| < \varepsilon$.

Esercizio 1.6. *Dimostrare che, per ogni $x_0 \in \mathbb{R}$, vale $\lim\limits_{x \to x_0} \cos x = \cos x_0$.*

Esempio 1.7. Vediamo un limite più complicato:

$$\lim_{x\to 0} e^x = 1.$$

Euristicamente il risultato è più che ragionevole, dato che, dalla definizione dell'esponenziale data in (13), segue

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

e ciascuno dei termini sommati tende a zero per $x \to 0$. Il problema è che i termini sommati sono infiniti! Per dimostrare in modo rigoroso la validità del limite bisogna, come sempre, stimare il termine $|f(x) - \ell| = |e^x - 1|$:

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

(nella riga precedente ci sono due sono disuguaglianze per serie... perché sono lecite?). Se scegliamo $|x| < 1$, si ha $e^{|x|} \leq e$ (si ricordi che la funzione e^x è crescente), quindi

$$|e^x - 1| \leq |x| e^{|x|} \leq e\,|x| \qquad \forall\, x \in (-1, 1),$$

da cui si arriva alla conclusione.

Osservazione 1.8. *$x \neq x_0$.* I punti x che intervengono nel limite per $x \to x_0$ sono, per definizione, *distinti da* x_0. In parole povere, il limite della funzione f per $x \to x_0$ è il comportamento che si prevede per la funzione f in x_0, in base al grafico della funzione vicino a x_0, ma *indipendentemente* da quello che succede nel punto limite.

A guardare bene, la Definizione 1.3 vale, così com'è per funzioni f definite in $(a, b) \setminus \{x_0\}$ (e $x_0 \in (a, b)$), cioè funzioni che <u>non</u> sono definite nel punto limite! Quello che conta è il punto limite x_0 sia "vicino" a punti in cui la funzione è definita: non ha senso calcolare il limite per $x \to 2$ di una funzione che è definita in $[0, 1]$!

Come si dimostra che un limite non esiste? Dalla definizione di limite, segue che, se $\lim_{x\to x_0} f(x) = \ell$, allora per ogni successione x_n, contenuta nell'insieme di definizione di f, e tale che x_n tende a x_0 per $n \to +\infty$, la successione $f(x_n)$ tende ad ℓ per $n \to +\infty$:

$$\lim_{n\to+\infty} x_n = x_0 \qquad \Rightarrow \qquad \lim_{n\to+\infty} f(x_n) = \ell. \tag{19}$$

Esercizio 1.9. *Dimostrare l'affermazione che avete appena letto.*

Dall'implicazione (19) discende il seguente

Criterio 1.10. <u>*Non esistenza del limite.*</u> *Se esistono due successioni x_n e ξ_n entrambe convergenti a x_0 e tali che*

$$\lim_{n\to+\infty} f(x_n) \neq \lim_{n\to+\infty} f(\xi_n),$$

la funzione f non può ammettere limite per $x \to x_0$.

Un esempio chiarirà meglio le idee. Consideriamo la funzione

$$\textsf{segno di } x: \qquad \operatorname{sgn} x := \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ +1 & x > 0 \end{cases}$$

e consideriamo $x_n = \frac{1}{n}$ e $\xi_n = -\frac{1}{n}$. È evidente che $\lim\limits_{n\to+\infty} x_n = \lim\limits_{n\to+\infty} \xi_n = 0$. Inoltre, per ogni n, $f(x_n) = 1$ e $f(\xi_n) = -1$, quindi

$$\lim_{n\to+\infty} f(x_n) = 1 \neq -1 \lim_{n\to+\infty} f(\xi_n).$$

Pertanto la funzione sgn non ammette limite per $x \to 0$.

Esercizio 1.11. *Dimostrare che* $\sin(1/x)$ *non ammette limite per* $x \to 0$.

Limite destro e limite sinistro. Quando si studia una funzione solo a destra o a sinistra del punto limite x_0 si parla di limite destro e di limite sinistro.

Definizione 1.12. *La funzione f ha* limite destro *uguale ad ℓ per x che tende a x_0, e si scrive $\lim\limits_{x\to x_0^+} f(x) = \ell$, se*

$$\forall\, \varepsilon > 0 \quad \exists\, \delta > 0 \quad t.c. \quad |f(x) - \ell| < \varepsilon \quad \forall\, x \in (a,b),\ x_0 < x < x_0 + \delta.$$

Analogamente per il limite sinistro*, che si indica con $\lim\limits_{x\to x_0^-} f(x) = \ell$.*

Esercizio 1.13. *Dimostrare che* $\lim\limits_{x\to 0^-} \operatorname{sgn} x = -1, \quad \lim\limits_{x\to 0^+} \operatorname{sgn} x = +1.$

Esercizio 1.14. *Calcolare i seguenti limiti*

$$\lim_{x\to 0^-} \arctan\left(\frac{1}{x}\right), \qquad \lim_{x\to 0^+} \arctan\left(\frac{1}{x}\right).$$

Dalla definizione di limite destro e sinistro si deduce (con poca fatica) il seguente:

Criterio 1.15. <u>*Esistenza del limite.*</u> *Una funzione ammette limite in un punto se e solo se esistono sia il limite destro che quello sinistro e coincidono.*

Di conseguenza, se uno tra i limiti destro e sinistro non esiste, o se entrambi esistono, ma non coincidono, la funzione non ha limite per $x \to x_0$. Avendo risolto l'Esercizio 1.14, sapete dire se $\arctan(1/x)$ ammette limite per $x \to 0$?

Limiti e operazioni razionali. Limiti di somme, differenze, prodotti e rapporti di funzioni godono delle stesse proprietà viste per le successioni:

$$\left.\begin{array}{l}\lim\limits_{x\to x_0} f(x)=\ell\\ \lim\limits_{x\to x_0} g(x)=m\end{array}\right\}\quad\Rightarrow\quad\begin{cases}\lim\limits_{x\to x_0}\big(f(x)\pm g(x)\big)=\ell\pm m,\\ \lim\limits_{x\to x_0} f(x)g(x)=\ell\, m,\\ \lim\limits_{x\to x_0}\dfrac{f(x)}{g(x)}=\dfrac{\ell}{m} & \text{se } m\neq 0.\end{cases}\tag{20}$$

Esempio 1.16. Per calcolare

$$\lim_{x\to 1}\frac{3x^2+x-1}{x^5+x^3}$$

non è una buona idea usare la definizione! Basta applicare le regole su descritte:

$$\begin{aligned}\lim_{x\to 1}\frac{3x^2+x-1}{x^5+x^3}&=\frac{\lim\limits_{x\to 1}(3x^2+x-1)}{\lim\limits_{x\to 1}(x^5+x^3)}=\frac{\lim\limits_{x\to 1}3x^2+\lim\limits_{x\to 1}x-\lim\limits_{x\to 1}1}{\lim\limits_{x\to 1}x^5+\lim\limits_{x\to 1}x^3}\\&=\frac{(\lim\limits_{x\to 1}3)(\lim\limits_{x\to 1}x)^2+\lim\limits_{x\to 1}x-\lim\limits_{x\to 1}1}{(\lim\limits_{x\to 1}x)^5+(\lim\limits_{x\to 1}x)^3}=\frac{3+1-1}{1+1}=\frac{3}{2}.\end{aligned}$$

Ora, se volete, provate a dimostrare il risultato usando solo la definizione di limite...

Esempio 1.17. Fissato $x_0\in\mathbb{R}$, calcoliamo

$$\lim_{x\to x_0}\frac{x^2-x_0^2}{x-x_0}.$$

Qui non è possibile applicare direttamente le regole viste, perché il denominatore tende a zero per $x\to x_0$. Ma basta una riga di conto per risolvere il problema:

$$\lim_{x\to x_0}\frac{x^2-x_0^2}{x-x_0}=\lim_{x\to x_0}\frac{(x-x_0)(x+x_0)}{x-x_0}=\lim_{x\to x_0}(x+x_0)=\lim_{x\to x_0}x+\lim_{x\to x_0}x_0=2x_0.$$

Analogamente, si dimostra che

$$\lim_{x\to x_0}\frac{x^3-x_0^3}{x-x_0}=3x_0^2,$$

utilizzando l'identità $x^3-x_0^3=(x^2+xx_0+x_0^2)(x-x_0)$. In generale, vale

$$\lim_{x\to x_0}\frac{x^n-x_0^n}{x-x_0}=nx_0^{n-1}\qquad n\in\mathbb{N},$$

infatti

$$\frac{x^n-x_0^n}{x-x_0}=x^{n-1}+x^{n-2}x_0+\cdots+xx_0^{n-2}+x_0^{n-1}.$$

Limiti e disequazioni. Anche per il rapporto tra limiti e disequazioni valgono le stesse regole già viste nel caso delle successioni: supponiamo che le funzioni f e g abbiano limite per $x \to x_0$, allora

$$f(x) < g(x) \qquad (\text{o} \quad f(x) \leq g(x)) \qquad \Rightarrow \quad \lim_{x \to x_0} f(x) \leq \lim_{x \to x_0} g(x).$$

La disuguaglianza stretta diviene una disuguaglianza debole. Le dimostrazioni sono analoghe a quelle per le successioni.

Da queste proprietà discende la seguente proposizione (analoga al Teorema 2.3).

PROPOSIZIONE 1.18. *Siano f, g, h tre funzioni tali che $f(x) \leq g(x) \leq h(x)$ per tutti i valori x in un intorno di x_0. Allora*

$$\lim_{x \to x_0} f(x) = \ell = \lim_{x \to x_0} h(x) \qquad \Longrightarrow \qquad \lim_{x \to x_0} g(x) = \ell.$$

Omettiamo la dimostrazione.

Zeri a denominatore ed uso degli infiniti. Anche per quanto riguarda quest'argomento, quello che c'è da capire è interamente contenuto nel caso delle successioni. In particolare, le forme indeterminate che si incontrano con più frequenza sono: $\frac{0}{0}$, $\infty \cdot 0$, $+\infty - \infty$, $\frac{\infty}{\infty}$.

Alcuni limiti notevoli. Il fatto che il limite sia compatibile con le operazioni di somma e prodotto fa in modo che nel calcolo effettivo dei limiti, nella maggior parte dei casi, non si debba utilizzare direttamente la definizione (con conseguente calcolo di ε e $\delta(\varepsilon)$, spesso tremendamente complicato), ma ci si possa ricondurre a limiti già noti. Il problema, a questo punto, è che di limiti noti ne abbiamo pochini... Corriamo al mercato ad acquistarne un po'.

ESEMPIO 1.19. Partiamo da un limite che non può mancare nella casa di nessuno:

$$\lim_{x \to 0} \frac{\sin x}{x} = 1, \tag{21}$$

(il valore x, come sempre, è calcolato in radianti). Dal significato geometrico di $\sin x$ si deduce immediatamente che

$$\sin x < x < \tan x \qquad \forall\, x \in (0, \pi/2).$$

Ne segue che, per ogni $x \in (0, \frac{\pi}{2})$,

$$1 < \frac{x}{\sin x} < \frac{1}{\cos x} \quad \Rightarrow \quad \cos x < \frac{\sin x}{x} < 1$$

Dato che $\cos x$ tende a $\cos 0 = 1$ per $x \to 0$, il rapporto $\dfrac{\sin x}{x}$ tende ad 1 per $x \to 0^+$. Lo stesso vale anche per $x \to 0^-$, dato che la funzione $\frac{\sin x}{x}$ è una funzione pari (verificare!). Quindi, per il Criterio 1.15, il gioco è fatto.

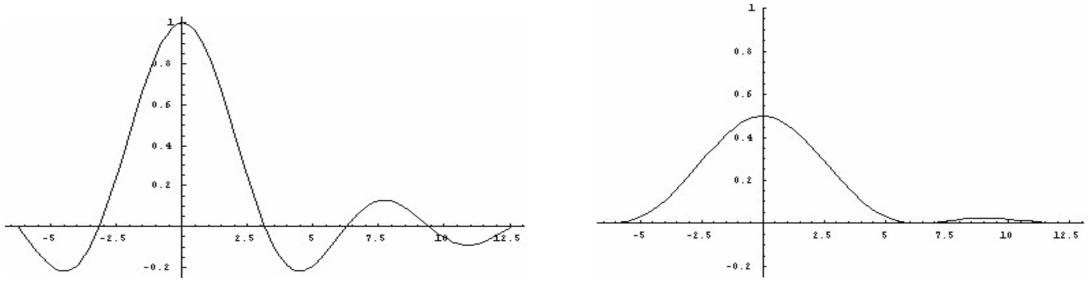

FIGURA 2. (a) $y = \frac{\sin x}{x}$, (b) $y = \frac{1-\cos x}{x^2}$

ESERCIZIO 1.20. *Utilizzando (21), dimostrare*

$$\lim_{x\to 0} \frac{\tan x}{x} = 1, \qquad \lim_{x\to 0} \frac{1-\cos x}{x} = 0, \qquad \lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2}.$$

Soluzione. Per il primo, basta ricordare la definizione di $\tan x$ e usare le proprietà dei limiti

$$\lim_{x\to 0} \frac{\tan x}{x} = \lim_{x\to 0} \frac{\sin x}{x} \frac{1}{\cos x} = \lim_{x\to 0} \frac{\sin x}{x} \frac{1}{\lim\limits_{x\to 0} \cos x} = 1.$$

Per i restanti due, si può utilizzare l'uguaglianza

$$1 - \cos x = \frac{1-\cos^2 x}{1+\cos x} = \frac{\sin^2 x}{1+\cos x},$$

da cui seguono

$$\lim_{x\to 0} \frac{1-\cos x}{x} = \lim_{x\to 0} \frac{\sin x}{x} \frac{1}{1+\cos x} \sin x = 0.$$
$$\lim_{x\to 0} \frac{1-\cos x}{x^2} = \lim_{x\to 0} \left(\frac{\sin x}{x}\right)^2 \frac{1}{1+\cos x} = \frac{1}{2}.$$

e quindi il risultato.

ESEMPIO 1.21. Una coppia di limiti molto importanti è

$$\lim_{x\to 0} \frac{e^x - 1}{x} = 1, \qquad \lim_{x\to 0} \frac{\ln(1+x)}{x} = 1. \tag{22}$$

Per dimostrare il primo limite di (22), notiamo che

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

dove si è usato che $e^x = \sum\limits_{n=0}^{\infty} \frac{x^n}{n!}$. Quindi

$$|f(x)| \le |x| \sum_{n=0}^{\infty} \frac{|x|^n}{(n+2)!} \le |x| \sum_{n=0}^{\infty} \frac{|x|^n}{n!} = |x| e^{|x|}.$$

Dato che $0 < e^{|x|} \le e$ per tutti i valori $x \in [-1, 1]$, si ha $|f(x)| \le e|x|$ che tende a zero per $x \to 0$.

Il secondo limite in (22) si può ottenere dal primo ponendo $y = e^x - 1$:

$$\lim_{y\to 0} \frac{y}{\ln(1+y)} = \lim_{x\to 0} \frac{e^x - 1}{x} = 1,$$

e passando agli inversi si ha la conclusione.

OSSERVAZIONE 1.22. Nel calcolo di quest'ultimo limite si è utilizzato il cambiamento di variabile per dedurre il valore del limite a partire dal precedente. La giustificazione rigorosa di questo procedimento può essere fatta, con un po' di attenzione, ma senza troppa difficoltà, a partire dalla definizione di limite.

I limiti appena presentati sono utili come esempi, ma allo stesso tempo, sono fondamentali per riuscire a calcolare altri limiti. Altri limiti importanti, di cui non diamo la dimostrazione, sono

$$\lim_{x\to 0} x^\alpha \ln x = 0, \qquad \lim_{x\to +\infty} \frac{x^\alpha}{a^x} = 0, \qquad \lim_{x\to +\infty} \frac{\log_a x}{x^\alpha} = 0 \qquad \forall\, a > 1,\ \alpha > 0.$$

Il significato di ciascuno di questi è particolarmente interessante. Nel primo limite, la funzione x^α tende a 0 per $x \to 0$, mentre $\ln x$ tende a $-\infty$. Non è chiaro a priori quale sia il comportamento della funzione prodotto dato che sono presenti due termini contrastanti. Il fatto che il limite valga zero vuol dire che la funzione x^α tende a zero tanto rapidamente da riuscire a dominare la divergenza $-\infty$ del termine $\ln x$. Allo stesso modo, il secondo limite indica che l'esponenziale a^x, con $a > 1$ diverge più rapidamente di x^α, e il terzo esprime che, al contrario, il logaritmo $\log_a x$, con $a > 1$, diverge più lentamente di x^α. Sulle questioni di ordini di infinito e di infinitesimo ritorneremo più avanti.

2. Continuità

Il concetto di limite è collegato a quello di *continuità*. Intuitivamente la continuità significa che piccoli cambiamenti nella variabile indipendente x provocano piccoli cambiamenti nella variabile dipendente $y = f(x)$. Al contrario un grafico costituito da due parti separate da una "frattura" in corrispondenza dell'ascissa x_0 esibisce (in quel punto) una *discontinuità di salto* (ad esempio, la funzione $\operatorname{sgn} x$ ha una discontinuità di salto in $x_0 = 0$).

L'idea di continuità è implicita nell'uso quotidiano della matematica elementare. Quando una funzione $y = f(x)$ è descritta da tabelle (come nel caso dei logaritmi o delle funzioni trigonometriche), i valori di y possono essere dati solo per un insieme "discreto" di valori della variabile indipendente x, ad esempio in intervalli di lunghezza 10^{-3} (un millesimo) o 10^{-6} (un milionesimo). Però potrebbe essere utile conoscere il valore della funzione per valori intermedi. In questo caso, si assume tacitamente che il valore $f(x_0)$ cercato, corrispondente ad un valore x_0 non presente nella tabella, sia approssimativamente lo stesso di $f(x)$ per un x che appaia nella tabella e che sia "vicino" ad x_0.

DEFINIZIONE 2.1. *Sia I un intervallo di $\mathbb{R}$. La funzione $f : I \to \mathbb{R}$ è* continua in *$x_0 \in I$ se ha limite per $x \to x_0$ esiste e tale limite coincide con il valore di f in x_0:*

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

cioè (ricordando la definizione di limite) se

$$\forall\, \varepsilon > 0, \quad \exists\, \delta > 0 \quad t.c. \quad |f(x) - f(x_0)| < \varepsilon \quad \forall\, x \in I,\ |x - x_0| < \delta.$$

Se una funzione f è continua in ogni punto $x_0 \in I$ allora f è continua in *I.*

Sia $P_0 = (x_0, y_0)$ un punto nel grafico. I punti (x, y) tali che $y_0 - \varepsilon < y < y_0 + \varepsilon$ costituiscono una striscia orizzontale J che contiene P_0. La continuità di f in x_0 significa che per ogni striscia di questo genere J (di qualsiasi ampiezza) è possibile determinare una striscia verticale K data da $x_0 - \delta < x < x_0 + \delta$ sufficientemente piccola tale che tutti i punti del grafico di f che sono in K giacciono anche in J.

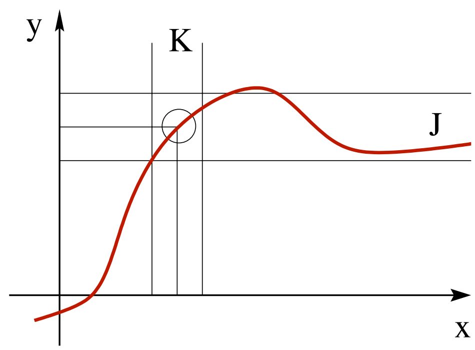

FIGURA 3. Significato geometrico della continuità

ESEMPIO 2.2. Per la funzione affine $f(x) = 5x + 3$ abbiamo

$$|f(x) - f(x_0)| = |(5x + 3) - (5x_0 + 3)| = 5|x - x_0|,$$

che esprime il fatto che la funzione $y = 5x + 3$ *dilata* le distanze di un fattore 5. In questo caso, ovviamente $|f(x) - f(x_0)| < \varepsilon$ per tutti valori x per cui $|x - x_0| < \varepsilon/5$. La condizione di continuità di f nel punto x_0 è soddisfatta scegliendo $\delta = \varepsilon/5$. Chiaramente è possibile scegliere un qualsiasi valore positivo tale che $\delta \leq \varepsilon/5$.

OSSERVAZIONE 2.3. Nella definizione di continuità, la condizione $|f(x) - f(x_0)| < \varepsilon$ è soddisfatta anche per x_0, a differenza della definizione di limite dove si chiede $|f(x) - \ell|$ per valori x vicini a x_0, ma diversi da x_0 stesso.

Osservazione 2.4. *Scommettiamo che...* Per chiarire ulteriormente il significato di continuità, spieghiamo le regole di un gioco per due persone. Supponiamo assegnata una funzione f ed il punto x_0 nel suo insieme di definizione. Il giocatore B può scegliere un qualsiasi numero $\varepsilon > 0$ a suo gusto e piacimento. Per ogni scelta di ε compiuta da B, A deve essere in grado di determinare $\delta > 0$ in modo che tutti i valori immagine $f(x)$, per x che dista da x_0 meno di δ, distino da $f(x_0)$ meno di ε. Se il giocatore B trova un $\varepsilon > 0$ per cui A non possa rispondere, vince; viceversa, se per ogni ε, A è in grado di trovare δ opportuno, vince il giocatore A. Il giocatore A vince se e solo se la funzione f è continua in x_0.

Se la funzione è $\sin(x^2)$ ed il punto $x_0 = 1$, quale giocatore vorreste essere: il giocatore A o il giocatore B?

Ora che abbiamo una definizione chiara di continuità, vorremmo sapere quante e quali funzioni tra quelle che conosciamo sono continue. Dalle proprietà dei limiti di somma, prodotto, quoziente discende che

la somma, la differenza, il prodotto e il rapporto di funzioni continue danno luogo a funzioni continue (prudenza nel quoziente![1]).

Anche le operazioni di composizione e di inversione conservano la continuità:

la composizione $f \circ g$ di funzioni f e g continue è continua

l'inversa f^{-1} di una funzione f continua è una funzione continua

La prima delle due proprietà discende dalla catena di implicazioni

$$x \to x_0 \quad \Rightarrow \quad g(x) \to g(x_0) \quad \Rightarrow \quad f(g(x))) \to f((g(x_0))).$$

La continuità della funzione inversa è geometricamente evidente, una volta ricordato che il grafico di f^{-1} si può ottenere da quello di f tramite un ribaltamento attorno alla bisettrice del primo e terzo quadrante.

Ma tutte queste bellissime proprietà non servono a nulla fino a che non si conosca per lo meno <u>una</u> funzione continua. Passiamo quindi ad analizzare qualche esempio di base.

Esempio 2.5. *Le funzioni costanti sono continue.* Banale! Infatti se $f(x) = c$ per ogni x, allora $|f(x) - f(x_0)| = |c - c| = 0$ sempre e comunque.

Esempio 2.6. *La funzione $f(x) = x$ è continua.* Anche questo è facile, dato che, fissato $x_0 \in \mathbb{R}$, si ha $|f(x) - f(x_0)| = |x - x_0|$, quindi basta scegliere $\delta(\varepsilon) = \varepsilon$ nella definizione di continuità per giungere alla conclusione.

[1]Come sempre nel caso della divisione, bisogna stare attenti al fatto che la divisione per zero non ha senso. Perciò se si hanno due funzione continue f e g, la funzione rapporto è una funzione continua dove è definito, cioè dove la funzione g non si azzera.

Esempio 2.7. *I polinomi sono funzioni continue.* Qui basta combinare le proprietà dei limiti (20), con la definizione di continuità e con i due esempi precedenti. Se $p(x) = a_0 + a_1 x + \cdots + a_n x^n$ per $a_0, a_1, \ldots, a_n \in \mathbb{R}$ dati, allora

$$\begin{aligned}\lim_{x\to x_0} p(x) &= \lim_{x\to x_0} (a_0 + a_1 x + \cdots + a_n x^n)\\ &= \lim_{x\to x_0} a_0 + \lim_{x\to x_0} a_1 \lim_{x\to x_0} x + \cdots + \lim_{x\to x_0} a_n \big(\lim_{x\to x_0} x\big)^n\\ &= a_0 + a_1 x_0 + \cdots + a_n x_0^n = p(x_0).\end{aligned}$$

Esempio 2.8. *La funzione $f(x) = \sin x$ è una funzione continua.* Lo abbiamo già visto nell'Esempio 1.5. Stesso dicasi per $\cos x$ (avete risolto l'Esercizio 1.6?).

Esempio 2.9. Cosa dire dell'esponenziale e^x? L'Esempio 1.7 ne garantisce la continuità in $x_0 = 0$. Da questa è possibile dedurre la continuità anche negli altri punti, utilizzando la proprietà $e^{x+y} = e^x e^y$. Infatti:

$$\lim_{x\to x_0} e^x = \lim_{x\to x_0} e^{x-x_0+x_0} = \lim_{x\to x_0} e^{x-x_0} e^{x_0} = \lim_{x\to x_0} e^{x-x_0} \lim_{x\to x_0} e^{x_0} = \lim_{h\to 0} e^h \lim_{x\to x_0} e^{x_0} = e^{x_0}.$$

Una volta che abbiamo questi mattoni fondamentali, ecco a cascata una quantità impressionante di funzioni continue:
– le funzioni razionali,
– le funzioni trigonometriche,
– esponenziali e logaritmi,
– tutte le loro composizioni e inverse.

Esercizio 2.10. *Perché le funzioni $f(x) = 10^x$ e $g(x) = \log_{10} x$ sono continue?*

Estensione per continuità. Quando una funzione f non è definita in x_0, ma esiste il limite $\lim_{x\to x_0} f(x) = \ell$, è naturale definire una nuova funzione come segue

$$F(x) := \begin{cases} f(x) & x \neq x_0\\ \ell & x = x_0.\end{cases}$$

La funzione F si chiama estensione per continuità di f, dato che, per costruzione, F è continua in x_0. La domanda "è possibile estendere per continuità in x_0 una assegnata funzione f?" equivale a "esiste il limite di f per $x \to x_0$?"

Esempio 2.11. La funzione $\frac{\sin x}{x}$ non è definita in $x = 0$, ma ammette limite per $x \to 0$. Quindi può essere *estesa per continuità* in $x = 0$ attribuendole il valore 1. La nuova funzione (continua in $\mathbb{R}$) è

$$f(x) = \begin{cases} \dfrac{\sin x}{x} & x \neq 0,\\ 1 & x = 0.\end{cases}$$

Esercizio 2.12. *(a) Dire quale delle seguenti funzioni può essere estesa per continuità in $x = 0$*

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

(b) Sia f una funzione continua in $x = 0$ e tale che $f(0) = 0$. E' vero che la funzione $f(x)\sin(1/x)$ può essere estesa per continuità in $x = 0$?

Funzioni lipschitziane. Una funzione $f : I \subset \mathbb{R} \to \mathbb{R}$ è lipschitziana se esiste una costante $L > 0$ tale che

$$|f(x_1) - f(x_2)| \leq L|x_1 - x_2| \qquad \forall x_1, x_2 \in I.$$

La lipschitzianità corrisponde al fatto che il *rapporto incrementale*, cioè il coefficiente della retta secante passante per i punti del grafico di f di coordinate $(x_1, f(x_1))$ e $(x_2, f(x_2))$

$$\frac{f(x_1) - f(x_2)}{x_1 - x_2},$$

è limitato in valore assoluto da un fissato valore finito L.

Esempi di funzioni lipschitziane sono le funzioni affini $f(x) = ax + b$. Un altro esempio è $f(x) = \sin x$, infatti, come già osservato in precedenza,

$$|\sin x - \sin x_0| \leq |x - x_0|.$$

Tutte le funzioni lipschitziane sono continue: dato $\varepsilon > 0$, per avere $|f(x) - f(x_0)| < \varepsilon$ basta scegliere $\delta = \varepsilon/L$,

$$|f(x) - f(x_0)| \leq L|x - x_0| < L\delta = \varepsilon.$$

Esercizio 2.13. *Dimostrare le seguenti affermazioni.*
(i) Se f, g sono funzioni lipschitziane, allora anche $f + g$ è lipschitziana.
(ii) Se f, g sono lipschitziane e limitate, *allora fg è lipschitziana.*
(iii) Se in (ii) si rimuove l'ipotesi di limitatezza, la conclusione non è vera.

Soluzione. (i) Indicate con L_f, L_g, due costanti per cui è soddisfatta la condizione di Lipschitz per f e g rispettivamente, allora

$$|(f(x) + g(x)) - (f(y) - g(y))| \leq |f(x) - f(y)| + |g(x) - g(y)| \leq (L_f + L_g)|x - y|.$$

(ii) Indichiamo con L_f, L_g, due costanti per cui è soddisfatta la condizione di Lipschitz per f e g rispettivamente, e sia $|f(x)| \leq M_f$ e $|g(x)| \leq M_g$, allora

$$\begin{aligned} |f(x)g(x) - f(y)g(y)| &= |f(x)g(x) - f(y)g(x) + f(y)g(x) - f(y)g(y)| \\ &\leq |f(x) - f(y)||g(x)| + |f(y)||g(x) - g(y)| \leq (L_f M_g + M_f L_g)|x - y|. \end{aligned}$$

(iii) Ad esempio, si può scegliere $f(x) = g(x) = x$: il prodotto è la funzione x^2 che non è lipschitziana dato che

$$\sup_{x \neq y} \frac{|x^2 - y^2|}{|x - y|} = \sup_{x \neq y} |x + y| = +\infty.$$

Chiaro, no?

Esercizio 2.14. *Una funzione f è* hölderiana *se esistono $L, \alpha > 0$ tali che*

$$|f(x_1) - f(x_2)| \leq L|x_1 - x_2|^\alpha \qquad \forall x_1, x_2.$$

Dimostrare che se una funzione è hölderiana allora è anche continua.

Soluzione. Infatti dato $\varepsilon > 0$, per avere $|f(x) - f(x_0)| < \varepsilon$ basta scegliere $\delta = L^{-1/\alpha}\varepsilon^{1/\alpha}$:

$$|f(x) - f(x_0)| \leq L|x - x_0|^\alpha < L\delta^\alpha = \varepsilon,$$

per giungere alla conclusione sani e salvi.

3. Esempi di discontinuità

Un modo per chiarire ulteriormente la definizione di continuità è "in negativo", cioè dando esempi per cui non è soddisfatta.

Esempio 3.1. Riprendiamo l'esempio $f(x) = \operatorname{sgn} x$. Chiaramente, in ogni punto $x_0 \neq 0$, questa funzione è continua (qual è la scelta di δ in funzione di ε dato?). In $x_0 = 0$ la funzione, invece, non è continua. Infatti non è possibile determinare nessun δ quando ε sia minore di 1, dato che $|f(x) - f(0)| = |f(x)| = 1$ per ogni $x \neq 0$.

La funzione $\operatorname{sgn} x$ è l'esempio più semplice di discontinuità in un punto x_0 detto discontinuità di salto: la funzione f si avvicina, per x che tende a x_0 da destra e da sinistra, a valori limite che non coincidono con il valore di f in x_0.

Esempio 3.2. Un esempio di discontinuità in cui non ci siano limiti né da destra né da sinistra è dato dalla funzione

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

Il grafico della funzione f può essere dedotto da quello della funzione $\sin x$ attraverso un "passaggio al reciproco" nella variabile indipendente. Grossolanamente parlando, tutte le oscillazioni (infinite!) della funzione $\sin x$ per $x > 1$ vengono compresse nell'intervallo limitato $(0, 1)$ e si accumulano sul segmento del piano (x, y) di estremi $(0, -1)$ e $(0, 1)$ e non c'è alcuna speranza che la funzione possa essere continua in $x = 0$. Una figura chiarisce più di mille parole (Fig.4(a)).

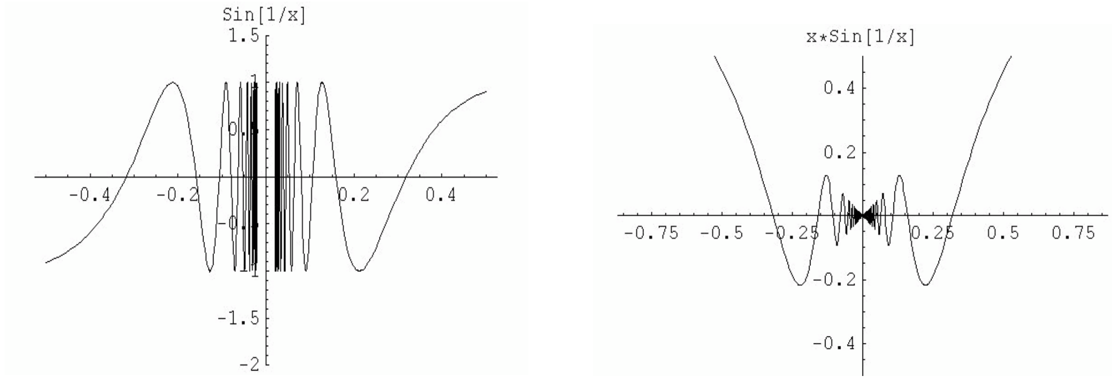

FIGURA 4. (a) Il grafico di $\sin(1/x)$; (b) Il grafico di $g(x)$.

Piccole varianti della funzione precedente possono condurre ad una funzione continua. Ad esempio consideriamo la funzione g seguente

$$g(x) = \begin{cases} x \sin\left(\dfrac{1}{x}\right) & x \neq 0, \\ 0 & x = 0. \end{cases}$$

Questa funzione (vedi Fig.4(b)) è continua in 0, infatti

$$|g(x) - 0| = \left| x \sin\left(\frac{1}{x}\right) \right| \leq |x| \to 0 \qquad \text{per} \quad x \to 0.$$

Sapete dire se è continua in 0 la funzione $(x^2 + 1)f(x)$, dove f è data nell'Esempio 3?

ESERCIZIO 3.3. *Sia $f : [a, b] \to \mathbb{R}$ una funzione nondecrescente e discontinua in $x_0 \in (a, b)$. Che tipo di discontinuità ha la funzione f in x_0?*

4. Teoremi sulle funzioni continue

Ora che abbiamo a disposizione un campionario vasto di funzioni continue e non, passiamo a stabilire alcune proprietà fondamentali che discendono dalla continuità: il *teorema dei valori intermedi* e il *teorema di Weierstrass* (che concerne il problema dell'esistenza di massimo e minimo). Entrambi discendono dal fatto che l'insieme dei numeri reali è *completo*, proprietà che traduce il fatto che la retta reale non ha buchi e che, rigorosamente, si basa sul postulato degli intervalli incapsulati e sull'assioma di Archimede. Nelle pagine che seguono ci dedichiamo prima a capire l'enunciato di questi due Teoremi fondamentali e solo successivamente ne vedremo le dimostrazioni.

Teorema del valore intermedio. Intuitivamente non c'è dubbio che se una funzione è continua, e quindi non ha salti, non può passare da un valore ad un altro senza passare per tutti i valori intermedi. Pensiamo ad un esempio banale: se il signor Lafcadio fa una passeggiata in montagna e ci comunica che è partito da un rifugio che si trova a 2200 metri s.l.m. ed è arrivato in cima ad una montagna alta 3000 metri s.l.m., è vero che ad un certo punto si è trovato ad un'altitudine di 2800 metri? E più in generale, si è mai trovato ad una qualsiasi quota η compresa tra 2200 e 3000? La risposta (intuitiva) è "SI", a meno che non abbia utilizzato il teletrasporto...

TEOREMA 4.1. <u>*Teorema del valore intermedio.*</u> *Sia $f : [a, b] \subset \mathbb{R} \to \mathbb{R}$ continua Allora, per ogni $\eta \in [f(a), f(b)]$, esiste $x_0 \in [a, b]$ tale che $f(x_0) = \eta$.*

Questo teorema dà condizioni sufficienti perchè l'equazione $f(x) = \eta$ abbia soluzione. Geometricamente, afferma che se i due punti $(a, f(a))$ e $(b, f(b))$ del grafico della funzione (continua) f giacciono su parti opposte rispetto alla retta $y = \eta$, allora il grafico di f interseca la retta in un punto intermedio.

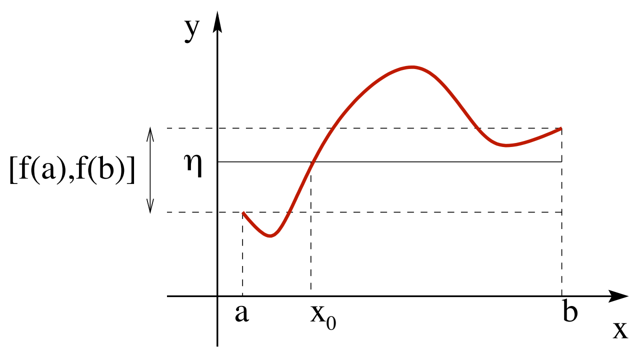

FIGURA 5. Il Teorema del valore intermedio

Controesempio 1. *f non è continua.* Nel caso di una funzione non continua la conclusione, in generale, è falsa. Ad esempio per la funzione

$$\text{segno di } x: \qquad \operatorname{sgn} x := \begin{cases} -1 & x < 0 \\ 0 & x = 0 \\ +1 & x > 0, \end{cases}$$

non esistono soluzioni di $\operatorname{sgn} x = \eta$ per ogni $\eta \notin \{0, \pm 1\}$.

Controesempio 2. *f definita in unione di intervalli disgiunti.* Consideriamo la funzione $f : I = [-1, 0) \cup (0, 1] \to \mathbb{R}$ definita da $f(x) = \frac{1}{x}$. Allora, nonostante $0 \in [-1, 1] = [f(-1), f(1)]$, l'equazione $\frac{1}{x} = 0$ non ammette soluzioni! Analogamente per $g : I = [0, 1] \cup [2, 3] \to \mathbb{R}$ definita da $g(x) = x$, ci sono dei valori $\eta \in [g(0), g(3)] = [0, 3]$ tali che l'equazione $x = \eta$ non ammette soluzioni *in I*.

Qui si è persa una proprietà fondamentale degli intervalli: la *connessione*, cioè la garanzia che se x_1, x_2 appartengono all'intervallo I, allora $[x_1, x_2] \subset I$. In qualche modo si può immaginare che una funzione continua non generi "strappi" o "buchi" nella trasformazione del dominio di partenza in quello di arrivo. E' chiaro però che se il dominio di partenza è "già strappato", cioè sconnesso (come nel caso di due intervalli chiusi disgiunti), è possibile che ci siano buchi anche nel dominio di arrivo.

Controesempio 3. *L'importanza di essere reale (razionale non basta!).* Consideriamo la funzione $f : \mathbb{Q} \cap [0,2] \to \mathbb{Q}$ definita da $f(x) = x^2$. Allora $f(0) = 0$, $f(2) = 4$ ed è sensato domandarsi se ci siano soluzioni $x \in \mathbb{Q} \cap [0,2]$ al problema $x^2 = 2 \in (0,4)$. Come abbiamo già visto non c'è nessun valore razionale il cui quadrato sia 2. Quindi *il Teorema del valore intermedio non vale nei razionali!*

Esercizio 4.2. *Sia $f : \mathbb{R} \to \mathbb{R}$ una funzione continua tale che*

$$\lim_{x\to-\infty} f(x) = -\infty \qquad e \qquad \lim_{x\to+\infty} f(x) = +\infty.$$

Dimostrare, utilizzando il Teorema del valore intermedio, che $f(\mathbb{R}) = \mathbb{R}$. Cosa si può concludere se, invece, si suppone $\lim_{x\to-\infty} f(x) = \ell_- \in \mathbb{R}$ e $\lim_{x\to+\infty} f(x) = \ell_+ \in \mathbb{R}$?

Conseguenza del teorema del valore intermedio è il cosiddetto Teorema di esistenza degli zeri.

Corollario 4.3. *Teorema di esistenza degli zeri.* *Sia $f : [a,b] \to \mathbb{R}$ continua. Allora, se $f(a)f(b) < 0$ (cioè se $f(a)$ e $f(b)$ hanno segno discorde), esiste $x_0 \in (a,b)$ tale che $f(x_0) = 0$.*

Esercizio 4.4. *Utilizzare il Teorema di esistenza degli zeri per dimostrare che ogni polinomio $p = p(x)$ di grado dispari ha sempre almeno uno zero.*

Soluzione. Se p è un polinomio di grado dispari

$$\lim_{x\to\pm\infty} p(x) = \pm\infty \qquad \text{oppure} \qquad \lim_{x\to\pm\infty} p(x) = \mp\infty,$$

e quindi esiste certamente $L > 0$ per cui $p(-L)p(L) < 0$. Di conseguenza, per il Teorema di esistenza degli zeri, esiste $x_0 \in (-L, L)$ che azzera il polinomio.

Una funzione strettamente monotòna, cioè tale che

$$x < x' \iff f(x) < f(x') \quad \text{oppure} \quad x < x' \iff f(x) > f(x'),$$

essendo iniettiva, è invertibile. In generale non è vero il viceversa: esistono funzioni invertibili che non sono monotòne (sapete trovarne un esempio?). Invece, nel caso di funzioni continue definite in un intervallo, la stretta monotonia è una condizione necessaria e sufficiente di invertibilità. La dimostrazione è conseguenza del Teorema del valore intermedio.

Corollario 4.5. *Sia f una funzione continua in $[a,b]$. Allora f è strettamente monotòna se e solo se f è invertibile.*

Dimostrazione. Basta dimostrare che se $f : [a,b] \to \mathbb{R}$ è continua ed invertibile, allora è anche strettamente monotòna. Supponiamo per assurdo che non lo sia, allora esisterebbero $x_1, x_2, x_3 \in [a,b]$ tali che $x_1 < x_2 < x_3$ per cui o $f(x_1) < f(x_2)$ e $f(x_3) < f(x_2)$ oppure $f(x_1) > f(x_2)$ e $f(x_3) > f(x_2)$. Supponiamo di essere nel primo caso (l'altro si tratta in modo analogo) e scegliamo η tale che $\max\{f(x_1), f(x_3)\} < \eta < f(x_2)$. Applicando il teorema del valore intermedio agli intervalli $[x_1, x_2]$ e $[x_2, x_3]$ si ottiene che esistono $\xi_1 \in (x_1, x_2)$ e $\xi_2 \in (x_2, x_3)$ per cui $f(\xi_1) = f(\xi_2) = \eta$ che contraddice l'ipotesi di invertibilità di f. □

Se si sostituisce l'ipotesi di "f strettamente monotona" con "f non strettamente monotona" la conclusione non è più vera. L'esempio più banale che si può pensare è quello di una funzione costante.

Teorema di Weierstrass. Un'altra proprietà fondamentale di una funzione continua f definita in un intervallo $[a,b]$ è l'*esistenza del (valore) massimo e del (valore) minimo.*

Teorema 4.6. Teorema di Weierstrass. *Sia $f : [a,b] \to \mathbb{R}$ continua. Allora esistono $x_0, x_1 \in [a,b]$ tali che $f(x_0) \le f(x) \le f(x_1)$ per ogni $x \in [a,b]$.*

Perseveriamo con la buona abitudine di cercare controesempi che mostrino il ruolo delle ipotesi del Teorema.

Controesempio 1. *f non è continua.* Se non è richiesta la continuità della funzione, è facile costruire casi di non esistenza di massimo/minimo. Ad esempio consideriamo $f : [-1,1] \to \mathbb{R}$ definita da

$$f(x) = \begin{cases} x^2 & x \neq 0, \\ 1 & x = 0. \end{cases}$$

Chiaramente $\inf\limits_{[-1,1]} f(x) = 0$, ma $f(x) \neq 0$ per ogni x. Analogamente si possono costruire casi in cui non c'è valore massimo.

Controesempio 2. $I = (a,b)$ *(intervallo aperto).* Anche in questo caso si possono trovare molti esempi che mostrano che le conclusioni del Teorema non sono vere. Ad esempio, $f(x) = x^2$ in $(-1,1)$ non ammette massimo (l'estremo superiore è 1), oppure $g(x) = \sin x$ in $(0, \pi/2)$ non ammette né massimo né minimo (l'estremo superiore è 1 e quello inferiore è 0). Si noti che in entrambi questi esempi, quello che si vorrebbe essere punto di massimo/minimo è uno degli estremi dell'intervallo, che però non appartiene ad I visto che l'intervallo è considerato aperto.

Controesempio 3. *I illimitato.* L'esempio più facile è $f(x) = x$ per $x \in \mathbb{R}$ che non ammette né massimo né minimo. Esistono anche funzioni *limitate* in domini illimitati che non ammettono né massimo né minimo, ad esempio, $f(x) = \arctan x$.

Se si combinano insieme il Teorema del valore intermedio ed il Teorema di Weierstrass si può dimostrare la seguente affermazione.

COROLLARIO 4.7. *Sia $f : [a, b] \to \mathbb{R}$ una funzione continua. Allora l'insieme immagine $f([a, b])$ è un intervallo chiuso e limitato.*

5. Gli intervalli incapsulati: *"divide et impera"*

Come abbiamo già detto nella presentazione naïf dei numeri reali, i fatti fondamentali che accettiamo come assiomi sono

> Postulato degli intervalli incapsulati. *Per ogni successione di intervalli $I_1, I_2, \ldots, I_n, \ldots$ 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.*
>
> Assioma di Archimede. *Per ogni numero reale a, esiste un numero naturale n più grande di a: in simboli,*
>
> $$\forall\, a \in \mathbb{R}, \qquad \exists\, n \in \mathbb{N}, \qquad \textit{tale che} \quad a \leq n.$$

Si ricordi che una delle conseguenze dell'assioma di Archimede è

$$x \leq \varepsilon \qquad \forall\, \varepsilon > 0 \qquad \Rightarrow \qquad x \leq 0.$$

Daremo ora le dimostrazioni del Teorema del Valore Intermedio e del Teorema di Weierstrass a partire dal seguente risultato.

COROLLARIO 5.1. *Sia $I_n = [a_n, b_n] \subset \mathbb{R}$ una successione di intervalli tali che*

(i) $I_{n+1} \subset I_n$ (cioè $a_n \leq a_{n+1} \leq b_{n+1} \leq b_n$) per ogni $n \in \mathbb{N}$;
(ii) $\lim\limits_{n\to+\infty} b_n - a_n = 0$.

Allora, esiste un unico $x_0 \in \mathbb{R}$ tale che $\bigcap\limits_{n=1}^{\infty} I_n = \{x_0\}$, cioè $a_n \leq x_0 \leq b_n$ per ogni $n \in \mathbb{N}$. Inoltre $\lim\limits_{n\to+\infty} a_n = \lim\limits_{n\to+\infty} b_n = x_0$.

Il Corollario indica che se la successione degli intervalli incapsulati ha la proprietà aggiuntiva che la lunghezza $|I_n| = b_n - a_n$ è infinitesima per $n \to +\infty$, l'intersezione degli I_n (non vuota per il Postulato degli Intervalli Incapsulati) è costituita da un solo elemento.

Dimostrazione del Corollario 5.1. La proprietà (i) implica $\bigcap_{n=1}^{\infty} I_n \neq \emptyset$ per il Postulato degli Intervalli Incapsulati; resta da dimostrare che tale intersezione è composta da un solo elemento. Siano $x_0, x_1 \in \bigcap_{n=1}^{\infty} I_n$ con $x_0 \leq x_1$. Allora $a_n \leq x_0 \leq x_1 \leq b_n$ per ogni n. Da questa relazione segue che $0 \leq x_1 - x_0 \leq b_n - a_n$. Dato che la successione $b_n - a_n$ è infinitesima, per ogni $\varepsilon > 0$, si ha $0 \leq x_1 - x_0 \leq b_n - a_n < \varepsilon$ per ogni n sufficientemente grande. In definitiva, per ogni $\varepsilon > 0$, si ha $0 \leq x_1 - x_0 < \varepsilon$, da cui segue $x_0 = x_1$, che conclude la prima parte del Teorema.

Inoltre, dato che $a_n \leq x_0 \leq b_n$, si ha anche $0 \leq x_0 - a_n \leq b_n - a_n \to 0$ per $n \to \infty$, quindi $a_n \to x_0$ per $n \to \infty$. Per b_n, basta notare che $b_n = a_n + (b_n - a_n) \to x_0 + 0 = x_0$. □

Divide et impera. Per dimostrare i Teoremi useremo sempre la strategia del *Divide et Impera*. Il punto chiave è definire una successione di intervalli incapsulati $[a_n, b_n]$ di misura $b_n - a_n$ infinitesima per $n \to +\infty$. Tipicamente, costruiremo la successione di intervalli I_n, scegliendo un primo intervallo opportuno $I_0 = [a, b]$, poi prendendo il punto intermedio dell'intervallo $\frac{a+b}{2}$ e scegliendo (secondo un criterio che dipende da caso a caso) come intervallo I_1 una delle due metà di I_0. Iterando il procedimento otterremo una successione di $I_n = [a_n, b_n]$ tale che

$$a_n \leq a_{n+1} \leq b_{n+1} \leq b_n, \quad \text{per ogni } n \in \mathbb{N}, \qquad b_n - a_n = \frac{b-a}{2^n},$$

cioè soddisfacente le ipotesi del Corollario 5.1.

Dimostrazione del Teorema del valore intermedio (Teorema 4.1). Supponiamo $f(a) < f(b)$ e $\eta \in (f(a), f(b))$. Se $f(b) < f(a)$, si può ragionare in modo simile. Come I_0 scegliamo l'intervallo di partenza $[a, b]$ e consideriamo il punto intermedio $\frac{a+b}{2}$. Procediamo come segue:

(i) se $f\left(\frac{a+b}{2}\right) = \eta$, siamo arrivati alla conclusione;
(i) se $f\left(\frac{a+b}{2}\right) > \eta$, poniamo $a_1 = a$ e $b_1 = \frac{a+b}{2}$;
(ii) se $f\left(\frac{a+b}{2}\right) < \eta$, poniamo $a_1 = \frac{a+b}{2}$ e $b_1 = b$.

In questo modo o siamo giunti alla conclusione, o abbiamo costruito un intervallo $I_1 = [a_1, b_1]$ tale che $f(a_1) < \eta < f(b_1)$ e $b_1 - a_1 = \frac{b-a}{2}$. Iteriamo il procedimento: scegliamo il punto $\frac{a_1+b_1}{2}$, calcoliamo $f\left(\frac{a_1+b_1}{2}\right)$ e procediamo come sopra.

Così facendo, o si è dimostrata la conclusione dopo un numero finito di passi, o si è costruita una successione di intervalli incapsulati $I_n = [a_n, b_n]$ tale che $b_n - a_n = \frac{b-a}{2^n}$. Applicando il Corollario 5.1, deduciamo che esiste $x_0 \in [a, b]$ tale che

$$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = x_0.$$

Per la scelta di a_n, b_n, si ha $f(a_n) < \eta < f(b_n)$ per ogni n. Dato che f è continua in x_0 e $a_n, b_n \to x_0$, per $n \to \infty$ si deduce $f(x_0) \le \eta \le f(x_0)$, cioè $f(x_0) = \eta$. $\square$

Dimostrazione del Teorema di Weierstrass (Teorema 4.6). Sia Λ l'estremo superiore di $\{f(x) : x \in [a,b]\}$. Per ora non sappiamo se Λ sia finito o no.

Poniamo $I_0 = [a,b]$. Dividiamo I_0 in due parti uguali, tramite il punto medio $\frac{a+b}{2}$. In almeno uno dei due sottointervalli $\left[a, \frac{a+b}{2}\right]$ e $\left[\frac{a+b}{2}, b\right]$ l'estremo superiore della funzione f è ancora uguale a Λ. Battezziamo il sottointervallo con questa proprietà I_1 e i suoi estremi con a_1 e b_1. Nel caso in cui entrambi gli intervalli vadano bene ne scegliamo uno a nostro piacere. Iteriamo il procedimento, dividendo il sottointervallo tramite il suo punto medio. In questo modo, otteniamo una successione di intervalli incapsulati $I_n = [a_n, b_n]$ tali che $b_n - a_n = \frac{b-a}{2^n}$. Applicando il Teorema 5.1, deduciamo che esiste x_1 tale che $\lim\limits_{n\to\infty} a_n = \lim\limits_{n\to\infty} b_n = x_1$. Vogliamo a questo punto mostrare che Λ è finito e, inoltre, $f(x_1) = \Lambda$.

La funzione f è continua in x_1, quindi per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che $f(x_1) - \varepsilon < f(x) < f(x_1) + \varepsilon$ per ogni $x \in (x_1 - \delta, x_1 + \delta)$. Fissato ε, e di conseguenza δ, scelgo $n \in N$ tale che $I_n \subset (x_1 - \delta, x_1 + \delta)$. Allora

$$f(x) < f(x_1) + \varepsilon \qquad \forall\, x \in I_n, \tag{23}$$

che esprime il fatto che $f(x_1) + \varepsilon$ è un maggiorante di $\{f(x) : x \in I_n\}$. Dato che $\sup\{f(x) : x \in I_n\} = \Lambda$ per costruzione, ne segue che Λ è finito, che $f(x_1) \le \Lambda$ e che $\Lambda \le f(x_1) + \varepsilon$. Ma, in quest'ultima relazione, ε può essere scelto arbitrariamente, quindi

$$0 \le \Lambda - f(x_1) \le \varepsilon \qquad \forall\, \varepsilon > 0.$$

Perciò $\Lambda = f(x_1)$ e la dimostrazione è completa. $\square$

Si noti che, sebbene la costruzione porti a determinare un singolo punto di massimo, ce ne potrebbero essere anche molti altri! In effetti, ad ogni passo, nella scelta del sottointervallo ci può essere libertà di scelta, nel caso in cui in entrambi i sottointervalli l'estremo superiore della funzione f sia ancora uguale a Λ.

Per concludere, utilizziamo la strategia del *divide et impera* per dimostrare l'esistenza di estremo superiore ed inferiore (risultato concettualmente del tutto indipendente dai concetti di continuità di funzioni reali di variabile reale).

Ricordiamo che il valore $\Lambda \in \mathbb{R}$ è l'estremo superiore di $E \subset \mathbb{R}$ se

(i) Λ è un maggiorante di E, cioè per ogni $y \in E$, si ha $y \le \Lambda$;

(ii) Λ è il più piccolo dei maggioranti, cioè se L è un maggiorante di E, allora $\Lambda \le L$.

Teorema di esistenza dell'estremo superiore. *Se $E \neq \emptyset$ è limitato superiormente, allora esiste $\Lambda = \sup E \in \mathbb{R}$.*

Dimostrazione. Sia $a \in E$ (E non è vuoto) e sia b un maggiorante di E (E è limitato superiormente). Indichiamo con I_0 l'intervallo chiuso e limitato $[a, b]$ e prendiamo il punto intermedio $\frac{a+b}{2}$. Allora $I_1 := [a_1, b_1]$, dove

(i) se $\frac{a+b}{2}$ è un maggiorante di E, poniamo $a_1 = a$ e $b_1 = \frac{a+b}{2}$,
(ii) se $\frac{a+b}{2}$ non è un maggiorante di E, poniamo $a_1 = \frac{a+b}{2}$ e $b_1 = b$.

In entrambi i casi in $I_1 = [a_1, b_1]$ c'è almeno un elemento di E e b_1 è un maggiorante di E. Iterando il procedimento, otteniamo la solita successione di intervalli incapsulati con $b_n - a_n = \frac{b-a}{2^n}$. Quindi, sempre per il Corollario 5.1, esiste Λ tale che $\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = \Lambda$. Dato che $b_n \geq y$ per ogni $y \in E$, la stessa proprietà vale al limite: $\Lambda \geq y$ per ogni $y \in E$. Inoltre, per costruzione, ci sono elementi di E arbitrariamente vicini a Λ, quindi anche la condizione (ii) è soddisfatta. $\square$

CAPITOLO 5

Derivate, derivate e derivate

Prima di definire rigorosamente il concetto di derivabilità, prendiamoci il tempo di discutere un paio di punti di vista che indicano quale sia il significato di questo nuovo oggetto matematico: la derivata.

La derivata come velocità. Consideriamo un punto che si muova lungo l'asse y con posizione $y = f(t)$ all'istante t Se la funzione f è affine, ossia $f(t) = At+B$, si parla di moto uniforme. La velocità A è il rapporto tra la distanza percorsa nell'intervallo di tempo $[t_0, t_1]$ e la durata di questo intervallo:

$$A = \frac{\Delta f}{\Delta t} = \frac{f(t) - f(t_0)}{t - t_0}.$$

Il moto è *uniforme* perchè la velocità è costante e, di conseguenza, in intervalli di tempo uguali, vengono percorse distanze uguali.

Se il moto non è uniforme, la quantità $\frac{f(t)-f(t_0)}{t-t_0}$ esprime la velocità media del punto nell'intervallo di tempo $[t_0, t]$. Se la velocità media tende ad un limite finito per $t \to t_0$, il valore del limite è detto velocità (istantanea):

$$v(t_0) = \lim_{t \to t_0} \frac{f(t) - f(t_0)}{t - t_0}.$$

Se il limite non esiste, la velocità istantanea non è definita.

Un esempio semplice è il moto di un corpo in caduta libera, cioè sottoposto alla sola forza di gravità. Sperimentalmente, la distanza percorsa al tempo t da un corpo, lasciato cadere da fermo al tempo $t = 0$, è proporzionale a t^2; si rappresenta quindi con una funzione della forma

$$y = f(t) = at^2 \qquad (a > 0).$$

La velocità v all'istante t si ottiene quindi calcolando

$$v(t_0) = \lim_{t \to t_0} \frac{f(t) - f(t_0)}{t - t_0} = \lim_{t \to t_0} \frac{at^2 - at_0^2}{t - t_0} = \lim_{t \to t_0} a(t + t_0) = 2at_0.$$

Quindi la velocità di un corpo in caduta libera cresce in modo proporzionale al tempo.

Nello studio del moto di un punto è utile osservare anche la variazione di velocità. Il procedimento è simile al precedente. L'accelerazione media è il rapporto tra la variazione di velocità nell'intervallo di tempo $[t_0, t]$ e la durata dell'intervallo, cioè è data da

$(v(t)-v(t_0))/(t-t_0)$. L'accelerazione (istantanea) a è il limite dell'accelerazione media per $t \to t_0$

$$a(t_0)=\lim_{t\to t_0}\frac{v(t)-v(t_0)}{t-t_0}.$$

Nel caso di moto uniforme $f(t)=At+B$,

$$v(t)=A \qquad \forall\, t \qquad \Rightarrow \qquad a(t_0)=\lim_{t\to t_0}\frac{A-A}{t-t_0}=0,$$

cioè l'accelerazione è nulla; nel caso del corpo in caduta libera $f(t)=at^2$,

$$v(t)=2at \qquad \forall\, t \qquad \Rightarrow \qquad a(t_0)=\lim_{t\to t_0}\frac{2at-2at_0}{t-t_0}=2a,$$

cioè il moto è *uniformente accelerato.*

La derivata come approssimazione lineare. In generale, supponiamo di esaminare l'evoluzione di una quantità (la posizione di un punto in movimento, la temperatura dell'acqua sul fuoco, o altro...) descritta all'istante t, dal numero reale $y=f(t)$. Fissiamo un istante iniziale t_0 e misuriamo il valore di $y=f(t_0)$. Per controllare quello che succederà da t_0 in poi, dobbiamo studiare la variazione di f, cioè la quantità

$$\Delta f(t;t_0):=f(t)-f(t_0).$$

Se la funzione f è costante, non c'è evoluzione: $\Delta f=0$ per ogni scelta di t_0 e t. Se f è una funzione affine, cioè se $f(t)=At+B$ per qualche $A,B\in\mathbb{R}$, allora

$$\Delta f(t;t_0)=(At+B)-(At_0+B)=A(t-t_0)=A\Delta t,$$

dove $\Delta t=t-t_0$ rappresenta l'intervallo di tempo trascorso dall'istante iniziale t_0 a quello finale t. Come si vede, se f è affine, la funzione Δf è lineare nell'incremento Δt della variabile indipendente t, cioè Δf è proporzionale a Δt. La costante di proporzionalità è A ed è data da $A=\Delta f/\Delta t$.

Proviamo con un polinomio di secondo grado in t: $f(t)=at^2+bt+c$ con $a,b,c\in\mathbb{R}$. In questo caso, scrivendo $t=t_0+\Delta t$

$$\begin{aligned}\Delta f(t_0+\Delta t;t_0)&=\left[a(t_0+\Delta t)^2+b(t_0+\Delta t)+c\right]-\left[at_0^2+bt_0+c\right]\\&=(2at_0+b)\,\Delta t+a\,(\Delta t)^2\,.\end{aligned}$$

Questa volta l'incremento Δf non è lineare in Δt, dato che compare il termine quadratico $A\,(\Delta t)^2$. Però Δf ha la gentilezza di decomporsi in due parti:

$$\Delta f(t_0+\Delta t;t_0)=(\text{termine lineare in }\Delta t)+(\text{resto})\,.$$

Quanto è grande il resto $a\,(\Delta t)^2$? Consideriamo un caso semplice: $a=1, b=0$, c qualsiasi, $t_0=1$. In questo caso:

$$f(t)=t^2+c, \qquad \Delta f(1+\Delta t;1)=2\Delta t+(\Delta t)^2,$$

la parte lineare è $2\Delta t$ ed il resto è $(\Delta t)^2$. Vediamo i valori di questi due termini per diverse scelte di Δt:

Δt	1	$0,1$	$0,01$	$0,001$	$0,0001$	$0,00001$
parte lineare: $2\Delta t$	2	$0,2$	$0,02$	$0,002$	$0,0002$	$0,00002$
resto: $(\Delta t)^2$	1	$0,01$	$0,0001$	$0,000001$	$0,00000001$	$0,0000000001$

Come si vede dalla tabella, sia il termine lineare che il resto diminuiscono per $\Delta t \to 0$ (sono infinitesimi). Ma c'è una differenza fondamentale: il resto $(\Delta t)^2$ diviene piccolo molto più rapidamente del termine lineare. Dunque è ragionevole approssimare, per $\Delta t \to 0$, l'incremento Δf tramite una funzione lineare in Δt:

$$\Delta f(t_0 + \Delta t; t_0) \approx (2at_0 + b)\,\Delta t \qquad \text{per} \quad \Delta t \to 0.$$

In generale, data f qualsiasi, se è possibile scrivere l'incremento Δf nella forma $\Delta f(t_0 + \Delta t; t_0) = A\Delta t + R$ con R che tende a zero più rapidamente del termine lineare $A\Delta t$ ha senso utilizzare l'approssimazione

$$\Delta f(t_0 + \Delta t; t_0) \approx A\Delta t \qquad \text{per} \quad \Delta t \to 0.$$

Tutte le volte che questa operazione è possibile, la funzione f si dice *derivabile* e il valore A è la *derivata prima* di f in t_0. La derivata, dunque, dà un'informazione sulla variazione Δf della funzione f quando la variabile indipendente t subisca una variazione Δt piccola.

Restano un paio di perplessità: che vuol dire la frase "il resto tende a zero più rapidamente del termine lineare"? E, in concreto, data una funzione f come stabilire se esiste e come calcolare il valore A? La risposta alla prima domanda permette magicamente di risolvere anche il secondo angoscioso quesito. Dire che il resto R tende a zero più rapidamente del termine lineare vuol dire richiedere che valga

$$\lim_{\Delta t \to 0} \frac{R}{\Delta t} = 0.$$

In questo limite ci si trova di fronte ad una forma indeterminata del tipo $\frac{0}{0}$, e la richiesta è che il resto R tende a zero tanto rapidamente da dominare l'effetto del termine infinitesimo a denominatore.

Dalla condizione sul resto si deduce un modo per calcolare A: se esiste A tale che $\Delta f = A\Delta t + R$ con R che soddisfa $\lim\limits_{\Delta t \to 0} R/\Delta t = 0$, allora, dividendo per Δt,

$$A = \frac{\Delta f}{\Delta t} - \frac{R}{\Delta t},$$

e passando al limite per $\Delta t \to 0$, si ha

$$A = \lim_{\Delta t \to 0} \frac{\Delta f}{\Delta t},$$

che dà una maniera rigorosa di definire la derivata di f e, allo stesso tempo, una maniera per calcolarne il valore.

1. Definizione di derivata

Riprendiamo il discorso da capo e mettiamo ordine nel *brainstorming* fatto fin qui.

DEFINIZIONE 1.1. Derivabilità. *Una funzione $f : [a,b] \to \mathbb{R}$ è* derivabile in $x_0 \in (a,b)$ *se esiste* finito *il limite*

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0}. \tag{24}$$

Se esiste, il limite si indica con $f'(x_0)$ e si dice derivata (prima) della funzione f in x_0. *Se f è derivabile in tutti i punti di $[a,b]$, si dice che f* è derivabile in $[a,b]$.

Per la derivata si usano anche altri simboli:

$$f' = \frac{df}{dx} = Df = \frac{dy}{dx} = \dot{y} = \cdots ,$$

e il limite (24) può essere scritto in maniere equivalenti

$$\lim_{x\to x_0} \frac{f(x)-f(x_0)}{x-x_0} = \lim_{h\to 0} \frac{f(x+h)-f(x)}{h} = \lim_{\Delta x\to 0} \frac{\Delta f}{\Delta x} = \cdots$$

Dato che la derivata f' dipende dal punto di derivazione, f' *è essa stessa una funzione*, il cui insieme di definizione è contenuto nell'insieme di definizione della funzione f (non è detto che i due domini di definizione coincidano).

Significato geometrico. Data una funzione $y = f(x)$, consideriamo il problema di determinare la retta tangente al grafico della funzione nel punto $P_0 = (x_0, f(x_0))$. L'idea è la seguente: dato un secondo punto $P = (x, f(x))$ sul grafico di f, per P_0 e P passa un'unica retta, detta retta secante. Se, muovendo P verso P_0, la retta secante

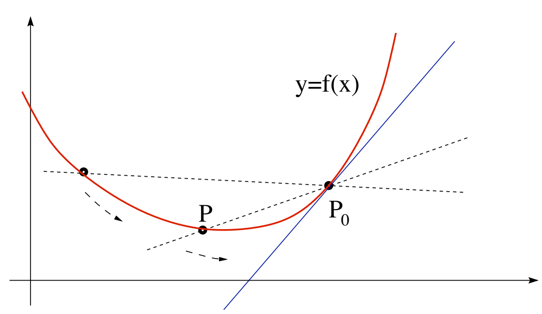

FIGURA 1. Il grafico di una funzione con tangente e secanti.

tende ad una posizione limite, tale retta limite è la retta tangente. Formuliamo ora, in maniera rigorosa, il processo geometrico di limite che abbiamo appena raccontato. Il coefficiente angolare della retta secante per P_0 e P è

$$\textsf{rapporto incrementale:} \qquad \frac{f(x)-f(x_0)}{x-x_0}.$$

Se la funzione f è derivabile in x_0, esiste il limite del rapporto incrementale e vale $f'(x_0)$, quindi il valore della derivata prima in x_0 rappresenta il coefficiente angolare della retta tangente al grafico della funzione nel punto $P_0 = (x_0, f(x_0))$.

Osservazione 1.2. Determinare una derivata vuol dire fare (con successo) un limite: i limiti si fanno nei punti interni ad un intervallo di definizione. Negli estremi si fanno al più limiti sinistri o limiti destri. In punti isolati non si fanno neanche i limiti... Chi penserebbe di fare la tangente in un singolo punto?

La derivata è dunque il limite di una funzione opportuna, il rapporto incrementale. Vediamo come calcolare *esplicitamente* tale funzione derivata. Partiamo da alcuni casi semplici:
– se $f(x) = c \in \mathbb{R}$ per ogni x, si ha

$$\frac{f(x+h)-f(x)}{h} = \frac{c-c}{h} = 0 \quad \Longrightarrow \quad f'(x) = \lim_{h\to 0} 0 = 0;$$

– se $f(x) = x$, vale

$$\frac{f(x+h)-f(x)}{h} = \frac{(x+h)-x}{h} = 1 \quad \Longrightarrow \quad f'(x) = \lim_{h\to 0} 1 = 1;$$

– infine, se $f(x) = x^2$, si ha

$$\frac{f(x+h)-f(x)}{h} = \frac{2xh+h^2}{h} = 2x+h \quad \Rightarrow \quad f'(x) = \lim_{h\to 0}(2x+h) = 2x.$$

Passiamo ora ad un esempio meno facile: sia $f(x) = \sqrt{x}$ per $x \geq 0$. Il rapporto incrementale in $x \neq 0$ è

$$\frac{\sqrt{x+h}-\sqrt{x}}{h} = \frac{\sqrt{x+h}-\sqrt{x}}{h}\,\frac{\sqrt{x+h}+\sqrt{x}}{\sqrt{x+h}+\sqrt{x}} = \frac{1}{\sqrt{x+h}+\sqrt{x}}$$

Passando al limite per $h \to 0$ si ottiene

$$\lim_{h\to 0} \frac{\sqrt{x+h}-\sqrt{x}}{h} = \frac{1}{2\sqrt{x}} \qquad\qquad x > 0.$$

Nel punto $x = 0$ la funzione ha una *singolarità* (come è fatto il grafico di $\sqrt{x}$?): pur essendo definita e continua, si ha

$$\lim_{h\to 0^+} \frac{\sqrt{x+h}-\sqrt{x}}{h} = +\infty.$$

Ecco altri due esempi di funzioni continue, ma non derivabili in $x = 0$:

$$f(x) = |x| \qquad \text{e} \qquad g(x) = \begin{cases} x \sin\left(\frac{1}{x}\right) & x \neq 0, \\ 0 & x = 0 \end{cases}$$

Per la funzione f, la non derivabilità in 0 è dovuta al fatto che i limiti destro e sinistro del rapporto incrementale esistono finiti ma non coincidono (il rapporto incrementale ha una discontinuità di salto in 0)

$$\lim_{x\to 0^+} \frac{|h|}{h} = 1 \neq -1 = \lim_{x\to 0^-} \frac{|h|}{h}.$$

Nel grafico, un comportamento di questo genere si traduce nella presenza di un *punto angoloso*. Nel caso della funzione g, il rapporto incrementale ha l'espressione

$$\frac{g(h) - g(0)}{h} = \frac{h\sin(1/h) - 0}{h} = \sin\left(\frac{1}{h}\right).$$

Come si è già visto, questa funzione non ha limite (né destro né sinistro) per $h \to 0$. In termini di grafico (controllare di persona!), questa funzione ha delle variazioni sempre più rapide di pendenza man mano che ci sia avvicina ad $x = 0$.

Due conseguenze. Vediamo cosa si può dedurre in un soffio dalla derivabilità.

1. *Derivabilità* $\Rightarrow$ *Continuità.* Se una funzione f è derivabile in x_0, allora è anche continua in x_0. Infatti la continuità della funzione f nel punto x_0 è equivalente all'affermazione $\lim\limits_{x\to x_0} \big(f(x) - f(x_0)\big) = 0$, e, dato che

$$f(x) - f(x_0) = \frac{f(x) - f(x_0)}{x - x_0}\,(x - x_0),$$

passando al limite per $x \to x_0$ si ottiene la conclusione.

2. *Equazione della retta tangente.* Data $f : [a, b] \to \mathbb{R}$, sia $x_0 \in [a, b]$ un punto in cui f è derivabile, la retta tangente è, per definizione, la retta passante per il punto $(x_0, f(x_0))$, il cui coefficiente angolare è pari a $f'(x_0)$

$$\textsf{equazione della retta tangente:} \qquad y = f(x_0) + f'(x_0)(x - x_0).$$

Fissato il punto x_0, il polinomio di primo grado in x a secondo membro può essere visto come un'approssimazione della funzione f *vicino al punto* x_0.

Nel sostituire la funzione con la sua retta tangente l'errore R_{x_0}, è pari a

$$R(x; x_0) = f(x) - f(x_0) - f'(x_0)(x - x_0).$$

Per $x \to x_0$, l'errore che si commette tende a zero, cioè

$$\lim_{x\to x_0} R(x; x_0) = \lim_{x\to x_0} \big(f(x) - f(x_0) - f'(x_0)(x - x_0)\big) = 0. \tag{25}$$

Ma (attenzione!) lo stesso è vero per qualsiasi altra retta per il punto $(x_0, f(x_0))$, infatti

$$\lim_{x\to x_0}\big(f(x)-f(x_0)-m(x-x_0)\big)=0 \qquad \forall m\in\mathbb{R}.$$

Quindi la proprietà (25) non è indicativa! Il fatto *fondamentale* è che per $R(x;x_0)$ vale

$$\lim_{x\to x_0}\frac{R(x;x_0)}{x-x_0}=\lim_{x\to x_0}\frac{f(x)-f(x_0)-f'(x_0)(x-x_0)}{x-x_0}=0. \tag{26}$$

Questa condizione è più restrittiva della precedente e, tra le funzioni affini, è verificata *solo* da quella che rappresenta la retta tangente ad f in x_0. In maniera equivalente, avremmo potuto dire che una funzione è derivabile in x_0 se esiste un valore $\ell\in\mathbb{R}$ per cui

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)-\ell(x-x_0)}{x-x_0}=0.$$

Il valore ℓ è pari a $f'(x_0)$.

Prime formule di derivazione. Applichiamo ora la definizione per calcolare esplicitamente le derivate di alcune funzioni semplici.

Polinomi e potenze. Si è già visto che valgono le regole di derivazione

$$\big(c\big)'=0,\qquad \big(x\big)'=1,\qquad \big(x^2\big)'=2x.$$

Per un generico polinomio di grado 2, $f(x)=ax^2+bx+c$ si può procedere in modo analogo. Il rapporto incrementale è

$$\frac{f(x+h)-f(x)}{h}=\frac{a(x+h)^2+b(x+h)+c-ax^2-bx-c}{h}=2ax+b+h.$$

Quindi, passando al limite per $h\to 0$, si ottiene

$$\big(ax^2+bx+c\big)'=\lim_{h\to 0}(2ax+b+h)=2ax+b.$$

In modo simile è possibile derivare un qualsiasi polinomio. Calcoliamo prima di tutto la derivata di $f(x)=x^n$ dove $n\in\mathbb{N}$. Il rapporto incrementale si può scrivere come

$$\frac{f(x_1)-f(x)}{x_1-x}=\frac{x_1^n-x^n}{x_1-x}=x_1^{n-1}+x_1^{n-2}x+\cdots+x^{n-1},$$

dato che $x_1^n-x^n=(x_1-x)(x_1^{n-1}+x_1^{n-2}x+\cdots+x^{n-1})$ per ogni $x_1,x\in\mathbb{R}$. Passando al limite per $x_1\to x$, ciascuno dei termini tende a x^{n-1} e quindi, dato che si tratta di n termini, si ottiene

$$\big(x^n\big)'=nx^{n-1}\qquad \forall n\in\mathbb{N}, \tag{27}$$

(per $n=1,2$ si ottengono le relazioni già note per x e x^2).

Una volta noto che è possibile calcolare esplicitamente la derivata di un qualsiasi polinomio, è naturale chiedersi se sia possibile fare lo stesso per funzioni razionali. Partiamo dal caso più semplice:

$$f(x)=\frac{1}{x}\qquad (x\neq 0)\qquad \Rightarrow\qquad \frac{f(x_1)-f(x)}{x_1-x}=\frac{\frac{1}{x_1}-\frac{1}{x}}{x_1-x}=\frac{x-x_1}{x_1x(x_1-x)}=-\frac{1}{x_1x}.$$

Quindi passando al limite $x_1\to x$, si ottiene la formula

$$\left(\frac{1}{x}\right)'=-\frac{1}{x^2}\qquad\qquad (x\neq 0).$$

Allo stesso modo è possibile trattare funzioni $f(x)=\frac{1}{x^\beta}$ con $\beta\in\mathbb{N}$ $(x\neq 0)$:

$$\frac{f(x_1)-f(x)}{x_1-x}=\frac{\frac{1}{x_1^\beta}-\frac{1}{x^\beta}}{x_1-x}=\frac{x^\beta-x_1^\beta}{x_1^\beta x^\beta(x_1-x)}=-\frac{x_1^{\beta-1}+x_1^{\beta-2}x+\cdots+x^{\beta-1}}{x_1^\beta x^\beta}.$$

Passando al limite per $x_1\to x$, si ottiene

$$(28)\qquad \left(x^{-\beta}\right)'\equiv\left(\frac{1}{x^\beta}\right)'=-\frac{\beta}{x^{\beta+1}}\equiv-\beta x^{-\beta-1}\qquad \forall\beta\in\mathbb{N},\quad \forall\, x\neq 0.$$

Vedremo più avanti come si possa calcolare la derivata di una generica funzione razionale.

Le formule (27) e (28) si possono sintetizzare nell'unica formula

$$(29)\qquad (x^\alpha)'=\alpha x^{\alpha-1}\qquad\qquad \forall\,\alpha\in\mathbb{Z}.$$

Dimostriamo che è possibile scegliere $\alpha\in\mathbb{Q}$ ottenendo ancora la formula (29). Supponiamo la funzione $f(x)=x^\alpha$ con $\alpha=p/q$ con p e q interi $(q\neq 0)$. Consideriamo, per semplicità, il caso $p,q>0$. Il rapporto incrementale è

$$\frac{x_1^\alpha-x^\alpha}{x_1-x}=\frac{x_1^{p/q}-x^{p/q}}{x_1-x}.$$

Ponendo $x_1^{1/q}=\xi_1$ e $x^{1/q}=\xi$, otteniamo

$$\frac{x_1^\alpha-x^\alpha}{x_1-x}=\frac{\xi_1^p-\xi^p}{\xi_1^q-\xi^q}=\frac{\xi_1^{p-1}+\xi_1^{p-2}\xi+\cdots+\xi^{p-1}}{\xi_1^{q-1}+\xi_1^{q-2}\xi+\cdots+\xi^{q-1}}.$$

Passando al limite per $x_1\to x$, cioè per $\xi_1\to\xi$, si ottiene

$$\lim_{x_1\to x}\frac{x_1^\alpha-x^\alpha}{x_1-x}=\lim_{\xi_1\to\xi}\frac{\xi_1^{p-1}+\xi_1^{p-2}\xi+\cdots+\xi^{p-1}}{\xi_1^{q-1}+\xi_1^{q-2}\xi+\cdots+\xi^{q-1}}=\frac{p\,\xi^{p-1}}{q\,\xi^{q-1}}=\frac{p}{q}\,\xi^{p-q}=\frac{p}{q}x^{\frac{p}{q}-1},$$

cioè la formula (29) per α razionale positivo.

In generale si può dimostrare che (29) vale *per ogni* $\alpha\in\mathbb{R}$, cioè

$$(30)\qquad (x^\alpha)'=\alpha x^{\alpha-1}\qquad\qquad \forall\,\alpha\in\mathbb{R}\quad \forall\, x\neq 0.$$

Funzioni trigonometriche. Grazie alle formule di addizione è possibile scrivere i rapporti incrementali di $\sin x$ e $\cos x$ come

$$\frac{\sin(x+h)-\sin x}{h}=\frac{\sin x\cos h+\cos x\sin h-\sin x}{h}=\sin x\,\frac{\cos h-1}{h}+\cos x\,\frac{\sin h}{h},$$
$$\frac{\cos(x+h)-\cos x}{h}=\frac{\cos x\cos h-\sin x\sin h-\cos x}{h}=\cos x\frac{\cos h-1}{h}-\sin x\,\frac{\sin h}{h}.$$

Passando al limite per $h\to 0$ e ricordando che $\lim\limits_{h\to 0}\frac{\cos h-1}{h}=0$ e $\lim\limits_{h\to 0}\frac{\sin h}{h}=1$,

$$\big(\sin x\big)'=\cos x \qquad \text{e} \qquad \big(\cos x\big)'=-\sin x.$$

Esponenziale e logaritmo. Come ultimo esempio, consideriamo le funzioni e^x e $\ln x$. Nel caso dell'esponenziale, il rapporto incrementale è

$$\frac{f(x+h)-f(x)}{h}=\frac{e^{x+h}-e^x}{h}=e^x\,\frac{e^h-1}{h}.$$

Passando al limite per $h\to 0$ e usando il limite notevole $\lim\limits_{h\to 0}\frac{e^h-1}{h}=1$,

$$\big(e^x\big)'=e^x,$$

che esprime una proprietà notevole dell'esponenziale: *la derivata di e^x è e^x*. In effetti, la funzione e^x è l'unica funzione f che verifica l'equazione (differenziale) $f'=f$ e la condizione $f(0)=1$.
Il rapporto incrementale del logaritmo naturale si riscrive come

$$\frac{f(x+h)-f(x)}{h}=\frac{\ln(x+h)-\ln x}{h}=\frac{1}{h}\ln\left(\frac{x+h}{x}\right)=\frac{1}{h}\ln\left(1+\frac{h}{x}\right).$$

Quindi, ponendo $t=h/x$ (x è fissato) e usando il limite notevole $\lim\limits_{t\to 0}\frac{\ln(1+t)}{t}=1$,

$$\lim_{h\to 0}\frac{f(x+h)-f(x)}{h}=\lim_{h\to 0}\frac{1}{h}\ln\left(1+\frac{h}{x}\right)=\lim_{t\to 0}\frac{1}{x}\,\frac{\ln(1+t)}{t}=\frac{1}{x}.$$

2. Regole fondamentali di derivazione

Dalla definizione dell'operazione di derivazione, discendono alcune regole basilari che permettono di derivare una classe ampia di funzioni, a partire da una classe più ristretta di derivate note.

<u>Linearità.</u> *Dati $\alpha,\beta\in\mathbb{R}$ e f,g derivabili, allora anche $\alpha f+\beta g$ è derivabile e*

$$\phi(x)=\alpha f(x)+\beta g(x) \qquad\Longrightarrow\qquad \phi'(x)=\alpha f'(x)+\beta g'(x).$$

Basta infatti osservare che il rapporto incrementale di ϕ si può riscrivere come

$$\frac{\phi(x+h)-\phi(x)}{h}=\alpha\,\frac{f(x+h)-f(x)}{h}+\beta\,\frac{g(x+h)-g(x)}{h},$$

e passare al limite per $h \to 0$, applicando le proprietà note dei limiti.

Ad esempio, la derivata di un polinomio $p(x) = a_n x^n + a_{n-1}x^{n-1} + \cdots + a_0$ si può calcolare senza bisogno di passare per il limite del rapporto incrementale, ma semplicemente usando la linearità della derivazione e la formula $(x^k)' = kx^{k-1}$:

$$\begin{aligned}(p(x))' &= a_n(x^n)' + a_{n-1}(x^{n-1})' + \cdots + a_1(x)' + (a_0)' \\ &= na_n x^{n-1} + (n-1)a_{n-1}x^{n-2} + \cdots + a_1.\end{aligned}$$

Derivata di un prodotto. *Date f, g derivabili, allora anche fg è derivabile e*

$$\phi(x) = f(x)g(x) \quad \Longrightarrow \quad \phi'(x) = f(x)g'(x) + f'(x)g(x).$$

Per dimostrare la formula, scriviamo il rapporto incrementale

$$\begin{aligned}\frac{\phi(x+h) - \phi(x)}{h} &= \frac{f(x+h)g(x+h) - f(x)g(x)}{h} \\ &= f(x+h)\,\frac{g(x+h) - g(x)}{h} + \frac{f(x+h) - f(x)}{h}\,g(x),\end{aligned}$$

(si è aggiunto e sottratto a numeratore la quantità $f(x+h)g(x)$). Per $h \to 0$, la conclusione.

Ad esempio, per calcolare la derivata della funzione $\phi(x) = x \sin x$,

$$(x \sin x)' = x(\sin x)' + (x)' \sin x = x \cos x + \sin x,$$

avendo usato le formule di derivazione per x e $\sin x$.

Derivata di un rapporto. *Se f e g sono derivabili ($g(x) \neq 0$ per ogni x), allora anche il rapporto f/g è derivabile e vale la formula*

$$\phi(x) = \frac{f(x)}{g(x)} \quad \Longrightarrow \quad \phi'(x) = \frac{f'(x)g(x) - f(x)g'(x)}{[g(x)]^2}.$$

La dimostrazione discende dalla struttura del rapporto incrementale per la funzione rapporto. Niente di sorprendente. Si ha:

$$\begin{aligned}\frac{\phi(x+h) - \phi(x)}{h} &= \frac{1}{h}\left[\frac{f(x+h)}{g(x+h)} - \frac{f(x)}{g(x)}\right] = \frac{f(x+h)g(x) - f(x)g(x+h)}{hg(x)g(x+h)} \\ &= \frac{f(x+h)g(x) - f(x)g(x) + f(x)g(x) - f(x)g(x+h)}{hg(x)g(x+h)} \\ &= \frac{1}{g(x)g(x+h)}\left[\frac{f(x+h) - f(x)}{h}\,g(x) - f(x)\,\frac{g(x+h) - g(x)}{h}\right].\end{aligned}$$

Per $h \to 0$, si ottiene la conclusione.

Osservazione 2.1. La formula di derivazione del rapporto è stata scritta nel caso in cui $g(x) \neq 0$ per ogni x. Ripercorrendo la dimostrazione ci si convince che basta supporre $g(x) \neq 0$ nel punto considerato. Infatti, se g è derivabile in x è anche continua nel punto, e quindi, c'è tutto un intorno I di x in cui g non si azzera.

Ad esempio, la derivata di $f(x) = \tan x$ è data da

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

Anche per derivare funzioni razionali basta applicare la formula di derivazione del rapporto. Ad esempio,

$$\left(\frac{x^2}{x+1}\right)' = \frac{2x(x+1) - x^2 \cdot 1}{(x+1)^2} = \frac{x(x+2)}{(x+1)^2}.$$

<u>*Derivata di una funzione composta.*</u> *Siano g, h derivabili, allora la funzione composta $f = h \circ g$ è derivabile e vale la formula (in inglese, nota come* chain rule*)*

$$f(x) = h(g(x)) \quad \Longrightarrow \quad f'(x) = h'(g(x))\, g'(x). \tag{31}$$

Usare concretamente questa regola è molto più semplice di quel che possa sembrare. Vediamo, ad esempio, come calcolare la derivata di $f(x) = e^{x^2}$.

i. Riconosciamo la struttura di funzione composta:

$$f(x) = h(g(x)) \qquad \text{dove} \qquad g(x) = x^2, \quad h(s) = e^s.$$

ii. Dato che $g(x) = x^2$ e $h(s) = e^s$, si ha $g'(x) = 2x$ e $h'(s) = e^s$.

iii. Ora occorre fare il prodotto delle derivate, *calcolando h' in $s = g(x) = x^2$*:

$$D\big(e^{x^2}\big) = 2xe^{x^2}.$$

Analogamente, dato che $D(\sin x) = \cos x$ e $D(\sqrt{s}) = 1/(2\sqrt{s})$,

$$D\left(\sqrt{1 + \sin x}\right) = \frac{\cos x}{2\sqrt{1 + \sin x}}.$$

Se la funzione è composta da più di due funzioni, si itera il procedimento:

$$D\left(h(g(f(x)))\right) = h'(g(f(x))) \cdot g'(f(x)) \cdot f'(x).$$

Ad esempio,

$$D\left(\sqrt{1 + \sin^2 x}\right) = \frac{1}{2\sqrt{1 + \sin^2 x}} \cdot 2 \sin x \cdot \cos x = \frac{\sin x \cos x}{\sqrt{1 + \sin^2 x}}.$$

Per dimostrare la formula (31), scriviamo il rapporto incrementale

$$\frac{\Delta f}{\Delta x} = \frac{\Delta h}{\Delta x} = \begin{cases} 0 & \text{se} \quad \Delta g = 0, \\ \dfrac{\Delta h}{\Delta g}\,\dfrac{\Delta g}{\Delta x} & \text{se} \quad \Delta g \neq 0, \end{cases} \tag{32}$$

dove

$$\Delta x = x_2 - x_1 \qquad\qquad \Delta f = f(x_2) - f(x_1)$$

$$\Delta h = h(g(x_2)) - h(g(x_1)) \qquad \Delta g = g(x_2) - g(x_1).$$

Se, per x_2 vicino ad x_1, si ha $\Delta g \neq 0$, la conclusione segue da

$$\lim_{\Delta x \to 0} \frac{\Delta f}{\Delta x} = \lim_{\Delta x \to 0} \frac{\Delta h}{\Delta g}\,\frac{\Delta g}{\Delta x} = \lim_{\Delta g \to 0} \frac{\Delta h}{\Delta g}\, \lim_{\Delta x \to 0} \frac{\Delta g}{\Delta x} = h'(g(x_1))\, g'(x_1),$$

dato che $\Delta g \to 0$ quando $\Delta x \to 0$. Se in ogni intorno di x_1 ci sono punti per cui $\Delta g = 0$, la derivata di g in x_1 deve essere nulla (come si dimostra?), e quindi vale la conclusione, dato che entrambe le rappresentazioni di $\Delta f/\Delta x$ in (32) tendono a zero per $\Delta x \to 0$.

Applicando la formula (31) è possibile ottenere le formule per le derivate di

$$x^\alpha \quad (\alpha \in \mathbb{R}) \qquad \text{e} \qquad a^x \qquad (a > 0).$$

Per entrambe è utile osservare utilizzare la formula

$$a^b = e^{b \ln a} \qquad \forall\, a > 0,\ b \in \mathbb{R}. \tag{33}$$

Usando la formula (33),

$$D(x^\alpha) = D(e^{\alpha \ln x}) = \frac{\alpha}{x}\, e^{\alpha \ln x} = \frac{\alpha x^\alpha}{x} = \alpha\, x^{\alpha - 1} \qquad \forall \alpha \in \mathbb{R},$$

$$D(a^x) = D(e^{x \ln a}) = e^{x \ln a}\, \ln a = a^x\, \ln a \qquad \forall\, a > 0.$$

<u>Derivata di una funzione inversa.</u> Una conseguenza della formula di derivazione di funzione composta è la formula della derivata dell'inversa di una funzione. La prima domanda naturale da porsi è: *se la funzione f è invertibile e derivabile, lo è anche la funzione inversa?* La risposta è immediata se si pensa a come si ottiene il grafico della funzione inversa a partire da quello della funzione originale e se si ricorda il significato geometrico della derivabilità: la funzione f è derivabile in x se in tale punto il grafico ammette tangente e tale retta tangente *non è verticale* (quando la tangente al grafico è verticale, il rapporto incrementale tende ad ∞). Il grafico di f^{-1} si può ottenere da quello della f tramite un ribaltamento attorno alla bisettrice del primo e terzo quadrante. In questa operazione di ribaltamento, rette orizzontali diventano verticali e viceversa. Quindi un punto in cui la tangente al grafico di f è orizzontale (cioè $f'(x) = 0$), corrisponde, nel grafico di f^{-1}, ad un punto in cui la tangente è verticale e viceversa (Fig.2). Questo significa che:

Se $f'(x) \neq 0$, la funzione inversa f^{-1} è derivabile nel punto $y = f(x)$.

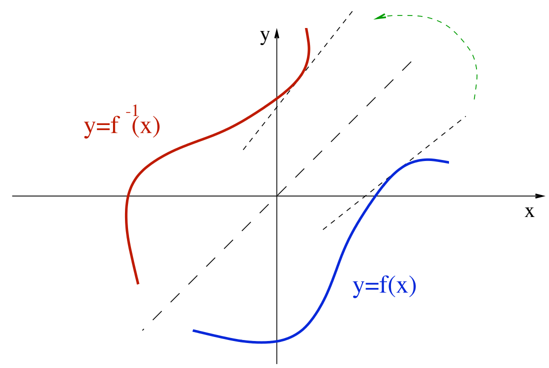

FIGURA 2. Una funzione e la sua inversa, con le relative tangenti.

Come calcolare la derivata della'inversa f^{-1}? Dato che $f(f^{-1}(x)) = x$ per ogni x, derivando membro a membro tramite la formula di derivazione di funzione composta,

$$f(f^{-1}(x)) = x \qquad \Longrightarrow \qquad f'(f^{-1}(x))(f^{-1})'(x) = 1.$$

Esplicitando $(f^{-1})'(x)$, si ottiene la formula di derivazione della funzione inversa

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}. \tag{34}$$

Verifichiamo questa formula, calcolando di nuovo la derivata della funzione $f(x) = \ln x$ (in precedenza la formula si è ottenuta in modo diverso). In questo caso

$$\begin{cases} f(t) = e^t \\ f^{-1}(x) = \ln x \end{cases} \Rightarrow \begin{cases} f'(t) = e^t \\ f'(f^{-1}(x)) = e^{\ln x} = x \end{cases} \Rightarrow \quad (\ln x)' = \frac{1}{x}.$$

Consideriamo le inverse delle funzioni trigonometriche e calcoliamone le derivate. Dato che $\cos t = \sqrt{1 - \sin^2 t}$ per $t \in [-\pi/2, \pi/2]$ e $\sin t = \sqrt{1 - \cos^2 t}$ per $t \in [0, \pi]$, si ha

$$\begin{cases} f(t) = \sin t \\ f^{-1}(x) = \arcsin x \end{cases} \Rightarrow \begin{cases} f'(t) = \cos t \\ f'(f^{-1}(x)) = \cos(\arcsin x) = \sqrt{1 - x^2} \end{cases}$$
$$\Rightarrow \quad (\arcsin x)' = \frac{1}{\sqrt{1 - x^2}} \qquad \forall x \in (-1, 1),$$

$$\begin{cases} f(t) = \cos t \\ f^{-1}(x) = \arccos x \end{cases} \Rightarrow \begin{cases} f'(t) = -\sin t \\ f'(f^{-1}(x)) = -\sin(\arccos x) = -\sqrt{1 - x^2} \end{cases}$$
$$\Rightarrow \quad (\arccos x)' = -\frac{1}{\sqrt{1 - x^2}} \qquad \forall x \in (-1, 1).$$

Per quanto riguarda la funzione $\arctan x$, è utile ricordare che

$$D(\tan t) = \frac{\cos^2 t + \sin^2 t}{\cos^2 t} = \frac{1}{\cos^2 t} = 1 + \tan^2 t.$$

Quindi

$$\begin{cases} f(t) = \tan t \\ f^{-1}(x) = \arctan x \end{cases} \Rightarrow \begin{cases} f'(t) = 1 + \tan^2 t \\ f'(f^{-1}(x)) = 1 + \tan^2(\arctan x) = 1 + x^2 \end{cases}$$
$$\Rightarrow \quad (\arctan x)' = \frac{1}{1+x^2}.$$

Ultima, ma non ultima, la formula della derivata di $f^{-1}(x) = \log_a x$ con $a > 0$ qualsiasi:

$$\begin{cases} f(x) = a^x \\ f^{-1}(x) = \log_a x \end{cases} \Rightarrow \begin{cases} f'(t) = a^t \ln a \\ f'(f^{-1}(x)) = a^{\log_a x} \ln a = x \ln a \end{cases} \Rightarrow \quad (\log_a x)' = \frac{1}{x \ln a}.$$

3. Derivate successive

L'operazione di derivazione porta da una funzione f ad una nuova funzione f'. E' naturale chiedersi se la funzione derivata f' sia a sua volta derivabile.

Definizione 3.1. *Sia $f : [a,b] \to \mathbb{R}$ derivabile e sia $x \in [a,b]$. Se esiste finito*

$$\lim_{h \to 0} \frac{f'(x+h) - f'(x)}{h}, \tag{35}$$

la funzione f è derivabile due volte in x, *il limite si indica con $f''(x)$ e si chiama* derivata seconda di f in x. *Come sempre, se f è derivabile due volte in tutti i punti dell'intervallo $[a,b]$, si dice che f è* derivabile due volte in $[a,b]$.

Per la derivata seconda si usano anche le notazioni

$$f'' = \frac{d^2 f}{dx^2} = D^2 f = \frac{d^2 y}{dx^2} = \cdots ,$$

Analogamente, nel caso di una funzione derivabile due volte, è possibile domandarsi se esista la derivata terza f'''. Iterando il procedimento si può parlare di derivata $n-$esima, che si indica con $f^{(n)}$. Simboli equivalenti sono

$$f^{(n)} \equiv D^n f \equiv \frac{d^n f}{dx^n} = \frac{d^n y}{dx^n}.$$

Qualche volta si indica la funzione f come la sua derivata $0-$esima: $f^{(0)} \equiv f$.

La maniera *operativa* di calcolare derivate successive è semplicemente di iterare le formule note per la derivazione. Ad esempio,

$$f(x) = x^3 + x \quad \Rightarrow \quad f'(x) = 3x^2 + 1 \quad \Rightarrow \quad f''(x) = 6x \quad \Rightarrow \quad f'''(x) = 6.$$

Le derivate di ordine superiore al terzo della funzione $f(x) = x^3 + x$ esistono e sono tutte nulle. In generale, un polinomio p di grado n è infinitamente derivabile (cioè ammette derivate di qualsiasi ordine), e le sue derivate di ordine maggiore o uguale ad $n+1$ sono tutte nulle. Anche le funzioni $\sin x$ e $\cos x$ sono infinitamente derivabili:

$$D(\sin x) = \cos x,\ D^2(\sin x) = -\sin x,\ D^3(\sin x) = -\cos x,\ D^4(\sin x) = \sin x$$
$$D(\cos x) = -\sin x,\ D^2(\cos x) = -\cos x,\ D^3(\cos x) = \sin x,\ D^4(\cos x) = \cos x.$$

Le derivate successive ripetono lo stesso schema in modo periodico, ossia

$$\begin{matrix} D^{2n-1}(\sin x) = (-1)^{n+1}\cos x, & D^{2n}(\sin x) = (-1)^n \sin x, \\ D^{2n-1}(\cos x) = (-1)^n \sin x, & D^{2n}(\cos x) = (-1)^n \cos x, \end{matrix} \qquad \forall n \in \mathbb{N}.$$

Pensando al caso di polinomi e funzioni trigonometriche, si potrebbe essere indotti a credere che tutte le funzioni siano infinitamente derivabili. Un esempio di funzione che sia derivabile due volte in un punto, ma non tre volte è $f(x) = x^{5/2}$. Infatti $f''(x) = \frac{15}{4}\sqrt{x}$ che, come sappiamo, non è derivabile in zero.

Esercizio 3.2. *Se $f(x) = \cos x$, quanto vale l'espressione $f''(x) + f(x)$? E se, dato $\lambda \in \mathbb{R}$, $g(x) = e^{\lambda x}$, quanto vale $g''(x) - \lambda^2 g(x)$?*

Notazioni. Comunememente si usano le notazioni (qui I è un intervallo aperto e $k \in \mathbb{N}$)

$$\begin{aligned} &C(I) \equiv C^0(I) := \{\text{funzioni continue in } I\} \\ &C^1(I) := \{\text{funzioni derivabili in } I \text{ e con } f' \in C(I)\} \\ &C^k(I) := \{\text{funzioni derivabili } k \text{ volte in } I \text{ e con } f^{(k)} \in C(I)\} \\ &C^\infty(I) := \{\text{funzioni infinitamente derivabili in } I\}. \end{aligned}$$

4. Il Teorema di Lagrange

Dato che il rapporto incrementale è determinato dai valori della funzione in due punti distinti, esso riflette proprietà della funzione "in grande". Invece, la derivata, che si ottiene con un procedimento di limite, riflette solo proprietà "in piccolo". E' molto utile poter dedurre proprietà *globali* della funzione (cioè "in grande") a partire da proprietà *locali* (cioè "in piccolo") date dalla derivata prima della funzione. Lo strumento più utile per questa operazione è il *teorema di Lagrange* (o *teorema del valor medio del calcolo differenziale*).[1]

Graficamente il Teorema di Lagrange afferma che data una funzione f continua nell'intervallo chiuso $[x_1, x_2]$ e derivabile nell'intervallo aperto (x_1, x_2), la retta passante per i punti $(x_1, f(x_1))$ e $(x_2, f(x_2))$ (detta retta secante) è parallela alla retta tangente

[1] Nei testi americani, spesso il Teorema di Lagrange è denominato "mean value theorem of differential calculus" o "intermediate value theorem".

al grafico nel punto $(\xi, f(\xi))$ per almeno un valore $\xi \in (x_1, x_2)$. Dato che il coefficiente angolare della secante è

$$\frac{\Delta f}{\Delta x} = \frac{f(x_2) - f(x_1)}{x_2 - x_1},$$

per questo valore intermedio ξ vale la relazione $f'(\xi) = [f(x_2) - f(x_1)]/[x_2 - x_1]$.

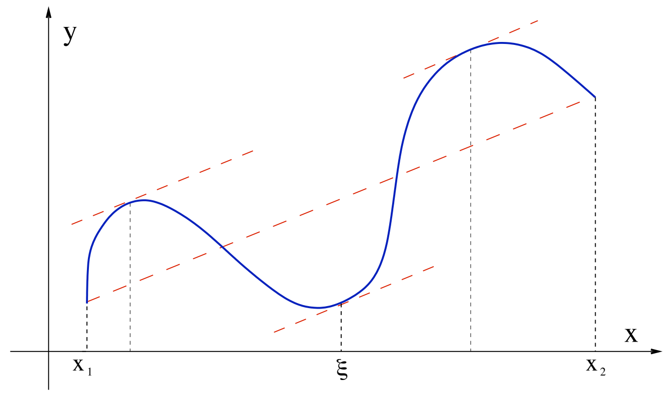

FIGURA 3. Il teorema di Lagrange

TEOREMA 4.1. <u>*Teorema di Lagrange.*</u> *Sia f continua in $[x_1, x_2]$ e derivabile in (x_1, x_2). Allora esiste $\xi \in (x_1, x_2)$ tale che*

$$f'(\xi) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

La tesi del Teorema equivale ad affermare che esiste $\theta \in (0, 1)$ per cui

$$f'(x_1 + \theta(x_2 - x_1)) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Le due formulazioni sono equivalenti dato che il punto intermedio ξ può sempre essere scritto nella forma $\xi = x_1 + \theta(x_2 - x_1)$ per $\theta \in (0, 1)$ opportuno. Oppure, sostituendo x_1 con x e x_2 con $x + h$, possiamo scrivere

$$\frac{f(x + h) - f(x)}{h} = f'(x + \theta h), \qquad \theta \in (0, 1).$$

OSSERVAZIONE 4.2. Il punto di partenza nella definizione di derivabilità è dare solidità ad approssimazioni del tipo

$$\Delta f \approx f'(x_0)\Delta x \qquad \text{per} \quad x \approx x_0.$$

dove $\Delta f = f(x) - f(x_0)$ e $\Delta x = x - x_0$. Il Teorema di Lagrange garantisce che $\Delta f = f'(\xi)\Delta x$ per qualche ξ compreso tra x e x_0. Quindi se si è disposti a pagare il prezzo di calcolare la derivata di f in un misterioso punto ξ, anziché in x_0, l'errore commesso è nullo (ma non si dimentichi che ξ <u>dipende</u> da x e da x_0).

Controesempio 4.3. *"Datemi un punto (interno) di non derivabilità, e vi darò un controesempio."* Se la funzione f non è derivabile in tutti i punti dell'intervallo aperto (x_1, x_2), non è detto che valga la conclusione del Teorema di Lagrange: può capitare che nessuna parallela della secante che congiunge gli estremi del grafico sia tangente al grafico stesso. Consideriamo la funzione $f(x) = |x|$ nell'intervallo $[-1, 1]$. Questa funzione è derivabile per ogni $x \neq 0$ e si ha

$$D(|x|) = \begin{cases} -1 & -1 \leq x < 0, \\ +1 & 0 < x \leq 1, \end{cases}$$

ma, dato che

$$\frac{f(1) - f(-1)}{1 - (-1)} = \frac{1-1}{2} = 0 \neq D(|x|) \qquad \forall x,$$

la conclusione del Teorema non vale.

Il Teorema di Lagrange è conseguenza del seguente risultato.

Teorema 4.4. Teorema di Rolle. *Sia ϕ continua in $[x_1, x_2]$ e derivabile in (x_1, x_2). Se $\phi(x_1) = \phi(x_2)$, allora esiste $\xi \in (x_1, x_2)$ tale che $\phi'(\xi) = 0$.*

Geometricamente, il Teorema di Rolle afferma che, se $\phi(x_1)$ e $\phi(x_2)$ coincidono allora il grafico di ϕ ha tangente orizzontale in un punto interno dell'intervallo (x_1, x_2).

Dimostrazione del Teorema 4.4. Sia $\ell = \phi(x_1) = \phi(x_2)$. Dato che la funzione ϕ è continua in $[x_1, x_2]$, per il Teorema di Weierstrass, esistono sia il massimo M che il minimo m di ϕ in $[x_1, x_2]$. Chiaramente, $m \leq \ell \leq M$.

Se $M = m$, deve essere $\phi(x) = M$ in tutto l'intervallo $[x_1, x_2]$, quindi $\phi'(x) = 0$ in tutti i punti dell'intervallo.

Se $M \neq m$, almeno uno dei due valori deve essere diverso da ℓ. Supponiamo che sia $M \neq \ell$ (l'altro caso si tratta in modo simile). Allora $M > \ell$ ed esiste $\xi \in [x_1, x_2]$ tale che $\phi(\xi) = M$. Inoltre visto che $\phi(x_1) = \phi(x_2) = \ell \neq M$, $\xi \neq x_1, x_2$, ossia $\xi \in (x_1, x_2)$. Dato che $\phi(x) \leq M = \phi(\xi)$ per ogni $x \in [x_1, x_2]$,

$$\frac{\phi(x) - \phi(\xi)}{x - \xi} \quad \begin{cases} \leq 0 & \forall\, x > \xi, \\ \\ \geq 0 & \forall\, x < \xi, \end{cases}$$

Passando al limite per $x \to \xi$ da destra e da sinistra e, sapendo che i limiti destro e sinistro esistono e coincidono, si ha

$$\phi'(\xi) = \lim_{x \to \xi^+} \frac{\phi(x) - \phi(\xi)}{x - \xi} \leq 0 \quad \text{e} \quad \phi'(\xi) = \lim_{x \to \xi^-} \frac{\phi(x) - \phi(\xi)}{x - \xi} \geq 0$$

da cui $0 \leq \phi'(\xi) \leq 0$, e quindi $\phi'(\xi) = 0$. $\square$

Dimostrazione del Teorema 4.1. Definiamo la funzione (ausiliaria) ϕ

$$\phi(x) := f(x) - f(x_1) - \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1),$$

che rappresenta la distanza verticale tra il punto $(x, f(x))$ del grafico della funzione e la retta secante passante per i suoi estremi. La funzione ϕ soddisfa le ipotesi di regolarità del Teorema di Rolle (cioè è continua in $[x_1, x_2]$ e derivabile in (x_1, x_2)). Inoltre

$$\phi(x_1) = f(x_1) - f(x_1) - \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x_1 - x_1) = 0,$$
$$\phi(x_2) = f(x_2) - f(x_1) - \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x_2 - x_1) = 0.$$

Quindi esiste un valore $\xi \in (x_1, x_2)$ tale che $\phi'(\xi) = 0$. Dato che

$$\phi'(x) = f'(x) - \frac{f(x_2) - f(x_1)}{x_2 - x_1} \qquad x \in (x_1, x_2),$$

si deduce che

$$\phi'(\xi) = f'(\xi) - \frac{f(x_2) - f(x_1)}{x_2 - x_1} = 0,$$

cioè la conclusione. □

Conseguenze del Teorema di Lagrange. L'apparentemente innocuo Teorema di Lagrange è un'arma estremamente potente. Vediamo perché.

a. <u>Funzioni monotòne.</u> Sia f derivabile in (a, b). Allora

$$f'(x) > 0 \quad \forall\, x \in (a, b) \quad \Longrightarrow \quad f \quad \text{strettamente crescente in } (a, b).$$

Infatti, supponiamo $f'(x) > 0$ per ogni $x \in (a, b)$ e siano x_1, x_2 in (a, b) tali che $x_1 < x_2$. Per il Teorema di Lagrange, esiste $\xi \in (x_1, x_2) \subset (a, b)$ tale che

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1).$$

Dato che $f'(\xi) > 0$ per ipotesi, ne segue $f(x_2) > f(x_1)$.

Analogamente si dimostra che

$$f'(x) < 0 \quad \forall\, x \in (a, b) \quad \Longrightarrow \quad f \quad \textit{strettamente decrescente in } (a, b).$$

Se invece dell'informazione $f'(x) > 0$ o $f'(x) < 0$, si ha l'informazione più debole $f'(x) \geq 0$ o $f'(x) \leq 0$, la conclusione va sostituita con le analoghe proprietà di monotonía deboli (nondecrescente/noncrescente).

Consideriamo, come esempio, la funzione

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

e studiamone la monotonía. Da quanto si è appena detto, basta studiare il segno della derivata prima:

$$\left(\frac{1}{1+x^2}\right)' = -\frac{2x}{(1+x^2)^2},$$

Dato che $f'(x)$ è strettamente positiva per $x < 0$ e strettamente negativa per $x > 0$, la funzione è crescente in $(-\infty, 0]$ ed è decrescente $[0, +\infty)$.

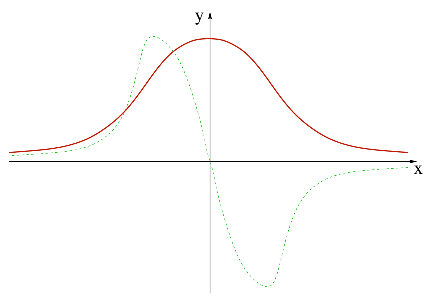

FIGURA 4. Grafico (qualitativo) di $1/(1+x^2)$ e della sua derivata

Vediamo un secondo esempio. Consideriamo la funzione $f(x) = \frac{1}{x}$. Dato che la derivata di questa funzione è

$$f'(x) = -\frac{1}{x^2} < 0 \qquad \forall x \neq 0,$$

concludiamo che la funzione f è decrescente... Se però calcoliamo la differenza del valore della funzione in 1 e in -1, otteniamo una contraddizione: $f(1) - f(-1) = 1 + 1 > 0$. Cosa sta succedendo? Bisogna stare attenti al fatto che le conclusioni sulla monotonìa delle funzioni seguono dal Teorema di Lagrange che vale su intervalli, cioè su insiemi "senza buchi" (si dicono *insiemi connessi*). Se togliamo dall'enunciato del Teorema l'ipotesi di "assenza di buchi", la conclusione non è più vera.[2] Nel caso della funzione $1/x$ stiamo applicando il Teorema all'insieme $(-\infty, 0) \cup (0, +\infty)$ che invece ha un buco: non contiene il punto 0. Ecco l'errore. Quindi la funzione $f(x) = 1/x$ NON è decrescente in $\mathbb{R} \setminus \{0\}$! Possiamo invece correttamente applicare i risultati sulla monotonìa alle semirette $(-\infty, 0)$ e $(0, \infty)$ *separatamente* e concludere che $\frac{1}{x}$ è decrescente in $(-\infty, 0)$ ed è decrescente in $(0, +\infty)$.

OSSERVAZIONE 4.5. Cogliamo l'occasione per far notare una sottigliezza. Se $f'(x) > 0$ in un intervallo, necessariamente la funzione f è crescente. Cosa succede se $f'(x_0) > 0$ *nel solo punto* x_0? La possibilità di tracciare la retta tangente (che è crescente) suggerirebbe il fatto che la funzione f sia crescente, per lo meno in un intorno di x_0. Invece

[2]Da cui il noto modo di dire, attribuito a N. Barbecue, "Non tutti i Teoremi riescono col buco"...

questa affermazione è falsa! Consideriamo la funzione

$$f(x) = \begin{cases} \frac{1}{2}x + x^2 \left(2 - \sin\left(\frac{1}{x}\right)\right) & x \neq 0, \\ 0 & x = 0, \end{cases}$$

Questa funzione è derivabile dappertutto e (c'è da dirlo? ...verificare!)

$$f'(x) = \begin{cases} 1/2 + \cos\left(\frac{1}{x}\right) + 2x\left(2 - \sin\left(\frac{1}{x}\right)\right) & x \neq 0, \\ 1/2 & x = 0, \end{cases}$$

Quindi $f'(0) = \frac{1}{2} > 0$, ma in ogni intorno di $x = 0$ cadono punti in cui la derivata è negativa: si tratta dei punti in cui $\cos\left(\frac{1}{x}\right)$ è uguale a -1. Quindi *non è vero che f è crescente in un intorno dell'origine.* Il problema sta nel fatto che f' non è continua in 0. Se fosse stata continua, $f'(x_0) > 0$ avrebbe implicato $f'(x) > 0$ in un intorno di x_0 e quindi la monotonia in tale intorno.

b. *Funzioni a derivata nulla.* Una seconda conseguenza del Teorema di Lagrange:

$$f'(x) = 0 \quad \forall\, x \in (a, b) \quad \Longrightarrow \quad f \quad \textit{è costante in } (a, b).$$

Infatti, per ogni coppia di valori $x_1, x_2 \in (a, b)$, esiste un valore ξ, compreso tra i due, per cui $f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1)$. Dato che $f'(x) = 0$ per ogni $x \in (a, b)$, si avrà, in particolare, $f'(\xi) = 0$, cioè

$$f(x_2) - f(x_1) = f'(\xi)(x_2 - x_1) = 0 \quad \Longrightarrow \quad f(x_2) = f(x_1).$$

Si noti che, anche qui, ha un ruolo fondamentale il fatto che si lavori su intervalli. Ad esempio, la funzione f definita da

$$f(x) = \begin{cases} 0 & x \in [0, 1], \\ 1 & x \in [2, 3], \end{cases}$$

è derivabile nel suo insieme di definizione $[0, 1] \cup [2, 3]$ e la sua derivata è ovunque nulla, ma la funzione si guarda bene dall'essere costante.

c. *Lipschitzianità di funzioni a derivata limitata.* Siano I l'intervallo aperto (α, β) e $f : I \to \mathbb{R}$ una funzione derivabile in I. Se f' è limitata in I, cioè se

$$\exists\, L > 0 \quad \text{tale che} \quad |f'(x)| \leq L \quad \forall\, x \in I.$$

allora dal Teorema di Lagrange segue

$$|f(x_2) - f(x_1)| = |f'(\xi)(x_2 - x_1)| \leq L|x_2 - x_1| \qquad \forall x_1, x_2 \in (\alpha, \beta).$$

Quindi *una funzione derivabile con derivata limitata è lipschitziana.*

In particolare, se $f \in C^1(I)$ (sempre con $I = (\alpha, \beta)$), la derivata prima è continua in $[a, b]$ per ogni $[a, b] \subset I$ e quindi, per il Teorema di Weierstrass, è limitata in $[a, b]$. Se ne deduce che le funzioni in $C^1(I)$ sono **localmente lipschitziane**, cioè lipschitziane in ogni intervallo chiuso e limitato contenuto in I.

d. <u>*Approssimazione lineare.*</u> Un'ulteriore applicazione interessante del Teorema di Lagrange è la stima dell'errore che si commette approssimando una funzione con la sua tangente in un punto. Sia f derivabile in $[a, b]$. Supponiamo conoscere il valore della funzione f e della sua derivata prima f' in un punto assegnato $x_0 \in [a, b]$. Si può pensare che il valore della funzione f in un qualsiasi altro punto sia dato *approssimativamente* dal valore della funzione lineare che definisce la tangente al grafico di f in x_0, cioè

$$f(x) \approx f(x_0) + f'(x_0)(x - x_0).$$

Questo corrisponde ad approssimare il grafico della funzione f con quello della sua tangente. *E' possibile stimare l'errore che commettiamo facendo questa approssimazione?* Consideriamo un esempio concreto. Vogliamo calcolare, in modo approssimato, il valore di $\sin(1/10)$. Dato che $1/10$ è ragionevolmente vicino a 0, possiamo pensare di approssimare la funzione $\sin x$ con la sua tangente in $x = 0$, cioè $\sin(x) \approx x$. Calcolando in $x = 0,1$ otteniamo l'approssimazione richiesta

$$\sin(0,1) \approx 0,1.$$

Il problema fondamentale è: quant'è grande l'errore commesso? In altri termini, è possibile stimare $|\sin(0,1) - 0,1|$?

Torniamo al caso generale. Supponiamo di lavorare con una funzione f che sia *derivabile due volte* nell'intervallo (a, b) e supponiamo che la derivata seconda f'' sia limitata, cioè esista $M > 0$ tale che $|f''(x)| \leq M$ per ogni $x \in [a, b]$. Dato $x_0 \in [a, b]$, vogliamo stimare *il valore assoluto* della quantità

$$R(x; x_0) = f(x) - f(x_0) - f'(x_0)(x - x_0).$$

Applicando il Teorema di Lagrange otteniamo l'espressione

$$R(x; x_0) = f'(\xi)(x - x_0) - f'(x_0)(x - x_0) = \big(f'(\xi) - f'(x_0)\big)(x - x_0),$$

dove ξ è compreso tra x e x_0. Applicando il Teorema di Lagrange a $f'(\xi) - f'(x_0)$

$$R(x; x_0) = f''(\eta)(\xi - x_0)(x - x_0),$$

dove η è tra ξ e x_0. Quindi il valore assoluto dell'errore $R(x; x_0)$ è stimato da

$$|R(x; x_0)| = |f''(\eta)||\xi - x_0||x - x_0| \leq M|x - x_0|^2, \tag{36}$$

dove si è usata la limitatezza della derivata seconda f'' e il fatto che $|\xi - x_0| \leq |x - x_0|$. Nel caso-modello di $f(x) = \sin x$, $x_0 = 0$ e $x = 1/10$, si ha $M = 1$, pertanto

$$\left|R\left(10^{-1}; 0\right)\right| \leq \frac{1}{100},$$

dove si è usato che $|f''(x)| = |-\sin x| \leq 1$ e $|x - x_0| = 1/10$. Quindi

$$0.09 < \sin(0.1) < 0.11$$

e. *Derivabilità tramite il limite della derivata.* In alcune situazioni, capita di lavorare con funzioni definite tramite formule diverse in diversi intervalli. Consideriamo come caso modello una funzione della forma

$$f(x) = \begin{cases} f_1(x) & x < x_0, \\ \ell & x = x_0, \\ f_2(x) & x > x_0, \end{cases}$$

dove $\ell \in \mathbb{R}$, e f_1, f_2 sono funzioni note. La domanda naturale è se la funzione f sia derivabile nel punto x_0 oppure no. Come abbiamo già visto, la derivabilità implica la continuità, quindi, prima di tutto, deve essere verificata la condizione

$$\lim_{x \to x_0^-} f_1(x) = \ell = \lim_{x \to x_0^+} f_2(x).$$

Se questa condizione non è verificata, la funzione non è continua in x_0 e quindi, a maggior ragione, non è neanche derivabile in x_0. Nel caso in cui la funzione sia continua in x_0, per stabilirne la derivabilità occorre calcolare il limite del rapporto incrementale in x_0. Dato che la funzione f è definita da espressioni diverse a seconda che ci si trovi a destra o a sinistra di x_0, è sensato calcolare il limite del rapporto incrementale da destra e da sinistra.[3] Per definizione, *la funzione f è derivabile in x_0 se e solo se questi limiti esistono e coincidono*, ossia se e solo se

$$\lim_{x \to x_0^-} \frac{f_1(x) - \ell}{x - x_0} = \lim_{x \to x_0^+} \frac{f_2(x) - \ell}{x - x_0}.$$

La derivata in x_0 è il valore comune di questi due limiti.

In molte situazioni le f_1 e f_2 sono funzioni derivabili in tutto il loro insieme di definizione ed è possibile calcolare esplicitamente la funzione derivata. Invece di calcolare il limite del rapporto incrementale, può essere più semplice calcolare le derivate f_1' e f_2' nei rispettivi domini e calcolare il limite di queste funzioni derivate. Quale informazione dà questa procedura?

Proposizione 4.6. *Dato $x_0 \in \mathbb{R}$ e $r > 0$, sia f continua in x_0 e derivabile in $(x_0 - r, x_0 + r) \setminus \{x_0\}$ e supponiamo che esistano finiti il limite destro e sinistro $\lim\limits_{x \to x_0^\pm} f'(x) = \ell^\pm$. Allora f è derivabile in x_0 se e solo se $\ell^+ = \ell^-$.*

[3]Se esiste il limite destro del rapporto incrementale di una funzione f in x_0, si dice che *f è derivabile da destra in x_0*. Analogamente per il limite sinistro. Per indicare il limite destro/sinistro del rapporto incrementale (qualora esistano), cioè per indicare la *derivata destra/sinistra* si usa il simbolo $D_\pm f(x_0)$, o varianti.

Dimostrazione. Grazie al Teorema di Lagrange,

$$\frac{f(x)-f(x_0)}{x-x_0}=\begin{cases} f'(\xi_-) & \text{se } x<x_0, \qquad (x<\xi_-<x_0) \\ \\ f'(\xi_+) & \text{se } x>x_0, \qquad (x_0<\xi_+<x) \end{cases}$$

dove $\xi_\pm$ sono punti opportuni tra x e x_0. Passando al limite per $x\to x_0$ da sinistra, dato che il limite sinistro della derivata f' esiste ed è uguale ad ℓ^-,

$$\lim_{x\to x_0^-}\frac{f(x)-f(x_0)}{x-x_0}=\lim_{x\to x_0^-}f'(\xi_-)=\ell^-.$$

Analogamente per il limite destro. Quindi, nelle ipotesi della Proposizione 4.6, i limiti destro e sinistro del rapporto incrementale esistono e sono uguali, rispettivamente, a ℓ^+ e ℓ^-. A questo punto, la conclusione è evidente. □

Ad esempio, studiamo la derivabilità in 0 della funzione $f(x)=x|x|$. Dato che

$$x|x|=\begin{cases} -x^2 & x<0, \\ x^2 & x\geq 0, \end{cases}$$

la funzione è certamente derivabile per $x\neq 0$ e

$$D\left(x|x|\right)=\begin{cases} -2x & x<0, \\ 2x & x>0. \end{cases}$$

Dato che $\lim\limits_{x\to 0^-}-2x=\lim\limits_{x\to 0^+}2x=0$, la funzione è derivabile in 0.

Consideriamo, invece, la funzione $e^{-|x|}$. In questo caso

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

La funzione è derivabile per $x\neq 0$ e

$$D\left(e^{-|x|}\right)=\begin{cases} e^x & x<0, \\ -e^{-x} & x>0. \end{cases}$$

Dato che $\lim\limits_{x\to 0^-}e^x=1\neq -1=\lim\limits_{x\to 0^+}-e^{-x}$, la funzione non è derivabile in 0.

La verifica della derivabilità in x_0 tramite il calcolo del limite della derivata a destra e a sinistra di x_0 è lecita *solo quando la derivata ammetta limiti destro e sinistro in* x_0. Quando questi limiti non esistano, il criterio non è più valido. La funzione può essere derivabile o può non esserlo. Ad esempio, consideriamo la funzione

$$f(x)=\begin{cases} x^2\sin\left(\frac{1}{x}\right) & x\neq 0, \\ 0 & x=0. \end{cases}$$

Per $x\neq 0$, la derivata prima f' di questa funzione è

$$f'(x)=2x\sin\left(\frac{1}{x}\right)-\cos\left(\frac{1}{x}\right).$$

Per $x \to 0$, il primo dei due termini è infinitesimo, mentre il secondo non ammette limite, quindi non esiste $\lim_{x \to 0^\pm} f'(x)$. La Proposizione 4.6 non è applicabile. Per studiare la derivabilità in zero, calcoliamo direttamente il limite del rapporto incrementale

$$\lim_{h \to 0} \frac{h^2 \sin(1/h) - 0}{h - 0} = \lim_{h \to 0} h \sin\left(\frac{1}{h}\right) = 0.$$

Quindi la funzione è derivabile in 0 e $f'(0) = 0$.

Tabella delle derivate

funzione f	derivata prima f'	funzione f	derivata prima f'
costante	0	x^α	$\alpha x^{\alpha-1}$
$\sin x$	$\cos x$	$\cos x$	$-\sin x$
e^x	e^x	$\ln x$	$\frac{1}{x}$
$\tan x$	$\frac{1}{\cos^2 x} = 1 + \tan^2 x$	$\cot x$	$-\frac{1}{\sin^2 x} = -1 - \cot^2 x$
a^x	$a^x \ln a$	$\arctan x$	$\frac{1}{1+x^2}$
$\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$	$\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$
$\sinh x$	$\cosh x$	$\cosh x$	$\sinh x$

CAPITOLO 6

Analisi locale e analisi globale

Una proprietà di una funzione f è *locale* se dipende dal comportamento della funzione nell'intorno di un punto x. Continuità e derivabilità in x sono proprietà locali. Una proprietà di una funzione f è *globale* se vale in tutto l'insieme di definizione della f. Ad esempio le funzioni $e^x, \arctan x, x^3, \dots$ sono funzioni globalmente monotòne crescenti e pertanto (globalmente) invertibili. Nella prima parte di questo capitolo approfondiamo l'uso della derivazione per determinare proprietà locali di funzioni: massimi/minimi relativi, punti di singolarità,... Torneremo più avanti sulle proprietà globali, concentrandoci sul problema di determinare massimi e minimi (assoluti) di una funzione assegnata.

1. Punti stazionari

Abbiamo già definito massimo e minimo di una funzione: data $f : D \to \mathbb{R}$, un punto $x_0 \in D$ è punto di massimo di f se $f(x) \leq f(x_0)$ per ogni $x \in D$. Il valore $f(x_0) = \max_{x \in D} f(x)$ è il massimo della funzione f in D. Analogo per i minimi.

L'esistenza di massimo e/o minimo è una proprietà globale della funzione. E' utile introdurre un analogo locale del concetto di massimo e di minimo.

Definizione 1.1. Massimo e minimo locale. *Il punto $x_0 \in D$ è un* punto di massimo locale *(o* relativo*) e il valore $f(x_0)$ è un* massimo locale *(o* relativo*) di f se esiste un intorno $(x_0 - \delta, x_0 + \delta)$ del punto x_0 tale che $f(x_0)$ è il massimo di f in $(x_0 - \delta, x_0 + \delta)$:*

$$\exists\, \delta > 0 \quad \textit{tale che} \quad f(x) \leq f(x_0) \qquad \forall x \in D \cap (x_0 - \delta, x_0 + \delta).$$

Analogamente per il minimo locale.

Un punto x_0 che sia o di massimo o minimo locale è un punto di estremo locale.

Per distinguere in modo più chiaro il massimo e il minimo dagli analoghi concetti locali, si parla di massimo globale (o assoluto) e di minimo globale (o assoluto). Dalla definizione segue immediatamente che se x_0 è punto di massimo globale, allora è anche punto di massimo locale. Il viceversa invece non è vero, come nel caso del grafico rappresentato in Figura 1. Per un esempio analitico, si può considerare la funzione

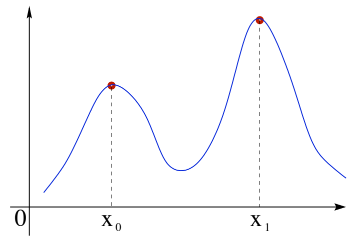

FIGURA 1. Il punto x_0 è di massimo locale, ma non globale; il punto x_1 è di massimo globale.

$f(x) = x^4 - x^2$. Dato che $f(0) = 0$ e $f(x) \leq 0$ per $x \in (-1, 1)$, il punto $x = 0$ è un punto di massimo locale, ma non è di massimo globale dato che $\lim_{x \to \pm\infty} f(x) = +\infty$.

DEFINIZIONE 1.2. *Sia $D \subset \mathbb{R}$. Un punto $x_0 \in \mathbb{R}$ è* interno a *D se esiste un intorno di x_0 interamente contenuto in D, cioè se esiste $\delta > 0$ per cui $(x_0 - \delta, x_0 + \delta) \subset D$.*

Se una funzione f ha un massimo o un minimo locale in corrispondenza di un punto x_0 interno all'insieme di definizione e in cui la funzione è derivabile, necessariamente

$$f'(x_0) = 0.$$

Basta infatti pensare alla necessaria posizione orizzontale della retta tangente (Fig.2(a)). Per una dimostrazione analitica, sia x_0 un punto di massimo locale interno e supponiamo f derivabile in x_0. Dato che $f(x) \leq f(x_0)$ per ogni $x \in (x_0 - \delta, x_0 + \delta)$

$$\frac{f(x) - f(x_0)}{x - x_0} \begin{cases} \leq 0 & \forall\, x_0 < x < x_0 + \delta, \\ \\ \geq 0 & \forall\, x_0 - \delta < x < x_0, \end{cases}$$

Passando al limite per $x \to x_0$ si deduce $f'(x_0) = 0$.

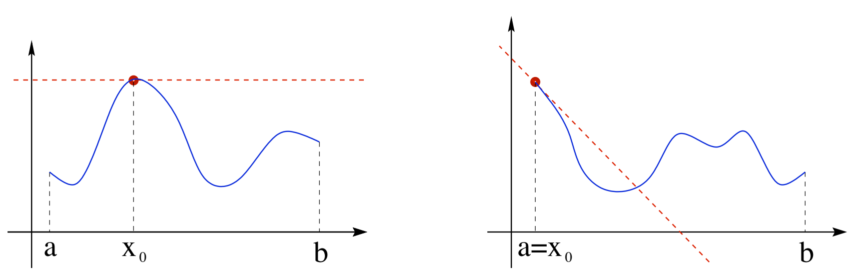

FIGURA 2. Se il punto di massimo locale è interno, la tangente è orizzontale. Se, invece, si trova sul bordo... non è detto!

Nel caso in cui il punto x_0 non sia interno al dominio, non è detto che la retta tangente sia orizzontale (Fig.2(b)). Conclusioni analoghe per i punti di minimo.

Pertanto, risolvere l'equazione $f'(x) = 0$ permette di determinare i possibili candidati a punti di minimo o massimo locale *interno* in cui f è *derivabile*. È comunque possibile che un estremo locale cada in un punto in cui la funzione non è derivabile; ad esempio $f(x) = |x|$ ha un minimo (globale) in $x = 0$ dove non ha retta tangente.

DEFINIZIONE 1.3. Punto stazionario. *Sia $f : D \to \mathbb{R}$. Un punto x_0, interno a D, tale che $f'(x_0) = 0$ si dice* punto stazionario[1] *(o* punto critico*) della funzione f.*

Equivalentemente, si può affermare che i punti critici di f sono i punti x per cui la tangente al grafico di f in $(x, f(x))$ è orizzontale.

ESERCIZIO 1.4. *Determinare i punti critici di $f(x) = x^7 + 14x^4 + 1$.*

Classificazione dei punti stazionari. Se x_0 è un punto di massimo o di minimo locale interno e f è derivabile in x_0, necessariamente x_0 è un punto critico, cioè $f'(x_0) = 0$. *Il viceversa non è vero*: esistono punti x_0 tali che $f'(x_0) = 0$, ma che non sono né punti di massimo locale, né punti di minimo locale. Ad esempio, la funzione $f(x) = x^3$ è strettamente crescente (quindi non ha né punti di massimo né punti di minimo in $\mathbb{R}$), ma $f'(x) = 3x^2$ si azzera nel punto $x = 0$.

Conoscendo il segno della derivata prima alla destra e alla sinistra del punto in questione, grazie al legame tra monotonia e segno di f', si può individuare quando un punto stazionario sia di massimo o di minimo. Supponiamo f derivabile in $(x_0-\delta, x_0+\delta)$ con $\delta > 0$, allora

$$f'(x) \begin{cases} \geq 0 & x_0 - \delta < x < x_0, \\ \\ \leq 0 & x_0 < x < x_0 + \delta, \end{cases} \qquad \Longrightarrow \qquad x_0 \quad \text{punto di massimo locale.}$$

Analogamente, per il minimo, vale

$$f'(x) \begin{cases} \leq 0 & x_0 - \delta < x < x_0, \\ \\ \geq 0 & x_0 < x < x_0 + \delta, \end{cases} \qquad \Longrightarrow \qquad x_0 \quad \text{punto di minimo locale.}$$

ESERCIZIO 1.5. *Determinare i punti critici della funzione $f(x) = x^2(3x^2 - 8x + 6)$ e dire quali di essi sono punti di massimo o di minimo.*

Soluzione. La derivata prima della funzione è

$$f'(x) = 2x(3x^2 - 8x + 6) + x^2(6x - 8) = 12x(x^2 - 2x + 1) = 12x(x - 1)^2.$$

I punti critici sono $x = 0$ e $x = 1$; il punto $x = 0$ è punto di minimo, mentre il punto $x = 1$ non è né di massimo né di minimo. Il grafico qualitativo della funzione f è in Figura 3.

---

[1] Il termine *stazionario* è ereditato dalla cinematica. Se x è la posizione di un punto su cui agisce una forza conservativa con potenziale dato dalla funzione $f = f(x)$, allora l'accelerazione del punto è proporzionale a f'. Se il punto viene collocato a riposo nella posizione x_0 e $f'(x_0) = 0$, allora, dato che l'accelerazione (cioè la variazione di velocità) è nulla, il punto *stazionerà* nella posizione x_0 per ogni tempo successivo

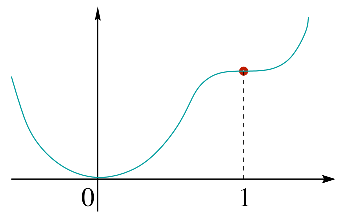

FIGURA 3. Il grafico di $f(x) = x^2(3x^2 - 8x + 6)$ dell'Esercizio 1.5.

Se x_0 è un punto critico di f, per riconoscere se f' cambia segno traversando x_0, basta considerare il segno di f'', qualora esista:

$$f'(x_0) = 0, \quad f''(x_0) > 0 \quad \Longrightarrow \quad x_0 \quad \text{punto di minimo locale;}$$

$$f'(x_0) = 0, \quad f''(x_0) < 0 \quad \Longrightarrow \quad x_0 \quad \text{punto di massimo locale.}$$

Si noti che si tratta solo di condizioni sufficienti: ad esempio, la funzione $f(x) = x^4$ ha un punto di minimo in 0, ma $f''(0) = 0$.

ESEMPIO 1.6. *Una raffinatezza per buongustai.* In genere, si immagina il grafico di una funzione vicino al punto di minimo x_0 con $f'(x) \leq 0$ per $x_0 - \delta < x < x_0$ e $f'(x) \geq 0$ per $x_0 < x < x_0 + \delta$. Esistono però anche situazioni in cui una funzione alla sinistra del punto di minimo non è decrescente e alla destra non è crescente. Scetticismo? Ecco un esempio:

$$f(x) = \begin{cases} x^2\left(2 - \sin\left(\dfrac{1}{x}\right)\right) & x \neq 0, \\ 0 & x = 0. \end{cases}$$

La funzione f è derivabile in tutto $\mathbb{R}$ e

$$f(x) \geq 0 \qquad \forall\, x \in \mathbb{R}, \qquad \text{e} \qquad f(x) = 0 \quad \Longleftrightarrow \quad x = 0.$$

Il punto $x = 0$ è punto di minimo globale, e quindi di minimo locale. Necessariamente $f'(0) = 0$ (come si può ottenere anche tramite il calcolo del limite del rapporto incrementale). La derivata prima di f nei punti $x \neq 0$ è

$$f'(x) = 2x\left(2 - \sin\left(\frac{1}{x}\right)\right) + \cos\left(\frac{1}{x}\right),$$

quindi, per $x \approx 0$, si ha $f'(x) \approx \cos\left(\frac{1}{x}\right)$, che assume valori sia positivi che negativi.

2. Analisi al microscopio

Nello studio dell'andamento qualitativo del grafico di una funzione, è interessante approfondire quello che succede in prossimità di certi punti significativi. Qui consideriamo come punti "significativi" quelli che corrispondono ad una delle seguenti situazioni:

(a) punti x_0 che non sono nell'insieme di definizione di f, ma che sono sul "bordo" (ad esempio, se $f : (a,b] \to \mathbb{R}$, il punto $x_0 = a$, oppure se $f : [a,b] \setminus \{c\} \to \mathbb{R}$, $x_0 = c$);

(b) punti x_0 dell'insieme di definizione in cui f non è continua;

(c) punti x_0 in cui f è continua, ma non derivabile.

Nei casi (b) e (c) si parla talvolta di *punti di singolarità*.

Asintoti verticali. Sia nel caso (a) che nel caso (b), si calcola il limite

$$\lim_{x \to x_0} f(x).$$

Se il limite è $+\infty$ o $-\infty$, la funzione ha in $x = x_0$ un asintoto verticale. Lo stesso è vero

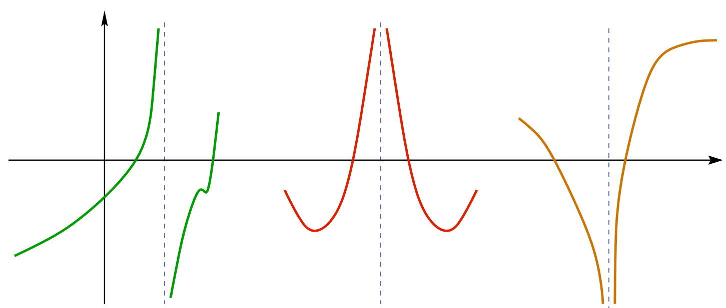

FIGURA 4. Alcuni esempi di asintoti verticali.

nel caso in cui sia il limite destro che il limite sinistro tendano a $+\infty$ o $-\infty$, ma con segni opposti. In generale, zeri del denominatore di una funzione razionale (che non siano anche zeri del numeratore), corrispondono a punti di asintoto verticale.

Ci sono situazioni più esotiche: il limite potrebbe non esistere oppure potrebbero esistere i limiti destro e sinistro, ma con valori diversi,... A voi individuare possibili esempi e corrispondenti grafici.

ESERCIZIO 2.1. *Studiare la funzione* $f(x) = \arctan(1/x)$ *vicino al punto* $x = 0$.

Punti angolosi e cuspidi. Consideriamo il caso (c), quindi supponiamo x_0 tale che la funzione f sia continua in x_0, ma non derivabile. Se esistono finiti i limiti destro e sinistro della derivata prima

$$\lim_{x \to x_0^\pm} f'(x) = \ell^\pm,$$

dato che f non è derivabile in x_0, deve essere $\ell^+ \neq \ell^-$. Un punto di questo genere si chiama punto angoloso (o spigolo). Per disegnarlo correttamente è possibile tracciare le rette tangenti destra e sinistra, cioè le rette di equazione $y = f(x_0) + \ell^\pm(x - x_0)$.

Nel caso in cui i limiti destro e sinistro siano $\pm\infty$ si possono avere due situazioni differenti. Se entrambi sono $+\infty$ (o $-\infty$), cioè se

$$\lim_{x\to x_0^\pm} f'(x) = +\infty \quad (-\infty),$$

il punto x_0 è un punto a tangente verticale. Se invece i limiti destro e sinistro sono $\pm\infty$, ma con segni opposti, il punto x_0 è una cuspide del grafico di f. Per un esempio di cuspide, si consideri la funzione $f(x) = \sqrt{|x|}$. In questo caso

$$\lim_{x\to 0^+} f'(x) = \lim_{x\to 0^+} \frac{1}{2\sqrt{x}} = +\infty, \qquad \lim_{x\to 0^-} f'(x) = \lim_{x\to 0^-} -\frac{1}{2\sqrt{-x}} = -\infty.$$

Ovviamente sono possibili comportamenti analoghi a quelli descritti, ma misti: ad

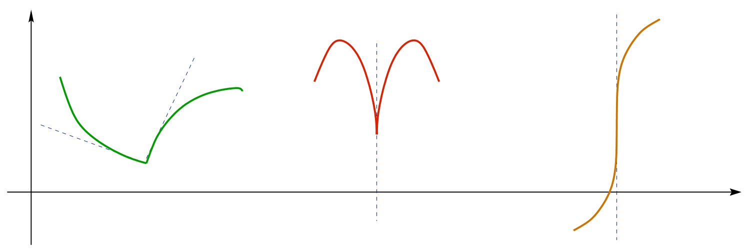

Figura 5. Da sinistra: un punto angoloso, una cuspide e un punto a tangente verticale.

esempio, una funzione può avere derivata prima che tende ad un valore dato da destra e che diverge da sinistra, o tutte le varianti che la mente è in grado di inventare.

3. Comportamento asintotico

Se la funzione f è definita in insiemi illimitati, è interessante studiarne il comportamento per $x \to \pm\infty$. Per fissare le idee, consideriamo una funzione f definita su una semiretta $[a, +\infty)$ con $a \in \mathbb{R}$. In questo caso si vuole stabilire cosa succeda per $x \to +\infty$, cioè determinare il *comportamento asintotico per* $x \to +\infty$. Considerazioni analoghe valgono per il caso di semirette del tipo $(-\infty, a]$, per $\mathbb{R}$, e, in generale, per domini illimitati.

La prima operazione sensata è il calcolo del limite per $x \to +\infty$. Se

$$\exists \lim_{x\to+\infty} f(x) = \ell \in \mathbb{R},$$

si dice che la funzione f tende asintoticamente ad ℓ, oppure che f ha un asintoto orizzontale (di equazione $y = \ell$) per $x \to +\infty$. Il grafico della funzione f si avvicina alla retta di equazione $y = \ell$ per x sempre più grandi. Per un disegno più preciso, si può studiare il segno della funzione $f(x) - \ell$, che indica se il grafico della funzione f sia al di sopra o al di sotto dell'asintoto.

Esempio 3.1. Consideriamo la funzione
$$f(x) = \frac{2x^2}{x^2+1} \qquad x \in [1, +\infty).$$
In questo caso,
$$\lim_{x\to+\infty} \frac{2x^2}{x^2+1} = 2,$$
quindi la funzione ha l'asintoto orizzontale di equazione $y = 2$. Dato che
$$f(x) - \ell = \frac{2x^2}{x^2+1} - 2 = -\frac{2}{x^2+1} < 0,$$
quindi f tende a $y = 2$ dal basso. A voi il gusto di tracciare il grafico di questa funzione.

Invece, la funzione
$$f(x) = \frac{2x^2 \sin x}{x^2+1} \qquad x \in [1, +\infty),$$
non ha limite per $x \to +\infty$ e quindi non ha asintoto orizzontale.

Se il limite della funzione f esiste, ma è $+\infty$ o $-\infty$, evidentemente non c'è asintoto orizzontale. Che cosa si può dire in questo caso? È possibile che la funzione tenda ad un **asintoto obliquo**, ossia è possibile che esistano $a, b \in \mathbb{R}$ tali che

$$\lim_{x\to+\infty} \big[f(x) - (ax+b)\big] = 0. \tag{37}$$

Questa proprietà indica che il grafico della funzione f si avvicina al grafico della retta $y = ax+b$ per $x \to +\infty$. Il problema è: *come determinare (qualora esistano) le costanti a e b?* Supponiamo che valga (37), allora
$$\lim_{x\to+\infty} \frac{f(x)}{x} = \lim_{x\to+\infty} \frac{f(x) - ax}{x} + a = a.$$
Una volta noto a, è possibile determinare b (qualora esista) calcolando
$$\lim_{x\to+\infty} \big[f(x) - ax\big] = b.$$
Ecco, quindi, le istruzioni per determinare la presenza di un asintoto obliquo:

i. calcolare $\lim\limits_{x\to+\infty} f(x)$: se il limite esiste finito, c'è un asintoto orizzontale (fine dello studio a $+\infty$), se il limite non esiste, non c'è né asintoto obliquo, né asintoto orizzontale (fine dello studio a $+\infty$), se il limite è $+\infty$ o $-\infty$ si va al punto (ii);

ii. calcolare $\lim\limits_{x\to+\infty} f(x)/x$: se il limite esiste finito, il suo valore è a e si va al punto (iii), se il limite non esiste o se vale $\pm\infty$, non c'è asintoto obliquo (fine dello studio a $+\infty$);

iii. calcolare $\lim\limits_{x\to+\infty} \big[f(x) - ax\big]$: se il limite esiste finito, il suo valore è b, la funzione ha asintoto obliquo di equazione $y = ax + b$, se il limite non esiste o se vale $\pm\infty$, non c'è asintoto obliquo (fine dello sudio a $+\infty$).

ESEMPIO 3.2. Consideriamo la funzione
$$f(x) = \frac{x^2 - 1}{3x + 1} \qquad x \in [0, +\infty).$$
Si ha
$$\begin{aligned}
\lim_{x\to+\infty} f(x) &= \lim_{x\to+\infty} \frac{x^2-1}{3x+1} = +\infty,\\
\lim_{x\to+\infty} \frac{f(x)}{x} &= \lim_{x\to+\infty} \frac{x^2-1}{3x(x+1)} = \frac{1}{3} =: a,\\
\lim_{x\to+\infty} f(x) - \frac{x}{3} &= \lim_{x\to+\infty} \frac{x^2-1}{3x+1} - \frac{x}{3} = \lim_{x\to+\infty} \frac{-3-x}{3(3x+1)} = -\frac{1}{9} =: b.
\end{aligned}$$
Quindi la funzione ha un asintoto obliquo di equazione $y = \frac{1}{3}x - \frac{1}{9}$. Anche in questo caso, per disegnare un grafico più preciso, si può studiare il segno della funzione
$$f(x) - (ax + b) = \frac{x^2-1}{3x+1} - \left(\frac{1}{3}x - \frac{1}{9}\right) = -\frac{8}{9(3x+1)} < 0 \qquad \forall\, x > -\frac{1}{3}.$$
La differenza è negativa, quindi la funzione tende all'asintoto dal basso.

Dopo il punto (i), se esiste finito $\lim_{x\to+\infty} f'(x)$, allora a è uguale al valore di questo limite e si può proseguire direttamente dal punto (iii). Se invece il limite di f' non esiste, bisogna necessariamente seguire il procedimento esposto sopra. Ad esempio, per $f(x) = x + \dfrac{\sin(x^2)}{x}$, si ha
$$f'(x) = 1 + 2\cos(x^2) - \frac{\sin(x^2)}{x^2},$$
che non ammette limite per $x \to +\infty$, ma è facile vedere che la funzione ha un asintoto obliquo per $x \to +\infty$ di equazione $y = x$.

ESERCIZIO 3.3. *Sia* $f(x) = \frac{p(x)}{q(x)}$ *con* p *polinomio di grado* k *e* q *polinomio di grado* h*. Dimostrare che:*
(i) f *ha un asintoto orizzontale se e solo se* $k \le h$*;*
(ii) f *ha un asintoto obliquo, non orizzontale, se e solo se* $k = h + 1$*.*

Altri profili asintotici. Il caso dell'asintoto obliquo è solo una situazione molto particolare: può capitare che una funzione tenda asintoticamente ad una funzione, che non sia un polinomio di primo grado. Consideriamo, ad esempio,
$$f(x) = \frac{x^3 + 1}{x - 2}.$$

Grazie all'algoritmo di divisione di polinomi, possiamo riscrivere questa funzione come

$$f(x) = x^2 + 2x + 4 + \frac{9}{x-2}.$$

Da questa espressione è immediato vedere che

$$\lim_{x\to\pm\infty} \big[f(x) - (x^2 + 2x + 4)\big] = 0,$$

e quindi il grafico di f tende asintoticamente alla parabola $y = x^2 + 2x + 4$.

Riconsideriamo la funzione

$$f(x) = \frac{2x^2 \sin x}{x^2+1} \qquad x \in [1, +\infty).$$

Dato che $\frac{2x^2}{x^2+1} \to 2$ per $x \to +\infty$, è sensato immaginare che questa funzione "assomigli" alla funzione $f(x) = 2\sin x$ per $x \to +\infty$. Calcoliamo la differenza tra $f(x)$ e $2\sin x$ e vediamo se è infinitesima:

$$\left|\frac{2x^2 \sin x}{x^2+1} - 2\sin x\right| = \frac{2\,|\sin x|}{1+x^2} \le \frac{2}{1+x^2} \to 0 \qquad \text{per} \quad x \to +\infty.$$

Quindi

$$f(x) = \frac{2x^2 \sin x}{x^2+1} = 2\sin x + h(x) \qquad \text{con} \qquad \lim_{x\to+\infty} h(x) = 0.$$

In generale se siamo in grado di riscrivere la funzione f nella forma

$$f(x) = g(x) + h(x)$$

con g funzione di cui si conosce il grafico e $h \to 0$ per $x \to \infty$, il grafico della funzione f tende verso quello della funzione g. Non esiste alcuna strategia generale per determinare

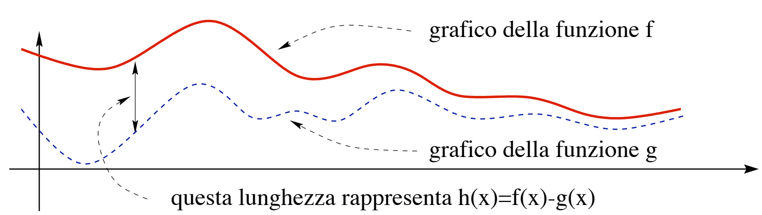

FIGURA 6. Il grafico di f con $f(x) = g(x) + h(x)$ e h infinitesima per $x \to +\infty$.

una decomposizione di questo genere.

4. Funzioni convesse

Come già detto, ripetuto ed usato ampiamente, il segno della derivata prima dà informazioni relative alla monotonia della funzione f. Quale ruolo gioca, invece, il segno della derivata seconda? Partiamo da una definizione.

DEFINIZIONE 4.1. *Una funzione* $f : [a,b] \to \mathbb{R}$ *è* convessa in $[a,b]$ *se*

$$f(tx + (1-t)y) \leq tf(x) + (1-t)f(y) \qquad \forall\, x, y \in [a,b] \quad \forall\, t \in (0,1). \tag{38}$$

Una funzione per cui valga la disuguaglianza opposta si dice concava.

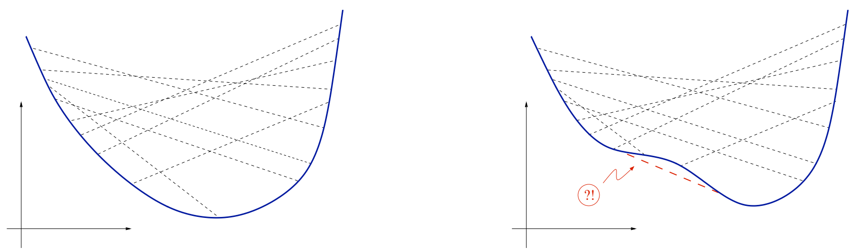

FIGURA 7. Una funzione convessa ed una non convessa.

Dalla definizione segue che se f è concava, allora $-f$ è convessa, e viceversa. Quindi studiare la convessità è sufficiente per comprendere anche la concavità.

OSSERVAZIONE 4.2. Esiste una maniera diversa di introdurre il concetto di convessità. Un sottoinsieme E del piano è convesso se scelta una qualsiasi coppia di punti P e Q appartenenti ad E, il segmento che li congiunge è interamente contenuto in E. Una funzione $f : [a,b] \to \mathbb{R}$ è convessa in $[a,b]$ se l'insieme $E := \{(x,y) : y \geq f(x)\}$, composto dai punti che si trovano sopra il suo grafico (detto epigrafico) è un insieme convesso. Le due definizioni sono comunque equivalenti (sapete dimostrarlo?).

Cerchiamo di capire il significato geometrico della condizione (38). Fissiamo $x = \bar{x}$ e $y = \bar{y}$ con $\bar{x} < \bar{y}$. Per $t \in (0,1)$, definiamo $z(t) := t\bar{x} + (1-t)\bar{y} \in (\bar{x}, \bar{y})$. Scriviamo la retta che passa per $(\bar{x}, f(\bar{x}))$ e $(\bar{y}, f(\bar{y}))$:

$$\Phi(x) = f(\bar{x}) + \frac{f(\bar{y}) - f(\bar{x})}{\bar{y} - \bar{x}}(x - \bar{x}),$$

e calcoliamo questa funzione in $z(t)$. Dato che

$$\begin{aligned}
\Phi(z(t)) &= f(\bar{x}) + \frac{f(\bar{y}) - f(\bar{x})}{\bar{y} - \bar{x}}(t\bar{x} + (1-t)\bar{y} - \bar{x}) \\
&= f(\bar{x}) + \frac{f(\bar{y}) - f(\bar{x})}{\bar{y} - \bar{x}}(1-t)(\bar{y} - \bar{x}) \\
&= f(\bar{x}) + (f(\bar{y}) - f(\bar{x}))(1-t) = tf(\bar{x}) + (1-t)f(\bar{y}),
\end{aligned}$$

la condizione (38), si può riscrivere come

$$f(z(t)) \leq \Phi(z(t)) \qquad \forall\, x, y \in [a,b] \quad \forall\, t \in (0,1).$$

Questa scrittura ha un interpretazione in termini di grafico immediata: *una funzione f è convessa, se per ogni scelta di x e y, il grafico di f giace al di sotto della retta secante che congiunge i punti $(x, f(x))$ e $(y, f(y))$ nell'intervallo di estremi x e y.*

Esercizio 4.3. *Se f e g sono due funzioni convesse in $[a,b]$, allora una tra* $\max\{f,g\}$ e $\min\{f,g\}$ *è convessa. Sapete dire quale?*

Proposizione 4.4. *Una funzione $f : [a,b] \to \mathbb{R}$ è convessa in $[a,b]$ se e solo se*

$$\frac{f(z)-f(x)}{z-x} \leq \frac{f(y)-f(x)}{y-x} \leq \frac{f(y)-f(z)}{y-z} \tag{39}$$

per ogni x, y, z tali che $a \leq x < z < y \leq b$.

La dimostrazione della formula (39) si ottiene riscrivendo in termini di rapporti incrementali la formula (38). I dettagli sono lasciati alla buona volontà del lettore.

La proprietà (39) può essere interpretata graficamente in termini di monotonía delle pendenza delle secanti: fissato y, la funzione

$$\phi(x) := \frac{f(x)-f(y)}{x-y}$$

è crescente in x. Quando la funzione è derivabile, questa proprietà diviene una richiesta di monotonia della derivata prima f'. Nel caso in cui la funzione f sia derivabile due volte, la monotonía della funzione f' può essere tradotta in termini di segno della derivata seconda f''.

Teorema 4.5. *Sia $f : [a,b] \to \mathbb{R}$. Allora*

(i) se f è derivabile una volta, la funzione f è convessa in $[a,b]$ se e solo se f' è non decrescente in $[a,b]$;

(ii) se f è derivabile due volte, la funzione f è convessa in $[a,b]$ se e solo se $f''(x) \geq 0$ per ogni $x \in [a,b]$.

Per le funzioni concave, vale un risultato analogo sostituendo a "f' crescente" la frase "f' non crescente" e a "$f'' \geq 0$" la frase "$f'' \leq 0$".

Dal Teorema 4.5(i) discende un'altra interpretazione geometrica della convessità: *se la funzione f è derivabile e convessa, il suo grafico è interamente al di sopra di ogni retta tangente ad esso.* Infatti, scriviamo la differenza tra f e la retta tangente in $(x_0, f(x_0))$

$$R(x; x_0) = f(x) - f(x_0) - f'(x_0)(x - x_0),$$

con l'obiettivo di dimostrare che se f è convessa, la funzione $R(x;x_0)$ è positiva. Applichiamo il Teorema di Lagrange e riscriviamo $R(x;x_0)$ come

$$R(x;x_0) = f'(\xi)(x-x_0) - f'(x_0)(x-x_0) = \big(f'(\xi) - f'(x_0)\big)(x-x_0).$$

Se $x > x_0$ allora $\xi > x_0$ e quindi, essendo f' crescente, $f'(\xi) > f'(x_0)$. Ne segue che il termine a destra è positivo perché prodotto di termini positivi. Se $x < x_0$ allora $\xi < x_0$ e, sempre per la monotonía di f', $f'(\xi) < f'(x_0)$. Questa volta i due termini sono entrambi negativi, ma comunque il loro prodotto è positivo.

Dimostrazione del Teorema 4.5. (i) Supponiamo che f sia convessa, allora vale la (39). Quindi, passando al limite per $z \to x^+$ si ottiene

$$f'(x) \le \frac{f(y)-f(x)}{y-x}.$$

Analogamente, passando al limite nella (39) per $z \to y^-$,

$$\frac{f(y)-f(x)}{y-x} \le f'(y).$$

Ne segue che $f'(x) \le f'(y)$ per ogni $x \le y$.

Viceversa, supponiamo che la funzione f' sia non decrescente e dimostriamo la (38) studiando la funzione differenza

$$F(t) := tf(x) + (1-t)f(y) - f(tx+(1-t)y), \qquad t \in [0,1],$$

con x, y fissati. Consideriamo il caso $y < x$ (l'altro è analogo). Calcolando la derivata di F e applicando il Teorema di Lagrange, si deduce che esiste $\xi \in (y,x)$ tale che

$$F'(t) = f(x) - f(y) - f'(tx+(1-t)y)(x-y) = \Big[f'(\xi) - f'(tx+(1-t)y)\Big](x-y).$$

Dato che, per $t \in [0,1]$, il punto $tx+(1-t)y$ descrive l'intervallo $[y,x]$, esiste t^* tale che $t^*x + (1-t^*)y = \xi$. Inoltre, dato che f' è non decrescente, $f'(tx+(1-t)y) \le f'(\xi)$ per $t \in [0,t^*]$ e $f'(tx+(1-t)y) \ge f'(\xi)$ per $t \in [t^*,1]$. Perciò:

$$F'(t) \ge 0 \quad \text{per} \quad t \in [0,t^*] \qquad \text{e} \qquad F'(t) \le 0 \quad \text{per} \quad t \in [t^*,1].$$

Perciò il minimo globale della funzione F è assunto in uno degli estremi $t=0$ o $t=1$ e, dato che $F(0)=F(1)=0$, $F(t) \ge 0$ per ogni $t \in [0,1]$, cioè la formula (38). □

Esercizio 4.6. *Sia f una funzione convessa e derivabile in (a,b). Se esistono $x_0, x_1 \in (a,b)$ con $x_0 \neq x_1$ tali che $f'(x_0) = f'(x_1) = 0$, che cosa si può dedurre sulla funzione f?*

Definizione 4.7. *Sia $f : (a,b) \to \mathbb{R}$ derivabile due volte. Se x_0 è tale che f'' cambia segno in x_0 (cioè è negativa da una parte e positiva dall'altra), allora x_0 si chiama* punto di flesso *della funzione f.*

Grazie al Teorema 4.5, se f'' ha segno opposto alla destra e alla sinistra di x_0, necessariamente x_0 è un punto di flesso.

ESEMPIO 4.8. Consideriamo la funzione $f(x) = \sin x$. La sua derivata seconda è $f''(x) = -\sin x$, quindi tutti punti della forma $x = k\pi$ per $k \in \mathbb{Z}$ sono punti di flesso.

Per la funzione $f(x) = \frac{1}{1+x^2}$, si ha

$$f'(x) = -\frac{2x}{(1+x^2)^2} \qquad \Longrightarrow \qquad f''(x) = \frac{2(3x^2-1)}{(1+x^2)^3},$$

e quindi i punti di flesso di f sono $x = \pm 1/\sqrt{3}$. La funzione f è convessa in $(-\infty, -1/\sqrt{3})$ e in $(1/\sqrt{3}, +\infty)$ e concava in $(-1/\sqrt{3}, 1/\sqrt{3})$.

La convessità è utile per determinare l'esistenza di minimi di una funzione. Infatti, vale la seguente implicazione

$$f \quad \text{convessa}, \quad f'(x_0) = 0 \qquad \Longrightarrow \qquad x_0 \quad \text{punto di minimo globale.}$$

La dimostrazione è lasciata per esercizio. Analogamente, per le funzioni concave ed i punti di massimo. Chiaramente se la convessità è solo locale (cioè in un intorno del punto x_0), x_0 è punto di minimo locale.

ESERCIZIO 4.9. *Dimostrare la seguente implicazione*

$$f\,:\,[a,b] \to \mathbb{R} \quad \textit{convessa} \quad \Rightarrow \quad f \quad \textit{continua in} \quad (a,b).$$

5. A caccia di massimi e minimi assoluti

Problema 1. Abbiamo già considerato il problema di determinare il cilindro di volume $V = k =$ costante con superficie totale S minima, con l'obiettivo (malcelato) di diventare ricchi grazie all'uso della matematica, applicando il risultato alla costruzione di scatole di fagioli, o, più in generale, di confezioni cilindriche con minima spesa di materiali. La speranza si era presto infranta quando ci siamo resi conto che per via elementare non riuscivamo a determinare il minimo della funzione S, cioè a risolvere il problema (r =raggio della base del cilindro)

$$\text{determinare il minimo di } S(r) = 2\pi\left(r^2 + \frac{k}{\pi r}\right) \qquad r > 0.$$

Torniamo al problema con la conoscenza delle derivate e studiamo la monotonìa di S:

$$\frac{dS}{dr} = 2\pi\left(2r - \frac{k}{\pi r^2}\right) = \frac{4\pi}{r^2}\left(r^3 - \frac{k}{2\pi}\right).$$

Perciò $S'(r) \geq 0$ se e solo se $r \geq r^*$ dove $r^* = (k/2\pi)^{1/3}$. Quindi la funzione S è decrescente in $(0, r^*)$ ed è crescente in (r^*, ∞). Ne segue che il punto di minimo richiesto esiste ed è proprio $r = r^*$ (Fig.8(a)). Problema risolto, corriamo in fabbrica!

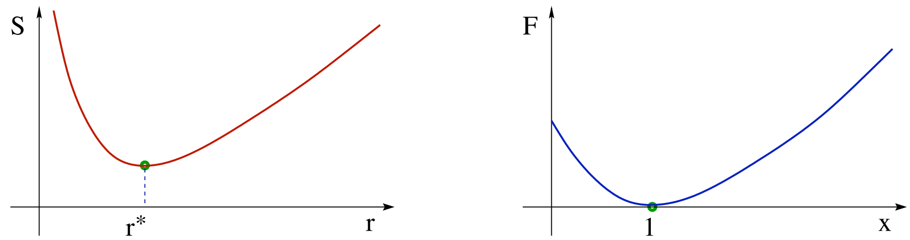

FIGURA 8. (a) Il grafico della funzione $S(r) = 2\pi\left(r^2 + \frac{k}{\pi r}\right)$; (b) il grafico della funzione $F(x) = x^p - 1 - p(x-1)$, $p > 1$.

Problema 2. Vogliamo dimostrare la disequazione

$$x^p - 1 \geq p(x-1) \qquad \forall p > 1, \quad \forall x \geq 0.$$

Fissiamo $p > 1$ e consideriamo la funzione

$$F(x) = x^p - 1 - p(x-1) \qquad x \geq 0.$$

Dato che $F'(x) = p(x^{p-1} - 1)$, $F'(x) < 0$ per $x \in (0,1)$ e $F'(x) > 0$ per $x \in (1, +\infty)$. Quindi la funzione F è decrescente in $[0,1)$ e crescente in $(1,+\infty)$ e $x = 1$ è un punto di minimo. Ne segue che $F(x) \geq F(1) = 0$, da cui la conclusione (Fig.8(b)).

Problema 3. Siano $a_1, a_2, \ldots, a_n \in \mathbb{R}$ assegnati. Supponiamo di voler determinare $x \in \mathbb{R}$ tale che sia minima la quantità

$$\sum_{i=1}^{n} (a_i - x)^2. \tag{40}$$

Possiamo immaginare che i valori a_i provengano da misurazioni di un fenomeno sotto osservazione e che si stia cercando un valore medio per questi numeri, che minimizzi l'errore commesso misurato dal valore in (40). Consideriamo la funzione $F(x) = \sum\limits_{i=1}^{n}(a_i - x)^2$ e calcoliamone la derivata:

$$F'(x) = -2\sum_{i=1}^{n}(a_i - x) = -2\left[\sum_{i=1}^{n} a_i - \sum_{i=1}^{n} x\right] = 2n\left[x - \frac{1}{n}\sum_{i=1}^{n} a_i\right].$$

La funzione F è decrescente a sinistra di $\frac{1}{n}\sum\limits_{i=1}^{n} a_i$ e crescente a destra. Il valore

$$x = \frac{1}{n}\sum_{i=1}^{n} a_i$$

è il punto di minimo (e coincide con la media aritmetica di $a_1, \ldots, a_n$).

Esercizio 5.1. *Dati $a_1, a_2, \ldots, a_n \in \mathbb{R}$, determinare x che minimizzi $\sum_{i=1}^{n} \lambda_i (a_i - x)^2$ dove $\lambda_1, \ldots, \lambda_n > 0$ sono pesi (positivi) assegnati.*

I problemi che abbiamo appena presentato mostrano alcune tra le miriadi di situazioni in cui si pone il problema: data una funzione f, come determinarne massimo e minimo globali (qualora esistano)? Proviamo ad affrontare il problema in generale.

Supponiamo di lavorare con una funzione f definita nell'intervallo $[a, b]$ e continua. Grazie al teorema di Weierstrass, l'ipotesi di continuità garantisce l'esistenza del massimo e del minimo assoluti. Abbiamo già visto che le soluzioni di $f'(x) = 0$ (cioè i punti critici di f) permettono di determinare i possibili candidati a punti di minimo o massimo locale *interno derivabile*. Chiaramente, è possibile che un estremo locale cada in un punto in cui la funzione non è derivabile. Quindi, la strategia per individuare il massimo ed il minimo di una funzione continua in $[a, b]$ è la seguente:

- $\star$ determinare l'insieme $\mathcal{S}$ dei punti stazionari in (a, b);
- $\star$ determinare l'eventuale insieme $\mathcal{N}$ dei punti in cui f non è derivabile;
- $\star$ calcolare la funzione in $\mathcal{S}$, in $\mathcal{N}$ e negli estremi dell'intervallo a e b.
- $\star$ individuare il più grande e il più piccolo tra i valori calcolati.

Esercizio 5.2. *Determinare il massimo ed il minimo assoluti di*

$$f(x) = (x^2 - 5x + 7)e^x \qquad x \in [0, 2].$$

Soluzione. La funzione f è derivabile dappertutto. Per determinare i punti singolari:

$$f'(x) = (2x - 5)e^x + (x^2 - 5x + 7)e^x = (x^2 - 3x + 2)e^x = (x - 2)(x - 1)e^x.$$

Quindi $f'(x) = 0$ se e solo se $x = 1$ o $x = 2$. L'insieme dei punti critici *interni* è $\mathcal{S} = \{1\}$. Dato che $f(0) = 7 < f(2) = e^2 < f(1) = 3e$, si ha

$$\min_{x \in [0,2]} f(x) = f(0) = 7, \qquad \max_{x \in [0,2]} f(x) = f(1) = 3e,$$

che è quanto richiesto dall'esercizio.

Spesso è utile conoscere il massimo del modulo di una funzione assegnata f, cioè risolvere il problema

$$\text{data } f : [a, b] \to \mathbb{R} \quad \text{continua}, \quad \text{calcolare } \max_{x \in [a,b]} |f(x)|.$$

In questo caso, si può procedere come detto sopra, o, alternativamente, determinare il massimo ed il minimo della funzione f in $[a, b]$ e poi sfruttare la relazione (evidente?)

$$\max_{x \in [a,b]} |f(x)| = \max\Big\{ \big|\max_{x \in [a,b]} f(x)\big|, \big|\min_{x \in [a,b]} f(x)\big| \Big\}.$$

Esercizio 5.3. *Calcolare* $\max\{|x^2 - 1| : x \in [-1, 2]\}$.

Nel caso in cui si studi una funzione f continua, ma definita su un dominio illimitato (ad esempio, $f : [a, +\infty) \to \mathbb{R}$), le ipotesi del Teorema di Weierstrass non sono soddisfatte e quindi non è detto che esistano il massimo ed il minimo della funzione. Comunque ha senso domandarsi: *quanto valgono l'estremo superiore e l'estremo inferiore? Nel caso in cui siano finiti, si tratta di massimo o di minimo?* La strategia per risolvere questo problema è simile a quanto appena visto. Il punto che bisogna modificare è quello relativo al calcolo della funzione negli estremi dell'intervallo. In questo caso almeno uno degli estremi dell'intervallo sarà $+\infty$ o $-\infty$ e le espressioni $f(+\infty)$ e $f(-\infty)$, in generale, non hanno senso, ma vanno sostituite con $\lim_{x\to\pm\infty} f(x)$. Vediamo la procedura negli esercizi che seguono.

Esercizio 5.4. *Determinare l'estremo superiore e l'estremo inferiore della funzione $f(x) = e^{x^2}$ in $\mathbb{R}$ e dire se si tratta di massimo e minimo.*

Soluzione. L'estremo superiore è presto detto: dato che $\lim_{x\to+\infty} f(x) = +\infty$, è chiaro che $\sup_{x\in\mathbb{R}} f = +\infty$. Che possiamo dire sull'estremo inferiore? Visto che la funzione f è continua su $\mathbb{R}$, esistono massimo e minimo di f in $[-M, M]$ per ogni scelta di $M > 0$. Quindi

$$\inf_{x\in\mathbb{R}} f(x) = \min\left\{ \min_{|x|\le M} f(x), \inf_{|x|>M} f(x) \right\}.$$

Inoltre, dato che $f(x) \to +\infty$ per $x \to \pm\infty$, per M grande, $\min_{|x|\le M} f(x) \le \inf_{|x|>M} f(x)$. Non resta che cercare i punti di minimo relativo in $[-M, M]$. Derivando, $f'(x) = 2xe^{x^2}$ che si azzera se e solo se $x = 0$, quindi l'unico punto di minimo relativo è $x = 0$ che, per quanto detto, è anche minimo assoluto: $\min_{x\in\mathbb{R}} e^{x^2} = e^0 = 1$.

Esercizio 5.5. *Sia $f : \mathbb{R} \to \mathbb{R}$ una funzione continua tale che*

$$\lim_{x\to+\infty} f(x) = \lim_{x\to-\infty} f(x) = +\infty.$$

Dimostrare che la funzione f ammette minimo assoluto su $\mathbb{R}$.

Esercizio 5.6. *Determinare l'estremo superiore e l'estremo inferiore della funzione $f(x) = e^{-x^2}$ in $\mathbb{R}$ e dire se si tratta di massimo e minimo.*

Soluzione. La funzione è derivabile su tutto $\mathbb{R}$ e la derivata vale $f'(x) = -2xe^{-x^2}$. Quindi c'è un unico punto critico $x = 0$ in cui la funzione vale $f(0) = 1$. Inoltre

$$\lim_{x\to-\infty} e^{-x^2} = \lim_{x\to+\infty} e^{-x^2} = 0.$$

Confrontando i valori deduciamo che

$$\inf_{x\in\mathbb{R}} e^{-x^2} = 0 \qquad \sup_{x\in\mathbb{R}} e^{-x^2} = f(0) = 1.$$

Dato che l'estremo superiore fa parte dell'insieme immagine, l'estremo superiore è massimo. Invece l'estremo inferiore non è minimo, perchè la funzione f è strettamente positiva.

Analogamente nel caso di funzioni continue definite in insiemi non chiusi, cioè $f : (a, b) \to \mathbb{R}$ oppure $f : [a, b) \to \mathbb{R}$, o varianti, non si applica il Teorema di Weierstrass. Anche in questi casi, per determinare l'estremo superiore/inferiore bisogna considerare i limiti agli estremi.

Concludiamo la Sezione, analizzando altri due problemi di massimo e minimo.

Un problema di statica: la puleggia di De L'Hôpital. Consideriamo gli assi cartesiani (x, y) posti in modo che l'asse y sia in verticale rispetto al suolo e che la forza di gravità sia direzionata nel verso delle y negative. Indichiamo con $O = (0, 0)$ e con $A = (a, 0)$ dove $a > 0$ è una lunghezza fissata. Nel punto O, fissiamo una corda di lunghezza b e all'estremità B di questa corda fissiamo una puleggia. Lasciamo pendere per fatti suoi la puleggia e fissiamo una seconda corda di lunghezza ℓ al punto A. Dopo aver fatto passare la seconda corda attraverso la puleggia (quindi facendola passare per il punto B) fissiamo un peso M all'altra estremità. La configurazione finale è disegnata in Figura 9. Il problema è: in quale punto (x, y) si posizionerà il peso M?

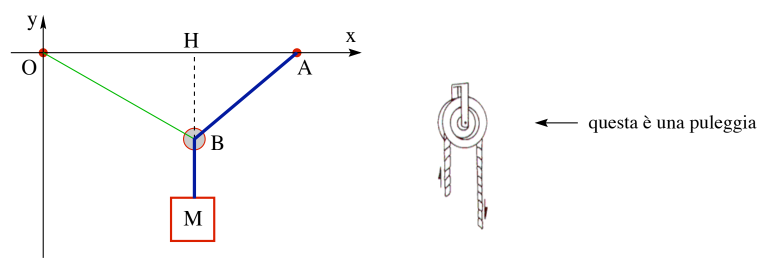

FIGURA 9. La puleggia di De L'Hôpital.

Qui stiamo supponendo che l'unica forza esterna che agisce sul sistema è la forza di gravità. Se, inoltre, supponiamo che il peso delle corde e della puleggia sia trascurabile rispetto al peso di M, e che non sia presente nessun tipo di attrito, il peso M si collocherà nella posizione più bassa possibile, cioè minimizzerà il valore y. Quest'affermazione discende dal principio seguente: *la posizione d'equilibrio (stabile) del sistema minimizza l'energia potenziale di M*. Dato che l'energia potenziale di M è della forma $Ay + B$ con $A > 0$, minimizzare l'energia potenziale equivale a minimizzare l'altezza y.

Riscriviamo per chiarezza i dati del problema:

$$\overline{OA} = a, \qquad \overline{OB} = b, \qquad \overline{AB} + \overline{BM} = \ell.$$

L'incognita è la posizione di $M = (x, y)$. Supponiamo inoltre $0 < b < a < \ell$. La soluzione del problema può essere divisa in due passi:

passo 1. assegnata la coordinata x di M, determinare la coordinata $y = f(x)$;

passo 2. calcolare (se esiste) il minimo della funzione f.

Da brave persone ordinate, partiamo dal primo passo. Indichiamo con H la proiezione ortogonale di M sull'asse x, cioè $H = (x, 0)$. Allora

$$y = -\left(\overline{HB} + \overline{BM}\right) = -\left(\overline{HB} + \ell - \overline{AB}\right) = \overline{AB} - \overline{HB} - \ell.$$

Si tratta ora di scrivere le lunghezze $\overline{AB}$ e $\overline{HB}$ in funzione di x. Basta usare il Teorema di Pitagora per ottenere:

$$\overline{HB} = \sqrt{\overline{OB}^2 - \overline{OH}^2} = \sqrt{b^2 - x^2}, \qquad \overline{AB} = \sqrt{\overline{HB}^2 + \overline{HA}^2} = \sqrt{b^2 - x^2 + (a-x)^2}$$

Quindi la funzione da studiare è

$$y = f(x) = \sqrt{b^2 - x^2 + (a-x)^2} - \sqrt{b^2 - x^2} - \ell \qquad x \in [0, b].$$

L'intervallo di variazione di x si deduce direttamente dal problema considerato.

Secondo passo: qual è la scelta di x che minimizza y? Dato che la funzione è continua in $[0, b]$ e l'intervallo è chiuso e limitato, il Teorema di Weierstrass ci assicura che il problema ha soluzione, ma non ci dà nessuna informazione su quale sia il punto di minimo. Implementiamo, quindi, la strategia proposta poche pagine fa.

Prima di tutto, notiamo che la funzione f non è derivabile nei punti in cui si azzera l'argomento di una delle due radici. Dato che $b < a$, il termine $b^2 - x^2 + (a-x)^2$ è sempre non nullo. La seconda radice $\sqrt{b^2 - x^2}$ si azzera per $x = \pm b$. Di questi due punti, $-b$ va scartato perché è fuori da $[0, b]$ e b è uno degli estremi dell'intervallo.

Determiniamo l'insieme $\mathcal{S}$ dei punti stazionari di f in $(0, b)$. Dato che

$$f'(x) = -\frac{a}{\sqrt{b^2 - x^2 + (a-x)^2}} + \frac{x}{\sqrt{b^2 - x^2}}$$

i punti stazionari $x \in (0, b)$ verificano

$$\frac{a}{\sqrt{b^2 - x^2 + (a-x)^2}} = \frac{x}{\sqrt{b^2 - x^2}},$$

e, passando ai quadrati,

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

Ora ci vogliono un po' di conti: l'equazione precedente è equivalente a

$$2ax^3 - 2a^2x^2 - b^2x^2 + a^2b^2 = 0 \quad \Longleftrightarrow \quad 2ax^2(x-a) - b^2(x+a)(x-a) = 0.$$

Dividendo per $x - a \neq 0$,

$$2ax^2 - b^2x - ab^2 = 0 \quad \Longleftrightarrow \quad x = x_\pm := \frac{b^2 \pm \sqrt{b^4 + 8a^2b^2}}{4a}$$

Dato che $x_- < 0 < x_+ < b$ (verificare!), c'è un unico punto critico in $(0, b)$: $\mathcal{S} = \{x_+\}$.

Ricapitolando, la nostra strategia propone tre punti di minimo assoluto possibili: $0, x_+, b$. Per determinare chi di questi sia il punto di minimo bisognerebbe confrontare i tre valori $f(0)$, $f(x_+)$ e $f(b)$. Fattibile, ma non particolarmente semplice. Seguiamo una strada diversa. Dato che

$$f'(0) = -\frac{a}{a-b} < 0$$

il punto 0 non può essere di minimo relativo e dunque nemmeno di minimo assoluto! La lotta rimane tra x_+ e b. In b non possiamo ragionare come in 0 dato che la funzione f non è derivabile in b, quindi $f'(b)$ non ha senso. Non perdiamoci d'animo: dato che

$$\lim_{x \to b^-} f'(x) = \lim_{x \to b^-} -\frac{a}{\sqrt{b^2 - x^2 + (a-x)^2}} + \frac{x}{\sqrt{b^2 - x^2}} = +\infty$$

la funzione f è crescente in un intorno (sinistro) di b, quindi nemmeno b può essere il punto di minimo richiesto. Perfetto: resta un unico sopravvisuto x_+ che è il punto di

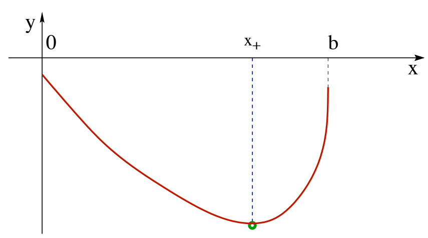

FIGURA 10. Il grafico della funzione f.

minimo cercato. Il peso M si collocherà nella posizione di coordinate $(x_+, f(x_+))$. In alternativa, per stabilire che x_+ è il punto di minimo assoluto, si sarebbe potuto anche notare che $f'' \geq 0$, quindi la funzione f è convessa e, necessariamente, il suo punto critico x_+ è di minimo assoluto.

Il principio di Fermat. Passiamo ora a considerare due problemi di *ottica geometrica* che si traducono nella ricerca del punto di minimo assoluto di certe funzioni. In quel che segue, considereremo un raggio di luce in maniera naïf: come un qualcosa che viaggia da un punto ad un altro, si riflette sugli oggetti, entra nell'occhio...

Il problema fondamentale è stabilire quale sia il tragitto percorso dal raggio per passare da un punto A ad un punto B. Il principio, proposto da Fermat, è il segunte: *il tragitto prescelto è quello che* <u>*minimizza*</u> *il tempo di percorrenza.* Se il raggio viaggia sempre nello stesso mezzo, la sua velocità v è costante, quindi minimizzare il tempo di percorrenza T equivale a minimizzare la lunghezza del percorso. Perciò il raggio percorre linee rette. Che succede in situazioni un po' più complicate?

Riflessione. Consideriamo un raggio di luce che parta dal punto $A = (0, a)$ e e che si diriga verso l'asse delle x dove immaginiamo collocato uno specchio. Il raggio viene riflesso nel punto $P = (x, 0)$ e da lì arriva nel punto $B = (b, c)$. I tragitti da A a P e da P a B sono percorsi lungo segmenti. Supponendo assegnati $a, b, c > 0$ e quindi assegnati i punti A e B, qual'è il punto P prescelto dal raggio luminoso?

Ad ogni scelta di $P = (x, 0)$, corrisponde un certo tempo di percorrenza $T = T(x)$. Il principio di Fermat afferma che il punto di riflessione $(x_0, 0)$ è tale che $T(x_0) = \min T(x)$. Bene, non resta che determinare l'espressione esplicita di $T = T(x)$ e trovare in quale punto sia assunto il minimo. Indicando con T_{AP} e T_{PB}, il tempo impiegato dal raggio per andare da A a P e da P a B rispettivamente, e con v la velocità della luce nel mezzo in considerazione,

$$T(x) = T_{AP}(x) + T_{PB}(x) = \frac{\overline{AP}}{v} + \frac{\overline{PB}}{v}.$$

Quindi, grazie al teorema di Pitagora,

$$T(x) = \frac{1}{v}\left(\sqrt{x^2 + a^2} + \sqrt{(b - x)^2 + c^2}\right) \qquad x \in [0, b].$$

Dato che la funzione T è derivabile infinite volte, applicando la strategia per il calcolo di massimi e minimi assoluti, il punto di minimo x_0 sarà o 0, o b, o tale che $T'(x_0) = 0$. Cerchiamo quindi i punti critici di T. La derivata prima T' è esplicitamente data da

$$T'(x) = \frac{1}{v}\left(\frac{x}{\sqrt{x^2 + a^2}} - \frac{b - x}{\sqrt{(b - x)^2 + c^2}}\right).$$

Essa si azzera se e solo se $x = x_* := ab/(a + c)$.

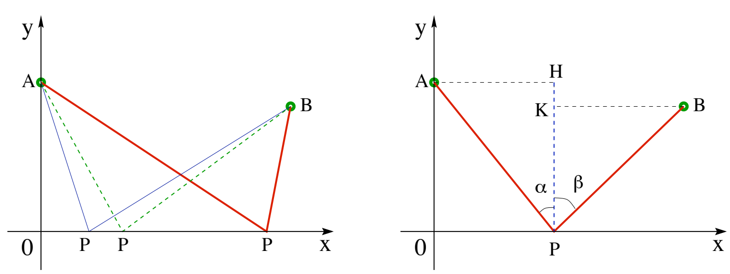

FIGURA 11. Quale sarà il punto P prescelto da un raggio riflesso in uno specchio? Quello che fa in modo che gli angoli di incidenza α e di riflessione β coincidano.

Invece di confrontare i valori di T per $x = 0, x_*, b$, si può notare che

$$T'(0) = -\frac{b}{v\sqrt{b^2 + c^2}} < 0 \qquad \text{e} \qquad T'(b) = \frac{b}{v\sqrt{b^2 + a^2}} > 0,$$

quindi nessuno dei due estremi dell'intervallo è di minimo. Pertanto il punto x_* è il punto di minimo assoluto.

Il punto di riflessione P è individuato da una condizione geometrica semplice. Sia α l'angolo, detto di *incidenza*, determinato dal segmento AP e dalla semiretta da P, perpendicolare all'asse x, contenuta nel semipiano $y > 0$, e sia β l'angolo, detto di *riflessione*, determinato dal segmento PB e dalla stessa semiretta di prima. Allora, indicando con H il punto di coordinate (x_*, a) e con K il punto di coordinate (x_*, c),

$$\tan\alpha = \frac{\overline{AH}}{\overline{PH}} = \frac{x_*}{a} = \frac{b}{a+c} \qquad \tan\beta = \frac{\overline{BK}}{\overline{PK}} = \frac{b - x_*}{c} = \frac{ab+bc-ab}{a+c}\frac{1}{c} = \frac{b}{a+c}$$

Quindi $\tan\alpha = \tan\beta$ e, dato che $\alpha, \beta \in [0, \pi/2]$, ne segue che $\alpha = \beta$. In definitiva, il principio di Fermat implica che *l'angolo di riflessione coincide con l'angolo di incidenza.*

Rifrazione. Cambiamo tipo di esperimento. Consideriamo un raggio luminoso che viaggi in due mezzi differenti in cui la sua velocità è v_+ e v_-. Per semplicità, supponiamo che il mezzo in cui la velocità è v_+ corrisponda alla regione di piano con $y > 0$ e quello in cui la velocità è v_- corrisponda a $y < 0$. Se un raggio parte da $A = (0, a)$ con $a > 0$ ed arriva a $B = (b, c)$ con $c < 0 < b$, che tragitto sceglie?

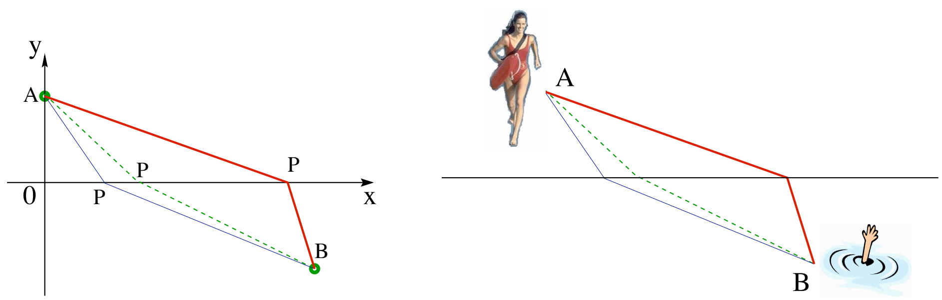

FIGURA 12. Quale percorso scegliere? Supponiamo che nel semipiano $y > 0$ ci sia la terra ferma e il mare nel semipiano $y < 0$. Nel punto A c'è la prestante bagnina di Baywatch pronta ad intervenire per salvare la vita di un affogando sito nel punto B. Sapendo che la soccorritrice quando corre sulla spiaggia va a velocità v_+ e quando nuota in mare va a velocità v_-, qual'è il percorso che le permette di soccorrere il malcapito nel minor tempo possibile?

Esattamente come prima, utilizziamo il principio di Fermat. Indicando con T_{AP} e T_{PB}, il tempo impiegato dal raggio per andare da A a P e da P a B rispettivamente, il tempo impiegato per andare da A a B è

$$T(x) = T_{AP}(x) + T_{PB}(x) = \frac{\overline{AP}}{v_+} + \frac{\overline{PB}}{v_-}.$$

e, di nuovo per il teorema di Pitagora,

$$T(x) = \frac{\sqrt{x^2 + a^2}}{v_+} + \frac{\sqrt{(b-x)^2 + c^2}}{v_-} \qquad x \in [0, b].$$

Anche questa funzione è derivabile infinite volte in $[0, b]$. La sua derivata prima è

$$T'(x) = \frac{x}{v_+\sqrt{x^2+a^2}} - \frac{b-x}{v_-\sqrt{(b-x)^2+c^2}}.$$

Quindi $T'(0) < 0 < T'(b)$ e anche in questo caso gli estremi non sono punti di minimo. Perciò il minimo è in $(0, b)$. Inoltre, con un po' di pazienza, si ottiene

$$T''(x) = \frac{a^2}{v_+\left(x^2+a^2\right)^{3/2}} + \frac{c^2}{v_-\left[(b-x)^2+c^2\right]^{3/2}} > 0$$

Dato che la derivata seconda è (strettamente) positiva, la derivata prima T' è strettamente crescente, quindi esiste un unico x_* tale che $T'(x_*) = 0$ (si ricordi che $T'(0) < 0 < T'(b)$). Necessariamente x_* è il punto di minimo che andiamo cercando.

Come individuarlo? Per ora sappiamo solo che x_* è individuato in maniera univoca dalla relazione $T'(x_*) = 0$, cioè x_* è l'unico valore per cui

$$\frac{1}{v_+}\frac{x_*}{\sqrt{x_*^2+a^2}} = \frac{1}{v_-}\frac{b-x_*}{\sqrt{(b-x_*)^2+c^2}}. \tag{41}$$

Come nel caso della riflessione, ragioniamo in termini di angoli. Sia α, *angolo di*

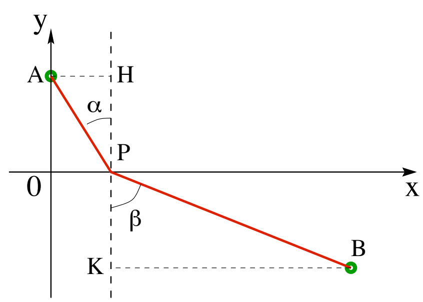

FIGURA 13. Gli angoli di incidenza e di rifrazione.

incidenza, l'angolo formato dal segmento AP con la semiretta verticale per P contenuta in $y > 0$, e sia β, *angolo di rifrazione*, l'angolo formato dal segmento PB con la semiretta verticale per P contenuta in $y < 0$. Se H e K sono le proiezioni di A e B sulla retta $x = x_*$, i triangoli APH e BPK sono rettangoli e quindi

$$\sin\alpha = \frac{\overline{AH}}{\overline{AP}} = \frac{x_*}{\sqrt{x_*^2+a^2}} \qquad \sin\beta = \frac{\overline{BK}}{\overline{BP}} = \frac{b-x_*}{\sqrt{(b-x_*)^2+c^2}}.$$

Sostituendo nella formula (41), si deduce che il punto di rifrazione P è scelto in modo che valga

$$\frac{\sin\alpha}{v_+} = \frac{\sin\beta}{v_-}.$$

Tale relazione è nota come legge di rifrazione di Snell. .

CAPITOLO 7

Ordini di grandezza e la formula di Taylor

1. Verso lo zero e ad un passo dall'infinito

Le funzioni con cui ci si trova a dover lavorare possono avere una struttura complicata, e anche molto. Se si è interessati al comportamento della funzione solo per determinati regimi, cioè per valori dell'incognita in opportune regioni, può bastare conoscere quali siano i termini dominanti all'interno della funzione. Ad esempio, se il valore $f(t)$ rappresenta la posizione di una particella all'istante t, si potrebbe essere interessati solo al comportamento della funzione f per valori grandi di t. Se $f(t) = e^t + \sin t$, è chiaro che saremo soddisfatti di un'approssimazione del tipo $f(t) \approx e^t$ per $t \to +\infty$, dato che questo termine diverge a $+\infty$, mentre l'altro rimane limitato. Ma se invece $f(t) = e^t + t$? Anche qui l'approssimazione sensata, per $t \to +\infty$, è $f(t) \approx e^t$, dato che l'esponenziale cresce più rapidamente del termine di primo grado t. Come formalizzare in modo preciso la frase "cresce ben più rapidamente"?

Lo stesso tipo di problema sorge nel caso di quantità infinitesime. Come confrontare termini che diventano molto piccoli (tendenti a zero)?

Ordine di infinito per $x \to +\infty$. Consideriamo qui funzioni f tali che

$$\lim_{x \to +\infty} |f(x)| = +\infty$$

(il caso $x \to -\infty$ è analogo). Come distinguere tra funzioni di questo genere quelle che divergono "più rapidamente" e quelle che divergono "meno rapidamente"? Ad esempio le funzioni x^α, $\ln x, e^x, a^x$ (con $\alpha > 0$ e $a > 1$) divergono per $x \to +\infty$ in modi essenzialmente differenti. Quale di queste funzioni cresce più rapidamente delle altre?

Dato che

x	1	10	100	1000
x^2	1	100	10000	1000000
x^3	1	1000	1000000000	1000000000000

ci aspettiamo che x^3 tenda all'infinito più rapidamente di x^2. La maniera rigorosa per esprimere questo concetto è studiare il rapporto delle due quantità. Dato che

$$\lim_{x \to +\infty} \frac{x^3}{x^2} = +\infty,$$

la grandezza di x^3 relativamente a quella di x^2 è maggiore. Questa proprietà si esprime dicendo che x^3 *tende a* $+\infty$ *più rapidamente di* x^2 *per* $x \to +\infty$.

Allo stesso modo, x^α cresce più rapidamente di x^β per $\alpha > \beta$

$$\lim_{x\to+\infty} \frac{x^\alpha}{x^\beta} = +\infty \qquad \forall\, \alpha > \beta.$$

DEFINIZIONE 1.1. Ordine di infinito I. *Siano f e g tali che*

$$\lim_{x\to+\infty} |f(x)| = \lim_{x\to+\infty} |g(x)| = +\infty. \tag{42}$$

Si dice che: f ha ordine di infinito superiore rispetto a g per $x \to +\infty$ *se*

$$\lim_{x\to+\infty} \frac{|f(x)|}{|g(x)|} = +\infty.$$

Analogamente, f ha ordine di infinito inferiore rispetto a g per $x \to +\infty$ *se*

$$\lim_{x\to+\infty} \frac{|f(x)|}{|g(x)|} = 0.$$

Può capitare che due funzioni abbiano lo stesso tipo di andamento all'infinito. Ad esempio: x e $2x$ sono entrambi polinomi di grado 1 e vale:

$$\lim_{x\to+\infty} \frac{x}{2x} = \frac{1}{2}.$$

Il fatto che il limite del rapporto sia una costante non zero, suggerisce che le due funzioni hanno lo stesso ordine di grandezza. Si sarebbe tentati di dire che due funzioni hanno lo stesso ordine di grandezza se e solo se

$$\lim_{x\to+\infty} \frac{|f(x)|}{|g(x)|} = \ell > 0.$$

Questa definizione, seppure ragionevole, è troppo restrittiva: non è in grado di coprire casi con termini oscillanti. Un esempio: per $f(x) = x(1 + \sin^2 x)$ e $g(x) = x$, si ha

$$\frac{|f(x)|}{|g(x)|} = \frac{|x||1 + \sin^2 x|}{|x|} = 1 + \sin^2 x \in [1, 2],$$

che non ammette limite per $x \to +\infty$. Ma, dato che $x \le f(x) \le 2x$, è ragionevole affermare che f tende all'infinito con la stessa velocità di x.

DEFINIZIONE 1.2. Ordine di infinito II. *Si dice che le funzioni f e g, soddisfacenti* (42)*, hanno* lo stesso ordine di infinito per $x \to +\infty$ *se esistono $C_1, C_2 > 0$ tali che*

$$0 < C_1 \le \frac{|f(x)|}{|g(x)|} \le C_2 \qquad \text{per } x \text{ sufficientemente grande.} \tag{43}$$

In generale, per verificare la condizione (43) occorre stimare il rapporto $|f|/|g|$ e non sempre questa operazione è facile. Però, se esiste ed è diverso da zero il limite

$$\lim_{x\to+\infty} \frac{|f(x)|}{|g(x)|} = \ell \neq 0,$$

la condizione (43) è automaticamente soddisfatta. Ad esempio, consideriamo le funzioni $f(x) = 2x^2 - 1$ e $g(x) = x^2 + x + 3$:

$$\lim_{x\to+\infty} \frac{2x^2-1}{x^2+x+3} = 2 \neq 0,$$

quindi hanno lo stesso ordine di infinito. In generale, se f è un polinomio di grado m e g un polinomio di grado p, allora: se $m > p$, f è di ordine superiore a g; se $m = p$, f e g sono dello stesso ordine; se $m < p$, f è di ordine inferiore a g.

Osservazione 1.3. Se f è di ordine superiore rispetto a g, allora la funzione somma $f + g$ ha lo stesso ordine di f, infatti

$$\lim_{x\to+\infty} \frac{f(x)+g(x)}{f(x)} = \lim_{x\to+\infty} \left(1 + \frac{g(x)}{f(x)}\right) = 1.$$

Ad esempio, la funzione $x + \ln x$ ha lo stesso ordine di infinito di x per $x \to +\infty$

$$\lim_{x\to+\infty} \frac{x+\ln x}{x} = \lim_{x\to+\infty} \left(1 + \frac{\ln x}{x}\right) = 1.$$

Definizione 1.4. *Se per una funzione f esiste un valore $\alpha > 0$ tale che f è dello stesso ordine di infinito di $|x|^\alpha$ per $x \to +\infty$ si dice che f è un* infinito di ordine α. *La funzione $|x|$ è detta* infinito campione per $x \to +\infty$.

Ad esempio, la funzione $\sqrt{1+x^2}$ è tale che

$$\lim_{x\to+\infty} \frac{\sqrt{1+x^2}}{x} = \lim_{x\to+\infty} \sqrt{\frac{1}{x^2}+1} = 1,$$

quindi ha ordine di infinito uguale ad 1 per $x \to +\infty$. La funzione $\dfrac{4x^3+1}{x-5}$ è tale che

$$\lim_{x\to+\infty} \frac{(4x^3+1)/(x-5)}{x^2} = \lim_{x\to+\infty} \frac{4x^3+1}{x^3-5x^2} = 4$$

quindi ha ordine di infinito 2. In generale, l'ordine di infinito di una funzione razionale con numeratore di grado m e denominatore di grado p (con $m > p$) è $m - p$ (dimostratelo!).

Osservazione 1.5. Se una funzione f ha ordine di infinito α, allora

$$\lim_{x\to+\infty} \frac{f(x)}{x^{\alpha+\varepsilon}} = \lim_{x\to+\infty} \frac{f(x)}{x^\alpha}\frac{1}{x^\varepsilon} = 0 \qquad \forall\, \varepsilon > 0;$$

cioè ha ordine di infinito inferiore rispetto ad $x^{\alpha+\varepsilon}$ per ogni $\varepsilon > 0$. Analogamente

$$\lim_{x\to+\infty} \frac{f(x)}{x^{\alpha-\varepsilon}} = \lim_{x\to+\infty} \frac{f(x)}{x^{\alpha}}\, x^{\varepsilon} = \infty \qquad \forall\, \varepsilon > 0,$$

quindi ha ordine di infinito superiore rispetto ad $x^{\alpha-\varepsilon}$ per ogni $\varepsilon > 0$.

Si potrebbe pensare di introdurre una "scala assoluta" di ordine di grandezza delle funzioni, attribuendo a ciascuna funzione divergente la corrispondente potenza che la rappresenta. Questa scala, però, non adempie il compito richiesto: ci sono funzioni la cui "velocità di divergenza" non corrisponde a quella di nessuna potenza e quindi che, in questo senso, non possono essere classificate. I due casi più rilevanti sono dati dalla funzione $\ln x$ e da e^x per cui vale

$$\lim_{x\to+\infty} \frac{\ln x}{x^{\alpha}} = 0, \qquad \lim_{x\to+\infty} \frac{e^{x}}{x^{\alpha}} = +\infty \qquad \forall\, \alpha > 0. \tag{44}$$

Per dimostrare (44), osserviamo preliminarmente che vale la disequazione[1]

$$\ln t \le t \qquad \forall\, t > 0. \tag{45}$$

Ora, dato $\alpha > 0$, scegliamo $\varepsilon \in (0, \alpha)$. Applicando la disequazione (45) per $t = x^{\alpha-\varepsilon}$ e usando le proprietà del logaritmo:

$$\ln x \le \frac{1}{\alpha-\varepsilon}\, x^{\alpha-\varepsilon} \qquad \forall\, x > 0,\ 0 < \varepsilon < \alpha.$$

Dividendo entrambi i membri per x^{α} e passando al limite,

$$0 \le \frac{\ln x}{x^{\alpha}} \le \frac{1}{(\alpha-\varepsilon)x^{\varepsilon}} \qquad \forall\, x > 0 \qquad \Rightarrow \qquad \lim_{x\to+\infty} \frac{\ln x}{x^{\alpha}} = 0.$$

Per il secondo limite in (44), ponendo $y = e^x$,

$$\lim_{x\to+\infty} \frac{e^{x}}{x^{\alpha}} = \lim_{y\to+\infty} \frac{y}{(\ln y)^{\alpha}} = \lim_{y\to+\infty} \left(\frac{y^{1/\alpha}}{\ln y}\right)^{\alpha} = +\infty.$$

Le formule in (44) affermano che $\ln x$ diverge per $x \to +\infty$ più lentamente di qualsiasi potenza x^{α} e che e^x diverge più rapidamente di qualsiasi potenza x^{α}.

Osservazione 1.6. Con le funzioni esponenziali e con i logaritmi, è possibile costruire funzioni che divergono a velocità sempre più grandi o a velocità sempre più piccole. Ad esempio,

$$\lim_{x\to+\infty} \frac{\ln(\ln x)}{\ln x} = \lim_{y\to+\infty} \frac{\ln y}{y} = 0, \qquad \lim_{x\to+\infty} \frac{e^{(e^x)}}{e^x} = \lim_{x\to+\infty} \frac{e^{y}}{y} = +\infty,$$

e quindi $\ln(\ln x)$ è un infinito di ordine inferiore a $\ln x$ e $e^{(e^x)}$ è di ordine superiore a e^x.

[1] Infatti, detta $F(t) = t - \ln t$, allora $F'(t) = 1 - 1/t$ e quindi F ha un punto di minimo assoluto per $t = 1$. Perciò $F(t) \ge F(1) = 1 > 0$.

OSSERVAZIONE 1.7. Cosa succede per a^x e $\log_a x$ con $a > 1$ (e diverso da e)? La funzione $\log_a x$ si può scrivere in termini della funzione $\ln x$:

$$\log_a x = \frac{\ln x}{\ln a}$$

(sapete giustificare questa formula?). Quindi

$$\lim_{x\to+\infty} \frac{\log_a x}{x^\alpha} = \lim_{x\to+\infty} \frac{\ln x}{x^\alpha \ln a} = 0 \qquad \forall\, \alpha > 0.$$

Procedendo come nel passaggio dal limite riguardante il logaritmo naturale al limite per l'esponenziale con base e, deduciamo che

$$\lim_{x\to+\infty} \frac{a^x}{x^\alpha} = +\infty \qquad \forall\, \alpha > 0,\ a > 1.$$

E' interessante confrontare tra loro esponenziali e logaritmi con basi diverse:

$$\lim_{x\to+\infty} \frac{\log_a x}{\log_b x} = \lim_{x\to+\infty} \frac{\ln x}{\ln a}\frac{\ln b}{\ln x} = \frac{\ln b}{\ln a} \qquad \forall a, b > 1,$$

quindi logaritmi con basi diverse hanno lo stesso ordine di infinito. Per gli esponenziali, invece, dati $a, b > 1$

$$\lim_{x\to+\infty} \frac{a^x}{b^x} = \lim_{x\to+\infty} \left(\frac{a}{b}\right)^x = \begin{cases} 0 & a < b, \\ 1 & a = b, \\ +\infty & b < a, \end{cases}$$

che mostra che esponenziali con base maggiore hanno ordine di infinito maggiore.

Ordine di infinito e comportamento asintotico. È possibile che per una funzione f valga una decomposizione del tipo

$$f(x) = g(x) + h(x) \qquad \text{con} \lim_{x\to+\infty} h(x) = 0, \tag{46}$$

dove la funzione g è una funzione "nota" (ad esempio, una funzione con un asintoto obliquo). Se la funzione $|g|$ diverge a $+\infty$ per $x \to +\infty$, allora

$$\lim_{x\to+\infty} |f(x)| = \lim_{x\to+\infty} |g(x) + h(x)| = +\infty.$$

Come sono collegati gli ordini di infinito di f e g? La risposta è semplice: *le funzioni f e g hanno lo stesso ordine di infinito*, infatti

$$\lim_{x\to+\infty} \frac{f(x)}{g(x)} = \lim_{x\to+\infty} \frac{g(x) + h(x)}{g(x)} = \lim_{x\to+\infty} \left(1 + \frac{h(x)}{g(x)}\right) = 1. \tag{47}$$

Ad esempio, tutte le funzioni che possiedono un asintoto obliquo (non orizzontale) per $x \to +\infty$ hanno ordine di infinito 1.

E' vero il viceversa? Supponiamo di sapere che una funzione f abbia lo stesso ordine di infinito di una funzione g:

$$\lim_{x\to+\infty} |f(x)| = \lim_{x\to+\infty} |g(x)| = +\infty \quad \text{e} \quad 0 < C_1 \leq \frac{|f(x)|}{|g(x)|} \leq C_2.$$

E' vero che vale una rappresentazione come quella data in (46)? La risposta, in generale, è negativa. Ad esempio, per le funzioni $f(x) = x + \ln x$ e $g(x) = x$

$$\lim_{x\to+\infty} \frac{x+\ln x}{x} = \lim_{x\to+\infty} 1 + \frac{\ln x}{x} = 1,$$

ma la differenza tra f e g se ne guarda bene dal tendere a zero per $x \to +\infty$

$$h(x) = f(x) - g(x) = (x - \ln x) - x = \ln x \qquad \Rightarrow \qquad \lim_{x\to+\infty} h(x) = +\infty.$$

Ordine di infinito per $x \to x_0$. Così come si confrontano i comportamenti delle funzioni per $x \to +\infty$, è possibile confrontare funzioni che divergono in un punto x_0. La terminologia è analoga alla precedente.

Definizione 1.8. *Ordine di infinito III. Date due funzioni f e g tali che*

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

si dice che f è un infinito di ordine superiore rispetto a g per $x \to x_0$ *se*

$$\lim_{x\to x_0} \frac{|f(x)|}{|g(x)|} = +\infty.$$

Se il limite è 0, f è un infinito di ordine inferiore rispetto a g per $x \to x_0$. *Infine, se esistono C_1, C_2 e $\delta > 0$ per cui*

$$0 < C_1 \leq \frac{|f(x)|}{|g(x)|} \leq C_2 \qquad \textit{per} \quad 0 < |x - x_0| < \delta \tag{48}$$

si dice che f e g hanno stesso ordine di infinito per $x \to x_0$.

Come nel caso di $x \to +\infty$, se $\lim\limits_{x\to x_0} f(x)/g(x) = \ell \neq 0$, allora è automaticamente verificata la condizione (48) e, quindi, f e g sono dello stesso ordine. Ad esempio

$$\lim_{x\to 0} \left|\frac{1}{x}\right| = \lim_{x\to 0} \left|\frac{1}{\sin x}\right| = +\infty \quad \text{e} \qquad \lim_{x\to 0} \frac{1/x}{1/\sin x} = \lim_{x\to 0} \frac{\sin x}{x} = 1,$$

quindi $f(x) = 1/x$ e $g(x) = 1/\sin x$ sono "infiniti" dello stesso ordine per $x \to 0$.

Definizione 1.9. *Se la funzione f ha lo stesso ordine di infinito di $1/|x - x_0|^\alpha$ per qualche $\alpha > 0$, si dice che f è un* infinito di ordine α per $x \to x_0$. *La funzione $1/|x - x_0|$ è detta* infinito campione per $x \to x_0$.

Anche qui si può ripetere quanto detto in precedenza: esistono funzioni che non hanno un ordine di infinito per $x \to x_0$. Un esempio? Con il cambio di variabile $y = -\ln x$, si deduce che

$$\lim_{x\to 0^+} \frac{\ln x}{1/x^\alpha} = \lim_{y\to+\infty} \frac{y}{e^{\alpha y}} = 0,$$

che mostra che $\ln x$ ha un ordine di infinito in 0 inferiore a qualsiasi potenza.

Ordine di infinitesimo. Così come si confrontano infiniti, è possibile confrontare funzioni infinitesime per $x \to x_0$ con $x_0 \in \mathbb{R}$ oppure per $x \to \pm\infty$. Qui, per abbreviare l'esposizione, scriviamo x_0 per indicare o un numero reale, o uno dei due simboli $\pm\infty$.

DEFINIZIONE 1.10. <u>*Ordine di infinitesimo.*</u> *Siano f e g infinitesime per $x \to x_0$. Si dice che f è un* infinitesimo di ordine superiore a g per $x \to x_0$ *se*

$$\lim_{x\to x_0} \frac{|f(x)|}{|g(x)|} = 0. \tag{49}$$

Se il limite è $+\infty$, f è un infinitesimo di ordine inferiore a g. *Infine, f e g hanno lo* stesso ordine di infinitesimo *se in un intorno di x_0 (nel caso di $x_0 = \pm\infty$ si intende per valori sufficientemente grandi),*

$$0 < C_1 \le \left|\frac{f(x)}{g(x)}\right| \le C_2 \qquad \textit{per qualche } C_1, C_2 > 0.$$

Come nel caso degli infiniti, si introducono infinitesimi campione.

– Se $x_0 \in \mathbb{R}$, si dice che f è un infinitesimo di ordine α, se è dello stesso ordine di $|x - x_0|^\alpha$. La funzione $|x - x_0|$ è l'infinitesimo campione per $x \to x_0$.
– Se $x_0 = \pm\infty$, si dice che f è un infinitesimo di ordine α, se è dello stesso ordine di $1/|x|^\alpha$. La funzione $1/|x|$ è l'infinitesimo campione per $x \to \pm\infty$.

Qualche esempio (tanto per gradire):

$$\lim_{x\to 0} \frac{\sin x}{x} = 1 \qquad \Rightarrow \qquad \sin x \text{ è infinitesimo di ordine 1 per } x \to 0$$

$$\lim_{x\to\pm\infty} \frac{1/(1+x^2)}{1/x^2} = 1 \qquad \Rightarrow \qquad \frac{1}{1+x^2} \text{ è infinitesimo di ordine 2 per } x \to \pm\infty$$

$$\lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2} \qquad \Rightarrow \qquad 1-\cos x \text{ è infinitesimo di ordine 2 per } x \to 0$$

$$\lim_{x\to\pm\infty} \frac{\arctan(1/x)}{1/x} = 1 \qquad \Rightarrow \qquad \arctan\left(\frac{1}{x}\right) \text{ è infinitesimo di ordine 1 per } x \to \pm\infty$$

I simboli di Landau: "O" e "o". Per indicare che una funzione f ha un ordine di grandezza inferiore a quello di un'altra funzione g si usa la notazione $f = o(g)$ per $x \to x_0$ (si legge f *è un "o piccolo" di* g). Il significato di questa affermazione è che f/g tende a zero per $x \to x_0$. Ad esempio,

$$x^\alpha = o(x^\beta) \qquad \forall\, \alpha < \beta \qquad\qquad \text{per } x \to +\infty.$$

Abbiamo anche visto che

$$\begin{aligned} &\ln x = o(x^\alpha) && \forall\, \alpha > 0 && \text{per } x \to +\infty, \\ &x^\alpha = o(e^x) && \forall\, \alpha > 0 && \text{per } x \to +\infty, \\ &1 - \cos x = o(x) && && \text{per } x \to 0. \end{aligned}$$

Esercizio 1.11. *Verificare la validità di*

$$\frac{1}{\sqrt{1+4x^2}} = \frac{1}{2x} + o\left(\frac{1}{x}\right) \qquad\qquad \textit{per } x \to +\infty.$$

Analogamente si introduce la notazione $f = O(g)$ (si legge f *è un "O grande" di* g) per indicare che la funzione f ha al più l'ordine di grandezza di g, ossia se il rapporto f/g è limitato in un intorno di x_0

$$\left|\frac{f(x)}{g(x)}\right| \le C,$$

per qualche $C > 0$, in un intorno di x_0. Ad esempio,

$$\sqrt{10x-1} = O(\sqrt{x}) \qquad\qquad \text{per } x \to +\infty.$$

Infatti, dato che

$$\lim_{x\to+\infty} \frac{\sqrt{10x-1}}{\sqrt{x}} = \lim_{x\to+\infty} \sqrt{10 - \frac{1}{x}} = \sqrt{10},$$

il rapporto di $\frac{\sqrt{10x-1}}{\sqrt{x}}$ è limitato per valori della x sufficientemente grandi.

Dai limiti notevoli

$$\lim_{x\to 0} \frac{e^x - 1}{x} = 1, \quad \lim_{x\to 0} \frac{\ln(1+x)}{x} = 1, \quad \lim_{x\to 0} \frac{\sin x}{x} = 1, \quad \lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2},$$

segue che, per $x \to 0$,

$$e^x = 1 + O(x), \quad \ln(1+x) = O(x), \quad \sin x = O(x), \quad \cos x = 1 + O(x^2).$$

Derivabilità con i simboli di Landau. Tramite i simboli di Landau si può riscrivere la derivabilità di una funzione in modo diverso. La derivabilità di una funzione f in x può essere scritta nella forma

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

Quindi, se una funzione è derivabile in x, vale la relazione (particolarmente significativa)

$$f(x+h) - f(x) = f'(x)h + o(h) \qquad \text{per } x \to 0. \tag{50}$$

L'interpretazione di questa formula è che l'incremento di f si può rappresentare come un termine $f'(x)h$, lineare nell'incremento h, più un resto che ha un ordine di grandezza inferiore ad h per $h \to 0$. Si usa la terminologia:

$$\textsf{differenziale di } f\text{:} \qquad df(x;h) := f'(x)h.$$

Fissato il valore di x, il differenziale $df(x;h)$ rappresenta un'approssimazione (valida a meno di $o(h)$) dell'incremento $\Delta f(x;h) := f(x+h) - f(x)$:

$$\Delta f(x;h) \approx df(x;h) \qquad \text{per } h \to 0.$$

Come si precisa il senso del simbolo $\approx$? Proprio tramite i simboli di Landau:

$$|\Delta f(x;h) - df(x;h)| = o(h) \qquad \text{per } h \to 0,$$

che esprime il fatto che l'errore che si commette sostituendo all'incremento Δf, il differenziale df è un infinitesimo di ordine superiore al primo.

2. Il Teorema di de l'Hôpital

Supponiamo f e g continue nell'intervallo (a,b) e sia $x_0 \in (a,b)$ tale che

$$f(x_0) = g(x_0) = 0.$$

In questa situazione non è evidente se esista e quanto valga il limite

$$\lim_{x \to x_0} \frac{f(x)}{g(x)}.$$

Sostituendo, formalmente, a f e g il valore nel punto limite si ottiene l'espressione $\frac{0}{0}$ che non ha senso. L'esistenza o meno del limite è legata alla rapidità con cui le funzioni f e g tendono a 0 per $x \to x_0$, ossia alla relazione che c'è tra i loro ordini di infinitesimo. Come risolvere un limite del genere? Se le funzioni f e g sono derivabili in x_0, si può pensare di approssimare le funzioni f e g con la loro retta tangente nel punto x_0:

$$\frac{f(x)}{g(x)} \approx \frac{f(x_0) + f'(x_0)(x-x_0)}{g(x_0) + g'(x_0)(x-x_0)} = \frac{f'(x_0)(x-x_0)}{g'(x_0)(x-x_0)} = \frac{f'(x_0)}{g'(x_0)}.$$

Come rendere rigorosa tale affermazione? Utilizzando (50), si ha

$$\frac{f(x)}{g(x)} = \frac{f(x_0) + f'(x_0)(x-x_0) + o(x-x_0)}{g(x_0) + g'(x_0)(x-x_0) + o(x-x_0)} = \frac{f'(x_0) + \frac{o(x-x_0)}{x-x_0}}{g'(x_0) + \frac{o(x-x_0)}{x-x_0}} \tag{51}$$

avendo utilizzato la proprietà $f(x_0) = g(x_0) = 0$. Passando al limite per $x \to x_0$, nel caso in cui $g'(x_0) \neq 0$, si ottiene

$$\lim_{x \to x_0} \frac{f(x)}{g(x)} = \frac{f'(x_0)}{g'(x_0)} \qquad (g'(x_0) \neq 0).$$

Questa formula è nota come regola di de l'Hôpital[2]. Lavorando in maniera più raffinata, si dimostra una variante della precedente formula di de l'Hôpital, che non richiede la derivabilità delle funzioni f e g nel punto limite, ma solo l'esistenza del limite del rapporto delle derivate. Anche tale variante è nota sotto lo stesso nome.

Teorema 2.1. Regola di de l'Hôpital. *Siano f e g due funzioni derivabili tali che $f(x_0) = g(x_0) = 0$. Se esiste finito*

$$\lim_{x \to x_0} \frac{f'(x)}{g'(x)} = \ell \in \mathbb{R},$$

allora esiste anche il limite $\lim_{x \to x_0} f(x)/g(x)$ ed ha lo stesso valore ℓ.

La conclusione è valida anche nel caso in cui il rapporto delle derivate abbia limite $+\infty$ o $-\infty$. Non si può dedurre nessuna conclusione nel caso in cui il rapporto delle derivate non abbia limite.

Esistono in commercio anche altre versioni del Teorema di de l'Hôpital che si applicano a casi differenti: forme indeterminate del tipo $0/0$ per $x \to \pm\infty$ o del tipo ∞/∞ per $x \to x_0$ e per $x \to \pm\infty$

$$\lim_{x \to \pm\infty} f(x) = \lim_{x \to \pm\infty} g(x) = 0, \qquad \lim_{x \to \pm\infty} \frac{f'(x)}{g'(x)} = \ell \quad \Rightarrow \quad \lim_{x \to \pm\infty} \frac{f(x)}{g(x)} = \ell$$

$$\lim_{x \to x_0} |f(x)| = \lim_{x \to x_0} |g(x)| = \infty, \qquad \lim_{x \to x_0} \frac{f'(x)}{g'(x)} = \ell \quad \Rightarrow \quad \lim_{x \to x_0} \frac{f(x)}{g(x)} = \ell$$

$$\lim_{x \to \pm\infty} |f(x)| = \lim_{x \to \pm\infty} |g(x)| = \infty, \qquad \lim_{x \to \pm\infty} \frac{f'(x)}{g'(x)} = \ell \quad \Rightarrow \quad \lim_{x \to \pm\infty} \frac{f(x)}{g(x)} = \ell.$$

Il simbolo ℓ può essere sia un numero reale sia $+\infty$ o $-\infty$.

Il principio è sempre lo stesso: nel caso di una forma indeterminata $0/0$ o ∞/∞, si può calcolare il limite del rapporto delle derivate. Se tale limite esiste, allora dà anche il valore del limite iniziale. Se, invece, il rapporto delle derivate non esiste, non si può

[2]Questa regola porta il nome il nome del matematico francese Guillaume Francois Antoine, marchese de L'Hôpital (1661 – 1704), che pubblicò la formula nel suo libro *Analyse des infiniment petits pour l'intelligence des lignes courbes* (1692). La regolar è in realtà dovuta a Jean Bernoulli, a cui L'Hôpital pagava una pensione di 300 franchi annui in cambio delle informazioni relative ai suoi progressi nel calcolo infinitesimale e della risoluzione di alcuni problemi posti dal de L'Hôpital (tra cui quello di determinare il limite di forme indeterminate). L'Hôpital, riconoscendo che parte del contenuto del suo trattato era dovuta a Bernoulli, preferì pubblicarlo in forma anonima. Ciò nonostante, una volta scoperto l'autore del libro, la formula fu associata al suo nome.

concludere nulla. Nel caso in cui il limite del rapporto delle derivate dia luogo, esso stesso, ad una forma indeterminata $0/0$ o ∞/∞, si può applicare di nuovo il Teorema di de l'Hôpital e (provare a) calcolare il limite del rapporto delle derivate seconde.

Esempio 2.2. Per calcolare il limite

$$\lim_{x\to 0} \frac{x - \arctan x}{x - \sin x}, \tag{52}$$

studiamo il limite del rapporto delle derivate

$$\lim_{x\to 0} \frac{1 - \frac{1}{1+x^2}}{1 - \cos x} = \lim_{x\to 0} \frac{x^2}{(1+x^2)(1-\cos x)} = 2,$$

dato che $\lim\limits_{x\to 0}(1-\cos x)/x^2 = 1/2$. Quindi il limite (52) esiste e vale 2.

Esempio 2.3. Calcoliamo

$$\lim_{x\to+\infty} \left(\frac{\pi}{2} - \arctan x\right) x.$$

Questo limite è della forma $0 \cdot \infty$, ma si può ricondurre alla tipologia trattabile con il teorema di de l'Hôpital riscrivendolo come

$$\lim_{x\to+\infty} \frac{\pi/2 - \arctan x}{1/x}.$$

Il rapporto delle derivate ha limite:

$$\lim_{x\to+\infty} \frac{-1/(1+x^2)}{-1/x^2} = \lim_{x\to+\infty} \frac{x^2}{1+x^2} = 1.$$

Quindi, il limite richiesto esiste e vale 1.

Esercizio 2.4. *Calcolare il limite*

$$\lim_{x\to 0} \frac{e^x - 1 - x}{\cos x - 1}.$$

La dimostrazione del Teorema di de L'Hôpital. Per cominciare, enunciamo (e dimostriamo) una variante del Teorema di Lagrange, nota come Teorema di Cauchy.

Teorema 2.5. *Teorema di Cauchy. Siano f e g due funzioni continue in $[a,b]$ e derivabili in (a,b). Allora esiste $\xi \in (a,b)$ tale che*

$$\det \begin{pmatrix} f(b) - f(a) & f'(\xi) \\ g(b) - g(a) & g'(\xi) \end{pmatrix} = 0, \tag{53}$$

cioè $(f(b) - f(a))g'(\xi) = f'(\xi)(g(b) - g(a))$.

Interpretazione geometrica. Date le funzioni f e g, consideriamo la funzione *vettoriale* (f,g) che associa ad un punto dell'intervallo $[a,b]$ il punto del piano di coordinate (f,g). Il Teorema di Cauchy asserisce che esiste sempre un valore $\xi \in (a,b)$ tale che il vettore incremento $(f(b)-f(a), g(b)-g(a))$ e il vettore "derivata" $(f'(x), g'(x))$ calcolato in $x=\xi$ sono paralleli.

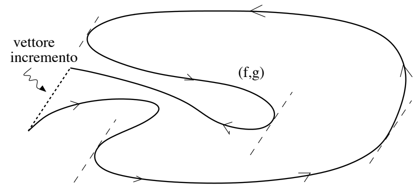

FIGURA 1. L'interpretazione geometrica del Teorema di Cauchy.

DIMOSTRAZIONE DEL TEOREMA 2.5. Consideriamo la funzione

$$\Phi(x) := \det\begin{pmatrix} f(b)-f(a) & f(x) \\ g(b)-g(a) & g(x) \end{pmatrix} \equiv \big[f(b)-f(a)\big]\, g(x) - f(x)\big[g(b)-g(a)\big].$$

La funzione Φ è continua in $[a,b]$ e derivabile in (a,b). Inoltre, si ha

$$\Phi(a) = \det\begin{pmatrix} f(b)-f(a) & f(a) \\ g(b)-g(a) & g(a) \end{pmatrix} = \det\begin{pmatrix} f(b) & 0 \\ g(b) & 0 \end{pmatrix} = 0.$$

$$\Phi(b) = \det\begin{pmatrix} f(b)-f(a) & f(b) \\ g(b)-g(a) & g(b) \end{pmatrix} = \det\begin{pmatrix} -f(a) & 0 \\ -g(a) & 0 \end{pmatrix} = 0.$$

Quindi, per il Teorema di Rolle, esiste $\xi \in (a,b)$ tale che $\Phi'(\xi) = 0$. Dato che $\Phi'(x) = (f(b)-f(a))g'(x) - f'(x)(g(b)-g(a))$, segue la conclusione. □

Armati del precedente risultato, si può dimostrare il Teorema di de l'Hôpital.

DIMOSTRAZIONE DEL TEOREMA 2.1. Scegliamo, nella formula (53), $a = x_0$ e $b = x$. Dato che, per ipotesi, $f(x_0) = g(x_0) = 0$, si ha

$$\frac{f(x)}{g(x)} = \frac{f'(\xi)}{g'(\xi)}, \tag{54}$$

dove ξ è compreso tra x_0 e x. Per $x \to x_0$, necessariamente $\xi \to x_0$, e il termine $f'(\xi)/g'(\xi)$ tende ad ℓ per ipotesi. Dato che questo termine è uguale a $f(x)/g(x)$, ne segue la conclusione □

Approssimazioni polinomiali. Utilizziamo adesso il Teorema di de l'Hôpital per dedurre delle approssimazioni polinomiali di funzioni con un errore che sia infinitesimo di ordine sempre più alto. Scegliamo come cavia la funzione $\sin x$. Dato che è derivabile in 0 e la sua derivata è 1,

$$\sin x = x + o(x) \qquad \text{per } x \to 0 \tag{55}$$

o, equivalentemente,

$$\lim_{x\to 0} \frac{\sin x - x}{x} = 0.$$

Per dedurre un'approssimazione per $\sin x$ con un errore che sia $o(x^2)$, calcoliamo

$$\lim_{x\to 0} \frac{\sin x - x}{x^2}.$$

Per applicare il Teorema di de l'Hôpital, studiamo il limite del rapporto delle derivate

$$\lim_{x\to 0} \frac{\cos x - 1}{2x} = -\frac{1}{2} \lim_{x\to 0} \frac{1 - \cos x}{x} = 0.$$

Dato che tale limite esiste finito, anche il limite di partenza esiste e vale 0. Quindi

$$\sin x = x + o(x^2) \qquad \text{per } x \to 0 \tag{56}$$

La formula (56) dice che l'errore che si commette approssimando $\sin x$ con x è un infinitesimo di ordine superiore al secondo per $x \to 0$. Questa informazione è migliore di quella data da (55), che ci garantiva solamente un errore di ordine superiore al primo.

Per ottenere un approsimazione con errore di ordine superiore al terzo, ragioniamo come in precedenza e calcoliamo

$$\lim_{x\to 0} \frac{\sin x - x}{x^3} = \lim_{x\to 0} \frac{\cos x - 1}{3x^2} = -\frac{1}{3} \lim_{x\to 0} \frac{1 - \cos x}{x^2} = -\frac{1}{6},$$

che implica $\sin x = x + O(x^3)$. Portando il termine $-1/6$ a primo membro, otteniamo

$$\lim_{x\to 0} \frac{\sin x - x + \frac{1}{6}x^3}{x^3} = 0$$

cioè la funzione $\sin x$ è pari a $x - \frac{1}{6}\,x^3$ più un errore superiore a x^3

$$\sin x = x - \frac{1}{6}\,x^3 + o(x^3) \qquad \text{per } x \to 0. \tag{57}$$

Possiamo iterare il procedimento e calcolare

$$\lim_{x\to 0} \frac{\sin x - x + \frac{1}{6}x^3}{x^4} = \lim_{x\to 0} \frac{\cos x - 1 + \frac{1}{2}x^2}{4x^3} = \lim_{x\to 0} \frac{-\sin x + x}{12x^2} = 0,$$

quindi

$$\sin x = x - \frac{1}{6}\,x^3 + o(x^4) \qquad \text{per } x \to 0. \tag{58}$$

Ripetiamo l'esperimento su una cavia diversa: e^x. Il fatto che e^x sia derivabile in $x = 0$ e la derivata valga 1 si traduce nella formula

$$e^x = 1 + x + o(x) \qquad \text{per } x \to 0. \tag{59}$$

Per migliorare l'espressione, calcoliamo

$$\lim_{x\to 0} \frac{e^x - 1 - x}{x^2} = \lim_{x\to 0} \frac{e^x - 1}{2x} = \frac{1}{2},$$

cioè $e^x = 1 + x + O(x^2)$. Il limite può essere riscritto come

$$\lim_{x\to 0} \frac{e^x - 1 - x}{x^2} - \frac{1}{2} = 0 \qquad \Longrightarrow \qquad \lim_{x\to 0} \frac{e^x - 1 - x - \frac{1}{2}x^2}{x^2} = 0,$$

cioè $e^x - 1 - x - \frac{1}{2}x^2 = o(x^2)$, o anche

$$e^x = 1 + x + \frac{1}{2}x^2 + o(x^2).$$

Allo stesso modo

$$\lim_{x\to 0} \frac{e^x - 1 - x - \frac{1}{2}x^2}{x^3} = \lim_{x\to 0} \frac{e^x - 1 - x}{3x^2} = \frac{1}{6}, \tag{60}$$

La relazione (60) si può riscrivere come

$$\lim_{x\to 0} \frac{1}{x^3} \left(e^x - 1 - x - \frac{1}{2}x^2 - \frac{1}{6}x^3 \right) = 0 \iff e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + o(x^3)$$

Cosa stiamo facendo iterando questo procedimento? Stiamo ottenendo delle approssimazioni ad ordini sempre più alti di una funzione data. La relazione $e^x = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3 + o(x^3)$ esprime il fatto che la funzione e^x si può approssimare, per $x \to 0$ con il polinomio $1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3$ commettendo un errore (la differenza tra e^x e il polinomio) che tende a zero per $x \to 0$ con ordine superiore a 3 (cioè più rapidamente di x^3).

L'iterazione dell'algoritmo che abbiamo visto conduce direttamente a quello che si chiama *polinomio di Taylor*.

3. La formula di Taylor

Replichiamo, in generale, l'esperimento fatto su $\sin x$ e e^x alla fine del paragrafo precedente. Se f è una funzione derivabile in x_0, si ha

$$\lim_{x\to x_0} \frac{f(x) - [f(x_0) + f'(x_0)(x - x_0)]}{x - x_0} = 0,$$

o, equivalentemente,

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + o(|x - x_0|) \qquad \text{per} \quad x \to x_0,$$

che esprime che la funzione f, vicino ad x_0, si approssima con la funzione $f(x_0) + f'(x_0)(x - x_0)$, con un errore che è un infinitesimo di ordine superiore ad 1.

Per ottenere un'approssimazione più precisa, supponendo che la funzione f si derivabile due volte, possiamo calcolare il limite

$$\lim_{x\to x_0}\frac{f(x)-[f(x_0)+f'(x_0)(x-x_0)]}{(x-x_0)^2}=\lim_{x\to x_0}\frac{f'(x)-f'(x_0)}{2(x-x_0)}=\frac{1}{2}f''(x_0),$$

avendo applicato il Teorema di de l'Hôpital. Il precedente limite si può scrivere come

$$\lim_{x\to x_0}\frac{f(x)-\left[f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2}f''(x_0)(x-x_0)^2\right]}{(x-x_0)^2}=0$$

ossia

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2}f''(x_0)(x-x_0)^2+o(|x-x_0|^2).$$

Così abbiamo scoperto che la funzione $f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2}f''(x_0)(x-x_0)^2$ approssima f, vicino ad x_0, con un errore di ordine superiore a 2. Il grafico della funzione $p(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2}f''(x_0)(x-x_0)^2$ rappresenta la parabola che meglio approssima la funzione f per $x\to x_0$.

Iterando ancora una volta il procedimento e supponendo che la funzione f sia derivabile tre volte in x_0, si ottiene

$$f(x)=f(x_0)+f'(x_0)(x-x_0)+\frac{1}{2}f''(x_0)(x-x_0)^2+\frac{1}{3!}f'''(x_0)(x-x_0)^3+o(|x-x_0|^3).$$

E in generale?

Teorema 3.1. Formula di Taylor. *Sia $f:(a,b)\to\mathbb{R}$ derivabile n volte in (a,b) e sia $x_0\in(a,b)$. Dato $n\in\mathbb{N}$, posto*

$$T_n(x;x_0):=f(x_0)+f'(x_0)(x-x_0)+\cdots+\frac{1}{n!}f^{(n)}(x_0)(x-x_0)^n,$$

si ha

$$\lim_{x\to x_0}\frac{f(x)-T_n(x;x_0)}{(x-x_0)^n}=0,$$

cioè vale la decomposizione $f(x)=T_n(x;x_0)+o(|x-x_0|^n)$.

Definizione 3.2. *Il polinomio $T_n(x;x_0)$ si chiama* polinomio di Taylor[3] di grado n della funzione f nel punto x_0 *e rappresenta un'approssimazione di f vicino ad x_0.*

La peculiarità della formula di Taylor sta nel fatto che il resto R_n, definito da

$$R_n(x;x_0):=f(x)-T_n(x;x_0),$$

è un infinitesimo di ordine superiore ad $|x-x_0|^n$ per $x\to x_0$.

[3]Se $x_0=0$, il polinomio p_n viene, a volte, chiamato polinomio di McLaurin.

Dimostrazione. Applichiamo il Teorema di de l'Hôpital al limite

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)-f'(x_0)(x-x_0)-\cdots-\frac{1}{(n-1)!}f^{(n-1)}(x_0)(x-x_0)^{n-1}}{(x-x_0)^n}$$
$$=\lim_{x\to x_0}\frac{f'(x)-f'(x_0)-f''(x_0)(x-x_0)-\cdots-\frac{1}{(n-2)!}f^{(n-1)}(x_0)(x-x_0)^{n-2}}{n(x-x_0)^{n-1}}.$$

Dato che sia il numeratore che il denominatore sono infinitesimi, possiamo applicare nuovamente il Teorema di de l'Hôpital, ottenendo

$$\lim_{x\to x_0}\frac{f''(x)-f''(x_0)-f'''(x_0)(x-x_0)\cdots-\frac{1}{(n-3)!}f^{(n-1)}(x_0)(x-x_0)^{n-3}}{n(n-1)(x-x_0)^{n-2}}.$$

Iterando $n-1$ volte il procedimento si ottiene

$$\lim_{x\to x_0}\frac{f^{(n-1)}(x)-f^{(n-1)}(x_0)}{n!(x-x_0)}=\frac{1}{n!}f^{(n)}(x_0).$$

Quindi vale

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)-\cdots-\frac{1}{(n-1)!}f^{(n-1)}(x_0)(x-x_0)^{n-1}}{(x-x_0)^n}=\frac{1}{n!}f^{(n)}(x_0),$$

da cui segue

$$\lim_{x\to x_0}\frac{f(x)-f(x_0)-\cdots-\frac{1}{(n-1)!}f^{(n-1)}(x_0)(x-x_0)^{n-1}}{(x-x_0)^n}-\frac{1}{n!}f^{(n)}(x_0)=0,$$

che porta alla conclusione. □

Esempio 3.3. Polinomi. Assegnati $a_0,\dots,a_p$, sia

$$f(x)=a_0+a_1x+a_2x^2+\cdots+a_px^p.$$

Consideriamo, prima di tutto, lo sviluppo in $x_0=0$. Dato che

$$\begin{aligned}
f'(x)&=a_1+2a_2\,x+\cdots+pa_px^{p-1}\\
f''(x)&=2a_2+3\cdot 2a_3\,x+\cdots+p(p-1)a_px^{p-2}\\
&\ \vdots\\
f^{(p)}(x)&=p(p-1)\cdots 2\cdot 1\,a_p\\
f^{(k)}(x)&=0 \qquad\qquad k>p
\end{aligned}$$

si ha

$$f'(0)=a_1,\quad f''(0)=2a_2,\quad \dots\quad f^{(p)}(0)=p!\,a_p,\quad f^{(k)}(0)=0\qquad k>p.$$

Quindi,

$$T_n(x;0)=\begin{cases}a_0+a_1x+\cdots+a_nx^n & n<p,\\ a_0+a_1x+\cdots+a_px^p & n\ge p.\end{cases}$$

Come era naturale aspettarsi, il polinomio di Taylor di f in $x_0 = 0$ e di grado n si ottiene considerando i termini del polinomio con grado minore o uguale ad n.

Per il polinomio di Taylor in $x_0 \neq 0$, occorre riscrivere il polinomio in termini di potenze di $h := x - x_0$. In questo modo si otterrà un'espressione del tipo

$$f(x) = b_0 + b_1(x - x_0) + \cdots + b_p(x - x_0)^p$$

con $b_0, b_1, \dots, b_p$ opportuni. Il polinomio di Taylor è dato da

$$T_n(x; x_0) = \begin{cases} b_0 + b_1(x - x_0) + \cdots + b_n(x - x_0)^n & n < p, \\ b_0 + b_1(x - x_0) + \cdots + b_p(x - x_0)^p & n \geq p. \end{cases}$$

Consideriamo, ad esempio, la funzione $f(x) = x + x^3$. Fissato $x_0 \in \mathbb{R}$, per scriverla in termini di potenze di $h = x - x_0$ calcoliamo

$$f(x_0 + h) = (x_0 + h) + (x_0 + h)^3 = x_0 + x_0^3 + (1 + 3x_0^2)h + 3x_0h^2 + h^3.$$

Quindi vale l'identità

$$x + x^3 = x_0 + x_0^3 + (1 + 3x_0^2)(x - x_0) + 3x_0(x - x_0)^2 + (x - x_0)^3.$$

Ad esempio, il polinomio di Taylor di grado 2 in x_0 è

$$T_2(x; x_0) = x_0 + x_0^3 + (1 + 3x_0^2)(x - x_0) + 3x_0(x - x_0)^2.$$

Lo stesso, evidentemente, si ottiene applicando direttamente la formula: da

$$f(x) = x + x^3, \qquad f'(x) = 1 + 3x^2, \qquad f''(x) = 6x,$$

segue

$$T_2(x; x_0) = x_0 + x_0^3 + (1 + 3x_0^2)(x - x_0) + 3x_0(x - x_0)^2.$$

Esempio 3.4. *Esponenziale in $x_0 = 0$.* Siano

$$f(x) = e^x \qquad \text{e} \qquad x_0 = 0.$$

Dato che $f^{(k)}(x) = e^x$ per ogni $k \in \mathbb{N}$, si ha $f^{(k)}(0) = e^0 = 1$ per ogni k, quindi il polinomio di Taylor di grado n è

$$T_n(x; 0) = \sum_{k=0}^{n} \frac{x^k}{k!} = 1 + x + \frac{x^2}{2} + \cdots + \frac{x^n}{n!}.$$

Questa formula è coerente con la definizione di esponenziale data in precedenza.

Che succede se $x_0 \neq 0$? I conti non sono molto diversi:

$$T_n(x; 0) = e^{x_0} \sum_{k=0}^{n} \frac{(x - x_0)^k}{k!} = e^{x_0} \left[1 + (x - x_0) + \frac{(x - x_0)^2}{2} + \cdots + \frac{(x - x_0)^n}{n!}\right].$$

ESEMPIO 3.5. <u>*Seno e coseno in $x_0 = 0$.*</u> Sia $f(x) = \sin x$, allora

$$f^{(2k)}(x) = (-1)^k \sin x, \qquad f^{(2k+1)}(x) = (-1)^k \cos x \qquad \forall\, k = 0, 1, 2, \ldots.$$

Calcolando in $x_0 = 0$, otteniamo

$$f^{(2k)}(0) = 0, \qquad f^{(2k+1)}(0) = (-1)^k \qquad \forall\, k = 0, 1, 2, \ldots.$$

Ne segue che

$$T_n(x; 0) = x - \frac{x^3}{3!} + \frac{x^5}{5!} - \ \ldots\ + (-1)^k \frac{x^{2k+1}}{(2k+1)!}$$

è il polinomio di Taylor di grado n con $n = 2k+1$ o $2k+2$. In effetti si può dimostrare che vale l'uguaglianza

$$\sin x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \qquad \forall\, x \in \mathbb{R}.$$

Analogamente se consideriamo la funzione $f(x) = \cos x$ abbiamo

$$f^{(2k)}(x) = (-1)^k \cos x, \qquad f^{(2k+1)}(x) = (-1)^{k+1} \sin x \qquad \forall\, k = 0, 1, 2, \ldots.$$

Calcolando in $x_0 = 0$, otteniamo

$$f^{(2k)}(0) = (-1)^k, \qquad f^{(2k+1)}(0) = 0 \qquad \forall\, k.$$

Ne segue che

$$T_n(x; 0) = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \frac{x^6}{6!} + \ \ldots\ + (-1)^k \frac{x^{2k}}{(2k)!}$$

è il polinomio di Taylor di $\cos x$ centrato in 0 di grado n con $n = 2k$ o $2k+1$. Anche per il coseno vale un'uguaglianza analoga alla precedente:

$$\cos x = \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \qquad \forall x \in \mathbb{R}.$$

ESERCIZIO 3.6. *Calcolare il polinomio di Taylor di grado 4 della funzione* $\sin x$ *centrato in $x_0 = \pi/2$ e quello centrato in $x_0 = \pi/4$.*

ESEMPIO 3.7. Siano $f(x) = \dfrac{1}{1-x}$ e $x_0 = 0$. Le derivate di f sono

$$f'(x) = \frac{1}{(1-x)^2}, \qquad f''(x) = \frac{2}{(1-x)^3}, \qquad \ldots, \qquad f^{(k)}(x) = \frac{k!}{(1-x)^{k+1}}.$$

Perciò $f(0) = 1$, $f'(0) = 1$, $f''(0) = 2$,... $f^{(k)}(0) = k!$. Quindi il polinomio di Taylor è

$$T_n(x; 0) = 1 + x + x^2 + \cdots + x^n = \sum_{k=0}^{n} x^k$$

coerente con l'espressione nota per la serie geometrica.

Esempio 3.8. Come ultimo esempio, consideriamo
$$f(x) = \ln(1+x) \qquad \text{e} \qquad x_0 = 0.$$
Le derivate di f sono
$$f'(x) = \frac{1}{1+x}, \qquad f''(x) = -\frac{1}{(1+x)^2}, \qquad \ldots, \qquad f^{(k)}(x) = (-1)^{k+1}\frac{(k-1)!}{(1+x)^k},$$
e quindi $f(0) = 0$, $f'(0) = 1$, $f''(0) = -1$, $f^{(k)}(0) = (-1)^{k+1}(k-1)!$. Il polinomio di Taylor di grado n in $x_0 = 0$ è
$$T_n(x;0) = x - \frac{x^2}{2} + \cdots + (-1)^{n+1}\frac{x^n}{n} = \sum_{k=1}^{n}(-1)^{k+1}\,\frac{x^k}{k}.$$

4. Espressioni del resto

Data una funzione f, con un buon numero di derivate, sappiamo determinare un polinomio che la approssimi vicino ad un punto assegnato x_0. In questa approssimazione viene commesso un errore pari a
$$R_n(x;x_0) := f(x) - \sum_{k=0}^{n}\frac{1}{k!}f^{(k)}(x_0)(x-x_0)^k.$$
Quali proprietà conosciamo su R_n? Per ora sappiamo solo che
$$R_n(x;x_0) = o(|x-x_0|^n) \quad \text{cioè} \quad \lim_{x\to x_0}\frac{R_n(x;x_0)}{(x-x_0)^n} = 0.$$
Questa è solo un'informazione sul comportamento al limite, quindi non dice nulla di preciso sulla grandezza della quantità R_n in punti $x \neq x_0$. Per poter stimare l'errore occorre una rappresentazione migliore di R_n. Ecco il nostro nuovo obiettivo.

Teorema 4.1. Resto in forma di Lagrange. *Se la funzione f è derivabile $n+1$ volte, esiste ξ, tra x_0 e x, tale che*
$$f(x) = \sum_{k=0}^{n}\frac{1}{k!}f^{(k)}(x_0)(x-x_0)^k + \frac{1}{(n+1)!}\,f^{(n+1)}(\xi)\,(x-x_0)^{n+1},$$

Dimostrazione. Sia f derivabile $n+1$ volte e consideriamo le funzioni
$$F(x) := R_n(x;x_0) = f(x) - \sum_{k=0}^{n}\frac{1}{k!}f^{(k)}(x_0)(x-x_0)^k \qquad \text{e} \qquad G(x) := (x-x_0)^{n+1}.$$
Dato che $F(x_0) = G(x_0) = 0$, applicando il Teorema di Cauchy a F e G, segue
$$\frac{F(x)-F(x_0)}{G(x)-G(x_0)} = \frac{F(x)}{G(x)} = \frac{F'(\xi_1)}{G'(\xi_1)},$$

per qualche ξ_1, compreso tra x_0 e x. Derivando le espressioni di F e G, otteniamo

$$F'(x) = f'(x) - \sum_{k=0}^{n-1} \frac{1}{k!} f^{(k+1)}(x_0)(x-x_0)^k \qquad \text{e} \qquad G'(x) = (n+1)(x-x_0)^n.$$

Se $n \geq 1$ è possibile riapplicare il Teorema di Cauchy, trovando ξ_2, compreso tra x_0 e ξ_1 e quindi anche tra x_0 e x, per cui

$$\frac{F(x)-F(x_0)}{G(x)-G(x_0)} = \frac{F'(\xi_1)}{G'(\xi_1)} = \frac{F''(\xi_2)}{G''(\xi_2)}.$$

Iterando $n+1$ volte il procedimento, si dimostra l'esistenza di ξ_{n+1} tra x_0 e x tale che

$$\frac{F(x)-F(x_0)}{G(x)-G(x_0)} = \frac{F^{(n+1)}(\xi_{n+1})}{G^{(n+1)}(\xi_{n+1})}.$$

Dato che $F^{(n+1)}(x) = f^{(n+1)}(x)$ e $G^{(n+1)}(x) = (n+1)!$, si deduce (qui $\xi = \xi_{n+1}$)

$$F(x) - F(x_0) = \frac{1}{(n+1)!} f^{(n+1)}(\xi)\,(G(x) - G(x_0))$$

da cui, ricordando le definizioni di F e G,

$$R_n(x; x_0) = \frac{1}{(n+1)!} f^{(n+1)}(\xi)\,(x-x_0)^{n+1},$$

cioè la conclusione. □

A partire da questa espressione del resto, possiamo stimare l'errore commesso quando si approssimi una funzione f con il suo polinomio di Taylor: se $M > 0$ è tale che $|f^{(n+1)}(t)| \leq M$ per ogni t tra x e x_0, allora

$$|R_n(x;x_0)| = \left| \frac{f^{(n+1)}(\xi)}{(n+1)!}(x-x_0)^{n+1} \right| = \frac{|f^{(n+1)}(\xi)|}{(n+1)!}|x-x_0|^{n+1} \leq \frac{M}{(n+1)!}|x-x_0|^{n+1}.$$

Calcolo approssimato di $\sin(1/10)$ *con stima dell'errore.* Abbiamo già considerato questo problema proponendo come "candidato" per l'approssimazione il valore $1/10$. In quell'occasione avevamo stimato l'errore commesso con $1/100$. Il procedimento era basato sul Teorema di Lagrange e sull'approssimazione della funzione $\sin x$ con la sua tangente nell'origine:

$$\sin x \approx x \qquad\qquad \text{per } x \to 0.$$

Detta $f(x) = \sin x$, la stima dell'errore discendeva da

$$\begin{aligned} |f(x) - f(x_0) - f'(x_0)(x-x_0)| &= |(f'(\xi) - f'(x_0))(x-x_0)| \\ &= |f''(\eta)(\xi - x_0)(x-x_0)| \leq |f''(\eta)||x-x_0|^2. \end{aligned}$$

dove $x = 1/10$ e $x_0 = 0$. Dato che $f''(x) = -\sin x$, la stima è immediata.

Come ottenere stime migliori? La scelta naturale è approssimare la funzione $\sin x$ con un suo polinomio di Taylor di grado opportuno. Ad esempio,

$$\sin x \approx x - \frac{x^3}{6} \qquad \text{per } x \to 0.$$

Quindi un'approssimazione migliore della precedente è

$$\sin\left(\frac{1}{10}\right) \approx \frac{1}{10} - \frac{1}{6000} = 0,0998\bar{3}.$$

Scriviamo l'errore con la forma di Lagrange $R_3(x; x_0) = \frac{1}{4!} f^{(4)}(\xi)(x - x_0)^4$, cioè

$$\sin\frac{1}{10} - \left(\frac{1}{10} - \frac{1}{6000}\right) = R_3(1/10; 0) = \frac{\sin\xi}{4!}\,\frac{1}{10^4},$$

quindi

$$|R_3(1/10; 0)| \leq \frac{1}{24 \cdot 10^4} = 4,1\bar{6} \times 10^{-6}.$$

In realtà il polinomio $x - \frac{x^3}{6}$ è anche il polinomio di Taylor di $\sin x$ in 0 di grado 4, quindi il resto può anche essere scritto come

$$\sin\frac{1}{10} - \frac{1}{10} + \frac{1}{6000} = R_4(1/10; 0) = \frac{\cos\xi}{5!}\,\frac{1}{10^5},$$

quindi

$$|R_4(1/10; 0)| \leq \frac{1}{120 \cdot 10^5} = 8,\bar{3} \times 10^{-8}.$$

In definitiva

$$\sin\left(\frac{1}{10}\right) = 0,0998\bar{3} \pm 8,\bar{3} \times 10^{-8}.$$