Università degli studi di Roma "La Sapienza"
a.a. 2003/04

**Note per il corso**

# Derivate ed Integrali

## I parte

versione 2.0 (October 7, 2003)

L. Lamberti & C. Mascia

con l'indispensabile consulenza di
C. Cassisa, E. Montefusco e V. Nesi

Corsi di laurea in Fisica, Fisica ed Astrofisica,
Tecnologie Fisiche e dell'Informazione

# Indice


Capitolo 1. I numeri reali 1
1. Numeri naturali, interi, razionali e irrazionali 1
2. Ordine e diseguaglianze 11
3. La struttura metrica: il modulo 13
4. La verità sui reali 16
5. Estremo superiore ed estremo inferiore 19
Appendice A. La formula del binomio 23

Capitolo 2. Funzioni: anno zero 25
1. Ingredienti di base 25
2. Operazioni elementari su grafici 31
3. Funzioni invertibili e funzioni monotòne 35
4. Classi di funzioni più o meno comuni 39
5. Problemi di massimo e minimo 44

Capitolo 3. Incontri ravvicinati con i limiti: le successioni 47
1. Limite di successioni 48
2. Il limite entra in società 55
3. Calcolo di alcuni limiti 61
4. Successioni monotòne 63
5. Serie numeriche 64
Appendice A. M'approssimo in un denso 71
Appendice B. La serie armonica e la serie armonica generalizzata 75

Capitolo 4. Le funzioni continue 77
1. Limite di funzioni 77
2. Continuità 86
3. Esempi di discontinuità 91
4. Teoremi sulle funzioni continue 92
5. Gli intervalli incapsulati: *"divide et impera"* 96
Appendice A. Altre applicazioni del *divide et impera* 99

Capitolo 5. L'integrale 101
1. L'area di un sottografico: strategia d'attacco 101
2. Ma cos'è quest'integrale? 108
3. Istruzioni per l'uso 111
4. Classi di funzioni integrabili 117
5. Teorema della media integrale 119
6. Funzioni definite tramite integrali 121

Appendice A. Quale senso ha la parola "area"? 123
Appendice B. Il volume di solidi di rotazione 126

Capitolo 6. Derivate, derivate e derivate 129
1. Definizione di derivata 132
2. Regole fondamentali di derivazione 137
3. Derivate successive 142
4. Il Teorema di Lagrange 143
5. Il Teorema fondamentale del calcolo integrale 152
Appendice A. Derivate parziali 156
Appendice B. Tabella delle derivate 158

CAPITOLO 1

# I numeri reali

Come ogni buona storia, cominciamo con la presentazione dei protagonisti di tutta la vicenda che andiamo a raccontare. Signori e signore... i numeri reali. Va precisato fin dall'inizio, per non deludere nessuno, che l'obiettivo che ci proponiamo è di introdurre i numeri reali in maniera "intuitivamente convincente" e non "rigorosamente soddisfacente", in modo da dedurne le proprietà fondamentali. Non fa parte dello scopo di queste Note una costruzione rigorosa dei reali. Senza offesa per nessuno.

## 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*. Ogni passaggio corrisponde ad un vuoto da colmare, ad una proprietà mancante nella classe dei numeri con cui ci si trova a lavorare che spinge ad introdurre oggetti che generalizzano la classe precedente e per cui la proprietà richiesta è soddisfatta.

**Numeri naturali.** I numeri naturali nascono da uno dei problemi primordiali dell'uomo (specie se insonne): il "conteggio". Cioè nascono dall'esigenza di rispondere alla domanda: "quanti?" Uno, nessuno, centomila... risponderebbe Pirandello. Più in generale, ci va bene una qualsiasi risposta del tipo: $0, 1, 2, 3, \ldots$. L'insieme di questi numeri, detti NUMERI NATURALI, si indica con il simbolo $\mathbb{N}$:

$$\text{NUMERI NATURALI:} \qquad \mathbb{N} = \{0, 1, 2, 3, \cdots\}.$$

Nell'insieme $\mathbb{N}$ sono definite le operazioni di addizione e moltiplicazione, per cui valgono le ben note *leggi commutative, associative, distributiva* (dato che si tratta di proprietà valide anche nell'insieme dei numeri reali, ne riscriviamo l'enunciato tra qualche pagina, pazientate).

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. Anche chi ha due fratelli, sa

bene che se avanzano due fette di torta, il problema è grave... a meno che non si abbia un coltello!

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*. Alla base del concetto di misura, c'è questa procedura:

– scegliere un'unità di misura (ad esempio, il metro campione);

– contare il numero di copie del segmento 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 $\mathbb{N}$. Ma c'è 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 $\mathbb{N}$. Anche in $\mathbb{Z}$ non è sempre possibile dividere! Perciò introduciamo un terzo insieme di numeri: l'insieme dei NUMERI RAZIONALI, cioè numeri che sono rapporto di numeri interi

$$\text{NUMERI RAZIONALI:} \quad \mathbb{Q} = \left\{ \frac{p}{q} \,:\, p \in \mathbb{Z},\ q \in \mathbb{N},\ q \neq 0,\ p, q \text{ senza fattori comuni} \right\}.$$

Attenzione! I numeri razionali possono essere scritti come rapporto di interi 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. Come già sapete e fate, i numeri del tipo $\frac{p}{1}$ si indicano semplicemente con $p$.

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 (qui sta il bello!) divisione, eccetto la divisione per zero. La maniera precisa di formulare il fatto che sia possibile "dividere per un qualsiasi numero razionale non nullo" è dire 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 con proprietà annesse e connesse 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 $R$ (a nostro gusto) come rappresentante di 0, che chiameremo ORIGINE, e un altro punto arbitrario come il punto 1. Definiamo la direzione da 0 a 1 come DIREZIONE POSITIVA, in questo modo, la retta $R$ diviene una retta *orientata*, cioè con una direzione preferenziale (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 $2, 3, \ldots$. 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 $\frac{p}{q}$ su $R$ per cui $p, q \in \mathbb{N}$ con $p$ più piccolo di $q$. Facile: dato che il numero $\frac{p}{q}$ è, moralmente, "$p$ volte $1/q$", si divide l'intervallo unitario

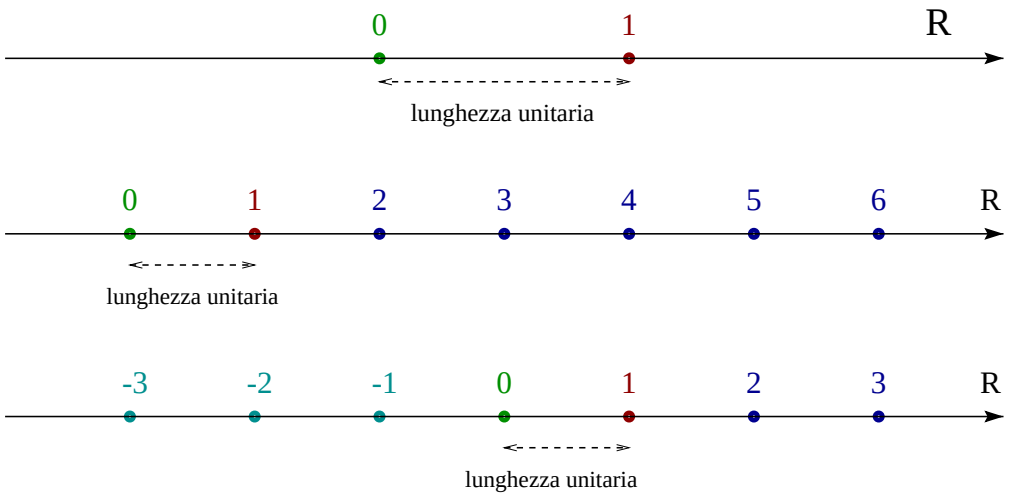



FIGURA 1. Dall'alto verso il basso: *(i)* i punti 0 e 1 sulla retta $R$ determinano il segmento di lunghezza unitaria; *(ii)* con "copie" del segmento unitario, si determina la posizione dei naturali sulla retta $R$; *(iii)* ripetendo il procedimento, questa volta alla sinistra di 0, si ottiene la rappresentazione 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.

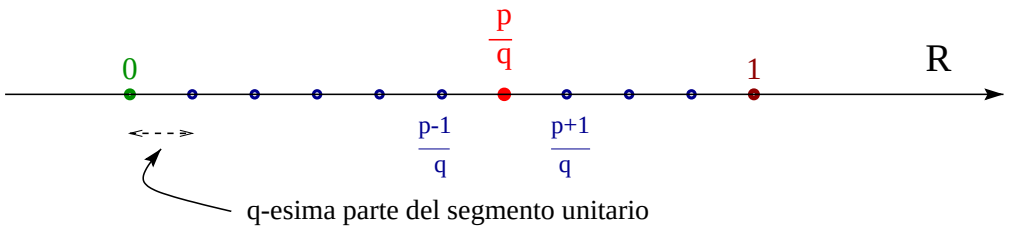



FIGURA 2. La rappresentazione su $R$ di un numero razionale $\frac{p}{q}$ per cui $p, q \in \mathbb{N}$ e $p < q$.

(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 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 $x$. Quindi, ad esempio, se $y$ si trova alla sinistra di 0, allora il punto $x + y$ è alla sinistra di $x$.

**Ordine, modulo e distanza nei numeri razionali.** Grazie al fatto che tutti i numeri razionali sono ora rappresentati su una retta $R$ e questa retta è orientata (la direzione da 0 a 1 è "positiva"), 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 *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 $x \geq 0$.*

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...

A questo punto, siamo in grado di introdurre un oggetto di fondamentale importanza: il *modulo*. Non si può sperare di uscire a testa alta da queste Note, senza aver acquisito un po' di familiarità con questo simbolo.

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

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

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 $y$.

Ricapitoliamo rapidamente il punto a cui siamo arrivati: partendo dai numeri naturali e passando per gli interi relativi, siamo arrivati a considerare i numeri razionali. Che struttura abbiamo trovato/costruito su $\mathbb{Q}$?

– struttura algebrica: con somma e prodotto (e relative proprietà), $\mathbb{Q}$ è un *campo*;

– relazione d'ordine: con l'introduzione di $\leq$, $\mathbb{Q}$ è *totalmente ordinato*;

– struttura metrica: grazie alla definizione del modulo, abbiamo introdotto in $\mathbb{Q}$ una nozione di distanza: $\mathbb{Q}$ è *dotato di metrica*.

La nostra fiducia (mal riposta) in $\mathbb{Q}$ ci spinge a sognare che questo insieme ci permetta di misurare tutte le lunghezze possibili...

**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 di stoffa (o di

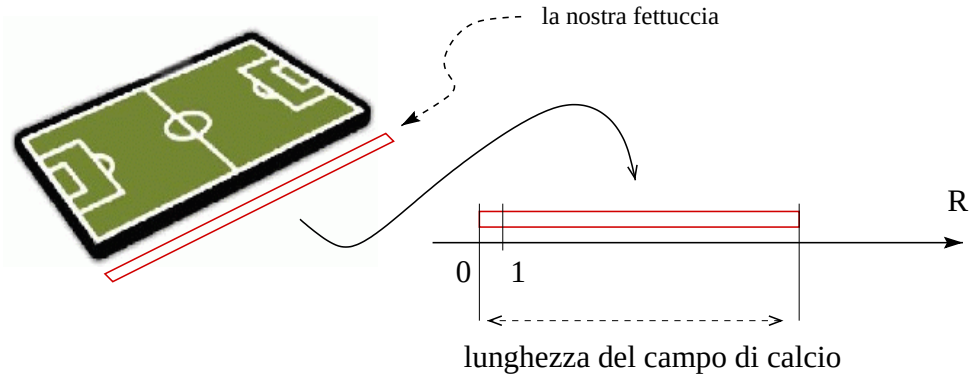



FIGURA 3. Il procedimento di misurazione: un campo di calcio.

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 in natura 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, $R \setminus \mathbb{Q} \neq \emptyset$. Una scoperta sensazionale (tanto da lasciarli sconvolti)! L'esempio più elementare di lunghezza non razionale è la lunghezza $\ell$ della diagonale di un quadrato di lato unitario. Infatti, per il teorema di Pitagora, $\ell^2 = 2$, ma *non esiste nessun numero razionale il cui quadrato sia* 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

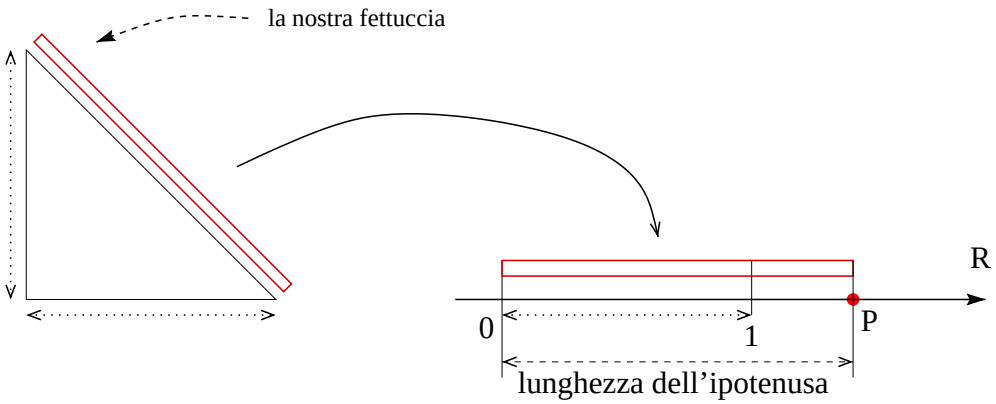



FIGURA 4. Il procedimento di misurazione dell'ipotenusa di un triangolo rettangolo di lato unitario: il punto $P$ non corrisponde a nessun numero razionale!

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 $\ell$ non può essere razionale.

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?*

Catastrofe! 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 $\mathbb{Q}$. Dato che "misurare" è il nostro obiettivo principe, non possiamo arrenderci proprio qui. 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:

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

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, non rigorosa e andrebbe precisata. A questo livello, però, ci accontentiamo di questa versione naïf. [1]

Il problema successivo è come sia possibile trattare con gli elementi di $\mathbb{R}$, ossia: quale rappresentazione possiamo dare ad un qualsiasi numero reale? Cosa significano le operazioni di somma e prodotto in $\mathbb{R}$? C'è ancora ordine? Si eredita il concetto di 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 $y$.

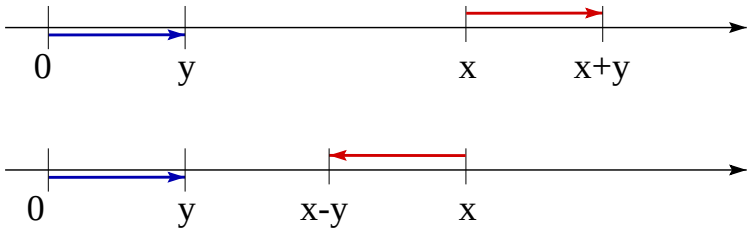



FIGURA 5. Somma e differenza di numeri reali.

Le operazioni di somma e prodotto in $\mathbb{Q}$ possono essere estese ad $\mathbb{R}$, di modo che anche $\mathbb{R}$ sia un CAMPO. Definire queste operazioni a partire dall'approssimazione tramite numeri razionali è possibile, ma troppo lungo per i tempi che ci sono concessi. 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 $a,b,c\in\mathbb{R}$;
*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.

Come somma/prodotto, anche ordinamento e distanza si possono estendere da $\mathbb{Q}$ ad $\mathbb{R}$.

[1] 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.

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 $x < y$.*

Anche 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... Bucato? Non bucato? La terminologia qui si fa un po' fumosa, e non per caso! Cosa vuol dire rigorosamente che "$\mathbb{R}$ non ha buchi"? Intuitivamente l'idea è chiara: $\mathbb{R}$ vuole essere la rappresentazione della nostra idea di "oggetto continuo" (pensiamo alla retta). Ma precisamente cosa si intende? Ci dedicheremo tra una manciata di pagine a spiegare in maniera più precisa come trasformare questa idea intuitiva in un oggetto matematicamente e logicamente chiaro. Non è comunque criminoso proseguire (per lo meno per un po') pensando ad $\mathbb{R}$ come un'intuitiva "retta senza buchi".

**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, indicandolo 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:

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

La validità di 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 $I := \{1, 2, 3, \ldots, 31\}$) e il numero del mese (che appartiene all'insieme $J := \{1, 2, 3, \ldots, 12\}$), 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 un piano, come si rappresentano gli insiemi

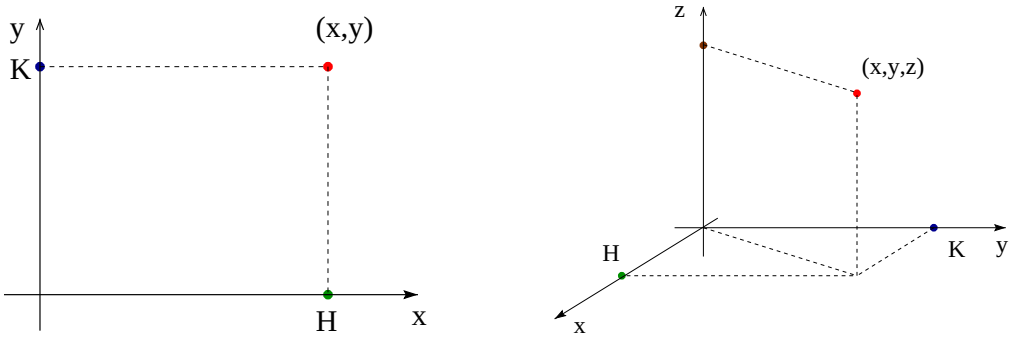



FIGURA 6. Il piano cartesiano $\mathbb{R}^2$ e lo spazio cartesiano $\mathbb{R}^3$.

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

Analogamente $\mathbb{R} \times \mathbb{R} \times \mathbb{R}$, o $\mathbb{R}^3$, indica l'insieme di terne <u>ordinate</u> di numeri reali $(x, y, z)$. L'insieme $\mathbb{R}^3$ si rappresenta con lo spazio, tramite una scelta di tre assi coordinati ortogonali tra loro.

Simile la situazione per $\mathbb{R}^n$ (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

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

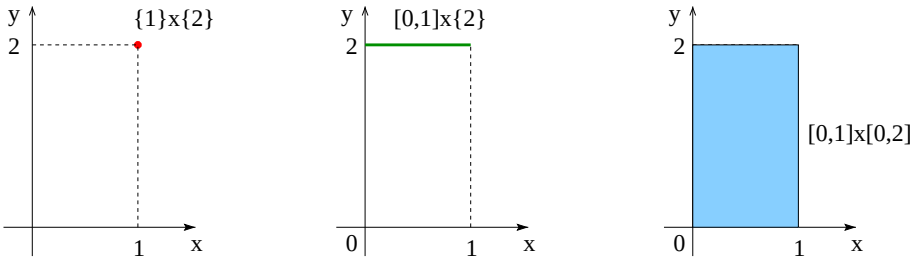



FIGURA 7. Gli insiemi $\{1\} \times \{2\}$, $[0,1] \times \{2\}$ e $[0,1] \times [0,2]$.

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.

## 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. Prima di tutto occorre ricordare che $a$ è maggiore di $b$ se e solo se nella retta reale orientata si trova alla sua destra, o, equivalentemente,

$$a > b \qquad \Longleftrightarrow \qquad a - b > 0.$$

*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

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

(iii) Le diseguaglianze possono essere sommate[2]: $a > b,\ c > d \quad \Rightarrow \quad a + c > b + d.$

(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 $\leq$.

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

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. Che succede in questi casi? Vediamo il caso dell'elevazione a potenza e delle radici $n-$esime.

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

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

Quindi l'operazione "elevare al quadrato" preserva l'ordine dei numeri positivi[3]: 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*

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

**Dimostrazione.** Dato che, nella scomposizione

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

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

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

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

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

(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:

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

## 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:

(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 DISEGUAGLIANZA 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 $-(a+b) \le |a| + |b|$, che discende dalla somma delle disequazioni $-a \le |a|$ e $-b \le |b|$. ∎

ESERCIZIO 3.1. *Verificare la validità dell'uguaglianza:* $|x| = \max\{x, -x\}$.

**Soluzione.** Se $x \ge 0$, allora $-x \le 0 \le x$, 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 $|a-b| \le |a| + |b|$. E' facile trovare esempi che mostrano che per certe scelte di $a$ e $b$ la diseguaglianza $|a-b| \le |a| - |b|$ è falsa.

Applicando due volte la diseguaglianza triangolare, si ottiene

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

Allo stesso modo, si ottiene la diseguaglianza più generale

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

o, equivalentemente (tramite il simbolo di sommatoria[4])

$$\left| \sum_{k=1}^{n} a_k \right| \le \sum_{k=1}^{n} |a_k| \qquad \forall a_1, \ldots, a_n \in \mathbb{R}, \tag{3.1}$$

Alcune conseguenze dirette della diseguaglianza triangolare sono molto utili.

PROPOSIZIONE 3.2. *Per ogni* $a, b \in \mathbb{R}$, *valgono*

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

*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

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

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

ESERCIZIO 3.3. *Dimostrare che, per ogni* $a > 0$,

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

---

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

**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 3.4. *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 3.5. *Dire quali tra i seguenti insiemi sono limitati e quali non lo sono:*

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

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

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

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 sempre 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 3.6. *Dimostrare le proprietà (a)-(b)-(c) della distanza utilizzando le proprietà (i)-(ii)-(iii) del modulo.*

DEFINIZIONE 3.7. *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 tramite $|x - x_0| < r$:

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

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?*

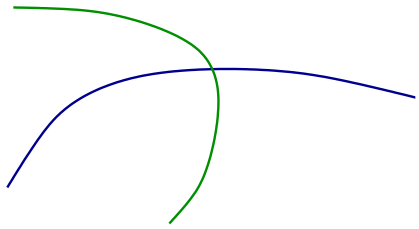


FIGURA 8. Due curve nel piano.

Un sondaggio del 2003 dà queste percentuali di risposta: il 78% degli intervistati dice "SI, SEMPRE", il 5% dice "QUALCHE VOLTA", il 3% dice "QUASI MAI", il 10% risponde "NON SO" e il 4% fugge scappando per timore di fare brutta figura. Evidentemente la risposta è racchiusa in quello che succede vicino al punto di incrocio delle due curve. Proponiamo tre maniere diverse di ragionare.

*Versione "atomistica" (o "alla Democrito").* Il nome è un po' altisonante, ma è bello per questo. La versione atomistica è basata sull'immaginare la curva come costituita 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, ormai polverizzate in tanti punti, non si incontrano. La risposta in questo caso è "QUASI MAI".

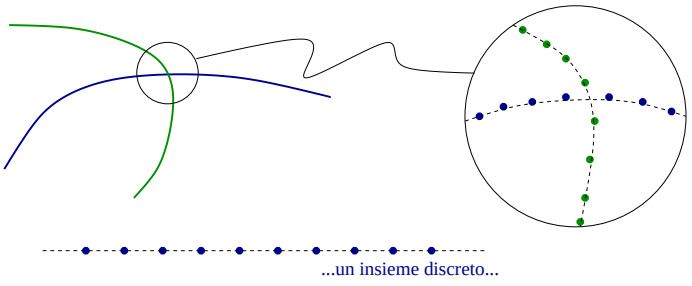



FIGURA 9. Versione "quantizzata": una retta è un'unione di punti a distanza fissata.

*Versione "razionale".* In questo caso si immagina la curva come una retta deformata, pensando la retta come unione dei soli punti razionali. Qui il disegno non è facile: dato che i punti sono densi, cioè 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 $x^2 = 2$, che si traduce nel fatto che 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

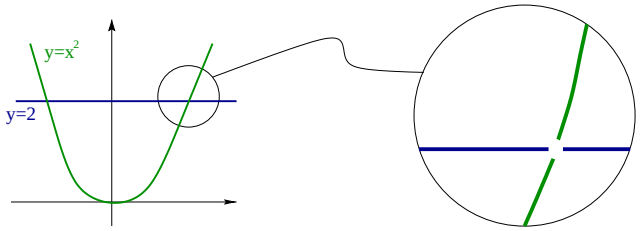



FIGURA 10. Versione "razionale": la parabola $y = x^2$ e la retta $y = 2$ nel piano $(x, y)$ con $x, y \in \mathbb{R}$ non si intersecano mai.

di pensiero risponde "QUALCHE VOLTA". Che traduce il fatto che se le due curve non si intersecano, è comunque possibile spostarle di (arbitrariamente) poco, in modo che si intersechino.

*Versione "reale".* Infine c'è la versione reale, la più intuitiva di tutte: rette e curve sono un "continuo" di punti, non ci sono buchi. Dunque le due curve si intersecano sempre. Ogni ingrandimento del punto di incontro delle curve dà sempre e comunque lo stesso tipo di figura.

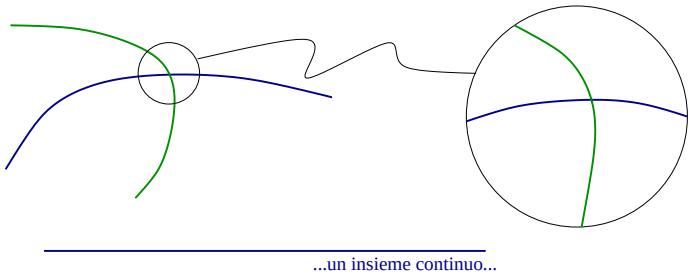



FIGURA 11. Versione "reale": l'idea intuitiva di "continuo".

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, punto e basta. *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 probabilmente sbagliata a livello microscopico!), è la più naturale.

Impareremo ad apprezzare l'insieme dei numeri reali lavorandoci, ma, prima di tutto, occorre rendere chiaro, attraverso alcuni assiomi, cosa voglia dire la frase "l'insieme dei reali non ha buchi". Ci sono vari modi per esprimere attraverso un assioma la *completezza dei numeri reali*, cioè di trasformare l'idea naïf di "non avere buchi" in una proprietà chiara da poter utilizzare tutte le volte che sia necessario. 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 (ad esempio, E.Giusti, "Esercizi e Complementi di Analisi Matematica, vol. I", Bollati Boringhieri, cap.2).

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

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

L'assioma si può anche esprimere così:

$$I_0, I_1, \ldots, I_n, \ldots \text{ intervalli chiusi e limitati}, \quad I_{n+1} \subset I_n \quad \forall\, n \in \mathbb{N} \quad \Rightarrow \quad \bigcap_{n \in \mathbb{N}} I_n \neq \emptyset.$$

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 $\ell$, corrispondente alla lunghezza dell'ipotenusa di un triangolo rettangolo con lati di lunghezza 1 ($\ell = \sqrt{2}$, per intendersi!). Se si sceglie la sequenza di intervalli incapsulati con lunghezza che diventa sempre più piccola in modo che l'intersezione degli intervalli sia data dal solo punto $\ell$, 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,*

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

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! ∎

*Il pescatore che aveva tempo da perdere. Un aneddoto dovrebbe illuminare sul significato della Proposizione. C'era una volta un pescatore che non aveva proprio niente da fare (per via dell'inquinamento, non c'era più nulla da prendere all'amo). Per vie traverse, era venuto a conoscenza di una barca ancorata nel punto $x$ dell'asse reale, che, sia con l'alta che con la bassa marea, si trovava sempre sull'acqua. Non riuscendo ad avvistarla, decise di fare qualche esperimento per capire più o meno dove si trovasse. Dopo aver piazzato nel punto 0 un palo di riferimento, il pescatore annoiato notò che, in qualsiasi posizione collocasse una bandierina, purché alla destra del palo di riferimento, in almeno un momento della giornata, essa si trovava in secca. Da questo concluse che la barca si trovava certamente alla sinistra del palo di riferimento. Siete d'accordo?*

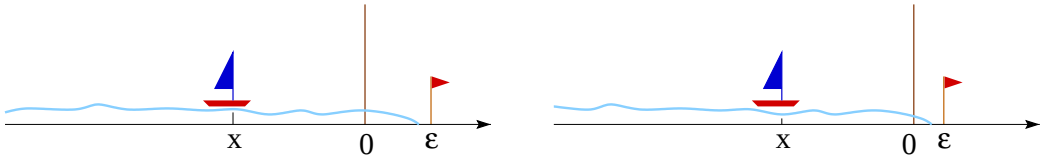



FIGURA 12. Due esperimenti del pescatore.

## 5. Estremo superiore ed estremo inferiore

La completezza dei numeri reali si porta appresso una serie di conseguenze notevoli. 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.

"Quanti anni ha Francesco Totti?" Domanda a bruciapelo a cui non so rispondere... sicuramente tra i 20 e i 30... Non è un'informazione precisa, ma meglio di niente. Cosa

rappresentano i numeri 20 e 30 rispetto alla reale età $x$ di Totti? Il primo è una stima per difetto e la seconda è una stima per eccesso: non so quanto vale $x$, ma posso con certezza assicurare che $20 \leq x \leq 30$. "Qual'è la fascia d'età dei giocatori della serie A di calcio?" Anche a questa domanda non so dare una risposta precisa, ma sono sicuro che abbiano tutti più di 15 anni e meno di 40: se $E$ è l'insieme delle età di tutti i giocatori di serie A, affermo che $15 \leq x \leq 40$ per ogni $x \in E$. I numeri 15 e 40 sono, rispettivamente, un minorante ed un maggiorante per l'insieme $E$. Ecco la definizione generale.

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 $\Lambda$ è una stima per eccesso di tutti gli elementi dell'insieme $E$ e un minorante $\lambda$ ne è una stima per difetto. Se, in qualche modo, siamo in grado di procurarci un minorante $\lambda$ ed un maggiorante $\Lambda$, 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 3.4) se e solo se è limitato superiormente e limitato inferiormente. Fatelo.

Il problema è che per uno stesso insieme $E$ ci sono 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 $\Lambda$/minorante $\lambda$ è elemento dell'insieme $E$ è chiaro che $\Lambda/\lambda$ dà la migliore stima per eccesso (per difetto). In questo caso si parla di *massimo/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 di tutte: $M$ è parte dell'insieme, quindi abbassare la soglia al di sotto di $M$ vuol dire non controllare più la collocazione di $E$.

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

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

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

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

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

La proprietà (ii) dell'estremo superiore $\Lambda$ garantisce che *non esiste una stima per eccesso migliore di* $\Lambda$. Ogni altro possibile maggiorante dell'insieme, necessariamente è maggiore (o uguale) a $\Lambda$. Similmente, la proprietà (ii) dell'estremo inferiore $\lambda$ garantisce che *non esiste una stima per difetto migliore di* $\lambda$. 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. Una specie di gioco delle tre carte. Infatti, rivedendo la definizione ci si accorge che la definizione di estremo superiore discende da questa strategia:

– dato l'insieme $E$, costruirne l'insieme $F = \{\Lambda \in \mathbb{R};\ x \leq \Lambda \quad \forall\, x \in E\}$ 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 /minorante dell'insieme $E$, allora *esiste sempre l'estremo superiore/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* $\inf \left\{ \frac{1}{n} \,:\, n = 1, 2, 3, \dots \right\} = 0$.

**Soluzione.** Bisogna dimostrare che $0 = \max E$ dove $E = \{x \in \mathbb{R} \,:\, x \leq \frac{1}{n} \quad \forall n \in \mathbb{N}\}$, cioè $0 \in E$ e 0 è un maggiorante per $E$. Dato che tutti i numeri $1/n$ sono positivi, $0 \in E$.

Resta da verificare che $x \leq 0$ per ogni $x \in E$. Sia $x \in E$. Per la proprietà di Archimede, per ogni $\varepsilon > 0$ esiste $n_\varepsilon$ tale che $n_\varepsilon \geq \frac{1}{\varepsilon}$, quindi $x \leq \frac{1}{n_\varepsilon} \leq \varepsilon$. In definitiva: per ogni $\varepsilon > 0$, $x \leq \varepsilon$. Perciò $x \leq 0$. $\square$

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!

## Appendice A. La formula del binomio

Questa Appendice è dedicata solo ed esclusivamente ad enunciare e dimostrare la FORMULA DEL BINOMIO

$$(x+y)^n = \sum_{k=0}^{n} \left( \begin{array}{c} n \\ k \end{array} \right) x^{n-k} y^k \qquad \forall x, y \in \mathbb{R}, \quad n \in \mathbb{N},$$

dove

$$\text{binomio di Newton} \quad : \qquad \left( \begin{array}{c} n \\ k \end{array} \right) := \frac{n!}{k!(n-k)!}$$

Uso il procedimento di *induzione*. Questo procedimento è utile nel caso di "affermazioni" in cui compaia un parametro naturale $n$. Consiste di due passi: il primo è di dimostrare l'affermazione per $n = 1$, il secondo è di dimostrare l'affermazione per $n+1$ supponendola vera per $n$. In questo modo, applicando il secondo passo ad $n = 1$ (verificata nel primo passo) si ottiene la proprietà per $n = 2$. Applicando il secondo passo con $n = 2$ si ottiene la validità per $n = 3$ e così via.

*Passo 1. Verifica per $n = 1$.*

Ponendo $n = 1$ si ha

$$n = 1\,; \quad \sum_{k=0}^{1} \left( \begin{array}{c} 1 \\ k \end{array} \right) x^{1-k} y^k = \left( \begin{array}{c} 1 \\ 0 \end{array} \right) x^1 y^0 + \left( \begin{array}{c} 1 \\ 1 \end{array} \right) x^0 y^1 = x + y = (x+y)^1.$$

*Passo 2. Supponendo la formula vera per $n$, la dimostro per $n+1$.*

Voglio dimostrare la formula per $n+1$, quindi, in sostanza, voglio rappresentare $(x+y)^{n+1}$ in termini di $x^{n+1-k}y^k$. Procedo cosi'

$$(x+y)^{n+1} = (x+y)(x+y)^n = (x+y) \sum_{k=0}^{n} \left( \begin{array}{c} n \\ k \end{array} \right) x^{n-k} y^k,$$

avendo usato nella seconda uguaglianza l'ipotesi induttiva che garantisce la validità della rappresentazione per $(x+y)^n$. Usando la proprietà distributiva un paio di volte, viene fuori che

$$(x+y)^{n+1} = x \sum_{k=0}^{n} \left( \begin{array}{c} n \\ k \end{array} \right) x^{n-k} y^k + y \sum_{k=0}^{n} \left( \begin{array}{c} n \\ k \end{array} \right) x^{n-k} y^k =$$

$$= \sum_{k=0}^{n} \left( \begin{array}{c} n \\ k \end{array} \right) x^{n+1-k} y^k + \sum_{k=0}^{n} \left( \begin{array}{c} n \\ k \end{array} \right) x^{n-k} y^{k+1}.$$

Ora, passaggio delicato per chi è poco familiare con l'uso delle sommatorie. Cambio nome all'indice della seconda somma. Introduco un nuovo indice che chiamo $h$ e che è definito da $h = k+1$. Riscrivo la seconda somma sostituendo sistematicamente $k$ con $h$ in maniera consistente con la definizione di $h$. Quindi

$$\sum_{k=0}^{n} \left( \begin{array}{c} n \\ k \end{array} \right) x^{n-k} y^{k+1} = \sum_{h=1}^{n+1} \left( \begin{array}{c} n \\ h-1 \end{array} \right) x^{n+1-h} y^h.$$

Nell'espressione a secondo membro posso (senza colpo ferire) risostituire l'indice $h$ con un nuovo indice $k$ per definizione uguale ad $h$, cioè $k = h$. Reinserendo tutto nella formula precedente

$$(x+y)^{n+1} = \sum_{k=0}^{n} \left( \begin{array}{c} n \\ k \end{array} \right) x^{n+1-k} y^k + \sum_{k=1}^{n+1} \left( \begin{array}{c} n \\ k-1 \end{array} \right) x^{n+1-k} y^k.$$

Procedo estraendo dalla prima somma il primo termine e dalla seconda somma l'ultimo termine.

$$(x+y)^{n+1} = x^{n+1} + \sum_{k=1}^{n} \left( \begin{array}{c} n \\ k \end{array} \right) x^{n+1-k} y^k + \sum_{k=1}^{n} \left( \begin{array}{c} n \\ k-1 \end{array} \right) x^{n+1-k} y^k + y^{n+1},$$

e raccogliendo i termini nelle due sommatorie centrali

$$(x+y)^{n+1} = x^{n+1} + \sum_{k=1}^{n} \left[ \left( \begin{array}{c} n \\ k \end{array} \right) + \left( \begin{array}{c} n \\ k-1 \end{array} \right) \right] x^{n+1-k} y^k + y^{n+1},$$

Dato che

$$\left( \begin{array}{c} n \\ k \end{array} \right) + \left( \begin{array}{c} n \\ k-1 \end{array} \right) = \frac{n!}{k!(n-k)!} + \frac{n!}{(k-1)!(n-k+1)!} =$$

$$= \frac{n!\big[(n-k+1)+k\big]}{k!(n+1-k)!} = \frac{n!(n+1)}{k!(n+1-k)!} = \frac{(n+1)!}{k!(n+1-k)!} = \left( \begin{array}{c} n+1 \\ k \end{array} \right),$$

ottengo

$$(x+y)^{n+1} = x^{n+1} + \sum_{k=1}^{n} \left( \begin{array}{c} n+1 \\ k \end{array} \right) x^{n+1-k} y^k + y^{n+1},$$

quindi, usando che $\left( \begin{array}{c} n+1 \\ 0 \end{array} \right) = \left( \begin{array}{c} n+1 \\ n+1 \end{array} \right) = 1$,

$$(x+y)^{n+1} = \left( \begin{array}{c} n+1 \\ 0 \end{array} \right) x^{n+1} + \sum_{k=1}^{n} \left( \begin{array}{c} n+1 \\ k \end{array} \right) x^{n+1-k} y^k + \left( \begin{array}{c} n+1 \\ n+1 \end{array} \right) y^{n+1} =$$

$$= \sum_{k=0}^{n+1} \left( \begin{array}{c} n+1 \\ k \end{array} \right) x^{n+1-k} y^k.$$

La dimostrazione è conclusa. Convinti?

CAPITOLO 2

# Funzioni: anno zero

## 1. Ingredienti di base

Praticamente 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 $a, b, c, \ldots$, dette *variabili dipendenti*, sono determinate da altre quantità $x, y, z, \ldots$, dette *variabili indipendenti*, si dice che $a, b, c, \ldots$ "sono funzioni di" $x, y, z, \ldots$ o, in modo equivalente che $a, b, c, \ldots$ "dipendono da" $x, y, z, \ldots$. 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 $\ell$ è data da $A = \ell^2$, quindi la variabile dipendente *area* $A$ è funzione della variabile indipendente *lato* $\ell$.

(ii) Quanto è la lunghezza $\ell$ dell'ipotenusa di un triangolo rettangolo sapendo che i cateti sono lunghi rispettivamente $a$ e $b$? Mettendo alle strette Pitagora si arriva alla famosa risposta: *la lunghezza dell'ipotenusa è pari alla radice quadrata della somma dei quadrati delle lunghezze dei cateti*, o, in formule, $\ell = \sqrt{a^2 + b^2}$. Qui la variabile dipendente è $\ell$, mentre le variabili indipendenti sono due: $a$ 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 $P$.

Altre miriadi di esempi si possono dare assegnando delle regole per associare ad un numero dato un altro numero, senza bisogno che le variabili abbiano un significato concreto preciso (area, lunghezza, ...): numeri e basta.

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`.

$$\texttt{Input} \longrightarrow \boxed{\text{FUNZIONE}} \longrightarrow \texttt{Output}$$

*In queste Note ci limitiamo al caso di funzioni che associano ad un numero reale un altro numero reale*, come nel caso dell'esempio (i). Si usa comunemente 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 $f$

$$f(x) = x^2 \qquad \text{per} \quad 0 < x < 2$$

non coincide con la funzione $g$ data da

$$g(x) = x^2 \qquad \text{per} \quad x \in \mathbb{R},$$

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 $f$.

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

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

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

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 $I = \mathbb{R}$ (le operazioni richieste hanno sempre senso), mentre la funzione $f(x) = \frac{1}{x}$ è definita in $I = \mathbb{R} \setminus \{0\}$ (la divisione per 0 in $\mathbb{R}$ non ha senso).

**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 (Fig.1(a)).

(ii) $y$ è inversamente proporzionale a $x$ (vedi Fig.1(b)),

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

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

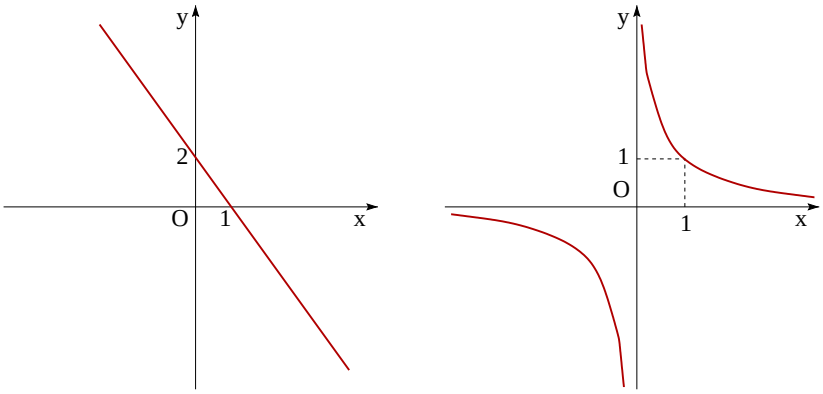



FIGURA 1. (a) La retta $y = -2x + 2$; (b) L'iperbole $y = 1/x$.

(iii) $y$ è il quadrato di $x$,

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

come è ben noto questa funzione ha per grafico una parabola (vedi Fig.2(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: la parte della bisettrice del primo e terzo quadrante nel semipiano $\{x \geq 0\}$ e la parte della bisettrice del secondo e quarto quadrante in $\{x < 0\}$ (vedi Fig.2(b)).

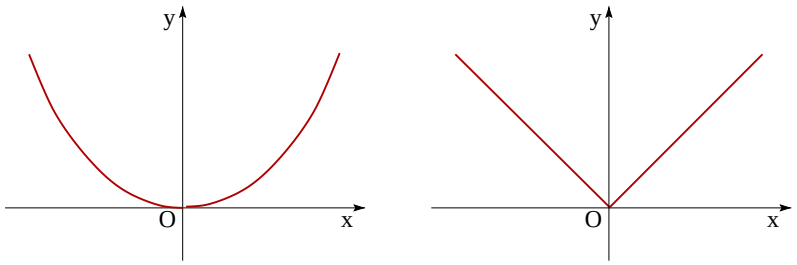



FIGURA 2. (a) La parabola $y = x^2$; (b) Il modulo: $y = |x|$.

**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.

Il polinomio più semplice è la *funzione affine*

$$y = ax + b,$$

il cui grafico è rappresentato da una retta. I valori $a$ e $b$ hanno un significato geometrico ben preciso. Siano $(x_1, y_1)$ e $(x_2, y_2)$ due punti del grafico di $f$, allora $y_1 = ax_1 + b$ e $y_2 = ax_2 + b$. Sottraendo le due equazioni membro a membro, si ottiene

$$y_1 - y_2 = a(x_1 - x_2) \quad \Rightarrow \quad a = \frac{y_1 - y_2}{x_1 - x_2} \quad \text{coeff. angolare.}$$

Quindi la costante $a$, detta COEFFICIENTE ANGOLARE, indica la pendenza del grafico. Per quanto riguarda $b$, detto TERMINE NOTO, individua l'intersezione della retta con l'asse delle $y$, e, in definitiva, individua l'"altezza" a cui giace la retta. Bisogna assolutamente acquisire familiarità con coefficiente angolare e termine noto. L'alternativa è emigrare verso qualche altra disciplina.

I polinomi di secondo grado sono funzioni della forma

$$y = ax^2 + bx + c \qquad a \neq 0,$$

ed hanno per grafico quello di una parabola. Per quanto riguardo il problema dell'esistenza degli zeri (ossia delle intersezioni del grafico della funzione con l'asse delle $x$), reminiscenze

liceali dovrebbero suggerire che ci sono tre casi possibili: due zeri, uno zero, nessuno zero. Più precisamente poniamo $\Delta = b^2 - 4ac$, allora

$$ax^2 + bx + c = 0 \quad \text{ammette} \quad \left\{ \begin{array}{c} \text{due soluzioni} \\ \text{una soluzione} \\ \text{nessuna soluzione} \end{array} \right\} \quad \text{se} \quad \left\{ \begin{array}{l} \Delta > 0 \\ \Delta = 0 \\ \Delta < 0 \end{array} \right.$$

Il valore $a$ incide sull'apertura della parabola (se $a > 0$ la concavità è verso l'alto, se $a < 0$ la concavità è verso il basso, nel primo caso la funzione si dice *convessa*, nel secondo caso *concava*). Più $|a|$ è grande e più la parabola ha una crescita "rapida". La retta $x = -b/2a$ è l'asse (di simmetria) della parabola. Con queste informazioni alla mano è possibile disegnare il grafico della parabola in modo approssimativo, ma soddisfacente.

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 di grado maggiore di 2. 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.

L'esempio più semplice di funzione razionale che non sia un polinomio è dato dalla funzione $y = 1/x$ che, come si è già visto, ha per grafico un'iperbole.

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 $x \geq -3/2$ e per $x < -1$ e negativa nel resto dell'insieme. Il grafico è in Figura 3(a).

Un altro esempio di funzione razionale facile è $f(x) = 1/x^2$, disegnata in Figura 3(b).

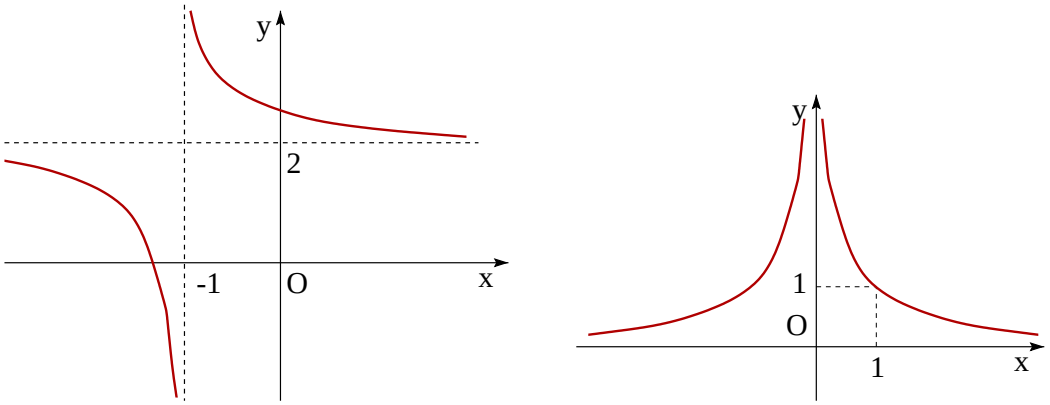



FIGURA 3. (a) La funzione $y = (2x+3)/(x+1)$; (b) la funzione $y = 1/x^2$.

**Funzioni trigonometriche.** Non è possibile in poche righe ricordare tutto il necessario sulle funzioni trigonometriche. Qui ci limitiamo alle proprietà principali, ma il lettore non creda che sia sufficiente! 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

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

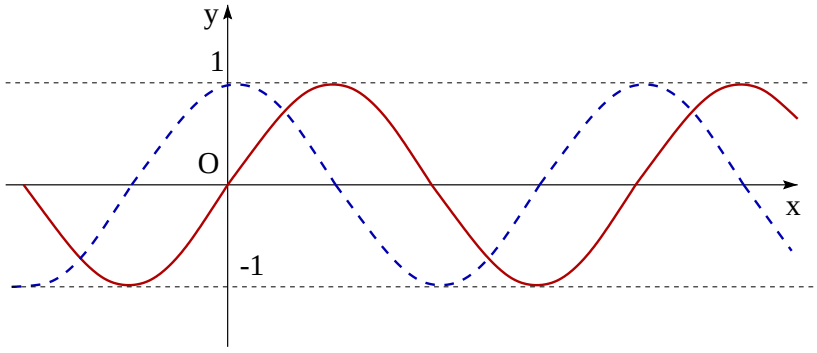



FIGURA 4. Il grafico della funzione $\sin x$ (linea continua) e della funzione $\cos x$ (linea tratteggiata).

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

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

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

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

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

**Soluzione.** Poniamo $\xi = \frac{x+y}{2}$ e $\eta = \frac{x-y}{2}$. Allora $x = \xi + \eta$ e $y = \xi - \eta$. Dunque
$\sin x - \sin y = \sin(\xi+\eta) - \sin(\xi-\eta) = \sin\xi\cos\eta + \cos\xi\sin\eta - \sin\xi\cos\eta + \cos\xi\sin\eta = 2\cos\xi\sin\eta$,
e ricordando la definizione di $\xi$ ed $\eta$ 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 $x \neq \frac{\pi}{2} + k\pi$ per $k \in \mathbb{Z}$ e $\cot x$ è definita per $x \neq k\pi$ per $k \in \mathbb{Z}$.

## 2. Operazioni elementari su grafici

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

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

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

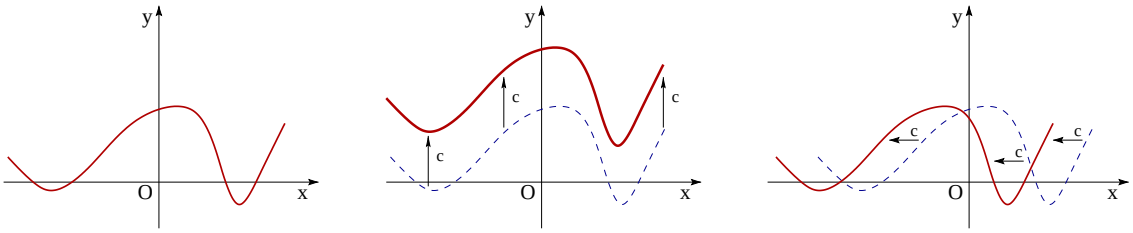



FIGURA 5. I grafici di (a) $y = f(x)$, (b) $y = f(x) + c$; e (c) $y = f(x + c)$.

(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 $m(x) = \frac{1}{2}\big||x| - 1\big|$.

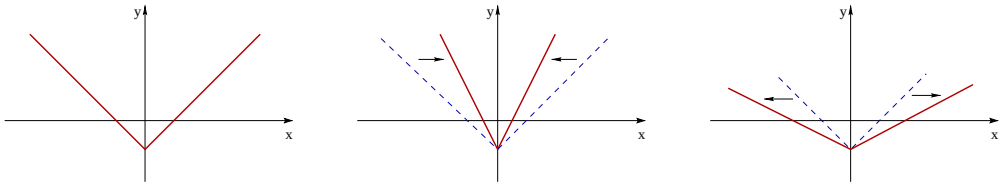



FIGURA 6. I grafici di (a) $y = f(x) = |x| - 1$, (b) $g(x) = |2x| - 1$, (c) $h(x) = \left|\frac{x}{2}\right| - 1$.

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 $\pm 1$ 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.

(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

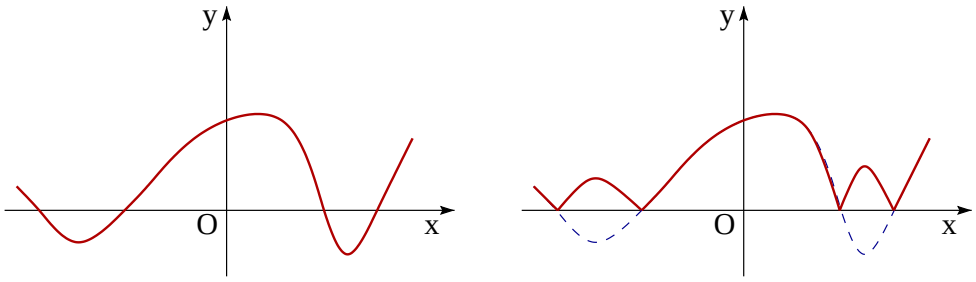



FIGURA 7. I grafici di (a) $y = f(x)$, (b) $y = |f(x)|$.

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 $y \geq 0$.

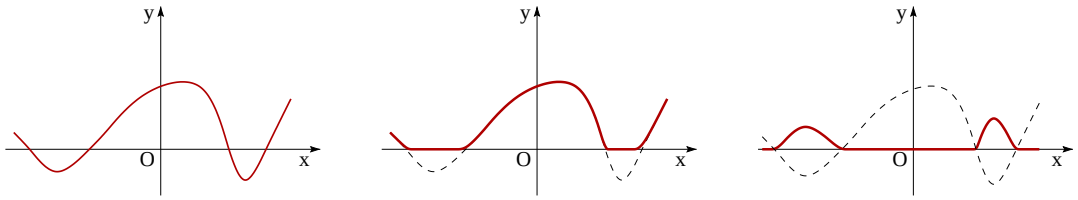



FIGURA 8. I grafici di (a) $y = f(x)$, (b) $y = \max\{f(x), 0\}$, (c) $y = \max\{-f(x), 0\}$,.

**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?

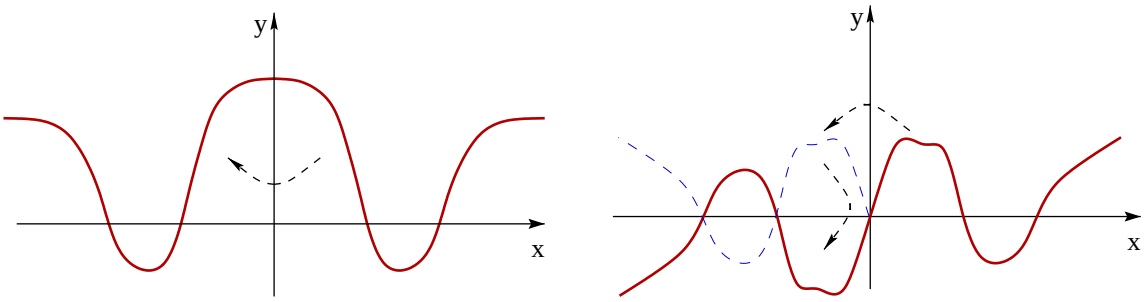



FIGURA 9. (a) una funzione pari, (b) una funzione dispari.

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

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

Il più piccolo valore $T$ per cui vale questa proprietà è il PERIODO[2] DI $f$ 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 $\pi$, infatti

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

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

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

$$\text{PARTE INTERA di } x: \qquad [x] := z \in \mathbb{Z},$$

[2]Avete notato che il periodo è il minimo di un insieme di numeri reali? Ma il minimo non sempre esiste... In effetti sull'esistenza del periodo ci si potrebbe dilungare, ma non lo faremo.

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

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 $y \in f(I)$. Ma quante sono?

**Iniettività e funzioni inverse.** Problema: sia $f$ una funzione assegnata,

*dato $y \in f(I)$, 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.

Mettiamola in altri termini: sia $I$ l'insieme degli esseri umani ed $f$ la funzione che associa ad ogni persona la sua data di nascita. Si tratta di una funzione iniettiva? Ovvero, scelta una data di nascita c'è un'unica persona nata quel giorno? Assolutamente no! Invece la funzione che associa ad ogni macchina la sua targa è iniettiva: macchine diverse hanno targhe diverse.

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.

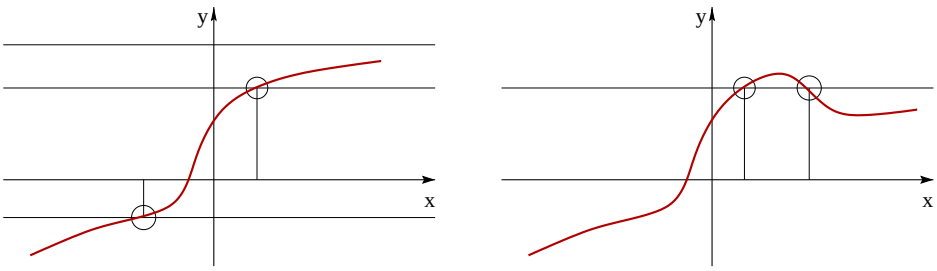



FIGURA 10. (a) una funzione iniettiva, (b) una funzione non iniettiva.

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 $x_1 = x_2$:

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

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

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

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

$$\frac{x_1+1}{x_1+2} = \frac{x_2+1}{x_2+2} \quad \Longleftrightarrow \quad (x_2+2)(x_1+1) = (x_2+1)(x_1+2)$$
$$\Longleftrightarrow \quad 2x_1 + x_2 = 2x_2 + x_1 \quad \Longleftrightarrow \quad x_1 = x_2.$$

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:

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

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 unico valore $x$ dato proprio dalla soluzione dell'equazione $f(x) = y$. In altre parole, se una funzione è iniettiva è possibile definire una nuova funzione che permette di "tornare indietro", cioè che associa ad ogni $y$ dell'insieme immagine, l'unico valore $x$ di cui è immagine.

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

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

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

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

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

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

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

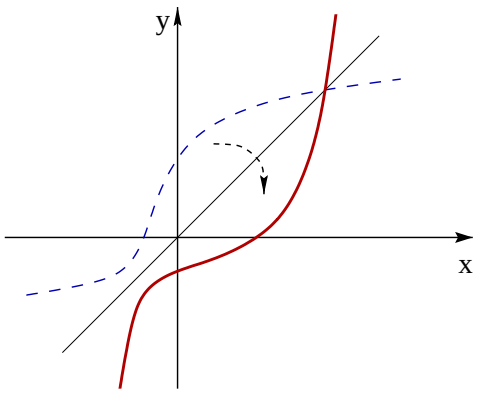



FIGURA 11. Il grafico della funzione inversa di una funzione assegnata.

**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 $x, x' \in I$,

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

la funzione è MONOTÒNA STRETTAMENTE DECRESCENTE o semplicemente STRETTAMENTE DECRESCENTE. Equivalentemente, si può scrivere

$$\begin{aligned} f \text{ strettamente crescente} \quad &\Longleftrightarrow \quad \big[f(x) - f(x')\big](x - x') > 0 \qquad \forall x \neq x' \\ f \text{ strettamente decrescente} \quad &\Longleftrightarrow \quad \big[f(x) - f(x')\big](x - x') < 0 \qquad \forall x \neq x' \end{aligned}$$

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 $\mathbb{R}$. Anche la funzione $|x|$ non è monotòna su $\mathbb{R}$.

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

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

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. Ecco un esempio interessante di quel che si accennava nella Sezione precedente: nonostante sia chiaro che la funzione $f$ è invertibile e, quindi, che esiste la funzione $f^{-1}$, è assolutamente fuori discussione poter determinare l'espressione esplicita per $f^{-1}$: vorrebbe dire risolvere rispetto ad $x$ l'equazione $y = x + x^n$...

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:

$$\begin{aligned} f \text{ non decrescente} &\iff \big[f(x) - f(x')\big](x - x') \geq 0 \qquad \forall x, x' \\ f \text{ non crescente} &\iff \big[f(x) - f(x')\big](x - x') \leq 0 \qquad \forall x, x' \end{aligned}$$

Attenzione, 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. Se però si considerano delle restrizioni opportune di queste funzioni, si ottengono delle funzioni monotone...

**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.

*Un conto da saltimbanco.* Per invertire la funzione $\cot x$ in $(0,\pi)$ è sufficiente la funzione arctan e un po' di dimestichezza con il suo uso. Consideriamo prima di tutto $x \in (0,\frac{\pi}{2})$, allora

$$y = \cot x = \frac{1}{\tan x} \quad \Longleftrightarrow \quad \frac{1}{y} = \tan x \quad \Longleftrightarrow \quad x = \arctan\left(\frac{1}{y}\right),$$

che dà l'espressione dell'inversa di $\cot x$ in $(0,\frac{\pi}{2})$. Se $x \in (\frac{\pi}{2},\pi)$,

$$y = \cot x = \frac{1}{\tan x} \quad \Longleftrightarrow \quad \frac{1}{y} = \tan x = \tan(x-\pi) \quad \Longleftrightarrow \quad x = \pi + \arctan\left(\frac{1}{y}\right).$$

Infine $\cot\left(\frac{\pi}{2}\right) = 0$, quindi l'inversa di $\cot x$ in $(0,\pi)$ è

$$\cot^{-1} x = \begin{cases} \pi + \arctan(1/x) & x \in (-\infty, 0), \\ \pi/2 & x = 0, \\ \arctan(1/x) & x \in (0, +\infty). \end{cases}$$

OSSERVAZIONE 4.1. Sia per le potenze pari che per le funzioni trigonometriche si usa lo stesso "trucco" per definire le funzioni inverse. Pensiamo, per semplicità, al caso della funzione $x^2$. Questa, se considerata in tutto $\mathbb{R}$, è una funzione non iniettiva e dunque non invertibile. Se invece consideriamo la restrizione di $x^2$ al sottoinsieme $I =$

$[0, +\infty)$, questa è strettamente monotona e dunque invertibile. Lo stesso procedimento viene utilizzato anche in altre situazioni: se una funzione non è invertibile, se ne considera una restrizione opportuna che è strettamente monotona e si considera l'inversa di questa "nuova" funzione.

> *Perché questo procedimento è utile? Perché le funzioni che stiamo considerando hanno delle proprietà di simmetria! Ad esempio, la funzione $x^2$ è pari, quindi se conosciamo ciò che succede per $x \geq 0$, basta ribaltare attorno alla retta $x = 0$ per ottenere anche quello che succede per $x < 0$, cioè, in definitiva, cambiare di segno. Non ci credete? Proviamo a cercare le soluzioni dell'equazione $x^2 = 4$. Dato che la funzione $x^2$ non è invertibile, non possiamo applicare nessuna funzione a questa equazione, in modo da ottenere una espressione del tipo $x = g(4)$. Non scoraggiamoci. Riduciamoci a cercare solo le soluzioni positive, cioè restringiamo la nostra attenzione al caso $x \geq 0$. Qui si che la funzione $x^2$ è invertibile! Quindi le soluzioni di $x^2 = 4$ con $x \geq 0$, si possono ottenere applicando la radice quadrata: per $x \geq 0$, $x = \sqrt{x^2} = \sqrt{4} = 2$. Bene. E per le $x < 0$? Semplice basta cambiare di segno... quindi le soluzioni di $x^2 = 4$ sono $x = +2$ e $x = -2$. Provate a ripetere il ragionamento per cercare le soluzioni di $x^4 = 81$ e di $\sin x = 1/2$.*

**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. Qui non ci curiamo di trasformare questa costruzione (solo vagamente accennata) in una definizione rigorosa e prendiamo per buono che abbia senso.

Per le funzioni esponenziali valgono le proprietá

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

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

$(i)$ $\log_a x$ è definito per $x > 0$, $(ii)$ $\log_a 1 = 0$,

$(iii)$ $\log_a x + \log_a y = \log_a(xy)$, $(iv)$ $\log_a(x^\alpha) = \alpha \log_a x$

$(v)$ $\log_a x$ è $\begin{cases} \text{decrescente} & \text{se } 0 < a < 1 \\ \text{crescente} & \text{se } a > 1 \end{cases}$

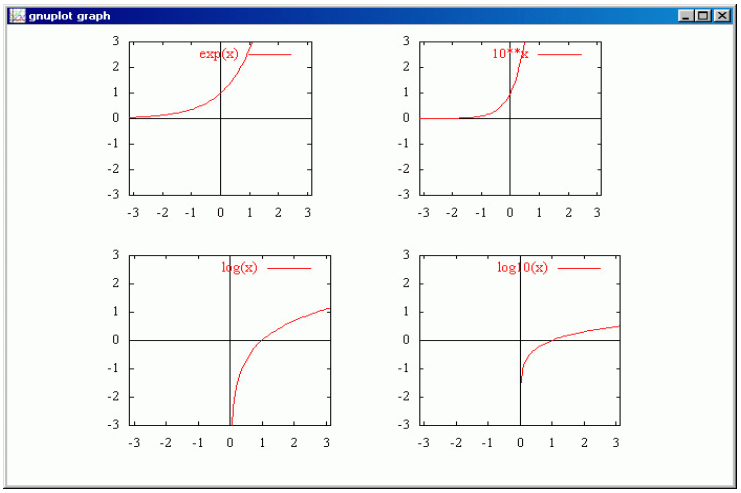



FIGURA 12. (a) $y = e^x$, (b) $y = 10^x$, (c) $y = \ln x$, (d) $y = \log_{10} x$

OSSERVAZIONE 4.2. 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 $\phi$" o anche "$g$ dopo $\phi$")[3].

[3]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 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 $\phi$ 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 $\phi$ è iniettiva su $\mathbb{R}$ e la sua funzione inversa $\phi^{-1}$ è anch'essa definita su tutto $\mathbb{R}$, *$\phi^{-1}$ 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 $J = (0, +\infty)$. 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 $\phi$ e $g$ e definire una nuova funzione $g \circ \phi$, il dominio di $g$ deve includere almeno un parte dell'immagine di $\phi$. 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 $\phi$ 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. (i)** Si possono assegnare funzioni definite in maniera più originale: ad esempio, (vedi Fig.13)

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

**(ii)** E' interessante anche la FUNZIONE CARATTERISTICA di un insieme assegnato. Sia $I \subset \mathbb{R}$, allora si definisce

$$\text{FUNZIONE CARATTERISTICA DI } I\text{:} \qquad \chi_I(x) = \begin{cases} 1 & \text{se } x \in I, \\ 0 & \text{se } x \notin I. \end{cases}$$

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.

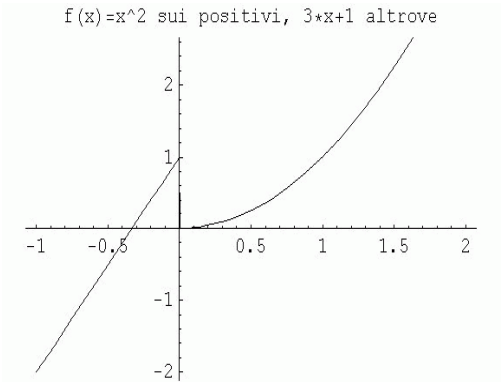



FIGURA 13. Il grafico di $y = f(x)$.

**(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 $g$,

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

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

## 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 $f(x) \ge f(x_0)$ 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 $f(x) \le f(x_1)$ 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* $\inf_{x\in I} f(x) \in \mathbb{R}$ *e* $\sup_{x\in I} f(x) \in \mathbb{R}$ *(quindi non sono* $\pm\infty$*) si dice* LIMITATA.

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

**Un problema escatologico: la fine dei tempi.** La dimensione $D$ dell'universo è funzione del tempo $t$: $D = D(t)$. E' noto che l'universo (a partire dall'istante $t = 0$ del *Big Bang*) si sta espandendo, quindi il grafico della funzione $D$ è quello di una funzione crescente. Uno dei problemi interessanti trattati dagli astrofisici è quello di prevedere quale sarà l'evoluzione di $D$ nel futuro. In particolare ci sono tre situazioni plausibili: l'universo si espanderà per sempre in modo illimitato, l'universo si espanderà sempre avvicinandosi

(asintoticamente) ad una soglia critica, l'universo raggiungerà un giorno un punto di massima espansione e da quel momento in poi tenderà a contrarsi fino al cosiddetto *Big Crunch*. In termini di estremo superiore e massimo le tre situazioni corrispondono a tre casi diversi: rispettivamente, l'estremo superiore è $+\infty$, l'estremo superiore è finito, ma non si tratta di massimo, la funzione $D$ ammette massimo.

**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 totale.

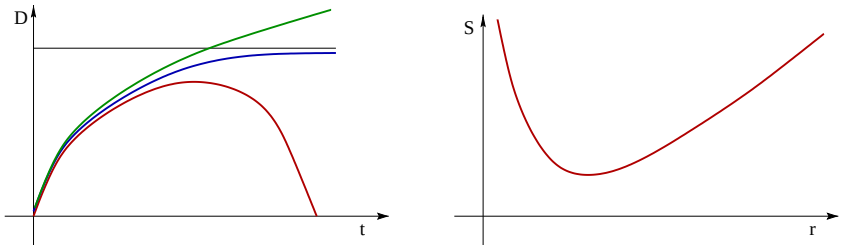



FIGURA 14. (a) Tre possibili evoluzioni per l'universo; (b) il grafico di $S(r) = 2\pi\left(r^2 + \frac{V}{\pi r}\right)$.

Allora $S = 2\pi r^2 + 2\pi hr = 2\pi\left(r^2 + hr\right)$ 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

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

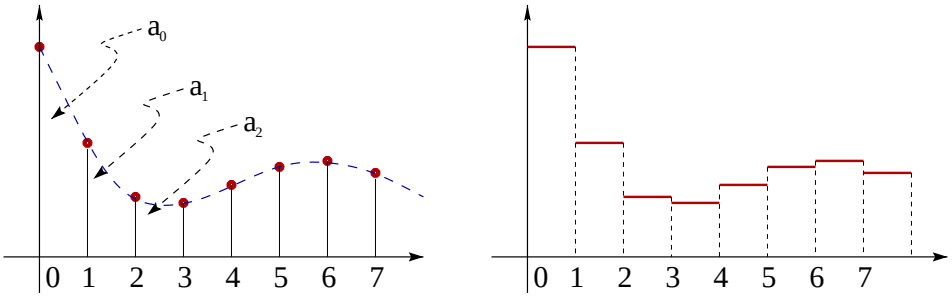



FIGURA 1.

**Due parole sulle 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 $a_0$.

Un altro esempio di successione definita per ricorrenza è l'ALGORITMO DI ERONE, che permette di ottenere approssimazioni della radice quadrata di un assegnato numero $x > 0$. L'algoritmo consiste nello scegliere un qualsiasi valore iniziale $a_1 = \alpha$ e poi costruire la successione

$$a_{n+1} = \frac{1}{2}\left(a_n + \frac{x}{a_n}\right).$$

Ad esempio, cerchiamo la radice di $x = 2$ e partiamo da $a_0 = 1$, allora i primi termini della successione sono $a_0 = 1$, $a_1 = 1,5$, $a_2 = 1,4167$. In una forma più generale questo algoritmo è noto come *metodo di Newton* (o delle tangenti).

## 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 $\ell$ fissato? In caso affermativo, si dice che la successione ammette limite $\ell$ 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 $1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$. 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 $n_\varepsilon$ in poi ($n_\varepsilon$ dipende da $\varepsilon$) la distanza di $a_n$ da 0 (che è data da $|a_n - 0|$) è minore di $\varepsilon$, cioè

$$\forall\, \varepsilon > 0 \quad \exists n_\varepsilon \in \mathbb{N} \quad \text{t.c.} \quad |a_n - 0| < \varepsilon \quad \forall\, n > n_\varepsilon. \tag{1.1}$$

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 (1.1).

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:

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

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 $n \to +\infty$ *se*

$$\forall\, \varepsilon > 0 \quad \exists\, n_\varepsilon \in \mathbb{N} \quad t.c. \qquad |a_n - \ell| < \varepsilon \quad \forall\, n > n_\varepsilon. \tag{1.2}$$

*Si scrive* $\lim_{n\to+\infty} a_n = \ell$ *o anche* $a_n \to \ell$ *per* $n \to \infty$. *Il valore $\ell$ è il* LIMITE DELLA SUCCESSIONE *$a_n$.*

*Se $a_n \to 0$ per $n \to +\infty$ (cioè se vale (1.1)), 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 $\ell$ meno di $\varepsilon$, tranne al più un numero finito (quelli con indice da 1 ad $n_\varepsilon$). Quindi una maniera equivalente di dire che $\ell$ è il limite di $a_n$ è affermare che *ogni intorno di $\ell$ contiene tutti i valori della successione $a_n$ tranne al più un numero finito.*

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

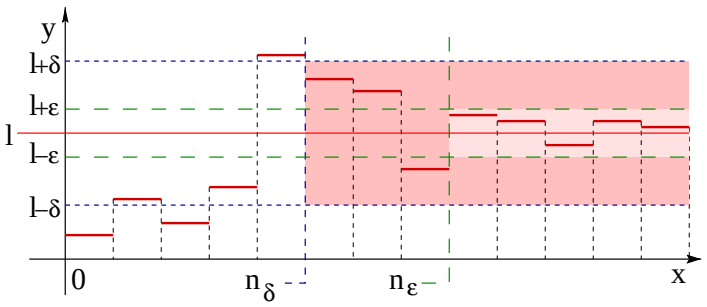



FIGURA 2.

*Per quale motivo si richiede che $\varepsilon$ possa essere scelto arbitrariamente? Non basterebbe scegliere un fissato $\varepsilon$ sufficientemente piccolo, ad esempio $\varepsilon$ 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 $\varepsilon$ rende la definizione "universale", cioè indipendente dalla personale idea di piccolo o grande.*

ESERCIZIO 1.2. *Sia $a_n$ una successione convergente ad $\ell$ 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 $\ell$ 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

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

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

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

**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 $\ell$ 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 e proviamo a dimostrarlo. Qual'è l'affermazione racchiusa nella definizione di limite? la distanza di $a_n$ da $\ell$ 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 $\varepsilon$ se scegliamo $n > n_\varepsilon$ dove *abbiamo completa libertà di scelta per* $n_\varepsilon$. 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 1.3. *Dimostrare a partire dalla definizione la validità 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 1.4. *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 $a_5, a_3, a_{12}, a_{101}, \ldots$ 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 1.5. *Dire quale delle seguenti espressioni possono essere i primi elementi di una sottosuccessione di* $a_n = n$

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

ESERCIZIO 1.6. *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 1.7. *Sia* $a_n$ *una successione convergente ad* $\ell$ *per* $n \to +\infty$. *Allora ogni sua sottosuccessione* $a_{n_k}$ *converge allo stesso limite* $\ell$ *per* $k \to +\infty$.

La dimostrazione è lasciata per esercizio...

La Proposizione 1.7 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 1.7 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 estremamente più semplice di come si è fatto finora, dimostriamo alcune proprietà generali delle successioni convergenti.

PROPOSIZIONE 1.8. *Se una successione è convergente, allora il suo limite è unico.*

**Dimostrazione.** Mostriamo che se la successione $a_n$ tende sia ad $\ell$ che ad $\ell'$, 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

$$0 \le |\ell - \ell'| < 2\varepsilon \qquad \forall\, \varepsilon > 0,$$

quindi $|\ell - \ell'| = 0$, cioè $\ell = \ell'$. ∎

DEFINIZIONE 1.9. *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 1.9 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 1.10. *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 $a_{N+1}, a_{N+2}, \ldots$ 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). ∎

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 1.9).

**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 1.11. *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*

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

*In modo equivalente si può scrivere*

$$\forall M \quad \exists n_M \in \mathbb{N} \quad t.c. \quad a_n \geq M \quad (risp.\ a_n \leq M) \quad \forall n > n_M. \tag{1.3}$$

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 1.11 (così come $n_\varepsilon$ della Definizione 1.1), non è definito in maniera univoca, tutt'altro!

ESERCIZIO 1.12. *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 $\ell$, 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 $\ell$ 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 1.13. *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 $\ell$ e quindi si avvicinano tra loro. La proprietà notevole è che vale anche il viceversa: se gli elementi della successione si avvicinano, allora la successione converge. Non è questa la sede per approfondire ulteriormente questo criterio, ma non si può mancare di dire che si tratta di una pietra miliare nella costruzione dei numeri reali a partire dai numeri razionali.

*Piccolo glossario per le successioni*

Se una successione $a_n \dots$

$\dots$ è convergente o divergente (a $\pm\infty$), allora è REGOLARE;

$\dots$ non è convergente né divergente, allora è NON REGOLARE;

$\dots$ tende a 0 per $n \to +\infty$, allora è INFINITESIMA;

$\dots$ 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;

$\dots$ 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.*

A mo' d'esempio, dimostriamo la proprietà del prodotto (in maniera analoga, il lettore desideroso di imparare può (provare a) dimostrare le altre proprietà). Supponiamo $a_n \to a$ e $b_n \to b$ per $n \to +\infty$. Allora[2]

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

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 $\varepsilon$ 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:

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

dato che nella seconda espressione è possibile passare al limite sia nel numeratore che nel denominatore.

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

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

Resta fuori un caso particolarmente importante: se

$$\lim_{n\to+\infty} b_n = 0$$

è possibile dire qualche cosa sul comportamento al limite della successione $\dfrac{a_n}{b_n}$? Torneremo tra breve sulla questione.

[2] Il valore di $n_\varepsilon$ 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 $b_n$.

**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* $\lim_{n\to+\infty} a_n = a$ *e* $\lim_{n\to+\infty} b_n = b$. *Se, per qualche* $N \in \mathbb{N}$, *vale* $a_n < b_n$ *(oppure* $a_n \le b_n$*) per ogni* $n \ge N$, *allora* $a \le b$.

**Dimostrazione.** Dato che $a_n$ tende a $a$ e $b_n$ tende a $b$, si ha che per ogni $\varepsilon > 0$, $a - \varepsilon < a_n \le b_n < b + \varepsilon$ 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, $b - a \ge 0$. ∎

Il Teorema 2.2 afferma che l'operazione di limite è "monotòna non decrescente", nel senso che vale l'implicazione

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

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

In maniera analoga, si dimostra il seguente risultato utilissimo!

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

*(i) esiste* $N \in \mathbb{N}$ *tale che* $a_n \le b_n \le c_n$ *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 di 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 $a_n \le b_n \le c_n$ per ogni $n > N$, si ottiene

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

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 $(1+h)^n \ge 1 + nh$ per ogni $n \in \mathbb{N}$ (dimostrarlo!),

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

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*

$$\lim_{n\to+\infty} \frac{a_{n+1}}{a_n} = \lambda \tag{2.1}$$

*Allora, se $\lambda \in [0,1)$, la successione $a_n$ è infinitesima, cioè $\lim_{n\to+\infty} a_n = 0$.*

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

$$a_{n+1} \le \sigma a_n \qquad \forall\, n \ge N. \tag{2.2}$$

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 (2.2) scegliendo $N = n_\varepsilon$.

*Passo 2.* Dimostriamo che $a_n$ è infinitesima. Per semplicità, supponiamo $N = 0$ (cioè che (2.2) valga per ogni $n$), allora

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

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

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

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 $(1+h)^n \ge 1 + nh$, segue

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

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 $\lim_{n\to+\infty} b_n = 0$?

– 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*

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

*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 $\infty$, 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:

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

(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:

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

Il prodotto di successioni divergenti non crea nessun problema particolare (bisogna solo stare attenti al segno di $\infty$, 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:

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

## 3. Calcolo di alcuni limiti

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 $\lim_{n\to+\infty} h_n = 0$ 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 $\pm 1$ 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.

$$\lim_{n\to+\infty} \frac{x^n}{n!} = 0 \qquad \forall\, x > 1. \tag{3.1}$$

Applichiamo di nuovo il Corollario 2.6

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

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

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

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. *Successioni monotòne.* *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.*

**Soluzione.** Bisogna verificare se vale la disequazione $a_{n+1} < a_n$ per ogni $n$:

$$\frac{n+1}{1+(n+1)^2} < \frac{n}{1+n^2} \quad \Longleftrightarrow \quad (1+n)(1+n^2) < n[1+(1+n)^2]$$
$$\Longleftrightarrow \quad 1+n+n^2+n^3 < 2n+2n^2+n^3 \quad \Longleftrightarrow \quad 1 < n+n^2.$$

La successione è strettamente decrescente per $n \geq 1$. □

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. Ecco il perché.

TEOREMA 4.4. *Regolarità delle successioni monotòne. 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 $\ell$ 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 $\ell$ (è un maggiorante). Inoltre, dato che $\ell$ 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, $\ell$ 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. Abbiate pietà. ∎

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

Passiamo ora alle serie numeriche. L'intenzione qui è di dare solo un'infarinatura di base sull'argomento, avendo come obiettivo principale la presentazione di un'altra serie fondamentale (oltre alla serie geometrica di cui già si è detto): la *serie esponenziale*.

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 (che abbiamo già utilizzato nell'Esempio 3.7) è di considerare la $S_n$ delle SOMME PARZIALI definita da

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

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*:

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

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. In altre parole, il fatto che il termine generico $a_k$ sia infinitesimo è una condizione necessaria, ma non sufficiente.

**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 $\sum_{k=0}^{\infty} a_k$ 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: bisogna stabilire se il decadimento a zero di $x^k/k!$ è sufficientemente rapido. 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 (3.1),

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

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!)

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

Per dimostrarla, introduciamo due nuovi oggetti. Dato $a \in \mathbb{R}$, la PARTE POSITIVA $a^+$ e la PARTE NEGATIVA $a^-$ di $a$ sono definite da

$$a^+ = \max(a, 0) \quad , \quad a^- = \max(-a, 0).$$

Quindi, se $a = 2$, $a^+ = 2$ e $a^- = 0$, mentre se $a = -3$, $a^+ = 0$ e $a^- = 3$. Dalla definizione discendono le proprietà seguenti:

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

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

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

Che collegamento c'è tra la convergenza assoluta e la convergenza della serie delle parti positive e delle parti negative? Semplice... si tratta della stessa cosa!

PROPOSIZIONE 5.9. *Una serie di termine generico* $a_k$ *converge assolutamente se e solo se convergono (semplicemente) le serie di termine generico* $a_k^+$ *e* $a_k^-$.

**Dimostrazione.** Supponiamo che la serie $\sum a_k$ sia assolutamente convergente. Allora è possibile 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... ∎

Grazie a questa caratterizzazione della convergenza assoluta, dimostrare che se una serie converge assolutamente, allora converge semplicemente è un gioco da ragazzi...

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.** Anche questa è una conseguenza immediata della linerità! 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. ∎

Se una serie converge assolutamente, allora vale la seguente versione (per somme infinite) della disuguaglianza triangolare:

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

Non sarò certo io a togliervi il gusto di dimostrarla da soli.

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

$$\text{(5.1)} \qquad \text{FUNZIONE ESPONENZIALE:} \qquad e^x := \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad x \in \mathbb{R}.$$

Questa posizione è del tutto rigorosa dato che la serie è sempre convergente. Bisognerebbe verificare che questa definizione sia compatibile con la costruzione naïf e con le proprietà viste in precedenza per le funzioni esponenziali. Tutto ciò ci porterebbe lontano dall'obiettivo del corso e quindi lasciamo perdere (sperando nella clemenza del lettore!)

**Approssimazione del numero di Nepero.** A partire dalla formula

$$e = \sum_{n=0}^{\infty} \frac{1}{n!},$$

è possibile ottenere approssimazioni del valore di $e$. Prima di tutto stimiamo l'errore che si commette sostituendo il numero $e$ con la somma parziale $s_n = \sum_{k=0}^{n} \frac{1}{k!}$, cioè la differenza $e - s_n$

$$\begin{aligned} 0 < e - s_n &= \frac{1}{(n+1)!} + \frac{1}{(n+2)!} + \cdots = \frac{1}{(n+1)!}\left(1 + \frac{1}{n+2} + \frac{1}{(n+2)(n+3)} + \cdots\right) \\ &< \frac{1}{(n+1)!}\left(1 + \frac{1}{n+1} + \frac{1}{(n+1)^2} + \cdots\right) = \frac{1}{(n+1)!} \cdot \frac{1}{1 - \frac{1}{n+1}} = \frac{1}{n!\,n}, \end{aligned}$$

quindi vale

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

Da questa stima si può determinare il numero di elementi che occorre sommare per ottenere una approssimazione di $e$ con errore prescritto. Ad esempio, supponiamo di ammettere un errore massimo di $10^{-4}$. Dato che

| $n$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---|---|---|---|---|---|---|---|
| $n!\,n$ | 1 | 4 | 18 | 96 | 600 | 4320 | 35280 |
| $1/(n!\,n)$ | 1 | 0,25 | $0,0\overline{5}$ | 0,010412 | $0,001\overline{6}$ | 0,000231 | 0,000028 |

va bene scegliere $n = 7$. Quindi l'approssimazione richiesta è

$$e \approx 1 + 1 + \frac{1}{2} + \frac{1}{6} + \frac{1}{24} + \frac{1}{120} + \frac{1}{720} + \frac{1}{5040} = 2,718254.$$

## Appendice A. M'approssimo in un denso

**Rappresentazione decimale dei numeri reali.** I numeri reali possono essere rappresentati tramite una *successione di decimali*. L'idea è di associare ad ogni numero $x \in \mathbb{R}$, una successione numerica $x_n$ che tenda ad $x$ per $n \to +\infty$ e i cui numeri $x_n$ siano "facilmente" rappresentabili sull'asse reale. Vediamo in che maniera.

Fissato $x \in \mathbb{R}$, esiste $c_0 \in \mathbb{Z}$ per cui $c_0 \leq x < c_0 + 1$ (grazie al principio di Archimede):

$$\forall\, x \in \mathbb{R} \quad \exists\, c_0 \in \mathbb{Z} \quad \text{tale che} \quad x \in I_0 := [c_0, c_0 + 1).$$

Se dividiamo l'intervallo $[c_0, c_0 + 1)$ in dieci parti uguali tramite i punti $c_0 + \frac{1}{10}, c_0 + \frac{2}{10}, \ldots, c_0 + \frac{9}{10}$, c'è un intero $c_1$, compreso tra $0, 1, 2, \ldots, 9$ (nel seguito, diciamo che un intero $c$ compreso in $\{0, 1, \ldots, 9\}$ è una CIFRA) tale che $x$ appartiene all'intervallo $I_1$ dove

$$I_1 := \left[c_0 + \frac{c_1}{10},\ c_0 + \frac{c_1}{10} + \frac{1}{10}\right).$$

Quindi vale la seguente affermazione

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

Di nuovo, dividiamo $I_1$ in dieci parti e troviamo una cifra $c_2$ tale che $x \in I_2$ dove

$$I_2 := \left[c_0 + \frac{c_1}{10} + \frac{c_2}{100},\ c_0 + \frac{c_1}{10} + \frac{c_2}{100} + \frac{1}{100}\right).$$

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

$$I_n := \left[c_0 + \frac{c_1}{10} + \cdots + \frac{c_n}{10^n},\ c_0 + \frac{c_1}{10} + \cdots + \frac{c_n}{10^n} + \frac{1}{10^n}\right).$$

In modo equivalente possiamo dire

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

Per ogni $n$, sia $x_n$ l'estremo sinistro di $I_n$, cioè

$$x_n := c_0 + \frac{c_1}{10} + \cdots + \frac{c_n}{10^n} = \sum_{k=0}^{n} \frac{c_k}{10^k}.$$

L'errore che si commette approssimando il numero $x$ con $x_n$ è minore o uguale alla lunghezza dell'intervallo $I_n$, cioè $10^{-n}$. Quindi, per $n \to +\infty$, $x_n$ tende a $x$ e, pertanto, i numeri $c_0, c_1, c_2, \ldots$ individuano il valore $x$. In altre parole, un qualsiasi numero reale è descritto completamente con una successione infinita di interi $c_k$. Nella notazione decimale ordinaria, la corrispondenza tra $x$ e la successione $c_0, c_1, c_2, \ldots$ è indicata scrivendo

$$x = c_0, c_1 c_2 c_3 \cdots.$$

dove anche l'intero $c_0$ viene espresso in notazione decimale.

La definizione di $\{x_n\}$ corrisponde al percorso da fare per individuare il valore $x$: parti da $c_0$, fai $c_1$ passi a destra ciascuno di lunghezza $1/10$, poi $c_2$ passi a destra ciascuno di lunghezza $1/100$, e così via. La sequenza $c_0, c_1, \ldots, c_n, \ldots$ corrisponde (in qualche

senso) al "numero civico" dell'abitante $x$ di "via dei $\mathbb{R}$eali": ad ogni decimale in più che considera, un alacre postino determina la collocazione di $x$ con un errore più piccolo di un fattore dieci. Per determinare esattamente la collocazione del numero $x$, il poveretto dovrà considerare un numero infinito di cifre!

OSSERVAZIONE 5.12. Qui abbiamo scelto gli intervalli $I_n$ chiusi a sinistra ed aperti a destra, e abbiamo utilizzato come approssimante $x_n$ del numero $x$, l'estremo sinistro di $I_n$. E' possibile, però, fare scelte diverse e, quindi, costruire diverse rappresentazioni decimali dello stesso numero reale. Per esempio, il numero 1 è rappresentato sia da $0,9999999999\ldots$ che da $1,0000000000\ldots$.

OSSERVAZIONE 5.13. Nella rappresentazione decimale dei numeri reali, il ruolo speciale giocato dal numero dieci è puramente accidentale ed è legato ad un problema di dita (dieci nel caso degli esseri umani). L'unica ragione dell'ampio uso del sistema decimale è proprio dovuto alla facilità di contare fino a dieci con le nostre mani. Non per niente, in inglese "cifra" si dice "digit", che manifesta in modo più immediato l'origine "manuale" del conteggio.... Avremmo potuto utilizzare un qualsiasi altro intero $p > 1$ e procedere in modo analogo, ottenendo una rappresentazione diversa, ma altrettanto valida. Ad esempio, scegliendo $p = 3$, quale rappresentazione decimale si ottiene per $\frac{1}{2}$? Per la risposta, provate ad immaginarvi postino che è in grado di percorrere passi di lunghezza $1/3^n$. Segnate sulla retta il numero $1/2$. Partite da 0: il numero di passi di lunghezza $1 = 1/3^0$ che potete percorrere senza superare il numero $1/2$ è $c_0$. Partendo ora da $c_0$, quanti passi di lunghezza $1/3$ potete percorrere? Procedete in questo modo con passi da $1/3^2$, $1/3^3$ e così via dicendo...

**Densità di $\mathbb{Q}$ in $\mathbb{R}$.** Sebbene i numeri razionali non coprano tutta la retta dei numeri reali $\mathbb{R}$ per via della presenza di numeri irrazionali, è vero che ci vanno molto vicini... Comunque si fissa una soglia di errore $\varepsilon$, è possibile approssimare un numero reale $x$ con un numero razionale $y$ commettendo un errore più piccolo della soglia $\varepsilon$ consentita, cioè

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

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

La costruzione fatta poco più su per la rappresentazione decimale di un numero reale $x$ fornisce una maniera pratica per determinare una tale approssimazione. Per costruzione $x_n \in \mathbb{Q}$ per ogni $n$ e $\lim_{n\to+\infty} x_n = x$, quindi dato $\varepsilon > 0$, basta scegliere $y = x_n$ con $n$ sufficientemente grande.

Dal punto di vista della misurazione concreta, la densità è una proprietà notevole! In linea di principio, tenendo conto che c'è sempre un errore di misurazione, ogni misura può essere compiuta in $\mathbb{Q}$ (ed è quello che si fa in pratica); infatti un qualsiasi elemento di $\mathbb{R}$ può essere individuato con un arbitrario grado di precisione usando punti razionali.

Ad esempio, anche un qualsiasi computer ragiona sempre e solo con (un sottoinsieme dei) numeri razionali.

**Costruzione dei numeri reali a partire dai numeri razionali.** L'obiettivo delle (poche) pagine che restano è di accennare come "costruire" l'insieme $\mathbb{R}$ a partire da $\mathbb{Q}$ (c'è un prerequisito: conoscere il significato di "classi di equivalenza" e di "insieme quoziente"), senza intenzione di esaurire l'argomento. Che si intende con la frase "costruire l'insieme $\mathbb{R}$"? Non è già chiaro chi sia? Finora ci siamo appoggiati su un'idea del tutto intuitiva: $\mathbb{R}$ è una retta senza "buchi". Se si vuole dare robustezza ad $\mathbb{R}$, occorre una definizione più rigorosa. Usiamo come punto di partenza $\mathbb{Q}$ con la sua struttura di campo totalmente ordinato. In $\mathbb{Q}$ è possibile dare senso alla nozione di successione e di convergenza esattamente nello stesso modo usato per $\mathbb{R}$. Lo stesso dicasi per le successioni di Cauchy. Nel seguito $x_n, y_n$ indicheranno sempre successioni di numeri razionali.

DEFINIZIONE 5.14. *La successione $x_n \in \mathbb{Q}$ è una* SUCCESSIONE DI CAUCHY *se:*

$$\forall\, \varepsilon \in \mathbb{Q}, \quad \varepsilon > 0, \quad \exists\, n_\varepsilon \in \mathbb{N}, \quad \textit{tale che} \quad |x_m - x_n| < \varepsilon \quad \forall\, m, n \geq n_\varepsilon.$$

Le successioni di Cauchy in $\mathbb{Q}$ godono di un certo numero di buone proprietà.

1. *Ogni successione convergente è di Cauchy.*
2. *Ogni successione di Cauchy è limitata.*
3. *Se $x_n$ e $y_n$ sono di Cauchy, allora anche $x_n + y_n$ e $x_n \cdot y_n$ sono di Cauchy.*
4. *Se $x_n$ e $y_n$ sono di Cauchy, allora anche $|x_n - y_n|$ è di Cauchy.*

Lasciamo la dimostrazione di queste proprietà per esercizio.

L'idea nella costruzione di $\mathbb{R}$ a partire da $\mathbb{Q}$ tramite le successioni di Cauchy è che un numero reale è individuato da una successione che lo approssimi con un errore piccolo quanto si vuole. Ad esempio, per approssimare/definire il numero (irrazionale) $\pi$ si può considerare la successione di numeri (razionali)

$$x_0 = 3 \quad x_1 = 3,1 \quad x_2 = 3,14 \quad x_3 = 3,141 \quad x_4 = 3,1415 \quad x_5 = 3,14159 \quad \ldots,$$

cioè l'elemento $x_n$ è quello che si ottiene considerando i primi $n$ decimali[4]. Quindi il punto di partenza è che un numero può essere definito, in qualche modo, come una successione di Cauchy. Consideriamo quindi l'insieme

$$S = \Big\{ \{x_n\} \text{ successione di Cauchy in } \mathbb{Q} \Big\}.$$

Il problema di questo insieme è che consta di troppi elementi: ci sono molti modi per rappresentare lo stesso numero! Ad esempio, le due successioni

$$x_0 = 1, \quad x_1 = 1, \quad x_2 = 1, \quad x_3 = 1, \quad x_4 = 1 \quad \ldots$$

[4]Una successione scelta in questo modo è di Cauchy, dato che

$$|x_{n+p} - x_n| \leq 10^{-n} \qquad \forall\, p \in \mathbb{N}.$$

$$y_0 = 0, \quad y_1 = 0,9, \quad y_2 = 0,99, \quad y_3 = 0,999, \quad y_4 = 0,9999 \quad \dots$$

rappresentano entrambe il numero 1. Occorre quindi raggruppare gli elementi di $S$ in classi di equivalenza, in modo che ogni singola classe rappresenti un numero reale. La relazione d'equivalenza si definisce così: date $\{x_n\}$ e $\{y_n\}$ in $S$

$$\{x_n\} \sim \{y_n\} \qquad \text{se} \qquad \lim_{n\to+\infty}(x_n - y_n) = 0.$$

E' facile verificare che si tratta di una relazione di equivalenza (riflessività e simmetria sono banali, la transitività discende dalla disuguaglianza triangolare). A questo punto il gioco è fatto: l'insieme dei numeri reali è definito da

$$\mathbb{R} := S/\sim .$$

Un numero reale quindi è una classe d'equivalenza di successioni di Cauchy. A rigore, quindi, andrebbe indicato con un simbolo del tipo $\big[\{x_n\}\big]$. Ora si tratta di definire la somma, il prodotto, l'ordinamento e poi controllare che valgano tutte le proprietà che devono valere. Avrete già capito che non ho alcuna intenzione di fare queste verifiche. Mi limiterò a qualche cenno.

**Somma e prodotto.** Date due successioni di Cauchy, la loro somma (cioè la somma elemento per elemento) è ancora di Cauchy. Quindi, nell'insieme $S$ è definita un'operazione di somma. Lo stesso dicasi per il prodotto. La stessa definizione è valida per $\mathbb{R}$? Bisogna verificare che la somma sia indipendente dalla scelta del rappresentante: se $\{x_n\} \sim \{x'_n\}$ e $\{y_n\} \sim \{y'_n\}$ è vero che $\{x_n + y_n\} \sim \{x'_n + y'_n\}$ e $\{x_n y_n\} \sim \{x'_n y'_n\}$? La risposta è "SI": infatti

$$\lim_{n\to+\infty}(x_n + y_n) - (x'_n + y'_n) = \lim_{n\to+\infty}(x_n - x'_n) + (y_n - y'_n) = 0,$$
$$\lim_{n\to+\infty}(x_n y_n - x'_n y'_n) = \lim_{n\to+\infty} y_n(x_n - x'_n) + x'_n(y_n - y'_n) = 0.$$

**Ordinamento.** Definisco prima i numeri positivi:

$$\big[\{x_n\}\big] > 0 \qquad \text{se} \qquad \exists\, \varepsilon > 0,\, n_\varepsilon \in \mathbb{N} \quad \text{tale che} \quad x_n \geq \varepsilon \quad \forall\, n \geq N.$$

Il senso è che un numero è positivo se una successione di Cauchy che lo rappresenta è maggiore di un numero fissato $\varepsilon > 0$ da un certo indice in poi. Facile è vedere che la definizione è ben posta (cioè non dipende dalla scelta del rappresentante). Allora

$$\big[\{x_n\}\big] \leq \big[\{y_n\}\big] \qquad \text{se} \qquad \big[\{y_n - x_n\}\big] > 0 \text{ oppure } \big[\{x_n\}\big] = \big[\{y_n\}\big].$$

Anche qui bisognerebbe verificare che valgano le proprietà richieste sull'ordinamento. Evito, così come evito di mostrare che anche questo procedimento permette di costruire $\mathbb{R}$, il campo ordinato completo che sempre abbiamo sognato.

## Appendice B. La serie armonica e la serie armonica generalizzata

Una delle serie numeriche più famose è quella di termine generico $\frac{1}{n}$, che è detta SERIE ARMONICA. Questa serie è divergente

$$\sum_{n=1}^{\infty} \frac{1}{n} = +\infty. \tag{5.2}$$

Infatti, sia $S_n := \sum_{k=1}^{n} \frac{1}{k}$ e consideriamo la sottosuccessione $S_{2^k}$ di $S_n$. Allora si ha

$$S_1 = 1, \quad S_2 = 1+\frac{1}{2}, \quad S_4 = 1+\frac{1}{2}+\left(\frac{1}{3}+\frac{1}{4}\right), \quad S_8 = 1+\frac{1}{2}+\left(\frac{1}{3}+\frac{1}{4}\right)+\left(\frac{1}{5}+\frac{1}{6}+\frac{1}{7}+\frac{1}{8}\right)$$

e così via. Obiettivo? Mostrare che

$$\lim_{k\to+\infty} S_{2^k} = +\infty.$$

Teniamo 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 $S_{2^{k+1}}$:

$$\begin{aligned} S_{2^{k+1}} &= S_{2^k}+\left(\frac{1}{2^k+1}+\frac{1}{2^k+2}+\cdots+\frac{1}{2^{k+1}}\right) \\ &\geq S_{2^k}+\left(\frac{1}{2^{k+1}}+\frac{1}{2^{k+1}}+\cdots+\frac{1}{2^{k+1}}\right) = S_{2^k}+\frac{2^k}{2^{k+1}} = S_{2^k}+\frac{1}{2} \end{aligned}$$

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 (vedi Esercizio 4.5).

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*, che è uno dei concetti fondamentali dell'analisi matematica. E non solo.

## 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. Nel caso visto nel Capitolo precedente, l'oggetto è il numero reale $a_n$ (che dipende dal parametro $n$), l'esistenza del limite per $n \to +\infty$ manifesta il fatto che i valori $a_n$ tendono a stabilizzarsi attorno ad un valore $n$ quando si considerano valori della variabile $n$ tendenti a $+\infty$.

**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

$$\forall\, \varepsilon > 0 \quad \exists\, M \quad \text{tale che} \qquad |d(x)| < \varepsilon \quad \forall\, x > M. \tag{1.1}$$

Il numero $M \in \mathbb{R}$ dipende dalla scelta di $\varepsilon$ (come $n_\varepsilon$ 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 (1.1) 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 $\ell$ 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 $x \to +\infty$

$$\lim_{x \to +\infty} f(x) = \ell, \tag{1.2}$$

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

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

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

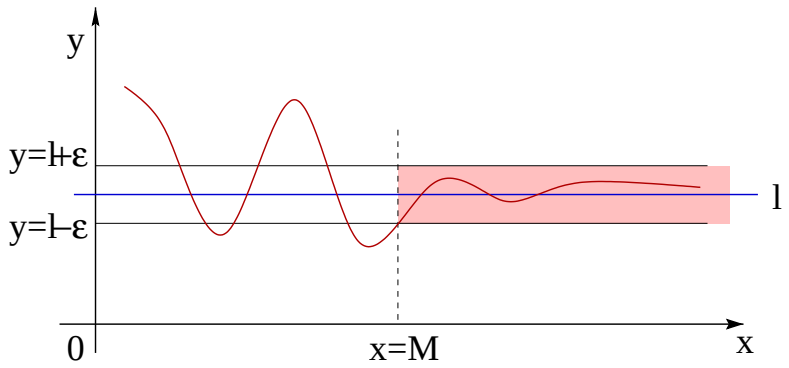



FIGURA 1. Una funzione che tende ad un limite per $x \to +\infty$.

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 $\mathbb{R}$. Piccole modifiche permettono di estendere il concetto a funzioni $f : (a, +\infty) \to \mathbb{R}$ per qualche $a > 0$, infatti per "mandare $x$ a $+\infty$" basta che la funzione sia definita da un certo valore $a$ in poi. Precisamente, se $f : (a, +\infty) \to \mathbb{R}$, allora

$$\lim_{x\to+\infty} f(x) = \ell \qquad \Longleftrightarrow \qquad \begin{cases} \forall\, \varepsilon > 0 \quad \exists\, M > a \quad \text{tale che} \\ \qquad |f(x) - \ell| < \varepsilon \quad \forall\, x > M. \end{cases}$$

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: $\lim_{x\to+\infty} f(x) = +\infty$ o $= -\infty$;

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

**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

$$\forall\, \varepsilon > 0 \quad \exists\, \delta = \delta(\varepsilon) > 0 \quad \text{t.c.} \quad |d(x)| < \varepsilon \quad \forall\, x \in (a, b),\ 0 < |x - x_0| < \delta. \tag{1.3}$$

Rispetto alla definizione di funzione infinitesima per $x \to +\infty$, l'unica differenza (non da poco!) 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 $\delta > 0$. Una volta formalizzato il concetto di "infinitesimo", la definizione di convergenza ad un valore limite è immediata.

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*

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

*In questo caso, si scrive*

$$\lim_{x \to x_0} f(x) = \ell. \tag{1.4}$$

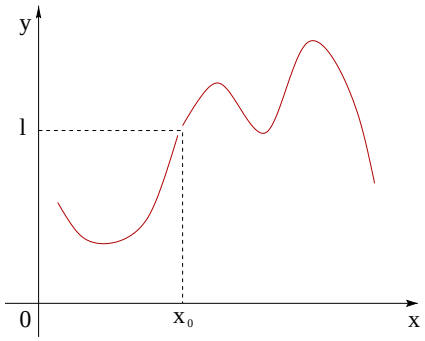



FIGURA 2. Limite in un punto

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

$$|x - x_0| \to 0 \quad \Rightarrow \quad d(x) := |f(x) - \ell| \to 0,$$

cioè $f(x) - \ell$ tende a 0 quando $x - x_0$ tende a 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 $\delta$ in funzione di $\varepsilon$, così come richiesto dalla definizione, basta osservare

che se $|x-1|<\delta$, allora $|f(x)-\ell|<3\delta$, quindi dato $\varepsilon>0$, basta scegliere $\delta$ in modo che $3\delta=\varepsilon$, cioè $\delta=\varepsilon/3$.

ESEMPIO 1.5. Fissato $x_0\in\mathbb{R}$, calcoliamo $\lim_{x\to x_0}\sin x$. 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,

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

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

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

da cui segue

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

Volendo determinare esplicitamente $\delta$, 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* $\lim_{x\to x_0}\cos x=\cos x_0$.

**Suggerimento.** Utilizzare $\cos x=\sin\left(\frac{\pi}{2}-x\right)$, oppure le formule di prostaferesi.

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 (5.1), 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 $|f(x)-\ell|=|e^x-1|$:

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

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

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

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 non 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 $\ell$ per $n \to +\infty$:

$$\lim_{n\to+\infty} x_n = x_0 \qquad \Rightarrow \qquad \lim_{n\to+\infty} f(x_n) = \ell. \tag{1.5}$$

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

Dall'implicazione (1.5) discende il seguente

CRITERIO 1.10. *Non esistenza del limite.* *Se esistono due successioni $x_n$ e $\xi_n$ 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

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

e consideriamo $x_n = \frac{1}{n}$ e $\xi_n = -\frac{1}{n}$. È evidente che $\lim_{n\to+\infty} x_n = \lim_{n\to+\infty} \xi_n = 0$. 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 la funzione* $\sin(1/x)$ *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 $\ell$ per $x$ che tende a $x_0$, e si scrive* $\lim_{x\to x_0^+} f(x) = \ell$, *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* $\lim_{x\to x_0^-} f(x) = \ell$.

ESERCIZIO 1.13. *Dimostrare che* $\lim_{x\to 0^-} sgn\, x = -1, \quad \lim_{x\to 0^+} sgn\, x = +1.$

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. *Esistenza del limite. 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 $x \to x_0$. Alla luce di questa osservazione e 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:

$$\left.\begin{array}{l}\lim_{x\to x_0} f(x) = \ell \\ \lim_{x\to x_0} g(x) = m\end{array}\right\} \quad \Rightarrow \quad \begin{cases} \lim_{x\to x_0} \big(f(x) \pm g(x)\big) = \ell \pm m, \\ \lim_{x\to x_0} f(x)g(x) = \ell\, m, \\ \lim_{x\to x_0} \dfrac{f(x)}{g(x)} = \dfrac{\ell}{m} \qquad \text{se } m \neq 0. \end{cases} \tag{1.6}$$

ESEMPIO 1.16. Per calcolare

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

non è una buona idea utilizzare la definizione! Basta applicare le regole appena descritte:

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

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

Ad essere rigorosi, i passaggi precedenti vanno letti dall'ultimo al primo. Infatti, (1.6) va letta così: "**se** i limiti di $f$ e $g$ esistono, **allora** il limite della somma è la somma dei limiti, ecc. ecc. ecc." Quindi, a rigore, bisogna prima osservare che i limiti dei singoli termini esistono e poi procedere a ritroso...

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

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

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 $f(x) \le g(x) \le h(x)$ 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 in cui si può incappa più spesso 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 $\varepsilon$ 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:

$$\lim_{x\to 0} \frac{\sin x}{x} = 1, \tag{1.7}$$

(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.

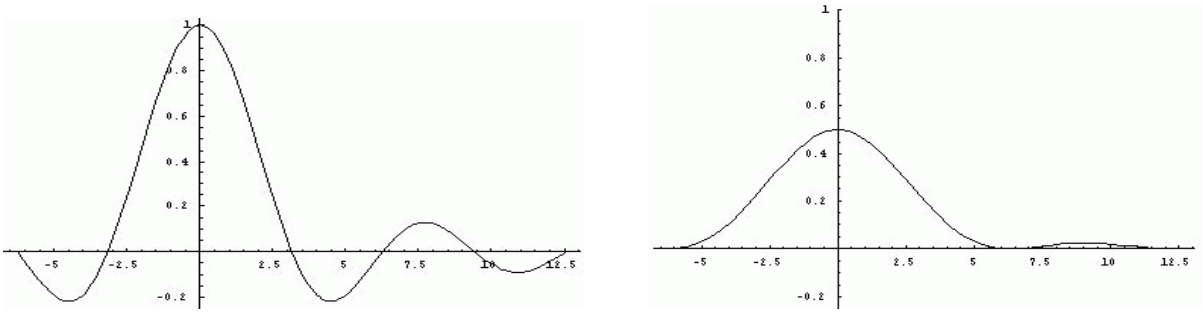



FIGURA 3. (a) $y = \frac{\sin x}{x}$, (b) $y = \frac{1-\cos x}{x^2}$

ESERCIZIO 1.20. *Utilizzando (1.7), 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

$$\lim_{x\to 0}\frac{e^x-1}{x}=1, \qquad \lim_{x\to 0}\frac{\ln(1+x)}{x}=1. \tag{1.8}$$

Per dimostrare il primo limite di (1.8), 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 $e^x=\sum_{n=0}^{\infty}\frac{x^n}{n!}$. 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 (1.8) 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.

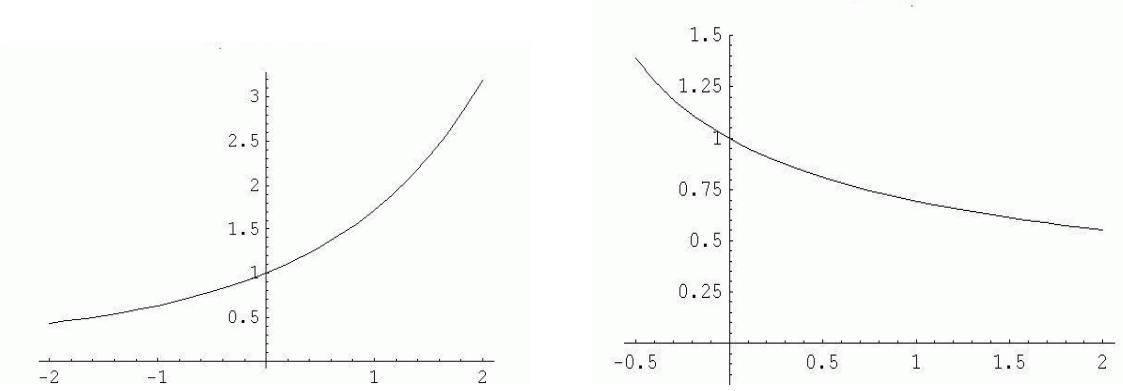



FIGURA 4. (a) $y=\frac{e^x-1}{x}$, (b) $y=\frac{\ln(1+x)}{x}$

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

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

Il significato di ciascuno di questi è particolarmente interessante. Nel primo limite, la funzione $x^\alpha$ 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 $x^\alpha$ 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 $x^\alpha$, e il terzo esprime che, al contrario, il logaritmo $\log_a x$, con $a > 1$, diverge più lentamente di $x^\alpha$. 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 ammette 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 $J$.

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

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

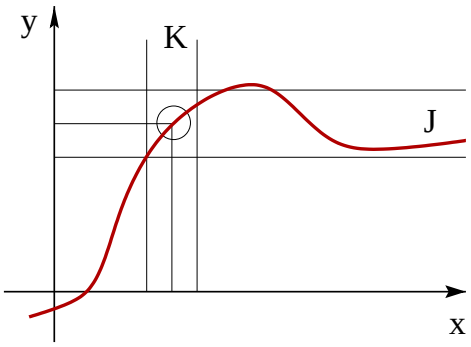



FIGURA 5. Significato geometrico della continuità

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.

> *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 $\varepsilon$ 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 $\delta$, distino da $f(x_0)$ meno di $\varepsilon$. Se il giocatore $B$ trova un $\varepsilon > 0$ per cui $A$ non possa rispondere, vince; viceversa, se per ogni $\varepsilon$, $A$ è in grado di trovare $\delta$ 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*

[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.

La prima delle due proprietà discende dalla catena di implicazioni

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

Come sempre, una dimostrazione rigorosa richiederebbe un uso meno spregiudicato delle implicazioni $\Rightarrow$...

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 una funzione continua!

ESEMPIO 2.4. *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 2.5. *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.

ESEMPIO 2.6. *I polinomi sono funzioni continue.* Qui basta combinare le proprietà dei limiti (1.6), 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

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

ESEMPIO 2.7. *La funzione $f(x) = \sin x$ è una funzione continua.* Lo abbiamo già visto nell'Esempio 1.5. Stesso dicasi per $\cos x$, a patto che abbiate risolto l'Esercizio 1.6.

ESEMPIO 2.8. 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 2.9. *Perché le funzioni $f(x) = 10^x$ e $g(x) = \log_{10} x$ sono continue?*

**Soluzione.** La funzione $f$ può essere ricondotta alla funzione $e^x$. Infatti, in generale, vale la formula $a^b = e^{b \ln a}$, dunque $10^x = e^{x \ln 10}$. La funzione $f$ è quindi una composizione di funzioni continue (moltiplicazione per $\ln 10$ e funzione esponenziale), quindi è una funzione continua. La funzione $g$ è, per definizione, la funzione inversa della funzione $f$, quindi è anch'essa una funzione continua. □

**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 2.10. 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 2.11. *(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 $L$. In altre parole, la trasformazione determinata dalla funzione non dilata le distanze oltre il fattore $L$:

$$\operatorname{dist}(f(x_2), f(x_1)) \leq L \operatorname{dist}(x_2, x_1).$$

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 2.12. *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 2.13. *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

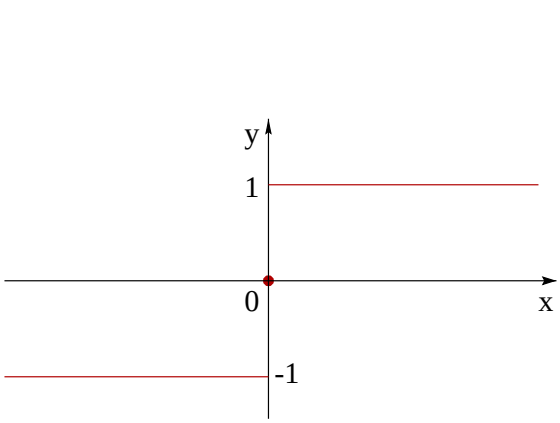
Un modo per chiarire ulteriormente la definizione di continuità è "in negativo", cioè dando esempi per cui non è soddisfatta.

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 $\delta$ in funzione di $\varepsilon$ dato?). In $x_0 = 0$ la funzione, invece, non è continua. Infatti non è possibile determinare nessun $\delta$ quando $\varepsilon$ 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

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



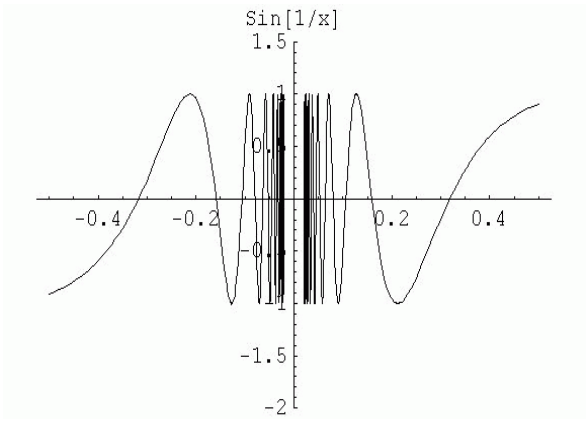




FIGURA 6. (b) La funzione *segno*; (b) Il grafico di $\sin(1/x)$.

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 $(0, 1)$. Chiaramente una figura chiarisce più di mille parole (Fig.6(b)). E' evidente che non c'è alcuna speranza che la funzione possa essere continua in $x = 0$ (si noti che, invece, la funzione $f$ è continua in ogni $x \neq 0$).

Piccole varianti della funzione precedente possono condurre ad una funzione continua. Ad esempio consideriamo la funzione $g$ ottenuta moltiplicando la $f$ dell'Esempio 3 per $x$, cioè

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

Questa funzione (vedi Fig.7) è continua in 0, infatti

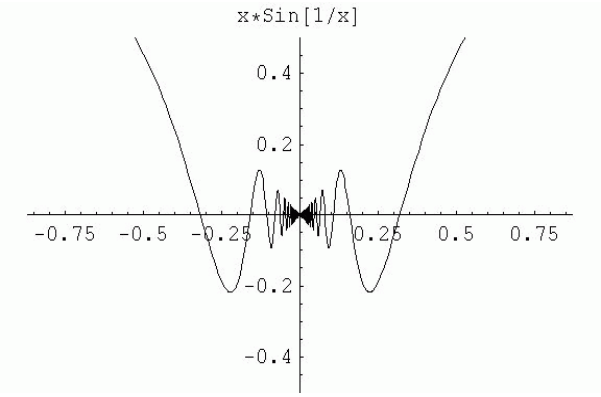



FIGURA 7. Il grafico di $g(x)$

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

Sapete dire se è continua in 0 la funzione $(x^2+1)f(x)$, dove $f$ è ancora quella dell'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 $\eta$ compresa tra 2200 e 3000? La risposta (intuitiva) è "SI", a meno che non abbia utilizzato il teletrasporto...

TEOREMA 4.1. *Teorema del valore intermedio.* *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.

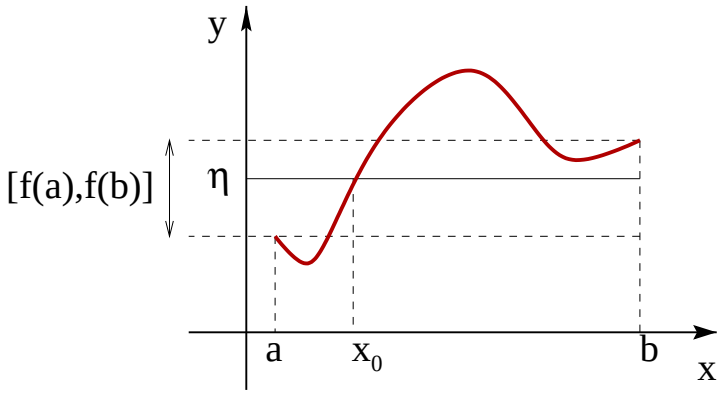



FIGURA 8. Il Teorema del valore intermedio

**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 banale 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 $\eta$ 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, che, chiaramente non è invertibile. In realtà basta che la funzione $f$ sia costante in un intervallo qualsiasi del suo insieme di definizione per non essere invertibile.

**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 $f(x_0) \leq f(x) \leq f(x_1)$ 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 $\inf_{[-1,1]} f(x) = 0$, 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.*

**Dimostrazione.** ...sapete scriverla da soli? ∎

## 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 $I_1, I_2, \dots, I_n, \dots$ chiusi e limitati che siano incapsulati, cioè tali che $I_{n+1} \subset I_n$ per ogni $n \in \mathbb{N}$, esiste sempre almeno un punto $x_0 \in \mathbb{R}$ tale che $x_0 \in I_n$ per ogni $n$.*

ASSIOMA DI ARCHIMEDE. *Per ogni numero reale $a$, esiste un numero naturale $n$ più grande di $a$: in simboli,*

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

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

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

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 $a_n \le a_{n+1} \le b_{n+1} \le b_n$) per ogni $n \in \mathbb{N}$;*
*(ii) $\lim_{n\to+\infty} b_n - a_n = 0$.*

*Allora, esiste un unico $x_0 \in \mathbb{R}$ tale che $\bigcap_{n=1}^{\infty} I_n = \{x_0\}$, cioè*

$$a_n \le x_0 \le b_n \qquad \text{per ogni} \quad n \in \mathbb{N}.$$

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

Il contenuto del Corollario è chiaro: 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 $x_0 \le x_1$. Allora $a_n \le x_0 \le x_1 \le b_n$ per ogni $n$. Da questa relazione segue che $0 \le x_1 - x_0 \le b_n - a_n$. Dato che la successione $b_n - a_n$ è infinitesima, per ogni $\varepsilon > 0$, si ha $0 \le x_1 - x_0 \le b_n - a_n < \varepsilon$ per ogni $n$ sufficientemente grande. In definitiva, per ogni $\varepsilon > 0$, si ha $0 \le x_1 - x_0 < \varepsilon$, da cui segue $x_0 = x_1$, che conclude la prima parte del Teorema.

Inoltre, dato che $a_n \le x_0 \le b_n$, si ha anche $0 \le x_0 - a_n \le b_n - a_n \to 0$ 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 $n \to +\infty$. Grazie al risultato appena visto è possibile dedurre l'esistenza del punto $x_0$ che, se la scelta degli $I_n$ è operata in maniera astuta, soddisfa le proprietà richieste. 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 che $f(a) < f(b)$ 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 $f(x_0) \leq \eta \leq f(x_0)$, cioè $f(x_0) = \eta$. ∎

**Dimostrazione del Teorema di Weierstrass (Teorema 4.6).** Sia $\Lambda := \sup_{x\in[a,b]} f(x)$. Per ora non abbiamo nessuna garanzia che $\Lambda$ sia finito.

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 $\Lambda$. 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

$\lim_{n\to\infty} a_n = \lim_{n\to\infty} b_n = x_1$. Vogliamo a questo punto mostrare che $\Lambda$ 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 $\varepsilon$, e di conseguenza $\delta$, scelgo $n \in N$ tale che $I_n \subset (x_1 - \delta, x_1 + \delta)$. Allora

$$f(x) < f(x_1) + \varepsilon \qquad \forall\, x \in I_n, \tag{5.1}$$

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 $\Lambda$ è finito, che $f(x_1) \leq \Lambda$ e che $\Lambda \leq f(x_1) + \varepsilon$. Ma, in quest'ultima relazione, $\varepsilon$ può essere scelto arbitrariamente, quindi

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

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

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 $\Lambda$.

## Appendice A. Altre applicazioni del *divide et impera*

In quest'ultima parte, diamo altre applicazioni del metodo di bisezione. Queste applicazioni nulla hanno a che vedere con le funzioni continue. Buon divertimento!

**Esistenza dell'estremo superiore.** Ricordiamo prima di tutto la definizione di estremo superiore di un insieme $E$. Il valore $\Lambda \in \mathbb{R}$ è l'ESTREMO SUPERIORE DI $E \subset \mathbb{R}$ se
(i) $\Lambda$ è un maggiorante di $E$, cioè per ogni $y \in E$, si ha $y \leq \Lambda$;
(ii) $\Lambda$ è il più piccolo dei maggioranti, cioè se $L$ è un maggiorante di $E$, allora $\Lambda \leq L$.

**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 $\Lambda$ 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 $\Lambda$, quindi anche la condizione (ii) è soddisfatta. ∎

**Calcolo approssimato di $\sqrt{2}$.** Tramite il *divide et impera*, possiamo anche costruire una successione di razionali che approssimi il numero irrazionale $\sqrt{2}$. Ecco la strategia:

*Passo 1.* Determiniamo una stima per eccesso ed una per difetto di $\sqrt{2}$. Questo può essere fatto trovando due numeri (razionali) $a_1$ e $b_1$ tali che $a_1^2 < 2 < b_1^2$. Ad esempio $1^2 = 1 < 2 < 4 = 2^2$, quindi possiamo scegliere $a_1 = 1$ e $b_1 = 2$.

*Passo 2.* Determiniamo il punto medio $\frac{a_1+b_1}{2}$ e calcoliamone il quadrato. Se tale numero è più grande di 2, poniamo $a_2 = a_1$ e $b_2 = \frac{a_1+b_1}{2}$. Se invece il quadrato di $\frac{a_1+b_1}{2}$ è più piccolo di 2, definiamo $a_2 = \frac{a_1+b_1}{2}$ e $b_2 = b_1$. Nel nostro caso,

$$\frac{1+2}{2} = \frac{3}{2} \quad \Rightarrow \quad \left(\frac{3}{2}\right)^2 = \frac{9}{4} > 2 \quad \Rightarrow \quad a_2 = 1,\ b_2 = \frac{3}{2}.$$

*Passo 3.* Iteriamo il procedimento, bisecando l'intervallo $[a_2, b_2]$ e procedendo con la stessa regola di prima per definire $a_3, b_3$.

Le due successioni $a_n$ e $b_n$ godono della proprietà $a_n < \sqrt{2} < b_n$ per ogni $n$. Inoltre, dato che $b_n - a_n = \frac{b-a}{2^{n-1}}$ è infinitesima, per $n \to \infty$ convergono allo stesso valore $x_0$, che deve necessariamente coincidere con $\sqrt{2}$.

Nel caso di $a_1 = 1$ e $b_1 = 2$ le stime per eccesso e per difetto sono

| | | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| $a_n$ | 1 | 1 | 1,25 | 1,375 | 1,375 | 1,40625 | 1,40625 | 1,4140625 |
| $b_n$ | 2 | 1,5 | 1,5 | 1,5 | 1,4375 | 1,4375 | 1,421875 | 1,421875 |

Come si vede, questo metodo, pur essendo efficace, non è particolarmente efficiente, nel senso ad ogni interazione permette di migliorare la stima solo della metà della lunghezza dell'intervallo. Dopo 8 iterazioni, siamo ancora lontani da una stima soddisfacente (si riesce a conoscere $\sqrt{2}$ con un errore minore di $1,421875 - 1,4140625 = 0,0078125$). Si può fare di meglio...

Confrontiamo con un altro algoritmo: il metodo di Erone. Scelto $a_1$ in qualche modo (diciamo $a_1 = 2$), determiniamo la successione per ricorrenza definita da

$$a_{n+1} = \frac{1}{2}\left(a_n + \frac{2}{a_n}\right).$$

Ecco un tabulato delle prime 5 iterazioni:

| | | | | | |
|---|---|---|---|---|---|
| $a_n$ | 2 | 1,5 | $1,41\overline{6}$ | 1,414215686 | 1,414213562 |

Provate a confrontare col valore di $\sqrt{2}$ dato da una calcolatrice...

Il fatto che l'algoritmo di Erone dia luogo ad un'approssimazione di $\sqrt{2}$ molto più precisa di quella data dal *divide et impera* può essere visto in maniera ancora più esplicita. Come si è detto il *divide et impera* permette di dimezzare l'errore commesso al passo precedente. Quindi, se $r_n$ indica l'errore di approssimazione al passo $n-$esimo, vale una relazione del tipo $r_{n+1} \approx \frac{1}{2} r_n$. Nel caso dell'algoritmo di Erone, invece, se $R_n = a_n - \sqrt{2}$, si ha

$$R_{n+1} = a_{n+1} - \sqrt{2} = \frac{1}{2a_n}(a_n - \sqrt{2})^2 = \frac{1}{2a_n} R_n^2.$$

Per $a_n$ vicino a $\sqrt{2}$, vale una relazione approssimata del tipo $R_{n+1} \approx \frac{1}{2\sqrt{2}} R_n^2$. Che succede se l'errore commesso al passo $n-$esimo è dell'ordine di grandezza di $10^{-2}$? Con il *divide et impera*, al passo successivo l'errore è $0,5 \cdot 10^{-2}$, mentre con l'algoritmo di Erone l'errore al passo successivo è dell'ordine di grandezza di $10^{-4}$.

Questo esempio mostra che non tutte le approssimazioni sono "equivalenti": in molte situazioni concrete, algoritmi che convergono più rapidamente sono preferibili ad altri che se la prendono troppo comoda...

CAPITOLO 5

# L'integrale

Il problema della misurazione delle lunghezze di segmenti ci ha condotti per un sentiero particolarmente interessante. Ricordate? Dai naturali, ai relativi, ai razionali ed infine ai numeri reali. E se ci viene la voglia di misurare aree di regioni del piano? Nel caso di un rettangolo, l'area è nota: è il prodotto delle lunghezze dei lati. Nel caso di una regione che sia unione finita di rettangoli che si toccano al più lungo il perimetro, basta sommare le aree dei singoli rettangoli. Ma per regioni più generali?

## 1. L'area di un sottografico: strategia d'attacco

Tra tutte le miriadi di regioni a cui possiamo pensare, alcune suonano più familiari a chi ha ormai acquisito una qualche dimestichezza con le funzioni di una variabile: *i sottografici di funzioni non negative*, cioè tali che $f(x) \geq 0$ per ogni $x \in [a,b]$.

DEFINIZIONE 1.1. *Data una funzione* $f : [a,b] \to [0,+\infty)$*, l'insieme*

$$S_f = \{(x,y) \mid \quad a \leq x \leq b, \quad 0 \leq y \leq f(x)\}$$

*si dice* SOTTOGRAFICO DI $f$ *nell'intervallo* $[a,b]$.

Si noti che, date due funzioni $f$ e $g$ non negative, il sottografico $S_f$ di $f$ è contenuto in $S_g$, sottografico di $g$, se e solo se $f \leq g$.

L'area del sottografico delle funzioni costanti è elementare: se $f(x) = C \geq 0$, il sottografico $S_f$ di $f$ in $[a,b]$ è un rettangolo, la cui area è, da sempre, $\mathcal{A}(S_f) = (b-a)\,C$. Altrettanto facile è l'area del sottografico di funzioni *costanti a tratti*.

DEFINIZIONE 1.2. *Dato l'intervallo* $[a,b]$*, un insieme* $P = \{x_0, x_1, \ldots, x_n\}$ *tale che* $a = x_0 < x_1 < \cdots < x_n = b$ *è una* PARTIZIONE DI $[a,b]$*. L'*AMPIEZZA *della partizione* $P$ *è il numero* $|P| := \max\{x_i - x_{i-1} \,:\, i = 1, \ldots, n\}$.

*Una funzione* $f : [a,b] \to \mathbb{R}$ *è una* FUNZIONE COSTANTE A TRATTI *(o* FUNZIONE A SCALA*) se esiste una partizione* $P = \{x_0, x_1, \ldots, x_n\}$ *di* $[a,b]$ *tale che* $f$ *è costante su ogni intervallo* $[x_{i-1}, x_i)$ *per* $i = 1, \ldots, n-1$ *e su* $[x_{n-1}, x_n]$*, ossia*

$$f(x) = \begin{cases} \alpha_i & x \in [x_{i-1}, x_i) \qquad i = 1, \ldots, n-1 \\ \alpha_n & x \in [x_{n-1}, x_n] \end{cases}$$

*dove* $\alpha_1, \ldots, \alpha_n$ *sono numeri reali.*

Per una funzione $f$ costante a tratti e non negativa il sottografico non è altro che un'unione finita di rettangoli, vedi Figura 1. L'area di $S_f$ è data dalla somma delle aree di questi rettangoli:

$$\mathcal{A}(S_f) = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) = \sum_{i=1}^{n} \alpha_i (x_i - x_{i-1})$$

dove $\xi_i$ è un qualsiasi punto di $[x_{i-1}, x_i)$ (tanto $f$ è costante in ogni intervallo $[x_{i-1}, x_i)$!).[1]

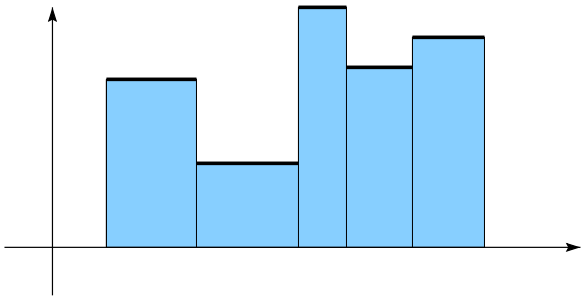


FIGURA 1. Una funzione costante a tratti non negativa ed il suo sottografico.

Facciamo conto che per ogni funzione $f$ non negativa sia ben definito un certo numero $\mathcal{A}(S_f)$ che è l'area del sottografico $S_f$ (questa affermazione si rivelerà in realtà falsa, nel senso che per i sottografici di alcune funzioni l'area non è definita!). Come determinare $\mathcal{A}(S_f)$? Una buona idea, che non si discosta di molto dall'idea del "ricoprire un pavimento con delle mattonelle", è quella di cercare di ricondursi, tramite approssimazioni, ad oggetti di cui si sappia già calcolare l'area, ad esempio i sottografici di funzioni costanti a tratti. Inoltre, vale la pena di considerare due approssimazioni: una fatta tramite un sottografico contenuto in quello di cui si cerca l'area ed una con sottografico che lo contiene. Le due approssimazioni daranno una stima dal basso ed una dall'alto del valore che andiamo cercando. Poi si cercherà di rendere le approssimazioni sempre più precise.

Quindi, data $f : [a, b] \to \mathbb{R}$ non negativa, consideriamo approssimazioni del sottografico $S_f$ date da sottografici di funzioni costanti a tratti $g$ ed $h$, con $g \leq f \leq h$ in $[a, b]$ (vedi Figura 2(a)). Una prima osservazione salta all'occhio: per poter stimare per eccesso la funzione $f$ con una funzione costante a tratti non negativa *occorre che la funzione $f$ sia superiormente limitata*[2]. Se $\sup_{[a,b]} f(x) = +\infty$, non è possibile trovare una funzione costante a tratti $h$ per cui $f \leq h$ (vedi Figura 2(b)). Per la stima dal basso con funzioni costanti a tratti, serve anche che $f$ sia inferiormente limitata, ipotesi che nel

[1] Si noti, *en passant*, che se si modifica una funzione costante a tratti $f$ in un numero finito di punti, il valore di $\mathcal{A}(S_f)$ rimane invariato. Infatti modifiche di questo genere corrispondono a modificare il sottografico della funzione di partenza di un numero finito di segmenti verticali, ciascuno dei quali ha area nulla.

[2] Per gentilezza, ricordiamo che $f$ è *superiormente limitata* se esista $M > 0$ per cui $f(x) \leq M$ per ogni $x \in [a, b]$.

nostro presente caso è automaticamente soddisfatta dato che $f$ è non negativa. L'ipotesi di limitatezza della funzione $f$ è fondamentale qui, così come in tutto il Capitolo e in tutta la definizione dell'integrale definito. Appuntarselo.

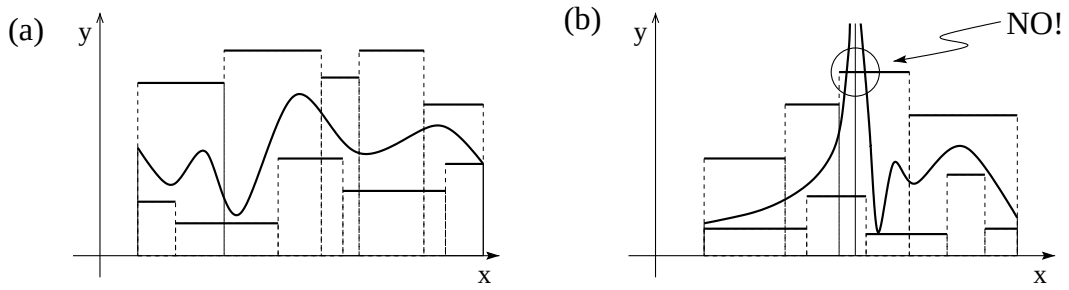



FIGURA 2. (a) Una funzione approssimata per eccesso e per difetto con due funzioni costanti a tratti. (b) Un tentativo (fallito) di approssimazione per eccesso di una funzione non limitata superiormente.

Operativamente è estremamente più comodo scegliere le due funzioni approssimanti in maniera astuta: prima di tutto facendo in modo che si tratti di funzioni costanti su intervalli definiti dalla stessa partizione, e poi scegliendo $g$ e $h$ in modo che in ciascuno di questi intervalli il loro valore sia pari all'estremo inferiore e all'estremo superiore della funzione $f$ considerata. Con questa seconda richiesta, la definizione di $g$ ed $h$ è ottimale una volta fissata la partizione.

Precisamente, data una partizione $P = \{x_0, x_1, \ldots, x_n\}$ dell'intervallo $[a, b]$, dividiamo l'intervallo $[a, b]$ negli $n$ sottointervalli $[x_{i-1}, x_i]$ per $i = 1, \ldots, n$. Così facendo, il sottografico $S_f$ si decompone nell'unione di $n$ sottografici, ognuno relativo ad uno degli intervalli $[x_{i-1}, x_i]$:

$$S_f = \bigcup_{i=1}^{n} S_f^i \qquad \text{dove} \quad S_f^i = \{(x, y) \,:\, x_{i-1} \le x \le x_i, \quad 0 \le y \le f(x)\}\,.$$

Insiemi $S_f^i$ con $i$ consecutivo hanno in comune un segmento verticale, ma, dato che i segmenti hanno area nulla, possiamo affermare con tranquillità che l'area di $S_f$ è la somma delle aree di $S_f^i$.

Se consideriamo due funzioni $g$ e $h$ costanti negli intervalli definiti da $P$ e tali che $g \le f \le h$, stiamo approssimiamo l'area $\mathcal{A}(S_f^i)$ con quella di due rettangoli che hanno come base il sotto–intervallo $(x_{i-1}, x_i)$ e come altezze un minorante e un maggiorante di $f$ in $(x_{i-1}, x_i)$, rispettivamente. La scelta migliore, cioè quella per cui l'errore commesso nella stima dell'area di ogni $E_i$ è minima, è quella del "più grande dei minoranti" e del "più piccolo dei maggioranti", cioè dell'estremo inferiore e dell'estremo superiore di $f$ in $(x_{i-1}, x_i)$.

Quindi, data una partizione $P$ ci sono due funzioni $\underline{f}$ e $\overline{f}$, costanti a tratti negli intervalli definiti da $P$ e tali che $\underline{f} \le f \le \overline{f}$, che danno la migliore stima per difetto e la

migliore per eccesso. Le funzioni $\underline{f}$ e $\overline{f}$ sono definite da

$$\underline{f}(x) = \alpha_i := \inf_{[x_{i-1},x_i]} f(x) \qquad \forall\, x \in [x_{i-1}, x_i),$$
$$\overline{f}(x) = \beta_i := \sup_{[x_{i-1},x_i]} f(x) \qquad \forall\, x \in [x_{i-1}, x_i).$$

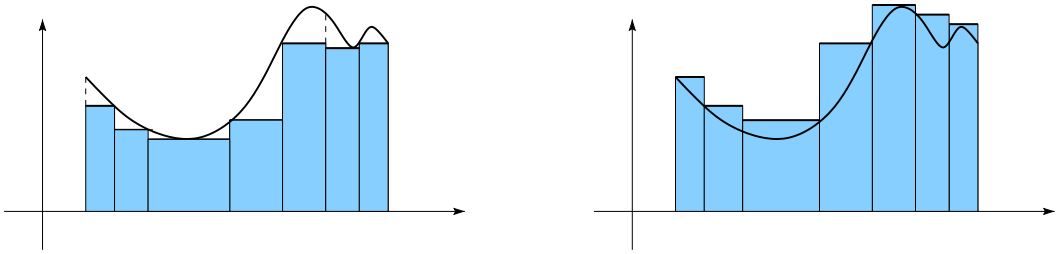


FIGURA 3. Due approssimazioni per $S_f$ tramite le funzioni $\underline{f}$ e $\overline{f}$.

In definitiva, data una partizione $P$ dell'intervallo $[a, b]$, sono ben definiti

$$\underline{\mathcal{A}}(S_f; P) := \mathcal{A}(S_{\underline{f}}) = \sum_{i=1}^{n} \alpha_i (x_i - x_{i-1}) \qquad \overline{\mathcal{A}}(S_f; P) := \mathcal{A}(S_{\overline{f}}) = \sum_{i=1}^{n} \beta_i (x_i - x_{i-1}),$$

con $\alpha_i = \inf_{(x_{i-1},x_i)} f(x)$, $\beta_i = \sup_{(x_{i-1},x_i)} f(x)$ e (tutte le volte che $\mathcal{A}(S_f)$ ha senso)

$$\underline{\mathcal{A}}(S_f; P) \le \mathcal{A}(S_f) \le \overline{\mathcal{A}}(S_f; P).$$

Dopo, si procede a migliorare l'approssimazione tramite la scelta di una partizione con un'ampiezza più piccola. Una possibilità è scegliere una successione di partizioni $P_n$ con ampiezza $|P_n|$ che tenda a zero per $n \to \infty$. In questo modo, passando al limite nelle due successioni numeriche $\underline{\mathcal{A}}(S_f; P_n)$ e $\overline{\mathcal{A}}(S_f; P_n)$ si dovrebbe (incrociando le dita) ottenere proprio l'area richiesta.

Negli esempi che seguono, utilizziamo la strategia appena proposta:

– dato $n \in \mathbb{N}$, scegliamo una partizione $P_n$ e costruiamo i numeri $\underline{\mathcal{A}}(S_f; P_n)$ e $\overline{\mathcal{A}}(S_f; P_n)$;
– calcoliamo i limiti

$$\lim_{n\to+\infty} \underline{\mathcal{A}}(S_f; P_n) \qquad \text{e} \qquad \lim_{n\to+\infty} \overline{\mathcal{A}}(S_f; P_n);$$

– se i limiti esistono e coincidono, il valore comune dei limiti è l'area cercata $\mathcal{A}(S_f)$, se i limiti non esistono o non coincidono, rimaniamo con un palmo di naso...

ESEMPIO 1.3. Sia $f(x) = x$ per $x \in [a, b]$ con $0 \le a < b$. Il sottografico della funzione $f$ nell'intervallo $[a, b]$ ha una forma estremamente familiare: nel caso in cui $a = 0$, si tratta di un triangolo, e nel caso di $a > 0$ si tratta di un trapezio. In entrambi casi, la geometria elementare fornisce una formula per il calcolo dell'area, che è data da:

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

(ricordate? "(base maggiore più base minore) per altezza diviso due"). Sembrerebbe quindi che non ci sia più nulla da dire. La questione invece si pone: cosa succede se si calcola l'area attraverso il procedimento di approssimazione per eccesso e per difetto proposto in precedenza?

Dividiamo l'intervallo $[a,b]$ in $n$ parti di uguale lunghezza tramite la partizione $P_n = \{x_k = a + kh \,:\, k = 0,\dots,n\}$ dove $h = (b-a)/n$. Dato che $f(x) = x$ è crescente,

$$\alpha_k = \inf_{[x_{k-1},x_k]} x = x_{k-1} = a + (k-1)h, \quad \beta_k = \sup_{[x_{k-1},x_k]} x = x_k = a + kh,$$

e $x_k - x_{k-1} = h$, valgono

$$\underline{\mathcal{A}}(S_f; P_n) = ah + \cdots + (a + (n-1)h)h = h\sum_{k=0}^{n-1}(a+kh) = hna + h^2\sum_{k=0}^{n-1} k,$$

$$\overline{\mathcal{A}}(S_f; P_n) = (a+h)h + \cdots + (a+nh)h = h\sum_{k=1}^{n}(a+kh) = hna + h^2\sum_{k=1}^{n} k$$

Tenendo conto della formula $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ e di $h = \frac{b-a}{n}$.

$$\underline{\mathcal{A}}(S_f; P_n) = hna + h^2\frac{n(n-1)}{2} = (b-a)a + (b-a)^2\frac{n(n-1)}{2n^2}$$

$$\overline{\mathcal{A}}(S_f; P_n) = hna + h^2\frac{n(n+1)}{2} = (b-a)a + (b-a)^2\frac{n(n+1)}{2n^2}.$$

Per $n \to +\infty$, le due quantità tendono allo stesso limite:

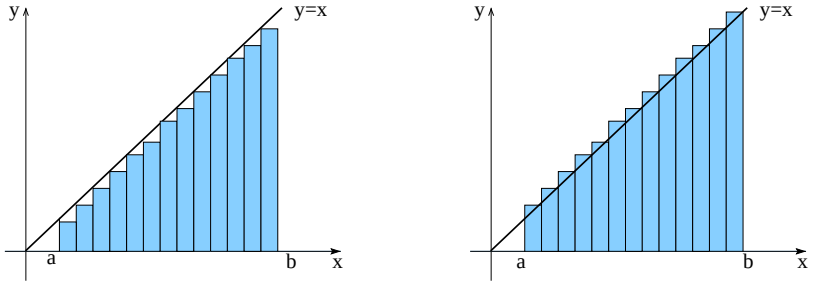



FIGURA 4. La funzione $f(x) = x$ e le approssimazioni determinate dai punti $a + \frac{k(b-a)}{n}$ con $k = 0,\dots,n$.

$$\lim_{n\to+\infty} \underline{A}(S_f; P_n) = \lim_{n\to+\infty} \overline{A}(S_f; P_n) = (b-a)a + \frac{(b-a)^2}{2} = \frac{b^2}{2} - \frac{a^2}{2},$$

che concorda con le formule note dalla geometria elementare (meno male!).

ESEMPIO 1.4. Passiamo a considerare $f(x) = x^2$ in $[a,b]$ con $0 \le a < b$. Stavolta la geometria elementare non ci aiuta. Qual è l'espressione dell'area per il sottografico di $x^2$?

Scegliamo la stessa partizione $P_n$ di prima. Dato che $x^2$ è crescente su $[a,b]$ per $a \ge 0$,

$$\alpha_i = \inf_{[x_{k-1},x_k]} x^2 = \big(a + (k-1)h\big)^2, \quad \beta_i = \sup_{[x_{k-1},x_k]} x^2 = \big(a + kh\big)^2,$$

da cui segue

$$\underline{\mathcal{A}}(S_f;P_n) = a^2h + (a+h)^2h + (a+2h)^2h + \cdots + (a+(n-1)h)^2h$$
$$\overline{\mathcal{A}}(S_f;P_n) = (a+h)^2h + (a+2h)^2h + \cdots + (a+nh)^2h$$

Svolti i quadrati e tenuto presente che $h = \frac{b-a}{n}$, si ottiene

$$\begin{aligned}\underline{\mathcal{A}}(S_f;P_n) &= h\left\{na^2 + 2ah\sum_{k=0}^{n-1}k + h^2\sum_{k=0}^{n-1}k^2\right\}\\ &= (b-a)\left\{a^2 + \frac{2a(b-a)}{n^2}\sum_{k=0}^{n-1}k + \frac{(b-a)^2}{n^3}\sum_{k=0}^{n-1}k^2\right\}\\ \overline{\mathcal{A}}(S_f;P_n) &= h\left\{na^2 + 2ah\sum_{k=1}^{n}k + h^2\sum_{k=1}^{n}k^2\right\}\\ &= (b-a)\left\{a^2 + \frac{2a(b-a)}{n^2}\sum_{k=1}^{n}k + \frac{(b-a)^2}{n^3}\sum_{k=1}^{n}k^2\right\}\end{aligned}$$

Utilizzando le formule (sapete dimostrarle?)

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

e sostituendo nelle precedenti,

$$\underline{\mathcal{A}}(S_f;P_n) = (b-a)\left\{a^2 + a(b-a)\,\frac{n-1}{n} + (b-a)^2\,\frac{(n-1)(2n-1)}{6n^2}\right\}$$
$$\overline{\mathcal{A}}(S_f;P_n) = (b-a)\left\{a^2 + a(b-a)\,\frac{n+1}{n} + (b-a)^2\,\frac{(n+1)(2n+1)}{6n^2}\right\}$$

Quindi, passando al limite per $n \to \infty$,

$$\lim_{n\to+\infty}\underline{\mathcal{A}}(S_f;P_n) = \lim_{n\to+\infty}\overline{\mathcal{A}}(S_f;P_n) = (b-a)a^2 + a(b-a)^2 + \frac{(b-a)^3}{3} = \frac{b^3}{3} - \frac{a^3}{3} =: \mathcal{A}(S_f).$$

ESEMPIO 1.5. Sperimentiamo la tecnica per una funzione più bizzarra:

$$\textit{funzione di Dirichlet}: \qquad D(x) := \begin{cases} 0 & x \in [0,1]\setminus\mathbb{Q},\\ 1 & x \in [0,1]\cap\mathbb{Q},\end{cases}$$

Data una qualsiasi partizione $P = \{a = x_0 < x_1 < \cdots < x_n = b\}$, in ognuno dei sottointervalli $[x_{k-1}, x_k]$ cadono sia numeri razionali che numeri irrazionali, quindi

$$\alpha_i = \inf_{[x_{k-1},x_k]} D(x) = 0, \qquad \beta_i = \sup_{[x_{k-1},x_k]} D(x) = 1.$$

Pertanto, *per ogni partizione* $P$, $\underline{\mathcal{A}}(S_D;P) = 0$ e $\overline{\mathcal{A}}(S_D;P) = 1$ e quindi

$$\sup_P \underline{\mathcal{A}}(S_D;P) = 0 < 1 = \inf_P \overline{\mathcal{A}}(S_D;P).$$

Anche scegliendo partizioni $P$ con ampiezza sempre più piccola, le stime per difetto e quelle per eccesso restano sempre ben lontane le une dalle altre. L'interpretazione che diamo di questa situazione è che *esistono sotto insiemi del piano a cui non è possibile associare un'area*, ossia, per alcuni insiemi $S$, l'espressione $\mathcal{A}(S)$ non ha senso!

ESEMPIO 1.6. Ancora un esempio: $f(x) = e^x$ in $[a, b]$ con $a < b$. Scegliamo ancora una volta la partizione $P_n = \{x_k = a + kh \ : \ k = 0, \dots, n\}$ dove $h = (b-a)/n$. La funzione $e^x$ è crescente su $[a, b]$ quindi

$$\alpha_i = \inf_{[x_{k-1}, x_k]} e^x = e^{x_{k-1}} = e^a\, e^{(k-1)h}, \quad \beta_i = \sup_{[x_{k-1}, x_k]} e^x = e^{x_k} = e^a\, e^{kh}.$$

Perciò

$$\underline{\mathcal{A}}(S_f; P_n) = e^a\, h\, \left[1 + e^h + e^{2h} + \cdots + e^{(n-1)h}\right] = e^a\, h\, \left[1 + e^h + (e^h)^2 + \cdots + (e^h)^{n-1}\right]$$
$$\overline{\mathcal{A}}(S_f; P_n) = e^a\, h\, \left[e^h + e^{2h} + e^{3h} + \cdots + e^{nh}\right] = e^{a+h}\, h\, \left[1 + e^h + (e^h)^2 + \cdots + (e^h)^{n-1}\right]$$

Dato che $1 + e^h + (e^h)^2 + \cdots + (e^h)^{n-1} = (e^{nh} - 1)/(e^h - 1)$ e $nh = b - a$,

$$\underline{\mathcal{A}}(S_f; P_n) = \frac{e^a\, h\, (e^{b-a} - 1)}{e^h - 1} = \frac{h}{e^h - 1}\, (e^b - e^a)$$
$$\overline{\mathcal{A}}(S_f; P_n) = \frac{e^{a+h}\, h\, (e^{b-a} - 1)}{e^h - 1} = \frac{h}{e^h - 1}\, e^h\, (e^b - e^a)$$

Quindi, passando al limite per $n \to \infty$,

$$\lim_{n \to +\infty} \underline{\mathcal{A}}(S_f; P_n) = \lim_{n \to +\infty} \overline{\mathcal{A}}(S_f; P_n) = e^b - e^a$$

ESEMPIO 1.7. Come ultimo esempio, calcoliamo l'area del sottografico di $f(x) = \cos x$ nell'intervallo $[0, \pi/2]$. Prima di tutto dimostriamo il seguente Lemma.

LEMMA 1.8. *Vale*

$$\sum_{k=0}^{n} \cos(ky) = \frac{1}{2} \left[ \frac{\sin(n + \frac{1}{2})y}{\sin(\frac{y}{2})} + 1 \right] \qquad \forall\, y \in \mathbb{R}, \quad \forall\, n \in \mathbb{N}. \tag{1.1}$$

**Dimostrazione.** *(Per induzione)* Per $n = 0$, l'identità è banalmente verificata. Supponiamola vera per $n - 1$ e dimostriamola per $n$:

$$\sum_{k=0}^{n} \cos(ky) = \frac{1}{2} \left[ \frac{\sin(n - \frac{1}{2})y}{\sin(\frac{y}{2})} + 1 \right] + \cos(ny) = \frac{1}{2} \left[ \frac{\sin(n - \frac{1}{2})y + 2\cos(ny)\sin(\frac{y}{2})}{\sin(\frac{y}{2})} + 1 \right].$$

Dato che

$$\sin\left(n \pm \frac{1}{2}\right) y = \sin(ny) \cos\left(\frac{y}{2}\right) \mp \cos(ny) \sin\left(\frac{y}{2}\right),$$

ne segue la conclusione. ∎

Per calcolare l'area di

$$S_f = \left\{ (x, y) \,\middle|\, \quad 0 \le x \le \frac{\pi}{2}, \quad 0 \le y \le \cos x \right\},$$

fissato $n \in \mathbb{N}$, scegliamo la partizione $P_n$ determinata dai punti $x_k := \frac{k\pi}{2n}$ per $k = 0, \dots, n$. La funzione $\cos x$ è decrescente in $[0, \pi/2]$, quindi

$$\alpha_i = \inf_{[x_{k-1}, x_k]} \cos x = \cos\left(\frac{k\pi}{2n}\right), \qquad \beta_i = \sup_{[x_{k-1}, x_k]} \cos x = \cos\left(\frac{(k-1)\pi}{2n}\right).$$

Perciò, utilizzando la formula (1.1) con $y = \frac{\pi}{2n}$, ecco che otteniamo

$$\underline{\mathcal{A}}(S_f; P_n) = \frac{\pi}{2n} \sum_{k=1}^{n} \cos\left(\frac{k\pi}{2n}\right) = \frac{\pi}{4n} \left[\frac{\sin(n + \frac{1}{2})\frac{\pi}{2n}}{\sin(\frac{\pi}{4n})} - 1\right] = \frac{\pi}{4n} \left[\frac{\sin(\frac{\pi}{2} + \frac{\pi}{4n})}{\sin(\frac{\pi}{4n})} - 1\right]$$

$$\overline{\mathcal{A}}(S_f; P_n) = \frac{\pi}{2n} \sum_{k=0}^{n-1} \cos\left(\frac{k\pi}{2n}\right) = \frac{\pi}{4n} \left[\frac{\sin(n + \frac{1}{2})\frac{\pi}{2n}}{\sin(\frac{\pi}{4n})} + 1\right] = \frac{\pi}{4n} \left[\frac{\sin(\frac{\pi}{2} + \frac{\pi}{4n})}{\sin(\frac{\pi}{4n})} + 1\right]$$

Dato che $\sin(\frac{\pi}{2} + \alpha) = \cos\alpha$,

$$\underline{\mathcal{A}}(S_f; P_n) = \frac{\pi/4n}{\tan(\pi/4n)} - \frac{\pi}{4n}, \qquad \overline{\mathcal{A}}(S_f; P_n) = \frac{\pi/4n}{\tan(\pi/4n)} + \frac{\pi}{4n}$$

Per $n \to +\infty$, $\lim_{n\to+\infty} \underline{\mathcal{A}}(S_f; P_n) = \lim_{n\to+\infty} \overline{\mathcal{A}}(S_f; P_n) = 1$ e quindi $\mathcal{A}(S_f) = 1$.

## 2. Ma cos'è quest'integrale?

Tiriamo le fila di quello che abbiamo fatto fin qui: senza preoccuparci troppo di immergerci nel significato della parola "area di una regione del piano", abbiamo considerato il caso di sottografici di funzioni non negative, proponendo un algoritmo per il calcolo dell'area: approssimare per difetto e per eccesso l'area richiesta tramite aree di sottografici di funzioni costanti a tratti e, facendo tendere l'ampiezza della partizione a 0, ottenere il valore dell'area cercata. Tramite un certo numero di esempi, ci siamo resi conto che il procedimento è ragionevole (si arriva spesso ad un risultato numerico), ma che in alcune situazioni non porta a nessuna conclusione (la funzione di Dirichlet).

Scordiamoci ora del problema dell'area e invertiamo il procedimento: dichiariamo "integrabili" le funzioni per cui il procedimento di approssimazione con ("aree" di sottografici di) funzioni costanti a tratti converge e "non integrabili" quelle per cui non converge. Nel caso in cui la funzione sia integrabile, il valore comune del limite delle approssimazioni per eccesso e di quelle per difetto dà la definizione di integrale definito della funzione. Rispetto al paragrafo precedente, c'è una differenza essenziale:

*le funzioni che consideriamo possono avere segno qualsiasi.*

Sia $f : [a, b] \to \mathbb{R}$ limitata e $P = \{x_0, x_1, \dots, x_n\}$ una partizione di $[a, b]$. Poniamo

$$\alpha_i := \inf_{[x_{i-1}, x_i]} f(x) \qquad \text{e} \qquad \beta_i := \sup_{[x_{i-1}, x_i]} f(x).$$

DEFINIZIONE 2.1. *Somme integrali.* *Si chiamano* SOMMA INTEGRALE PER DIFETTO/PER ECCESSO DI $f$ RELATIVAMENTE ALLA PARTIZIONE $P$, *i valori delle somme*

$$\underline{S}(f; P) := \sum_{i=1}^{n} \alpha_i (x_i - x_{i-1}) \qquad e \qquad \overline{S}(f; P) := \sum_{i=1}^{n} \beta_i (x_i - x_{i-1}).$$

Non fa male sottolineare che per ogni funzione $f$ limitata e per ogni partizione $P$, le somme integrali per difetto e per eccesso sono sempre ben definite dato che tutti i valori $\alpha_i, \beta_i$ sono finiti.

DEFINIZIONE 2.2. *Integrale definito.* *Una funzione* $f : [a,b] \to \mathbb{R}$ *limitata è* INTEGRABILE (SECONDO RIEMANN) *se vale l'uguaglianza*

$$\sup\{\underline{S}(f;P) : P \text{ partizione di } [a,b]\} = \inf\left\{\overline{S}(f;P) : P \text{ partizione di } [a,b]\right\}. \tag{2.1}$$

*Il valore comune è l'*INTEGRALE DEFINITO DI $f$ IN $[a,b]$ *e si indica con*

$$\int_a^b f(x)\,dx.$$

OSSERVAZIONE 2.3. Il simbolo dell'integrale è una variante del simbolo di somma indicato da una lunga $S$ come si usava al tempo di Leibnitz. Il simbolo $dx$ è l'erede della lunghezza dell'intervallo $x_i - x_{i-1}$. L'uso della "d" minuscola ricorda che le approssimazioni migliori dell'integrale di $f$ si ottengono considerando partizioni la cui ampiezza sia molto piccola: $\Delta x := x_i - x_{i-1} \approx dx$.

La lettera usata per indicare la variabile di integrazione è indifferente: al posto di $\int_a^b f(x)\,dx$, si può scrivere $\int_a^b f(t)\,dt$ o $\int_a^b f(u)\,du$ (come nel caso delle sommatorie, in cui il nome dato all'indice è assolutamente ininfluente).

OSSERVAZIONE 2.4. Nella definizione di integrale, non viene fatta nessuna richiesta di positività della funzione. Se la *funzione integranda* $f$ è positiva in $[a,b]$ ed è integrabile, l'integrale coincide con l'area del sottografico di $f$ in $[a,b]$, anzi... ne è la definizione:

$$f \geq 0 \qquad \Rightarrow \qquad \mathcal{A}(\{(x,y) : a \leq x \leq b,\ 0 \leq y \leq f(x)\}) := \int_a^b f(x)\,dx.$$

Se $f$ è negativa in tutto o in parte dell'intervallo, il significato dell'integrale non è più quello di un'area: *l'integrale è somma di termini positivi e negativi, gli uni e gli altri in corrispondenza delle zone in cui il grafico è sopra o sotto l'asse* $x$ (vedi Figura 5). Ad esempio, l'integrale della funzione $x$ nell'intervallo $[-1,1]$ è zero (convincersene con una figura). Non c'è motivo per aversene a male, basta averlo ben chiaro fin dal principio.

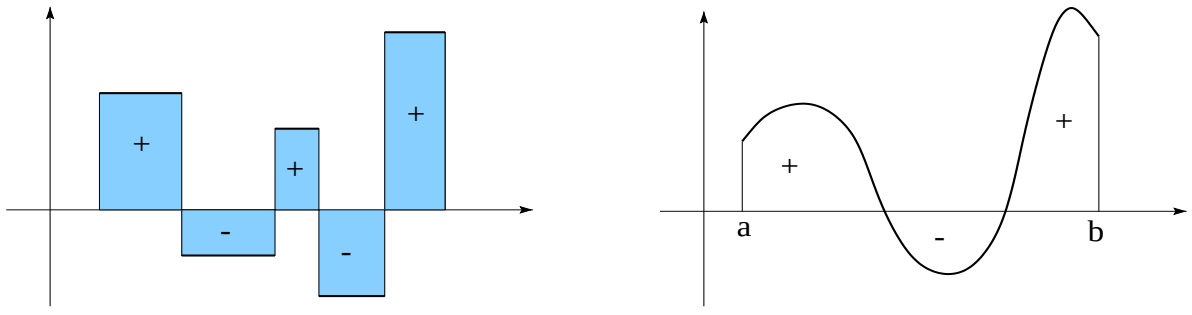



FIGURA 5. L'integrale non è l'area: (a) una funzione costante a tratti, (b) una funzione qualsiasi.

La condizione di integrabilità data dalla definizione è chiara e limpida da un punto di vista di rigore matematico. Meno chiaro è come rispondere concretamente alla domanda: quali classi di funzioni sono integrabili? Ecco il motivo della Proposizione che segue.

PROPOSIZIONE 2.5. *Una funzione $f : [a,b] \to \mathbb{R}$ è integrabile in $[a,b]$ se e solo se*

$$\forall\, \varepsilon > 0 \qquad \exists\, P_\varepsilon \qquad t.c. \qquad \overline{S}(f;P_\varepsilon) - \underline{S}(f;P_\varepsilon) < \varepsilon. \tag{2.2}$$

OSSERVAZIONE 2.6. La definizione di integrabilità e le proprietà dell'estremo superiore e dell'estremo inferiore permettono immediatamente di concludere che per ogni $\varepsilon > 0$, esistono partizioni $P'_\varepsilon$ e $P''_\varepsilon$ per cui $\overline{S}(f;P'_\varepsilon) - \underline{S}(f;P''_\varepsilon) < \varepsilon$. La Proposizione 2.5 garantisce che la stessa stima è vera con la stessa partizione!

Inoltre la condizione espressa in (2.2) è sufficiente per l'integrabilità. Ricordando la definizione di $\overline{S}(f;P_\varepsilon)$ e $\underline{S}(f;P_\varepsilon)$, la condizione (2.2) può essere riscritta

$$\overline{S}(f;P_\varepsilon) - \underline{S}(f;P_\varepsilon) = \sum_{i=1}^{n} (\beta_i - \alpha_i)\,(x_i - x_{i-1}) < \varepsilon$$

dove $\alpha_i$ e $\beta_i$ sono gli stessi definiti in precedenza. Quindi, dimostrare l'integrabilità di una funzione passa per la stima della differenza $\beta_i - \alpha_i$, ovvero per la stima di $|f(y) - f(x)|$ al variare di $x, y$ in $[x_{i-1}, x_i]$, per partizioni con ampiezza sufficientemente piccola.

**Dimostrazione della Proposizione 2.5.**
*Passo 1.* Dimostriamo prima di tutto che

$$\underline{S}(f;P_1) \le \underline{S}(f;P_2) \qquad \text{e} \qquad \overline{S}(f;P_2) \le \overline{S}(f;P_1) \qquad \forall\, P_2 \subset P_1. \tag{2.3}$$

Dato che $P_2 \subset P_1$, si può costruire $P_2$ a partire da $P_1$ aggiungendo un numero finito di punti. Quindi basta studiare il caso in cui $P_2 = P_1 \cup \{\xi\}$. Il caso generale si ottiene iterando un numero finito di volte il procedimento. Inoltre consideriamo solo le somme per difetto, l'altra parte è analoga.

Supponiamo $P_1 = \{x_0, x_1, \dots, x_n\}$ e $\xi \in (x_{k-1}, x_k)$ per un opportuno $k \in \{1, \dots, n\}$. Le espressioni di $\underline{S}(f;P_1)$ e di $\underline{S}(f;P_2)$ coincidono in tutti i termini tranne in quelli relativi all'intervallo $[x_{k-1}, x_k]$ e quindi, indicando con

$$\alpha = \inf_{[x_{k-1},x_k]} f(x), \qquad \alpha' = \inf_{[x_{k-1},\xi]} f(x), \qquad \alpha'' = \inf_{[\xi,x_i]} f(x),$$

vale, dato che $\alpha \le \alpha'$ e $\alpha \le \alpha''$,

$$\begin{aligned}
\underline{S}(f;P_2) - \underline{S}(f;P_1) &= \Big[\alpha'(x_k - \xi) + \alpha''(\xi - x_{k-1})\Big] - \alpha(x_k - x_{k-1}) \\
&= \Big[\alpha'(x_k - \xi) + \alpha''(\xi - x_{k-1})\Big] - \Big[\alpha(x_k - \xi) + \alpha(\xi - x_{k-1})\Big] \\
&= (\alpha' - \alpha)(x_k - \xi) + (\alpha'' - \alpha)(\xi - x_{k-1}) \ge 0,
\end{aligned}$$

*Passo 2.* Ora dimostriamo che

$$\underline{S}(f;P') \le \overline{S}(f;P'') \qquad \forall\, P',\ P'' \text{ partizioni.}$$

Infatti, per costruzione, $\underline{S}(f;P) \le \overline{S}(f;P)$ per ogni partizione $P$. Quindi, se consideriamo la partizione $P = P' \cup P''$ e utilizziamo (2.3):

$$\underline{S}(f;P') \le \underline{S}(f;P) \le \overline{S}(f;P) \le \overline{S}(f;P'').$$

*Passo 3.* Infine, dimostriamo la Proposizione.

Per ogni partizione $P$, vale

$$0 \le \inf_P \overline{S}(f;P) - \sup_P \underline{S}(f;P) \le \overline{S}(f;P) - \underline{S}(f;P)$$

Se per ogni $\varepsilon > 0$ esiste una partizione $P_\varepsilon$ che verifica (2.2), allora

$$0 \le \inf_P \overline{S}(f;P) - \sup_P \underline{S}(f;P) < \varepsilon \qquad \forall \varepsilon > 0,$$

cioè $\inf_P \overline{S}(f;P) = \sup_P \underline{S}(f;P)$, e $f$ è integrabile.

Se invece supponiamo che la funzione $f$ sia integrabile, per ogni $\varepsilon > 0$ esistono due partizioni $P'_\varepsilon$ e $P''_\varepsilon$ per cui vale $\overline{S}(f;P'_\varepsilon) - \underline{S}(f;P''_\varepsilon) < \varepsilon$. Scegliendo $P_\varepsilon = P'_\varepsilon \cup P''_\varepsilon$, grazie a $\underline{S}(f;P''_\varepsilon) \le \underline{S}(f;P_\varepsilon)$ e $\overline{S}(f;P_\varepsilon) \le \overline{S}(f;P'_\varepsilon)$,

$$\overline{S}(f;P_\varepsilon) - \underline{S}(f;P_\varepsilon) \le \overline{S}(f;P'_\varepsilon) - \underline{S}(f;P''_\varepsilon) < \varepsilon,$$

cioè la (2.2). ∎

Se si costruisce una successione di partizioni $P_n$ per cui

$$\lim_{n\to+\infty} \underline{S}(f;P_n) = \lim_{n\to+\infty} \overline{S}(f;P_n) = \ell,$$

a maggior ragione si avrà che: per ogni $\varepsilon > 0$ esiste una partizione $P_n$ per cui vale

$$0 \le \overline{S}(f;P_n) - \underline{S}(f;P_n) = \Big[\overline{S}(f;P_n) - \ell\Big] + \Big[\ell - \underline{S}(f;P_n)\Big] < 2\varepsilon$$

cioè la funzione $f$ è integrabile e l'integrale è proprio il valore $\ell$. In particolare, gli Esempi 1.3, 1.4, 1.6 e 1.7 indicano che le funzioni $x$, $x^2$ e $e^x$ sono integrabili in $[a,b]$, $\cos x$ è integrabile in $[0,\pi/2]$ e che valgono

$$\int_a^b C\,dx = C(b-a), \quad \int_a^b x\,dx = \frac{b^2}{2} - \frac{a^2}{2}, \quad \int_a^b x^2\,dx = \frac{b^3}{3} - \frac{a^3}{3}, \quad \int_0^{\pi/2} \cos x\,dx = 1.$$

L'Esempio 1.5 mostra che *esistono anche funzioni non integrabili!*

## 3. Istruzioni per l'uso

L'integrale che abbiamo appena definito gode di tre proprietà fondamentali, ereditate direttamente dal fatto che un integrale definito, nella sostanza, è il "limite" di sommatorie:

*linearità*, *additività* e *monotonìa*. Prima di descriverle per l'integrale, vediamo di cosa si tratta nel caso delle più familiari sommatorie: dati $a_1, \dots, a_n, b_1, \dots, b_n, \alpha, \beta \in \mathbb{R}$

$$\textit{linearità:} \qquad \sum_{k=1}^{n}(\alpha a_k + \beta b_k) = \alpha \sum_{k=1}^{n} a_k + \beta \sum_{k=1}^{n} b_k$$

$$\textit{additività:} \qquad \sum_{k=1}^{n} a_k = \sum_{k=1}^{m} a_k + \sum_{k=m+1}^{n} a_k$$

$$\textit{monotonìa:} \qquad a_k \le b_k \qquad \Rightarrow \qquad \sum_{k=1}^{n} a_k \le \sum_{k=1}^{n} b_k.$$

Vediamo come si traducono queste tre proprietà nel caso degli integrali definiti. Le dimostrazioni sono posticipate a fine paragrafo.

*Linearità*. *Una combinazione lineare di funzioni integrabile è integrabile: dati $c_1, c_2 \in \mathbb{R}$ e $f, g$ integrabili in $[a, b]$, la funzione $c_1 f + c_2 g$ è integrabile e vale*

$$\int_a^b [c_1 f(x) + c_2 g(x)]\, dx = c_1 \int_a^b f(x)\, dx + c_2 \int_a^b g(x)\, dx \tag{3.1}$$

*Additività*. *Sia $f$ integrabile in $[a, b]$ e sia $c \in (a, b)$. Allora $f$ è integrabile in $[a, c]$ e in $[c, b]$ e vale*

$$\int_a^b f(x)\, dx = \int_a^c f(x)\, dx + \int_c^b f(x)\, dx \qquad \forall\, a < c < b. \tag{3.2}$$

*In particolare, se $f$ integrabile in $[a, b]$, allora è integrabile in ogni sottointervallo di $[a, b]$.*

*Monotonìa*. *Per ogni coppia di funzioni $f, g$ integrabili in $[a, b]$,*

$$f(x) \le g(x) \quad \forall\, x \in [a, b] \quad \Longrightarrow \quad \int_a^b f(x)\, dx \le \int_a^b g(x)\, dx. \tag{3.3}$$

Per ora abbiamo definito $\int_a^b f(x)\, dx$ solo nel caso $a < b$. E' usanza diffusa in tutto il globo terracqueo, *definire l'integrale nel caso di $a = b$ o $a > b$, in modo che sia preservata la regola dell'additività*. Scrivendo (3.2) con $c = a$, si ottiene

$$\int_a^b f(x)\, dx = \int_a^a f(x)\, dx + \int_a^b f(x)\, dx,$$

da cui segue

$$\int_a^a f(x)\, dx := 0,$$

coerente con l'eventuale interpretazione in termini di aree.

Scriviamo (3.2) per $b = a$, allora

$$0 = \int_a^a f(x)\, dx = \int_a^c f(x)\, dx + \int_c^a f(x)\, dx$$

da cui segue la definizione

$$\int_c^a f(x)\,dx := -\int_a^c f(x)\,dx \qquad a < c,$$

dove il membro destro ha il significato definito nel paragrafo precedente. Ad esempio,

$$\int_1^0 x\,dx = -\int_0^1 x\,dx = -\frac{1}{2}.$$

Dalla proprietà di monotonìa dell'integrale discende una proprietà che è, sostanzialmente, una "disuguaglianza triangolare per integrali": *se $f$ è una funzione integrabile in $[a,b]$, allora anche $|f|$ è integrabile in $[a,b]$ e vale*

$$\left|\int_a^b f(x)\,dx\right| \le \int_a^b |f(x)|\,dx. \tag{3.4}$$

Per la dimostrazione, si voli a fine paragrafo.

**Dimostrazioni in libertà.** Dedichiamoci ora a dimostrare le proprietà dell'integrale. Le proprietà di linearità e di additività contengono due parti. Una prima parte concerne il fatto che dall'integrabilità di alcune funzioni se ne deduce l'integrabilità di certe altre. La seconda parte mette in relazione tra loro gli integrali definiti delle varie funzioni.

*Linearità.* Siano $f$ e $g$ integrabili in $[a,b]$ e $c_1, c_2 \in \mathbb{R}$. Data $P = \{x_0, x_1, \dots, x_n\}$, partizione di $[a,b]$, battezziamo

$$\alpha_i = \inf_{[x_{i-1},x_i]} [c_1 f(x) + c_2 g(x)] \qquad \text{e} \qquad \beta_i = \sup_{[x_{i-1},x_i]} [c_1 f(x) + c_2 g(x)]\,,$$

e, analogamente,

$$\alpha_i^f = \inf_{[x_{i-1},x_i]} f(x), \quad \beta_i^f = \sup_{[x_{i-1},x_i]} f(x), \qquad \alpha_i^g = \inf_{[x_{i-1},x_i]} g(x), \quad \beta_i^g = \sup_{[x_{i-1},x_i]} g(x).$$

Per ogni $x, y \in (x_{i-1}, x_i)$, si ha

$$\begin{aligned}[c_1 f(x) + c_2 g(x)] - [c_1 f(y) + c_2 g(y)] &= c_1\,[f(x) - f(y)] + c_2\,[g(x) - g(y)] \\ &\le |c_1| \left[\beta_i^f - \alpha_i^f\right] + |c_2|\,[\beta_i^g - \alpha_i^g]\,,\end{aligned}$$

quindi, passando all'estremo superiore in $x$ e all'estremo inferiore in $y$, si deduce che

$$\beta_i - \alpha_i \le |c_1|\big([\beta_i^f - \alpha_i^f\big) + |c_2|\big(\beta_i^g - \alpha_i^g\big).$$

Perciò, per ogni partizione $P$ di $[a,b]$,

$$\begin{aligned}\overline{S}(c_1 f + c_2 g; P) - \underline{S}(c_1 f + c_2 g; P) &= \sum_{i=1}^n [\beta_i - \alpha_i]\,(x_i - x_{i-1}) \\ &\le |c_1| \sum_{i=1}^n \left[\beta_i^f - \alpha_i^f\right](x_i - x_{i-1}) + |c_2| \sum_{i=1}^n [\alpha_i^g - \beta_i^g]\,(x_i - x_{i-1}) \\ &= |c_1| \left[\overline{S}(f;P) - \underline{S}(f;P)\right] + |c_2| \left[\overline{S}(g;P) - \underline{S}(g;P)\right].\end{aligned}$$

Dato che la funzione $f$ è integrabile, grazie alla Proposizione 2.5, esiste una partizione $P_1$ tale che $\overline{S}(f;P_1) - \underline{S}(f;P_1) < \varepsilon$. Analogamente, esiste una partizione $P_2$ per cui $\overline{S}(g;P_1) - \underline{S}(g;P_1) < \varepsilon$. Scegliendo la partizione $P = P_1 \cup P_2$, entrambe le disequazioni sono soddisfatte e quindi

$$\overline{S}(c_1 f + c_2 g; P) - \underline{S}(c_1 f + c_2 g; P) < (|c_1| + |c_2|)\,\varepsilon.$$

Dalla Proposizione 2.5, segue che la funzione $c_1 f + c_2 g$ è integrabile in $[a,b]$.

Resta da dimostrare la formula (3.1). Data la solita partizione $P = \{x_0, \dots, x_n\}$, scegliamo $\xi_1, \xi_2, \dots, \xi_n$ con $\xi_i \in (x_{i-1}, x_i)$ per $i = 1, \dots, n$. Per definizione di $\overline{S}(f;P)$ e $\underline{S}(f;P)$ valgono

$$\underline{S}(f;P) \le \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) \le \overline{S}(f;P), \qquad \underline{S}(f;P) \le \int_a^b f(x)\,dx \le \overline{S}(f;P),$$

e quindi

$$\left| \int_a^b f(x)\,dx - \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) \right| \le \overline{S}(f;P) - \underline{S}(f;P)$$

Analoghe relazioni valgono sostituendo alla funzione $f$ la funzione $g$ e la funzione $c_1 f + c_2 g$. Sfruttando queste stime si deduce che

$$\begin{aligned}
&\left| \int_a^b [c_1 f + c_2 g]\; dx - c_1 \int_a^b f\,dx - c_2 \int_a^b g\,dx \right| \\
&\qquad \le \left| \int_a^b [c_1 f + c_2 g]\; dx - \sum_{i=1}^{n} [c_1 f(\xi_i) + c_2 g(\xi_i)]\,(x_i - x_{i-1}) \right| \\
&\qquad\quad + |c_1| \left| \int_a^b f\,dx - \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}) \right| + |c_2| \left| \int_a^b g\,dx - \sum_{i=1}^{n} g(\xi_i)(x_i - x_{i-1}) \right| \\
&\qquad \le \left[\overline{S}(c_1 f + c_2 g; P) - \underline{S}(c_1 f + c_2 g; P)\right] \\
&\qquad\quad + |c_1| \left[\overline{S}(f;P) - \underline{S}(f;P)\right] + |c_2| \left[\overline{S}(g;P) - \underline{S}(g;P)\right].
\end{aligned}$$

Leggendo il primo e l'ultimo termine, siamo arrivati a stabilire che, per ogni partizione $P$, vale la disuguaglianza

$$\begin{aligned}
&\left| \int_a^b [c_1 f + c_2 g]\; dx - c_1 \int_a^b f\,dx - c_2 \int_a^b g\,dx \right| \\
&\qquad \le \left[\overline{S}(c_1 f + c_2 g; P) - \underline{S}(c_1 f + c_2 g; P)\right] \\
&\qquad\quad + |c_1| \left[\overline{S}(f;P) - \underline{S}(f;P)\right] + |c_2| \left[\overline{S}(g;P) - \underline{S}(g;P)\right].
\end{aligned}$$

Per dimostrare la formula (3.1), scegliamo un partizione $P$ per cui

$$\overline{S}(f;P) - \underline{S}(f;P) < \varepsilon, \quad \overline{S}(g;P) - \underline{S}(g;P) < \varepsilon, \quad \overline{S}(c_1 f + c_2 g; P) - \underline{S}(c_1 f + c_2 g; P) < \varepsilon;$$

con questa scelta di $P$

$$\left| \int_a^b [c_1 f + c_2 g]\; dx - c_1 \int_a^b f\,dx - c_2 \int_a^b g\,dx \right| \le (1 + |c_1| + |c_2|)\,\varepsilon.$$

Dato che $\varepsilon$ è arbitrariamente piccolo, non c'è scampo: (3.1) è dimostrata. ∎

La linearità garantisce, in particolare, che la somma di funzioni integrabili è integrabile. Lo stesso vale per il prodotto, come enunciato nell'esercizio che segue.

ESERCIZIO 3.1. *Se $f$ e $g$ sono integrabili in $[a,b]$, anche il prodotto $fg$ lo è.*

**Soluzione.** Le funzioni $f$ e $g$ sono limitate, quindi esistono $M_f, M_g > 0$ tali che $|f(x)| \le M_f$ e $|g(x)| \le M_g$ per ogni $x$. Data la partizione $P = \{x_0, x_1, \dots, x_n\}$ indichiamo con

$$\alpha_i^{fg} = \inf_{[x_{i-1},x_i]} f(x)g(x) \qquad \text{e} \qquad \beta_i^{fg} = \sup_{[x_{i-1},x_i]} f(x)g(x),$$

con significati analoghi per $\alpha_i^f, \alpha_i^g, \beta_i^f$ e $\beta_i^g$. Dati $x, y \in (x_{i-1}, x_i)$, si ha

$$f(x)g(x) - f(y)g(y) = \Big[f(x) - f(y)\Big]\, g(x) + f(y)\Big[g(x) - g(y)\Big] \le M_g(\beta_i^f - \alpha_i^f) + M_f(\beta_i^g - \alpha_i^g).$$

Prendendo l'estremo superiore in $x$ e l'estremo superiore in $y$, si ottiene

$$\beta_i^{fg} - \alpha_i^{fg} \le M_g(\beta_i^f - \alpha_i^f) + M_f(\beta_i^g - \alpha_i^g).$$

Perciò

$$\overline{S}(fg;P) - \underline{S}(fg;P) \le M_g\Big[\overline{S}(f;P) - \underline{S}(f;P)\Big] + M_f\Big[\overline{S}(g;P) - \underline{S}(g;P)\Big].$$

Dato che il termine a secondo membro può essere reso arbitrariamente piccolo per $P$ opportuno, la conclusione segue dalla solita Proposizione 2.5. ∎

Senza offesa per nessuno, sottolineiamo qui che non è vero che l'integrale del prodotto è il prodotto degli integrali!

*Additività.* Se si pensa all'idea geometrica di partenza, la proprietà di additività sembra abbastanza naturale: per calcolare l'area possiamo dividere la regione in due parti e sommare i valori delle aree delle due sottoregioni. Ma qui la cosa è diversa: prima di tutto abbiamo una definizione analitica da rispettare e ogni affermazione deve discendere rigorosamente da quella definizione. In più c'è un particolare non banale: chi garantisce che se una funzione è integrabile in $[a,b]$, allora è anche integrabile in $[a,c]$ e $[c,b]$ per $c \in (a,b)$? Seppur ragionevole, quest'affermazione è tutta da verificare.

Per comodità, introduciamo la funzione $\chi_{_E} : \mathbb{R} \to \mathbb{R}$

FUNZIONE CARATTERISTICA DI $E$: $$\chi_{_E}(x) := \begin{cases} 1 & x \in E, \\ 0 & x \notin E. \end{cases}$$

Se l'insieme $E$ è un intervallo, la funzione $\chi_{_E}$ è integrabile. Quindi, se $f : [a,b] \to \mathbb{R}$ è integrabile in $[a,b]$ e $c \in [a,b]$, allora sono integrabili anche le funzioni prodotto $f\chi_{_{[a,c]}}(x)$ e $f\chi_{_{[c,b]}}(x)$ e vale l'uguaglianza

$$f(x) = f(x)\chi_{_{[a,c]}}(x) + f(x)\chi_{_{[c,b]}}(x),$$

Inoltre, dato che moltiplicare per una funzione caratteristica si traduce nel "troncare a zero" la funzione $f$ fuori dall'insieme corrispondente

$$\int_a^b f(x)\chi_{_{[a,c]}}(x)\,dx = \int_a^c f(x)\,dx \qquad \text{e} \qquad \int_a^b f(x)\chi_{_{[c,b]}}(x)\,dx = \int_c^b f(x)\,dx$$

(queste uguaglianze sono meno evidenti di quel che possano sembrare e andrebbero in realtà dimostrate in modo rigoroso...) Quindi, grazie alla linearità,

$$\int_a^b f(x)\,dx = \int_a^b f(x)\chi_{[a,c]}(x)\,dx + \int_a^b f(x)\chi_{[c,b]}(x)\,dx = \int_a^c f(x)\,dx + \int_c^b f(x)\,dx.$$

*Monotonìa.* Dimostrare la monotonìa dell'integrale è particolarmente facile. Se $f(x) \le g(x)$ per ogni $x \in [a,b]$, allora

$$\inf_{x\in[x_{i-1},x_i]} f(x) \le \inf_{x\in[x_{i-1},x_i]} g(x) \qquad \forall [x_{i-1},x_i] \subset [a,b].$$

Quindi, per ogni partizione $P$ di $[a,b]$, vale $\underline{S}(f;P) \le \underline{S}(g;P)$, e passando all'estremo superiore si ottiene la conclusione.

Rimane da dimostrare l'ultima proprietà dichiarata: se $f$ è integrabile in $[a,b]$, allora anche la funzione $|f|$ è integrabile in $[a,b]$ e vale

$$\left|\int_a^b f(x)\,dx\right| \le \int_a^b |f(x)|\,dx.$$

La parte più complicata della dimostrazione sta nel verificare che effettivamente la funzione $|f|$ sia integrabile. Infatti la stima finale segue da $\pm f \le |f|$, che, grazie alla monotonìa dell'integrale, implica

$$\pm \int_a^b f(x)\,dx \le \int_a^b |f(x)|\,dx.$$

Con un minimo di memoria della definizione del modulo, si giunge alla conclusione.

ESERCIZIO 3.2. *Se $f$ è integrabile in $[a,b]$, allora anche $|f|$ lo è.*

**Soluzione.** Il problema è sempre lo stesso: data $P = \{x_0,\dots,x_n\}$ partizione di $[a,b]$ e posti $A_k = \inf_{[x_{k-1},x_k]} |f(x)|$, $B_k = \sup_{[x_{k-1},x_k]} |f(x)|$, quale stima possiamo recuperare per $B_k - A_k$?

Siano $\alpha_k = \inf_{[x_{k-1},x_k]} f(x)$, $\beta_k = \sup_{[x_{k-1},x_k]} f(x)$, dati $x, y$ in $(x_{k-1},x_k)$, si ha

$$\big||f(x)| - |f(y)|\big| \le |f(x) - f(y)| \le \beta_k - \alpha_k,$$

e prendendo l'estremo superiore in $x$ e l'inferiore in $y$ si ottiene $B_k - A_k \le \beta_k - \alpha_k$.

Con questa stima alla mano, si deduce che

$$\overline{S}(|f|;P) - \underline{S}(|f|;P) \le \sum_{k=1}^n (\beta_k - \alpha_k)(x_k - x_{k-1}) = \overline{S}(f;P) - \underline{S}(f;P).$$

e grazie alla Proposizione 2.5 si arriva alla conclusione. □

ESERCIZIO 3.3. *Date $f$ e $g$ integrabili, dimostrare che anche $f_+(x) := \max\{f(x),0\}$, $f_-(x) := -\min\{f(x),0\}$, $\max\{f(x),g(x)\}$ e $\min\{f(x),g(x)\}$ sono integrabili.*

## 4. Classi di funzioni integrabili

Una volta dato senso al concetto di integrale e determinate le proprietà principali, bisogna dedicarsi a determinare un certo numero di funzioni che siano effettivamente integrabili, altrimenti l'oggetto appena definito risulterebbe sostanzialmente inutile. Come si è detto, per dimostrare l'integrabilità di $f$, basta mostrare

$$\forall\, \varepsilon > 0 \qquad \exists\, P_\varepsilon \text{ partizione} \quad \text{tale che} \quad \overline{S}(f; P_\varepsilon) - \underline{S}(f; P_\varepsilon) \leq \varepsilon.$$

Dunque seguiremo questa strategia: fissata una partizione $P$, mostreremo che la differenza $\overline{S}(f;P) - \underline{S}(f;P)$ può essere resa arbitrariamente piccola, a patto di scegliere una partizione $P$ la cui ampiezza $|P|$ sia sufficientemente piccola.

**Integrabilità delle funzioni monotone.** La prima classe che consideriamo è quella delle funzioni monotòne. Ricordiamo un certo numero di proprietà di questa classe di funzioni particolarmente interessanti per quello che vedremo a breve.

TEOREMA 4.1. *Sia $f : [a,b] \to \mathbb{R}$ monotòna, allora è integrabile in $[a,b]$.*

**Dimostrazione.** Si osservi prima di tutto che la monotonìa della funzione $f$ nell'intervallo chiuso $[a,b]$ ne garantisce automaticamente la limitatezza, dato che

$$\min\{f(a), f(b)\} \leq f(x) \leq \max\{f(a), f(b)\} \qquad \forall\, x \in [a,b].$$

Per fissare le idee, supponiamo che $f$ sia una funzione crescente. Data una partizione $P = \{x_0, \ldots, x_n\}$, gli estremi inferiori e superiori di $f$ in $[x_{i-1}, x_i]$ sono

$$\alpha_i = \inf_{[x_{i-1},x_i]} f(x) = f(x_{i-1}) \qquad \text{e} \qquad \beta_i = \sup_{[x_{i-1},x_i]} f(x) = f(x_i).$$

Le somme integrali per difetto e per eccesso sono date da

$$\underline{S}(f;P) = \sum_{i=1}^{n} f(x_{i-1})(x_{i-1} - x_i) \qquad \text{e} \qquad \overline{S}(f;P) = \sum_{i=1}^{n} f(x_i)(x_{i-1} - x_i).$$

e la loro differenza è stimata da

$$\overline{S}(f;P) - \underline{S}(f;P) = \sum_{i=1}^{n} \left[f(x_i) - f(x_{i-1})\right] (x_{i-1} - x_i)$$

Indicando con $|P|$ l'ampiezza di $P$, cioè il massimo delle lunghezze $x_{i-1} - x_i$,

$$\overline{S}(f;P) - \underline{S}(f;P) \leq |P| \sum_{i=1}^{n} \left[f(x_i) - f(x_{i-1})\right] = |P| \left[f(b) - f(a)\right].$$

Per ogni $\varepsilon > 0$, è possibile scegliere $|P|$ sufficientemente piccola, in modo che la differenza $\overline{S}(f;P) - \underline{S}(f;P)$ sia minore di $\varepsilon$, pertanto, grazie (ancora una volta!) alla Proposizione 2.5, la funzione è integrabile. ∎

A partire dalle funzioni monotone e grazie al fatto che combinazioni lineari di funzioni integrabili sono integrabili, è possibile costruire una classe ancora più ampia di funzioni.

Partiamo da un esempio più facile: sia $f$ una funzione definita in $[a,b]$ tale che, per qualche $c \in [a,b]$, la funzione $f$ è crescente in $[a,c]$ e decrescente in $[c,b]$. Allora la funzione $f$ si può riscrivere come differenza di due funzioni crescenti (vedi Figura 6):

$$f = f_1 - f_2$$

dove

$$f_1(x) = \begin{cases} f(x) & x \in [a,c] \\ f(c) & x \in [c,b] \end{cases} \qquad \text{e} \qquad f_2(x) = \begin{cases} 0 & x \in [a,c] \\ f(c) - f(x) & x \in [c,b] \end{cases}$$

Quindi, anche funzioni con un cambio di monotonia sono integrabili. Con una costruzione

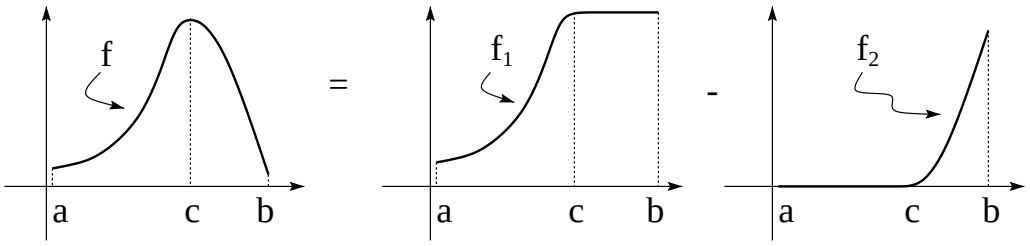



FIGURA 6. Una funzione $f$ con un cambio di monotonia, decomposta come differenza delle funzioni crescenti $f_1$ e $f_2$.

analoga, si mostra che *tutte le funzioni con un numero finito di cambi di monotonia sono integrabili.* Ad esempio, tutti i polinomi e, più in generale, tutte le funzioni razionali sono integrabili.

ESERCIZIO 4.2. *Conoscete una funzione definita in un intervallo limitato che abbia un numero infinito di cambi di monotonìa?*

**Integrabilità delle funzioni lipschitziane.** La seconda classe che consideriamo è quella dele funzioni lipschitziane. Tanto per non sbagliare, ricordiamo che una funzione $f : [a,b] \to \mathbb{R}$ è LIPSCHITZIANA, se esiste $L > 0$ tale che

$$\exists\, L > 0 \qquad \text{t.c.} \qquad |f(x) - f(y)| \le L|x - y| \qquad \forall\, x, y \in [a,b].$$

TEOREMA 4.3. *Sia $f : [a,b] \to \mathbb{R}$ lipschitziana, allora è integrabile in $[a,b]$.*

**Dimostrazione.** Fissata la beneamata partizione $P = \{x_0, x_1 \ldots, x_n\}$,

$$\overline{S}(f;P) - \underline{S}(f;P) = \sum_{i=1}^{n} (\beta_i - \alpha_i)(x_i - x_{i-1}),$$

con il solito significato per $\alpha_i$ e $\beta_i$. Dato che $f$ è lipschitziana, essa è anche continua in $[a,b]$, quindi ammette massimo e minimo in ogni intervallo $[x_{i-1}, x_i]$ e vale

$$\alpha_i = \inf_{(x_{i-1},x_i)} f = \min_{[x_{i-1},x_i]} f(x) = f(\eta_i) \qquad \text{e} \qquad \alpha_i = \sup_{(x_{i-1},x_i)} f = \max_{[x_{i-1},x_i]} f(x) = f(\xi_i)$$

con $\eta_i$ e $\xi_i$, rispettivamente, un punto di massimo ed uno di minimo della funzione $f$ in $[x_{i-1}, x_i]$. Sostituendo nella relazione precedente, si ottiene

$$\overline{S}(f;P) - \underline{S}(f;P) = \sum_{i=1}^{n} (f(\xi_i) - f(\eta_i))(x_i - x_{i-1}) \leq L \sum_{i=1}^{n} |\xi_i - \eta_i|(x_i - x_{i-1}).$$

Dato che $\xi_i, \eta_i \in [x_{i-1}, x_i]$, la differenza $|\xi_i - \eta_i|$ è minore o uguale dell'ampiezza $|P|$ della partizione. Quindi

$$\overline{S}(f;P) - \underline{S}(f;P) \leq L|P| \sum_{i=1}^{n} (x_i - x_{i-1}) \leq L(b-a)|P|.$$

Scegliendo $P$ tale che $|P| < \varepsilon/L(b-a)$ , la differenza $\overline{S}(f;P) - \underline{S}(f;P)$ è strettamente minore di $\varepsilon$ e, grazie alla Proposizione 2.5, segue la conclusione. ∎

Più in generale si può dimostrare il fondamentale

TEOREMA 4.4. *Integrabilità delle funzioni continue. Sia $f : [a,b] \to \mathbb{R}$ una funzione continua, allora $f$ è integrabile in $[a,b]$.*

La dimostrazione passa per un altro risultato (il Teorema di Heine-Cantor) che garantisce l'*uniforme continuità* della $f$. Come dovrebbe intuirsi dai passaggi per il caso lipschitziano, l'uniforme continuità gioca il suo ruolo nella stima (in modo uniforme, appunto!) della differenza $\beta_i - \alpha_i = f(\xi_i) - f(\eta_i)$. Dettagli omessi per carità di patria.

Prima di tirare avanti, facciamo il punto della situazione sulle classi di funzioni che sappiamo essere integrabili:
– le funzioni monotone (in realtà, tutte le funzioni con un numero finito di cambi di monotonia);
– le funzioni lipschitziane (in realtà, tutte le funzioni continue);
– la combinazione lineare di funzioni integrabili;
– il prodotto di funzioni integrabili.

ESERCIZIO 4.5. *Dimostrare che la composizione di una funzione lipschitziana con una funzione integrabile dà luogo ad una funzione integrabile.*

Ve ne viene in mente qualche altra?

## 5. Teorema della media integrale

Tutte le funzioni $f : [a,b] \to \mathbb{R}$ integrabili sono, per definizione, funzioni limitate. In particolare questo significa che data una funzione integrabile, il suo integrale definito può sempre essere stimato in maniera "rude": per le proprietà di monotonìa dell'integrale, se $m \leq f(x) \leq M$ per ogni $x \in [a,b]$,

$$m(b-a) = \int_a^b m\,dx \leq \int_a^b f(x)\,dx \leq \int_a^b M\,dx = M(b-a),$$

Questa formula è intuitivamente ovvia: se pensiamo ad una funzione non negativa e all'integrale come area, le quantità $M(b-a)$ e $m(b-a)$ rappresentano le aree di una rettangolo circoscritto ed inscritto nel sottografico di $f$ (vedi Figura 7).

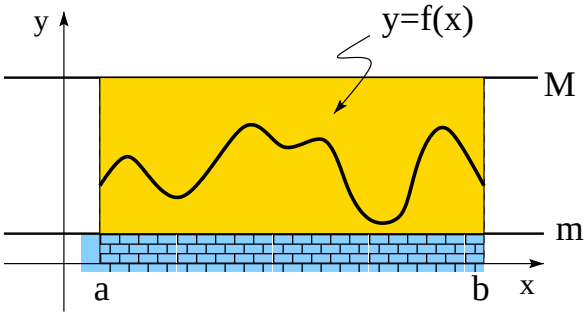



FIGURA 7. Significato geometrico della stima $m(b-a) \le \int_a^b f(x)\,dx \le M(b-a)$.

La formula precedente si può riscrivere come $m \le \langle f \rangle \le M$ dove $\langle f \rangle$, detto MEDIA INTEGRALE DI $f$ IN $[a,b]$, è definito da

$$\langle f \rangle := \frac{1}{b-a}\int_a^b f(x)\,dx. \tag{5.1}$$

Quindi, espresso a parole, *la media integrale di una funzione $f$ è sempre compresa tra un qualsiasi minorante ed un qualsiasi maggiorante di $f$.*

Se $f$ è continua in un intervallo $[a,b]$, si può dire qualcosa di più. Per il teorema di Weierstrass, esistono due punti $\xi, \eta \in [a,b]$ tali che

$$\min_{[a,b]} f(x) = f(\xi) \le f(x) \le f(\eta) = \max_{[a,b]} f(x) \qquad \forall\, x \in [a,b].$$

Quindi $\langle f \rangle$, definita in (5.1) è compresa tra il massimo ed il minimo della funzione $f$ in $[a,b]$. Grazie al Teorema del valore intermedio, si può concludere che *per una funzione $f$ continua in $[a,b]$, la media integrale fa sempre parte dell'insieme immagine $f([a,b])$.*

TEOREMA 5.1. *Teorema della Media Integrale. Sia $f : [a,b] \to \mathbb{R}$ continua. Allora esiste $\xi \in [a,b]$ tale che $f(\xi) = \langle f \rangle$ con $\langle f \rangle$ definito in (5.1).*

Nel caso di una funzione non negativa $f$, il teorema equivale ad affermare che esiste un rettangolo di base $[a,b]$ ed una altezza $f(\xi)$ opportuna con la stessa area del sottografico di $f$ in $[a,b]$.

Non è particolarmente sconvolgente (ma più avanti servirà) osservare che la formula (5.1) vale anche nel caso $b < a$, infatti

$$\frac{1}{b-a}\int_a^b f(x)\,dx = \frac{1}{a-b}\int_b^a f(x)\,dx,$$

e a quest'ultimo termine si può applicare il Teorema della Media Integrale.

CONTROESEMPIO 5.2. Se la funzione $f$ non è continua in tutto $[a,b]$, non è detto che valga (5.1). Ad esempio, si consideri la funzione

$$f(x) = \operatorname{sgn} x = \begin{cases} -1 & x < 0, \\ 0 & x = 0, \\ +1 & x > 0, \end{cases}$$

nell'intervallo $[-1,2]$. Allora $f$ è integrabile (per almeno due motivi... quali?)

$$\mu = \frac{1}{2-(-1)} \int_{-1}^{2} \operatorname{sgn} x \, dx = \frac{1}{3}(-1+2) = \frac{1}{3},$$

che non fa parte dell'immagine della funzione $\operatorname{sgn} x$.

La *media aritmetica* dei numeri $f_1, f_2, \ldots, f_n$ è, per definizione $(f_1 + f_2 + \cdots + f_n)/n$. La media integrale di una funzione in un intervallo è, in un certo senso, limite di medie aritmetiche di valori assunti dalla funzione. Eccone la genesi. Data la solita funzione $f$ definita in $[a,b]$, consideriamo la partizione di $[a,b]$ in sottointervalli di lunghezza $\Delta x_i = (b-a)/n$ per ogni $i$. Scelti $\xi_i \in (x_{i-1}, x_i)$, la media aritmetica di $f(\xi_1), \ldots, f(\xi_n)$ è pari a

$$\frac{f(\xi_1) + \cdots + f(\xi_n)}{n} = \frac{1}{b-a} \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}).$$

Per $n \to \infty$, il termine a destra converge proprio alla media integrale di $f$ in $[a,b]$.

ESERCIZIO 5.3. *Siano $f \in C([a,b])$ e $p \in C([a,b])$ tale che $p(x) > 0$ per ogni $x$. Dimostrare che esiste $\xi \in [a,b]$ tale che*

$$f(\xi) = \frac{\int_a^b f(x)p(x)\,dx}{\int_a^b p(x)\,dx}.$$

**Soluzione.** Se $m$ e $M$ indicano il minimo ed il massimo di $f$ in $[a,b]$, grazie al fatto che $p(x) > 0$, si ha $mp(x) \le f(x)p(x) \le Mp(x)$ per ogni $x$. Integrate in $[a,b]$, mescolate e concludete... □

## 6. Funzioni definite tramite integrali

Una volta fissata la funzione $f$, l'integrale definito dipende dall'intervallo di integrazione o, detto in altri termini, è una funzione degli estremi di integrazione $a$ e $b$. Per studiare questa dipendenza, supponiamo l'estremo inferiore fissato al valore $a$ e indichiamo l'estremo superiore (variabile) con $x$: consideriamo, quindi, la FUNZIONE INTEGRALE

$$\phi(x) = \int_a^x f(t)\,dt. \tag{6.1}$$

La formula (6.1) può essere utilizzata per "generare nuove funzioni" a partire da una funzione integrabile $f$. Ad esempio, si può definire

$$\phi(x) := \int_1^x \frac{1}{t}\,dt \qquad x > 0,$$

che ha perfettamente senso, dato che la funzione $\frac{1}{t}$ è una funzione continua in ogni intervallo $[a,b]$ con $0 < a \le b$. Tale funzione, come vedremo, si dimostra essere uguale

al logaritmo naturale e la formula precedente può quindi essere scelta come definizione analitica di $\ln x$. Oppure provate, ad esempio, a considerare

$$\phi(x) = \int_0^x [t]dt$$

essendo $[t]$ la funzione *parte intera*, disegnando il grafico di $\phi(x)$. O tante altre. Il fatto importante è che, in molte situazioni di quotidiana normalità, compare all'orizzonte una funzione che non è data da un polinomio, o da una funzione razionale, o quant'altro, ma è una funzione integrale. E con un oggetto del genere si vorrebbe lavorare così come si lavora con altre funzioni più familiari. Ad esempio, una funzione integrale è continua?

*Se $f$ è integrabile in $[a,b]$, la funzione $\phi$, definita in (6.1), è lipschitziana.*

Infatti, dato che $f$ è integrabile, essa è anche limitata. Sia $M > 0$ tale che $|f(t)| \le M$ per ogni $t \in [a,b]$, allora,

$$\text{se } x < y, \quad |\phi(x) - \phi(y)| = \left| \int_x^y f(t)\,dt \right| \le \int_x^y |f(t)|\,dt \le M(y-x);$$

$$\text{se } x > y, \quad |\phi(x) - \phi(y)| = \left| -\int_y^x f(t)\,dt \right| \le \int_y^x |f(t)|\,dt \le M(x-y).$$

Quindi

$$|f(t)| \le M \qquad \Longrightarrow \qquad |\phi(x) - \phi(y)| \le M|x-y|.$$

In particolare, se $f$ è integrabile, la funzione $\phi$ è una funzione continua.

ESERCIZIO 6.1. *Consideriamo la funzione $f(x) = \operatorname{sgn} x$ che è continua tratti (ha un punto di salto in $x = 0$). Chi è la funzione*

$$\phi(x) = \int_0^x sgn\,(t)\,dt?$$

**Soluzione.** Dato che

$$x > 0 \quad \Rightarrow \quad \phi(x) = \int_0^x \operatorname{sgn}(t)\,dt = \int_0^x 1\,dt = x$$

$$x < 0 \quad \Rightarrow \quad \phi(x) = \int_0^x \operatorname{sgn}(t)\,dt = \int_0^x (-1)\,dt = -x,$$

la funzione $\phi(x)$ è la funzione modulo: $\phi(x) = |x|$. Sorpresi? E se definiamo

$$\psi(x) = \int_1^x \operatorname{sgn}(t)\,dt,$$

cosa otteniamo? □

ESERCIZIO 6.2. *Sia $f : \mathbb{R} \to \mathbb{R}$ integrabile in ogni intervallo $[a,b]$ di $\mathbb{R}$ e sia*

$$\phi(x) := \int_0^x f(t)\,dt$$

*(i) Dimostrare che, se $f(x) \ge 0$ per ogni $x \in \mathbb{R}$, allora $\phi$ è non decrescente.*
*(ii) Disegnare qualitativamente il grafico di $\phi$ nel caso in cui $xf(x) > 0$ per ogni $x \ne 0$.*

## Appendice A. Quale senso ha la parola "area"?

Il capitolo è iniziato discutendo la questione del calcolo dell'area del sottografico di una funzione non negativa. Vale la pena di ritornare sul problema del significato dell'area delle regioni di piano, senza intenzione di entrare nel dettaglio tecnico della questione.

Dalla geometria elementare è noto che a certe figure piane è associato un numero, detto *area*. Ma che senso ha la parola "area" per domini del piano del tutto generali? E come calcolarla nel caso di figure che non rientrino in quella misera classe di regioni di cui (per tradizione) ci è stata tramandata la formula dell'area? Occorre cominciare da qualche parte e il punto d'avvio più ragionevole è accettare, come legge assiomatica o come prescrizione giuridica (fate voi), che

*l'area di un rettangolo è data dalla famosa formula "base per altezza".*

Mettiamoci d'accordo con i nomi: un insieme $R \subset \mathbb{R}^2$ è un RETTANGOLO se è il prodotto cartesiano di due intervalli limitati. Quindi sono rettangoli oggetti del tipo

$$[a,b] \times [c,d], \quad (a,b) \times (c,d), \quad [a,b) \times (c,d], \quad (a,b] \times (c,d), \quad \cdots,$$

dove $a \leq b$ e $c \leq d$. Per definizione, stabiliamo che *l'area* $\mathcal{A}(R)$ *di un rettangolo* $R$ *è il prodotto della lunghezza dei due intervalli da cui è definito.* Ad esempio,

$$\mathcal{A}\big([a,b] \times [c,d]\big) = \mathcal{A}\big((a,b) \times (c,d)\big) = (b-a)(d-c).$$

Occorre subito fare alcune osservazioni su quanto stabilito fin qui:

– abbiamo ammesso come rettangoli solo quelli con lati paralleli agli assi coordinati, quindi per ora non abbiamo nessuna informazione sull'area di rettangoli (nel senso naïf del termine) con lati obliqui;

– tra i rettangoli abbiamo ammesso anche oggetti "degeneri": segmenti orizzontali e verticali nel piano e persino l'insieme vuoto. Dalla definizione di $\mathcal{A}(R)$ segue che ciascuno di questi "rettangoli degeneri" ha area zero;

– l'area del rettangolo non varia se si aggiungono o si tolgono dei lati, ad esempio i rettangoli $[0,1] \times [0,1]$, $(0,1) \times (0,1)$, $(0,1) \times (0,1]$, $[0,1] \times (0,1]$ hanno tutti la stessa area.

Ora introduciamo una seconda regola per le aree:

*se una regione limitata del piano è divisa in più parti (senza sovrapposizioni), l'area di tutta la regione è la somma delle aree delle singole parti.*

Grazie a questa legge, la parola area ha senso anche per regioni del piano che sono unione finita di rettangoli. Diamo loro un nome: un insieme $R \subset \mathbb{R}^2$ è un PLURIRETTANGOLO se è unione finita di rettangoli. Un plurirettangolo $R$ può sempre essere decomposto come

unione finita di rettangoli disgiunti, cioè esistono $N$ rettangoli $R_i$ tali che

$$R = \bigcup_{i=1}^{N} R_i \qquad \text{con} \quad R_i \cap R_j = \emptyset \quad \forall\, i \neq j.$$

L'area di un plurirettangolo siffatto è definita dalla somma delle aree dei rettangoli $R_i$

$$\mathcal{A}(R) = \sum_{i=1}^{N} \mathcal{A}(R_i).$$

Al punto a cui siamo arrivati, abbiamo una classe abbastanza ampia (tutti i plurirettangoli del piano) di insiemi di cui sappiamo dire cosa è l'area. Come comportarsi per una regione $S$ più generale? L'idea è molto antica ed è, nella sostanza, quella del *metodo di esaustione* (proposto già dall'antichità): approssimare la regione $S$ tramite plurirettangoli contenuti in $S$ e plurirettangoli contenenti $S$ e poi... cercare la migliore approssimazione possibile!

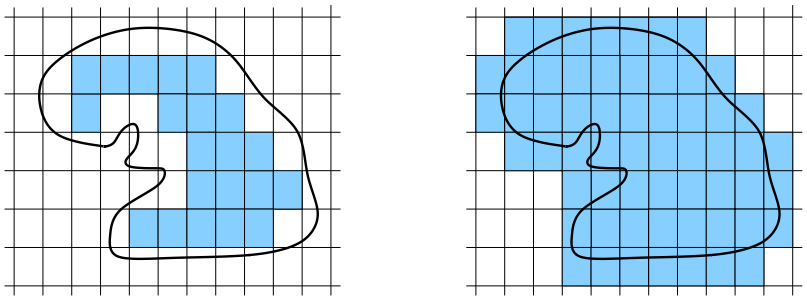


FIGURA 8. La regione $S$ e due plurirettangoli approssimanti dall'interno e dall'esterno.

DEFINIZIONE 6.3. *Sia $S \subset \mathbb{R}^2$ un insieme del piano contenuto in un rettangolo. L'insieme $S$ è* MISURABILE *se*

$$\sup\{\mathcal{A}(R) \,:\, R \textit{ plurirett.}, R \subset S\} = \inf\{\mathcal{A}(R) \,:\, R \textit{ plurirett.}, S \subset R\}.$$

*In questo caso, il valore comune si indica con $\mathcal{A}(S)$ ed è detto* AREA DI $S$.

L'idea è più che ragionevole: se $\mathcal{A}(S)$ ha senso, la migliore approssimazione del valore $\mathcal{A}(S)$ dall'interno è data dall'estremo superiore delle aree dei plurirettangoli contenuti in $S$, e la migliore approssimazione di $\mathcal{A}(S)$ dall'esterno è data dall'estremo inferiore delle aree dei plurirettangoli contenenti $S$. Tanto vale prendere la proprietà di coincidenza di questi due numeri come definizione di misurabilità, cioè di esistenza dell'area!

Come si può far diminuire l'errore di un'approssimazione dall'interno e/o dall'esterno? Semplice, basta considerare plurirettangoli composti da rettangoli di lato sempre più piccolo. In un certo senso si vuole compiere un'operazione di "limite", facendo tendere la lunghezza dei lati dei singoli rettangoli a zero (vedi Figura 9). Per ricollegare queste definizioni a quanto fatto ad inizio capitolo per i sottografici di funzioni positive, bisogna dimostrare che

$$\sup\{\mathcal{A}(R) \,:\, R \text{ plurirett.}, R \subset S_f\} = \sup\{\underline{\mathcal{A}}(S_f, P) \,:\, P \text{ partizione}\},$$

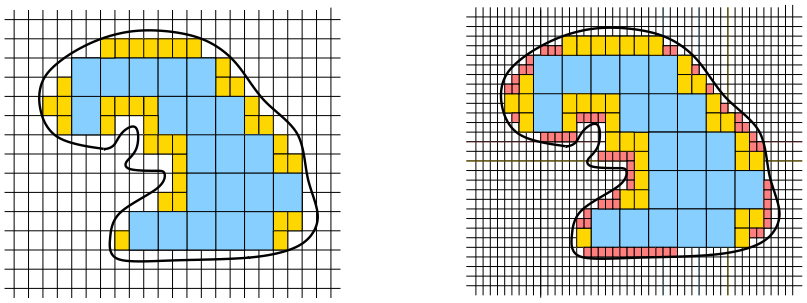


FIGURA 9. Plurirettangoli composti da rettangoli di lato più piccolo danno approssimazioni migliori.

$$\inf\{\mathcal{A}(R) \,:\, R \text{ plurirett.}, S_f \subset R\} = \inf\{\overline{\mathcal{A}}(S_f, P) \,:\, P \text{ partizione}\}.$$

Qualche disegno vi convincerà della validità di queste uguaglianze. Provateci.

## Appendice B. Il volume di solidi di rotazione

L'integrale è utile per il calcolo delle aree, ma anche, in alcune situazioni speciali, per il calcolo di volumi di solidi. Consideriamo una funzione $f : [a, b] \to [0, +\infty)$ continua e disegniamone il grafico $\Gamma_f$. Se si fa ruotare il grafico $\Gamma_f$ attorno all'asse $x$, si ottiene la superficie laterale di un solido $\Sigma$. Qual'è il suo volume? Come per il calcolo delle aree, anche nel calcolo dei volumi bisogna partire da una formula nota. Qui, diamo per buona la regola che ci è stata insegnata da bambini: il volume di un cilindro di raggio di base $r$ e altezza $h$ è $\pi r^2 h$. Se volessimo (ma non è questa la sede!) sviluppare una teoria completa dei volumi di solidi bisognerebbe giustificare in maniera più convincente il senso di questa affermazione.

Ragionando in maniera simile a quanto fatto per la determinazione dell'area di un sottografico, approssimiamo il solido con l'unione di oggetti di cui conosciamo il volume. La scelta più ragionevole è l'unione di cilindri ottenuti tramite una rotazione di rettangoli con lati paralleli agli assi $x$ e $y$ come in Figura 10(b) Un'approssimazione di questo genere

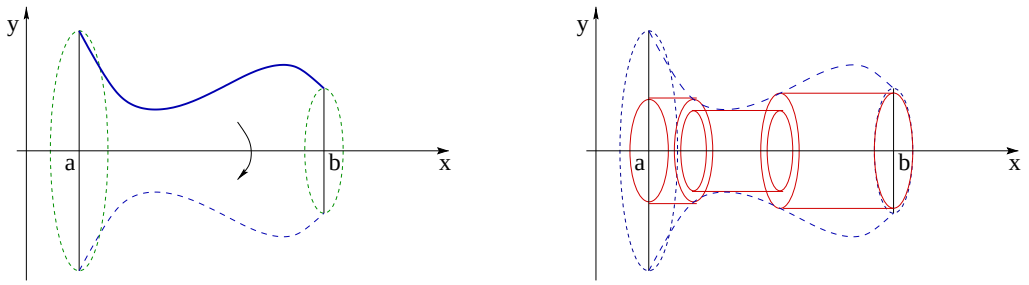



FIGURA 10. (a) Un solido di rotazione; (b) Una sua approssimazione con cilindri.

di $\Sigma$, si ottiene tramite la scelta di una partizione $P = \{a \equiv x_0 < x_1 < \cdots < x_n = b\}$ dell'intervallo $[a, b]$ e di $n$ punti $\xi_1, \ldots, \xi_n$ tali che $\xi_i \in [x_{i-1}, x_i]$. Quindi l'approssimazione del volume $\mathcal{V}(\Sigma)$ di $\Sigma$ è data da

$$\mathcal{V}(\Sigma) \approx \sum_{i=1}^{n} \pi f^2(\xi_i)(x_i - x_{i-1}).$$

Volendo si potrebbe scegliere $\xi_i$ in modo che $f(\xi_i) = \min_{[x_{i-1}, x_i]} f(x)$ ed ottenere in questo modo una stima per difetto, e, analogamente, una stima per eccesso. Che succede quando l'ampiezza della partizione tende a zero? Semplice, l'errore di approssimazione tende a zero e si ottiene così la formula desiderata per il volume di $\Sigma$

$$\mathcal{V}(\Sigma) = \pi \int_a^b f^2(x)\, dx.$$

Collaudiamo la formula con un paio di esempi.

ESEMPIO 6.4. **Il volume del cono.** Un cono di altezza $h$ e raggio di base $r$, si può ottenere tramite una rotazione del grafico della funzione

$$f(x) = \frac{rx}{h} \qquad x \in [0, h].$$

Quindi

$$\mathcal{V}(\Sigma) = \pi \int_0^h \left(\frac{rx}{h}\right)^2 dx = \frac{\pi\, r^2}{h^2} \int_0^h x^2\, dx = \frac{\pi\, r^2}{h^2}\, \frac{h^3}{3} = \frac{1}{3}\, \pi\, r^2\, h,$$

che è la tradizionale formula del volume del cono.

ESEMPIO 6.5. **Il volume della sfera.** Una sfera di raggio $r$ si ottiene con la rotazione del grafico di

$$f(x) = \sqrt{r^2 - x^2} \qquad x \in [-r, r].$$

Applicando la formula del volume

$$\mathcal{V}(\Sigma) = \pi \int_{-r}^r \left(\sqrt{r^2 - x^2}\right)^2 dx = \pi \int_{-r}^r \left(r^2 - x^2\right) dx = \pi \left\{2r^3 - \frac{2}{3} r^3\, dx\right\} = \frac{4}{3}\pi r^3.$$

Anche questa volta, giustamente, i conti tornano!

Il problema fondamentale a questo punto è che conosciamo le formule di integrazione di poche funzioni e, di conseguenza, non possiamo utilizzare la formula per il volume dei solidi di rotazione in molte altre situazioni. Ma si tratta solo di pazientare: nell'arco di qualche Capitolo, grazie alla definizione di derivata ed al Teorema fondamentale del calcolo integrale, saremo in grado di calcolare esplicitamente gli integrali definiti di una classe ben più ampia di funzioni.

ESERCIZIO 6.6. *Qual'è il volume del solido ottenuto facendo ruotare il grafico di* $e^x$ *per* $x \in [0, 1]$ *attorno all'asse* $x$*?*

**Soluzione.** La formula che abbiamo visto dice che il volume del solido $\Sigma$ è

$$\mathcal{V}(\Sigma) = \pi \int_0^1 e^{2x}\, dx.$$

Per calcolare l'integrale si può procedere nello stesso modo utilizzato nell'Esempio 1.6. Sia $f(x) = e^{2x}$. Dato che sappiamo già che $f$ è integrabile in $[0, 1]$ basta calcolare il limite per $n \to +\infty$ di uno solo tra $\underline{\mathcal{A}}(S_f; P_n)$ e $\overline{\mathcal{A}}(S_f; P_n)$. Consideriamo il primo dei due.

Scegliendo la partizione $P_n = \{x_k = kh \,:\, k = 0, \ldots, n\}$ con $h = 1/n$,

$$\underline{\mathcal{A}}(S_f; P_n) = h\, \left[1 + e^{2h} + (e^{2h})^2 + \cdots + (e^{2h})^{n-1}\right] = \frac{h\,(e^2 - 1)}{e^{2h} - 1} = \frac{2h}{e^{2h} - 1} \cdot \frac{e^2 - 1}{2}.$$

Per $n \to \infty$,

$$\int_0^1 e^{2x}\, dx = \lim_{n \to +\infty} \underline{\mathcal{A}}(S_f; P_n) = \frac{1}{2}\,(e^2 - 1),$$

e quindi il volume è $\frac{\pi}{2}\,(e^2 - 1)$. □

CAPITOLO 6

# Derivate, derivate e derivate

Il motivo principale per cui la derivazione è un'operazione fondamentale è che essa permette di controllare le variazioni di una funzione assegnata. Supponiamo di esaminare l'evoluzione di una quantità descritta, all'istante $t$, dal numero reale $y = f(t)$. Ad esempio, la funzione $f$ può rappresentare la posizione di un punto in movimento lungo l'asse $y$; ma si potrebbe trattare anche della temperatura dell'acqua messa sul fuoco per preparare un ottimo piatto di pasta. O altro. 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

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

Nel caso banale in cui $f$ 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 $\Delta f$ è lineare nell'incremento $\Delta t$ della variabile indipendente $t$, cioè $\Delta f$ è proporzionale a $\Delta t$. La costante di proporzionalità è $A$ ed è data da $A = \Delta f/\Delta t$. Quest'ultima formula indica che, se è noto *a priori* che la funzione $f$ è affine, ma non si sa di quale specifica funzione affine si tratti, per determinare il valore della costante $A$ basta compiere due misurazioni di $f$ (agli istanti $t_0$ e $t$) e calcolare il rapporto degli incrementi $\Delta f$ e $\Delta t$.

Difficilmente saremo così fortunati da lavorare sempre con funzioni $f$ affini. Cosa succede nel caso generale? 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 $\Delta f$ non è lineare in $\Delta t$, dato che compare il termine quadratico $A\,(\Delta t)^2$. Però $\Delta f$ 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 $\Delta t$:

| $\Delta t$ | 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 $\Delta f$ tramite una funzione lineare in $\Delta t$:

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

Questo per un polinomio di secondo grado. In generale, data $f$ qualsiasi, se è possibile scrivere l'incremento $\Delta f$ 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 $\Delta f$ della funzione $f$ quando la variabile indipendente $t$ subisca una variazione $\Delta t$ 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. Nel caso del polinomio di secondo grado il conto è facile: $R = a\,(\Delta t)^2$, quindi

$$\lim_{\Delta t \to 0} \frac{R}{\Delta t} = \lim_{\Delta t \to 0} \frac{a\,(\Delta t)^2}{\Delta t} = \lim_{\Delta t \to 0} a\Delta t = 0.$$

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 $\lim_{\Delta t \to 0} R/\Delta t = 0$, allora, dividendo per $\Delta t$,

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

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. Ottimo!

**La derivata come velocità.** Consideriamo un punto che si muova lungo l'asse $y$. Il moto del punto è noto se si conosce ad ogni istante $t$ la posizione $y = f(t)$. 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*.

## 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*

$$\lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}. \tag{1.1}$$

*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 (l'importante è capirsi!):

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

e il limite (1.1) 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 tende ad

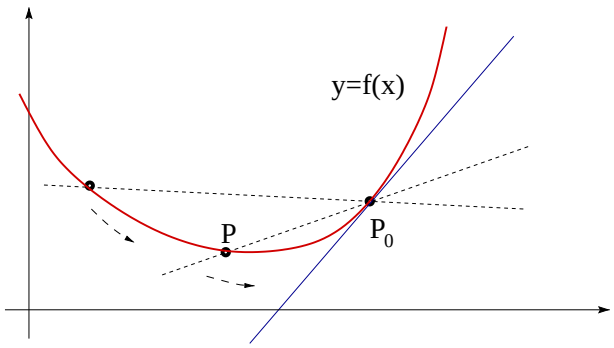



FIGURA 1. Il grafico di una funzione con tangente e secanti.

una posizione limite, tale retta limite è la RETTA TANGENTE. Formuliamo ora il processo geometrico di limite che abbiamo appena raccontato in modo rigoroso. Il coefficiente angolare della retta secante per $P_0$ e $P$ è

$$\text{RAPPORTO INCREMENTALE:} \qquad \frac{f(x) - f(x_0)}{x - x_0}.$$

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 $P_0 = (x_0, f(x_0))$.

> *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..., quindi non si parlerà mai di derivata, cosa del resto suggerita dall'idea di cercare la retta tangente: chi penserebbe di fare la tangente in un singolo punto? Ad esempio, la funzione $\sqrt{x-\sqrt{x}}$ è definita nel maxi intervallo $[1,+\infty)$ e nel povero singolo punto $x_0 = 0$. Sarà derivabile nel punto $x_0 = 0$ ? È una domanda priva di senso!*

La definizione analitica di derivata è chiara: si tratta semplicemente del limite di una funzione opportuna, il rapporto incrementale. Meno chiaro è come si possa calcolare *esplicitamente* la funzione derivata. In generale, non è possibile calcolare tale funzione semplicemente ponendo l'incremento $h = 0$, dato che questo vorrebbe dire "dividere per zero"! Partiamo da due casi semplici:

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

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

Nel caso di $f(x) = x^2$, si ha

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

Consideriamo $f(x) = \sqrt{x}$ per $x > 0$. Il rapporto incrementale 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*: pur essendo definita e continua, si ha

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

Provate a disegnare i grafici di $\sqrt{x}$ e di $\sqrt{|x|}$ (notate che le rette tangenti a questi due grafici in $x = 0$ sono verticali).

Ci sono funzioni non derivabili con comportamenti ben diversi da questo. Ad esempio

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

sono entrambe continue, ma non derivabili in 0. 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 $\lim_{x\to x_0}\big(f(x)-f(x_0)\big)=0$, 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 $f'(x_0)$

$$\text{EQUAZIONE DELLA RETTA TANGENTE:} \qquad y = f(x_0) + f'(x_0)(x-x_0).$$

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

$$\lim_{x\to x_0} R(x;x_0) = \lim_{x\to x_0}\big(f(x)-f(x_0)-f'(x_0)(x-x_0)\big) = 0. \tag{1.2}$$

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 (1.2) non è granché indicativa! Il fatto *fondamentale* è che per $R(x;x_0)$ vale

$$\lim_{x\to x_0} \frac{R(x;x_0)}{x-x_0} = \lim_{x\to x_0} \frac{f(x)-f(x_0)-f'(x_0)(x-x_0)}{x-x_0} = 0. \tag{1.3}$$

Questa condizione è più restrittiva della precedente (perché?) 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 $\ell$ è pari a $f'(x_0)$. Da questo punto di vista la derivabilità di una funzione descrive la proprietà di una funzione di potere essere "ben approssimata" con una funzione affine, cioé del tipo $y = mx + q$, nell'intorno di un punto $x_0$.

**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

$$\big(x^n\big)' = nx^{n-1} \qquad \forall n \in \mathbb{N}, \tag{1.4}$$

(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:

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

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

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

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

$$\left(x^{-\beta}\right)' \equiv \left(\frac{1}{x^\beta}\right)' = -\frac{\beta}{x^{\beta+1}} \equiv -\beta x^{-\beta-1} \qquad \forall \beta \in \mathbb{N}, \quad \forall\, x \neq 0. \tag{1.5}$$

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

Le formule (1.4) e (1.5) si possono sintetizzare nell'unica formula

$$(x^\alpha)' = \alpha x^{\alpha-1} \qquad \forall\, \alpha \in \mathbb{Z}. \tag{1.6}$$

Dimostriamo che è possibile scegliere $\alpha \in \mathbb{Q}$ ottenendo ancora la formula (1.6). 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 (1.6) per $\alpha$ razionale positivo.

In generale si può dimostrare che (1.6) vale *per ogni* $\alpha \in \mathbb{R}$, cioè

$$(x^\alpha)' = \alpha x^{\alpha-1} \qquad \forall\, \alpha \in \mathbb{R} \quad \forall\, x \neq 0. \tag{1.7}$$

*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 $\lim_{h\to 0} \frac{\cos h - 1}{h} = 0$ e $\lim_{h \to 0} \frac{\sin h}{h} = 1$,

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

*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 $\lim_{h\to 0} \frac{e^h-1}{h} = 1$,

$$\left(e^x\right)' = e^x,$$

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 $\lim_{t\to 0} \frac{\ln(1+t)}{t} = 1$,

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

## 2. Regole fondamentali di derivazione

Come nel caso degli integrali, 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*

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

Per la dimostrazione, osserviamo che il rapporto incrementale di $\phi$ si può riscrivere come

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

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 $(x^k)' = kx^{k-1}$:

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

<u>Derivata di un prodotto.</u> *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 quanto bizzaro possa sembrare, questa è la maniera corretta di derivare un prodotto di funzioni. Per convincersi (cioè 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*

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

Se l'altra formula pareva folle, questa non ne parliamo... Ma è un fatto della vita: per derivare un rapporto, bisogna procedere in questo modo. La dimostrazione discende dalla struttura del rapporto incrementale per la funzione rapporto. Niente di sorprendente. La strategia è di riscrivere questo rapporto incrementale in modo che compaiano quelli delle funzioni $f$ e $g$:

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

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. A questo si ripete il ragionamento fatto qualche riga più su, considerando valori $h$ tali che $x+h\in I$ e si arriva, incolumi, alla stessa conclusione.

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

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

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

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

Abbiamo sviluppato un certo numero di regole per derivare una classe ampia di funzioni: polinomi, funzioni razionali, alcune funzioni trigonometriche, alcune funzioni esponenziali... Vogliamo ora estendere ulteriormente le classi di funzioni di cui sappiamo calcolare *esplicitamente* la derivata. In particolare vogliamo sapere come derivare funzioni composte (ad esempio, $e^{x^2}$, $\sqrt{1+\sin^2 x}$, ...) e funzioni inverse (ad esempio, $\arcsin x$, $\arccos x$, $\arctan x$, ...).

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

$$f(x) = h(g(x)) \quad \Longrightarrow \quad f'(x) = h'(g(x))\, g'(x). \tag{2.1}$$

Quindi per derivare una funzione composta, bisogna derivare ciascuna funzione e fare il prodotto di ciascun termine, calcolato nel valore opportuno. 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 basta fare il prodotto delle derivate, *calcolando la funzione $h'$ in $s = g(x) = x^2$*:

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

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,

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

C'è un unico modo per imparare l'uso concreto di questa formula ed è di riempire un buon numero di fogli di derivate di funzioni composte, possibilmente con l'occhio vigile di qualcuno che controlli l'esattezza del procedimento.

Per dimostrare la formula (2.1), scriviamo il rapporto incrementale

$$\frac{\Delta f}{\Delta x} = \frac{\Delta h}{\Delta x} = \begin{cases} 0 & \text{se} \quad \Delta g = 0, \\ \dfrac{\Delta h}{\Delta g}\,\dfrac{\Delta g}{\Delta x} & \text{se} \quad \Delta g \neq 0, \end{cases} \tag{2.2}$$

dove

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

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 (2.2) tendono a zero per $\Delta x \to 0$.

Applicando la formula (2.1) 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 che $e^{b \ln a} = \left(e^{\ln a}\right)^b = a^b$. Questa maniera di scrivere un esponenziale in base diversa da $e$ tramite un esponenziale con base $e$ è estremamente utile. A eterna memoria, riscriviamo la formula

$$a^b = e^{b \ln a} \qquad \forall\, a > 0,\ b \in \mathbb{R}. \tag{2.3}$$

Usando la formula (2.3),

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

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

<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 $\infty$). 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 $y = f(x)$.*

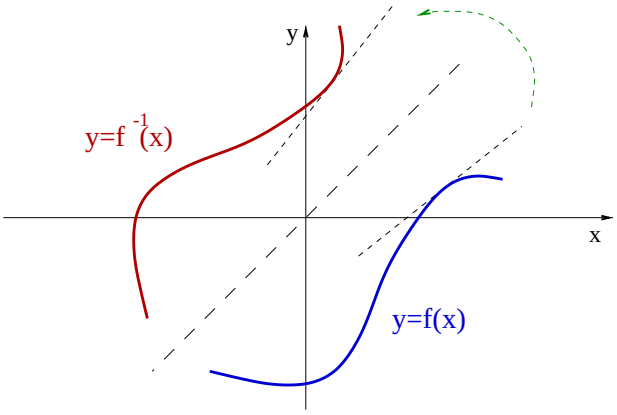



FIGURA 2. Una funzione e la sua inversa, con le relative tangenti.

Ad esempio, funzioni $f$ derivabili con $f'(x) > 0$ dappertutto, hanno sempre inversa derivabile in ogni punto del loro insieme di definizione.

Come si può calcolare la derivata della funzione 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

$$(f^{-1})'(x) = \frac{1}{f'(f^{-1}(x))}. \tag{2.4}$$

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

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

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

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

$$\Rightarrow \quad (\arcsin x)' = \frac{1}{\sqrt{1 - x^2}} \qquad \forall x \in (-1, 1),$$

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

$$\Rightarrow \quad (\arccos x)' = -\frac{1}{\sqrt{1 - x^2}} \qquad \forall x \in (-1, 1).$$

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:

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

Ora siamo soddisfatti: sappiamo derivare un numero di funzioni molto ampio.

## 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*

$$\lim_{h \to 0} \frac{f'(x+h) - f'(x)}{h}, \tag{3.1}$$

*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

$$\begin{aligned} D^{2n-1}(\sin x) &= (-1)^{n+1}\cos x, & D^{2n}(\sin x) &= (-1)^n \sin x, \\ D^{2n-1}(\cos x) &= (-1)^n \sin x, & D^{2n}(\cos x) &= (-1)^n \cos x, \end{aligned} \qquad \forall n \in \mathbb{N}.$$

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 $k \in \mathbb{N}$)

$$C(I) \equiv C^0(I) := \{\text{funzioni continue in } I\}$$
$$C^1(I) := \{\text{funzioni derivabili in } I \text{ e con } f' \in C(I)\}$$
$$C^k(I) := \{\text{funzioni derivabili } k \text{ volte in } I \text{ e con } f^{(k)} \in C(I)\}$$
$$C^\infty(I) := \{\text{funzioni infinitamente derivabili in } I\}.$$

*Ognuno dei simboli precedenti indica un insieme di funzioni. Gli insiemi di funzioni hanno più struttura di quel che si potrebbe immaginare a prima vista: è possibile sommare, moltiplicare funzioni, ottenendo nuove funzioni. A partire da questa osservazione, si sviluppano molti rami estremamente interessanti della matematica, ma, come spesso accade, il tempo è tiranno...*

## 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 al grafico

[1] Nei testi americani, spesso il Teorema di Lagrange è denominato "mean value theorem of differential calculus" o "intermediate value theorem".

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 $\xi$ vale la relazione $f'(\xi) = [f(x_2) - f(x_1)]/[x_2 - x_1]$.

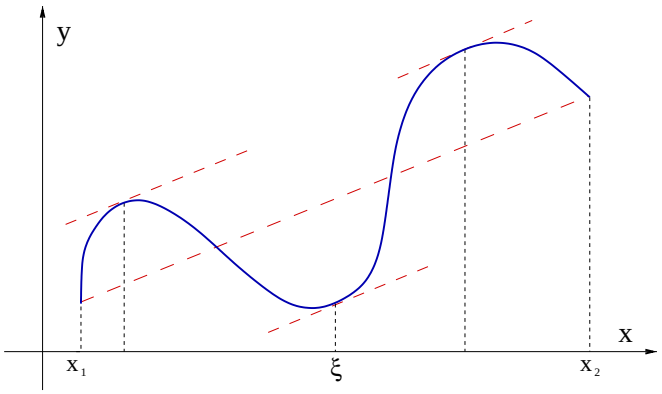



FIGURA 3. Il teorema di Lagrange

TEOREMA 4.1. *Teorema di Lagrange.* *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*

$$f'(\xi) = \frac{f(x_2) - f(x_1)}{x_2 - x_1}.$$

Per enunciare la tesi del Teorema si può anche 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 $\xi$ 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 $\xi$ 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 $\xi$, anziché in $x_0$, l'errore commesso è nullo! Attenzione, però, bisogna sempre ricordare che il punto $\xi$ dipende 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 $\phi$ 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 $\phi$ 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 $\phi$ è continua in $[x_1, x_2]$, per il Teorema di Weierstrass, esistono sia il massimo $M$ che il minimo $m$ di $\phi$ 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 $\ell$. 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 $\phi'(\xi) = 0$. ∎

**Dimostrazione del Teorema 4.1.** Definiamo la funzione (ausiliaria) $\phi$

$$\phi(x) := f(x) - f(x_1) - \frac{f(x_2) - f(x_1)}{x_2 - x_1}(x - x_1),$$

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 $\phi$ 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

$$f'(x) \quad \begin{cases} < 0 & \forall\, x > 0, \\ \\ > 0 & \forall\, x < 0, \end{cases}$$

la funzione è crescente in $(-\infty, 0]$ ed è decrescente $[0, +\infty)$. (Per disegnare un grafico della funzione $f$ si può sfruttare anche $f(0) = 1$, $f \geq 0$ e che $f$ è una funzione pari).

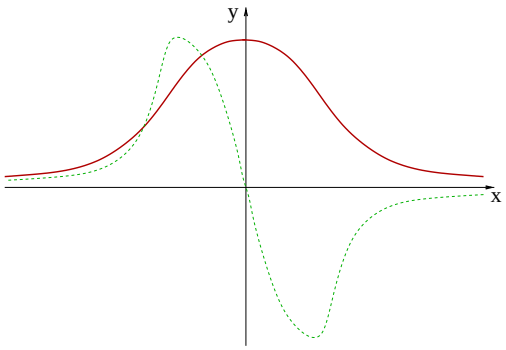



FIGURA 4. Grafico (qualitativo) di $1/(1+x^2)$ e della sua derivata

Vediamo un secondo esempio (potentemente istruttivo!). 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 $(0, +\infty)$.

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 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!)

$$f'(x) = \begin{cases} 1/2 + \cos\left(\frac{1}{x}\right) + 2x\left(2 - \sin\left(\frac{1}{x}\right)\right) & x \neq 0, \\ 1/2 & x = 0, \end{cases}$$

[2] Da cui il noto modo di dire, attribuito a N. Barbecue, "Non tutti i Teoremi riescono col buco"...

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:

$$f'(x) = 0 \quad \forall\, x \in (a,b) \quad \Longrightarrow \quad f \quad \text{è costante in } (a,b).$$

Infatti, per ogni coppia di valori $x_1, x_2 \in (a,b)$, esiste un valore $\xi$, 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.

> *La proprietà di* connessione *di un insieme è importante. Non è questa la sede per approfondire di più la questione, sottolineiamo soltanto che la scelta di lavorare principalmente con funzioni definite in intervalli (o semirette) permette di trarre alcune conclusioni che non sarebbero valide nel caso di funzioni definite in insiemi qualsiasi. E' compito del lettore attento individuare dove sia possibile estendere i risultati a casi più generali, e dove il lavorare su intervalli sia indispensabile.*

*c.* *Lipschitzianità di funzioni a derivata limitata.* La derivabilità di una funzione ne implica la continuità. Vediamo ora un'altra proprietà nella stessa linea di pensiero.

Siano $I$ l'intervallo aperto $(\alpha, \beta)$ 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. Approssimazione lineare.* 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 $\xi$ 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 $\eta$ è compreso tra $\xi$ e $x_0$. Quindi il valore assoluto dell'errore $R(x;x_0)$ è stimato da

$$|R(x;x_0)| = |f''(\eta)||\xi - x_0||x - x_0| \leq M|x - x_0|^2, \tag{4.1}$$

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 4.5. *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 $\lim_{x \to x_0^\pm} f'(x) = \ell^\pm$. Allora $f$ è derivabile in $x_0$ se e solo se $\ell^+ = \ell^-$.*

**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 $\ell^-$,

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

Analogamente per il limite destro. Quindi, nelle ipotesi della Proposizione 4.5, i limiti destro e sinistro del rapporto incrementale esistono e sono uguali, rispettivamente, a $\ell^+$ e $\ell^-$. A questo punto, la conclusione è evidente. ∎

[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.

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

$$D\left(x|x|\right) = \begin{cases} -2x & x < 0, \\ 2x & x > 0. \end{cases}$$

Ne segue che

$$\lim_{x \to 0^-} -2x = \lim_{x \to 0^+} 2x = 0,$$

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

$$D\left(e^{-|x|}\right) = \begin{cases} e^x & x < 0, \\ -e^{-x} & x > 0. \end{cases}$$

In questo caso

$$\lim_{x \to 0^-} e^x = 1 \neq -1 = \lim_{x \to 0^+} -e^{-x},$$

e quindi 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 4.5 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 $f'(0) = 0$.

## 5. Il Teorema fondamentale del calcolo integrale

Torniamo al problema di determinare formule di derivazione e consideriamo il caso delle funzioni integrali. Sia $f \in C([a,b])$ (cioè una funzione continua in $[a,b]$). Fissato $\alpha \in [a,b]$, definiamo la funzione $\phi$ come segue

$$\phi(x) = \int_\alpha^x f(t)\,dt \qquad x \in [a,b].$$

Abbiamo già visto che questa funzione è continua (è anche lipschitziana). Per studiarne la derivabilità, consideriamo il suo rapporto incrementale

$$\frac{\phi(x+h)-\phi(x)}{h} = \frac{1}{h}\left[\int_\alpha^{x+h} f(t)\,dt - \int_\alpha^x f(t)\,dt\right] = \frac{1}{h}\int_x^{x+h} f(t)\,dt$$

(l'ultima uguaglianza è conseguenza delle proprietà di additività dell'integrale). Dato che la funzione $f$ è continua, è possibile applicare il Teorema della Media Integrale per riscrivere il rapporto incrementale come

$$\frac{\phi(x+h)-\phi(x)}{h} = f(\xi) \qquad (\xi \quad \text{compreso tra } x \text{ e } x+h).$$

Dato che $f$ è continua e $\xi \to x$ quando $h \to 0$, si ha $\lim_{h\to 0} f(\xi) = f(x)$. Quindi

$$\phi'(x) = \lim_{h\to 0} \frac{\phi(x+h)-\phi(x)}{h} = \lim_{h\to 0} f(\xi) = f(x).$$

Abbiamo quindi dimostrato il seguente risultato.

TEOREMA 5.1. *Sia $f \in C([a,b])$, $\alpha \in [a,b]$ e $\phi(x) = \int_\alpha^x f(t)\,dt$ per $x \in [a,b]$. Allora $\phi$ è derivabile in $[a,b]$ e $\phi'(x) = f(x)$.*

Una prima conseguenza (pratica) notevole è che, dato che siamo in grado di calcolarne la derivata, possiamo dedurre molte proprietà qualitative importanti anche per una funzione che non sia espressa direttamente tramite funzioni elementari, ma come integrale di una funzione elementare. Ad esempio consideriamo la funzione

$$\text{FUNZIONE DEGLI ERRORI:} \quad \mathrm{Erf}(x) = \frac{1}{\sqrt{2\pi}}\int_0^x e^{-t^2/2}\,dt,$$

(a meno di costanti moltiplicative). Chiaramente questa funzione è ben definita su tutto $\mathbb{R}$ (la funzione $e^{-x^2}$ è continua in tutto $\mathbb{R}$ e quindi integrabile su ogni intervallo). Dal Teorema 5.1 deduciamo che

$$D\left(\mathrm{Erf}(x)\right) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2} > 0.$$

Perciò le rette tangenti al grafico di questa funzione hanno sempre coefficiente angolare positivo. Quindi $\mathrm{Erf}(x)$ è crescente.

Inoltre, il Teorema 5.1 risolve un problema interessante:

*data $f$, trovare una funzione $F$ che risolva l'equazione $F' = f$.*

L'equazione $F' = f$ è un'*equazione differenziale* in cui il dato è la funzione $f$ e l'incognita è la funzione $F$. Una soluzione $F$ di questa equazione si dice PRIMITIVA DI $f$. Il Teorema 5.1 afferma che se $f \in C([a,b])$ il problema $F' = f$ ammette almeno una soluzione (data dalla funzione integrale $\phi$), cioè esiste sempre almeno una primitiva. Da questo punto di vista *si può intepretare l'operazione di integrazione come l'operazione inversa della derivazione.*

TEOREMA 5.2. *Sia $f : [a,b] \to \mathbb{R}$ e siano $F$ e $G$ due sue primitive. Allora esiste $c \in \mathbb{R}$ tale che*

$$F(x) - G(x) = c \qquad \text{per ogni} \quad x \in [a,b].$$

**Dimostrazione.** La derivata della funzione differenza $F - G$ è nulla:

$$(F(x) - G(x))' = F'(x) - G'(x) = f(x) - f(x) = 0.$$

Per quanto già visto (Sezione 4(b)), la differenza $F - G$ deve essere costante. ∎

Quindi, se $f \in C([a,b])$, l'equazione $F' = f$ è completamente risolta: tutte le soluzioni sono della forma

$$\int_\alpha^x f(t)\,dt + c \qquad c \in \mathbb{R}.$$

La classe delle primitive della funzione $f$ si indica con

$$\int f(x)\,dx,$$

e si chiama INTEGRALE INDEFINITO di $f$. Si noti bene che l'integrale indefinito di una funzione indica una classe di funzioni, e non una singola funzione.

Sintetizziamo i due risultati enunciati in un'unico Teorema.

TEOREMA 5.3. *Teorema Fondamentale del Calcolo Integrale. Data $f \in C([a,b])$, le soluzioni dell'equazione differenziale $F' = f$ sono tutte e sole della forma*

$$F(x) = \int_\alpha^x f(t)\,dt + c, \qquad \text{con} \quad \alpha \in [a,b] \quad e \quad c \in \mathbb{R}.$$

Se si cerca una primitiva $F$ di una funzione $f$ con la richiesta aggiuntiva che la funzione $F$ valga in un punto assegnato $x_0$ un valore dato $y_0$, cioè se si vuole risolvere

$$\text{(5.1)} \qquad \text{dati } f \in C([a,b]),\ x_0 \in [a,b],\ y_0 \in \mathbb{R}, \text{ trovare } F \text{ tale che} \quad \begin{cases} F'(x) = f(x), \\ F(x_0) = y_0, \end{cases}$$

la soluzione $F$ esiste, è unica ed è data da

$$\text{(5.2)} \qquad F(x) = y_0 + \int_{x_0}^x f(t)\,dt.$$

Il problema (5.1) rientra nella classe dei *problemi di Cauchy* per equazioni differenziali.

OSSERVAZIONE 5.4. Il problema di Cauchy può essere interpretato in termini di moto di un particella. Se $f(x)$ è la velocità della particella all'istante $x$ e $y_0$ la sua posizione all'istante iniziale $x_0$, la soluzione $F = F(x)$, definita in (5.2), del problema di Cauchy (5.1), rappresenta la posizione della particella all'istante $x$.

**Primitive e calcolo degli integrali definiti.** Il Teorema fondamentale del calcolo ha una conseguenza interessante che riguarda il calcolo esplicito di integrali definiti. Fino ad adesso, data una funzione $f$ continua abbiamo visto che ha senso parlare del suo integrale definito nell'intervallo $[a, b]$

$$\int_a^b f(t)\,dt. \tag{5.3}$$

Abbiamo anche calcolato il valore di questo integrale definito per scelte particolari della funzione $f$. Il calcolo era basato sulla costruzione dell'integrale tramite somme approssimanti e sul fatto che queste somme approssimanti, grazie ad alcune proprietà delle funzioni che abbiamo considerato, possono essere riscritte in una forma che permette di passare al limite ed ottenere il valore esplicito dell'integrale. Concretamente questa strategia di calcolo non è affatto conveniente.

Supponiamo ora di voler calcolare (5.3) e supponiamo di conoscere già (per altre vie) una primitiva della funzione $f$, cioè una funzione $F$ tale che $F' = f$. Sappiamo che anche la funzione integrale definita da

$$\phi(x) = \int_a^x f(t)\,dt.$$

è una primitiva di $f$ e, quindi, per il Teorema Fondamentale del Calcolo (II parte), differisce da $F$ per una costante, cioè $\phi(x) = F(x) + c$ per qualche $c \in \mathbb{R}$. La costante $c$ può essere determinata, calcolando in $x = a$:

$$0 = \phi(a) = F(a) + c \qquad \Longrightarrow \qquad c = -F(a).$$

Si deduce quindi che $\phi(x) = F(x) - F(a)$ e quindi

$$\int_a^b f(t)\,dt = \phi(b) = F(b) - F(a).$$

Quindi, *se si conosce una primitiva $F$ della funzione $f$, l'integrale definito di $f$ in $[a, b]$ è uguale alla differenza dei valori della primitiva in $b$ e in $a$*, cioè

$$F' = f \qquad \Longrightarrow \qquad \int_a^b f(t)\,dt = F(b) - F(a).$$

La differenza $F(b) - F(a)$ si indica anche con $F(x)\Big|_a^b$, o $\Big[F(x)\Big]_a^b$.

Ad esempio, abbiamo visto che

$$\int_a^b x^2\,dx = \frac{b^3}{3} - \frac{a^3}{3}.$$

Per dimostrare questa formula, siamo passati per una rappresentazione (non banale) della somma dei quadrati dei primi $n$ numeri interi. Grazie a quanto abbiamo appena notato, per calcolare lo stesso oggetto, basta determinare una primitiva della funzione $x^2$. Dato che $D(x^3) = 3x^2$,

$$D\left(\frac{1}{3}\,x^3\right) = x^2,$$

quindi una primitiva di $x^2$ è $x^3/3$. Otteniamo perciò

$$\int_a^b x^2\,dx = \frac{1}{3}\,x^3\,\Big|_a^b = \frac{b^3}{3} - \frac{a^3}{3}.$$

Più in generale, dato che

$$D\left(\frac{1}{n+1}\,x^{n+1}\right) = x^n,$$

vale la formula

$$\int_a^b x^n\,dx = \frac{1}{n+1}\,x^{n+1}\,\Big|_a^b = \frac{1}{n+1}(b^{n+1} - a^{n+1}).$$

Nello stesso modo si ottengono formule di integrazione per altre funzioni. Ad esempio,

$$\int_a^b \sin x\,dx = -\cos x\Big|_a^b = \cos a - \cos b.$$

Attenzione! *L'integrazione esplicita tramite una primitiva nota è una scorciatoia notevole rispetto al calcolo del limite delle somme approssimanti, ma è una "fortuna occasionale". Esistono tante funzioni per cui non è possibile determinare una primitiva in forma elementare. Non bisogna perciò mai confondere la definizione di integrale definito (che ha senso per ogni funzione continua o monotona o...) e il calcolo tramite una primitiva nota (che si può realizzare solo in alcuni casi).*

## Appendice A. Derivate parziali

Molto spesso le funzioni con cui si tratta non dipendono da una sola variabile indipendente $x$, ma da più di una. Ad esempio, da quante variabili dipende $x^a$? Da una... e se pensassimo anche $a$ variabile? Allora sono due: $x$ e $a$. In effetti, nella prima risposta abbiamo immaginato $a$ come un parametro fissato, nella seconda abbiamo accettato una sua possibile variazione. In generale, una funzione di due variabili è un oggetto del tipo $f = f(x, y)$. Qui la variabile indipendente è la coppia $(x, y)$ che appartiene ad un appropriato insieme (dominio di definizione). Se si tiene costante la variabile $y$, cioè la si considera un parametro fissato, e si fa variare *solo* la variabile $x$, si ottiene una funzione di una variabile. E' possibile domandarsi se questa funzione sia o meno derivabile.

DEFINIZIONE 5.5. *La funzione $f = f(x, y)$ è* DERIVABILE PARZIALMENTE RISPETTO AD *$x$ nel punto $(x, y)$ se esiste il limite*

$$\lim_{h\to 0} \frac{f(x+h, y) - f(x, y)}{h}.$$

*Il valore del limite si indica con uno dei simboli seguenti*

$$\frac{\partial f}{\partial x}(x, y) = f_x(x, y) = D_x f(x, y).$$

*Analogamente, $f$ è* DERIVABILE PARZIALMENTE RISPETTO AD *$y$ nel punto $(x, y)$ se esiste*

$$\frac{\partial f}{\partial y}(x, y) = f_y(x, y) = D_y f(x, y) = \lim_{k\to 0} \frac{f(x, y+k) - f(x, y)}{h}.$$

Per indicare le derivate parziali si usa il simbolo $\partial$ ("d storto") per ricordare che ci sono molte direzioni di derivabilità e che l'informazione che si ottiene con una singola derivata è una informazione parziale. Questa è coerenza.

Se $f$ è derivabile rispetto ad $x$ e rispetto ad $y$, il vettore $\nabla f(x, y) = (f_x(x, y), f_y(x, y))$ è detto GRADIENTE DI $f$ (nel punto $(x, y)$)[4].

Per il calcolo concreto delle derivate parziali, si possono usare le stesse regole di derivazione di funzioni da $\mathbb{R}$ in $\mathbb{R}$, basta ricordarsi di considerare costanti le variabili rispetto a cui non si deriva. Ad esempio,

$$\begin{aligned} f(x, y) = x^2 + y^2 \quad &\Longrightarrow \quad f_x(x, y) = 2x, \quad f_y(x, y) = 2y, \\ f(x, y) = x^3 y + y^2 \quad &\Longrightarrow \quad f_x(x, y) = 3x^2 y, \ \ f_y(x, y) = x^3 + 2y. \end{aligned}$$

Analogamente al caso di funzioni da $\mathbb{R}$ in $\mathbb{R}$, è possibile definire le *derivate successive* di una funzione $f$. Una funzione di due variabili ammette (al più) due derivate prime (cioè $f_x$ e $f_y$) e (al più) quattro derivate seconde:

$$f_{xx} = \frac{\partial^2 f}{\partial x^2}, \qquad f_{xy} = \frac{\partial^2 f}{\partial x \partial y}, \qquad f_{yx} = \frac{\partial^2 f}{\partial y \partial x}, \qquad f_{yy} = \frac{\partial^2 f}{\partial y^2}.$$

[4]Il simbolo $\nabla$ si legge "nabla".

Ad esempio, consideriamo la funzione $f(x,y) = x^3y + y^2$. Allora

$$f_{xx} = 6x, \qquad f_{xy} = 3x^2, \qquad f_{yx} = 3x^2, \qquad f_{yy} = 2.$$

Chiaramente è possibile definire anche le derivate terze, quarte e compagnia bella.

## Appendice B. Tabella delle derivate

| funzione $f$ | derivata prima $f'$ | funzione $f$ | derivata prima $f'$ |
|---|---|---|---|
| costante | $0$ | $x^\alpha$ | $\alpha x^{\alpha-1}$ |
| $\sin x$ | $\cos x$ | $\cos x$ | $-\sin x$ |
| $e^x$ | $e^x$ | $\ln x$ | $\dfrac{1}{x}$ |
| $\tan x$ | $\dfrac{1}{\cos^2 x} = 1 + \tan^2 x$ | $\cot x$ | $-\dfrac{1}{\sin^2 x} = -1 - \cot^2 x$ |
| $a^x$ | $a^x \ln a$ | $\arctan x$ | $\dfrac{1}{1+x^2}$ |
| $\arcsin x$ | $\dfrac{1}{\sqrt{1-x^2}}$ | $\arccos x$ | $-\dfrac{1}{\sqrt{1-x^2}}$ |
| $\sinh x$ | $\cosh x$ | $\cosh x$ | $\sinh x$ |