Marco Bramanti Carlo D. Pagani Sandro Salsa

Analisi matematica 1

Questo E-Books deve essere cancellato entro le 24 ore successive al download nel caso non si possieda il libro originale. Nessuna garanzia di alcun tipo tutela il materiale fornito e l'utente si impegna ad utilizzarlo a suo rischio e pericolo. Il libro è a scopo informativo e non è assolutamente da considerarsi promozione della pirateria; a tal proposito è reato servirsene a scopo di lucro o di profitto.

ZANICHELLI

nichelli

ematica elementare

ale

ica e Mathematica
lisi matematica
matematica (volume 1º)

tematica
ematica (2ª edizione)

Bramanti, Pagani, Salsa **Analisi matematica 1**
Casolaro **Integrali** (Masson)
Davis **Il mondo dei grandi numeri**
De Marco **Analisi zero** (3ª edizione)
De Marco **Analisi Uno** (2ª edizione - Decibel)
De Marco **Analisi Due** (2ª edizione - Decibel)
De Marco, Mariconda **Esercizi di analisi uno** (2ª edizione - Decibel)
De Marco, Mariconda **Esercizi di analisi due** (Decibel)
De Marco, Mariconda **Esercizi di calcolo in una variabile** (Decibel)
De Marco, Mariconda **Esercizi di calcolo in più variabili** (Decibel)
De Marco **Matematica Uno** (Decibel)
Ghizzetti, Rosati **Analisi matematica** (Masson – 2ª edizione)
Maffei Migliori **Esercizi, appunti e note di istituzioni matematiche**
Minnaja **Matematica Due** (Decibel)
Pagani, Salsa **Analisi matematica** (Masson - 2 volumi)
Pagani, Salsa **Matematica** (Masson)
Pagani, Salsa **Serie di funzioni ed equazioni differenziali**
Ritelli, Bergamini, Trifone **Fondamenti di matematica**
Salsa, Squellati **Esercizi di matematica** (2 volumi)
Salsa, Squellati **Esercizi di analisi matematica II** (Masson - 3 volumi)
Thomas Jr., Finney **Analisi matematica**
Thomas Jr., Finney **Elementi di analisi matematica e geometria**
Torrigiani **Ripensare matematica. In preparazione alle facoltà universitarie scientifiche**

Analisi numerica e Fisica matematica

Bagarello **Fisica matematica**
Barozzi **Matematica per l'ingegneria dell'informazione** (ristampa aggiornata)
Bevilacqua, Bini, Capovani, Menchi **Introduzione alla matematica computazionale**
Bevilacqua, Bini, Capovani, Menchi **Metodi numerici**
Bini, Capovani, Menchi **Metodi numerici per l'algebra lineare**
Bordoni **Lezioni di meccanica razionale** (Masson)
Cercignani **Spazio, tempo, movimento. Introduzione alla meccanica razionale**
Codegone **Metodi matematici per l'ingegneria**
Fabrizio **Elementi di meccanica classica**
Finzi **Meccanica razionale** (2 volumi)
Levi-Civita **Caratteristiche dei sistemi differenziali e propagazione ondosa**
Levi-Civita, Amaldi **Compendio di meccanica razionale** - volume 1
Levi-Civita, Amaldi **Lezioni di meccanica razionale** (volume 2 parte 2)
Vivarelli **Appunti di meccanica razionale** (2ª edizione ampliata)

Probabilità, Statistica e Matematica finanziaria

Campostrini, Parpinel **Introduzione all'inferenza statistica** (Decibel)
Cantucci, Isola **Probabilità elementare**
Dall'Aglio **Calcolo delle probabilità** (3ª edizione)
De Giuli; Giorgi, Maggi, Magnani **Matematica per l'economia e la finanza**
Invernizzi, Rinaldi, Sgarro **Moduli di matematica e statistica**
Letta **Probabilità... elementare**
Monti Pierobon **Teoria della probabilità** (Decibel)
Nash **Giochi non cooperativi e altri scritti**
Ottaviani **Riassunto delle lezioni di matematica attuariale**
Pintacuda **Probabilità** (Decibel)
Scozzafava **Incertezza e probabilità**
Scozzafava **Probabilità e statistica** (Decibel – 2ª edizione)
Toscano **Training autogeno in probabilità** (Decibel/ Zanichelli, rist. aggiornata)

Algebra e Geometria

Abeasis **Complementi di algebra lineare e geometria**
Abeasis **Elementi di algebra lineare e geometria**
Abeasis **Geometria analitica del piano e dello spazio**
Broglia, Fortuna, Luminati **Problemi risolti di algebra lineare** (Decibel)
Carfagna, Piccolella **Complementi ed esercizi di geometria e algebra lineare** (2a edizione)
Cerasoli, Eugeni, Protasi **Elementi di matematica discreta**
Chirita, Ciarletta **Calcolo** (2 volumi)
Dedò **Trasformazioni geometriche**
Dedò **Forme**
Enriques **Lezioni di geometria proiettiva** (Ristampa anastatica)
Questioni riguardanti le matematiche elementari 2 voll. a cura di Enriques
Facchini **Algebra** (Decibel)
Jänich **Topologia**
Kosniowski **Introduzione alla topologia algebrica**
Maroscia **Geometria e algebra lineare**
Maroscia **Introduzione alla geometria e all'algebra lineare**
Maroscia **Problemi di geometria**
Piacentini Cattaneo **Algebra** (Decibel)
Procesi Ciampi, Rota **Esercizi di geometria e algebra**
Ragusa, Sparacino **Esercizi di algebra. Teoria degli insiemi, teoria dei gruppi, teoria degli anelli**
Salce **Lezioni sulle matrici** (Decibel)
Serafini **Ottimizzazione**
Steinhaus **Matematica per istantanee**
Vaccaro, Carfagna, Piccolella **Lezioni di geometria e algebra lineare** (Masson)
Ventre **Introduzione ai grafi planari**

Indice

Prefazione **vii**

1 Numeri **1**

1 Insiemi e logica 1

1.1 Concetti di base sugli insiemi 1

1.2 Un po' di logica elementare 9

2 Sommatorie e coefficienti binomiali 13

2.1 Il simbolo di sommatoria 13

2.2 Fattoriale di n 15

2.3 Coefficienti binomiali e formula di Newton 16

3 Campi ordinati 18

4 Numeri reali. Estremo superiore e assioma di continuità 20

4.1 Inadeguatezza dell'insieme dei razionali per misurare le lunghezze 20

4.2 Estremo superiore e assioma di continuità 21

4.3 Valore assoluto. Disuguaglianza triangolare 23

4.4 Intervalli 24

5 Radicali, potenze, logaritmi 25

5.1 Radici n-esime aritmetiche 25

5.2 Potenze a esponente reale 26

5.3 Logaritmi 27

5.4 Approssimazioni 28

6 Insiemi infiniti 28

7 Il principio di induzione 32

8 Numeri complessi 35

8.1 Definizione di $\mathbb{C}$ e struttura di campo 35

8.2 Coniugato e modulo 37

8.3 Forma trigonometrica 40

8.4 Radici n-esime 43

2 Funzioni di una variabile **49**

1 Il concetto di funzione 49

2 Funzioni reali di variabile reale 52

2.1 Generalità 52

2.2 Funzioni limitate 53

2.3 Funzioni simmetriche 54

2.4 Funzioni monotone 55

2.5 Funzioni periodiche 55
3 Funzioni elementari 56
3.1 Funzioni potenza 56
3.2 Funzioni esponenziali e logaritmiche 61
3.3 Funzioni trigonometriche 63
3.4 Fenomeni vibratori 64
3.5 Funzioni parte intera e mantissa 68
3.6 Funzioni iperboliche 69
3.7 Operazioni sui grafici 70
3.8 Funzioni definite a tratti 74
4 Funzioni composte e inverse 75
4.1 Funzioni composte 75
4.2 Funzioni invertibili; funzioni inverse 77
4.3 Le funzioni trigonometriche inverse 80
4.4 Le funzioni iperboliche inverse 82

3 Limiti e continuità 87
1 Successioni 87
1.1 Definizione di successione. Definizione di limite 87
1.2 Successioni monotone 93
1.3 Il calcolo dei limiti 96
1.4 Il numero e 101
1.5 Confronti e stime asintotiche 103
2 Limiti di funzioni, continuità, asintoti 110
3 Il calcolo dei limiti 121
3.1 Proprietà fondamentali di limiti e continuità 121
3.2 Limiti notevoli 128
3.3 Confronti e stime asintotiche 130
3.4 Stime asintotiche e grafici 132
4 Proprietà globali delle funzioni continue o monotone su un intervallo 136
4.1 Funzioni continue su un intervallo 136
4.2 Funzioni monotone su un intervallo 141
4.3 Continuità e invertibilità 143

4 Calcolo differenziale per funzioni di una variabile 147
1 Introduzione al calcolo differenziale 147
2 Derivata di una funzione 150
2.1 Derivata e retta tangente 150
2.2 Altre interpretazioni della derivata 153
2.3 Derivate di funzioni elementari 154
2.4 Punti angolosi, cuspidi, flessi a tangente verticale 157
3 Regole di calcolo delle derivate 160
3.1 Algebra delle derivate 161
3.2 Derivata di una funzione composta 162
3.3 Derivata di funzione inversa 167
4 Il teorema del valor medio e le sue conseguenze 171
4.1 Punti stazionari. Massimi e minimi locali 171

4.2 Teorema del valor medio. Test di monotonia 174
4.3 Soluzione di alcuni problemi di massimo e minimo 181
4.4 Il teorema di de l'Hospital 187
4.5 Limite della derivata e derivabilità 190
5 Derivata seconda 195
5.1 Significato geometrico della derivata seconda 195
5.2 Derivata seconda, concavità e convessità 196
6 Studio del grafico di una funzione 202
7 Calcolo differenziale e approssimazioni 208
7.1 Differenziale e approssimazione lineare. Il simbolo di "o piccolo" 208
7.2 Limiti notevoli e sviluppi 212
7.3 Formula di Taylor-MacLaurin con resto secondo Peano 213
7.4 Formula di Taylor-MacLaurin con resto secondo Lagrange 218
7.5 Risoluzione approssimata di equazioni: il metodo di Newton 221

5 Serie 229
1 Serie numeriche 229
1.1 Definizione e primi esempi 229
1.2 Serie a termini non negativi 233
1.3 Serie a termini di segno variabile 239
2 Serie di Taylor. Esponenziale complesso 245
2.1 Serie di Taylor delle trascendenti elementari 245
2.2 Serie nel campo complesso. Esponenziale complesso 249

6 Calcolo integrale per funzioni di una variabile 257
1 Introduzione al calcolo integrale 257
2 L'integrale come limite di somme 258
2.1 La definizione di integrale 258
2.2 Classi di funzioni integrabili 262
3 Proprietà dell'integrale 263
4 Il teorema fondamentale del calcolo integrale 266
5 Calcolo di integrali indefiniti e definiti 268
5.1 Integrali immediati, per scomposizione, per sostituzione 268
5.2 Integrazione delle funzioni razionali 273
5.3 Integrazione per parti 277
5.4 Integrazione delle funzioni trigonometriche 281
5.5 Integrazione delle funzioni irrazionali 285
5.6 Integrazione di funzioni discontinue 287
6 Alcune applicazioni fisiche e geometriche 289
7 Calcolo numerico approssimato di un integrale 294
8 Integrali generalizzati 296
8.1 Integrazione di funzioni non limitate 296
8.2 Criteri di integrabilità al finito 297
8.3 Integrazione su intervalli illimitati 300
8.4 Criteri di integrabilità all'infinito 303
9 Funzioni integrali 305
10 Convoluzione e sistemi fisici lineari 310

11 Teorema di Bolzano-Weierstrass, continuità uniforme e integrabilità delle funzioni continue 314
11.1 Alcuni risultati fondamentali per le successioni di numeri reali 314
11.2 Continuità uniforme 316
11.3 Integrabilità delle funzioni continue 318

7 Modelli dinamici discreti 323
1 Introduzione alla modellistica 323
1.1 Modello di Malthus 324
1.2 Modello logistico 326
1.3 Modello dell'acceleratore 327
2 Generalità sulle equazioni alle differenze 328
3 Equazioni lineari del prim'ordine a coefficienti costanti 329
4 Equazioni autonome non lineari 333
4.1 Orbite, diagrammi a gradino e punti fissi o d'equilibrio 333
4.2 Punti fissi e stabilità 335
4.3 Stabilità per linearizzazione 337
4.4 Orbite periodiche e stabilità 340
4.5 Esistenza di orbite periodiche 342
4.6 Comportamento caotico 343
4.7 Equazione logistica discreta 344
5 Equazioni lineari a coefficienti costanti del second'ordine 351
5.1 Equazioni omogenee 351
5.2 I numeri di Fibonacci 354
5.3 Equazioni non omogenee 356
5.4 Stabilità 358

A Formule utili 361
1 Costanti matematiche 361
2 Funzioni trigonometriche 361
3 Funzioni iperboliche 364
4 Derivate elementari 365
5 Regole di derivazione 365
6 Sviluppi di Mac Laurin delle principali funzioni 366
7 Tabella di primitive 367

B Grafici 369

Indice analitico 377

Prefazione

Negli ultimi otto anni i corsi universitari di matematica hanno subìto notevoli cambiamenti: il "nuovo ordinamento" degli studi universitari, basato sul modello di una laurea triennale seguita da un biennio specialistico, ha profondamente mutato le esigenze e le caratteristiche dell'insegnamento della matematica "di base". Ad una prima fase caratterizzata da una drastica riduzione nei contenuti, e contemporaneo mutamento nel taglio di questi insegnamenti, improntati ad uno stile più pragmatico e meno astratto, si sono succeduti via via vari aggiustamenti, che hanno cercato di tener conto sia dell'esperienza didattica che delle esigenze proprie dell'insegnamento di una disciplina con caratteristiche sue proprie, com'è la matematica. Ora siamo ad un nuovo punto di mutamento: l'università torna a sottolineare maggiormente la formazione di base, che era stata in parte sacrificata nella frammentazione di mille corsi brevi o compositi. L'esperienza del nostro testo "Matematica", scritto otto anni fa per venire incontro alle mutate esigenze didattiche, e rivisto dopo quattro anni, ci spinge quindi oggi ad un nuovo impegno, nella speranza di offrire dei testi universitari di matematica che possano essere pienamente utili a studenti e docenti.

La prima scelta, naturale per quanto detto sopra, è stata: a ciascun corso il suo libro di testo. Ecco quindi questa "Analisi 1", che sarà seguita a breve da una "Analisi 2", e affiancata da una "Geometria e Algebra", che sarà scritta da altri ma coordinata con questi nostri testi: questo allo scopo di poter dare ad ogni corso il giusto spazio di approfondimento, senza cadere nelle eccessive sintesi, ma nemmeno nella tentazione del "trattato" che rimane poi chiuso su uno scaffale. Il nostro volume unico "Matematica" continuerà ad esistere, almeno per il momento, e potrà essere utile nelle situazioni in cui rimangano corsi piuttosto compressi.

I contenuti di questo testo di Analisi 1 sono quelli fondamentali di un corso di Analisi Matematica per funzioni di una variabile, ed un'occhiata all'indice sarà sufficiente ad illustrarli. Merita un discorso a parte solo l'ultimo capitolo, "Modelli dinamici discreti", che presenta un argomento meno tradizionale con il quale si intende aprire per lo studente una finestra verso la modellistica, con la speranza di stimolarne la curiosità verso gli sviluppi successivi dell'Analisi e delle sue applicazioni. La controparte continua dei modelli dinamici, ossia le equazioni differenziali, sono state invece collocate nel secondo volume.

Pur nel maggiore approfondimento, i criteri didattici generali che ci ispirano in questo testo sono gli stessi che ci hanno guidato fin qui:

1. Anzitutto, introdurre il minimo di astrazione necessaria per raggiungere l'obiettivo di conoscere, comprendere e saper utilizzare i contenuti di base dell'Analisi Matematica, con particolare riguardo agli aspetti effettivamente utilizzati negli altri corsi della laurea triennale.

2. Mantenere un equilibrio tra sinteticità e chiarezza: "Things should be made as simple as possible, but not any simpler"[1] (Einstein). L'eccessiva brevità oscura le idee. La giustificazione del risultato, la dimostrazione, quando non richieda un apparato formale troppo pesante, e quindi non sia incompatibile con la sinteticità, rende più consapevoli dei nessi e perciò aiuta a comprendere.

3. Motivazione. In uno studio impegnativo come quello della matematica, la motivazione gioca un ruolo fondamentale. D'altro canto, lo studente che affronta un corso di matematica di base, di solito sta iniziando lo studio di una disciplina tecnico-scientifica, che costituisce il suo interesse principale. Perciò si è cercato di presentare ogni nuovo concetto attraverso esempi tratti dalle applicazioni più comuni e di sviluppare la teoria accompagnandola costantemente con riferimenti a problemi tratti dalle varie scienze, evidenziando ove possibile il ruolo dello strumento matematico nella modellizzazione scientifica.

4. Nessuna separazione tra "teoria" e "pratica". Non esiste sapere senza saper fare, e viceversa. Esempi, esercizi e applicazioni sono costantemente alternati alla presentazione teorica.

5. Modularità. I corsi di matematica di base sono variamente organizzati nei vari corsi di studio e nelle varie sedi. Inevitabilmente ogni docente dovrà scegliere quali parti del testo svolgere e quali no, nei propri corsi. Si è cercato di mantenere la massima modularità e indipendenza possibile, compatibilmente con la struttura logica del discorso matematico. In ogni capitolo la materia è stata organizzata raggruppando i concetti irrinunciabili in alcuni paragrafi. Tutto questo dovrebbe rendere agevole per il docente, e quindi per lo studente, un utilizzo parziale del libro.

In particolare, di quasi tutti i teoremi del calcolo differenziale e integrale in una variabile qui presentati è stata fornita una dimostrazione, ottenuta con un percorso logico che renda il più contenuto possibile l'investimento iniziale di tipo fondazionale, *ridotto sostanzialmente al solo teorema di esistenza del limite* per successioni monotone. Alla fine del capitolo sul calcolo integrale è stata inserita una sezione che, avendo come punto d'arrivo la dimostrazione dell'integrabilità delle funzioni continue, presenta sinteticamente i principali temi e risultati fondazionali tradizionali: teorema di Bolzano-Weierstrass, completezza di $\mathbb{R}$, teorema di Heine-Cantor. Svolgendo anche questa sezione si ha un quadro ragionevolmente completo dei fondamenti dell'Analisi Matematica di base, mentre omettendolo non è compromessa in alcun modo la comprensione delle restanti parti del testo.

Desideriamo anche in questa occasione ringraziare tutti i nostri colleghi e studenti che nell'arco di questi anni, con i loro suggerimenti, osservazioni e critiche, ci hanno stimolato a cercare di migliorare continuamente i nostri testi.

Giugno 2008 *Gli Autori*

[1] Le cose andrebbero rese il più semplice possibile, ma non troppo semplici.

1 Numeri

L'obiettivo principale di questo capitolo è introdurre gli *oggetti* più elementari del discorso matematico: i *numeri*. Sul concetto di numero, come vedremo, si basa quello, centrale in tutto questo corso, di *funzione*, che sarà introdotto nel prossimo capitolo e ampiamente ripreso in tutto il seguito.

Tuttavia, è necessario prima introdurre o puntualizzare alcuni concetti che hanno a che fare col *linguaggio matematico*. Come vedremo, non si tratta solo di questioni terminologiche, ma piuttosto di questioni logiche e di metodo: il linguaggio matematico infatti è strettamente legato alla *logica*. A sua volta, da più di cent'anni la logica matematica utilizza anche il linguaggio specifico degli *insiemi*.

Gli insiemi quindi si possono vedere per un verso come altri oggetti elementari del discorso matematico (come lo sono i numeri), per un altro verso come parte integrante del linguaggio logico con cui parliamo di matematica. Al linguaggio logico-insiemistico dedicheremo perciò il primo paragrafo.

1 INSIEMI E LOGICA

1.1 Concetti di base sugli insiemi

La teoria degli insiemi può essere presentata e studiata come teoria assiomatica. Storicamente questo è stato fatto a partire dal 1908, con Zermelo[1]. In modo più informale, la teoria veniva utilizzata già dagli anni 1880[2]; noi la utilizzeremo in questo modo informale, o, come si usa dire, "ingenuo".

Le parole chiave

Il linguaggio degli insiemi si basa su tre parole chiave:

1. Insieme
2. Elemento
3. Appartenenza

[1] Ernst Zermelo, 1871-1953, matematico tedesco.

[2] Ad esempio da Georg Cantor (1845-1918), normalmente considerato il padre del concetto di insieme.

Significato dei termini

Non *definiamo* queste tre parole, ma possiamo spiegarle in modo informale ed implicito. Quando si dice: l'insieme dei punti di un piano, l'insieme degli iscritti a un'università, l'insieme delle stelle di una galassia ... tutti comprendono il significato di queste frasi; la nozione di *insieme* è generalmente assunta come primitiva (cioè non riducibile a concetti più elementari). Useremo come sinonimi le espressioni: collezione, classe, aggregato, famiglia... Un insieme è determinato dai suoi *elementi*, nel senso che un insieme è definito quando abbiamo un criterio con cui stabilire se un dato oggetto è o non è elemento di questo insieme. Invece di dire "il tale oggetto è un elemento del tale insieme" si dice anche che quell'elemento *appartiene* all'insieme.

Simboli

Per indicare gli insiemi si usano solitamente lettere maiuscole, come

$$A, B, X, Y \ldots$$

Per indicare gli elementi di un insieme si usano solitamente lettere minuscole $a, b, \ldots$ $\ldots, x, y \ldots$, e per indicare che un elemento x appartiene all'insieme A scriviamo

$$x \in A$$

o anche $A \ni x$ ("Ad A appartiene x").

Come si specifica un insieme

La scrittura

$$A = \{1, 2, 5\}$$

significa che l'insieme A ha come elementi i numeri $1, 2, 5$ (e nient'altro): l'insieme è quindi ben definito. In questo caso si dice che l'abbiamo definito *per tabulazione*, ossia elencando gli elementi che vi apparetengono. Per questo scopo si usa sempre la scrittura $A = \{\ldots\}$ (cioè con le parentesi graffe che contengono l'elenco degli elementi, separati da virgole). Definire un insieme per tabulazione presuppone che l'insieme abbia un numero *finito* di elementi. Come vedremo, molto spesso in matematica abbiamo a che fare con insiemi infiniti; per definire questi insiemi ci vorrà quindi un altro metodo: vedremo in seguito quale.

Osservazioni sul concetto di insieme

Notiamo che

$$\{1, 2, 5\} = \{1, 5, 2\}$$

ossia: l'*ordine* in cui elenchiamo gli elementi è irrilevante, perché, come detto sopra, ciò che determina un insieme è sapere quali elementi appartengono o meno ad esso, e nient'altro. Il concetto di un *ordine* tra gli elementi è estraneo al concetto di insieme. Così pure gli è estraneo il concetto di *molteplicità* degli elementi: in altre parole, non ha senso dire che un certo elemento di un insieme è "contato due volte". Ad esempio, l'insieme delle soluzioni dell'equazione

$$(x - 1) = 0$$

(che ha come unico elemento il numero 1) è uguale all'insieme delle soluzioni dell'equazione

$$(x-1)^2 = 0.$$

Il fatto che la seconda equazione abbia la soluzione $x = 1$ che ha molteplicità algebrica 2, non cambia l'insieme delle sue soluzioni.

Relazioni tra insiemi: uguaglianza

Abbiamo già detto che due insiemi sono uguali quando possiedono gli stessi elementi. Detto più esplicitamente:

$$A = B \text{ significa:}$$

"Ogni elemento che appartiene ad A appartiene anche a B e ogni elemento che appartiene a B appartiene anche ad A". Si noti che dimostrare l'uguaglianza tra due insiemi comporta quindi dimostrare *due* proposizioni distinte. Formalmente:

$$\forall x\,(x \in A \Longrightarrow x \in B) \text{ e } \forall x\,(x \in B \Longrightarrow x \in A)\,. \tag{1.1}$$

Abbiamo introdotto due *simboli logici*:

il simbolo $\forall$ si chiama *quantificatore!universale* e si legge "per ogni", "per tutti", "per ciascuno";

il simbolo $\Longrightarrow$ si chiama *implicazione logica* e si legge "implica" o "se... allora" (ad esempio: "se $x \in A$ allora $x \in B$").

Riprenderemo entrambi più avanti.

Relazioni tra insiemi: inclusione

Può accadere che valga solo una delle due richieste espresse dalla definizione di uguaglianza (1.1). Ad esempio, se sappiamo solo che ogni elemento di A è anche elemento di B, potremo dire che A è *contenuto* in B. Si scrive:

$$A \subseteq B$$

(e si legge "A è contenuto in B") per indicare che

$$\forall x\,(x \in A \Longrightarrow x \in B)\,.$$

Per quanto detto, quindi, dire che $A = B$ equivale a dire che "$A \subseteq B$ e $B \subseteq A$". Quando $A \subseteq B$, diciamo anche che A è un *sottoinsieme* di B.

Se si afferma che $A \subseteq B$, non si esclude che sia $A = B$. Se invece vogliamo proprio affermare che A è contenuto in B ma non coincide con B, diremo che A è *strettamente contenuto* in B, e scriveremo

$$A \subsetneq B.$$

Si parla in tal caso di *inclusione stretta*. Esplicitamente, affermare che $A \subsetneq B$ significa affermare che

"Ogni elemento che appartiene ad A appartiene anche a B ed esiste un elemento di B che non appartiene ad A".

Formalmente:

$$\forall x\,(x \in A \Longrightarrow x \in B) \text{ e } \exists x \in B : x \notin A.$$

Così facendo abbiamo usato altri tre simboli logici:

- il simbolo $\exists$ si chiama *quantificatore esistenziale* e si legge "esiste", "esistono";
- il simbolo : si legge "tale che" (e tipicamente accompagna il simbolo "esiste" nella costruzione "*esiste* x appartenente ad A *tale che* x ha una certa proprietà");
- il simbolo $\notin$ si legge "non appartiene".

Si faccia attenzione a non confondere tra loro le relazioni "appartiene a" e "è contenuto in" (in simboli, $\in$ e $\subseteq$). In un certo senso, sono due modi per indicare che "qualcosa sta dentro qualcos'altro", ma hanno una differenza logica fondamentale: un *insieme* è *contenuto* in un altro insieme; un *elemento appartiene* a un insieme. L'appartenenza è una relazione tra due oggetti matematici che stanno su due piani gerarchici diversi, mentre l'inclusione è una relazione tra due oggetti logici che stanno sullo stesso livello gerarchico. Ad esempio:

$$\begin{aligned} 3 &\in \{1,3,4\}\,; \\ \{3\} &\subseteq \{1,3,4\}\,; \\ \{1,4\} &\subseteq \{1,3,4\}\,. \end{aligned}$$

In particolare, non si confonda 3 (che è un numero) con $\{3\}$, che è l'insieme che contiene come unico elemento il numero 3.

Talvolta si considerano insiemi che hanno per elementi altri insiemi; questo non contraddice l'osservazione appena fatta: in ogni caso, i simboli $\in$ e $\subseteq$ non sono intercambiabili, ma devono rispettare la distinzione dei livelli gerarchici che abbiamo spiegato. Ad esempio, se definiamo l'insieme

$$A = \{\{1\}, \{2\}, \{1,2\}\},$$

allora è corretto affermare che $\{2\} \in A$, perché $\{2\}$ è un insieme che svolge il ruolo di *elemento* di A, mentre sarebbe scorretto dire che $\{2\} \subseteq A$, perché questo significherebbe che $2 \in A$, mentre A ha come elemento $\{2\}$, ma non 2.

Insieme vuoto

Si definisce *insieme vuoto* l'insieme che non contiene alcun elemento, e si indica tale insieme con il simbolo

$$\emptyset.$$

L'insieme vuoto "ha zero elementi", ma *non è* il numero zero! Non è neanche l'insieme $\{0\}$ (cioè l'insieme che ha zero come unico elemento); difatti, $\{0\}$ ha un elemento, mentre $\emptyset$ non ne ha nessuno. (Il lettore quindi, d'ora in poi *non usi* il simbolo $\emptyset$ per indicare il numero zero).

Per qualsiasi insieme A, possiamo affermare che

$$\emptyset \subseteq A.$$

Per dimostrarlo, dovremmo provare che ogni elemento che appartiene a $\emptyset$, appartiene anche ad A; ma nessun elemento appartiene a $\emptyset$, per cui la dimostrazione è terminata.

Insieme delle parti

L'esempio più naturale di insieme che ha come elementi altri insiemi è il seguente. Fissiamo un insieme X; consideriamo ora l'insieme che ha per elementi tutti i sottoinsiemi di X. Questo insieme si chiama *insieme delle parti di* X, e si indica col simbolo

$$\mathcal{P}(X).$$

Ogni insieme X ha due *sottoinsiemi banali*, che sono X stesso e l'insieme vuoto $\emptyset$ (potrebbero coincidere, se X è vuoto): non bisogna dimenticarseli, quando si elencano tutti i sottoinsiemi di X. Ad esempio, se

$$X = \{1, 2, 3\},$$

allora

$$\mathcal{P}(X) = \{\emptyset, \{1\}, \{2\}, \{3\}, \{1, 2\}, \{2, 3\}, \{1, 3\}, X\}.$$

Per esercizio, il lettore provi a dimostrare che se X ha n elementi, allora $\mathcal{P}(X)$ ha 2^n elementi. Se X è infinito, anche $\mathcal{P}(X)$ è infinito, naturalmente. Si può dimostrare che, in un certo senso, anche in questo caso è vero che $\mathcal{P}(X)$ è "più numeroso" di X. Approfondiremo quest'ultima osservazione nel par. 6 sugli insiemi infiniti.

Insiemi numerici

Prima di continuare il discorso, è bene introdurre i principali *insiemi numerici* che incontreremo nel resto del corso. Per il momento non diremo niente sulle proprietà di questi insiemi: ci basta introdurre il nome e il simbolo con cui si denotano e richiamare allo studente chi sono queste sue vecchie conoscenze.

$\mathbb{N}$ è l'insieme dei *numeri naturali*, ossia i numeri con cui si conta: $0, 1, 2, 3, \ldots$

$\mathbb{Z}$ è l'insieme dei *numeri interi* (detti anche numeri relativi), ossia: $0, 1, -1, 2, -2, \ldots$;

$\mathbb{Q}$ è l'insieme dei *numeri razionali*,numeri!razionali ossia delle frazioni

$$\frac{n}{m}, \text{ dove } n, m \in \mathbb{Z}, \text{ e } m \neq 0.$$

I numeri razionali possono essere scritti anche in forma decimale. Ad esempio:

$$\frac{3}{4} = 0{,}75, \text{ mentre } \frac{1}{3} = 0{,}33333\cdots = 0{,}\overline{3}.$$

Gli esempi vogliono ricordare il seguente fatto: un numero razionale, scritto in forma decimale, dopo la virgola può presentare un numero finito di cifre (diverse da zero), oppure un numero infinito di cifre diverse da zero, che però si ripetono *periodicamente*. Si ricordi che $0,\overline{9} = 1$. (Per convincersene, il lettore si chieda quanto vale $1 - 0,\overline{9}$, e provi a rispondere ragionando).

$\mathbb{R}$ è l'insieme dei *numeri reali*, ossia quelli che, scritti in forma decimale, presentano dopo la virgola una successione qualsiasi di cifre diverse da zero, eventualmente anche infinita e non periodica. Che esistano numeri di quest'ultimo tipo (ossia reali ma non razionali), si capisce riflettendo su esempi come:

$$0{,}10110111011110\ldots$$

Il numero precedente dopo la virgola ha: una cifra uguale a 1, poi 0, poi due cifre uguali a 1, poi 0, poi tre cifre uguali a 1... e così via. È chiaro che questo criterio

 © 978-88-08-06485-1

definisce con precisione un numero decimale. D'altro canto, l'allineamento di cifre diverse da zero dopo la virgola non è né finito né periodico: questo numero perciò è *irrazionale*.

Nel corso incontreremo anche: $\mathbb{C}$, l'insieme dei numeri complessi, ossia del tipo $a + ib$, dove a, b sono numeri reali, e i è l'unità immaginaria, ossia un numero il cui quadrato è -1. Tra gli insiemi numerici ora ricordati, questo è l'unico che probabilmente non è una vecchia conoscenza per lo studente: dedicheremo a questo insieme il par. 8, definendolo rigorosamente. Per ora ci accontentiamo di averne introdotto il nome e il simbolo.

Tra gli insiemi numerici introdotti, valgono le inclusioni:

$$\mathbb{N} \subsetneqq \mathbb{Z} \subsetneqq \mathbb{Q} \subsetneqq \mathbb{R} \subsetneqq \mathbb{C}.$$

Come indicato dai simboli, tutte le inclusioni sono strette: esistono numeri interi non naturali (i numeri negativi), esistono numeri razionali non interi (le frazioni proprie), esistono numeri reali non razionali (i numeri irrazionali), esistono numeri complessi non reali (i numeri immaginari).

Operazioni su insiemi

Molto spesso, nel contesto di un certo discorso matematico, esiste un insieme X che svolge il ruolo di *universo*, nel senso che tutti gli insiemi di cui si parla in quel contesto sono sottoinsiemi di X. Ad esempio, in questioni di aritmetica potrebbe essere $X = \mathbb{N}$, mentre in questioni di analisi potrebbe essere $X = \mathbb{R}$. Perciò succede spesso di trovarsi a operare con insiemi che sono tutti sottoinsiemi di un certo universo comune, X. In queste condizioni, si possono definire certe *operazioni sugli insiemi*.

- *Intersezione*. L'intersezione di due insiemi A, B è definita da:

  $$A \cap B = \{x : x \in A \text{ e } x \in B\}.$$

  In altre parole, è l'insieme degli elementi che appartengono sia al primo sia al secondo insieme.

- *Unione*. L'unione di due insiemi A, B è definita da:

  $$A \cup B = \{x : x \in A \text{ o } x \in B\}.$$

  In altre parole, è l'insieme degli elementi che appartengono al primo o al secondo insieme, intendendo la "o" in modo non esclusivo (esplicitamente: $A \cup B$ è l'insieme degli elementi che appartengono ad A o a B *o a entrambi*).

- *Differenza*. La differenza tra due insiemi A, B è definita da:

  $$A \setminus B = \{x : x \in A \text{ e } x \notin B\}.$$

  In altre parole, è l'insieme degli elementi che appartengono al primo ma non al secondo insieme.

- *Complementare.* È un tipo particolare di differenza. Se $A \subseteq X$, il complementare di A rispetto a X è per definizione:

$$\complement_X A = X \setminus A.$$

Spesso X è sottointeso: si usa allora il simbolo A^c.

Le definizioni appena date delle operazioni tra insiemi sono visualizzate nei diagrammi in figura 1.1.

Figura 1.1. Diagrammi di Venn.

C'è poi un'altra operazione sugli insiemi, che può essere eseguita su due insiemi qualsiasi (cioè due insiemi non necessariamente contenuti nel medesimo universo):

- *Prodotto cartesiano.* Dati due insiemi non necessariamente distinti A e B, possiamo considerare un nuovo insieme costituito da tutte le *coppie ordinate* (a, b) con $a \in A$ e $b \in B$. Esso prende il nome di *prodotto cartesiano* di A per B e si indica col simbolo $A \times B$.

Una illustrazione è data nella figura 1.2, dove si può osservare anche che $A \times B$ in generale è diverso da $B \times A$.

Un tipico uso di prodotto cartesiano si ha con $\mathbb{R} \times \mathbb{R}$, che si abbrevia col simbolo $\mathbb{R}^2$, e denota l'insieme delle *coppie ordinate di numeri reali.* Analogamente, $\mathbb{R}^n$ (abbreviazione del prodotto cartesiano di n insiemi uguali ad $\mathbb{R}$) è l'insieme delle *n-uple ordinate di numeri reali*:

$$\mathbb{R}^n = \{(x_1, x_2, ..., x_n) : x_i \in \mathbb{R}\}.$$

Figura 1.2.

Operazioni tra insiemi e operazioni logiche

Nelle definizioni appena date, non sarà sfuggita al lettore la relazione tra *operazioni sugli insiemi* e *operazioni logiche*. Precisamente:

- L'intersezione insiemistica è definita mediante la "*e*" (congiunzione logica).
- L'unione insiemistica è definita mediante la "*o*" (disgiunzione logica).
- La differenza insiemistica e l'operazione di complementazione sono definite mediante il "*non*" (negazione logica).

Come si vede, linguaggio logico e linguaggio insiemistico sono due facce della stessa medaglia. Vedremo in seguito altre corrispondenze di questo tipo.

Proprietà delle operazioni su insiemi

Le operazioni sugli insiemi godono di una serie di proprietà, che discendono in sostanza da proprietà analoghe delle operazioni logiche (congiunzione, disgiunzione, negazione)[3]. Queste proprietà sono utili quando si vuole trasformare un'espressione astratta che coinvolge operazioni su insiemi in un'altra equivalente, analogamente a quanto facciamo comunemente nel calcolo letterale, quando trasformiamo un'espressione in un'altra. Si osservi che alcune proprietà delle operazioni insiemistiche non assomigliano a quelle delle operazioni sui numeri:

- *Proprietà dell'intersezione*:
  commutativa: $A \cap B = B \cap A$;
  associativa: $A \cap (B \cap C) = (A \cap B) \cap C$
  idempotenza: $A \cap A = A$
- *Proprietà dell'unione*:
  commutativa: $A \cup B = B \cup A$;
  associativa: $A \cup (B \cup C) = (A \cup B) \cup C$
  idempotenza: $A \cap A = A$
- *Proprietà distributive* (legano unione e intersezione):
  $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
  $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$

[3]Si tratta delle regole logiche con cui utilizziamo "e", "o", "non" in qualsiasi ragionamento deduttivo. A questo livello informale del discorso non è il caso nemmeno di esplicitarle.

- *Leggi di De Morgan* (legano il complementare con le altre operazioni):
  $(A \cup B)^c = A^c \cap B^c$
  $(A \cap B)^c = A^c \cup B^c$
  $(A^c)^c = A$

  Si noti come l'operazione di complementare scambi tra loro unione e intersezione.

- Operazioni con l'insieme vuoto e con l'insieme universo X:
  $A \cap \emptyset = \emptyset$
  $A \cap X = A$
  $A \cup \emptyset = A$
  $A \cup X = X$
  $X^c = \emptyset$
  $\emptyset^c = X$

- Relazioni tra l'inclusione e le operazioni di unione e intersezione:
  $A \subseteq B$ se e solo se $A \cap B = A$ se e solo se $A \cup B = B$.

1.2 Un po' di logica elementare

Predicati (o proprietà) e proposizioni (o enunciati)
Consideriamo la seguente affermazione:

(1.2) "il numero naturale n è dispari"

e chiediamoci se è vera. Ovviamente l'unica risposta sensata è: " dipende da n". Infatti nella (1.2) il simbolo n rappresenta una *variabile* che può assumere valori diversi e che possono rendere l'affermazione vera o falsa.

Una frase di questo tipo si chiama *predicato* o *proprietà*: la sua verità o falsità dipende dai valori della o delle variabili che in essa compaiono.

Esaminiamo ora l'affermazione

" Per ogni numero naturale n, se n è dispari allora n^2 è dispari"

che possiamo scrivere più formalmente nel modo seguente:

(1.3) "$\forall n \in \mathbb{N}\,(n \text{ dispari} \implies n^2 \text{ dispari})$" .

In questo caso non ha senso assegnare liberamente valori particolari ad n. Per esempio, per $n = 3$ si otterrebbe "Per ogni numero naturale 3,..." che è priva di significato. Si dice che la variabile n non è libera, ma vincolata dal *quantificatore* $\forall$.

Come conseguenza si ha che la (1.3) è *vera* o *falsa* " una volta per tutte" e prende il nome di *enunciato* o *proposizione*. In particolare, i componenti della (1.3) sono l'insieme $\mathbb{N}$, due predicati definiti su $\mathbb{N}$ dati da

$$p(n) : \text{"}n \text{ dispari"} \quad \text{e} \quad q(n) : \text{"}n^2 \text{dispari"}$$

e l'implicazione $p(n) \implies q(n)$. In generale un enunciato che presenti un insieme A, due predicati $p(x)$ e $q(x)$ il cui argomento x varia in A e la struttura logica

(1.4) "$\forall x \in A\,(p(x) \implies q(x))$"

prende il nome di *implicazione universale*.

Dimostrazioni e controesempi

La maggior parte dei teoremi è costituita da implicazioni universali, nelle quali il predicato p fa la parte dell'*ipotesi* e q quella della *tesi*. In particolare la (1.3) è un teorema. In questo caso ci si convince facilmente che l'asserto è vero, ma come si fa a dimostrarlo rigorosamente? Per esempio, è sufficiente osservare che

$$3 \text{ è dispari e } 3^2 \text{ è dispari}$$

per affermare che il teorema è vero?

Certamente no, poiché il teorema pretende che l'implicazione valga per *ogni* numero naturale. Tuttavia, i numeri dispari sono infiniti: come facciamo a provare un'implicazione per infiniti numeri? Il procedimento chiave è questo: si considera il generico n che soddisfa l'*ipotesi* (essere dispari) e si dimostra che n soddisfa la *tesi* (il suo quadrato è dispari). Vediamo come si opera.

DIMOSTRAZIONE. Sia n dispari. Osserviamo che qualunque numero dispari si può scrivere nella forma $2k+1$, con k opportuno. Sia dunque $n = 2k+1$ il generico numero dispari. Per dimostrare il teorema occorre poter scrivere n^2 come (intero pari $+1$). Si ha:

$$n^2 = (2k+1)^2 = 4k^2 + 4k + 1 = 2\left(2k^2 + 2k\right) + 1.$$

Poiché $2\left(2k^2+2k\right)$ è un intero pari n^2 risulta dispari. ◆

Evidenziamo il procedimento logico usato nella dimostrazione: per dimostrare un'implicazione universale come (1.4), si considera il *generico* x che soddisfa l'ipotesi $p(x)$ e si cerca di mostrare che $q(x)$ è vera.

Veniamo ora ai *controesempi*, che costituiscono un'importante risorsa per dimostrare la... falsità di un'implicazione universale. Chiediamoci, per esempio, se il seguente teorema è vero o falso:

TEOREMA 1.1 *Per ogni numero naturale n, se n è primo allora n è dispari.*

Un attimo di riflessione mostra che questo teorema è falso. Infatti, il numero 2 è primo ma è pari.

Si osservi che anche questo teorema esprime un'implicazione universale, ossia ha la stessa struttura logica di quello dell'esempio precedente. Riflettiamo: come mai per poter affermare che un'implicazione universale è vera è necessaria una dimostrazione (mentre un esempio non è sufficiente), mentre per mostrare che un'implicazione universale è falsa basta un esempio contrario?

Il punto è che l'implicazione universale pretende proprio che *ogni* x che soddisfa l'ipotesi soddisfi anche la tesi: perciò, se troviamo anche un solo esempio di x che soddisfa l'ipotesi ma non la tesi, questo significa che l'implicazione universale è falsa. Si badi: non "falsa in un caso", ma semplicemente "falsa", perché, come abbiamo già osservato, l'implicazione universale è vera o falsa una volta per tutte.

In generale, un esempio di oggetto x che soddisfa l'ipotesi ma non la tesi di una implicazione universale, e quindi ne mostra la falsità, si chiama *controesempio*. Il numero $n = 2$ è un controesempio al fatto che "per ogni naturale n, se n è primo allora n è dispari". Detto in un altro modo: la *negazione* della proposizione

"Per ogni $x \in A$, se vale $p(x)$ allora vale $q(x)$" è

"Esiste un $x \in A$ per cui vale $p(x)$ ma non vale $q(x)$". (Questo x particolare costituisce un controesempio).

Implicazione logica e inclusione insiemistica. Insiemi definiti da proprietà

Abbiamo visto in precedenza che esiste un parallelismo tra le operazioni di intersezione, unione, complementare e le operazioni logiche di congiunzione ("e"), disgiunzione ("o"), e negazione ("non"). Possiamo ora completare questo quadro osservando il parallelismo che esiste tra la relazione di inclusione insiemistica e l'implicazione logica. Per spiegarlo, consideriamo l'implicazione universale:

"Per ogni numero naturale n, se n è divisibile per 4 allora n è divisibile per 2".

Se indichiamo con

$D_4 = \{n \in \mathbb{N} : n \text{ è divisibile per } 4\}$ e $D_2 = \{n \in \mathbb{N} : n \text{ è divisibile per } 2\}$

possiamo osservare che l'implicazione universale scritta sopra è equivalente all'affermazione:

$$\text{“}D_4 \subseteq D_2\text{”}.$$

Infatti, questa inclusione significa che ogni elemento appartenente a D_4 appartiene anche a D_2, cioè che ogni numero naturale divisibile per 4 è anche divisibile per 2.

Astraendo dall'esempio, possiamo affermare che l'implicazione universale:

"Per ogni $x \in A$, se vale $p(x)$ allora vale $q(x)$" è equivalente all'inclusione insiemistica

$$\{x \in A : p(x) \text{ è vera}\} \subseteq \{x \in A : q(x) \text{ è vera}\}.$$

Questo ci porta anche a osservare esplicitamente il ruolo importante giocato dagli *insiemi definiti da proprietà*. Introducendo il concetto di insieme, abbiamo spiegato cosa significa definire un insieme per tabulazione dei suoi elementi, come quando scriviamo

$$A = \{1, 2, 5\}.$$

Abbiamo anche osservato che questo metodo ha senso solo quando l'insieme che si vuole definire possiede un numero finito di elementi (altrimenti, non possiamo *elencarli* tutti!). Ora possiamo esplicitare qual è l'altro modo di definire un insieme, modo che abbiamo già usato in vari esempi e che è tipico di quando si vuole definire un insieme infinito: definire un insieme mediante una *proprietà caratteristica*, come nell'esempio:

$$D_4 = \{n \in \mathbb{N} : n \text{ è divisibile per } 4\}$$

o, volendo enunciare in astratto questo schema logico:

$$E = \{x \in A : p(x) \text{ è vera}\}.$$

In queste definizioni, l'importante è che la proprietà $p(x)$ che si utilizza abbia senso per ogni x dell'insieme A, e quindi risulti vera o falsa (senza ambiguità di significato) per ogni particolare $x \in A$; l'insieme E consisterà allora di *tutti e soli* quegli x appartenenti ad A per cui la proprietà $p(x)$ è vera. A questo proposito, è fondamentale che sia specificato l'insieme universo A in cui si ragiona: ad esempio, "L'insieme dei *numeri naturali* divisibili per 4" e non invece "L'insieme dei *numeri* divisibili per 4", che avrebbe un significato ambiguo (il numero $\sqrt{2}$ è divisibile per 4?).

Negazioni e dimostrazioni indirette

Terminiamo questa rapida panoramica sui concetti logici e insiemistici puntualizzando alcuni fatti relativi alla negazione di una proposizione, e il suo utilizzo nelle dimostrazioni indirette.

Cominciamo ad osservare che l'implicazione universale

$$\text{“}\forall x \in A\,(p\,(x) \Longrightarrow q\,(x))\text{”} \tag{1.5}$$

è logicamente equivalente a

$$\text{“}\forall x \in A\,(\text{non } q\,(x) \Longrightarrow \text{non } p\,(x))\text{”}\ . \tag{1.6}$$

La seconda implicazione si dice la *controinversa* della prima. Ad esempio: poiché sappiamo che vale il teorema

$$\text{“}\forall n \in \mathbb{N}\,\big(n \text{ dispari} \Longrightarrow n^2 \text{ dispari}\big)\text{”},$$

allora possiamo anche affermare che

$$\text{“}\forall n \in \mathbb{N}\,\big(n^2 \text{ pari} \Longrightarrow n \text{ pari}\big)\text{”}. \tag{1.7}$$

Infatti: sia n^2 pari, e proviamo che allora n è pari; se n non è pari, allora è dispari, e quindi per la prima implicazione anche n^2 è dispari, contro l'ipotesi che n^2 sia pari. Di conseguenza n è pari e la seconda implicazione è dimostrata.

Un attimo di riflessione mostra che il ragionamento appena fatto ha una validità generale, e mostra appunto che se è vera la (1.5) allora è vera la (1.6); di più, se è vera la seconda allora è vera la prima (perché "non non p" è logicamente equivalente a p), per cui le due sono logicamente equivalenti.

L'equivalenza tra (1.5) e (1.6) è detta *legge delle controinverse* e un tipico metodo di dimostrazione indiretta consiste appunto nel provare la (1.6) per mostrare che è vera la (1.5).

L'altra tipica dimostrazione indiretta è la *dimostrazione per assurdo*, che esemplifichiamo col seguente teorema, che ci sarà utile in seguito.

Teorema 1.2 *Non esiste un numero razionale il cui quadrato è* 2.

La seguente dimostrazione si trova negli Elementi di Euclide (circa 300 a.C.).

Dimostrazione. Supponiamo per assurdo che esista un numero $r \in \mathbb{Q}$ tale che $r^2 = 2$. Possiamo scrivere $r = \frac{n}{m}$ con $n, m \in \mathbb{Z}, m \neq 0$. Inoltre, possiamo supporre che la frazione n/m sia già ridotta ai minimi termini, ossia "semplificata" (in altre parole: n, m non contengono fattori comuni). Abbiamo dunque la catena di implicazioni:

$$\begin{aligned}\left(\frac{n}{m}\right)^2 &= 2\\ n^2 &= 2m^2\end{aligned}$$

per cui n^2 è pari; ma allora per la (1.7) anche n è pari e possiamo scrivere $n = 2k$ per qualche $k \in \mathbb{Z}$. Allora la relazione $n^2 = 2m^2$ si riscrive

$$\begin{aligned}(2k)^2 &= 2m^2\\ 4k^2 &= 2m^2\\ m^2 &= 2k^2\end{aligned}$$

per cui m^2 è pari; ma allora per la (1.7) anche m è pari; dunque sia n che m sono pari, e questo è assurdo, perché avevamo supposto che la frazione n/m fosse già stata semplificata. ◆

In generale, la dimostrazione per assurdo consiste nel supporre vera l'ipotesi del teorema e la negazione della tesi, e dedurre da questi fatti una contraddizione di qualsiasi tipo.

Sia nelle dimostrazioni per assurdo sia nell'usare la legge delle controinverse, occorre naturalmente saper costruire la corretta *negazione* di una proposizione o proprietà data. Si osservi che questa è un'operazione puramente formale, sintattica, che non ha a che fare col contenuto della proposizione, né col fatto che essa sia vera o falsa.

Riportiamo schematicamente alcune regole tipiche con cui si costruisce la negazione di una proposizione o proprietà. Il lettore è invitato a rendersi conto della validità di queste regole ragionando su esempi concreti.

Se $p(x), q(x)$ sono due proprietà qualsiasi, allora:

- la negazione di "$p(x)$ e $q(x)$" è "non $p(x)$ o non $q(x)$";
- la negazione di "$p(x)$ o $q(x)$" è "non $p(x)$ e non $q(x)$";
- la negazione di "non $p(x)$" è $p(x)$;
- la negazione di "$\forall x$ vale $p(x)$" è "$\exists x$: non $p(x)$"
- la negazione di "$\exists x$: vale $p(x)$" è "$\forall x$ non $p(x)$"
- la negazione di "$\forall x\,(p(x) \Longrightarrow q(x))$" è "$\exists x : (p(x)$ e non $q(x))$".

2 SOMMATORIE E COEFFICIENTI BINOMIALI

2.1 Il simbolo di sommatoria

Definizione 1.1 Siano $a_1, a_2, \ldots, a_n$, n numeri reali. La loro somma

$$a_1 + a_2 + \ldots + a_n$$

si può indicare in forma compatta col simbolo di *sommatoria*:

$$\sum_{i=1}^{n} a_i$$

che si legge: "sommatoria per i da 1 a n di a_i". Il simbolo i si dice *indice di sommatoria*. ■

Il simbolo di sommatoria è dunque una pura e semplice stenografia, che tuttavia risulta molto utile quando i termini a_i sono definiti esplicitamente in funzione dell'indice i, ad esempio:

$$\sum_{i=1}^{10} \frac{1}{i} = 1 + \frac{1}{2} + \frac{1}{3} + \ldots + \frac{1}{10}$$

$$\sum_{i=3}^{n} i^2 = 3^2 + 4^2 + 5^2 + \ldots + n^2$$

Per usare agilmente il simbolo di sommatoria occorre capire bene l'utilizzo dell'indice di sommatoria. Anzitutto, l'indice di sommatoria è un *indice muto*. Questo vuol dire che se si sostituisce i con j, k o qualunque altro indice (in tutte le sue occorrenze) il senso dell'espressione non cambia:

$$\sum_{i=3}^{n} i^2 = \sum_{j=3}^{n} j^2$$

Invece,

$$\sum_{i=1}^{n} i^2 \neq \sum_{i=1}^{m} i^2$$

in quanto i due simboli indicano la somma, rispettivamente, dei primi n oppure dei primi m quadrati: se $n \neq m$ il risultato sarà diverso.

Le seguenti proprietà formali delle sommatorie sono facilmente comprensibili se si pensa a ciò che esse affermano in termini di somme scritte per esteso.

Proposizione 1.1 (proprietà formali delle sommatorie)

1. *Prodotto per una costante (proprietà distributiva):*

$$\sum_{k=1}^{n} (c \cdot a_k) = c \cdot \sum_{k=1}^{n} a_k$$

2. *Sommatoria con termine costante:*

$$\sum_{k=1}^{n} c = c \cdot n = c \cdot (\textit{numero di addendi della sommatoria})$$

3. *Somma di sommatorie:*

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

4. *Scomposizione:*

$$\sum_{k=1}^{n+m} a_k = \sum_{k=1}^{n} a_k + \sum_{k=n+1}^{n+m} a_k$$

5. *Traslazione di indici:*

$$\sum_{k=1}^{n} a_k = \sum_{k=1+m}^{n+m} a_{k-m}$$

6. *Riflessione di indici:*

$$\sum_{k=1}^{n} a_k = \sum_{k=1}^{n} a_{n-k+1} = \sum_{k=0}^{n-1} a_{n-k}$$

La dimostrazione di queste proprietà è un esercizio di trascrizione: basta, cioè, scrivere per esteso che cosa indica ciascuna sommatoria ed eseguire qualche passaggio elementare.

Somma di una progressione geometrica

Si dice che n termini sono in *progressione geometrica* se il rapporto tra ogni termine (a partire dal secondo) e il precedente è costante. Tale costante si dice *ragione* della progressione. Se il primo termine è a e la ragione è q, i termini successivi saranno aq, aq^2, aq^3, e così via.

Proposizione 1.2 *Per la somma dei primi termini della progressione geometrica di ragione $q \neq 1$ (e $a = 1$), vale la formula:*

$$\sum_{k=0}^{n} q^k = \frac{1 - q^{n+1}}{1 - q} \tag{2.1}$$

Se $q = 1$ la sommatoria scritta vale $n + 1$, ovviamente.

Dimostrazione. Proviamo l'identità nella forma equivalente:

$$(1 - q) \sum_{k=0}^{n} q^k = 1 - q^{n+1}.$$

Infatti, applicando le proprietà delle sommatorie 1, 5, 4 (v. proposizione 1.1) si ha:

$$\begin{aligned}(1 - q) \sum_{k=0}^{n} q^k &= \sum_{k=0}^{n} q^k - q \sum_{k=0}^{n} q^k = \\ &= \sum_{k=0}^{n} q^k - \sum_{k=0}^{n} q^{k+1} = \sum_{k=0}^{n} q^k - \sum_{k=1}^{n+1} q^k = \\ &= 1 + \sum_{k=1}^{n} q^k - \left(\sum_{k=1}^{n} q^k + q^{n+1} \right) = 1 - q^{n+1}\end{aligned}$$

2.2 Fattoriale di n

Un'espressione di uso frequente in Analisi è il *fattoriale* di n. Con ciò si intende il prodotto dei primi n interi: si indica con $n!$ (e si legge "n fattoriale")

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

Si pone, per definizione, $0! = 1$. Il numero $n!$ cresce molto rapidamente al crescere di n. Ecco i primi valori

n	0	1	2	3	4	5	6	7	8	9	10
$n!$	1	1	2	6	24	120	720	5 040	40 320	362 880	3 628 800

Alcune proprietà del fattoriale, di verifica immediata, sono:

$$n! = n\,(n - 1)!$$

$$\text{se } 0 < k < n \qquad \frac{n!}{(n - k)!} = n\,(n - 1)\,(n - 2) \ldots (n - k + 1)$$

L'ultimo prodotto scritto è il *prodotto di k fattori, partendo da n e decrescendo.* Per esempio,

$$\frac{100!}{95!} = 100 \cdot 99 \cdot 98 \cdot 97 \cdot 96$$

Si noti che una calcolatrice tascabile riesce a calcolare $100 \cdot 99 \cdot 98 \cdot 97 \cdot 96$ senza problemi, ma segnala "errore" non appena si digita 100! Conviene sempre semplificare il più possibile le espressioni che contengono il fattoriale, prima di calcolarle.

Una prima applicazione del fattoriale si ha nel problema seguente: in quanti modi è possibile ordinare n oggetti distinti?

Per esempio, se abbiamo tre oggetti: a, b, c, questi possono essere ordinati in $3! = 6$ modi:

$$abc, \quad acb, \quad bac, \quad bca, \quad cab, \quad cba$$

Ognuno di questi ordinamenti si chiama *permutazione* (o *sostituzione*) degli n oggetti. Si ottiene facilmente che il numero delle permutazioni di n oggetti distinti è $n!$

Basta infatti pensare di realizzare una permutazione collocando gli n oggetti in n scatole numerate da 1 a n. La prima scatola si può riempire in n modi. Riempita la prima scatola, si può riempire la seconda in $(n-1)$ modi con i restanti oggetti. Così il complesso delle prime due scatole si può riempire in $n(n-1)$ modi. Il complesso delle prime tre scatole si può riempire in $n(n-1)(n-2)$ modi e così via fino a riempire tutte le scatole.

2.3 Coefficienti binomiali e formula di Newton

Sviluppando la potenza n-esima di un binomio $(a+b)$ si trova

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

I numeri $c_{n,k}$ sono detti *coefficienti binomiali* e intervengono in molte questioni non solo di Analisi, ma anche di Probabilità, Statistica, etc. Essi si indicano col simbolo $\binom{n}{k}$ (si legga: n su k); si può dimostrare che hanno l'espressione seguente:

$$c_{n,k} = \binom{n}{k} = \frac{n!}{k!(n-k)!} \qquad 0 \le k \le n$$

Per quanto osservato poco sopra, si può scrivere anche:

$$\binom{n}{k} = \frac{n\,(n-1)\,(n-2)\dots(n-k+1)}{k!}$$

espressione che è più maneggevole per il calcolo effettivo.

La formula di Newton si scrive dunque:

$$\boxed{(a+b)^n = \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k}} \tag{2.2}$$

Valgono anche le seguenti proprietà, di facile verifica:

$$\binom{n}{k} = \binom{n}{n-k} \qquad \binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \tag{2.3}$$

Una dimostrazione delle proprietà (2.2)-(2.3) sarà data nel paragrafo 7. Quest'ultima relazione permette di calcolare i numeri $\binom{n}{k}$ per mezzo del cosiddetto *triangolo di Tartaglia*; in cima al triangolo si pone il numero $\binom{0}{0} = 1$ (per definizione); ai lati si pongono i numeri $\binom{n}{0} = \binom{n}{n} = 1$ per ogni $n \geq 1$ e allora, per $0 < k < n$, il numero $\binom{n}{k}$ viene scritto all'incrocio della n-esima riga e della k-esima colonna e risulta somma dei due numeri che si trovano nella riga precedente, quello sulla stessa colonna e quello sulla colonna precedente.

potenza	coefficienti					
$n = 0$	1					
1	1	1				
2	1	2	1			
3	1	3	3	1		
4	1	4	6	4	1	
5	1	5	10	10	5	1
.....						

Figura 1.3.

Per esempio, abbiamo $(a+b)^5 = a^5 + 5a^4b + 10a^3b^2 + 10a^2b^3 + 5ab^4 + b^5$.

Esercizi

1. Usando il triangolo di Tartaglia scrivere esplicitamente

$$(1+a)^8 \qquad (2a-3b)^7$$

2. Risolvere l'equazione: $4\binom{x}{4} = 15\binom{x-2}{3}$, $x \in \mathbb{N}$.

Dimostrare le seguenti identità, sfruttando le proprietà delle sommatorie e i suggerimenti forniti:

3. $$\sum_{i=1}^{n} i = \frac{n(n+1)}{2}$$

Suggerimento: calcolare $2\sum_{i=1}^{n} i = \sum_{i=1}^{n} i + \sum_{i=1}^{n} i$ eseguendo nella seconda sommatoria una riflessione di indici.

4. $$\sum_{k=0}^{n-1} (2k+1) = n^2$$

Suggerimento: sfruttare il risultato dell'esercizio precedente.

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

Suggerimento: $\sum_{k=0}^{n} (k+1)^3 = \sum_{k=0}^{n} \left(k^3 + 3k^2 + 3k + 1\right) = \sum_{k=1}^{n} k^3 + 3\sum_{k=1}^{n} k^2 + \ldots$ D'altra parte, $\sum_{k=0}^{n} (k+1)^3 = \sum_{k=1}^{n+1} k^3 = \sum_{k=1}^{n} k^3 + (n+1)^3$. Dal confronto tra le due espressioni...

6 Calcolare esplicitamente la seguente somma, semplificando opportunamente:

$$\sum_{k=1}^{100}\left(\frac{1}{k}-\frac{1}{k+1}\right)$$

7 Calcolare esplicitamente le seguenti somme, sfruttando la formula per una progressione geometrica:

$$\sum_{k=0}^{30}(-1)^k\,\frac{2^{3k+1}}{3^k} \qquad \sum_{k=2}^{100}3^{2-k}$$

3 CAMPI ORDINATI

Inizieremo ora a studiare più da vicino la struttura degli insiemi numerici che abbiamo introdotto in precedenza, ed in particolare l'insieme $\mathbb{Q}$ dei numeri razionali e l'insieme $\mathbb{R}$ dei numeri reali. Chiariamo subito che non intendiamo presentare una costruzione rigorosa di questi insiemi. Piuttosto, dando per presupposto che lo studente sappia già cosa sono i numeri razionali e reali (ovvero accontentandoci della descrizione ingenua di questi insiemi che abbiamo presentato nel paragrafo 1), puntualizzeremo con più precisione quali sono le *proprietà* di cui godono questi insiemi numerici.

Iniziamo quindi col richiamare le proprietà (ben note) dei numeri razionali; indichiamo con a, b, c... generici numeri razionali.

• R_1. È definita in $\mathbb{Q}$ un'operazione (detta *addizione* o *somma*) che ha le seguenti proprietà:

1. $\forall\, a, b$ $\qquad a+b=b+a$ $\qquad$ (proprietà commutativa)
2. $\forall\, a, b, c$ $\qquad (a+b)+c=a+(b+c)$ $\qquad$ (proprietà associativa)
3. Esiste un elemento (elemento neutro della somma, detto *zero* e indicato con 0) tale che $\forall\, a$, $a+0=a$
4. $\forall\, a$, esiste un elemento (l'inverso di a rispetto alla somma, detto *opposto* di a e indicato con $-a$) tale che $a+(-a)=0$.

• R_2. È definita in $\mathbb{Q}$ un'operazione (detta *moltiplicazione* o *prodotto*) che ha le seguenti proprietà:

1. $\forall\, a, b$ $\qquad a\cdot b=b\cdot a$ $\qquad$ (proprietà commutativa).
2. $\forall\, a, b, c$ $\qquad (a\cdot b)\cdot c=a\cdot(b\cdot c)$ $\qquad$ (proprietà associativa).
3. Esiste un elemento (elemento neutro del prodotto, detto *unità* e indicato con 1) tale che $\forall\, a$, $a\cdot 1=a$.
4. $\forall\, a\neq 0$, esiste un elemento (l'inverso di a rispetto al prodotto, detto *reciproco* di a e indicato con $\frac{1}{a}$ oppure a^{-1}) tale che $a\cdot a^{-1}=1$.

Le operazioni di somma e prodotto sono legate dalla seguente proprietà:

5. $\forall\, a, b, c$ $\qquad (a+b)\cdot c=a\cdot c+b\cdot c$ $\qquad$ (proprietà distributiva).

Si osservi che le proprietà delle due operazioni sono quasi le stesse; le uniche differenze sono: il fatto che, per il prodotto, non esiste l'inverso dell'elemento neutro della somma (cioè non esiste $1/0$), e la proprietà distributiva (in cui somma e prodotto giocano un ruolo asimmetrico). Questi fatti fanno sì che le due operazioni siano sostanzialmente diverse.

Dalle proprietà R_1 e R_2 discende la possibilità di eseguire senza restrizioni le quattro operazioni fondamentali: addizione e moltiplicazione (sopra definite), sottrazione (ponendo $a - b = a + (-b)$) e divisione (ponendo $a : b = ab^{-1}$ purché sia $b \neq 0$).

Rappresentazione geometrica

Una rappresentazione di tipo geometrico di $\mathbb{Q}$ si può ottenere associando a ogni numero razionale un punto della retta euclidea. Ad un punto di questa, scelto arbitrariamente, si associa 0 e a un altro, distinto dal primo, si associa 1, individuando così il segmento orientato 01 che costituisce l'unità di misura. A questo punto si ha una corrispondenza biunivoca tra i numeri razionali e quei punti P della retta che sono estremi dei segmenti orientati $0P$ commensurabili (nel senso classico della geometria elementare, cioè che hanno un multiplo comune) con 01.

Figura 1.4.

Per descrivere compiutamente la struttura dell'insieme $\mathbb{Q}$ osserviamo anche che esso è un *insieme ordinato*; cioè che nell'insieme $\mathbb{Q}$ è definita la relazione $\leq$ (minore o uguale di...; analogamente la relazione $\geq$: maggiore o uguale di ...), la quale è una *relazione d'ordine*, cioè verifica le tre proprietà seguenti:

1.	$\forall a$	$a \leq a$	riflessiva
2.	$\forall a, b$	se $a \leq b$ e $b \leq a$ allora $a = b$	antisimmetrica
3.	$\forall a, b, c$	se $a \leq b$ e $b \leq c$ allora $a \leq c$	transitiva

Inoltre, presi due qualsiasi numeri razionali a, b, è sempre possibile confrontarli per mezzo della relazione $\leq$, nel senso che vale necessariamente almeno una delle relazioni $a \leq b$ o $b \leq a$. Si esprime questo fatto dicendo che l'ordinamento di $\mathbb{R}$ è totale, o che $\mathbb{R}$ è un *insieme totalmente ordinato*.

Osservazione Si rifletta sulla natura generale astratta delle proprietà che definiscono una relazione d'ordine. Ad esempio, se X è un insieme qualsiasi (per esempio $\mathbb{Q}$), nell'insieme $\mathcal{P}(X)$ (insieme di tutti i sottoinsiemi di X) possiamo considerare la relazione $\subseteq$ di inclusione insiemistica; chiaramente, per ogni $A, B, C \subseteq X$, è vero che

1.	$A \subseteq A$		(proprietà riflessiva);
2.	$A \subseteq B$ e $B \subseteq A$	implica $A = B$	(proprietà antisimmetrica);
3.	$A \subseteq B$ e $B \subseteq C$	implica $A \subseteq C$	(proprietà transitiva).

Quindi $\mathcal{P}(X)$ è un insieme ordinato, rispetto alla relazione d'ordine $\subseteq$. Si osservi tuttavia che non è un insieme *totalmente* ordinato. Infatti, dati due insiemi qualsiasi $A, B \subseteq X$, in generale può non valere né $A \subseteq B$ né $B \subseteq A$. Quindi: non ogni insieme ordinato è totalmente ordinato.

Mettiamo in evidenza le proprietà dell'ordinamento dei numeri razionali:

• R_3. È definita in $\mathbb{Q}$ una relazione d'ordine totale $\leq$, compatibile con la struttura algebrica, cioè:

 © 978-88-08-06485-1

1. $\forall a, b, c$ se $a \leq b$ allora $a + c \leq b + c$
2. $\forall a, b, c$ con $c > 0$ se $a \leq b$ allora $ac \leq bc$.

Osserviamo che tutte le regole ben note del calcolo algebrico derivano dalle proprietà R_1, R_2, R_3. A titolo di esempio, i "principi di identità" usati nel risolvere le disequazioni, non sono altro che le proprietà di compatibilità della relazione d'ordine con la somma e col prodotto: tutte le usuali procedure con cui si risolvono le disequazioni sono quindi conseguenza di questi assiomi e delle proprietà algebriche della somma e del prodotto, espresse da R_1, R_2.

Astraendo ora dal caso particolare dell'insieme dei numeri razionali, diamo la seguente

Definizione 1.2 Chiameremo *campo ordinato* un insieme in cui sono definite due operazioni (che chiamiamo somma e prodotto) e una relazione d'ordine, che soddisfano tutte le proprietà R_1, R_2, R_3. Un insieme con le proprietà R_1, R_2 si dice *campo*. ■

Un momento di riflessione mostra che tutto ciò che abbiamo detto fin qui riguardo alle operazioni di somma e prodotto e alla relazione d'ordine $\leq$ nell'insieme dei numeri razionali, vale anche per l'insieme dei numeri reali. Possiamo quindi sintetizzare il punto fondamentale di questo paragrafo nella seguente affermazione:

Sia $\mathbb{Q}$ sia $\mathbb{R}$ sono campi ordinati.

Da questo punto di vista, quindi, $\mathbb{Q}$ ed $\mathbb{R}$ appaiono dotati di proprietà simili. Nel prossimo paragrafo invece evidenzieremo qual è la proprietà che distingue sostanzialmente $\mathbb{R}$ da $\mathbb{Q}$ e che rende $\mathbb{R}$ l'ambiente giusto per sviluppare l'analisi matematica.

4 NUMERI REALI. ESTREMO SUPERIORE E ASSIOMA DI CONTINUITÀ

4.1 Inadeguatezza dell'insieme dei razionali per misurare le lunghezze

La struttura di campo ordinato dei razionali assolve alla maggior parte degli scopi pratici del calcolo, nel senso che si può esprimere con un numero razionale la misura di ogni grandezza con sufficiente precisione. Tuttavia è noto che ci sono grandezze che non sono commensurabili tra loro: l'esempio classico è dato dalla diagonale e dal lato di un quadrato. Con riferimento alla figura 1.5, se il lato misura 1, l'ascissa d che misura la diagonale non è un numero razionale. Infatti, per il Teorema di Pitagora, $d^2 = 1^2 + 1^2 = 2$ ma, come abbiamo dimostrato nel par. 1.2 (Teorema 1.2), non esiste un numero razionale il cui quadrato è 2.

Dunque il punto d sulla retta non è il rappresentante di alcun numero razionale. Ciò significa che, dopo aver "occupato" i punti della retta con i numeri razionali, su

Figura 1.5.

di essa rimangono ancora dei posti vuoti. Significa anche che *l'insieme dei numeri razionali è inadeguato ad esprimere le lunghezze dei segmenti.* Ma la lunghezza è solo la più semplice delle grandezze che si incontrano in geometria e in fisica: un problema esattamente analogo si incontra nella misurazione di qualsiasi altra grandezza, come aree, volumi, tempi, velocità ecc.

Sorge spontanea allora la domanda: è possibile ampliare l'insieme dei razionali in modo da avere ancora un campo ordinato, i cui elementi (numeri) siano in *corrispondenza biunivoca* con i punti della retta euclidea? Il candidato naturale a questa estensione è l'insieme dei numeri reali. La domanda dunque diventa: l'insieme $\mathbb{R}$ può essere messo in corrispondenza biunivoca con i punti della retta? Vedremo che la risposta a questa domanda (pur di riformularla in modo puramente analitico) è positiva. Per renderci conto di questo dobbiamo però introdurre un nuovo concetto, che ci porterà alla comprensione della fondamentale *proprietà di continuità di* $\mathbb{R}$, che si rivelerà essenziale per l'analisi matematica.

4.2 Estremo superiore e assioma di continuità

Sia X un qualunque insieme totalmente ordinato. In pratica, ci interesserà il caso $X = \mathbb{Q}$ o $X = \mathbb{R}$, ma è utile non specificare per ora quale dei due ambienti consideriamo.

Sia E un insieme contenuto in X. Tale insieme si dice *limitato superiormente* se esiste un numero M per cui risulti $x \leq M$ per ogni elemento $x \in E$; *limitato inferiormente* se esiste un numero m per cui risulti risulta $x \geq m$ per ogni elemento $x \in E$; *limitato* se valgono entrambe le condizioni, ossia esistono due numeri, m e M, tali che

$$m \leq x \leq M \text{ per ogni } x \in E.$$

Introduciamo ora il concetto di elemento *massimo* (*minimo*) di un insieme. Diremo che un elemento $\overline{x}$ è *massimo* per E se:

i) $\overline{x} \in E$;
ii) $x \leq \overline{x}\ \forall x \in E$.

Analoga definizione per il minimo $\underline{x}$. È evidente che, affinché il massimo (minimo) esista, l'insieme deve essere superiormente (inferiormente) limitato.

Esempi 4.1

	Insieme E	Max	Min
I)	$\mathbb{N}$	non esiste	0
II)	Numeri pari relativi	non esiste	non esiste
III)	$\left\{1, \frac{1}{2}, \frac{1}{3}, \dots, \frac{1}{n}, \dots\right\}$	1	non esiste
IV)	$\left\{n \in \mathbb{N} : \frac{n-1}{n+1}\right\}$	non esiste	-1
V)	$\left\{x \in \mathbb{R} : x^3 \geq 27\right\}$	non esiste	3
VI)	$\left\{x \in \mathbb{Q} : x \geq 0, x^2 < 2\right\}$	non esiste	0

Si osserva, negli esempi precedenti, che talvolta, pur essendo l'insieme E limitato, esso non possiede massimo o minimo.

Introduciamo allora un concetto, fondamentale per l'Analisi matematica, che è sostitutivo del concetto di massimo o minimo: quello di *estremo superiore* (sup) ed *estremo inferiore* (inf).

Definizione 1.3 Sia $E \subseteq X$. Un numero $k \in X$ (non necessariamente $\in E$) si dice un *maggiorante* (*minorante*) di E se $k \geq x$ ($k \leq x$) $\forall x \in E$. Notiamo che un insieme superiormente (inferiormente) limitato ha molti maggioranti (minoranti). Chiameremo *estremo superiore* (*inferiore*) di E, e lo indicheremo con $\sup E$ ($\inf E$) il minimo (massimo) dei maggioranti (minoranti) di E (se esiste). ■

È evidente che, se l'insieme possiede massimo (minimo), questo coincide con l'estremo superiore (inferiore).

Riprendiamo gli esempi precedenti per vedere se, quando non esiste il massimo o il minimo, esiste però il sup o l'inf. Negli esempi I, II, V la risposta è negativa per l'evidente ragione che gli insiemi ivi considerati non sono limitati (da sopra e/o da sotto). Nell'esempio III è facile vedere che l'inf è 0 (il quale non è minimo poiché 0 non fa parte dell'insieme) e nell'esempio IV il sup è 1. Consideriamo l'esempio VI; si intuisce che il sup dovrebbe essere un numero il cui quadrato è 2; ma in $\mathbb{Q}$ tale numero non esiste; esiste però in $\mathbb{R}$ ed è $\sqrt{2}$. Questa circostanza non è casuale, ma illustra precisamente la differenza tra l'insieme $\mathbb{Q}$ dei numeri razionali e l'insieme $\mathbb{R}$ dei numeri reali. Per chiarirla, diamo la seguente definizione generale:

Si dice che l'insieme X (totalmente ordinato) possiede la *proprietà dell'estremo superiore*, se

- R_4. Ogni insieme $E \subset X$ non vuoto e limitato superiormente possiede estremo superiore in X.

Si prova facilmente che, se vale questa proprietà, allora è anche vero che ogni sottoinsieme di X non vuoto e inferiormente limitato ammette estremo inferiore.

Si faccia attenzione al contenuto della precedente definizione: non si richiede che X *stesso* abbia estremo superiore (ad esempio, nel caso $X = \mathbb{Q}$ o $\mathbb{R}$ questo è certamente falso!), ma che *ogni sottoinsieme non vuoto superiormente limitato di* X ne sia provvisto. L'esempio VI mostra che certamente $\mathbb{Q}$ *non ha* la proprietà dell'estremo superiore. Invece, $\mathbb{R}$ ha questa proprietà. Questo fatto non può essere dimostrato a partire dalla nostra "definizione ingenua" di $\mathbb{R}$, data in precedenza. Tuttavia nella presentazione assiomatica di $\mathbb{R}$, questa proprietà costituisce parte della definizione stessa di $\mathbb{R}$:

Definizione 1.4 (Definizione assiomatica di $\mathbb{R}$) Chiamiamo $\mathbb{R}$ un insieme che soddisfa le proprietà R_1, R_2, R_3, R_4, ossia un campo ordinato che ha la proprietà dell'estremo superiore. ■

Ricordiamo ancora, per confronto, che l'insieme $\mathbb{Q}$ soddisfa R_1, R_2, R_3, ma non R_4. La proprietà dell'estremo superiore prende anche il nome di *assioma di Dedekind*, o assioma di *continuità*, o di *completezza* e si può enunciare anche nella seguente forma equivalente.

Sia $\{A, B\}$ una partizione di $\mathbb{R}$ (cioè A e B sono insiemi non vuoti e disgiunti la cui unione è $\mathbb{R}$); essa si chiama *sezione* se: $\forall\, a \in A$ e $\forall\, b \in B$ risulta $a < b$. Allora si dimostra che:

- R_4'. Per ogni sezione $\{A, B\}$ di $\mathbb{R}$ esiste un unico numero reale s (detto elemento separatore) tale che

$$a \leq s \leq b \qquad \forall\, a \in A, \forall\, b \in B$$

(tale elemento separatore altro non è che $\sup A = \inf B$).

Ritorniamo ora al problema dell'incommensurabilità di lato e diagonale del quadrato. Quando, con riga e compasso, fissato il lato unitario del quadrato, ne costruiamo la diagonale e la riportiamo sulla retta di partenza, l'arco tracciato dal compasso "spezza la retta in due", generando quella che abbiamo chiamato una *sezione*; l'elemento separatore, che esiste per l'assioma R_4 rappresenta geometricamente il punto di intersezione, e quindi quel numero misura la diagonale. Quindi l'interpretazione geometrica della proprietà R_4 ci mostra come l'insieme $\mathbb{R}$ sia una rappresentazione adeguata della nostra idea intuitiva di retta, così adeguata che spesso in matematica si usa l'espressione "la retta reale" per indicare $\mathbb{R}$, confondendo l'insieme numerico con la sua rappresentazione geometrica naturale. Ricordiamo che l'assioma di continuità è stato messo in luce da R. Dedekind nel 1872, ossia circa 2400 anni dopo la scoperta pitagorica della incommensurabilità di lato e diagonale del quadrato.

Come vedremo nel resto del corso, le conseguenze *analitiche* dell'assioma di continuità sono rilevanti per l'Analisi matematica.

4.3 Valore assoluto. Disuguaglianza triangolare

Si dice *valore assoluto* del numero reale a (o *modulo* di a) il numero non negativo così definito:

$$|a| = \begin{cases} a & \text{se} \quad a \geq 0 \\ -a & \text{se} \quad a < 0 \end{cases} \tag{4.1}$$

Dalla definizione di valore assoluto segue immediatamente che:

$$\forall\, a \geq 0: \quad |x| \leq a \iff -a \leq x \leq a \tag{4.2}$$

Dalla (4.2) segue l'importante *disuguaglianza triangolare*:

$$\forall\, x, y \in \mathbb{R}: \quad |x + y| \leq |x| + |y| \tag{4.3}$$

Infatti, basta scrivere le relazioni:

$$-|x| \leq x \leq |x| \qquad \text{e} \qquad -|y| \leq y \leq |y|$$

quindi sommare membro a membro, ottenendo

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

da cui, per la (4.2), segue la (4.3).

La disuguaglianza triangolare è spesso usata nella forma seguente:

$$|a-b| \leq |a-c| + |b-c| \quad \forall\, a,b,c \in \mathbb{R} \tag{4.4}$$

Basta porre nella (4.3) $x = a - c$ e $y = c - b$. Se invece, sempre nella (4.3), si pone $x = a - b$ e $y = b$ si ottiene

$$|a| \leq |a-b| + |b| \quad \text{cioè} \quad |a| - |b| \leq |a-b|$$

Analogamente, scambiando a con b otteniamo: $|b| - |a| \leq |a-b|$ e dunque, ricordando la (4.2),

$$\big||a| - |b|\big| \leq |a-b| \tag{4.5}$$

L'espressione $|a-b|$ rappresenta geometricamente la distanza (nel senso della geometria elementare) dei due punti a e b sulla retta.

La (4.3) può facilmente estendersi al caso di n addendi:

$$\left|\sum_{k=1}^{n} x_k\right| \leq \sum_{k=1}^{n} |x_k| \tag{4.6}$$

Valgono anche le seguenti proprietà immediate:

$$|ab| = |a|\,|b| \qquad \left|\frac{a}{b}\right| = \frac{|a|}{|b|} \qquad |-a| = |a| \tag{4.7}$$

4.4 Intervalli

Dati due numeri reali a, b, si chiama *intervallo di estremi* a *e* b uno dei seguenti insiemi:

$[a,b]$, insieme dei numeri reali x, $a \leq x \leq b$

$[a,b)$, " " " " x, $a \leq x < b$

$(a,b]$, " " " " x, $a < x \leq b$

(a,b), " " " " x, $a < x < b$

Come si può notare la parentesi quadra (tonda) in corrispondenza di uno dei due estremi indica che quest'ultimo è incluso (escluso) nell'intervallo.

Gli intervalli $[a,b]$ si dicono *chiusi*; quelli (a,b) si dicono *aperti*. Tutti gli intervalli indicati sopra sono *limitati*; si chiamano intervalli (illimitati) anche le semirette, per esempio:

$(-\infty, b)$, insieme dei numeri reali x, $x < b$

$[a, +\infty)$, " " " " x, $x \geq a$

o l'intera retta $\mathbb{R} = (-\infty, +\infty)$.

Figura 1.6.

Si può dimostrare che gli intervalli, limitati o illimitati, sono tutti e soli i sottoinsiemi I di $\mathbb{R}$ che soddisfano la seguente proprietà caratteristica (detta *connessione*): presi comunque tre numeri reali

$$x_1 < x_2 < x_3, \text{ se } x_1, x_3 \in I, \text{ allora } x_2 \in I.$$

Nel seguito ci capiterà di considerare il prodotto cartesiano di due (o più) intervalli, cui si può dare il significato geometrico di rettangolo (in due dimensioni) o parallelepipedo (in tre dimensioni).

Esempio

4.2 Sia $A = [0, 1]$, $B = [1, 2]$; la figura 1.2 di pagina 8 illustra gli insiemi $A \times B$, $B \times A$, $A \times A$ indicato anche con A^2.

5 RADICALI, POTENZE, LOGARITMI

In conseguenza della proprietà R_4 possiamo eseguire, nel campo reale, operazioni che sono solo occasionalmente possibili nel campo razionale, come l'estrazione di radice o l'elevamento a potenza.

5.1 Radici n-esime aritmetiche

Teorema 1.3 *Sia $y \in \mathbb{R}, y > 0$ e n un intero ≥ 1. Esiste un unico numero reale positivo x tale che $x^n = y$.*

Tale numero si chiama *radice n-esima aritmetica di y* e si indica con uno dei simboli $\sqrt[n]{y}$ oppure $y^{1/n}$. Si noti che la radice n-esima aritmetica è non negativa, per esempio $\sqrt{4} = 2$, $\sqrt{9} = 3$ e più in generale $\sqrt{x^2} = |x|$.

Una semplice dimostrazione di questo teorema potrà essere data più avanti nel corso, utilizzando le proprietà delle *funzioni continue* (si veda il cap. 3, par. 4.1). Per ora ci limitiamo a mostrare, con un esempio, come si può costruire la rappresentazione decimale della radice n-esima. Cerchiamo l'allineamento decimale di $\sqrt{2}$; questo numero, non essendo razionale, sarà rappresentato da un allineamento infinito (non periodico). Si procede così: si costruisce una classe di numeri razionali della forma

$$\begin{aligned} 0 < & a_0, \\ & a_0, a_1 \\ & a_0, a_1 a_2 \\ & a_0, a_1 a_2 a_3 \\ & \dots\dots\dots\dots \\ & a_0, a_1 a_2 a_3 \dots a_n \\ & \dots\dots\dots\dots\dots\dots \end{aligned}$$

con la regola che: ognuno di questi numeri è il più grande tra quelli con lo stesso numero di decimali dopo la virgola il cui quadrato è minore di 2:

$$\begin{array}{ccc} 1 & 1^2 = 1 & 2^2 = 4 \\ 1{,}4 & (1,4)^2 = 1,96 & (1,5)^2 = 2,25 \\ 1,41 & (1,41)^2 = 1,9881 & (1,42)^2 = 2,0164 \\ \dots & \dots\dots\dots\dots & \dots\dots\dots\dots \end{array}$$

Questo insieme di numeri (chiamiamolo E_-) è limitato superiormente (ognuno è < 2); per la proprietà R_4 esso possiede estremo superiore; il numero $\sqrt{2}$ si definisce precisamente come $\sup E_-$.

Si sarebbe potuto anche costruire una classe di numeri E_+ come la precedente con la regola che ognuno di essi sia il più piccolo tra quelli con lo stesso numero di decimali dopo la virgola il cui quadrato è > 2; avremmo ottenuto

$$2 \qquad 1{,}5 \qquad 1{,}42 \qquad \dots$$

Questo insieme E_+ è limitato inferiormente (ogni elemento è > 1), perciò possiede estremo inferiore; si dimostra che $\inf E_+ = \sup E_-$. I numeri della classe E_- approssimano $\sqrt{2}$ per difetto, quelli della classe E_+ per eccesso.

5.2 Potenze a esponente reale

L'estrazione di radice n-esima è l'operazione inversa dell'elevamento a potenza intera. Si può estendere l'operazione di elevamento a potenza per *ogni* esponente razionale se la base è *positiva* (utilizzando il teorema precedente) ponendo

$$\boxed{\text{se} \quad r = \frac{m}{n} \quad a > 0 \quad a^r := (a^m)^{\frac{1}{n}} = \sqrt[n]{a^m}}$$

(si assume: $m \in \mathbb{Z}$, n positivo).

Se infine l'esponente è reale $b = b_0, b_1 b_2 \dots b_n \dots$ il numero a^b $(a > 0)$ sarà *individuato* dalla classe di numeri

$$a^{b_0} \qquad a^{b_0, b_1} \qquad a^{b_0, b_1 b_2} \qquad \dots$$

in un modo simile a quello del caso della radice.

Se la base a è negativa l'operazione di elevamento a potenza a^b è definita solo in certi casi: se l'esponente b è intero, oppure razionale n/m purché *non sia* n dispari ed m pari. Infatti: si pone $a^{n/m} = \sqrt[m]{a^n}$ e inoltre, se $c < 0$ e m dispari, si definisce $\sqrt[m]{c} = -\sqrt[m]{-c}$. Per esempio:

$$(-2)^{3/5} = \sqrt[5]{-8} - -\sqrt[5]{8}$$

$$(-2)^{2/5} = \sqrt[5]{4}$$

$$(-2)^{3/4} = \sqrt[4]{-8} \quad \textit{non esiste nel campo reale!}$$

Quando si dice "non esiste in $\mathbb{R}$" si intende che non è possibile definire tale operazione in modo da mantenere valide le usuali regole di calcolo. Qualcosa di più sull'argomento sarà detto nel paragrafo 8, parlando dei numeri complessi.

Le proprietà principali dell'elevamento a potenza sono:

siano a, b reali positivi, c, d reali qualsiasi

E_0 $\quad a^0 = 1 \quad \forall a \neq 0; \quad 1^c = 1 \quad \forall c$

E_1 $\quad a^c > 0 \quad \forall c; \quad a^c \lesseqgtr 1$ se $a \lesseqgtr 1$ e $c > 0$

E_2 $\quad a^{c+d} = a^c \cdot a^d$

E_3 $\quad (ab)^c = a^c \cdot b^c$

E_4 $\quad (a^b)^c = a^{bc}$

E_5 $\quad c < d \implies a^c \lesseqgtr a^d$ se $a \gtreqless 1$

E_6 $\quad 0 < a \leq b \implies a^c \leq b^c \quad \forall c > 0$

5.3 Logaritmi

Consideriamo l'equazione

$$a^x = y \qquad a > 0$$

con y assegnato e x incognito.

Anzitutto, se $a = 1$, essa è risolubile solo se $y = 1$ (e in tal caso ogni numero reale x è soluzione). Sia dunque $a \neq 1$. Se $y \leq 0$ essa non ha alcuna soluzione (cfr. E_1). Il seguente teorema ci dice che essa ha una sola soluzione per ogni $y > 0$.

Teorema 1.4 *Sia $a > 0$, $a \neq 1$, $y > 0$. Esiste un unico numero reale x tale che $a^x = y$.*

Tale numero prende il nome di *logaritmo in base a di y* e si indica col simbolo $\log_a y$. Quindi, per definizione: $a^{\log_a y} = y$. Le proprietà dei logaritmi, che si deducono da quelle degli esponenziali, sono:

siano x, y, a reali *positivi*, $a \neq 1$

L_1 $\quad \log_a xy = \log_a x + \log_a y$

L_2 $\quad \log_a \dfrac{x}{y} = \log_a x - \log_a y$

L_3 $\quad \log_a x^\alpha = \alpha \log_a x \qquad \forall \alpha \in \mathbb{R}$

L_4 $\quad \log_a x = \dfrac{1}{\log_x a} = -\log_{\frac{1}{a}} x \qquad (x \neq 1)$

L_5 $\quad \log_b x = \dfrac{\log_a x}{\log_a b} \quad \forall b > 0,\ b \neq 1$

Nel capitolo 2 introdurremo il concetto di *funzione* e vedremo che le operazioni sopra definite consentono di introdurre tre importanti classi di funzioni: le *potenze*, gli *esponenziali* e i *logaritmi*.

5.4 Approssimazioni

Concludiamo questo paragrafo con qualche osservazione sulle approssimazioni che si fanno lavorando coi numeri reali. Cominciamo col ricordare che un numero *razionale* può essere sempre espresso con precisione assoluta, sia ricorrendo alla scrittura frazionaria che a quella decimale (eventualmente con cifre periodiche). Invece non è possibile, naturalmente, scrivere *tutte* le cifre decimali di un numero irrazionale, dato che queste sono infinite e si susseguono senza periodicità. Cosa significa allora "conoscere" o "specificare" un numero irrazionale? Significa conoscere qualche algoritmo che ci consenta (almeno teoricamente) di scrivere tante cifre decimali esatte quante ne desideriamo. Nell'esempio (già considerato) del numero irrazionale

$$0{,}1010010001\ldots.$$

(formato in base alla regola: scrivere una cifra 1, una cifra 0, una cifra 1, due cifre 0, una cifra 1, tre cifre 0, e così via) è chiaro che potremmo scrivere tante cifre quante ne desideriamo. In altri casi, come $\sqrt{2}$ o $\log_2 3$, le cose sono più laboriose, richiedendo l'esecuzione di calcoli iterativi per determinare ogni successiva cifra decimale; tuttavia, questi calcoli sono effettivamente eseguibili. Altre volte, in Analisi matematica, un numero viene specificato in modo non costruttivo, denotandolo come l'unico numero che risolve un determinato problema (una volta che si sia dimostrato appunto che tale soluzione esiste ed è unica). Si tratta di un modo operativamente meno soddisfacente, ma teoricamente ineccepibile.

Veniamo ora agli aspetti pratici. Ogni volta che eseguiamo calcoli con una calcolatrice tascabile, questa scriverà solo numeri con un numero fissato di cifre decimali dopo la virgola (tipicamente 9). Questo significa che stiamo lavorando solo con numeri razionali, anzi con un sottoinsieme *finito* dell'insieme $\mathbb{Q}$! Lavorando con un computer le cose migliorano un po', ma rimaniamo comunque nell'ambito di sottoinsiemi finiti di $\mathbb{Q}$. Occorre naturalmente esserne consapevoli.

Oltre all'approssimazione generata dagli strumenti di calcolo, a volte siamo noi stessi a non essere interessati a troppe cifre decimali; introduciamo allora volutamente delle approssimazioni, scrivendo ad esempio

$$\sqrt{2} \simeq 1{,}414.$$

Si ricordi, a tale proposito, la regola di *arrotondamento* espressa dai seguenti esempi:

$$2{,}4138 \simeq 2{,}41$$
$$2{,}4152 \simeq 2{,}42.$$

Esplicitamente: l'ultima cifra che si scrive viene arrotondata all'unità inferiore (superiore) se la prima cifra che si trascura è da 0 a 4 (rispettivamente, da 5 a 9).

6 INSIEMI INFINITI

Accenniamo ora ad una questione fondamentale che ha a che fare con gli insiemi, e precisamente il problema di come si possa confrontare la "numerosità" degli insiemi infiniti. Avvertiamo che ciò che diremo in questo paragrafo non sarà direttamente uti-

lizzato nel seguito del corso; tuttavia, ci sembra giusto dare spazio a questo problema, al tempo stesso così naturale e difficile[4].

Affrontiamo il discorso a partire dagli insiemi numerici notevoli che abbiamo introdotto: $\mathbb{N}, \mathbb{Z}, \mathbb{Q}, \mathbb{R}$. Chiediamoci: quanti sono gli elementi di ciascuno di questi insiemi?

Intuitivamente, la risposta sembra ovvia: ciascuno di questi insiemi ha *infiniti* elementi; tuttavia i numeri razionali sono più numerosi dei numeri interi, essendo $\mathbb{Z} \subset \mathbb{Q}$, e per lo stesso motivo i numeri reali sono di più dei numeri razionali. Queste risposte pongono però un problema: come possiamo affermare, al tempo stesso, che due insiemi sono *entrambi* infiniti, ma uno è *più numeroso* dell'altro? Com'è possibile, in generale, confrontare la numerosità degli insiemi infiniti? Ad esempio, ci sono più punti in una sfera o in un piano? (In questo caso, nessuno dei due insiemi è contenuto nell'altro).

Per dar senso a queste domande, prima ancora che per sapervi rispondere, occorre definire cosa si intenda per uguale numerosità di due insiemi. Astraendo dall'esperienza del *contare* gli elementi di un insieme finito, si è giunti a identificare l'idea di uguale numerosità con quella di *corrispondenza biunivoca*.

Definizione 1.5 Due insiemi A, B si dicono di uguale *cardinalità* o *potenza* (termini che traducono l'idea intuitiva di *numerosità*) se possono essere messi in corrispondenza biunivoca tra loro, cioè se esiste una legge che associa ad ogni elemento di A uno e un solo elemento di B, e viceversa. ■

Si osservi, ad esempio, la seguente corrispondenza tra gli insiemi $\mathbb{Z}$ e $\mathbb{N}$:

$$\begin{array}{cccccccccc}
\mathbb{Z} & 0 & 1 & -1 & 2 & -2 & \dots & n & -n & \dots \\
 & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & & \updownarrow & \updownarrow & \\
\mathbb{N} & 0 & 1 & 2 & 3 & 4 & \dots & 2n-1 & 2n & \dots
\end{array}$$

Come si vede, la legge ora definita realizza una corrispondenza biunivoca tra $\mathbb{Z}$ e $\mathbb{N}$. Quindi, anche se dal punto di vista dell'inclusione $\mathbb{Z}$ ha "più elementi" di $\mathbb{N}$ (nel senso che ha tutti gli elementi di $\mathbb{N}$ più altri), gli insiemi hanno la stessa cardinalità (si dice anche che sono *equipotenti*). In questo senso, quindi, vanno pensati come *ugualmente numerosi*.

Definizione 1.6 Si dice *numerabile* un insieme che ha la stessa cardinalità di $\mathbb{N}$. ■

Ad esempio, $\mathbb{Z}$ è numerabile. Il termine numerabile indica che gli elementi dell'insieme si possono enumerare, ossia disporre in un elenco numerato (posto 1, posto 2, posto 3,...).

Di più, possiamo dimostrare che:

Teorema 1.5 *$\mathbb{Q}$ è numerabile.*

[4] Le ricerche sulla numerosità degli insiemi infiniti iniziarono nel 19° secolo con Bolzano, Dedekind, e soprattutto Cantor, il cui nome è legato in maniera speciale a questo tema.

DIMOSTRAZIONE. Cominciamo a dimostrare che l'insieme dei numeri razionali *positivi* è numerabile. Per far questo, rappresentiamo i numeri razionali positivi come frazioni n/m, con n, m interi positivi, e disponiamo queste frazioni in una tabella triangolare infinita, al modo seguente:

$\frac{n}{m}$, con $n+m=\dots$				
2	$\frac{1}{1}$			
3	$\frac{1}{2}$	$\frac{2}{1}$		
4	$\frac{1}{3}$	$\frac{2}{2}$	$\frac{3}{1}$	
5	$\frac{1}{4}$	$\frac{2}{3}$	$\frac{3}{2}$	$\frac{4}{1}$
...	...			

Si noti che: tutte le frazioni che rappresentano numeri razionali positivi compaiono in questa tabella almeno una volta; alcuni sono ripetuti più volte (ad es. $1/1 = 2/2$); ogni riga ha lunghezza finita. Quindi possiamo mettere in corrispondenza biunivoca $\mathbb{N}$ con l'insieme dei razionali positivi, percorrendo la tabella, riga dopo riga (e saltando un elemento quando è uguale ad uno già incontrato). Ad esempio:

1	2	3	4	5	6	7	8	9	...
$\frac{1}{1}$	$\frac{1}{2}$	$\frac{2}{1}$	$\frac{1}{3}$	$\frac{3}{1}$	$\frac{1}{4}$	$\frac{2}{3}$	$\frac{3}{2}$	$\frac{4}{1}$	...

(si noti che abbiamo saltato $2/2$ perché uguale a $1/1$, già incontrato). Questo dimostra che l'insieme dei razionali positivi è numerabile; allora anche $\mathbb{Q}$ è numerabile, e questo si può provare con una dimostrazione analoga a quella con cui abbiamo provato che $\mathbb{Z}$ è numerabile: detti $q_1, q_2, q_3, \dots$ i razionali positivi, si pone:

$$\begin{array}{cccccccccc}
\mathbb{N} & 1 & 2 & 3 & 4 & 5 & 6 & 7 & \dots \\
 & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \updownarrow & \\
\mathbb{Q} & 0 & q_1 & -q_1 & q_2 & -q_2 & q_3 & -q_3 & \dots
\end{array}$$

che realizza una corrispondenza biunivoca tra $\mathbb{N}$ e $\mathbb{Q}$. ♦

Aver scoperto che diversi insiemi infiniti, uno propriamente contenuto nell'altro $(\mathbb{N}, \mathbb{Z}, \mathbb{Q})$, hanno la stessa cardinalità, potrebbe far pensare che questo sia vero per *tutti* gli insiemi infiniti. Ciò non è vero. Infatti, si dimostra che:

TEOREMA 1.6 $\mathbb{R}$ *non è numerabile.*

DIMOSTRAZIONE. La tecnica di questa dimostrazione prende il nome di *procedimento diagonale di Cantor* ed è un'idea semplice ma profonda che si utilizza in varie parti della matematica.

Proveremo, per l'esattezza, che l'intervallo $[0, 1]$ ha una cardinalità non numerabile, da cui seguirà, a maggior ragione, la non numerabilità di $\mathbb{R}$. Supponiamo dunque per assurdo che $[0, 1]$ sia numerabile e disponiamo *tutti* i numeri reali dell'intervallo $[0, 1]$ in un elenco $r_1, r_2, r_3 \dots$ Scriviamo ora ogni numero r_i in forma decimale (cioè nella forma $0.a_1a_2a_3\dots$ dove gli a_i sono cifre da 0 a 9; in particolare, se le cifre sono tutte zero si ha $r_i = 0$ e se sono tutte nove si ha $r_i = 0.\overline{9} = 1$):

$$\begin{aligned}
r_1 &= 0.a_{11}a_{12}a_{13}\dots \\
r_2 &= 0.a_{21}a_{22}a_{23}\dots \\
r_3 &= 0.a_{31}a_{32}a_{33}\dots \\
&\dots\dots\dots\dots\dots
\end{aligned}$$

Definiamo ora il seguente numero decimale:

$$r = 0.b_1 b_2 b_3 \ldots$$

dove le cifre b_i sono definite con la seguente regola:

$$\text{se } a_{ii} \text{ è una cifra } 0, 1, 2, 3 \text{ o } 4, \text{ poniamo } b_i = 5;$$
$$\text{se } a_{ii} \text{ è una cifra } 5, 6, 7, 8 \text{ o } 9, \text{ poniamo } b_i = 4.$$

Con questa definizione risulta $b_i \neq a_{ii}$ per ogni i. Si noti che il numero r è stato costruito ragionando sulla *diagonale* della tabella infinita che ha per righe i numeri r_i, da cui il nome del procedimento.

Osserviamo ora che r è un numero reale appartenente all'intervallo $[0, 1]$ (perché è del tipo $0.b_1 b_2 b_3 \ldots$ con tutte le cifre uguali a 4 o 5) e d'altro canto non è uguale a nessuno dei numeri r_i dell'elenco, in quanto:

$r \neq r_1$ perché la prima cifra di r è diversa dalla prima cifra di r_1 (perché $b_1 \neq a_{11}$);
$r \neq r_2$ perché la seconda cifra di r è diversa dalla seconda cifra di r_2 (perché $b_2 \neq a_{22}$);
$r \neq r_3$ perché...
...

Ma questo porta a una contraddizione, avevamo supposto che gli r_i esaurissero completamente l'insieme dei numeri reali dell'intervallo $[0, 1]$. Dunque $[0, 1]$ non è numerabile. ♦

Poiché $\mathbb{N}$ non può essere messo in corrispondenza biunivoca con $\mathbb{R}$, ma può esser messo in corrispondenza biunivoca con un sottoinsieme proprio di $\mathbb{R}$ ($\mathbb{N}$ stesso!), diciamo che $\mathbb{N}$ ha cardinalità *minore* di $\mathbb{R}$. Ecco quindi com'è possibile dar senso al confronto tra la "numerosità" di due insiemi infiniti.

Notiamo anche che l'intervallo $[0, 1]$, anzi qualsiasi intervallo di $\mathbb{R}$ (aperto o chiuso, limitato o illimitato) ha la stessa cardinalità di $\mathbb{R}$. In altre parole, la retta *non ha* una cardinalità maggiore del segmento, ma la stessa. Questo fatto si può dimostrare ad esempio con una costruzione geometrica elementare, che realizza una corrispondenza biunivoca tra punti della retta e di un segmento come quella in figura 1.7.

Figura 1.7. I punti del segmento (aperto) AB sono messi in corrispondenza biunivoca con i punti della retta, tracciando opportuni segmenti dai due punti fissati P_1, P_2 alla retta stessa.

La cardinalità di $\mathbb{R}$ prende il nome di *potenza del continuo*. Non solo ogni intervallo di $\mathbb{R}$ ha questa stessa cardinalità. Si può dimostrare che lo stesso vale per il piano, per lo spazio tridimensionale e per i loro sottoinsiemi "continui": per esempio, un piano e una sfera hanno entrambi la potenza del continuo.

Abbiamo incontrato fin qui solo *due* livelli gerarchici di infinito: la potenza del numerabile e quella del continuo. Non si deve credere che esistano solo queste due: ad esempio, l'insieme di tutti i sottoinsiemi di $\mathbb{R}$ è ancora più numeroso di $\mathbb{R}$; anzi con questo procedimento si può sempre costruire un insieme più numeroso di un insieme dato:

Teorema 1.7 *Sia A un insieme qualsiasi (finito o infinito!) e $\mathcal{P}(A)$ il suo insieme delle parti. Allora la cardinalità di $\mathcal{P}(A)$ è sempre maggiore della cardinalità di A.*

Una dimostrazione di questo teorema è suggerita nell'Esercizio 7 dei Complementi in fondo a questo capitolo.

Perciò i livelli gerarchici delle cardinalità infinite sono anch'essi infiniti.

7 IL PRINCIPIO DI INDUZIONE

Presentiamo ora un potente metodo dimostrativo, detto *dimostrazione per induzione*. Questo procedimento si può applicare a teoremi che abbiano la struttura seguente:

"Per ogni $n \in \mathbb{N}$, $n \geq n_0$, vale la proprietà $p(n)$".

Il numero n_0 è il più piccolo intero per cui si vuole che la proprietà sia vera; se $n_0 = 0$ il teorema afferma semplicemente che la proprietà è vera per ogni $n \in \mathbb{N}$.

La dimostrazione per induzione consiste nei due passi seguenti:

1. Si dimostra che $p(n)$ è vero per $n = n_0$ (primo passo dell'induzione).
2. Si dimostra che, se n è un generico numero naturale $\geq n_0$, dal fatto che $p(n)$ sia vero segue che $p(n+1)$ è vero.

Si può allora concludere che per ogni $n \geq n_0$, $p(n)$ è vero. La validità di questo metodo dimostrativo, intuitivamente, si basa su questo fatto:

1. Per il punto 1, sappiamo che $p(n_0)$ è vera. Supponiamo ad esempio $n_0 = 1$: sappiamo quindi che $p(1)$ è vera (questo va dimostrato esplicitamente).
2. Per il punto 2, poiché è vera $p(1)$, sarà vera $p(2)$: infatti abbiamo dimostrato che qualunque sia n, se è vera $p(n)$ è vera anche $p(n+1)$. Ma allora, poiché è vera $p(2)$, sarà vera $p(3)$; ma allora è vera $p(4), \ldots$ e così via, dunque è vera $p(n)$ per ogni $n \geq 1$.

Concretamente, la dimostrazione si articola in due fasi.

1. Dimostrare direttamente $p(n_0)$;
2. Assumere come ipotesi $p(n)$ (ipotesi induttiva) e provare $p(n+1)$.

È questo il punto delicato, spesso soggetto a equivoci. "Assumere per ipotesi" $p(n)$ non significa assumere come ipotesi esattamente la nostra tesi? Non proprio. Quello che si deve provare è che:

per ogni $n \geq n_0$, se è vera $p(n)$ allora è vera anche $p(n+1)$

e *non*

se per ogni n è vera $p(n)$, allora è vera anche $p(n+1)$.

Lo studente rifletta sulla differenza tra le due frasi appena scritte. La sottigliezza della dimostrazione per induzione è tutta lì. Diamo un paio di esempi significativi di applicazione di questo metodo:

Teorema 1.8 (Disuguaglianza di Bernoulli) *Per ogni intero* $n \geq 0$, $x \in \mathbb{R}$, $x \geq -1$

$$(1+x)^n \geq 1 + nx.$$

Dimostrazione. Per induzione su n. Sia $n = 0$. Allora l'asserto diventa

$$(1+x)^0 \geq 1 + 0 \cdot x,$$

cioè $1 \geq 1$, evidentemente vero. Supponiamo che sia vero per n, e proviamolo per $(n+1)$.

$$(1+x)^{n+1} = (1+x) \cdot (1+x)^n \geq$$

(per ipotesi induttiva $(1+x)^n \geq 1 + nx$; inoltre per ipotesi $(1+x) \geq 0$ perché $x \geq -1$)

$$\geq (1+x)(1+nx) = 1 + (n+1)x + nx^2 \geq 1 + (n+1)x$$

dove nell'ultima disuguaglianza si è sfruttato il fatto che $nx^2 \geq 0$. La catena di disuguaglianze mostra che $(1+x)^{n+1} \geq 1 + (n+1)x$, pertanto la tesi è dimostrata.

Teorema 1.9 (Formula del binomio di Newton) *Per ogni intero* $n \geq 0$, *con* $a, b \in \mathbb{R}$,

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

Dimostrazione. (Questo risultato è stato già enunciato nel paragrafo 2.3; ora ne forniamo una dimostrazione). Per induzione su n. Per $n = 0$ l'asserto è

$$(a+b)^0 = \binom{0}{0} a^0 b^0, \text{ cioè } 1 = 1,$$

vero. Sia vero per n, e dimostriamolo per $(n+1)$.

$$(a+b)^{n+1} = (a+b)(a+b)^n =$$

per l'ipotesi induttiva

$$= (a+b) \sum_{k=0}^{n} \binom{n}{k} a^k b^{n-k} = \sum_{k=0}^{n} \binom{n}{k} a^{k+1} b^{n-k} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} =$$

eseguendo nella prima sommatoria una traslazione di indici

$$= \sum_{k=1}^{n+1} \binom{n}{k-1} a^k b^{n-k+1} + \sum_{k=0}^{n} \binom{n}{k} a^k b^{n+1-k} =$$

scorporando l'addendo per $k = 0$ dalla seconda sommatoria e quello per $k = n+1$ dalla prima

$$= a^{n+1} + \sum_{k=1}^{n} \binom{n}{k-1} a^k b^{n-k+1} + b^{n+1} + \sum_{k=1}^{n} \binom{n}{k} a^k b^{n+1-k} =$$

sommando le due sommatorie

$$= a^{n+1} + b^{n+1} + \sum_{k=1}^{n} \left[\binom{n}{k-1} + \binom{n}{k} \right] a^k b^{n-k+1} =$$

applicando la formula di ricorrenza $\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k}$

$$= a^{n+1} + b^{n+1} + \sum_{k=1}^{n} \binom{n+1}{k} a^k b^{n-k+1} =$$

$$= \sum_{k=0}^{n+1} \binom{n+1}{k} a^k b^{n-k+1},$$

che è esattamente l'asserto voluto, per $n+1$. ♦

Chiediamoci ora che cosa giustifichi la validità del procedimento di dimostrazione per induzione dal punto di vista formale, e non puramente intuitivo. Per rispondere, si consideri anzitutto la seguente riformulazione del principio di induzione nel linguaggio degli insiemi:

"Sia S un sottoinsieme di $\mathbb{N}$. Se $0 \in S$ e, per ogni $n \in \mathbb{N}$, se $n \in S$ anche $n+1 \in S$, allora $S = \mathbb{N}$".

Formulato in questo modo, il principio di induzione viene ad essere uno degli assiomi del sistema dei numeri naturali; in altre parole, questo significa che la validità di questo principio è parte della definizione dell'insieme $\mathbb{N}$. La situazione è analoga a quella che abbiamo incontrato parlando della proprietà dell'estremo superiore: la validità della proprietà dell'estremo superiore (assioma di continuità) è parte della definizione dell'insieme $\mathbb{R}$ dei numeri reali.

Esercizi

Lo studente dimostri per induzione le seguenti identità, alcune delle quali si sono già incontrate nel pararagrafo 1.2:

8 La formula per la somma dei primi n interi:

$$\sum_{k=1}^{n} k = \frac{n(n+1)}{2} \text{ per ogni intero } n \geq 1.$$

9 La formula per la somma geometrica:

$$\sum_{k=0}^{n} q^k = \frac{1-q^{n+1}}{1-q} \text{ per ogni intero } n \geq 0 \text{ e numero reale } q \neq 1.$$

10 Ricavare la formula per la somma dei primi n numeri pari:

$$\sum_{k=1}^{n} (2k) \text{ per ogni intero } n \geq 1.$$

11 Ricavare la formula perla somma dei primi n numeri dispari:

$$\sum_{k=0}^{n-1} (2k+1) \text{ per ogni intero } n \geq 1.$$

12 Dimostrare che se A è un insieme costituito da n elementi, l'insieme delle parti di A ha 2^n elementi.

13 Dimostrare l'identità dei coefficienti binomiali (utilizzata nella dimostrazione della formula del binomio di Newton):

$$\binom{n+1}{k} = \binom{n}{k-1} + \binom{n}{k} \text{ per } n, k \text{ interi, } 1 \leq k \leq n.$$

8 NUMERI COMPLESSI

Ci sono ragioni, principalmente di natura algebrica, che ci spingono ad ampliare ulteriormente il campo numerico; si tratta, in sostanza, di definire in tutta la sua generalità l'operazione di elevamento a potenza a^b (e la sua "inversa"), operazione che, nel campo reale, è definita per ogni esponente b solo se a è positivo, e, se a è negativo, soltanto per alcuni esponenti (interi o razionali $\frac{m}{n}$ con n dispari).

In questo paragrafo introduciamo quindi il campo dei numeri complessi e definiamo su di essi le operazioni algebriche elementari e la radice n-esima. Definiremo in seguito (cap. 5, par. 2.2) l'operazione di elevamento a potenza qualunque, così come i logaritmi e gli esponenziali nel campo complesso. Molte applicazioni che lo studente potrà incontrare nel corso dei suoi studi richiedono l'uso di questi strumenti.

8.1 Definizione di $\mathbb{C}$ e struttura di campo

Abbiamo indicato con $\mathbb{R}^2$ (abbreviazione di $\mathbb{R} \times \mathbb{R}$) l'insieme delle coppie ordinate (a, b) di numeri reali. Su queste definiamo direttamente le operazioni di somma e prodotto con le seguenti regole:

$$(a, b) + (c, d) := (a + c, b + d) \tag{8.1}$$

$$(a, b) \cdot (c, d) := (ac - bd, ad + bc) \tag{8.2}$$

Questa "somma" e questo "prodotto" verificano le proprietà commutativa, associativa e distributiva enunciate nel paragrafo 3. Osserviamo anche che, $\forall\,(a, b)$:

$$(a, b) + (0, 0) = (0, 0) + (a, b) = (a, b)$$

dunque la coppia $(0, 0)$ è *elemento neutro* per la somma.

$$(a, b) \cdot (1, 0) = (1, 0) \cdot (a, b) = (a, b)$$

dunque la coppia $(1, 0)$ è *elemento neutro* per il prodotto.

$$(a, b) + (-a, -b) = (0, 0)$$

dunque $(-a, -b)$ è l'*opposto* di (a, b).

Se

$$(a,b) \neq (0,0) \text{ allora } (a,b) \cdot \left(\frac{a}{a^2+b^2}, \frac{-b}{a^2+b^2}\right) = (1,0)$$

dunque la coppia $(a/(a^2+b^2), -b/(a^2+b^2))$ è il *reciproco* di (a,b).

Dunque le proprietà R_1, R_2 (vedi par. 3) sono verificate per la somma e il prodotto così definiti e perciò l'insieme $\mathbb{R}^2$ così strutturato è un campo, che chiameremo *campo dei numeri complessi* e indicheremo con $\mathbb{C}$.

Osserviamo ora che $\mathbb{C}$ contiene il sottoinsieme $\mathbb{C}_0$ delle coppie del tipo $(a,0)$; esso è un *sottocampo* di $\mathbb{C}$, poiché somma e prodotto di coppie di questo tipo sono ancora coppie dello stesso tipo; infatti si ha:

$$(a,0) + (b,0) = (a+b,0)$$

$$(a,0) \cdot (b,0) = (ab,0)$$

Inoltre $\mathbb{C}_0$ può essere ordinato ponendo $(a,0) < (b,0)$ se $a < b$. Se allora mettiamo in corrispondenza biunivoca l'insieme dei numeri reali $\mathbb{R}$ con $\mathbb{C}_0$, ponendo

$$(a,0) \longleftrightarrow a$$

possiamo *identificare* i numeri reali a con i numeri complessi del tipo $(a,0)$.

In questo senso il campo dei numeri complessi $\mathbb{C}$ è un ampliamento di quello dei numeri reali $\mathbb{R}$.

Consideriamo ora il numero $(0,1)$. Esso ha la singolare proprietà che:

$$(0,1) \cdot (0,1) = (-1,0)$$

cioè *il suo quadrato coincide col numero reale* -1! Per questa ragione la coppia $(0,1)$ merita di essere indicata con un simbolo speciale: la indicheremo con "i" e la chiameremo *unità immaginaria*.

Forma algebrica dei numeri complessi

A questo punto conviene semplificare la notazione. Osserviamo che, se scriviamo semplicemente a invece di $(a,0)$ ecc. abbiamo

$$(a,b) = (a,0) + (0,1) \cdot (b,0) = a + ib$$

Con questa notazione le regole (8.1) e (8.2) sono le ordinarie regole del calcolo letterale, ove si tenga conto che $i^2 = -1$:

$$(a+ib) + (c+id) = (a+c) + i(b+d)$$

$$(a+ib) \cdot (c+id) = (ac-bd) + i(ad+bc)$$

La scrittura

$$z = a + ib \tag{8.3}$$

è detta *forma algebrica* dei numeri complessi; a si chiama *parte reale* di z e si indica con $\mathrm{Re}(z)$, b si chiama *parte immaginaria* e si indica con $\mathrm{Im}(z)$.

Piano complesso

In un piano cartesiano, rappresentiamo i numeri complessi $a + ib$ come punti di coordinate (a, b): ecco una semplice e comoda immagine geometrica del campo complesso. In questo contesto, il piano viene detto *piano complesso* o *piano di Gauss*; gli assi x, y si dicono *asse reale* e *asse immaginario*: i punti sull'asse reale sono i numeri reali, i punti sull'asse immaginario sono i numeri immaginari puri (cioè del tipo ib).

La somma di due numeri complessi è il numero complesso che ha per coordinate la somma delle coordinate: il significato geometrico di questo fatto è che il punto $z+t$ si costruisce a partire dai punti z, t in base alla "regola del parallelogramma" (vedi fig. 1.8).

Figura 1.8.

Abbiamo visto che $\mathbb{C}$ soddisfa gli assiomi di campo; *non* soddisfa però quelli di *campo ordinato*, ovvero non è possibile definire una relazione $\leq$ tra i numeri complessi, in modo che valgano le proprietà elencate nel paragrafo 3 (in particolare, la R_3). Infatti si può dimostrare che da quelle proprietà segue che *il quadrato di un numero qualsiasi non è mai negativo*, e d'altra parte, *se un numero è positivo il suo opposto è negativo*. Ora, in $\mathbb{C}$ si ha:

$$1^2 = 1 \quad \text{e} \quad i^2 = -1$$

Abbiamo quindi due quadrati che sono l'uno l'opposto dell'altro. Nessuno dei due però può essere negativo (perché sono quadrati), e questo è assurdo (perché tra a e $-a$ uno dev'essere negativo, se $a \neq 0$). Concludiamo che $\mathbb{C}$ *non è un campo ordinato*.

8.2 Coniugato e modulo

Il numero complesso $a - ib$ si dice il *complesso coniugato* di $z = a + ib$ e si indica con $\bar{z}$. Evidentemente si ha:

$$z + \bar{z} = 2a = 2\,\mathrm{Re}(z)$$

$$z - \bar{z} = 2bi = 2i\,\mathrm{Im}(z)$$

L'operazione di *coniugio* ha le seguenti elementari proprietà rispetto alla somma e al prodotto:

$$\overline{(z_1 + z_2)} = \bar{z}_1 + \bar{z}_2$$

$$\overline{(z_1 z_2)} = \bar{z}_1 \cdot \bar{z}_2$$

$$\overline{\left(\frac{1}{z}\right)} = \frac{1}{\bar{z}}$$

Osserviamo ora che

$$z\bar{z} = (a + ib)(a - ib) = a^2 + b^2 \geq 0$$

Si chiama *modulo* di $z = a + ib$ il numero reale non negativo $\sqrt{a^2 + b^2}$, che si indica con $|z|$. Se $z = a$ è reale, il suo modulo si chiama *valore assoluto* e si indica sempre con $|a|$, come detto nel paragrafo 4. Valgono le seguenti proprietà:

a) $|z| \geq 0$ e $|z| = 0 \iff z = 0$

b) $|z| = |\bar{z}|$

c) $|\operatorname{Re}(z)| \leq |z| \quad |\operatorname{Im}(z)| \leq |z| \quad |z| \leq |\operatorname{Re}(z)| + |\operatorname{Im}(z)|$

(8.4) d) $|z_1 + z_2| \leq |z_1| + |z_2|$ (disuguaglianza triangolare)

(8.5) e) $|z_1 + z_2| \geq \big|\, |z_1| - |z_2| \,\big|$

Le proprietà a), b), c) si verificano immediatamente. Proviamo d), e).

Esse sono equivalenti alla seguente:

$$\big(|z_1| - |z_2|\big)^2 \leq |z_1 + z_2|^2 \leq \big(|z_1| + |z_2|\big)^2$$

Ponendo $z_1 = a + ib$, $z_2 = c + id$ otteniamo:

$$\left(\sqrt{a^2 + b^2} - \sqrt{c^2 + d^2}\right)^2 \leq (a + c)^2 + (b + d)^2 \leq \left(\sqrt{a^2 + b^2} + \sqrt{c^2 + d^2}\right)^2$$

Con calcoli elementari questa doppia disuguaglianza si riduce a

$$-\sqrt{a^2 + b^2} \cdot \sqrt{c^2 + d^2} \leq ac + bd \leq \sqrt{a^2 + b^2} \cdot \sqrt{c^2 + d^2}$$

che è equivalente alla seguente:

$$|ac + bd| \leq \sqrt{a^2 + b^2} \cdot \sqrt{c^2 + d^2}$$

Elevando al quadrato entrambi i membri si arriva a:

$$(ac + bd)^2 \leq (a^2 + b^2)(c^2 + d^2)$$

ovvero a

$$0 \leq -2acbd + a^2d^2 + b^2c^2 = (ad - bc)^2$$

che è vera per ogni a, b, c, $d \in \mathbb{R}$.

Geometricamente, $|z|$ rappresenta la *distanza* del punto (o numero complesso) z dall'origine; $|z_1 - z_2|$ rappresenta la distanza dei due punti z_1 e z_2; le disuguaglianze (8.4) e (8.5) traducono il noto teorema sulle lunghezze dei lati di un triangolo (vedi fig. 1.9).

Figura 1.9.

Utilizzando i concetti ora introdotti, possiamo rappresentare in forma algebrica il rapporto di due numeri complessi $\dfrac{a+ib}{c+id}$; basta moltiplicare numeratore e denominatore per $c - id$; abbiamo:

$$\frac{a+ib}{c+id} = \frac{(a+ib)(c-id)}{|c+id|^2} = \frac{ac+bd}{c^2+d^2} + i\frac{bc-ad}{c^2+d^2}$$

Vediamo come si può risolvere un'equazione nel campo complesso, quando questa coinvolge l'incognita $z = x + iy$ anche attraverso $\operatorname{Re} z$, $\operatorname{Im} z$, $\overline{z}$, $|z|$.

Esempio

8.1

$$z^2 + i\operatorname{Im} z + 2\overline{z} = 0$$

Poniamo $z = x + iy$, con x, y incognite reali, e trascriviamo l'equazione a questo modo

$$z^2 = (x+iy)^2 = x^2 - y^2 + 2ixy$$

$$i \operatorname{Im} z = iy$$

$$2\overline{z} = 2(x - iy) = 2x - 2iy$$

$$\left(x^2 - y^2 + 2ixy\right) + (iy) + (2x - 2iy) = 0$$

Ora, un numero complesso è zero se e solo se la sua parte reale e parte immaginaria sono zero. Perciò mettiamo in evidenza la parte reale e la parte immaginaria del primo membro

ed uguagliamo entrambe a zero:

$$\left(x^2 - y^2 + 2x\right) + i\left(2xy + y - 2y\right) = 0$$

$$\begin{cases} x^2 - y^2 + 2x = 0 \\ 2xy - y = 0 \end{cases}$$

Si è così trasformata l'equazione in una incognita complessa in un sistema di due equazioni in due incognite reali. Risolviamo il sistema. La seconda equazione dà:

$$y = 0 \quad \text{o} \quad x = \frac{1}{2}$$

Per $y = 0$ la prima equazione diventa

$$x^2 + 2x = 0$$

che dà

$$x = 0 \quad \text{o} \quad x = -2$$

Per $x = \frac{1}{2}$ la prima equazione diventa

$$-y^2 + \frac{5}{4} = 0$$

che ha soluzioni

$$y = \pm\frac{\sqrt{5}}{2}$$

Quindi le soluzioni sono:

$$z = 0 \qquad z = -2 \qquad z = \frac{1}{2} + i\frac{\sqrt{5}}{2} \qquad z = \frac{1}{2} - i\frac{\sqrt{5}}{2}$$

L'equazione ha 4 soluzioni.

Il metodo visto in quest'esempio (passare alla parte reale e immaginaria dell'equazione) *è applicabile in linea di principio ad ogni equazione in* $\mathbb{C}$. In pratica, un generico sistema di due equazioni in due incognite è quasi sempre insolubile per via algebrica. Perciò prima di mettersi su questa strada è bene osservare se non ce ne sia una più semplice.

8.3 Forma trigonometrica

Come è noto dalla Geometria, i punti del piano possono essere individuati, oltre che dalle loro coordinate cartesiane, anche dalle coordinate polari: ρ (*raggio polare*, cioè distanza del punto dall'origine) e θ (*angolo polare*, cioè l'angolo che la retta congiungente il punto con l'origine forma con l'asse delle ascisse positive, misurato in senso antiorario). È chiaro che una coppia ρ, θ, con $\rho > 0$, individua un ben determinato punto del piano; invece un punto del piano individua univocamente la coordinata ρ, ma l'angolo θ, misurato in radianti, è determinato solo a meno di multipli di 2π (fig. 1.10).

Dato un numero complesso z, il suo modulo $|z|$ coincide col raggio polare del punto che ne è l'immagine sul piano complesso. Chiamiamo *argomento* di z, e lo indicheremo con $\arg(z)$, uno qualsiasi degli angoli θ relativi al punto z. In questo modo l'argomento

Figura 1.10.

di z non è univocamente determinato. Spesso questa indeterminatezza non porta alcun inconveniente. Altre volte invece è preferibile assegnare un ben determinato argomento a un numero complesso. Ciò può ottenersi in infiniti modi, fissando un qualsiasi intervallo, di ampiezza 2π, entro il quale far variare l'angolo θ. Gli intervalli più comunemente usati a questo scopo sono $[0, 2\pi)$ e $(-\pi, \pi]$; allora l'argomento di z viene detto *argomento principale*.

Per esempio, il numero $-i$ ha come argomento $-\pi/2$ oppure $3\pi/2$ oppure qualunque altro valore della forma $-\pi/2 + 2k\pi$ con $k \in \mathbb{Z}$. Il suo argomento principale sarà $3\pi/2$ se si adotta la convenzione che $\theta \in [0, 2\pi)$, $-\pi/2$ con la convenzione che $\theta \in (-\pi, \pi]$. I numeri reali positivi hanno argomento principale 0 e quelli negativi π con entrambe le convenzioni. Per il numero 0 l'argomento non viene definito.

Dato il numero $z = a + ib$, dalla trigonometria ricaviamo immediatamente le relazioni tra le coordinate cartesiane a, b e quelle polari ρ, θ:

$$a = \rho \cos\theta \qquad b = \rho \sin\theta \tag{8.6}$$

cosicché il numero complesso z può anche scriversi nella forma

$$\boxed{z = \rho(\cos\theta + i \sin\theta)} \tag{8.7}$$

che è detta *forma trigonometrica* dei numeri complessi.

Le relazioni inverse di (8.6) sono:

$$\boxed{\rho = \sqrt{a^2 + b^2} \quad \cos\theta = \frac{a}{\sqrt{a^2 + b^2}} \quad \sin\theta = \frac{b}{\sqrt{a^2 + b^2}}} \tag{8.8}$$

Esempio

8.2 Scriviamo in forma trigonometrica il numero complesso:

$$\sqrt{3} + i$$

Abbiamo $\rho = \sqrt{a^2 + b^2} = \sqrt{3 + 1} = 2$;

$$\sqrt{3} + i = 2\left(\frac{\sqrt{3}}{2} + i\frac{1}{2}\right) = 2\left(\cos\frac{\pi}{6} + i \sin\frac{\pi}{6}\right)$$

Formule di De Moivre

La forma trigonometrica è comoda per esprimere prodotti e quozienti di numeri complessi. Se abbiamo infatti

$$z_1 = \rho_1(\cos\theta_1 + i\sin\theta_1) \qquad z_2 = \rho_2(\cos\theta_2 + i\sin\theta_2)$$

otteniamo

$$\begin{aligned} z_1 z_2 &= \rho_1\rho_2\{\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2 + i(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2)\} = \\ &= \rho_1\rho_2\{\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2)\} \end{aligned} \tag{8.9}$$

Se $z_2 \neq 0$, abbiamo

$$\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2}\cdot\frac{\cos\theta_1 + i\sin\theta_1}{\cos\theta_2 + i\sin\theta_2}$$

Moltiplicando numeratore e denominatore per $\cos\theta_2 - i\sin\theta_2$ e tenendo conto che $(\cos\theta_2)^2 + (\sin\theta_2)^2 = 1$ abbiamo:

$$\frac{z_1}{z_2} = \frac{\rho_1}{\rho_2}(\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 - i\sin\theta_2) = \frac{\rho_1}{\rho_2}\{\cos(\theta_1-\theta_2) + i\sin(\theta_1-\theta_2)\} \tag{8.10}$$

Pertanto il modulo del prodotto e del quoziente di due numeri complessi è, rispettivamente, il prodotto e il quoziente dei moduli; l'argomento è, rispettivamente, la somma e la differenza degli argomenti:

$$\boxed{\begin{array}{ll} |z_1\cdot z_2| = |z_1|\cdot|z_2| & \arg(z_1\cdot z_2) = \arg z_1 + \arg z_2 \\ |z_1/z_2| = |z_1|/|z_2| & \arg(z_1/z_2) = \arg z_1 - \arg z_2 \end{array}} \tag{8.11}$$

La (8.9) si generalizza al caso di un numero qualsiasi di fattori $z_1, z_2, \ldots, z_n$.

$$z_1 z_2 \ldots z_n = \rho_1\rho_2\ldots\rho_n\{\cos(\theta_1+\theta_2+\ldots+\theta_n) + i\sin(\theta_1+\theta_2+\ldots+\theta_n)\} \tag{8.12}$$

Se poi i fattori sono tutti uguali, otteniamo:

$$\boxed{z^n = \rho^n\big(\cos(n\theta) + i\sin(n\theta)\big)} \tag{8.13}$$

Le relazioni (8.9)-(8.13) vanno sotto il nome di *formule di De Moivre*. Vediamone alcune applicazioni.

Esempi

8.3 Scrivere in forma algebrica $(1+i)^7$.

Determiniamo modulo e argomento di $(1+i)$ e poi applichiamo la formula di De Moivre.

$$|1+i| = \sqrt{2} \quad \arg(1+i) = \frac{\pi}{4}$$

Quindi

$$|(1+i)^7| = \left(\sqrt{2}\right)^7 = 8\sqrt{2} \quad \arg(1+i)^7 = \frac{7}{4}\pi$$

Di conseguenza:

$$(1+i)^7 = 8\sqrt{2}\left(\frac{1}{\sqrt{2}} - i\frac{1}{\sqrt{2}}\right) = 8 - 8i$$

8.4 Risolvere l'equazione

$$z^3 - |z| = 0$$

Sostituire $z = x + iy$ e separare parte reale e parte immaginaria è possibile, ma porta a calcoli algebrici un po' pesanti. Se riscriviamo l'equazione nella forma

$$z^3 = |z|$$

possiamo invece notare che ambo i membri si esprimono facilmente se si pone $z = \rho(\cos\theta + i\sin\theta)$, ovvero se si usa la *forma trigonometrica*. Infatti:

$$z^3 = \rho^3\,(\cos 3\theta + i\sin 3\theta) \qquad |z| = \rho$$

e l'equazione è soddisfatta se e solo se i due membri hanno moduli uguali e argomenti che differiscono per multipli di 2π, ovvero (il secondo membro ha argomento 0):

$$\begin{cases} \rho^3 = \rho \\ 3\theta = 2k\pi \quad \text{con} k \in \mathbb{Z} \end{cases}$$

La prima equazione dà $\rho = 0$ e $\rho = 1$ (attenzione: ρ dev'essere ≥ 0 perché è il modulo del numero complesso; perciò $\rho = -1$ non è accettabile); la seconda dà: $\theta = \frac{2k\pi}{3}$. Si trova pertanto:

$$z = 0 \qquad z = \cos\frac{2k\pi}{3} + i\sin\frac{2k\pi}{3} \quad \text{per } k = 0, 1, 2$$

Esplicitamente:

$$z = 0 \qquad z = 1 \qquad z = -\frac{1}{2} + i\frac{\sqrt{3}}{2} \qquad z = -\frac{1}{2} - i\frac{\sqrt{3}}{2}$$

Le formule di De Moivre permettono di dare un'*interpretazione geometrica al prodotto di numeri complessi*. Sia z, per cominciare, un numero complesso di modulo 1, quindi del tipo $(\cos\theta + i\sin\theta)$. Allora, moltiplicare un numero per z significa sommare θ al suo argomento, cioè eseguire una *rotazione di angolo* θ. Se z ha modulo ρ anziché 1, oltre ad eseguire una rotazione si esegue una *dilatazione di coefficiente* ρ. Ad esempio:

- moltiplicare per i significa eseguire una rotazione di $\frac{\pi}{2}$;
- moltiplicare per -1 significa eseguire una rotazione di π;
- moltiplicare per $(1 + i)$ significa eseguire una dilatazione di coefficiente $\sqrt{2}$ e una rotazione di $\pi/4$.

8.4 Radici n-esime

Dato un numero complesso w, diremo che z è una *radice n-esima (complessa)* di w se risulta $z^n = w$.

Teorema 1.10 *Sia $w \in \mathbb{C}$, $w \neq 0$, e n intero ≥ 1. Esistono precisamente n radici n-esime complesse $z_0, z_1, \ldots, z_{n-1}$ di w; posto $w = r(\cos\varphi + i\sin\varphi)$ e $z_k = \rho_k(\cos\theta_k + i\sin\theta_k)$ abbiamo*

$$\boxed{\begin{aligned} &\rho_k = r^{1/n} \\ &\theta_k = \frac{\varphi + 2k\pi}{n} \qquad k = 0, 1, \ldots, n-1 \end{aligned}} \tag{8.14}$$

DIMOSTRAZIONE. I numeri z_k sono evidentemente radici di w, come risulta calcolando z_k^m mediante la formula di De Moivre. Mostriamo che non ve ne sono altre. Infatti, perché un numero $R(\cos\psi + i\sin\psi)$ sia radice n-esima di w, dovrebbe risultare

$$R^n = r \quad \text{e} \quad n\psi = \varphi + 2h\pi \quad \text{con} \quad h \in \mathbb{Z}$$

e cioè

$$R = r^{1/n} \quad \text{e} \quad \psi = \varphi/n + 2h\pi/n$$

Dando a h i valori $0, 1, \ldots, n-1$ troviamo appunto i numeri z_k. Dando a h un qualsiasi altro valore $\bar{h}$ diverso dai precedenti, questo può scriversi nella forma $\bar{h} = k + mn$ ($m \in \mathbb{Z}$ è il quoziente e k è il resto della divisione di $\bar{h}$ per n) per cui sarebbe

$$\psi = \frac{\varphi}{n} + \frac{2k\pi}{n} + 2m\pi = \theta_k + 2m\pi$$

e ritroveremmo ancora gli stessi z_k precedenti. ♦

Per le radici complesse si usa purtroppo una notazione un po' ambigua, la stessa in uso per indicare la radice aritmetica; si indica cioè con $\sqrt[n]{z}$ o $z^{1/n}$ l'insieme delle n radici complesse di z. Ciò può creare confusione quando z è reale. Infatti il simbolo $\sqrt{4}$, inteso come radice aritmetica di 4, è 2; inteso come radice complessa di 4 è l'insieme dei due numeri $+2$ e -2.

Esempio

8.5 Le tre radici cubiche di -1 sono i numeri z_k della forma: $\cos\theta_k + i\sin\theta_k$ con $\theta_k = \pi/3 + 2k\pi/3$, $k = 0, 1, 2$. Esplicitamente abbiamo:

$$z_0 = \frac{1}{2}(1 + \sqrt{3}i)$$

$$z_1 = -1$$

$$z_2 = \frac{1}{2}(1 - \sqrt{3}i)$$

Figura 1.11.

La disposizione delle radici cubiche dell'esempio precedente nel piano di Gauss non è casuale. Infatti se $w = r(\cos\varphi + i\sin\varphi)$ le radici n-esime $z_0, z_1, \ldots, z_{n-1}$ di w si trovano ai vertici del poligono regolare di n lati inscritto nella circonferenza di centro 0 e raggio $r^{1/n}$, con il vertice z_0 posto nel punto di argomento $\theta = \varphi/n$.
Nella figura 1.12 sono rappresentate le radici seste di i : $z_0, z_1, \ldots, z_5$.

Esempio

8.6 Calcolare $\sqrt[5]{1+i}$. Il numero $(1+i)$ ha modulo $\sqrt{2}$ e argomento $\pi/4$. Le radici quinte saranno quindi:

$$\sqrt[5]{\sqrt{2}}\left[\cos\left(\frac{\frac{\pi}{4} + 2k\pi}{5}\right) + i\sin\left(\frac{\frac{\pi}{4} + 2k\pi}{5}\right)\right] =$$

$$= \sqrt[10]{2}\left[\cos\left(\frac{\pi}{20} + k\frac{2}{5}\pi\right) + i\sin\left(\frac{\pi}{20} + k\frac{2}{5}\pi\right)\right] \quad \text{con} \quad k = 0, 1, 2, 3, 4$$

Gli angoli trovati non sono notevoli, ma volendo si possono calcolare seni e coseni in modo approssimato, usando una calcolatrice.

Figura 1.12.

Equazioni di secondo grado

Un'equazione di secondo grado

$$az^2 + bz + c = 0$$

con coefficienti $a, b, c \in \mathbb{C}$, si risolve con la "solita" formula

$$z = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}$$

a patto di intendere la radice quadrata in senso complesso. Il segno $\pm$ davanti al simbolo di radice è in realtà superfluo, perché il simbolo stesso $\sqrt{\cdot}$ nel campo complesso denota *due numeri*, uno opposto dell'altro.

Esempio

8.7

$$z^2 + 2iz - \sqrt{3}i = 0$$

$$z = -i \pm \sqrt{-1 + \sqrt{3}i} = -i \pm \sqrt{2\left(\cos\frac{2}{3}\pi + i\sin\frac{2}{3}\pi\right)} =$$

$$= -i \pm \sqrt{2}\left(\cos\frac{\pi}{3} + i\sin\frac{\pi}{3}\right) = \pm\frac{\sqrt{2}}{2} + i\left(-1 \pm \frac{\sqrt{6}}{2}\right)$$

Il teorema 1.10 ci dice che un polinomio del tipo $z^n + a$ (con a complesso) ha in $\mathbb{C}$ esattamente n radici; nel campo reale invece l'equazione $x^n + a = 0$ può avere due, una, o nessuna radice (esempi: $x^2 - 1 = 0$, $x^3 - 1 = 0$, $x^2 + 1 = 0$); ora sappiamo che tale equazione ha sempre n radici in $\mathbb{C}$, ma solo occasionalmente una o due di esse stanno in $\mathbb{R}$.

Il risultato è di portata ben più generale, come afferma il seguente teorema, di cui non riportiamo la dimostrazione.

TEOREMA 1.11 (TEOREMA FONDAMENTALE DELL'ALGEBRA) *Un'equazione polinomiale della forma*

$$a_0 + a_1 z + \ldots + a_n z^n = 0 \qquad (a_n \neq 0)$$

con coefficienti complessi qualsiasi ha precisamente n radici in $\mathbb{C}$, se ognuna di esse viene contata con la sua molteplicità.[5]

Esercizi

14 Scrivere in forma algebrica ($z = a + ib$, $a, b \in \mathbb{R}$) i seguenti numeri complessi:

$$\frac{(2+i)(1-i)}{3-2i} \qquad \frac{1}{i(3+2i)^2} \qquad \frac{\left(\sqrt{3}+i\sqrt{2}\right)^3}{\sqrt{2}-\sqrt{3}i}$$

15 Calcolare modulo e argomento dei seguenti numeri complessi:

$$-3i; \ -5; \ -\sqrt{3}+i$$

16 Calcolare modulo e argomento dei numeri:

$$\frac{1+i\sqrt{3}}{1-i} \qquad \frac{1+i}{1-i} \qquad \frac{1+i}{\sqrt{3}+i}$$

17 Disegnare nel piano complesso il luogo dei punti z tali che:

$$\text{i)} \ \ |z| = |z+i| \qquad \text{ii)} \ \ \mathrm{Re}\,(z^2) > 2 \qquad \text{iii)} \ \ \mathrm{Im}\,\frac{1}{z} = -1$$

18 Verificare che, $\forall\, z \in \mathbb{C}$:

$$|z| \geq \mathrm{Re}\,(z) \qquad |z| \geq \mathrm{Im}\,(z) \qquad |z| \leq |\mathrm{Re}\,(z)| + |\mathrm{Im}\,(z)|$$

19 Scrivere in forma algebrica $a + ib$ i numeri

$$\frac{1-i\sqrt{3}}{1+i} \quad \left(\frac{1-i}{1+i}\right)^3 \quad (1+i)^{20} \quad (1-i)^{11} \quad (1+i\sqrt{3})^n - (1-i\sqrt{3})^n$$

20 Calcolare le radici seste di -1 e rappresentarle sul piano di Gauss.

21 Risolvere le equazioni

$$\text{a)} \ \ (z-2)^3 = -i \qquad \text{b)} \ \ z^2 - (4+i)z + 4 + 2i = 0$$

22 Esprimere $\cos 5\theta$ come combinazione lineare di potenze di $\cos\theta$ e $\sin\theta$.

23 Esprimere $(\sin\theta)^4$ come combinazione lineare di termini del tipo $\sin(k\theta)$ e $\cos(k\theta)$.

24 Calcolare le seguenti radici n-esime nel campo complesso. Quando l'argomento di $\sqrt[n]{z}$ è un valore notevole, riscrivere in forma algebrica le radici (ad esempio: $2\left(\cos\frac{\pi}{6} + i\sin\frac{\pi}{6}\right) = \sqrt{3}+i$).

$$\sqrt[3]{i-1} \qquad \sqrt{-1+\sqrt{3}i} \qquad \sqrt[4]{-2-2\sqrt{3}i}$$

25 Disegnare nel piano complesso i seguenti insiemi:

$$A = \left\{z : 1 < |z| < 2; \frac{\pi}{6} < \arg z < \frac{\pi}{3}\right\} \qquad B = \{iz : z \in A\}$$

[5] Se $P(z)$ è un polinomio in z di grado n e z_0 una sua radice, si dice che z_0 è di molteplicità k (k intero, ≥ 1) se vale la formula $P(z) = (z - z_0)^k Q(z)$, dove Q è un polinomio tale che $Q(z_0) \neq 0$.

Determinare tutte le soluzioni delle seguenti equazioni nel campo complesso:

26 $z^2 + 2z + 3 = 0$

27 $z + 3i + (\operatorname{Re} z)\left(i + (\operatorname{Im} z)^2\right) = 0$

28 $z^2 + 2iz - 3 = 0$

29 $iz^3 = \overline{z}$

30 $iz^2 + (1+i)z + 1 = 0$

31 $z^6 + 2z^3 - 3 = 0$

32 $z^2 + \overline{z} = 0$

33 $(\overline{z})^4 = |z|$

34 $i \operatorname{Re} z + z^2 = |z|^2 + 1$

35 $2|z|^2 = z^3$

COMPLEMENTI

1 Nel paragrafo 1.1 si sono definite l'unione e l'intersezione di due (e quindi di un numero finito) di insiemi, e si è mostrata l'analogia tra unione e "o", intersezione ed "e". Consideriamo ora una famiglia infinita $A_1, A_2, A_3, \ldots$ di sottoinsiemi di un insieme fissato X. Come si potrebbero definire rigorosamente la loro unione e intersezione infinita, cioè

$$\bigcup_{n=1}^{\infty} A_n, \bigcap_{n=1}^{\infty} A_n \ ?$$

Suggerimento: non si possono usare infinite "e" od "o", bisognerà invece usare opportunamente i *quantificatori*.

2 Dimostrare che $\log_2 3$ è irrazionale.

3 Dimostrare le proprietà dei coefficienti binomiali (2.3) enunciate nel paragrafo 2.3 mediante un calcolo algebrico diretto (e non per induzione).

4 Dimostrare la disuguaglianza di Bernoulli $(1+x)^n \geq 1 + nx$ nel caso particolare $x \geq 0$, usando lo sviluppo del binomio di Newton (invece che per induzione, come fatto nel paragrafo 7).

5 Dimostrare che l'intervallo $[0,1]$ e il quadrato $[0,1] \times [0,1]$ hanno la stessa cardinalità. *Suggerimento*: rappresentare in forma decimale $0.a_1a_2a_3\ldots$ gli elementi di $[0,1]$, e trovare un algoritmo esplicito che associ, biunivocamente, ad ogni elemento di questo tipo una coppia ordinata di elementi di questo tipo.

6 Dimostrare che $\mathcal{P}(\mathbb{N})$ ha la stessa cardinalità dell'intervallo $[0,1]$.
Suggerimento : rappresentare i numeri dell'intervallo $[0,1]$ in forma decimale binaria, cioè del tipo $0.a_1a_2a_3\ldots$ con a_i cifre 0 o 1; quindi, trovare una corrispondenza biunivoca tra l'insieme dei numeri così scritti e $\mathcal{P}(\mathbb{N})$.

7 Si studi e si comprenda in dettaglio la seguente dimostrazione del teorema, enunciato nel paragrafo 6: "Per ogni insieme A, A e $\mathcal{P}(A)$ non hanno la stessa cardinalità".

DIMOSTRAZIONE. Per assurdo, supponiamo che esista una corrispondenza biunivoca f tra A e $\mathcal{P}(A)$, e definiamo

$$X = \{a \in A : a \notin f(a)\}$$

dove $f(a)$ indica l'elemento di A che corrisponde ad a. Sia ora $x \in A$ l'elemento che corrisponde a X. Allora si ha: $x \in f(x) \Longleftrightarrow x \in X \Longleftrightarrow x \notin f(x)$, assurdo. ♦

2 Funzioni di una variabile

1 IL CONCETTO DI FUNZIONE

Il concetto di *funzione* è centrale in matematica, ed estremamente generale. Dal punto di vista fisico, si può dire che questo concetto nasca per descrivere matematicamente quello di *grandezza variabile*. Esempi di grandezze variabili sono pressoché ovunque attorno a noi: la posizione di un oggetto mobile, o la sua velocità, la temperatura in una stanza, la densità di un oggetto, i prezzi delle merci ecc. Si potrebbe pensare che, dopo tutto, queste grandezze sono espresse da *numeri*, per cui non c'è bisogno di introdurre alcun nuovo concetto per descriverle. Tuttavia, dire che, ad esempio, la velocità di un'automobile è espressa da un numero è un'affermazione piuttosto imprecisa, perché un numero può indicare solo la velocità *in un istante fissato*. Più in generale: i numeri, di cui finora ci siamo prevalentemente occupati, sono adatti a esprimere *grandezze costanti*; per esprimere grandezze variabili ci vuole un'idea diversa. Ora, un'analisi logica del concetto di grandezza variabile mostra che questa coinvolge sempre una relazione tra *due* grandezze: la velocità di un'automobile varia *al variare del tempo*; la densità di un oggetto può variare *da punto a punto*; in altre parole:

> l'esistenza di una grandezza variabile sottointende l'esistenza di una relazione tra *due* grandezze (tempo e velocità; posizione e densità ecc.), ovvero la *dipendenza* di una grandezza da un'altra.

Questa dipendenza segue, di volta in volta, una certa *legge*. Vediamo alcuni esempi di questo tipo, che ci permetteranno di precisare una caratteristica fondamentale del concetto di funzione:

- Se si lascia cadere un oggetto pesante da una certa altezza, lo spazio percorso dall'oggetto varia col tempo secondo la formula

  $$s(t) = \frac{1}{2}gt^2 \qquad (g \text{ accelerazione di gravità})$$

  Cioè a t viene associato $s(t)$; in simboli:

  $$t \longmapsto s(t) = \text{spazio percorso al tempo } t\,.$$

Figura 2.1.

- Se un capitale unitario viene investito per un anno al tasso mensile i, il capitale finale dipende da i secondo la formula

$$k(i) = \left(1 + \frac{i}{12}\right)^{12}$$

Cioè:

$$i \longmapsto k(i) = \text{capitale alla fine dell'anno}$$

- La potenza P di una lente dipende dalla lunghezza focale f secondo la formula

$$P = \frac{1}{f}$$

(se f è misurata in metri, P è espressa in diottrie).

In ciascuno degli esempi precedenti, a un numero reale (ingresso) viene associato *univocamente* un altro numero reale (uscita). È proprio l'univocità della corrispondenza a caratterizzare una funzione. In generale, gli ingressi ammissibili per una data corrispondenza sono soggetti a restrizioni naturali, legate alla natura stessa della corrispondenza.

Nel primo esempio il numero reale "di partenza" ha il significato di tempo; immaginando di lasciar cadere l'oggetto a un tempo iniziale $t = 0$ è evidente che ci si dovrà limitare a tempi $t \geq 0$.
Nel secondo esempio, evidentemente dovrà essere $0 \leq i \leq 1$. Infine, nel terzo esempio, dal significato di f, deve essere $f \geq 0$, mentre la formula impone ulteriormente $f > 0$.

L'insieme dei valori "di partenza" ammissibili per una data funzione prende il nome di *dominio*.

Spesso si usano le locuzioni *variabile indipendente* per indicare un ingresso generico e *variabile dipendente* per indicare l'uscita.
I prossimi esempi sono di tipo un po' diverso:

- Consideriamo una sbarra metallica sottile di lunghezza L; possiamo rappresentarla con l'intervallo $[0, L]$. La temperatura T nel punto x della sbarra all'istante t, è espressa da un numero che dipende dalla *coppia* di numeri (x, t), secondo una certa legge. In simboli:

$$(x, t) \mapsto T(x, t).$$

Questa volta l'*ingresso* non è un numero, ma piuttosto una coppia ordinata di numeri (x, t), mentre l'uscita è ancora un numero. D'altro canto, se $x \in [0, L]$ e $t \geq 0$, la coppia (x, t) è un elemento dell'insieme $A = [0, L] \times [0, +\infty)$. Si può ancora dire, quindi, che ad ogni elemento di un certo *insieme* A viene associato univocamente un elemento di un altro insieme B (qui $B = \mathbb{R}$).

- Se A è l'insieme degli studenti di un certo corso e B è l'insieme dei loro nomi, la corrispondenza che associa ad ogni studente il suo nome è univoca (mentre il viceversa non è necessariamente vero: potrebbero esserci due studenti con lo stesso nome). In questo caso né A né B sono insiemi numerici, e tuttavia risulta ben definita una corrispondenza univoca tra questi due insiemi.

Si arriva così al concetto generale di funzione[1]:

[1]Questa definizione moderna di funzione è dovuta a Dirichlet, in una memoria del 1837.

Definizione 2.1 Dati due insiemi A, B qualsiasi, una *funzione f di dominio A a valori in B* (o anche "di *codominio B*") è una qualsiasi legge che ad ogni elemento di A associa uno e un solo elemento di B. Scriveremo:

$$f : A \to B$$

(che si legge "f definita da A a B"). La scrittura

$$f : x \mapsto f(x)$$

(che si legge "f associa $f(x)$ ad x") indica come la funzione f agisce sugli elementi. Il simbolo $f(x)$ indica il valore che la funzione f associa ad x, e non va confuso col simbolo f, che denota la funzione stessa. ■

Ad esempio, per la funzione k sopra definita si avrebbe:

$$k : [0, 1] \to \mathbb{R}$$
$$k : i \mapsto \left(1 + \frac{i}{12}\right)^{12}$$

Si usa anche la scrittura

$$y = f(x)$$

per indicare che la variabile y è funzione di x, ossia che y è il valore che f associa all'ingresso x. Di nuovo, la proprietà caratteristica di f, affinché la si possa chiamare funzione, è l'*univocità* della corrispondenza: assegnato l'elemento di "ingresso" $a \in A$, l'elemento di "uscita" $b = f(a)$ dev'essere univocamente determinato.

Si può pensare a una funzione come a una scatola nera che a ogni ingresso ammissibile x associa un'*unica* uscita $f(x)$:

Figura 2.2.

L'uscita corrispondente a x si chiama *immagine* di x; l'insieme delle possibili uscite si chiama *immagine* di D tramite f e si indica con il simbolo $f(D)$ o $\mathrm{Im} f$. Si noti che nella scrittura $f : A \to B$, il codominio B può essere più grande dell'immagine $f(A)$ (ossia in generale è $f(A) \subseteq B$). Ad esempio, se f ha valori reali, solitamente si scrive $f : A \to \mathbb{R}$ senza precisare quale sia l'effettiva immagine di f.

Alcune classi di funzioni di cui ci occuperemo in questo corso sono le seguenti:

- le funzioni $f : \mathbb{N} \to \mathbb{R}$, che si dicono *successioni*: le studieremo nel capitolo 3, paragrafo 1;
- le funzioni $f : \mathbb{R} \to \mathbb{R}$ (*funzioni reali di variabile reale*), di cui ci occuperemo a partire da questo capitolo;

- le funzioni $f : \mathbb{R}^n \to \mathbb{R}^m$ di tipo lineare, dette anche *trasformazioni lineari*, di cui si occupa l'algebra lineare;
- le funzioni $f : \mathbb{R}^n \to \mathbb{R}^m$ (non necessariamente lineari), di cui ci occuperemo nel secondo volume.

Infine, nello studio del calcolo differenziale e integrale, incontreremo anche alcuni esempi di funzioni definite tra insiemi che, a loro volta, hanno per elementi altre funzioni. Come si vede, il concetto di *funzione* (come quello di *insieme*, introdotto nel cap. 1) fa attualmente parte del linguaggio di base della matematica, che unifica tra loro concetti e oggetti molto diversi.

2 FUNZIONI REALI DI VARIABILE REALE

2.1 Generalità

Nel paragrafo precedente è stato introdotto il concetto di funzione nella sua generalità. Le funzioni di cui ci occuperemo a partire da questo capitolo, sono le *funzioni reali di variabile reale*, ossia le funzioni per cui sia la variabile di "ingresso" che quella di "uscita" sono numeri reali. La funzione ha dunque per *dominio* un sottoinsiene di $\mathbb{R}$ e per *codominio* $\mathbb{R}$; l'*immagine* di f sarà anch'essa un sottoinsieme di $\mathbb{R}$:

$$f : D \to \mathbb{R} \text{ con } D \subseteq \mathbb{R}$$

$$f : x \mapsto f(x)$$

Le funzioni più comuni, come vedremo, hanno come dominio e come immagine un *intervallo* (che eventualmente è tutto $\mathbb{R}$) o l'*unione di un numero finito di intervalli* (cfr. par. 4.4 del cap. 1).

La dipendenza di $f(x)$ da x si visualizza efficacemente disegnando il *grafico* di f, ossia l'insieme dei punti del piano di coordinate (x, y) con

$$y = f(x)$$

e x variabile nel dominio D (v. fig. 2.3). Per le funzioni più comuni, il grafico è una curva, nel senso intuitivo del termine.

La proprietà fondamentale che fa di f una funzione, ossia il fatto che ad ogni ingresso $x \in D$ faccia corrispondere *una e una sola* uscita $f(x) \in \mathbb{R}$, ha allora il seguente significato geometrico:

Figura 2.3. Grafico di una funzione di dominio $D = [a, b]$.

ogni retta parallela all'asse delle ordinate che taglia l'asse delle ascisse in un punto x del dominio D, interseca il grafico di f in uno e un sol punto.

Infatti, se la retta non intersecasse il grafico significherebbe che all'ingresso x non corrisponde alcuna uscita, mentre se lo intersecasse in più di un punto significherebbe che all'ingresso x corrispondono più uscite distinte, cadendo quindi l'univocità della funzione (v. fig. 2.4).

Figura 2.4. Questa curva non è il grafico di una funzione: all'ingresso x quale uscita $f(x)$ corrisponde?

Si noti che, invece, nulla impedisce che una retta parallela all'asse delle *ascisse* tagli il grafico di f in più punti (o in nessun punto).

Per le funzioni reali di variabile reale si possono introdurre alcune importanti nozioni, che non sempre hanno senso per funzioni di tipo più generale.

2.2 Funzioni limitate

Se il grafico di una funzione $f : D \longrightarrow \mathbb{R}$ è contenuto nel semipiano inferiore delimitato da una retta parallela all'asse delle ascisse, per esempio di equazione $y = M$, la funzione si dice *limitata superiormente*. Analiticamente significa che

$$f(x) \leq M \qquad \text{per ogni} \qquad x \in D$$

Analogamente, f si dirà *limitata inferiormente* se il suo grafico è contenuto nel semipiano superiore

$$f(x) \geq m \qquad \text{per ogni} \qquad x \in D$$

Una funzione si dirà *limitata* se è limitata sia inferiormente che superiormente. Il grafico di una funzione limitata è contenuto in una striscia orizzontale del piano xy.

Per esempio (v. fig. 2.5):

- $x \longmapsto x^3$, $x \in \mathbb{R}$, non è limitata né superiormente, né inferiormente.
- $x \longmapsto x^2$, $x \in \mathbb{R}$, è limitata inferiormente; infatti $x^2 \geq 0$, $\forall x \in \mathbb{R}$.
- $x \longmapsto \frac{1}{1+x^2}$, $x \in \mathbb{R}$, è limitata, poiché $0 < \frac{1}{1+x^2} \leq 1$, $\forall x \in \mathbb{R}$.

Equivalentemente, si può dire che una funzione è limitata superiormente (limitata inferiormente, limitata) se, rispettivamente, la sua *immagine* è un sottoinsieme di $\mathbb{R}$ limitato superiormente (limitato inferiormente, limitato).

$x \longmapsto x^3$
non limitata
(né superiormente,
né inferiormente)

$x \longmapsto x^2$
limitata inferiormente

$x \longmapsto \frac{1}{1+x^2}$
limitata

Figura 2.5.

2.3 Funzioni simmetriche

I grafici di alcune funzioni presentano particolari proprietà di simmetria. Per esempio vi sono funzioni il cui grafico è simmetrico rispetto all'asse delle ordinate. Queste funzioni si chiamano funzioni *pari*, hanno un dominio simmetrico rispetto a $x = 0$ (tipicamente un intervallo del tipo $(-a, a)$) e sono caratterizzate dalla relazione

$$f(-x) = f(x) \tag{2.1}$$

che esprime l'uguaglianza delle ordinate corrispondenti ai punti x e $-x$, simmetrici rispetto a $x = 0$.

Funzioni che hanno il grafico simmetrico rispetto all'origine si chiamano *dispari*; anch'esse hanno dominio simmetrico rispetto a $x = 0$ e sono caratterizzate dalla relazione

$$f(-x) = -f(x) \tag{2.2}$$

Per esempio, la funzione $x \longmapsto x^2$ è pari, mentre $x \longmapsto x^3$ è dispari. Più in generale, le potenze a esponente intero sono funzioni pari (dispari) se l'esponente è pari (dispari).

grafico di una funzione *pari*

grafico di una funzione *dispari*

Figura 2.6.

Notiamo che una funzione *non può*, invece, avere il grafico simmetrico *rispetto all'asse* x: per quanto osservato nel paragrafo 2.1, una curva con tale simmetria non sarebbe il grafico di una funzione, venendo meno l'univocità della corrispondenza.

2.4 Funzioni monotone

Una funzione si dice *monotona crescente* o *non decrescente* se per ogni coppia di punti x_1, x_2 nel dominio di f si ha:

$$x_1 > x_2 \quad \text{implica} \quad f(x_1) \geq f(x_2) \tag{2.3}$$

Se nella (2.3) vale il maggiore anziché il maggiore o uguale si dice che f è *strettamente crescente*.

Se invece nella (2.3) si sostituisce "$\geq$" con "$\leq$" f si dice *decrescente* o *non crescente*; se poi la (2.3) vale con "$<$" anziché con "$\leq$" f si dirà *strettamente decrescente*.

In altri termini: f è crescente (strettamente crescente) se, all'aumentare di x, l'ordinata corrispondente sul grafico non diminuisce (aumenta); f è decrescente (strettamente decrescente) se, all'aumentare di x, l'ordinata corrispondente non aumenta (diminuisce).

Figura 2.7.

Per esempio, $x \longmapsto x^3$ è strettamente crescente; la funzione costante $x \longmapsto k$ (che ha come grafico la retta di equazione $y = k$) è sia non decrescente sia non crescente. Tutte queste funzioni (crescenti o decrescenti, strettamente o non) si dicono *monotone*.

2.5 Funzioni periodiche

La funzione$f : D \longrightarrow \mathbb{R}$ (non costante) è *periodica* di periodo T, $T > 0$, se T è il più piccolo numero reale positivo tale che

$$f(x+T) = f(x) \quad \text{per ogni} \quad x \in D$$

Ogni intervallo di lunghezza T, contenuto in D, si chiama *intervallo di periodicità*. Basterà disegnare il grafico di f su un qualunque intervallo di periodicità, per conoscere il grafico di f su tutto il dominio (v. fig. 2.8).

Tipici esempi di funzioni periodiche sono le funzioni trigonometriche $x \longmapsto \sin x$ ($T = 2\pi$), $x \longmapsto \cos x$ ($T = 2\pi$), $x \longmapsto \operatorname{tg} x$ ($T = \pi$) sulle quali torneremo in un prossimo paragrafo.

Figura 2.8. Funzione periodica di periodo 2.

3 FUNZIONI ELEMENTARI

Esaminiamo in questo paragrafo alcune tra le più comuni funzioni numeriche, evidenziandone le proprietà principali e segnalandone uso e interpretazione in qualche contesto applicativo.

Lo studente conoscerà già, probabilmente, la maggior parte di queste funzioni. Si tenga presente che le proprietà che qui richiameremo brevemente (proprietà algebriche e loro utilizzo, grafici,...) devono essere possedute con padronanza, per poter procedere con efficacia nello studio del calcolo infinitesimale. Lo studente che trovasse difficoltà, ad esempio, nello svolgere gli esercizi riportati alla fine di questo capitolo, è invitato a dedicare qualche tempo allo studio di questi argomenti elementari.[2]

3.1 Funzioni potenza

Riserviamo la denominazione di *funzioni potenza* a funzioni del tipo seguente:

$$\boxed{f(x) = kx^{\alpha}} \qquad (\alpha \neq 0) \tag{3.1}$$

dove k e α sono numeri reali. In generale, queste funzioni sono definite solo per $x \geq 0$, se $\alpha > 0$, e per $x > 0$, se $\alpha < 0$.

La (3.1) indica che l'uscita $f(x)$ è proporzionale, secondo la costante di proporzionalità k, a x^{α}. Cominciamo a passare in rassegna alcuni esempi di funzioni potenza che si incontrano in applicazioni di vario tipo, quindi faremo alcune puntualizzazioni generali.

Negli esempi sottostanti usiamo lettere più appropriate alla situazione, per l'ingresso e l'uscita.

Moto rettilineo uniforme

La nota legge cinematica

$$\begin{array}{ccc} s(t) & = & v \cdot t \\ \downarrow & & \downarrow \\ \text{spazio} & & \text{tempo} \end{array}$$

esprime la proporzionalità dello spazio percorso rispetto al tempo impiegato a percorrerlo. La costante di proporzionalità è la velocità.

[2] Si rimanda, ad esempio, al testo: M. Bramanti: *PreCalculus*. Ed. Esculapio.

Capacità di un condensatore

In un comune condensatore, la quantità di carica Q presente sulle armature e la differenza di potenziale ΔV esistente tra queste sono legate dalla formula

$$Q = C \cdot \Delta V$$

La costante di proporzionalità C è la capacità del condensatore.

Nei due esempi precedenti abbiamo dunque leggi del tipo

$$\boxed{f(x) = kx} \qquad (\alpha = 1 \text{ nella } (3.1))$$

Il grafico al variare di k di f è ben noto:

Figura 2.9.

Trasformazioni isoterme

In una trasformazione a temperatura costante di una mole di un gas ideale, pressione e volume sono legati dalla legge

$$p = \frac{RT}{V}$$

Cioè, p è inversamente proporzionale a V secondo la costante RT, dove $R = 1{,}986$ cal/K e T è la temperatura assoluta.

Potenziale elettrico generato da una carica puntiforme q

Sia q una carica elettrica concentrata in un punto Q dello spazio. Se P è un altro punto a distanza r da Q, il potenziale in Q prodotto dalla carica q è dato da

$$V = h\frac{q}{r}$$

dove h è una costante che dipende dalle unità di misura usate.

Ancora, V è inversamente proporzionale alla distanza tra P e Q.

Nei due esempi precedenti, abbiamo leggi del tipo "proporzionalità inversa":

$$\boxed{f(x) = \frac{k}{x}} \qquad (\alpha = -1 \text{ nella } (3.1))$$

Al variare di k (che qui consideriamo > 0) si ottiene una famiglia di iperboli equilatere.

 © 978-88-08-06485-1

Figura 2.10.

Superficie e volume di una sfera in funzione del raggio

Per una sfera di raggio r, superficie e volume sono assegnati dalle formule seguenti:

$$S = 4\pi r^2 \qquad V = \frac{4}{3}\pi r^3$$

Cioè, la superficie è proporzionale al quadrato del raggio, mentre il volume è proporzionale al suo cubo.

Abbiamo dunque leggi del tipo

$$\boxed{f(x) = kx^n} \quad \text{con} \quad n = 2, n = 3$$

i cui grafici al variare di k sono rappresentati qui di seguito:

Figura 2.11.

Moto lungo un piano inclinato

Se una sferetta d'acciaio rotola senza attrito lungo un piano inclinato dal punto B al punto A come in figura, sotto la sola azione della gravità, giunge al punto A con

velocità data dalla formula

$$v = \sqrt{2gh} \quad (g = \text{accelerazione di gravità}).$$

Figura 2.12.

Si tratta di una legge del tipo

$$\boxed{f(x) = k\sqrt{x}}$$

($\alpha = \frac{1}{2}$ nella (3.1)), i cui grafici al variare di k sono illustrati sotto.

Figura 2.13.

Trasformazioni adiabatiche

In una trasformazione adiabatica di una mole di un gas ideale, pressione e volume sono legati dalla formula

$$p = \frac{c}{V^\gamma} \qquad (c \text{ dipende da } R \text{ e } T)$$

dove $\gamma = \frac{5}{3}$ per un gas monoatomico e $\gamma = \frac{7}{5}$ per un gas diatomico.

Si tratta di una legge del tipo

$$\boxed{f(x) = kx^{-\gamma}} \qquad (\alpha = -\gamma \text{ nella } (3.1))$$

Al variare di k, i grafici sono qualitativamente simili alle iperboli della figura 2.10 (limitandosi al primo quadrante se γ è irrazionale o se è una frazione ridotta ai minimi termini con denominatore pari).

Facciamo ora alcune osservazioni generali. Ricordiamo anzitutto che l'operazione di elevamento a potenza è ben definita per qualunque esponente se la base è positiva, ma può essere definita anche con base negativa se l'esponente è un intero oppure un razionale (frazione) con denominatore dispari.

Precisamente: la funzione $x^{m/n}$ con n intero dispari è definita anche per $x < 0$ ed è pari se m è un intero pari, dispari se m è un intero dispari. Ricapitoliamo i vari casi possibili dal punto di vista dei *grafici* di queste funzioni.

Cominciamo dalle *potenze a esponente razionale*:

$$f(x) = x^{m/n}$$

con m, n interi ridotti ai minimi termini. Le situazioni qualitativamente diverse, nel caso in cui l'esponente è positivo, sono schematizzate dagli esempi seguenti, che raccolgono tutti i casi possibili (a parte il caso banale $f(x) = x$):

Figura 2.14.

A commento dei grafici precedenti, aggiungiamo le seguenti osservazioni, di cui lo studente è invitato a rendersi conto:

- in $(0, +\infty)$ la funzione è strettamente crescente, mentre in $(-\infty, 0)$, se è definita, basta tener conto della simmetria di f per decidere se cresce o decresce;
- in $(0, +\infty)$ la funzione presenta un andamento simile a quello delle funzioni x^2 o $\sqrt{x}$, rispettivamente (e cioè tangente all'asse x o all'asse y, nell'origine), a seconda che l'esponente sia maggiore o minore di 1. Quest'ultima affermazione potrà essere dimostrata molto semplicemente una volta introdotto il concetto di derivata, nel capitolo 4.

Se l'esponente è negativo, le situazioni qualitativamente diverse sono schematizzate dagli esempi seguenti:

Figura 2.15.

In questi casi la funzione è strettamente decrescente in $(0, +\infty)$, mentre in $(-\infty, 0)$, se è definita, basta tener conto della simmetria di f per decidere se cresce o decresce.

Per le funzioni *potenza a esponente reale* (ma non razionale) la casistica si semplifica, perché $f(x)$ non è definita per $x < 0$. Precisamente, $f(x) = kx^\alpha$ è definita per $x \geq 0$ se $\alpha > 0$, per $x > 0$ se $\alpha < 0$. Le situazioni possibili sono le seguenti:

Figura 2.16. Grafico delle potenze a esponente reale.

3.2 Funzioni esponenziali e logaritmiche

In questo e nel prossimo paragrafo descriveremo due classi di funzioni notevoli: le funzioni esponenziali (e logaritmiche) e le funzioni circolari o trigonometriche.

L'importanza di queste due classi di funzioni deriva dal fatto che esse sono in un certo senso i prototipi utili a descrivere due corrispondenti gruppi di fenomeni frequentissimi in natura: i fenomeni di decadimento (o, al contrario, di crescita) e i fenomeni periodici. Esempi di fenomeni del primo tipo sono: il decadimento radioattivo, il processo di raffreddamento di un corpo, il diffondersi di un'infezione o il moltiplicarsi di una colonia di batteri. Esempi di fenomeni del secondo tipo sono: il moto dei pianeti, la propagazione di onde (meccaniche o elettromagnetiche), il moto di un pendolo, certi andamenti di malattie influenzali di carattere stagionale ecc. Ora, accade che spesso le funzioni esponenziali servano a descrivere i fenomeni del primo tipo mentre quelle trigonometriche siano utili a descrivere quelli del secondo tipo. Il motivo profondo di questo fatto sarà messo in luce nello studio delle equazioni differenziali di cui ci occuperemo nel secondo volume.

Se a è un numero reale *positivo* e *diverso da* 1, la funzione

$$f : (0, +\infty) \longmapsto \mathbb{R} \qquad f(x) = \log_a x$$

si chiama funzione *logaritmo in base* a, mentre la funzione

$$g : \mathbb{R} \longmapsto \mathbb{R} \qquad g(x) = a^x$$

si chiama funzione *esponenziale in base* a.

Figura 2.17. a) Grafico di $f(x) = \log_a x$; b) grafico di $g(x) = a^x$.

Le funzioni f e g sono legate dalla relazione fondamentale

$$\boxed{x = a^{\log_a x} \quad (x > 0, a > 0, a \neq 1)} \tag{3.2}$$

ossia,

$$y = \log_a x \quad \text{equivale a} \quad x = a^y \tag{3.3}$$

Confrontando invece le funzioni potenza (introdotte nel paragrafo 3.1) con le funzioni esponenziali, notiamo che la stessa operazione algebrica di elevamento a potenza nel campo reale, a^b, è alla base della definizione di queste due classi di funzioni:

- se l'esponente b è fissato e la base è variabile abbiamo le funzioni potenza:

$$x \mapsto x^b$$

- se la base a è fissata e l'esponente è variabile abbiamo le funzioni esponenziali:

$$x \mapsto a^x$$

La base naturale

Tra le funzioni esponenziali e logaritmiche, si incontrano con particolare frequenza quelle la cui base è *il numero e di Nepero*. Si tratta di una costante molto importante in matematica, che sarà definita e trattata ampiamente nel prossimo capitolo 3. Per ora è sufficiente sapere che e vale circa 2,7, perciò le funzioni e^x e $\log_e x$ (che scriveremo semplicemente $\log x$ sottointendendo la base e) hanno i grafici caratteristici delle funzioni esponenziali e logaritmiche di base maggiore di 1.

Notiamo che una funzione esponenziale a^x (di base qualsiasi, positiva e diversa da 1) si può sempre esprimere nella forma e^{bx}, scegliendo $b = \log a$.

Esempi

3.1 Se $[H^+]$ indica la concentrazione di ioni idrogeno (= numero di moli/cm^3) in una soluzione, si definisce come misura della sua acidità la quantità

$$\text{pH} = -\log_{10}[H^+]$$

Poiché per l'acqua pura $[H^+] = 10^{-7}$, il corrispondente pH è pari a $-\log_{10} 10^{-7} = 7$, equivalente a soluzione neutra. Se pH > 7 la soluzione è acida, se pH < 7 la soluzione è basica.

3.2 *Scarica di un condensatore.* In un circuito con capacità e resistenza, come quello indicato in figura 2.18a, la carica $q(t)$ è assegnata dalla formula

$$q(t) = Qe^{-t/\tau} \tag{3.4}$$

dove Q è la carica iniziale (al tempo $t = 0$) e $\tau = RC$.

Figura 2.18. La carica decresce secondo la legge esponenziale (3.4) come indicato in b).

La funzione in (3.4) si può scrivere nella forma $q(t) = Qa^t$ con $a = e^{-1/\tau} < 1$.

Nella tabella seguente sono raccolte le principali proprietà delle funzioni in oggetto. Le altre proprietà algebriche dei logaritmi sono state richiamate nel capitolo 1.

$f(x) = \log_a x$	$g(x) = a^x$
dominio: $(0, +\infty)$; immagine: $\mathbb{R}$	dominio: $\mathbb{R}$; immagine: $(0, +\infty)$
strettamente $\begin{cases} \text{crescente} & \text{se } a > 1 \\ \text{decrescente} & \text{se } a < 1 \end{cases}$	strettamente $\begin{cases} \text{crescente} & \text{se } a > 1 \\ \text{decrescente} & \text{se } a < 1 \end{cases}$
$\log_a x \cdot y = \log_a x + \log_a y$	$a^{x+y} = a^x \cdot a^y$

3.3 Funzioni trigonometriche

Le funzioni trigonometriche elementari sono:

$x \longmapsto \cos x$	$x \longmapsto \sin x$	$x \longmapsto \operatorname{tg} x$	$x \longmapsto \operatorname{cotg} x$
(coseno di x)	(seno di x)	(tangente di x)	(cotangente di x)

La variabile x ha il significato di *misura in radianti* di un angolo. I significati geometrici delle quattro funzioni sono illustrati nella figura 2.19.

Dalla figura si vede che $\cos x$ = ascissa di P, $\sin x$ = ordinata di P, $\operatorname{tg} x$ = = lunghezza con segno del segmento AT, $\operatorname{cotg} x$ = lunghezza con segno del segmento BS.

Valgono le relazioni fondamentali seguenti

$$\boxed{(\sin x)^2 + (\cos x)^2 = 1 \qquad \operatorname{tg} x = \frac{\sin x}{\cos x} \qquad \operatorname{cotg} x = \frac{\cos x}{\sin x} = \frac{1}{\operatorname{tg} x}}$$

La prima si ricava dal fatto che P si trova sulla circonferenza di equazione $u^2 + v^2 = 1$; la seconda segue dalla similitudine dei triangoli OPQ e OTA; la terza segue dalla similitudine dei triangoli OPQ e OBS.

Nella seguente tabella sono raggruppate le principali proprietà delle funzioni seno, coseno, tangente.

Figura 2.19. Definizione di seno, coseno, tangente e cotangente.

$x \longmapsto \sin x, x \longmapsto \cos x$	$x \longmapsto \operatorname{tg} x$	$x \longmapsto \operatorname{cotg} x$
dominio: $\mathbb{R}$; immagine: $[-1,1]$	dominio: $\mathbb{R} \setminus \{\frac{\pi}{2} + k\pi\}$; immagine: $\mathbb{R}$	dominio: $\mathbb{R} \setminus \{k\pi\}$; immagine: $\mathbb{R}$
periodiche di periodo $T = 2\pi$	periodica di periodo $T = \pi$	periodica di periodo $T = \pi$
$\sin(-x) = -\sin x$ (funzione dispari)	$\operatorname{tg}(-x) = -\operatorname{tg} x$ (funzione dispari)	$\operatorname{cotg}(-x) = -\operatorname{cotg} x$ (funzione dispari)
$\cos(-x) = \cos x$ (funzione pari)		

Poiché $\sin x = \cos\left(x - \frac{\pi}{2}\right)$ ricaviamo che $\sin x$ si ottiene da $\cos x$ con uno *sfasamento* di x pari a $\frac{\pi}{2}$, ossia, il grafico di $\sin x$ si ottiene da quello di $\cos x$ traslandolo a destra di $\frac{\pi}{2}$ (v. fig. 2.20).

3.4 Fenomeni vibratori

Abbiamo visto che le funzioni seno e coseno sono periodiche di periodo 2π. Più in generale le funzioni

$$t \longmapsto a \sin \omega t \qquad t \longmapsto b \cos \omega t \tag{3.5}$$

dove a, b, ω sono numeri reali positivi, sono periodiche di periodo $T = \frac{2\pi}{\omega}$. Infatti, per esempio

$$\sin\left[\omega\left(t + \frac{2\pi}{\omega}\right)\right] = \sin(\omega t + 2\pi) = \sin \omega t$$

Inoltre, essendo $|\sin \omega t| \leq 1$, $|\cos \omega t| \leq 1$ si ha

$$|a \sin \omega t| \leq a\,, \qquad |b \cos \omega t| \leq b$$

Le funzioni (3.5) descrivono vibrazioni elementari di *ampiezza* a e b, rispettivamente, di *pulsazione* $\omega = \frac{2\pi}{T}$ e *frequenza* $\nu = \frac{1}{T}$.

La funzione

$$h(t) = a \sin \omega t + b \cos \omega t$$

Figura 2.20. Grafici di $y = \cos x$, $y = \sin x$, $y = \operatorname{tg} x$ e $y = \operatorname{cotg} x$.

descrive la sovrapposizione delle due vibrazioni elementari di uguale pulsazione ω. Quest'ultima è ancora una vibrazione elementare *sfasata* rispetto alle precedenti. Infatti, posto $A = \sqrt{a^2 + b^2}$, si può scrivere

$$h(t) = A\frac{a}{\sqrt{a^2+b^2}}\sin\omega t + A\frac{b}{\sqrt{a^2+b^2}}\cos\omega t \tag{3.6}$$

Osserviamo ora che i numeri $\alpha = a/\sqrt{a^2+b^2}$ e $\beta = b/\sqrt{a^2+b^2}$ soddisfano le condizioni

$$-1 \le \alpha \le 1 \qquad -1 \le \beta \le 1 \qquad \alpha^2 + \beta^2 = 1$$

Esiste quindi un unico angolo φ tale che

$$\cos\varphi = \alpha\,, \qquad \sin\varphi = \beta$$

cosicché la (3.6) si può riscrivere nella forma seguente

$$h(t) = A\cos\varphi\sin\omega t + A\sin\varphi\cos\omega t = A\sin(\omega t + \varphi) \tag{3.7}$$

Dunque $h(t)$ rappresenta una vibrazione elementare di ampiezza A, pulsazione $\omega = \frac{2\pi}{T}$, sfasata di un angolo φ.

 © 978-88-08-06485-1

In sintesi

$$\boxed{\begin{array}{ll} a\sin\omega t + b\cos\omega t = A\sin(\omega t+\varphi) & \\ A = \sqrt{a^2+b^2} & \begin{cases} a = A\cos\varphi \\ b = A\sin\varphi \end{cases} \end{array}}$$

Sotto condizioni abbastanza generali, un fenomeno naturale periodico si potrà scrivere come sovrapposizione di un numero finito o infinito di vibrazioni elementari di frequenza diversa; ciò conduce al concetto di serie di Fourier, ovvero a somme del tipo

$$\sum a_n \cos n\omega t + b_n \sin n\omega t$$

che tratteremo nel secondo volume.

Figura 2.21. La vibrazione $h(t) = A\sin(\omega t+\varphi)$.

Moltiplicando una vibrazione elementare per potenze o esponenziali si possono modellizzare effetti di smorzamento o di amplificazione.

Per esempio, consideriamo la funzione

$$h(t) = t\sin\omega t \qquad (t \geq 0)$$

essendo $-1 \leq \sin\omega t \leq 1$, si ha

$$-t \leq t\sin\omega t \leq t$$

e quindi il grafico di h si trova tra i grafici delle rette di equazione $y = -t$, $y = t$. Nei punti in cui $\sin\omega t = 1$, cioè $t = \frac{\pi}{2\omega} + k\frac{2\pi}{\omega}$ $(k = 0, 1, 2, \ldots)$, il grafico di h tocca quello di $y = t$, mentre nei punti in cui $\sin\omega t = -1$, cioè $t = \frac{3\pi}{2\omega} + k\frac{2\pi}{\omega}$, il grafico di h tocca quello di $y = -t$.

Il grafico di h è perciò il seguente:

Figura 2.22. Grafico di $h(t) = t \sin \omega t$ (oscillazioni amplificate).

Dal grafico si vede come la moltiplicazione per t abbia l'effetto di una amplificazione della vibrazione, all'aumentare di t.
Analogamente, la funzione

$$k(t) = e^{-\alpha t} \sin \omega t \qquad (\alpha > 0)$$

modellizza una *vibrazione smorzata*.

Ricordando che $e^{-\alpha t} = \left(\frac{1}{e^{\alpha}}\right)^t$ è un'esponenziale con base minore di 1, considerazioni analoghe a quelle svolte per la funzione h, indicano che il grafico di k è compreso tra i grafici delle funzioni $y_1 = e^{-\alpha t}$ e $y_2 = -e^{-\alpha t}$, come in figura:

Figura 2.23. Grafico di $k(t) = e^{-\alpha t} \sin \omega t$ (oscillazioni smorzate).

3.5 Funzioni parte intera e mantissa

Due funzioni che tipicamente si incontrano nella scrittura di algoritmi sono la funzione *parte intera* e la funzione *mantissa* o *parte decimale*. La funzione "parte intera di x", talvolta indicata con $[x]$, è definita da:

$$[x] = \text{quell'intero } n \text{ tale che } n \leq x < n+1.$$

Ad esempio,

$$\begin{aligned} [2,38] &= 2; \\ [3] &= 3; \\ [-1,8] &= -2. \end{aligned}$$

Mentre per i numeri positivi, quindi, la parte intera si ottiene semplicemente "buttando via le cifre dopo la virgola", per i numeri negativi occorre prendere il massimo intero $\leq x$, che è diverso da quello che si ottiene buttando via le cifre dopo la virgola (tranne nel caso in cui x sia già un intero). Il grafico della funzione parte intera è:

Figura 2.24. Grafico della funzione parte intera.

La mantissa, o parte decimale di x, indicata con (x) oppure $\text{mant}(x)$, è definita da:

$$(x) = x - [x].$$

Ad esempio, si ha:

$$\begin{aligned} (2{,}38) &= 0{,}38; \\ (3) &= 0; \\ (-1{,}8) &= 0{,}2. \end{aligned}$$

La mantissa quindi non è un intero ma un numero reale, compreso in $[0,1)$. Per i numeri positivi, si ottiene semplicemente "buttando via le cifre prima della virgola", per i numeri negativi la mantissa è il complemento a uno del numero che si ottiene buttando via le cifre prima della virgola.

Il grafico della funzione mantissa è:

Figura 2.25. Grafico della funzione mantissa.

Si noti che la mantissa è una funzione periodica di periodo 1; si tratta di una semplice funzione periodica di tipo diverso da quelle trigonometriche.

3.6 Funzioni iperboliche

Nelle applicazioni della matematica sono importanti certe combinazioni delle funzioni e^x ed e^{-x} che ora definiamo:

$$\begin{array}{ll}
\mathrm{Sh} : \textit{seno iperbolico} & \mathrm{Sh}\, x := \dfrac{e^x - e^{-x}}{2} \\[2ex]
\mathrm{Ch} : \textit{coseno iperbolico} & \mathrm{Ch}\, x := \dfrac{e^x + e^{-x}}{2} \\[2ex]
\mathrm{Th} : \textit{tangente iperbolica} & \mathrm{Th}\, x := \dfrac{\mathrm{Sh}\, x}{\mathrm{Ch}\, x} = \dfrac{e^x - e^{-x}}{e^x + e^{-x}}\,.
\end{array}$$

Esse sono evidentemente definite su tutto $\mathbb{R}$. Le seguenti proprietà discendono immediatamente dalla definizione:

1) $\mathrm{Sh}\,(-x) = -\mathrm{Sh}\, x$ (funzione dispari)
$\mathrm{Ch}(-x) = \mathrm{Ch}\, x$ (funzione pari)
$\mathrm{Th}(-x) = -\mathrm{Th}\, x$ (funzione dispari)

2) $\mathrm{Sh}\,(0) = 0 \qquad \mathrm{Ch}(0) = 1 \qquad \mathrm{Th}(0) = 0$
$\mathrm{Sh}\, x \le \dfrac{1}{2} e^x \le \mathrm{Ch}\, x \qquad \forall x \in \mathbb{R}$

3)
$$(\mathrm{Ch}\, x)^2 - (\mathrm{Sh}\, x)^2 = 1$$
$$\mathrm{Sh}(x+y) = \mathrm{Sh}\, x\, \mathrm{Ch}\, y + \mathrm{Sh}\, y\, \mathrm{Ch}\, x$$
$$\mathrm{Ch}(x+y) = \mathrm{Ch}\, x \mathrm{Ch}\, y + \mathrm{Sh}\, y\, \mathrm{Sh}\, x$$
$$\mathrm{Th}(x+y) = \frac{\mathrm{Th}\, x + \mathrm{Th}\, y}{1 + \mathrm{Th}\, x\, \mathrm{Th}\, y}$$
$$\mathrm{Sh}(2x) = 2\, \mathrm{Sh}\, x\, \mathrm{Ch}\, x$$
$$\mathrm{Ch}(2x) = (\mathrm{Ch}\, x)^2 + (\mathrm{Sh}\, x)^2$$

I grafici di queste funzioni sono illustrati nella figura 2.26.

Figura 2.26. Grafici di $\operatorname{Sh} x = \frac{e^x - e^{-x}}{2}$, $\operatorname{Ch} x = \frac{e^x + e^{-x}}{2}$, $\operatorname{Th} x = \frac{e^x - e^{-x}}{e^x + e^{-x}}$.

La funzione $\operatorname{Ch} x$ compare nella soluzione di un semplice problema fisico: trovare la sagoma lungo la quale si dispone un filo pesante, omogeneo, fissato per le due estremità. In queste ipotesi il filo descriverà, in un piano verticale, il grafico della funzione

$$f(x) = \operatorname{Ch} x$$

(a patto di scegliere opportunamente il sistema di riferimento e le unità di misura sui due assi). Per questo motivo la curva grafico di $\operatorname{Ch} x$ è anche detta *catenaria*.

3.7 Operazioni sui grafici

Conoscendo il grafico di una funzione $y = f(x)$, mediante semplici trasformazioni geometriche è possibile disegnare il grafico delle seguenti funzioni:

$$\begin{aligned}
y_1 &= f(x) + a, \quad a \in \mathbb{R};\\
y_2 &= f(x + a), \quad a \in \mathbb{R};\\
y_3 &= k \cdot f(x), \quad k \in \mathbb{R};\\
y_4 &= f(kx), \quad k \in \mathbb{R};\\
y_5 &= |f(x)|;\\
y_6 &= f(|x|).
\end{aligned}$$

Inoltre, più operazioni di questi tipi possono essere effettuate in sequenza. Pertanto, a partire dalle funzioni elementari precedentemente introdotte, è possibile costruire mediante queste operazioni i grafici di una grande varietà di nuove funzioni. Illustriamo su esempi il significato e l'azione di queste trasformazioni.

• Consideriamo per esempio $f(x) = \log x$. Allora $y_1 = f(x) + a = (\log x) + a$. La funzione y_1 è definita per $x > 0$, come $\log x$, e il grafico di y_1 si ottiene da quello di y *con una traslazione di a unità* verso l'alto se $a > 0$, verso il basso se $a < 0$. Infatti, la quantità a viene aggiunta al *valore* della funzione $f(x)$ (non alla variabile x), dopo che questa è stata calcolata.

Figura 2.27. Grafico di $\log x + a$, $a > 0$.

• Consideriamo ancora $f(x) = \log x$. Allora $y_2 = f(x+a) = \log(x+a)$. Essendo $\log x$ definito per $x > 0$, $\log(x+a)$ sarà definito per $x + a > 0$ e cioè per $x > -a$. Poiché $\log x = 0$ se $x = 1$, si ha $\log(x+a) = 0$ per $x + a = 1$ e cioè $x = 1 - a$.

In altri termini il grafico di y_2 si ottiene da quello di y *con una traslazione di a unità* a sinistra se $a > 0$, a destra se $a < 0$.

Figura 2.28. Grafici di $y = \log(x-1)$ e $y = \log(x+2)$.

• Il grafico di $y_3 = kf(x)$ si ottiene da quello di f moltiplicando per k tutte le ordinate $f(x)$. In particolare se $k = -1$ le ordinate sono semplicemente cambiate di segno, cosicché il grafico di y_2 è simmetrico, rispetto all'asse x, a quello di f.

Per esempio, sia $f(x) = \sin x$. I grafici di $y = 3\sin x$, $y = \frac{1}{2}\sin x$, $y = -\sin x$ sono riportati in figura 2.29.

Osserviamo che se $k > 1$, il grafico si "stira" nella direzione verticale, dilatando verso l'alto le ordinate positive e verso il basso quelle negative. Al contrario, se $0 < k < 1$ il grafico si contrae, sempre in direzione verticale. Perciò l'operazione di

Figura 2.29. Grafici di $y = 3\sin x$, $y = \frac{1}{2}\sin x$ e $y = -\sin x$.

moltiplicazione di $f(x)$ per k ha il significato geometrico di *dilatazione* (se $|k| > 1$) o *contrazione* (se $|k| < 1$) *sull'asse delle* y, eventualmente accompagnata da una *rifiessione rispetto all'asse* x, se $k < 0$.

• Il grafico di $y_4 = f(kx)$ si ottiene da quello di $f(x)$ con un cambiamento di scala sull'asse x. Se $k > 1$, kx cresce più rapidamente di x e perciò il grafico di y_4 sarà simile a quello di f ma con oscillazioni più rapide, ovvero sarà "compresso" in direzione orizzontale, di un fattore $1/k$. Analogamente se $0 < k < 1$ il grafico apparirà "dilatato" in direzione orizzontale, con oscillazioni più dolci. Se $k < 0$, oltre ad una compressione (se $|k| > 1$) o dilatazione (se $|k| < 1$) sull'asse delle x, ci sarà una riflessione rispetto all'asse delle y.

Ad esempio, si confrontino i grafici di $\sin x, \sin 2x, \sin \frac{x}{2}$ su $[0, 2\pi]$:

Figura 2.30. Grafici di $y = \sin x$, $y = \sin \frac{x}{2}$ e $y = \sin 2x$.

Per quanto riguarda la riflessione per $k < 0$, si confrontino anche i grafici di $\log x$ e $\log(-x)$:

Figura 2.31. Grafici di $y = \log x$ e $y = \log(-x)$.

• Per disegnare il grafico di $y_5 = |f(x)|$ ricordiamo che

$$|f(x)| = \begin{cases} f(x) & \text{se } f(x) \geq 0 \\ -f(x) & \text{se } f(x) < 0 \end{cases}$$

Dunque nel passare dal grafico di f a quello di $|f|$ i punti a ordinata non negativa rimangono inalterati mentre quelli a ordinata negativa vengono trasformati nei loro simmetrici rispetto all'asse x.

Il grafico di $|f|$ si ottiene perciò da quello di f "ribaltando" simmetricamente rispetto all'asse delle ascisse la parte del grafico di f che si trova nel semipiano inferiore e lasciando inalterato il resto. Per esempio, per le funzioni $y = x$ e $y = \sin x$ si ha:

Figura 2.32. Dal grafico di f a quello di $|f|$.

• Infine, per tracciare il grafico di $y_6 = f(|x|)$, osserviamo che:
$|x| = x$ per $x \geq 0$, e quindi nel semipiano destro i due grafici coincidono;
$|-x| = |x|$, e quindi y_6 è una funzione pari, perciò simmetrica rispetto all'asse delle ordinate.

Di conseguenza il grafico di $y_6 = f(|x|)$ verrà tracciato lasciando inalterato il grafico di f nel semipiano destro e ribaltandolo simmetricamente rispetto all'asse delle ordinate. Per esempio, se $y = e^x$ si ha:

Figura 2.33. Dal grafico di $y = e^x$ a quello di $y = e^{|x|}$.

3.8 Funzioni definite a tratti

Abbiamo visto come traslazioni, dilatazioni e riflessioni di vario tipo consentano di costruire funzioni più complesse a partire da quelle elementari. Un altro modo tipico con cui a partire dalle funzioni elementari se ne costruiscono di nuove, è quello di usare definizioni analitiche diverse su intervalli diversi. Si consideri il prossimo:

Esempio 3.3 Sia

$$f(x) = \begin{cases} \log x \text{ se } x > 1 \\ x^2 \text{ se } 0 < x \leq 1 \\ x \text{ se } x \leq 0 \end{cases}$$

Un attimo di riflessione mostra che la funzione f è definita su tutto $\mathbb{R}$ e ha il grafico seguente:

Figura 2.34.

Il valore $f(x)$ è calcolato mediante "istruzioni" diverse a seconda dell'intervallo in cui cade la x. Una funzione di questo tipo si chiama solitamente "funzione definita a tratti". Dal punto di vista logico / algoritmico, si può dire che una funzione di questo tipo ha la particolarità di essere costruita utilizzando, come "ingredienti", non solo le funzioni matematiche elementari discusse fin qui, ma anche la *funzione logica* "se... allora".[3].

Si presti attenzione all'insieme di definizione di una funzione definita a tratti: la funzione $f(x)$ dell'esempio è definita anche per $x < 0$ non ostante la presenza di $\log x$. Infatti, l'istruzione "calcola $\log x$" interviene solo se $x > 1$.

4 FUNZIONI COMPOSTE E INVERSE

4.1 Funzioni composte

Nel paragrafo 3.7 abbiamo illustrato come si può costruire il grafico di $|f|$ a partire da quello di f. La funzione $x \longmapsto |f(x)|$ è in realtà "composta" di due funzioni: dato x, si calcola $f(x)$; calcolato $f(x)$, si calcola $|f(x)|$.

Si tratta di operare in serie con due scatole nere, la prima corrispondente a f, la seconda al suo valore assoluto, secondo lo schema seguente:

$$x \longmapsto \boxed{\quad f \quad} \longmapsto f(x) \longmapsto \boxed{\quad |\cdot| \quad} \longmapsto |f(x)|$$

In generale, date due funzioni

$$f : E \longmapsto \mathbb{R} \qquad g : F \longmapsto \mathbb{R}$$

se $f(E) \subseteq F$ (cioè se per ogni $x \in E$ si ha che $f(x) \in F$) si può definire la funzione $h : E \longmapsto \mathbb{R}$ *composta di* f e g (nell'ordine), denotata col simbolo $g \circ f$, mediante la formula

$$h(x) = (g \circ f)(x) = g[f(x)] \tag{4.1}$$

ossia

$$\begin{tikzcd}
x \arrow[r] \arrow[dr, bend right] & \boxed{\quad f \quad} \arrow[r] & f(x) \arrow[r] & \boxed{\quad g \quad} \arrow[r] & g[f(x)] \\
& & \boxed{g \circ f} \arrow[ur, bend right] & &
\end{tikzcd}$$

Può accadere che risultino ben definite sia la composizione $g \circ f$ che $f \circ g$; ma in generale sarà

$$f \circ g \neq g \circ f$$

[3]Per implementare la funzione dell'esempio in qualche linguaggio di programmazione, occorrerebbe annidare due funzioni "se... allora" una dentro l'altra, secondo uno schema del tipo:

"se $x > 1$ allora $\log x$, altrimenti (se $x > 0$ allora x^2, altrimenti x)"

(il prodotto di composizione non è commutativo!) L'operazione di composizione si può estendere a tre o più fattori. Si verifica che se la composizione $(f \circ g) \circ r$ esiste, allora esiste anche $f \circ (g \circ r)$ e sono uguali (proprietà associativa)

$$(f \circ g) \circ r = f \circ (g \circ r)$$

Esempi

4.1 $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = x^2$ e $g : \mathbb{R} \longrightarrow \mathbb{R}$, $g(x) = \cos x$.
Poiché g è definita su tutto $\mathbb{R}$, $h = g \circ f$ è ben definita su $\mathbb{R}$ e vale la formula

$$h(x) = g[f(x)] = \cos(x^2)$$

ossia

$$x \longmapsto \boxed{(\cdot)^2} \longmapsto x^2 \longmapsto \boxed{\cos(\cdot)} \longmapsto \cos(x^2)$$

Si noti che anche $k = f \circ g$ è ben definita ed è data dalla formula

$$k(x) = (\cos x)^2$$

Si osservi che queste due funzioni sono ben diverse tra loro: ad esempio, $(\cos x)^2$ è periodica e mai negativa, $\cos(x^2)$ non è periodica e ha segno oscillante. Si vedano i prossimi grafici:

Figura 2.35. Grafico di $y = \cos(x^2)$ e di $(\cos x)^2$.

4.2 $f : \mathbb{R} \longrightarrow \mathbb{R}$, $f(x) = e^x$ e $g : [0, +\infty) \longrightarrow \mathbb{R}$, $g(x) = \sqrt{x}$.
Poiché $f(x) > 0$ per ogni x in $\mathbb{R}$, $h = g \circ f$ è ben definita in $\mathbb{R}$ e si ha $h(x) = \sqrt{e^x}$.

Anche $k = f \circ g$ è ben definita in $\mathbb{R}_+$ e si ha $k(x) = e^{\sqrt{x}}$.

Se una funzione $f : D \longrightarrow \mathbb{R}$ è tale che $f(D) \subseteq D$, allora si può comporre con sé stessa; la funzione

$$f^2 = f \circ f \qquad \text{ossia} \quad f^2(x) = f[f(x)] \tag{4.2}$$

viene detta *iterata seconda* di f. Analogamente, $f^n(x) = f[f[\ldots[f(x)]\ldots]$ (n volte) si dice iterata n-esima di f.

4.2 Funzioni invertibili; funzioni inverse

Supponiamo che f abbia come dominio un insieme $D \subseteq \mathbb{R}$. Per ogni ingresso $x \in D$ esiste un'unica uscita $f(x)$. Se succede che per ogni uscita $y \in f(D)$ esiste *un solo* ingresso $x \in D$ tale che $f(x) = y$, allora f si dice *invertibile*, e realizza una corrispondenza biunivoca tra D e $f(D)$.

Più formalmente, si dice che $f : D \to \mathbb{R}$ è invertibile in D se vale una delle seguenti condizioni (equivalenti):

(i) $\forall x_1, x_2 \in D$, $x_1 \neq x_2 \Longrightarrow f(x_1) \neq f(x_2)$
(ii) $\forall x_1, x_2 \in D$, $f(x_1) = f(x_2) \Longrightarrow x_1 = x_2$
(iii) $\forall y \in f(D)\ \exists!\ x \in D$ tale che $f(x) = y$.

La funzione che associa a ogni uscita $y \in f(D)$ l'unico ingresso $x \in D$ tale che $f(x) = y$ si chiama *funzione inversa di* f e si indica con il simbolo f^{-1}.

In sintesi

$$\begin{cases} y = f(x) \\ x \in D \end{cases} \qquad \text{equivale a} \qquad \begin{cases} x = f^{-1}(y) \\ y \in f(D) \end{cases} \tag{4.3}$$

La scatola nera di f^{-1} lavora a ritroso rispetto a quella di f, secondo lo schema seguente

Partire da x e ritornare dopo un giro in x equivale a $f^{-1}[f(x)] = x$ per ogni $x \in I$; partire da y e tornare in y equivale a $f[f^{-1}(y)] = y$, per ogni $y \in f(I)$.

La condizione di invertibilità equivale a richiedere che il grafico di f sia intersecato al massimo in un punto da *ogni* retta parallela all'asse delle ascisse.

La funzione in figura 2.36a non è invertibile in quanto per il valore di y indicato esistono tre punti x_1, x_2, x_3 che hanno immagine y. La funzione in figura 2.36b è invece invertibile in quanto ogni retta parallela all'asse delle ascisse o non interseca il grafico di f o lo interseca esattamente in un punto.

Classi di funzioni che sicuramente *non risultano* invertibili sono: le funzioni simmetriche pari ($f(-x) = f(x)$), le funzioni periodiche ($f(x + T)) = f(x)$).

Figura 2.36. a) Funzione non invertibile: all'uscita y corrispondono 3 ingressi. b) Funzione invertibile.

Osserviamo invece che:

Teorema 2.1 *Una funzione $f : D \to \mathbb{R}$ strettamente monotona in D è invertibile in D. Inoltre, la sua inversa è ancora strettamente monotona.*

Dimostrazione. Supponiamo per fissare le idee che f sia strettamente crescente in D. Siano $x_1, x_2 \in D$ e proviamo che

$$x_1 \neq x_2 \Longrightarrow f(x_1) \neq f(x_2).$$

Se $x_1 \neq x_2$, allora o $x_1 < x_2$, oppure $x_1 > x_2$. Per la monotonia stretta di f, nel primo caso si ha $f(x_1) < f(x_2)$, nel secondo $f(x_1) > f(x_2)$; in entrambi i casi $f(x_1) \neq f(x_2)$, perciò f è invertibile. Sia ora $x = f^{-1}(y)$ la sua funzione inversa, e proviamo che f^{-1} è strettamente crescente. Sia dunque $y_1 < y_2$. Se fosse $x_1 \geq x_2$ (dove $x_i = f^{-1}(y_i)$), poiché f è crescente avremmo $y_1 \geq y_2$, assurdo; quindi $x_1 < x_2$, ossia $f^{-1}(y_1) < f^{-1}(y_2)$, e f^{-1} è strettamente crescente. ♦

Si noti, ad ogni modo, che una funzione può essere invertibile anche senza essere strettamente monotona, come mostra il prossimo esempio:

Figura 2.37.

Nel prossimo capitolo, parlando di funzioni continue, diremo qualcosa di più sulle condizioni di invertibilità per una funzione definita su un intervallo.

Grafico di f^{-1}

Le relazioni (4.3) indicano che se il punto (x_0, y_0) è sul grafico di f allora il punto (y_0, x_0) sta sul grafico di f^{-1}. Essendo i punti (x_0, y_0) e (y_0, x_0) simmetrici rispetto alla bisettrice di equazione $y = x$, si deduce che il grafico di f^{-1} si ricava da quello di f per simmetria rispetto alla bisettrice $y = x$.

Figura 2.38. Il grafico di f e della sua inversa.

Se è nota l'espressione analitica di f, invertibile, l'espressione analitica di f^{-1} si trova cercando di risolvere (quasi sempre però è impossibile!) rispetto a x l'equazione

$$f(x) = y \qquad \text{(Trovare l'ingresso } x \text{ che produce l'uscita } y\text{)}$$

Per esempio, se $f(x) = 2x + 3$, f è invertibile su $\mathbb{R}$ e l'equazione $2x + 3 = y$ dà per f^{-1} l'espressione analitica

$$f^{-1}(y) = \frac{y-3}{2}$$

Coppie f, f^{-1} sono le seguenti:

$$f : \begin{cases} y = x^2 \\ x \geq 0 \end{cases} \qquad f^{-1} : \begin{cases} x = \sqrt{y} \\ y \geq 0 \end{cases}$$

$$f : \begin{cases} y = x^3 \\ x \in \mathbb{R} \end{cases} \qquad f^{-1} : \begin{cases} x = \sqrt[3]{y} \\ y \in \mathbb{R} \end{cases}$$

$$f : \begin{cases} y = a^x \\ x \in \mathbb{R} \end{cases} \qquad f^{-1} : \begin{cases} x = \log_a y \\ y > 0 \end{cases}$$

Si consideri anche il prossimo esempio: $f(x) = x + e^x$.

Essendo somma di due funzioni strettamente crescenti in tutto $\mathbb{R}$, $f(x)$ è strettamente crescente e quindi invertibile su tutto $\mathbb{R}$. Tuttavia, cercheremmo inutilmente di risolvere rispetto a x l'equazione $x + e^x = y$. In altre parole, f^{-1} esiste, ma non si sa scrivere esplicitamente.

Si noti che *tutte* le funzioni potenza: $y = x^\alpha$, per $x > 0$, sono invertibili e le funzioni inverse $x = y^{1/\alpha}$ sono ancora potenze (con esponente reciproco di quello della funzione data). Lo studente controlli questa affermazione sulla famiglia dei grafici di x^α.

Osserviamo anche che le potenze pari: x^{2n} $(n = 1, 2, \ldots)$ *non* sono invertibili su tutta la retta, ma solo sulla semiretta $x \geq 0$, mentre le potenze dispari: x^{2n+1} e le potenze $x^{p/q}$ con q dispari essendo monotone strettamente crescenti sono invertibili da $-\infty$ a $+\infty$.

4.3 Le funzioni trigonometriche inverse

Essendo periodiche, le funzioni trigonometriche non possono essere invertibili. Infatti, per esempio, l'equazione

$$\sin x = y$$

ha infinite soluzioni se $-1 \leq y \leq 1$ (l'uscita y corrisponde a infiniti ingressi) oppure non ha soluzioni reali se $|y| > 1$.

Per parlare di funzioni inverse di seno, coseno e tangente occorrerà restringersi a intervalli nei quali queste funzioni siano strettamente monotone e perciò invertibili.

Un intervallo nel quale la funzione seno è invertibile è $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$. La funzione inversa si chiama *arcoseno* (arcsin)

Dunque:

$$\begin{cases} y = \sin x \\ x \in \left[-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right] \end{cases} \qquad \text{equivale a} \qquad \begin{cases} x = \arcsin y \\ y \in [-1, 1] \end{cases}$$

Il grafico dell'arcoseno si ottiene dall'arco di sinusoide ristretto all'intervallo $\left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, per simmetria rispetto alla bisettrice $y = x$, come mostrato in figura 2.39.

Figura 2.39. Seno e arcoseno.

Analogamente, un intervallo di monotonia del coseno è $[0, \pi]$, nel quale dunque esso è invertibile. La funzione inversa si chiama *arcocoseno* (arcos)

Pertanto:

$$\begin{cases} y = \cos x \\ x \in [0, \pi] \end{cases} \qquad \text{equivale a} \qquad \begin{cases} x = \arccos y \\ y \in [-1, 1] \end{cases}$$

Il grafico dell'arcocoseno è illustrato in figura 2.40.

Figura 2.40. Coseno e arcocoseno.

Infine, osserviamo che nell'intervallo $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$ la tangente è strettamente monotona e quindi invertibile. La funzione inversa si chiama *arcotangente* (arctg), ed è definita in $(-\infty, +\infty)$.

Dunque:

$$\begin{cases} y = \operatorname{tg} x \\ x \in \left(-\dfrac{\pi}{2}, \dfrac{\pi}{2}\right) \end{cases} \qquad \text{equivale a} \qquad \begin{cases} x = \operatorname{arctg} y \\ y \in (-\infty, +\infty) \end{cases}$$

Il grafico dell'arcotangente è mostrato in figura 2.41.

Figura 2.41. Tangente e arcotangente.

Per mezzo delle funzioni trigonometriche inverse, possiamo esprimere le soluzioni di un'equazione o disequazione trigonometrica, quando questa coinvolge angoli non notevoli. Lo studente rifletta sui prossimi esempi:

- Le soluzioni di $\sin x = \dfrac{1}{4}$ sono $x = \arcsin \dfrac{1}{4} + 2k\pi; x = \pi - \arcsin \dfrac{1}{4} + 2k\pi$
- Le soluzioni di $\cos x < \dfrac{1}{5}$ sono $\arccos \dfrac{1}{5} + 2k\pi < x < 2\pi - \arccos \dfrac{1}{5} + 2k\pi$
- Le soluzioni di $\operatorname{tg} x \geq 3$ sono $\operatorname{arctg} 3 + k\pi \leq x < \dfrac{\pi}{2} + k\pi$

(Per convincersi delle affermazioni fatte, ragionare sulla circonferenza trigonometrica e sulla definizione delle funzioni trigonometriche inverse).

4.4 Le funzioni iperboliche inverse

Consideriamo la funzione $y = \mathrm{Sh}x$. È definita e strettamente crescente in tutto $\mathbb{R}$, perciò è invertibile. Inaspettatamente, la sua funzione inversa è facile da scrivere esplicitamente. Infatti, poniamo

$$x = \mathrm{Sh}y = \frac{e^y - e^{-y}}{2}$$

e risolviamo rispetto ad y. Moltiplicando ambo i membri per e^y l'equazione si può riscrivere come:

$$e^{2y} - 2xe^y - 1 = 0$$

che è un'equazione di secondo grado nell'incognita e^y. Ricaviamo:

$$e^y = x \pm \sqrt{x^2 + 1}$$

poiché $e^y > 0$, la soluzione col segno $-$ va scartata. Rimane dunque:

$$y = \log\left(x + \sqrt{x^2 + 1}\right)$$

Questa è l'espressione analitica della funzione inversa di $\mathrm{Sh}\,x$, che prende il nome di *settore seno iperbolico*, e si indica anche con $\mathrm{SettSh}x$. È definita per ogni x reale.

La funzione $y = \mathrm{Ch}\,x$ è definita in $\mathbb{R}$, strettamente crescente per $x \geq 0$, decrescente per $x \leq 0$. Perciò non è invertibile su tutto $\mathbb{R}$. La sua restrizione a $x \geq 0$ però lo è. Vogliamo determinare l'espressione analitica della funzione inversa di questa restrizione.

Con calcoli analoghi ai precedenti, scriviamo

$$x = \mathrm{Ch}y = \frac{e^y + e^{-y}}{2}$$

e risolviamo rispetto ad y:

$$e^{2y} - 2xe^y + 1 = 0$$

$$e^y = x \pm \sqrt{x^2 - 1}$$

Questa volta entrambi i numeri $x \pm \sqrt{x^2 - 1}$ sono positivi; ricordiamo però che stiamo ragionando solo per $y \geq 0$, che equivale a scegliere il segno $+$. Pertanto:

$$y = \log(x + \sqrt{x^2 - 1})$$

Questa è l'espressione analitica della funzione inversa di $\mathrm{Ch}\,x$, che prende il nome di *settore coseno iperbolico*, e si indica anche con $\mathrm{SettCh}x$. Si noti che è definita per $x \geq 1$. Si osservi anche che, mentre ad esempio l'equazione

$$\mathrm{Sh}x = 2 \text{ ha l'unica soluzione } x = \mathrm{SettSh}2 = \log(2 + \sqrt{5})$$

l'equazione

$$\mathrm{Ch}x = 3 \text{ ha le due soluzioni } x = \pm\mathrm{SettCh}3 = \pm\log\left(3 + 2\sqrt{2}\right)$$

Figura 2.42. a) Grafico di $y = \operatorname{Sh} x$ e $y = 2$; b) grafico di $y = \operatorname{Ch} x$ e $y = 3$.

Le funzioni iperboliche inverse saranno utilizzate nell'*integrazione delle funzioni irrazionali* (cap. 6).

Esercizi

1. Determinare l'insieme di definizione delle seguenti funzioni:

a) $f(x) = \sqrt{e^{2x} - e^x}$ b) $f(x) = \log(1 - x^3)$ c) $f(x) = \arcsin(x^2 - 2)$

2. Utilizzando il grafico di $f(x) = e^x$ e i metodi del paragrafo 3.7, disegnare i grafici di

$$y_1 = e^{|x|} \qquad y_2 = 3e^x \qquad y_3 = e^x - 1 \qquad y_4 = |e^x - 1|$$

Verificare i risultati al computer.

3. Utilizzando il grafico di $f(x) = \log x$ e i metodi del paragrafo 3.7, disegnare i grafici di

$$y_1 = \log|x| \qquad y_2 = 3\log x \qquad y_3 = \log(x-2) \qquad y_4 = |\log x| \qquad y_5 = \big|\log|x|\big|$$

Verificare i risultati al computer.

4. Una vibrazione elementare sia descritta dalla funzione

$$f(x) = A\cos\omega x \qquad A = \text{ ampiezza} \qquad \omega/2\pi = \text{ frequenza}$$

Tracciare il grafico di f per $A = 1, -1, 3, -3$ e $\omega = 1, 2, 3$.
Visualizzare al computer le variazioni del grafico al variare di A e di ω.

5. Sia $f(x) = \sin x$. Tracciare il grafico e studiare, aiutandosi col computer, come esso si modifichi moltiplicando f per x, $\frac{1}{x}$, e^x, e^{-x}.

6. Determinare dominio e immagine delle seguenti funzioni. Nel caso in cui la funzione sia invertibile, determinare la funzione inversa

a) $y_1 = x^3 + 2$ b) $y_2 = e^{-|x|}$ c) $y_3 = x|x|$

7. Sia $f(x) = \frac{1}{x}$. Calcolare la funzione composta $f^2 = f \circ f$.

8 Determinare l'espressione analitica della funzione composta $f \circ g$ dove

$$f(x) = \begin{cases} 1 & \text{se } x \geq 0 \\ -1 & \text{se } x < 0 \end{cases} \qquad g(x) = \sin x$$

Disegnare il grafico di $f \circ g$.

9 Disegnare al computer le funzioni seguenti

$$y_1 = [x^2] = \text{ parte intera di } x^2 \quad \text{e} \quad y_2 = (x^2) = \text{ mantissa di } x^2$$

10 Dire se le seguenti funzioni sono simmetriche (pari o dispari):

$$\text{a) } \frac{x}{1+x^2} \quad \text{b) } x \sin 2x \quad \text{c) } \cos(x^3) \quad \text{d) } 2^{x^3} \quad \text{e) } \frac{x^2+1}{x^4+2} \quad \text{f) } \frac{x^3+1}{x+4}$$

11 Stabilire se le seguenti funzioni sono periodiche, e in caso affermativo determinarne il periodo:

$$\text{a) } \sin 2x + \cos 3x \quad \text{b) } (\sin x)^2 \quad \text{c) } \cos(2x+1) \quad \text{d) } \sin(x^3) \quad \text{e) } 2^{\sin x} \quad \text{f) } \operatorname{tg}\frac{x}{2} + (\cos x)^2$$

12 Stabilire con ragionamenti elementari se le seguenti funzioni sono monotone sul loro insieme di definizione:

$$\text{a) } \frac{1}{x^2} \quad \text{b) } e^{x^3+1} \quad \text{c) } \sin 2x \quad \text{d) } \log(1+2^{-x}) \quad \text{e) } \frac{1}{1+x^2}$$

13 Tracciare il grafico qualitativo delle seguenti funzioni potenza:

$$\text{a) } x^{-2/7} \quad \text{b) } x^{4/5} \quad \text{c) } x^{6/5} \quad \text{d) } x^{\pi} \quad \text{e) } x^{1/\sqrt{2}}$$

14 Risolvere le seguenti disequazioni:

$$\text{i) } |x-4| \geq x+2 \qquad \text{ii) } \frac{(2x-1)(x+1)}{x} \geq 0 \qquad \text{iii) } \frac{x}{|x|}(1-x) \leq 1+x$$

15 Risolvere le seguenti disequazioni:

$$\text{i) } \sqrt{|x-1|} < 2-x \qquad \text{ii) } \sqrt[3]{1+x^3} > x-4 \qquad \text{iii) } \frac{\sqrt{2x+1}-3}{x^2-4} \leq 0$$

16 Risolvere le disequazioni:

$$\text{i) } \log_2(x-4) - \log_2(x-1) > 1 \qquad \text{ii) } (\log_2 x)^2 - 3\log_2 x - 5 \geq 0$$

17 Trovare le soluzioni dell'equazione:

$$\log_{\frac{1}{2}}(1+x) = \log_2(2-x)$$

18 Risolvere le seguenti disequazioni:

$$\text{i) } 2(\sin x)^2 + \sin x - 3 < 0 \qquad \text{ii) } (\cos x)^2 + 2\cos x - 1 \geq 0$$

19 Risolvere le seguenti disequazioni, eseguendo anche, ove necessario, un confronto grafico, e localizzando i valori che non si possono determinare esattamente (Esempio: soluzioni $\alpha \leq x \leq 5$, con $3 < \alpha < 4$).

$$\text{i) } e^{\sqrt{x}-x+1} > 1 \qquad \text{ii) } \frac{e^x}{1-x^2} < 1$$

20 Dire quali sono il dominio e l'immagine delle seguenti funzioni:

$$f(x) = 2\arcsin(x-1)\,; \quad g(x) = \frac{1}{2}\arccos(2x+1)$$

21 Con lo stesso metodo usato per invertire la funzione $\operatorname{Sh} x$, scrivere la funzione inversa di $\operatorname{Th} x$.

COMPLEMENTI

1 Dimostrare che la funzione $f(x) = x^2$ è strettamente crescente per $x > 0$. (*Suggerimento: moltiplicare ogni membro della disuguaglianza* $0 < x_1 < x_2$ *una volta per* x_1 *e una volta per* x_2, *e confrontare le disuguaglianze ottenute*).

Utilizzando un'idea simile a quella suggerita, provare per induzione su n che x^n è strettamente crescente in $(0, +\infty)$.

2 Utilizzare l'identità

$$f(x) = \frac{f(x) + f(-x)}{2} + \frac{f(x) - f(-x)}{2}$$

per dimostrare che ogni funzione $f : \mathbb{R} \to \mathbb{R}$ si può scrivere come somma di una funzione pari p e una dispari d. Se si applica questo procedimento a $f(x) = e^x$, chi sono $p(x)$ e $d(x)$?

3 Siano A e B due intervalli. Una funzione crescente (o decrescente) sia in A che in B, è crescente (o decrescente) in $A \cup B$? Rispondere ragionando sull'esempio $f(x) = 1/x$.

4 Siano f, g funzioni monotone per cui è definita $f \circ g$. Allora $f \circ g$ è monotona, e precisamente:

	se f è crescente...	se f è decrescente...
...e g è crescente	allora $f \circ g$ è...	allora $f \circ g$ è...
...e g è decrescente	allora $f \circ g$ è...	allora $f \circ g$ è...

5 Siano f, g funzioni simmetriche (pari o dispari) per cui è definita $f \circ g$. Allora $f \circ g$ è simmetrica, e precisamente:

	se f è pari...	se f è dispari...
...e g è pari	allora $f \circ g$...	allora $f \circ g$...
...e g è dispari	allora $f \circ g$ è...	allora $f \circ g$ è...

6 Siano f, g funzioni per cui è definita $f \circ g$. Chiediamoci se $f \circ g$ è periodica, nei seguenti casi:

	se f è periodica...	se f non è periodica...
...e g è periodica	allora $f \circ g$...	allora $f \circ g$...
...e g non è periodica	allora $f \circ g$...	allora $f \circ g$...

7 Sia $f : \mathbb{R} \to \mathbb{R}$ una funzione periodica di periodo T. Si dica se le seguenti funzioni sono periodiche, indicandone il periodo, oppure non lo sono, esibendo in tal caso un contresempio esplicito:

$$f_1(x) = f(\omega x) \ (\text{con } \omega > 0); \qquad f_2(x) = f(x + c) \ (\text{con } c \in \mathbb{R});$$
$$f_3(x) = |f(x)|; \qquad f_4(x) = f(|x|);$$
$$f_5(x) = f\left(x^2\right)$$

8 Siano $f_1, f_2 : \mathbb{R} \to \mathbb{R}$ funzioni periodiche di periodo T_1, T_2, rispettivamente. Ci chiediamo se la funzione $f_1 + f_2$ è periodica o no, e in caso affermativo qual è il suo periodo. Se vogliamo che risulti

$$(f_1 + f_2)(x + T) = (f_1 + f_2)(x) \ \forall x,$$

dovrà essere:

$$T = n_1 T_1 = n_2 T_2$$

per due opportuni interi positivi n_1, n_2. Ne segue che T dev'essere un multiplo comune di T_1 e T_2. In particolare:

i) condizione necessaria affinché $f_1 + f_2$ sia periodica è che il rapporto tra i due periodi T_1/T_2 sia un numero *razionale*;

ii) se questa condizione è soddisfatta, un numero T che soddisfa la condizione è il *minimo comune multiplo di* T_1, T_2. Questo numero è il periodo di f (oppure un multiplo del periodo di f, che in tal caso è più piccolo).

Usando di questi fatti:

a) Si determinino i periodi delle seguenti funzioni:

$$\sin 4x + 3\cos 6x; \ \ \operatorname{tg} 6x - 2\cos 8x; \ \ 5\sin\frac{x}{3} + \cos\frac{x}{4}$$

e si controlli il risultato previsto tracciando il grafico di queste funzioni col computer.

b) Si dimostri che la funzione

$$\sin x + \sin(\pi x)$$

non è periodica (pur essendo somma di due funzioni periodiche), e ci si renda conto di questo fatto tracciando il grafico di questa funzione col computer, su intervalli di varia ampiezza.

9 Dimostrare, in base alla definizione, le proprietà 3) delle funzioni iperboliche enunciate nel paragrafo 3.6.

3 Limiti e continuità

In questo corso ci occuperemo prevalentemente del *calcolo infinitesimale*, disciplina matematica che affonda le sue radici nella Grecia del III secolo a.C. (Euclide, Archimede), ha un grande sviluppo a partire dal Seicento, parallelamente al nascere della scienza moderna, in particolare ad opera di Newton e Leibniz, tra il 1670 e il 1710 circa; viene quindi sottoposta a revisione critica e fondata rigorosamente nell'Ottocento, prima da Cauchy, nel 1821[1], poi da Weierstrass e da vari altri matematici (Heine, Cantor, Méray,...) intorno al 1870. Le idee e le tecniche di calcolo proprie di questa disciplina fanno oggi parte del bagaglio essenziale con cui scienza e tecnologia si esprimono e procedono.

Il fondamento concettuale del calcolo infinitesimale sta nella nozione di *limite* (D'Alembert 1765, Cauchy 1821), che quindi può a buon diritto considerarsi una pietra miliare nella storia del pensiero scientifico. Noi introdurremo questo concetto gradualmente, prima nel caso discreto (par. 1) e poi in quello continuo, in cui storicamente è nato (par. 2).

Nel contesto discreto, il limite si può vedere come un'*operazione* che, a differenza delle operazioni algebriche elementari (somma, prodotto), viene eseguita non su una *coppia* di numeri ma su una *successione* di infiniti numeri. Per prima cosa introdurremo quindi il concetto di successione. Questo argomento, oltre a servirci immediatamente per introdurre il concetto di limite di funzione, sarà ripreso nel capitolo 5 parlando di *serie numeriche*, e nel capitolo 7 discutendo i *modelli dinamici discreti*.

1 SUCCESSIONI

1.1 Definizione di successione. Definizione di limite

Consideriamo l'insieme $\mathbb{N}$ degli interi non negativi ordinato secondo l'ordine naturale

$$\mathbb{N} : 0, 1, 2, 3, \ldots, n, \ldots$$

[1] Corso di Analisi per l'Ecole Polytechnique di Parigi.

Questo è l'esempio canonico di successione. Stabiliamo ora una legge che associa, a ogni elemento di $\mathbb{N}$ (o da un certo intero in poi) un numero (reale):

$$n \longmapsto a_n$$

Chiameremo *successione* una tale corrispondenza.

Una successione può dunque vedersi come una funzione

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

$$f : n \mapsto a_n$$

(o eventualmente, $f : \{n \in \mathbb{N} : n \geq n_0\} \to \mathbb{R}$, per un certo n_0 fissato). Il fatto che il dominio della funzione f sia l'insieme dei naturali, rende possibile visualizzare la successione enumerando i suoi valori, nell'ordine in cui essi si succedono al crescere di n:[2]

$$a_0, a_1, a_2, \ldots, a_n, \ldots$$

Esempio

1.1

$n \geq 0$	$n \longmapsto n^2$	$0, 1, 4, 9, 16, \ldots$
$n \geq 0$	$n \longmapsto (-1)^n$	$1, -1, 1, -1, 1, -1, \ldots$
$n \geq 1$	$n \longmapsto 2^{1/n}$	$2, \sqrt{2}, \sqrt[3]{2}, \sqrt[4]{2}, \ldots$
$n \geq 1$	$n \longmapsto \frac{1}{n}$	$1, \frac{1}{2}, \frac{1}{3}, \frac{1}{4}, \ldots$
$n \geq 2$	$n \longmapsto \frac{n+1}{n-1}$	$3, 2, \frac{5}{3}, \frac{6}{4}, \frac{7}{5}, \ldots$
$n \geq 0$	$n \longmapsto 4$	$4, 4, 4 \ldots$ (successione costante)

Possiamo rappresentare graficamente questa corrispondenza con i punti del piano cartesiano di coordinate (n, a_n) (fig.3.1).

Sottolineiamo che la successione è nota quando è nota la legge che, dato l'intero n, determina il numero a_n associato a quell'intero. Per indicare una successione useremo i simboli

$$n \longmapsto a_n \quad \text{oppure} \quad \{a_n\}$$

precisando l'insieme in cui varia l'indice n (tutto $\mathbb{N}$ o da un certo intero in poi).

Una successione $\{a_n\}$ si dirà

limitata inferiormente	se esiste un numero m tale che	$a_n \geq m \; \forall n$
limitata superiormente	se esiste un numero M tale che	$a_n \leq M \; \forall n$
limitata	se esistono due numeri m e M tali che	$m \leq a_n \leq M \; \forall n$

[2] I puntini di sospensione... scritti nella formula seguente dopo a_n sono fondamentali: significano che non stiamo considerando soltanto i primi n termini della successione (cioè un insieme finito di numeri), ma l'intera successione di infiniti termini (in cui n gioca il ruolo di indice muto).

Figura 3.1.

Per esempio, la successione $\{(-1)^n\}$ è limitata; $\{n^2\}$ è limitata solo inferiormente; la successione $\{(-2)^n\}$ non è limitata (né inferiormente, né superiormente).

L'operazione che vogliamo definire (il *limite*) consente di rispondere in forma rigorosa alla domanda: *come si comportano i numeri* $\{a_n\}$ *quando* n *diventa sempre più grande?*

Cominciamo con l'introdurre un modo di dire molto utile.

Definizione 3.1 Diciamo che una successione $\{a_n\}$ possiede (o acquista) *definitivamente* una certa proprietà se esiste un $N \in \mathbb{N}$ tale che a_n soddisfa quella proprietà per ogni intero $n \geq N$.

Esempio

1.2 La successione $\{n - 10\sqrt{n}\}$ è definitivamente positiva; la successione $\{\frac{1}{n}\}$ $(n = 1, 2, \ldots)$ è definitivamente minore di $10^{-10^{10}}$.

Successioni convergenti

Definizione 3.2 Una successione $\{a_n\}$ si dice *convergente* se esiste un numero $l \in \mathbb{R}$ con questa proprietà: qualunque sia $\varepsilon > 0$ risulta definitivamente

$$|a_n - l| < \varepsilon. \tag{1.1}$$

In altre parole: per ogni $\varepsilon > 0$ si può trovare un intero N (che naturalmente dipenderà in generale da questo ε) tale che

$$|a_n - l| < \varepsilon \quad \text{per ogni} \quad n \geq N$$

Se la successione $\{a_n\}$ è convergente, ad essa è associato perciò il numero l. Si osservi che tale numero è unico, poiché, se ve ne fossero due, l_1 e l_2, associati alla medesi-

ma successione, risulterebbe definitivamente (applicando la disuguaglianza triangolare (4.4), cap. 1)

$$|l_1 - l_2| = |l_1 - a_n + a_n - l_2| \leq |l_1 - a_n| + |a_n - l_2| < 2\varepsilon$$

ma tale disuguaglianza, potendo noi scegliere ε come vogliamo, può sussistere solo se $l_1 = l_2$.

Definizione 3.3 Il numero l che compare nella (1.1) si chiama *limite* della successione $\{a_n\}$, e si scrive della successione

$$\lim_{n\to+\infty} a_n = l \quad \text{oppure} \quad a_n \to l \text{ per } n \to +\infty$$

(si legge, rispettivamente: il limite, per n tendente all'infinito, di a_n è l, oppure: a_n tende a l per n tendente a infinito).

Si noti che la disuguaglianza (1.1) corrisponde, più esplicitamente, alle seguenti due:

$$l - \varepsilon < a_n < l + \varepsilon \tag{1.2}$$

Rappresentando graficamente i punti della successione

Figura 3.2.

la condizione di convergenza significa che, fissata una striscia orizzontale $[l-\varepsilon, l+\varepsilon]$ "comunque stretta", da un certo indice in poi i punti della successione non escono più da questa striscia (v. fig. 3.2). Da questa osservazione risulta chiaramente che: *ogni successione convergente è limitata.*

Esempi

1.3 Mostriamo che $\lim_{n\to\infty} \frac{n+1}{n-1} = 1$ (cosa che si può facilmente sospettare osservando l'andamento della successione). Delle due disuguaglianze

$$1 - \varepsilon < \frac{n+1}{n-1} < 1 + \varepsilon$$

quella di sinistra è sempre soddisfatta, mentre quella di destra è soddisfatta per

$$n > \frac{2+\varepsilon}{\varepsilon}$$

Fissato $\varepsilon > 0$, basterà scegliere $N = (2+\varepsilon)/\varepsilon$ (o uguale al primo intero $> (2+\varepsilon)/\varepsilon$) per soddisfare la condizione richiesta dalla definizione di limite.

1.4 Per mostrare che $2^{1/n} \to 1$ per $n \to \infty$, si studiano le disuguaglianze

$$1 - \varepsilon < 2^{1/n} < 1 + \varepsilon$$

Quella di sinistra è sempre soddisfatta; quella di destra, prendendo il logaritmo (in base 2) di ambo i membri, si scrive

$$\frac{1}{n} < \log_2(1 + \varepsilon)$$

ed è soddisfatta se $n > 1/\log_2(1 + \varepsilon)$. Si sceglie perciò $N = \ldots$

Non risultano convergenti invece le prime due successioni dell'esempio 1.1. Esse sono però molto diverse tra loro e conviene introdurre definizioni che ne mettano in risalto la differenza.

Successioni divergenti. Successioni irregolari

Definizione 3.4 Quando, al crescere di n, una successione supera definitivamente qualunque numero $M > 0$ fissato, diremo che *diverge a* $+\infty$; se invece scende al di sotto di $-M$, diremo che *diverge a* $-\infty$. (Il simbolo ∞ si legge "infinito").

Diremo nei due casi, rispettivamente, che $+\infty$ e $-\infty$ sono i limiti della successione e scriveremo, rispettivamente:

$$\lim_{n \to +\infty} a_n = +\infty \quad \text{oppure} \quad \lim_{\to +\infty} a_n = -\infty\,.$$

Questi simboli, $+\infty$ e $-\infty$, non sono numeri. Se rappresentiamo i numeri reali sulla retta euclidea, ogni numero corrisponde a un punto e ogni punto a un numero. Con i simboli $+\infty$ e $-\infty$ conveniamo di indicare due "punti", uno $(+\infty)$ sta alla destra di ogni punto di $\mathbb{R}$ e l'altro $(-\infty)$ alla sinistra; a questi due punti non corrisponde però alcun numero (in altre parole, non possiamo definire sui simboli $+\infty$ e $-\infty$ le operazioni di somma e prodotto con le proprietà indicate in R_1 e R_2, anche se, come vedremo, potremo fare "parzialmente" queste operazioni).

L'insieme dei numeri reali $\mathbb{R}$ con l'aggiunta dei due elementi $\{+\infty\}$ e $\{-\infty\}$ sarà indicato con $\mathbb{R}^*$

$$\mathbb{R}^* = \mathbb{R} \cup \{-\infty\} \cup \{+\infty\}$$

Possiamo rappresentare "visivamente" l'insieme $\mathbb{R}^*$ mettendo in corrispondenza biunivoca (fig. 3.3) i punti della retta con quelli di una semicirconferenza, proiettando questi ultimi dal centro C sulla retta $\mathbb{R}$:

Figura 3.3.

Ai punti A e B non corrisponde su $\mathbb{R}$ alcun punto; diremo che $\{-\infty\}$ è il corrispondente del punto A e $\{+\infty\}$ il corrispondente di B.

L'operazione di limite risulta completamente significativa se ambientata in $\mathbb{R}^*$ invece che in $\mathbb{R}$; cioè il limite di una successione può essere un numero reale l oppure $+\infty$ oppure $-\infty$; le successioni il cui limite è un numero reale si dicono *convergenti*, quelle il cui limite è $+\infty$ oppure $-\infty$ si dicono *divergenti*. La successione canonica $\{n\}$ degli interi naturali evidentemente diverge a $+\infty$; così pure la successione $\{2^n\}$.

Infine osserviamo che ci sono successioni che non ricadono in nessuna delle categorie precedenti, cioè non sono convergenti né divergenti; per esempio la successione $\{(-1)^n\}$ oppure $\{(-2)^n\}$ (si noti che la prima è limitata e la seconda no). Tali successioni si diranno *irregolari* o *indeterminate*. Per esse l'operazione di limite non è definita, ovvero *il loro limite non esiste*.

Insiemi non limitati

È comodo adottare la convenzione introdotta per i limiti anche per il sup e per l'inf, estendendo la definizione di queste quantità nel modo seguente:

Definizione 3.5 Se l'insieme $E \subseteq \mathbb{R}$ non è limitato superiormente (inferiormente) diremo che

$$\sup E = +\infty \quad (\inf E = -\infty).$$

In questo modo la proprietà R_4 dei numeri reali può essere enunciata così:

> ogni insieme $E \subseteq \mathbb{R}$ non vuoto è dotato di estremo superiore e inferiore; $\sup E$ ($\inf E$) è un numero se E è limitato superiormente (inferiormente), altrimenti è $+\infty$ ($-\infty$).

Infinitesimi e infiniti

Una successione a_n tendente a zero si dice *infinitesima*. Ad esempio, sono infinitesime le successioni $\{\frac{1}{n}\}$, $\{\frac{1}{n^2}\}, \ldots$

Il concetto di infinitesimo gioca un ruolo centrale ed è fondamentale anche per avere un'*immagine intuitiva* corretta ed efficace dei concetti del calcolo infinitesimale. Vedremo nel par. 2 che il concetto di infinitesimo nel continuo (cioè parlando di funzioni) sarà perfettamente analogo. L'idea chiave a cui prestare attenzione è la seguente:

> "infinitesimo" non è un "numero infinitamente piccolo" (concetto privo di senso, se non si vuole che denoti semplicemente il numero 0) ma una *quantità variabile* (successione o, come vedremo, funzione), che *diviene indefinitamente piccola*.

Analogamente, una successione a_n tendente a $\pm\infty$ si dice *infinita*. Ad esempio, $\{n^2\}$, $\{n!\}$ sono infiniti.

Talvolta è possibile precisare se una successione convergente si avvicina al suo limite *per eccesso* o *per difetto*. Questo concetto è precisato dalla prossima

Definizione 3.6 Si dice che la successione $\{a_n\}$ tende a $l \in \mathbb{R}$ *per eccesso* (*per difetto*), e si scrive

$$\lim_{n\to+\infty} a_n = l^+ \text{ oppure } a_n \to l^+ \text{ per } n \to +\infty$$

(rispettivamente:

$$\lim_{n\to+\infty} a_n = l^- \text{ oppure } a_n \to l^- \text{ per } n \to +\infty)$$

se per ogni $\varepsilon > 0$ si ha che

$$0 \le a_n - l < \varepsilon \text{ definitivamente}$$

(rispettivamente:

$$0 \le l - a_n < \varepsilon \text{ definitivamente}).$$

In sostanza, dire che $a_n \to l^+$ significa affermare che $a_n \to l$ e inoltre $a_n \ge l$ definitivamente; dunque a_n si avvicina ad l "da sopra", ossia approssima l per eccesso.

Si rifletta sui prossimi esempi:

Esempi

1.5
$$\lim_{n\to+\infty} \frac{1}{n} = 0^+;$$

1.6
$$\lim_{n\to+\infty} \frac{n}{n+1} = 1^-;$$

1.7
$$\lim_{n\to+\infty} \frac{(-1)^n}{n} = 0,$$

ma in questo caso non si può affermare né che $a_n \to 0^+$ né che $a_n \to 0^-$.

1.2 Successioni monotone

Definizione 3.7 Una successione $\{a_n\}$ si dirà:
monotona crescente se $a_n \le a_{n+1}$; strettamente crescente se $a_n < a_{n+1}$ $\forall n$;
monotona decrescente se $a_n \ge a_{n+1}$; strettamente decrescente se $a_n > a_{n+1}$ $\forall n$.

Per esempio, la successione $\{n^2\}$ è monotona strettamente crescente, la successione $\{\frac{1}{n}\}$ è monotona strettamente decrescente, la successione $\{(-1)^n\}$ non è monotona; ogni successione costante è monotona (crescente o decrescente, non strettamente).
Riguardo all'operazione di limite, queste successioni hanno una importanza particolare; infatti esse non sono mai irregolari, ma sono convergenti oppure divergenti a seconda che siano limitate oppure no. Il risultato fondamentale è il seguente:

Teorema 3.1 (di monotonia) *Sia $\{a_n\}$ una successione monotona crescente e superiormente limitata.*
Allora $\{a_n\}$ è convergente, e il suo limite è uguale a $\sup\{a_n : n \in \mathbb{N}\}$.

Analogamente, se $\{a_n\}$ una successione monotona decrescente e inferiormente limitata, allora $\{a_n\}$ è convergente, e il suo limite è uguale a $\inf\{a_n : n \in \mathbb{N}\}$.

Ricordiamo che il simbolo $\sup\{a_n : n \in \mathbb{N}\}$ denota l'estremo superiore dell'insieme dei valori a_n assunti dalla successione.

DIMOSTRAZIONE. Consideriamo l'insieme dei valori assunti dalla successione, $\{a_n : n \in \mathbb{N}\}$. Poiché la successione è limitata superiormente, questo insieme è limitato superiormente, quindi (per la proprietà dell'estremo superiore di cui gode $\mathbb{R}$) esiste il suo estremo superiore,

$$\Lambda = \sup\{a_n : n \in \mathbb{N}\}, \Lambda \in \mathbb{R}.$$

Proviamo ora che

$$\lim_{n\to\infty} a_n = \Lambda.$$

Occorre mostrare che per ogni $\varepsilon > 0$ si ha

$$\Lambda - \varepsilon < a_n < \Lambda + \varepsilon \text{ definitivamente.}$$

La seconda disuguaglianza è ovvia: per ogni n è $a_n \leq \Lambda$ (e quindi $a_n < \Lambda + \varepsilon$), perché Λ è l'estremo superiore degli a_n, quindi in particolare è un maggiorante dell'insieme dei valori a_n. Per provare la prima disuguaglianza, consideriamo il numero $\Lambda - \varepsilon$. Ricordiamo che per definizione di estremo superiore, Λ è il *minimo* dei maggioranti dell'insieme $\{a_n : n \in \mathbb{N}\}$. Perciò, essendo $\Lambda - \varepsilon < \Lambda$, certamente $\Lambda - \varepsilon$ non è un maggiorante dell'insieme $\{a_n : n \in \mathbb{N}\}$. Questo significa che esiste un n_0 per cui

$$a_{n_0} > \Lambda - \varepsilon.$$

D'altro canto la successione è monotona crescente, perciò per ogni $n \geq n_0$ risulta $a_n \geq a_{n_0}$. Abbiamo quindi provato che

$$a_n \geq a_{n_0} > \Lambda - \varepsilon$$

per ogni $n \geq n_0$, ossia definitivamente. Quindi $\lim_{n\to\infty} a_n = \Lambda$ (fig 3.4). ◆

Figura 3.4.

Per esprimere anche simbolicamente che il limite è il sup (o l'inf) di una successione crescente (o decrescente) si usa la notazione (di evidente significato):

$$a_n \uparrow l \quad \text{oppure} \quad a_n \downarrow l.$$

Questo in particolare implica che $a_n \to l^-$ (rispettivamente, l^+), ma contiene una ulteriore informazione: la monotonia della successione.

Questo teorema è una conseguenza dell'assioma di continuità R_4 dei numeri reali (v. cap. 1, par. 5) e pertanto vale se l'ambiente che consideriamo è $\mathbb{R}$. Ad esempio, non

è vero che una successione crescente e limitata di numeri razionali ammette sempre limite razionale, cioè in $\mathbb{Q}$.

Esempio

1.8 Sia $\{a_n\}$ la successione così definita:

$$a_0 = 0$$
$$a_1 = 0,1$$
$$\dots$$
$$a_2 = 0,1011$$
$$a_3 = 0,10110111$$
$$a_4 = 0,1011011101111$$

e così via. (Al passo n si aggiunge al numero decimale ottenuto al passo precedente una cifra zero seguita da n cifre uguali a 1).

La successione $\{a_n\}$ è evidentemente crescente, e superiormente limitata (ad esempio, $a_n \leq 1$). In $\mathbb{R}$, la successione converge al numero $\sup\{a_n : n \in \mathbb{N}\}$, che dopo la virgola presenta un allineamento decimale illimitato e non periodico di cifre (una cifra 1, una cifra 0, due cifre 1, una cifra 0, tre cifre 1, una cifra 0, e così via all'infinito). Quindi il limite della successione è un numero irrazionale. Quest'esempio mostra che nell'insieme $\mathbb{Q}$ il teorema di monotonia è falso.

Vedremo in seguito che il teorema di monotonia sarà utilizzato per dimostrare importanti proprietà delle funzioni continue. Questo teorema quindi costituisce una delle motivazioni per cui è utile studiare l'analisi matematica nell'ambiente dei numeri reali, anziché in quello dei numeri razionali.

Il teorema di monotonia si può completare con il prossimo enunciato, che considera successioni limitate o illimitate:

Corollario 3.2 *Sia $\{a_n\}$ una successione monotona crescente. Allora esiste*

$$\lim_{n\to\infty} a_n = \sup\{a_n : n \in \mathbb{N}\}.$$

Esplicitamente: se $\{a_n\}$ è superiormente limitata, allora converge (e il suo limite è uguale all'estremo superiore dei suoi valori, che in questo caso è un numero reale); se invece $\{a_n\}$ è superiormente illimitata, allora a_n tende a $+\infty$ (che in questo caso è pari all'estremo superiore dei suoi valori).

Si può anche dire, sinteticamente: una successione monotona, converge o diverge (non può essere irregolare).

Dimostrazione. Se $\{a_n\}$ è superiormente limitata, l'enunciato è contenuto nel teorema di monotonia. Se invece $\{a_n\}$ è superiormente illimitata, questo significa che per ogni $K > 0$ esiste un n_0 tale che $a_{n_0} > K$. D'altro canto la successione è crescente, perciò per ogni $n \geq n_0$ si ha $a_n \geq a_{n_0} > K$. Abbiamo quindi provato che per ogni $K > 0$ è $a_n > K$ definitivamente. Questo significa che $a_n \to +\infty$. ♦

Esempio

1.9 (*Successione geometrica*) Si consideri la progressione geometrica di ragione q, $\{q^n\}$ (cfr. cap. 1, paragrafo 2):

$$1, q, q^2, q^3, \ldots, q^n, \ldots$$

Se $q > 1$ la successione è monotona crescente, illimitata superiormente.

Se $q = 1$ la successione è costante. Se $0 < q < 1$, la successione è monotona decrescente; è facile mostrare che tende a zero. Se q è negativo la successione non è più monotona. Lo studente verifichi le seguenti affermazioni:

$$\lim_{n \to +\infty} q^n = \begin{cases} +\infty & \text{se} \quad q > 1 \\ 1 & \text{se} \quad q = 1 \\ 0 & \text{se} \quad |q| < 1 \\ \text{non esiste} & \text{se} \quad q \leq -1 \end{cases}$$

1.3 Il calcolo dei limiti

In questo paragrafo passeremo in rassegna i teoremi basilari sul calcolo dei limiti. Le dimostrazioni si basano sulla definizione di limite, sull'uso di *disuguaglianze*, e sull'uso di proprietà *definitivamente vere*. In particolare, questi teoremi illustrano la relazione tra l'operazione di limite e le strutture algebriche e d'ordine presenti in $\mathbb{R}$.

Cominciamo ad esaminare le proprietà dell'operazione di limite rispetto alle *operazioni algebriche*. Quando il limite esiste finito, si dimostra un risultato semplice e naturale: l'operazione di limite commuta con queste operazioni, cioè:

Teorema 3.3 (**algebra dei limiti**) *Se $a_n \to a$ e $b_n \to b$ allora*

$$\boxed{\begin{array}{ll} a_n \pm b_n \to a \pm b & \\ a_n b_n \to ab & \\ \dfrac{a_n}{b_n} \to \dfrac{a}{b} & (b_n, b \neq 0) \\ a_n^{b_n} \to a^b & (a_n, a > 0) \end{array}}$$

Dimostrazione. Proviamo ad esempio che:

$$a_n \to a, b_n \to b \Rightarrow a_n + b_n \to a + b.$$

Fissiamo $\varepsilon > 0$ e consideriamo:

$$\begin{aligned} |(a_n + b_n) - (a + b)| &= |(a_n - a) + (b_n - b)| \leq \\ &\leq |a_n - a| + |b_n - b| \end{aligned}$$

(per la disuguaglianza triangolare, cap.1, par. 4.3). Poiché per ipotesi $a_n \to a$ e $b_n \to b$, si ha che

$$|a_n - a| < \varepsilon \text{ e } |b_n - b| < \varepsilon \text{ definitivamente,}$$

perciò concludiamo che

$$|(a_n + b_n) - (a + b)| < 2\varepsilon \text{ definitivamente,}$$

quindi $a_n + b_n \to a + b$.

Proviamo ora:

$$a_n \to a, b_n \to b \Rightarrow a_n b_n \to ab$$

Fissiamo $\varepsilon > 0$. Usando ancora la disuguaglianza triangolare si ha:

$$\begin{aligned} |a_n b_n - ab| &= |a_n (b_n - b) + b (a_n - a)| \leq \\ &\leq |a_n (b_n - b)| + |b (a_n - a)| = |a_n| \, |b_n - b| + |b| \, |a_n - a| \\ &\leq |a_n| \, |b_n - b| + |b| \, |a_n - a| \, . \end{aligned}$$

Poiché $a_n \to a$, $|a_n - a| < \varepsilon$ definitivamente; inoltre $|a_n| < |a| + \varepsilon$ definitivamente; poiché $b_n \to b$, $|b_n - b| < \varepsilon$ definitivamente. Quindi:

$$|a_n b_n - ab| < (|a| + \varepsilon) \, \varepsilon + |b| \, \varepsilon < \varepsilon \cdot \text{cost.}$$

definitivamente. Per l'arbitrarietà di ε segue la tesi.

Tralasciamo le dimostrazioni degli altri due casi. ♦

L'operazione di limite mantiene inoltre l'*ordinamento* cioè:

Teorema 3.4 (di permanenza del segno, 1ª forma) *Se $a_n \to a$ e $a > 0$ ($a < 0$) allora $a_n > 0$ definitivamente ($a_n < 0$ definitivamente).*

Dimostrazione. Fissato $\varepsilon > 0$, per definizione di limite abbiamo che

$$|a_n - a| < \varepsilon \text{ definitivamente,}$$

che riscriviamo nella forma:

$$a - \varepsilon < a_n < a + \varepsilon \text{ definitivamente.}$$

Poiché $a > 0$, possiamo scegliere $\varepsilon > 0$ in modo che sia anche $a - \varepsilon > 0$, allora la disuguaglianza $a - \varepsilon < a_n$ mostra che $a_n > 0$ definitivamente. In modo analogo si dimostra il caso $a < 0$. ♦

Teorema 3.5 (di permanenza del segno, 2ª forma) *Se $a_n \to a \in \mathbb{R}$, e $a_n \geq 0$ definitivamente, allora risulta $a \geq 0$. Più in generale:*

$$\textit{se } a_n \to a, b_n \to b \textit{ e } a_n \geq b_n \textit{ definitivamente, allora } a \geq b.$$

Dimostrazione. Segue dal teorema precedente. Infatti, se per assurdo fosse $a < 0$, dal teorema precedente si avrebbe $a_n < 0$ definitivamente, il che è incompatibile con l'ipotesi che sia $a_n \geq 0$ definitivamente. Questo dimostra la prima affermazione del teorema. Applicando ora questa proprietà alla differenza $a_n - b_n$ si ottiene anche la seconda affermazione. Infatti:

$$a_n \geq b_n \text{ definitivamente} \Longrightarrow a_n - b_n \geq 0 \text{ definitivamente,}$$

e poiché $a_n - b_n \to a - b$ (teorema sull'algebra dei limiti), si conclude $a - b \geq 0$, ovvero $a \geq b$. ♦

L'ultima relazione significa che in una disuguaglianza tra due successioni si può passare al limite ad ambo i membri, mantenendo il segno $\leq$. Si noti che in generale, invece, nel passaggio al limite non si conserva il segno di disuguaglianza *stretta*: ad esempio, anche se gli a_n sono strettamente positivi, il loro limite a è positivo o *nullo*, come mostra il semplice esempio di $\frac{1}{n} \to 0$.

È utile anche la seguente proposizione:

Teorema 3.6 (del confronto) *Se $a_n \leq b_n \leq c_n$ definitivamente e*

$$a_n \to l, c_n \to l \in \mathbb{R},$$

allora anche $b_n \to l$.

Dimostrazione. Fissiamo $\varepsilon > 0$. Allora definitivamente si ha

$$l - \varepsilon < a_n < l + \varepsilon; \quad l - \varepsilon < c_n < l + \varepsilon$$

da cui segue (definitivamente)

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

e quindi, definitivamente,

$$l - \varepsilon < b_n < l + \varepsilon$$

Dunque $b_n \to l$.

Casi particolari di questo teorema che si usano frequentemente sono espressi dal prossimo corollario, molto utile quando si studia il prodotto tra una successione oscillante (ma limitata) e una che tende a zero:

Corollario 3.7 *1. Se $|b_n| \leq c_n$ definitivamente e $c_n \to 0$, allora anche $b_n \to 0$.*
2. Se $c_n \to 0$ e b_n è limitata (ma non necessariamente convergente), allora $c_n b_n \to 0$.
Detto a parole: il prodotto di una successione infinitesima e una limitata è infinitesimo.

Dimostrazione.
1. Sappiamo che definitivamente è $-c_n \leq b_n \leq c_n$; d'altro canto se $c_n \to 0$ anche $-c_n \to 0$, quindi per il teorema precedente (con $a_n = -c_n$ e $l = 0$) si ha che $b_n \to 0$.
2. Se b_n è limitata, ossia $|b_n| \leq K$ per un certo $K > 0$ e per ogni n, possiamo scrivere

$$|b_n c_n| \leq K\,|c_n|\,.$$

Poiché $c_n \to 0$, anche $K\,|c_n| \to 0$, e per il punto 1 si conclude che $b_n c_n \to 0$.

Esempi

1.10 Applicando la definizione si dimostri che

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

Infatti, se $\alpha > 0$, fissato $M > 0$ risulta $n^\alpha > M$ per $n > M^{1/\alpha}$; perciò $n^\alpha \to +\infty$; se $\alpha < 0$, scrivendo $n^\alpha = 1/n^{|\alpha|}$, si osserva che $1/n^{|\alpha|} < \varepsilon$ per $n^{|\alpha|} > 1/\varepsilon$ ossia per $n > 1/\varepsilon^{1/|\alpha|}$, cioè definitivamente.

1.11 Limiti che si presentano nella forma di rapporto di due espressioni, ognuna costituita dalla somma di potenze di n:

$$\frac{n^{5/2} - 3n + 7}{n^3 + \sqrt{n} - 3n^2}$$

Si mette in evidenza a numeratore come a denominatore la potenza maggiore

$$\ldots = \frac{n^{5/2}\left(1 - \frac{3}{n^{3/2}} + \frac{7}{n^{5/2}}\right)}{n^3\left(1 + \frac{1}{n^{5/2}} - \frac{3}{n}\right)} = \frac{1}{\sqrt{n}} \cdot \left(\frac{1 - \frac{3}{n^{3/2}} + \frac{7}{n^{5/2}}}{1 + \frac{1}{n^{5/2}} - \frac{3}{n}}\right).$$

Ora per il teorema sull'algebra dei limiti e sapendo che potenze negative di n tendono a zero possiamo affermare che:

$$1 - \frac{3}{n^{3/2}} + \frac{7}{n^{5/2}} \to 1; \quad 1 + \frac{1}{n^{5/2}} - \frac{3}{n} \to 1.$$

Pertanto la successione tra parentesi tende a 1; d'altro canto $\frac{1}{\sqrt{n}} \to 0$, perciò la successione di partenza tende a zero.

1.12 La successione $\frac{\sin n}{n}$ è il prodotto di $\frac{1}{n}$ (convergente) e $\sin n$ (irregolare); quindi il teorema sull'algebra dei limiti non è applicabile. Tuttavia è applicabile l'ultimo corollario: la successione $\frac{1}{n}$ è infinitesima mentre $|\sin n| \leq 1$, perciò $\frac{\sin n}{n} \to 0$.

Fin qui abbiamo visto teoremi che operano su coppie di successioni entrambe convergenti o comunque limitate. Consideriamo ora il caso in cui i limiti sono $+\infty$ o $-\infty$. Supponiamo, per esempio, che $a_n \to a$ e $b_n \to +\infty$; allora è facile (e intuitivo) vedere che $a_n + b_n \to +\infty$. Abbrevieremo questa scrittura così: $a + \infty = +\infty$. Ragionando in maniera analoga possiamo compendiare le regole per il limite della somma (o differenza) di due successioni delle quali una o entrambe sono divergenti con le scritture seguenti:

$$\boxed{\begin{aligned} a + \infty &= +\infty \\ a - \infty &= -\infty \\ +\infty + \infty &= +\infty \\ -\infty - \infty &= -\infty \end{aligned}}$$

Analogamente per il prodotto (o il rapporto) abbiamo le regole seguenti; il segno di ∞ va determinato con la usuale regola dei segni:

$$\boxed{\begin{aligned} a \cdot \infty &= \infty \qquad (a \neq 0) \\ \frac{a}{0} &= \infty \qquad (a \neq 0) \\ \frac{a}{\infty} &= 0 \end{aligned}}$$

Esplicitamente, la seconda regola scritta significa ad esempio quanto segue:

$$\text{se } a_n \to a > 0 \text{ e } b_n \to 0^+, \text{ allora } \frac{a_n}{b_n} \to +\infty.$$

In modo analogo si procede, in base alla regola dei segni, se $a < 0$ o $b_n \to 0^-$; è comunque necessario, per applicare questo teorema, sapere che b_n tende a zero per eccesso o per difetto.

Le precedenti regole prendono il nome di "aritmetizzazione parziale del simbolo di infinito".

DIMOSTRAZIONE. Proviamo ad esempio che:

$$a_n \to +\infty, b_n \to b \in \mathbb{R} \Rightarrow \frac{b_n}{a_n} \to 0.$$

Fissiamo $\varepsilon > 0$. Poiché $a_n \to +\infty$, definitivamente si ha

$$a_n > \frac{1}{\varepsilon}$$

Poiché $b_n \to b$, definitivamente si ha

$$|b_n| < |b| + \varepsilon$$

Ne segue che definitivamente

$$\left|\frac{b_n}{a_n}\right| < \varepsilon\,(|b| + \varepsilon) < \varepsilon \cdot \text{cost}.$$

Per l'arbitrarietà di ε, segue la tesi. ◆

Lo studente noterà che mancano le regole relative a quattro operazioni, e cioè

$$+\infty - \infty\,, \quad 0 \cdot \infty\,, \quad \frac{0}{0}\,, \quad \frac{\infty}{\infty}$$

Queste espressioni si chiamano *forme di indecisione*, poiché nessuna regola può essere stabilita a priori per determinare il risultato, come vedremo negli esempi sotto illustrati.

Le regole sopra elencate (e la mancanza di regole per le forme di indecisione) confermano la natura particolare dei "punti" $+\infty$ e $-\infty$, che non possono essere considerati "numeri" poiché non rispettano le proprietà R_1 e R_2 del capitolo 1 paragrafo 3.

Esempi

1.13 $\sqrt{n+1} - \sqrt{n-1}$: forma di indecisione del tipo $+\infty - \infty$.
Moltiplicando e dividendo per $\sqrt{n+1} + \sqrt{n-1}$ si trova

$$\ldots = \frac{\left(\sqrt{n+1} - \sqrt{n-1}\right)\left(\sqrt{n+1} + \sqrt{n-1}\right)}{\sqrt{n+1} + \sqrt{n-1}} = \frac{2}{\sqrt{n+1} + \sqrt{n-1}} \to 0$$

1.14 Limiti di successioni che si presentano nella forma $a_n^{b_n}$ si trattano più facilmente considerando la successione dei logaritmi (per fissare le idee, in base 10)

$$\log_{10} a_n^{b_n} = b_n \log_{10} a_n$$

Si mostra che (cfr. anche più avanti, al paragrafo 3) se questa successione converge a l, diverge a $+\infty$, $-\infty$, o è indeterminata, la successione $a_n^{b_n}$, rispettivamente, converge a 10^l, $+\infty$, 0, è indeterminata.

Si può così osservare che le espressioni

$$1^{\pm\infty} \qquad 0^0 \qquad (+\infty)^0$$

sono altrettante forme di indecisione, corrispondenti (passando al logaritmo) alla forma di indecisione $0 \cdot \infty$.

1.4 Il numero e

Introdurremo ora un numero molto importante per l'Analisi, che sarà definito come limite di una particolare successione. Cominciamo a dimostrare il seguente risultato:

TEOREMA 3.8 *La successione*

$$a_n = \left(1 + \frac{1}{n}\right)^n$$

è convergente.

Si osservi che questa successione presenta una forma di indecisione 1^∞.

DIMOSTRAZIONE. Si proverà che la successione è monotona crescente

$$a_{n+1} > a_n$$

e limitata:

$$2 \leq a_n < 4$$

e perciò è convergente, per il Teorema di monotonia (par. 1.2).

Per provare che a_n è monotona crescente, studiamo, per $n \geq 2$, il rapporto:

$$\begin{aligned}
\frac{a_n}{a_{n-1}} &= \frac{\left(1+\frac{1}{n}\right)^n}{\left(1+\frac{1}{n-1}\right)^{n-1}} = \frac{\left(\frac{n+1}{n}\right)^n}{\left(\frac{n}{n-1}\right)^{n-1}} = \\
&= \left(\frac{\frac{n+1}{n}}{\frac{n}{n-1}}\right)^n \cdot \frac{1}{\left(\frac{n}{n-1}\right)^{-1}} = \left(\frac{n^2-1}{n^2}\right)^n \cdot \frac{1}{\left(\frac{n-1}{n}\right)} = \\
&= \frac{\left(1-\frac{1}{n^2}\right)^n}{1-\frac{1}{n}} \geq \frac{1-n\cdot\frac{1}{n^2}}{1-\frac{1}{n}} = 1,
\end{aligned}$$

dove nell'ultima disuguaglianza scritta si è applicata la *disuguaglianza di Bernoulli* (v. cap.1, par. 7)

$$(1+x)^n \geq 1 + nx$$

con $x = -\frac{1}{n^2}$. Questo mostra che $\frac{a_n}{a_{n-1}} \geq 1$, ossia $a_n \geq a_{n-1}$, e la successione è monotona crescente. In particolare, essendo $a_1 = 2$, segue $a_n \geq 2\ \forall\, n \geq 1$.

Consideriamo ora la successione

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

Si noti che

$$b_n = a_n \cdot \left(1 + \frac{1}{n}\right),$$

perciò $b_n > a_n$. Con calcoli simili a quelli appena svolti (lasciamo i dettagli al lettore), si mostra che che $b_n \leq b_{n-1}$. Poiché $b_1 = 4$, risulta quindi

$$a_n < b_n \leq b_1 = 4, \forall\, n \geq 1,$$

e a_n è limitata. ◆

Il limite della successione a_n appena studiata è un numero (irrazionale) molto importante in matematica, per varie ragioni che vedremo; esso viene indicato con la lettera e (*numero di Nepero*) e la sua rappresentazione decimale inizia così:

$$2.7182818284\ldots$$

Dunque, *per definizione*,

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

Questo numero viene molto spesso usato come base dei logaritmi, i quali, quando si usa questa base, vengono detti *naturali* o *neperiani* (dal nome del matematico John Napier) e indicati semplicemente col simbolo log (oppure ln) senza indicazione della base.

Esempio

1.15 *Il numero di Nepero e un problema... finanziario.* Supponiamo di possedere un capitale (per esempio, 1 milione di euro) e di investirlo al tasso di interesse annuale t (cioè con una rendita di t milioni all'anno).

Se l'interesse viene pagato annualmente, dopo un anno il capitale posseduto sarà $1+t$. Se l'interesse viene calcolato mensilmente avremo:

– dopo il primo mese un capitale pari a $1+\frac{t}{12}$;

– dopo il secondo mese, pari a $1+\frac{t}{12}+\frac{t}{12}\left(1+\frac{t}{12}\right)=\left(1+\frac{t}{12}\right)^2$;

– dopo il terzo mese, pari a $\left(1+\frac{t}{12}\right)^2\frac{t}{12}+\left(1+\frac{t}{12}\right)^2=\left(1+\frac{t}{12}\right)^3$;

– alla fine dell'anno avremo un capitale pari a $\left(1+\frac{t}{12}\right)^{12}$.

Se l'interesse viene calcolato ogni n-esimo di anno, avremo alla fine un capitale pari a

$$\left(1+\frac{t}{n}\right)^n$$

Per $t=1$ (rendita del 100%!) otteniamo esattamente la successione che definisce e.

Far tendere n a ∞ significa calcolare l'interesse dopo frazioni di anno sempre più piccole fino ad arrivare a calcolarlo con continuità. Il capitale che si ottiene alla fine dell'anno in quest'ultimo caso è esattamente pari a e.

Una volta *definito* il numero e come

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

si può *dimostrare* che risulta anche:

$$e = \lim_{n\to+\infty}\left(1-\frac{1}{n}\right)^{-n}. \tag{1.3}$$

Inoltre, a partire dalle due relazioni precedenti, si può provare il seguente

Teorema 3.9 *Sia $\{a_n\}$ una qualsiasi successione divergente (a $+\infty$ o $-\infty$). Allora esiste*

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

(Una traccia per la dimostrazione della (1.3) e di questo teorema sarà fornita nei Complementi alla fine del par. 1). Quest'ultimo teorema si rivela estremamente utile nel calcolo di limiti che coinvolgono la forma di indeterminazione $[1^\infty]$.

Esempio

1.16 Calcoliamo

$$\lim_{n\to+\infty}\left(\frac{n}{3+n}\right)^{5n}.$$

Si tratta di una forma di indeterminazione $[1^\infty]$. Scriviamo:

$$\left(\frac{n}{3+n}\right)^{5n+1}=\frac{1}{\left(1+\frac{3}{n}\right)^{5n+1}}=\frac{1}{\left[\left(1+\frac{3}{n}\right)^{\frac{n}{3}}\right]^{\frac{3(5n+1)}{n}}}\equiv\frac{1}{\left[\left(1+\frac{1}{a_n}\right)^{a_n}\right]^{b_n}}$$

con $a_n=\frac{n}{3}$. Per il teorema precedente, la successione entro parentesi quadre tende ad e, mentre l'esponente

$$b_n=\frac{3\,(5n+1)}{n}\to 15$$

perciò il limte cercato è $1/e^{15}$.

Altre situazioni di questo tipo saranno illustrate negli esercizi alla fine del par. 1.

1.5 Confronti e stime asintotiche

Abbiamo visto che una successione che tende a 0 è un *infinitesimo*; una successione che diverge (a $+\infty$, a $-\infty$) si dice *infinito*. Quando due successioni sono entrambe infinitesimi o entrambe infiniti, è utile poter stabilire un confronto tra di esse, per capire quale delle due tenda "più rapidamente" a 0 o all'infinito.

Esempi di infiniti sono le successioni seguenti:

$$\{\log n\}\quad\{\sqrt{n}\}\quad\{n^2\}\quad\{2^n\}$$

Esempi di infinitesimi si ottengono dalle successioni precedenti considerando gli elementi reciproci.

Siano $\{a_n\}$ e $\{b_n\}$ due infiniti. Consideriamo il limite del rapporto a_n/b_n; si hanno quattro possibilità:

$$\lim_{n\to\infty}\frac{a_n}{b_n}=\begin{cases}0 & \text{i)}\\ l \text{ finito e } \neq 0 & \text{ii)}\\ \pm\infty & \text{iii)}\\ \text{inesistente} & \text{iv)}\end{cases}$$

Diciamo che: i) $\{a_n\}$ è un infinito di *ordine inferiore* a $\{b_n\}$

ii) $\{a_n\}$ e $\{b_n\}$ sono infiniti dello *stesso ordine*

iii) $\{a_n\}$ è un infinito di *ordine superiore* a $\{b_n\}$

iv) $\{a_n\}$ e $\{b_n\}$ *non sono confrontabili.*

Se $\{a_n\}$ e $\{b_n\}$ sono due infinitesimi (e b_n è definitivamente $\neq 0$), ancora si considera il limite del rapporto a_n/b_n e, in corrispondenza delle 4 possibilità sopra elencate, diremo che:

i) $\{a_n\}$ è un infinitesimo di *ordine superiore* a $\{b_n\}$

ii) $\{a_n\}$ e $\{b_n\}$ sono infinitesimi *dello stesso ordine*

iii) $\{a_n\}$ è un infinitesimo di *ordine inferiore* a $\{b_n\}$

iv) $\{a_n\}$ e $\{b_n\}$ *non sono confrontabili.*

Il caso $\frac{a_n}{b_n} \to 1$ è particolarmente importante: si usa dire, in tal caso, che le due successioni $\{a_n\}$ e $\{b_n\}$ sono *asintotiche* e, per indicare questa circostanza, si scrive

$$a_n \sim b_n$$

(si legge: a_n è *asintotico* a b_n).

Il simbolo di asintotico è molto utile nel calcolo dei limiti per le seguenti proprietà:

Proposizione 3.1

1. *Se $a_n \sim b_n$, le due successioni hanno lo stesso comportamento: convergono allo stesso limite, o divergono entrambe a $\pm\infty$, o entrambe non hanno limite.*

2. *Si possono scrivere catene di relazioni asintotiche, cioè:*

$$\text{se } a_n \sim b_n \sim \ldots \sim c_n, \text{ allora } a_n \sim c_n$$

3. *Un'espressione composta da prodotto o quoziente di più fattori può essere stimata fattore per fattore:*

$$\text{se } a_n \sim a'_n, b_n \sim b'_n, c_n \sim c'_n, \text{ allora } \frac{a_n b_n}{c_n} \sim \frac{a'_n b'_n}{c'_n}$$

Attenzione: lo stesso *non vale* per le somme o per l'esponenziale.

Dimostrazione.

1. Dimostriamo la prima affermazione: se $a_n \sim b_n$, le due successioni hanno lo stesso comportamento. Sia $a_n \to l \in \mathbb{R}$; poiché $b_n = \frac{b_n}{a_n} \cdot a_n$ e $\frac{b_n}{a_n} \to 1$ (per definizione di asintotico), allora per il teorema sul limite del prodotto

$$b_n \to 1 \cdot l = l.$$

Lo stesso teorema (nel caso di limite infinito) permette di concludere che se $a_n \to \pm\infty$ e $b_n \sim a_n$, anche $b_n \to \pm\infty$. Osserviamo ora che la relazione di asintotico è simmetrica,

quindi quanto appena provato mostra anche che se b_n converge (diverge), anche a_n converge (diverge). Ne concludiamo che se a_n è irregolare, anche b_n è irregolare, perché se per assurdo non lo fosse, per quanto appena affermato anche a_n sarebbe convergente o divergente.
2. Proviamo la transitività della relazione di asintotico:

$$\text{se } a_n \sim b_n \text{ e } b_n \sim c_n \text{ allora } a_n \sim c_n.$$

Le nostre ipotesi significano che

$$\frac{a_n}{b_n} \to 1 \text{ e } \frac{b_n}{c_n} \to 1.$$

Allora per il teorema sul limite del prodotto,

$$\frac{a_n}{c_n} = \frac{a_n}{b_n} \cdot \frac{b_n}{c_n} \to 1.$$

Analogamente si prova la terza proprietà.

Un tipico modo per mostrare che $a_n \sim b_n$ consiste nello scrivere $a_n = b_n c_n$ con $c_n \to 1$. Ad esempio:

$$2n^2 + 3n + 1 = 2n^2 \left(1 + \frac{3}{n} + \frac{1}{2n^2}\right) \sim 2n^2$$

perché $\left(1 + \frac{3}{n} + \frac{1}{2n^2}\right) \to 1$.

Osservazione Il fatto che la relazione di asintotico soddisfi le 3 proprietà:

1. Riflessiva: $a_n \sim a_n$;
2. Simmetrica: se $a_n \sim b_n$ allora $b_n \sim a_n$;
3. Transitiva: se $a_n \sim b_n$ e $b_n \sim c_n$ allora $a_n \sim c_n$

si esprime dicendo che "asintotico" è una *relazione di equivalenza*. (Si noti che le prime due proprietà sono immediate, mentre la terza è stata dimostrata sopra). Più in generale, in matematica si chiama relazione d'equivalenza una relazione che soddisfa i 3 assiomi ora enunciati.

Mostriamo ora il seguente:

Teorema 3.10 (gerarchia degli infiniti)

$$\lim_{n\to+\infty} \frac{\log_a n}{n^\alpha} = 0$$

$$\lim_{n\to+\infty} \frac{n^\alpha}{a^n} = 0$$

per ogni $a > 1, \alpha > 0$.

Questi limiti descrivono la "velocità" con cui i logaritmi (con base > 1), le potenze, gli esponenziali (con base > 1) vanno all'∞: i logaritmi vanno più lentamente di qualsiasi potenza, le potenze più lentamente di qualsiasi esponenziale a base > 1.

DIMOSTRAZIONE. Per dimostrare la prima relazione, iniziamo a stabilire un'utile disuguaglianza tra un qualsiasi numero reale positivo e il suo logaritmo. Per $x \in \mathbb{R}, x > 0$, sia k la parte intera di x (v. cap. 2, par. 3.6), ossia l'intero $k \geq 0$ per cui è $k \leq x < k+1$. Si ha:

$$2^x \geq 2^k = (1+1)^k \geq 1 + k > x$$

dove: la prima disuguaglianza segue dalla monotonia della funzione esponenziale, la seconda dallo sviluppo del binomio di Newton (o dalla disuguaglianza di Bernoulli). Passando ai logaritmi in base a, otteniamo

$$\boxed{\log_a x < x \log_a 2}$$

per ogni $x > 0$. Applichiamo ora questa disuguaglianza al numero $x = n^{\alpha/2}$ e abbiamo:

$$\frac{\alpha}{2} \log_a n < n^{\alpha/2} \log_a 2$$

e quindi

$$\frac{\log_a n}{n^{\alpha/2}} < \frac{2}{\alpha} \log_a 2$$

$$\frac{\log_a n}{n^\alpha} = \frac{\log_a n}{n^{\alpha/2}} \cdot \frac{1}{n^{\alpha/2}} \leq \frac{2}{\alpha} \log_a 2 \cdot \frac{1}{n^{\alpha/2}}.$$

Per il teorema del confronto, $\frac{\log_a n}{n^\alpha} \to 0$, e la prima relazione è dimostrata.

Applichiamo ora questo risultato sostituendo all'intero n l'intero 2^n. Avremo:

$$0 = \lim_{n\to\infty} \frac{\log_a (2^n)}{(2^n)^\alpha} = \lim_{n\to\infty} \frac{n \log_a 2}{(2^\alpha)^n}.$$

Se ora $a > 1$ è fissato, scegliendo $\alpha > 0$ in modo che sia $2^\alpha = a$ otteniamo che $\frac{n}{a^n} \to 0$, che è la seconda relazione nel caso particolare in cui n è elevato ad esponente 1. Il caso generale segue dall'identità:

$$\frac{n^\alpha}{a^n} = \left(\frac{n}{a^{n/\alpha}}\right)^\alpha = \left(\frac{n}{(a^\alpha)^n}\right)^\alpha.$$

Infatti per il risultato precedente $\frac{n}{(a^\alpha)^n} \to 0$ (la base a^α è ancora un numero > 1), quindi anche $\left(\frac{n}{(a^\alpha)^n}\right)^\alpha \to 0$. ♦

Altri confronti tra infiniti si possono risolvere grazie al seguente criterio:

TEOREMA 3.11 (CRITERIO DEL RAPPORTO) *Sia a_n una successione positiva (cioè $a_n > 0$ per ogni n). Se esiste*

$$\boxed{\lim_{n\to\infty} \frac{a_{n+1}}{a_n} = l}$$

e $l < 1$, allora $a_n \to 0$; se $l > 1$ (ed eventualmente $l = +\infty$), allora $a_n \to +\infty$.

Il teorema precedente riconduce lo studio del carattere di una successione positiva, a_n, al calcolo del limite di un'altra successione, la successione (dei rapporti), $\frac{a_{n+1}}{a_n}$. In certi casi quest'ultima è più semplice da studiare di quella di partenza, come vedremo negli esempi. Si osservi che nel caso $l = 1$ il teorema non permette di concludere nulla.

DIMOSTRAZIONE. Supponiamo che

$$\frac{a_{n+1}}{a_n} \to l < 1.$$

Fissato $\varepsilon > 0$, definitivamente, ossia per ogni $n \geq n_0$ (con n_0 opportuno) si ha $\frac{a_{n+1}}{a_n} < l + \varepsilon$. Scegliamo ε abbastanza piccolo perché si abbia $l + \varepsilon < 1$. Possiamo scrivere la catena di disuguaglianze:

$$\begin{aligned} &a_{n_0+1} < (l+\varepsilon)\, a_{n_0}; \\ &a_{n_0+2} < (l+\varepsilon)\, a_{n_0+1} < (l+\varepsilon)^2\, a_{n_0} \\ &\quad \cdots \\ &a_{n_0+k} < (l+\varepsilon)^k\, a_{n_0}. \end{aligned}$$

Poiché $(l+\varepsilon) < 1$, si ha $(l+\varepsilon)^k \to 0$ per $k \to \infty$. D'altro canto a_{n_0} è fissato; dunque per k abbastanza grande il secondo membro (e quindi il primo) è piccolo quanto si vuole, il che dimostra che $a_n \to 0$.

Supponiamo ora che $\frac{a_{n+1}}{a_n} \to l > 1$. Fissato $\varepsilon > 0$, definitivamente si ha $\frac{a_{n+1}}{a_n} > l - \varepsilon$. Scegliamo ε abbastanza piccolo perché si abbia $l - \varepsilon > 1$. Analogamente a prima, possiamo ottenere la disuguaglianza:

$$a_{n_0+k} > (l-\varepsilon)^k\, a_{n_0}$$

per un certo n_0 fissato e qualsiasi k. Poiché $(l-\varepsilon) > 1$ e quindi $(l-\varepsilon)^k \to +\infty$, si conclude che $a_n \to +\infty$.

Esempi

1.17 Dimostriamo che

$$\lim_{n\to\infty} \frac{b^n}{n!} = 0 \text{ per ogni } b > 0.$$

Applichiamo il criterio del rapporto alla successione $a_n = \frac{b^n}{n!}$. Si ha:

$$\frac{a_{n+1}}{a_n} = \frac{b^{n+1}}{(n+1)!} \cdot \frac{n!}{b^n} = \frac{b}{n+1} \to 0.$$

Per il criterio del rapporto allora, $\frac{b^n}{n!} \to 0$. Abbiamo quindi ottenuto un nuovo caso nella gerarchia degli infiniti.

1.18 Si provi a calcolare

$$\lim_{n\to\infty} \frac{\log n}{n}$$

col criterio del rapporto, osservando che il metodo fallisce.

Ecco alcuni esempi di come si applicano tutte le osservazioni precedenti per risolvere alcune forme di indeterminazione. Quando avremo studiato un certo numero di *limiti notevoli* (par. 3.2) potremo risolvere mediante stime asintotiche situazioni più complesse di queste.

Esempi

1.19

$$\lim_{n\to+\infty} \frac{2n^3+4n+1}{5(n+1)^3} = \left[\frac{\infty}{\infty}\right]$$

Considerando solo le potenze di grado massimo a numeratore e denominatore, possiamo scrivere:

$$\frac{2n^3+4n+1}{5\,(n+1)^3} \sim \frac{2n^3}{5n^3} = \frac{2}{5}$$

e pertanto la successione tende a $2/5$.

1.20
$$\lim_{n\to+\infty} \frac{2^n+n}{2^{n+1}} = \left[\frac{\infty}{\infty}\right].$$

2^n è un infinito di ordine superiore rispetto ad n; possiamo scrivere quindi:

$$2^n + n = 2^n\left(1+\frac{n}{2^n}\right) \sim 2^n$$

perché $\frac{n}{2^n} \to 0$, e quindi $\left(1+\frac{n}{2^n}\right) \to 1$. Pertanto

$$\frac{2^n+n}{2^{n+1}} \sim \frac{2^n}{2^{n+1}} = \frac{1}{2}$$

e il limite è $\frac{1}{2}$.

1.21
$$\lim_{n\to+\infty} \sqrt[n]{n} = \left[\infty^0\right]$$

Scriviamo

$$\sqrt[n]{n} = n^{1/n} = e^{\log n^{1/n}} = e^{\frac{\log n}{n}}.$$

Usando il confronto di infiniti, $\frac{\log n}{n} \to 0$, e quindi si deduce

$$\lim_{n\to+\infty} \sqrt[n]{n} = e^0 = 1.$$

1.22
$$\lim_{n\to\infty} \frac{n!}{n^n} = \left[\frac{\infty}{\infty}\right].$$

Applicando il criterio del rapporto, consideriamo:

$$\frac{a_{n+1}}{a_n} = \frac{(n+1)!}{(n+1)^{n+1}} \cdot \frac{n^n}{n!} = \frac{(n+1)\,n^n}{(n+1)^n \cdot (n+1)} = \left(\frac{n}{n+1}\right)^n = \frac{1}{\left(1+\frac{1}{n}\right)^n} \to \frac{1}{e} < 1,$$

dove nell'ultimo passaggio si è usata la definizione di e; perciò $a_n \to 0$. Abbiamo quindi provato che $n!$ è un infinito di ordine inferiore rispetto a n^n.

Esercizi

1. Dare esempi di "infiniti" di ordine inferiore a $\{\log n\}$ e di ordine superiore a $\{2^n\}$.

2. Provare che
$$\log(n+1) \sim \log n$$

3. Calcolare
$$\lim_{n\to+\infty} \frac{a_{n+1}}{a_n}$$
per le seguenti successioni:
$$a_n = n^n \qquad a_n = n! \qquad a_n = 2^n \qquad a_n = n^2$$

4. Dare una stima asintotica delle seguenti successioni, mediante una successione "più semplice", e calcolare quindi il limite:
$$\frac{n^3+2n^2+\sin n}{n+\log n} \qquad \frac{n^2\log n + n}{\log 3n} \qquad \frac{n!-(n-1)!}{n+(n-2)!}$$

5 Lo studente, utilizzando una normale calcolatrice tascabile o un PC, calcoli $\left(1+\frac{1}{n}\right)^n$. Troverà, per esempio

n	$(1+1/n)^n$
10^3	2,7169239
10^4	2,7181459
$\vdots$	$\vdots$
10^{11}	1
10^{12}	1

Cerchi di spiegarne la ragione.

6 Un altro esempio: il limite

$$\lim_{n\to+\infty}(\sqrt{n^2+1}-n)n$$

si presenta sotto la forma di indecisione $\infty-\infty$. Moltiplicando e dividendo l'espressione per $\sqrt{n^2+1}+n$ tale limite prende la forma

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

Se si usa una calcolatrice per il calcolo del limite si ottiene (per esempio)

n	$(\sqrt{n^2+1}-n)n$	$n/(\sqrt{n^2+1}+n)$
10^2	0,4999875	0,4999875
10^3	0,5	0,499999
$\vdots$	$\vdots$	$\vdots$
10^7	0	0,5
10^8	0	0,5

Spiegare la ragione del diverso risultato.

7 Calcolare i limiti, per $n\to+\infty$, delle successioni seguenti:

$$\sqrt{n+1}-\sqrt{n} \qquad \sqrt{n^2+n}-n \qquad (\sqrt{n^4+1}-n^2)n$$

$$\log(n+1)-\log n \qquad \left(\frac{n}{n+1}\right)^n \qquad \left(1+\frac{1}{n^2}\right)^n \qquad \left(1+\frac{1}{n^7}\right)^{n^9}$$

COMPLEMENTI

1 Dimostrare il teorema sul limite del quoziente: se $a_n\to a, b_n\to b$, allora $\frac{a_n}{b_n}\to\frac{a}{b}$ purché $b_n, b\neq 0$.

Suggerimento: per maggiorare $\left|\frac{a_n}{b_n}-\frac{a}{b}\right|$ fare denominatore comune; ora il numeratore si maggiora in modo simile a quello visto nella dimostrazione del teorema sul limite del prodotto; per il denominatore occorre invece dimostrare che, ad esempio, $|b_n|\geq|b|/2$ definitivamente, e poi...

2 Dimostrare le relazioni che abbiamo chiamato "aritmetizzazione parziale di infinito", nei casi che non sono stati dimostrati.

3 Dimostrare che i valori assunti dalla successione $\sin n$ sono tutti diversi tra loro, ossia che $n, m \in \mathbb{N}, n \neq m$ implica $\sin n \neq \sin m$.

4 Provare che

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

Suggerimento: $\left(1-\frac{1}{n}\right)^{-n} = \left(\frac{n}{n-1}\right)^{n} = \ldots$; ricondursi al limite che definisce e.

5 Provare il Teorema 3.9 enunciato nel par. 1.4, ad esempio nel caso $a_n \to +\infty$.
Suggerimento: detta $[a_n]$ la parte intera di a_n, provare anzitutto le disuguaglianze:

$$\left(1+\frac{1}{[a_n]+1}\right)^{[a_n]} \leq \left(1+\frac{1}{a_n}\right)^{a_n} \leq \left(1+\frac{1}{[a_n]}\right)^{[a_n]+1}.$$

Quindi sfruttando opportunamente il fatto che $[a_n]$ è intero, ricondursi al limite già noto di $\left(1+\frac{1}{n}\right)^{n}$.

6 Dimostrare la proprietà 3 del simbolo di $\sim$ enunciata nel paragrafo 1.5 (Proposizione 3.1), usando la definizione di "asintotico" e di limite.

7 Si faccia un esempio di due successioni a_n, b_n tendenti a $+\infty$, per cui si ha:

$$a_n \sim b_n \text{ ma } e^{a_n} \textit{ non asintotico } \text{a } e^{b_n}.$$

Dunque il simbolo di asintotico non si può usare con gli esponenziali come si userebbe nei prodotti o quozienti.
Suggerimento: scegliere come a_n la somma di due infiniti di tipo diverso.

2 LIMITI DI FUNZIONI, CONTINUITÀ, ASINTOTI

L'operazione di limite si può estendere dalle successioni alle funzioni. Potremo così precisare il comportamento di una funzione quando la variabile indipendente si muove vicino a un determinato punto oppure diventa molto grande (in valore assoluto).

In questo paragrafo introdurremo i concetti fondamentali riguardanti i limiti e la continuità di funzioni; nel prossimo paragrafo 3 svilupperemo i teoremi e gli strumenti che permetteranno il calcolo effettivo dei limiti; infine, nel paragrafo 4 approfondiremo lo studio delle proprietà delle funzioni continue, incontrando alcuni dei più significativi teoremi dell'analisi delle funzioni di una variabile. Nei capitoli successivi, mediante l'operazione di limite introdurremo i concetti di derivata, di differenziale e di integrale per una funzione reale di variabile reale.

Consideriamo come caso tipico un intervallo I, un punto $c \in I$ e una funzione f a valori reali, definita in I, salvo al più nel punto c. L'intervallo I può essere limitato o illimitato, chiuso o aperto; il punto c può essere interno all'intervallo oppure uno dei suoi estremi (eventualmente $+\infty$ o $-\infty$).

Prendiamo ora una qualunque successione di punti x_n $(n = 1, 2, \ldots)$, nell'intervallo I e diversi da c, che tenda a c, per $n \to +\infty$.

In corrispondenza alla successione di ingressi x_n, consideriamo la successione delle uscite $f(x_n)$.

Se, qualunque sia la successione scelta, si ha che $f(x_n)$ tende al limite ℓ (finito o infinito) si dice che il *limite di $f(x)$ per x che tende a c è ℓ* e si scrive

$$\lim_{x \to c} f(x) = \ell \text{ oppure } f(x) \to \ell \text{ per } x \to c$$

(che si leggono: "il limite per x tendente a c di $f(x)$ è uguale a ℓ" oppure "$f(x)$ tende a ℓ per x tendente a c").

Sinteticamente:

Definizione 3.8 (successionale di limite) Si dice che

$$\lim_{x \to c} f(x) = \ell$$

(dove $c, \ell \in \mathbb{R}^*$) se per ogni successione $\{x_n\}$ di punti di I diversi da c, tale che $x_n \to c$, si ha che $f(x_n) \to c$ per $n \to \infty$.

Una funzione che per $x \to c$ tende a 0 si dice *infinitesima* per $x \to c$; analogamente una funzione che tende a $\pm\infty$ si dice *infinita*.

La definizione appena data di *limite di funzione* si dice definizione *successionale* di limite, in quanto riconduce il concetto di limite di funzione a quello di limite di successione. Non è l'unica definizione possibile, come vedremo, e a prima vista può sembrare una definizione poco operativa. Tuttavia, come dimostreranno sia il successivo sviluppo della teoria, sia gli esempi che faremo di calcolo di limiti, il fatto di avere già sviluppato le basi del calcolo dei limiti per le successioni, renderà molto vantaggiosa questa definizione, proprio dal punto di vista operativo. Ad esempio, possiamo subito affermare che vale il:

Teorema 3.12 (di unicità del limite) *Se esiste $\lim_{x \to c} f(x) = \ell$, tale limite ℓ è unico.*

Infatti, se esistessero due limiti ℓ_1, ℓ_2 diversi tra loro, presa una qualsiasi successione $x_n \to c$ si avrebbe

$$f(x_n) \to \ell_1 \text{ e } f(x_n) \to \ell_2$$

dunque la *successione* $f(x_n)$ avrebbe due limiti distinti, assurdo.

Osserveremo ora un po' più da vicino le varie situazioni che si possono verificare nella definizione di limite di funzione. È subito chiaro che la casistica sarà più ricca e complessa rispetto al caso delle successioni, perché mentre per una successione a_n il limite si calcola necessariamente per $n \to +\infty$, per una funzione $f(x)$ il limite si può calcolare per $x \to c$ dove c può essere qualsiasi elemento di $\mathbb{R}^*$.

Anzitutto chiariamo la terminologia: nella scrittura

$$\lim_{x \to c} f(x) = l$$

parleremo di

$$\text{limite} \begin{cases} \text{finito} \\ \text{infinito} \end{cases} \text{se} \begin{cases} l \in \mathbb{R} \\ l = +\infty, l = -\infty \end{cases}$$

e parleremo di

$$\text{limite} \begin{cases} \text{al finito} \\ \text{all'infinito} \end{cases} \text{se} \begin{cases} c \in \mathbb{R} \\ c = +\infty, l = -\infty \end{cases}$$

I prossimi esempi hanno lo scopo sia di mostrare come si utilizza concretamente la definizione successionale di limite, sia di illustrare le varie situazioni tipiche che si possono realizzare.

Limite finito all'infinito

Esempio

2.1 Dimostriamo che:

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

Applicando la definizione, si tratta di provare che per ogni successione $\{x_n\}$ tale che $x_n \to -\infty$, si ha

$$\lim_{n \to +\infty} e^{x_n} = 0.$$

A sua volta, per definizione di limite di successione, questo significa provare che fissato $\varepsilon > 0$ qualsiasi, risulti

$$|e^{x_n}| < \varepsilon \text{ definitivamente,}$$

ossia (essendo l'esponenziale sempre positivo),

$$e^{x_n} < \varepsilon \text{ definitivamente.}$$

L'ultima disuguaglianza è equivalente a:

$$x_n < \log \varepsilon.$$

Se $\varepsilon > 0$ è un numero piccolo (< 1), $\log \varepsilon$ sarà un numero negativo (e grande in valore assoluto); poniamo $K = -\log \varepsilon > 0$. Dobbiamo provare che, fissato questo numero $K > 0$, risulta $x_n < -K$ definitivamente; ma questo è proprio ciò che vale per ipotesi, perché $x_n \to -\infty$. Questo conclude la dimostrazione.

Asintoto orizzontale

Si dice che f ha un *asintoto orizzontale* di equazione $y = \ell$ ($\ell \in \mathbb{R}$) per $x \to +\infty$ oppure per $x \to -\infty$ se $\lim_{x \to +\infty} f(x) = \ell$ oppure $\lim_{x \to -\infty} f(x) = \ell$, rispettivamente.

Ogni situazione di limite finito all'infinito, quindi, corrisponde graficamente alla presenza di un asintoto orizzontale, ossia di una retta orizzontale a cui il grafico della funzione si avvicina sempre più (come nell'esempio precedente).

Limite per eccesso e per difetto

Come abbiamo visto per le successioni (par. 1.1), quando una funzione ha limite finito ℓ, talvolta è possibile precisare se questo limite viene raggiunto *per eccesso* (ℓ^+) o *per difetto* (ℓ^-). Graficamente, questo significa che il grafico della funzione si avvicina alla quota $y = \ell$ dall'alto o dal basso. La definizione precisa è la seguente:

Definizione 3.9 (Limite per eccesso o per difetto) Se $\ell \in \mathbb{R}$ e $c \in \mathbb{R}^*$, si dice che

$$\lim_{x \to c} f(x) = \ell^+ \text{ (rispettivamente, } \ell^-),$$

e in tal caso si dice che $f(x)$ tende a ℓ per eccesso (per difetto) per x tendente a c, se per ogni successione $\{x_n\}$ di punti di I diversi da c, tale che $x_n \to c$, si ha che $f(x_n) \to \ell^+$ (rispettivamente ℓ^-) per $n \to \infty$. Ricordiamo che, come visto nel paragrafo 1.1, affermare che $f(x_n) \to \ell^+$ significa che $f(x_n) \to \ell$ e inoltre $f(x_n) \geq \ell$ definitivamente.

Ad esempio,

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

Un errore comune consiste nel pensare che il simbolo 0^+ denoti un (misterioso) numero diverso da zero e "poco più grande di 0". Sbagliato: il limite della funzione è il "solito" numero zero che conosciamo; semplicemente, la scrittura 0^+ aggiunge una informazione: la funzione tende a zero per eccesso, ossia i valori di $f(x)$ tendono a zero mantenendosi non negativi.

Si osservi che non ogni limite finito è necessariamente assunto per eccesso o per difetto, come mostra la figura 3.5. Questa funzione tende a 1 per $x \to +\infty$, ma non si può affermare che $f(x) \to 1^+$, né che $f(x) \to 1^-$.

Figura 3.5.

Limite infinito all'infinito

Esempio

2.2 Dimostriamo che:

$$\lim_{x \to +\infty} \log_{1/2} x = -\infty.$$

Dobbiamo provare che per ogni successione $\{x_n\}$ tale che $x_n \to +\infty$, si ha

$$\lim_{n \to +\infty} \log_{1/2} x_n = -\infty.$$

A sua volta, questo significa provare che fissato $K > 0$ qualsiasi, risulti

$$\log_{1/2} x_n < -K \text{ definitivamente,}$$

il che equivale a:

$$x_n > \left(\frac{1}{2}\right)^{-K} = 2^K \text{ definitivamente.}$$

Ma per ipotesi $x_n \to +\infty$, quindi fissata la quantità positiva 2^K, certamente risulta $x_n > 2^K$ definitivamente, e l'asserto è dimostrato.

Asintoto obliquo

Nei casi in cui una funzione presenta limite infinito all'infinito, può accadere (ma non sempre accade) che esista una retta, obliqua, a cui il grafico della funzione si avvicina indefinitamente. Si parla in tal caso di *asintoto obliquo*. Precisamente: si dice che una funzione $f(x)$ *ha asintoto obliquo* $y = mx + q$ $(m \neq 0, q \in \mathbb{R})$ per $x \to +\infty$ (o per $x \to -\infty$) se

$$\lim_{x \to +\infty\,(-\infty)} [f(x) - (mx + q)] = 0.$$

Esempio

2.3 Sia $f(x) = 2x + 1 + e^x$. Studiamo la funzione per $x \to -\infty$. Si vede facilmente che $f(x) \to -\infty$; inoltre,

$$\lim_{x \to -\infty} [f(x) - (2x + 1)] = \lim_{x \to -\infty} e^x = 0.$$

Perciò, per definizione di asintoto obliquo, si riconosce che la retta $y = 2x + 1$ è asintoto obliquo per f, per $x \to -\infty$.

In casi meno elementari, anziché dover "indovinare" qual è l'asintoto obliquo è utile avere un criterio operativo per cercarlo. Vale in proposito la seguente:

Proposizione 3.2 *La funzione $f(x)$ ammette asintoto obliquo per $x \to +\infty$ se e solo valgono le seguenti due condizioni:*

1. *Esiste finito*

$$\lim_{x \to +\infty} \frac{f(x)}{x} = m \neq 0;$$

2. *Esiste finito*

$$\lim_{x \to +\infty} [f(x) - mx] = q$$

(dove m è il numero calcolato al punto 1). In tal caso l'asintoto è $y = mx + q$. Analogo criterio vale per $x \to -\infty$.

Si lascia per esercizio la facile dimostrazione. Vediamo un semplice esempio di applicazione; altri, più elaborati, si potranno affrontare dopo aver approfondito, nella prossima sezione, i metodi di calcolo dei limiti.

Esempio

2.4 Studiamo la funzione $f(x) = 3x + \sqrt{x}$ per $x \to +\infty$. Si vede subito che $f(x) \to +\infty$ per $x \to +\infty$. Vediamo se presenta un asintoto obliquo. Calcoliamo perciò (in base alla Proposizione precedente):

$$\lim_{x \to +\infty} \frac{3x + \sqrt{x}}{x} = \lim_{x \to +\infty} \left(3 + \frac{1}{\sqrt{x}}\right) = 3 \equiv m$$

(come si verifica facilmente applicando la definizione di limite). Calcoliamo ora

$$\lim_{x\to+\infty}[f(x)-mx]=\lim_{x\to+\infty}\left[\left(3x+\sqrt{x}\right)-3x\right]=\lim_{x\to+\infty}\sqrt{x}=+\infty.$$

Poiché questo limite è infinito (non esiste q), la funzione *non ammette* asintoto obliquo.

Limite infinito al finito

Esempio

2.5 Dimostriamo che:

$$\lim_{x\to 0}\frac{1}{x^2}=+\infty.$$

Dobbiamo provare che per ogni successione $\{x_n\}$ tale che $x_n\to 0$ e $x_n\neq 0\ \forall n$ si ha

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

(Si noti che la precisazione "$x_n\neq 0\ \forall n$", che è contenuta nella definizione di limite, era irrilevante negli esempi precedenti in cui $c=\pm\infty$, ma diventa ora importante). A sua volta, questo significa provare che fissato $K>0$ qualsiasi, risulti

$$\frac{1}{x_n^2}>K \text{ definitivamente,}$$

il che equivale a:

$$|x_n|<\frac{1}{\sqrt{K}} \text{ definitivamente.}$$

Ma per ipotesi $x_n\to 0$, quindi fissata la quantità positiva $\varepsilon=\frac{1}{\sqrt{K}}$, certamente risulta $|x_n|<\varepsilon$ definitivamente, e l'asserto è dimostrato.

Limite destro e sinistro

Talvolta una funzione si comporta diversamente (dal punto di vista del suo limite) a seconda che x si avvicini a c da destra o da sinistra. Ad esempio, per $x\to 0$ la funzione $1/x$ diventa sempre più grande in valore assoluto, ma con un segno diverso a seconda che sia $x\gtrless 0$. Per descrivere questo tipo di situazioni, si introduce il concetto di limite destro e sinistro:

Definizione 3.10 (Limite destro e sinistro) Se $c\in\mathbb{R}$ e $\ell\in\mathbb{R}^*$, si dice che

$$\lim_{x\to c^+}f(x)=\ell \text{ (rispettivamente, } \lim_{x\to c^-}f(x)=\ell),$$

e in tal caso si dice che il limite destro (sinistro) di $f(x)$ per x tendente a c è ℓ, se per ogni successione $\{x_n\}$ di punti di I tali che $x_n\to c^+$ ($x_n\to c^-$) per $n\to+\infty$ e $x_n\neq c\ \forall n$, si ha che $f(x_n)\to c$ per $n\to\infty$.

Ad esempio,

$$\lim_{x\to 0^+}\frac{1}{x}=+\infty \text{ e } \lim_{x\to 0^-}\frac{1}{x}=-\infty,$$

come si verifica facilmente. Si noti che, invece,

$$\lim_{x\to 0}\frac{1}{x} \text{ non esiste!}$$

Più in generale:

il limite $\lim_{x\to c} f(x) = \ell$ esiste se e solo se esistono, e sono entrambi uguali a ℓ, i limiti destro e sinistro $\lim_{x\to c^+} f(x)$, $\lim_{x\to c^-} f(x)$; tuttavia, può accadere che i limiti destro e sinistro esistano ma siano diversi tra loro, oppure uno solo dei due esista: in questi casi $\lim_{x\to c} f(x)$ non esiste.

Se f è definita in (a, b), il limite per $x \to a$ (rispettivamente $x \to b$) è automaticamente un limite destro (rispettivamente sinistro).

Asintoto verticale

Si dice che f ha un *asintoto verticale* di equazione $x = c$ ($c \in \mathbb{R}$) per $x \to c$ (oppure per $x \to c^+$ o $x \to c^-$) se $\lim_{x\to c} f(x) = +\infty$ o $-\infty$ (oppure questo accade per $x \to c^+$ o $x \to c^-$, rispettivamente).

Ogni situazione di limite infinito al finito, quindi, corrisponde graficamente alla presenza di un asintoto verticale, ossia di una retta verticale a cui il grafico della funzione si avvicina sempre più. Ad esempio, diciamo che

$x = 0$ è asintoto verticale per $\dfrac{1}{x^2}$ (per $x \to 0$);

$x = 0$ è asintoto verticale per $\dfrac{1}{x}$ (per $x \to 0^+$ e per $x \to 0^-$);

$x = 0$ è asintoto verticale per $\log x$ (per $x \to 0^+$).

Il lettore verifichi per esercizio l'ultima affermazione.

Limite finito al finito

Questo è, in un certo senso, il caso meno intuitivo dei 4 finora trattati; vediamo perché.

Esempi

2.6 Chiediamoci quanto vale:

$$\lim_{x\to 0} \sin x.$$

Chiunque, probabilmente, risponderebbe: zero. Dopotutto, $\sin 0 = 0$. Occorre rendersi conto, però, che anche se il risultato è proprio questo, la *motivazione* è errata. Infatti, la definizione di limite richiede che:

$\forall$ successione $\{x_n\}$ tale che $x_n \neq 0$ $\forall n$ e $x_n \to 0$ per $n \to \infty$, si abbia $\sin x_n \to 0$ per $n \to \infty$.

Quindi, per mostrare che $\lim_{x\to 0} \sin x = 0$ non si deve calcolare $\sin 0$! Si può procedere così, invece: dalla disuguaglianza elementare

$$|\sin x| \leq |x|, \text{ valida } \forall x \in \mathbb{R}$$

(si ragioni sulla circonferenza trigonometrica per rendersene conto), leggiamo che $|\sin x_n| \leq |x_n|$ e quindi, per il teorema del confronto, se $x_n \to 0$ anche $\sin x_n \to 0$. Pertanto

$$\lim_{x\to 0} \sin x = 0.$$

Si confronti col prossimo:

2.7 Sia

$$f(x) = \begin{cases} 1 \text{ se } x \neq 0 \\ 0 \text{ se } x = 0 \end{cases}$$

e si voglia calcolare

$$\lim_{x \to 0} f(x).$$

Applicando la definizione di limite si vede subito che il limite vale 1. Infatti, se $x_n \to 0$ per $n \to \infty$ ma $x_n \neq 0$ $\forall n$, risulta $f(x_n) = 1$ $\forall n$, e quindi $f(x_n) \to 1$ per $n \to \infty$. In questo caso, dunque (a differenza dell'esempio precedente), si ha:

$$\lim_{x \to 0} f(x) \neq f(0).$$

Continuità

Si rifletta sui due esempi appena visti. In entrambi i casi il limite al finito di una certa funzione esiste ed è finito. Nel primo caso, tale limite coincide col valore della funzione nel punto considerato, nel secondo caso no. Il buon senso suggerisce che la prima situazione è quella "normale", mentre la seconda è un po' "patologica". Questa idea è chiarita dalla prossima definizione, che introduce un nuovo concetto fondamentale.

Definizione 3.11 (Continuità) Se $f : I \to \mathbb{R}$ (I intervallo) e $c \in I$, si dice che f è *continua* in c se esiste

$$\lim_{x \to c} f(x) = f(c).$$

Si dice che f è continua in I se è continua in ciascun punto di I. Una funzione *non continua* in un punto c si dice *discontinua* in c.

Gli esempi precedenti si possono leggere dicendo che la funzione $\sin x$ è continua in $x = 0$, mentre la funzione $f(x)$ dell'esempio 2.7 è discontinua in $x = 0$. Si consideri anche il prossimo

Esempio

2.8

$$\lim_{x \to 0^{\pm}} \frac{x}{|x|} = \pm 1.$$

Questo esempio è interessante per due motivi. Anzitutto, in questo caso non sarebbe neppure possibile calcolare $f(0)$, quindi è chiaro a priori che il limite non si potrà calcolare semplicemente valutando la funzione nel punto 0. Inoltre, in questo caso i limiti destro e sinistro esistono finiti e sono diversi tra loro. Questo fatto merita una definizione:

Definizione 3.12 Si dice che c è un *punto di discontinuità a salto* per $f(x)$ quando *i limiti destro e sinistro in c esistono finiti*, ma diversi tra loro. Il salto è costituito dalla differenza dei limiti e precisamente:

$$\text{salto in } c = \lim_{x \to c^+} f(x) - \lim_{x \to c^-} f(x).$$

Se uno dei due limiti destro o sinistro per $x \to c$ coincide con $f(c)$, si dice che f è *continua da destra* o *da sinistra*, rispettivamente.

Nell'esempio precedente, la funzione $\frac{x}{|x|}$ ha un punto di discontinuità a salto in $x = 0$, con salto 2.

Le funzioni che presentano discontinuità a salto si prestano bene a modellizzare fenomeni che registrano bruschi cambiamenti.

Le funzioni continue, d'altro canto, devono la loro importanza al fatto che $\lim_{x \to x_0} f(x) = f(x_0)$ può essere interpretato dicendo che "se x è vicino a x_0" allora "$f(x)$ è vicino a $f(x_0)$", ossia

$$\text{piccole variazioni di } x \longmapsto \boxed{\quad f \quad} \longrightarrow \text{piccole variazioni di } f(x)$$

Per coglierne la rilevanza, pensiamo al seguente problema: noto x_0 e nota f, mediante una calcolatrice tascabile, calcolare $f(x_0)$.
Se x_0 è un numero irrazionale, la calcolatrice approssima x_0 con $x_0 + \Delta x$ (Δx è l'errore commesso nell'approssimazione) e calcola di conseguenza $f(x_0 + \Delta x)$, approssimando ulteriormente il risultato finale.

Se f è continua in x_0 e $|\Delta x| \ll 1$ (cioè l'errore assoluto $|\Delta x|$ è molto piccolo) è ragionevole considerare il risultato ottenuto come una buona approssimazione del risultato finale. Ma se f è discontinua, in corrispondenza di uno spostamento, anche piccolissimo, a destra o a sinistra di x_0, si trova un valore $f(x_0 + \Delta x)$ molto diverso da $f(x_0)$.

Non esistenza del limite

Il limite di una funzione può anche non esistere, naturalmente:

Esempio 2.9

$$\lim_{x \to +\infty} \sin x \quad \text{non esiste}$$

Per dimostrarlo è sufficiente trovare due successioni $\{x_n\}$ e $\{y_n\}$ divergenti a $+\infty$ tali che $\sin x_n$ e $\sin y_n$ tendano a due limiti diversi. Si può scegliere $x_n = n\pi$ e $y_n = \frac{\pi}{2} + 2n\pi$; in corrispondenza di tali successioni si ha $\sin x_n = 0$ e $\sin y_n = 1$ e quindi la definizione di limite non è soddisfatta.

Si rifletta sul fatto che, in base alle definizioni che abbiamo dato di limite di successione e limite di funzione, dimostrare rigorosamente che una funzione non ha limite è molto più facile che dimostrare che una *successione* non ha limite. Ad esempio, non è così banale dimostrare che non esiste $\lim_{n \to +\infty} \sin n$. Analogamente:

Esempio 2.10

$$\lim_{x \to 0} \sin \frac{1}{x} \quad \text{non esiste}$$

Infatti: sia $x_n = 1/(2n\pi)$, $y_n = 1/(\frac{\pi}{2} + 2n\pi)$. Entrambe le successioni tendono a 0, ma $\sin \frac{1}{x_n} = 0$, $\sin \frac{1}{y_n} = 1$. Poiché lungo successioni diverse che tendono a 0 la funzione ha limiti diversi, il limite della funzione non esiste.

La situazione è analoga all'esempio precedente, ma questa volta la funzione ha infinite oscillazioni in uno spazio finito (perciò non è possibile materialmente disegnarne il grafico in *tutto* l'intervallo $(0, 1)$, ad esempio; si veda la figura 3.6).

Figura 3.6.

Definizione topologica di limite

Concludiamo questa sezione presentando una diversa (ma equivalente) definizione di limite di funzione, che ha la caratteristica di essere indipendente dal concetto di successione e limite di successione, e che talvolta ci sarà utile.

Per precisare l'idea di "vicinanza a un punto" è comoda la nozione di *intorno*:

Definizione 3.13 (Intorni) Un *intorno* di un punto $x_0 \in \mathbb{R}$ è un intervallo aperto che contiene x_0. Useremo intorni del tipo $(x_0 - \delta, x_0 + \delta)$, centrati cioè in x_0. Dire che "x si muove in un intorno di x_0" significa affermare che $x \in (x_0 - \delta, x_0 + \delta)$, dove si pensa $0 < \delta \ll 1$. Si può parlare anche di intorni di $+\infty$ o di $-\infty$: sono intervalli della forma $(a, +\infty)$ o $(-\infty, b)$, rispettivamente.

Sarà utile anche introdurre la locuzione "definitivamente per $x \to c$", analogamente a quanto abbiamo fatto nel caso delle successioni:

Definizione 3.14 (Definitivamente) Diremo che una funzione $f(x)$ ha una certa proprietà *definitivamente* per $x \to c$ se esiste un intorno U di c tale che la proprietà vale per $f(x)$ per ogni $x \in U$, $x \neq c$.

Veniamo ora alla definizione topologica di limite.

Definizione 3.15 Sia $c \in \mathbb{R}^*$ e sia f una funzione definita almeno definitivamente per $x \to c$. Si dice che

$$\lim_{x \to c} f(x) = \ell \text{ (con } \ell \in \mathbb{R}^*)$$

se per ogni intorno U_ℓ di ℓ esiste un intorno V_c di c tale che

$$\forall x \in V_c, x \neq c \text{ si ha } f(x) \in U_\ell.$$

Si noti che, come anticipato, questa definizione di limite di funzione non fa ricorso al concetto di limite di successione. Anche questa definizione, come quella successionale, è una definizione unitaria, che contiene come casi particolari i 4 tipi di limiti che abbiamo studiato analiticamente (limite finito o infinito, al finito o all'infinito). Il concetto centrale che permette questa sintesi è in questo caso il concetto di *intorno*.

Poiché questo è un concetto base di quella branca della matematica che prende il nome di *topologia*, questa definzione di limite "mediante intorni" si dice anche "definizione topologica di limite".

Il lettore, a titolo di esercizio, verifichi che, particolarizzando il concetto di intorno di c ai casi $c \in \mathbb{R}$ o $c = \pm\infty$, la definizione topologica di limite si traduce nelle seguenti 4 definizioni:

Si dice che

$$\boxed{\lim_{x \to c} f(x) = \ell}$$

se:

1. (Limite finito al finito: $c, \ell \in \mathbb{R}$)

$$\forall \varepsilon > 0\ \exists \delta > 0 : \forall x \neq c, |x - c| < \delta \Longrightarrow |f(x) - \ell| < \varepsilon.$$

2. (Limite infinito al finito: $c \in \mathbb{R}, \ell = +\infty$)

$$\forall K > 0\ \exists \delta > 0 : \forall x \neq c, |x - c| < \delta \Longrightarrow f(x) > K.$$

3. (Limite finito all'infinito: $c = +\infty, \ell \in \mathbb{R}$)

$$\forall \varepsilon > 0\ \exists K > 0 : \forall x > K \Longrightarrow |f(x) - \ell| < \varepsilon.$$

4. (Limite infinito all'infinito: $c = +\infty, \ell = +\infty$)

$$\forall K > 0\ \exists H > 0 : \forall x > H \Longrightarrow f(x) > K.$$

Esercizi

8 Dimostrare in base alla definizione successionale di limite, i seguenti limiti di funzioni elementari:

$$\lim_{x \to +\infty} x^\alpha = \begin{cases} +\infty \text{ se } \alpha > 0 \\ 0^+ \text{ se } \alpha < 0 \end{cases}$$

$$\lim_{x \to +\infty} a^x = \begin{cases} +\infty \text{ se } a > 1 \\ 0^+ \text{ se } 0 < a < 1 \end{cases}$$

$$\lim_{x \to 0^+} \log_a x = \begin{cases} -\infty \text{ se } a > 1 \\ +\infty \text{ se } 0 < a < 1 \end{cases}$$

$$\lim_{x \to \left(\frac{\pi}{2}\right)^\pm} \operatorname{tg} x = \mp\infty$$

9 Dimostrare in base alla definizione successionale di limite che non esiste

$$\lim_{x \to +\infty} \operatorname{tg} x.$$

Più in generale, dimostrare che non esiste

$$\lim_{x \to +\infty} f(x)$$

se $f : \mathbb{R} \to \mathbb{R}$ è una funzione periodica non costante.

10 Dimostrare il criterio per la ricerca dell'asintoto obliquo contenuto nella Proposizione 3.2

11 Dimostrare in base alla definizione successionale di limite che non esiste (né finito, né infinito)

$$\lim_{x \to +\infty} x \sin x.$$

12 Sia

$$f(x) = \begin{cases} 1 & \text{se } x \text{ è razionale} \\ 0 & \text{se } x \text{ è irrazionale} \end{cases}$$

Dimostrare che f è discontinua in ogni punto.

13 Per ciascuna delle due funzioni parte intera di x e mantissa di x, si calcoli il limite per $x \to +\infty$, e si dica in quali punti la funzione è discontinua, precisando se in tali punti è (almeno) continua da destra o da sinistra.

14 Si determinino tutti i punti di discontinuità delle seguenti funzioni:

$$\operatorname{arctg} \frac{1}{x}; \quad \sin\left(\frac{1}{\sin \frac{1}{x}}\right); \quad [e^x] \text{ (dove } [\cdot] \text{ indica la parte intera).}$$

3 IL CALCOLO DEI LIMITI

Esaminiamo ora più da vicino le proprietà dei limiti e i principali metodi per il loro calcolo effettivo.

3.1 Proprietà fondamentali di limiti e continuità

Anzitutto possiamo enunciare i teoremi sui limiti di funzioni che discendono immediatamente dai corrispondenti teoremi sui limiti di successioni (che abbiamo dimostrato nel par. 1.3).

Utilizzeremo qui la nozione di "proprietà vera definitivamente per $x \to c$", introdotta alla fine del paragrafo precedente. Nei prossimi enunciati x_0 e ℓ saranno punti di $\mathbb{R}^*$, salvo avviso contrario.

Teorema 3.13 (del confronto) *Se:*

1. *Per $x \to c$, $f(x) \to \ell$ e $g(x) \to \ell$;*
2. *$f(x) \leq h(x) \leq g(x)$ definitivamente per $x \to c$*

allora anche $h(x) \to \ell$ per $x \to c$.

Corollario 3.14 *Se:*

1. *Per $x \to c$, $g(x) \to 0$;*
2. *$|h(x)| \leq g(x)$ definitivamente per $x \to c$*

allora anche $h(x) \to 0$ per $x \to c$.

Corollario 3.15 *Se $f(x) \to 0$ per $x \to 0$ e $g(x)$ è limitata definitivamente per $x \to c$, allora $f(x)\,g(x) \to 0$ per $x \to c$.*

Teorema 3.16 (di permanenza del segno, 1ª forma)
Se per $x \to c$, $f(x) \to \ell > 0$, allora $f(x) > 0$ definitivamente per $x \to c$.

Teorema 3.17 (di permanenza del segno, 2ª forma)
Se per $x \to c$, $f(x) \to \ell$, e $f(x) \geq 0$ definitivamente per $x \to c$, allora $\ell \geq 0$.

Teorema 3.18 (di permanenza del segno per funzioni continue)
Se f è continua in c e $f(c) > 0$, allora $f(x) > 0$ definitivamente per $x \to c$.

Teorema 3.19 (Algebra dei limiti, caso dei limiti finiti)
Se per $x \to c$, $f(x) \to \ell_1$ e $g(x) \to \ell_2$ ($\ell_1, \ell_2 \in \mathbb{R}$), allora per $x \to c$ si ha:

1. $f(x) \pm g(x) \to \ell_1 \pm \ell_2$;
2. $f(x)\,g(x) \to \ell_1 \ell_2$;
3. $f(x)/g(x) \to \ell_1/\ell_2$ *(purché $\ell_2 \neq 0, g(x) \neq 0$ definitivamente per $x \to c$).*

Teorema 3.20 (Aritmetizzazione parziale di ∞) *Valgono per i limiti di funzioni gli stessi risultati di "aritmetizzazione parziale di ∞" che valgono per i limiti di successioni (v. par. 1.3).*

Dimostrazione dei teoremi precedenti. Mostriamo in che modo i teoremi sui limiti di successioni dimostrati nel par. 1.3 e la definizione successionale di limite di funzione implicano questi teoremi sui limiti di funzioni. Come si vedrà, l'idea è semplice e applicata in modo ripetitivo.

Teorema del confronto Sia x_n una qualsiasi successione tale che $x_n \neq c\ \forall n$ e $x_n \to c$ per $n \to \infty$. Vogliamo provare che $h(x_n) \to \ell$ per $n \to \infty$. D'altro canto sappiamo che

$$f(x_n) \leq h(x_n) \leq g(x_n) \text{ definitivamente, e}$$
$$f(x_n) \to \ell \text{ e } g(x_n) \to \ell$$

(perché per ipotesi $f(x) \to \ell$ e $g(x) \to \ell$ e questo significa che per $x_n \to c$ si ha $f(x_n) \to \ell$ e $g(x_n) \to \ell$). Ma allora, per il teorema del confronto applicato alle successioni $f(x_n), h(x_n), g(x_n)$, si conclude che $h(x_n) \to \ell$, e il teorema è provato.

Dal teorema del confronto seguono i due corollari, esattamente come nel caso delle successioni.

Teorema di permanenza del segno. Sia x_n una qualsiasi successione tale che $x_n \neq c$ per ogni n e $x_n \to c$; allora $f(x_n) \to \ell > 0$ (per l'ipotesi), quindi per il teorema di permanenza del segno per successioni (applicato alla successione $f(x_n)$) si conclude che $f(x_n) > 0$ definitivamente. Poiché questo vale per ogni successione $x_n \to c$, si conclude che $f(x) > 0$ definitivamente per $x \to c$. (Si veda anche l'Esercizio 7, nei Complementi alla fine di questo capitolo.)

Da questo teorema si deduce il teorema nella sua seconda forma (come nel caso delle successioni), e anche il teorema di permanenza del segno per funzioni continue. Infatti: se f è continua in c, allora $f(c) = \lim_{x \to c} f(x)$, quindi l'ipotesi $f(c) > 0$ significa che $\lim_{x \to c} f(x) > 0$, e questo per il teorema di permanenza del segno (1ª forma) implica che $f(x) > 0$ definitivamente per $x \to c$.

TEOREMA SULL'ALGEBRA DEI LIMITI Sia x_n una qualsiasi successione tale che $x_n \neq c$ per ogni n e $x_n \to c$; allora per l'ipotesi si ha che $f(x_n) \to \ell_1$ e $g(x_n) \to \ell_2$; dal teorema sull'algebra dei limiti per successioni si conclude quindi che $f(x_n) \pm g(x_n) \to \ell_1 \pm \ell_2$, e quindi $f(x) \pm g(x) \to \ell_1 \pm \ell_2$. In modo perfettamente analogo si dimostrano i punti 2 e 3 del teorema, ed il teorema di aritmetizzazione parziale di ∞. ◆

Dal teorema dell'algebra dei limiti segue un analogo risultato per quanto riguarda le funzioni continue:

TEOREMA 3.21 (ALGEBRA DELLE FUNZIONI CONTINUE) *Siano f, g due funzioni definite almeno in un intorno di $x_0 \in \mathbb{R}$, e continue in x_0. Allora:*

1. *$f(x) \pm g(x)$ è continua in x_0;*
2. *$f(x)\, g(x)$ è continua in x_0;*
3. *$f(x)/g(x)$ è continua in x_0, purché $g(x_0) \neq 0$.*

DIMOSTRAZIONE. Proviamo ad esempio la 3, essendo le altre analoghe (e più semplici). Per ipotesi sappiamo che f e g sono continue in x_0, ossia:

$$f(x) \to f(x_0) \text{ e } g(x) \to g(x_0) \text{ per } x \to x_0.$$

Inoltre $g(x_0) \neq 0$ e quindi, per il teorema di permanenza del segno per funzioni continue, $g(x) \neq 0$ definitivamente per $x \to x_0$. Allora per il teorema sull'algebra dei limiti si conclude che

$$\frac{f(x)}{g(x)} \to \frac{f(x_0)}{g(x_0)} \text{ per } x \to x_0,$$

ossia f/g è continua in x_0. ◆

Vale anche il prossimo importante:

TEOREMA 3.22 (DI CONTINUITÀ DELLE FUNZIONI ELEMENTARI) *Le seguenti funzioni elementari sono continue in tutti i punti del proprio insieme di definizione:*

1. *Potenze a esponente intero, razionale o reale;*
2. *Funzioni esponenziali;*
3. *Funzioni logaritmiche;*
4. *Funzioni trigonometriche elementari* $(\sin x, \cos x)$.

Vedremo in seguito come, combinando tra loro in molti modi questi "componenti elementari" continui si ottengano ancora funzioni continue. Avremo quindi modo, molto spesso, di sapere *a priori* che una funzione è continua nel suo insieme di definizione (ossia senza applicare caso per caso la definizione di continuità). In questi casi, potremo allora sapere a priori che il limite di $f(x)$ per $x \to x_0$ si può effettivamente calcolare valutando semplicemente $f(x_0)$.

DIMOSTRAZIONE. Non dimostriamo il teorema in tutte le sue affermazioni. Ci limitiamo, a titolo d'esempio, a provare la continuità in tutto $\mathbb{R}$ delle funzioni $\sin x, \cos x$.

Sappiamo già (v. Esempio 2.6) che $\sin x$ è continuo in $x = 0$. Da questo segue che anche $\cos x$ è continuo in $x = 0$. Infatti, uno sguardo alla circonferenza trigonometrica mostra che, se x è un angolo nel primo quadrante, $\sin x + \cos x \geq 1$ (in quanto 1 è l'ipotenusa di un triangolo rettangolo di cateti $\sin x, \cos x$). Ne segue che

$$0 \leq 1 - \cos x \leq \sin x \text{ per } x \in \left(0, \frac{\pi}{2}\right), \text{ e quindi}$$
$$0 \leq 1 - \cos x \leq |\sin x| \text{ per } x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right).$$

Allora per il teorema del confronto $1 - \cos x \to 0$ per $x \to 0$, quindi $\cos x \to 1$ per $x \to 0$, e perciò $\cos x$ è continua in $x = 0$.

Per provare ora la continuità di $\sin x$ in un generico punto x_0 consideriamo la catena di relazioni:

$$\begin{aligned} |\sin(x_0 + h) - \sin x_0| &= |\sin x_0 \cosh + \cos x_0 \sinh - \sin x_0| \leq \\ &\leq |\sin x_0| \, |\cosh - 1| + |\cos x_0| \, |\sinh| . \end{aligned}$$

Ora $|\sinh|$ e $|\cosh - 1|$ tendono a zero per $h \to 0$, per quanto appena dimostrato, mentre $|\sin x_0|$ e $|\cos x_0|$ sono costanti, dunque $\sin(x_0 + h) - \sin x_0 \to 0$, ossia $\sin(x_0 + h) \to \sin x_0$ per $h \to 0$, e $\sin x$ è continua in x_0. Una analoga catena di disuguaglianze dimostra la continuità di $\cos x$. ♦

Esempio

3.1 Per il teorema 3.21 possiamo affermare ad esempio che:

1. Le funzioni potenza a esponente intero sono continue (in tutto $\mathbb{R}$), in quanto $f(x) = x$ è ovviamente continua, e $f(x) = x^n$ è il prodotto di n funzioni continue.
2. I polinomi sono funzioni continue, in quanto ottenuti sommando funzioni del tipo cx^n, continue per il punto precedente.
3. Le funzioni razionali (cioè i quozienti di polinomi) sono funzioni continue, tranne nei punti in cui si annulla il denominatore (che è un polinomio, quindi si annulla in un numero finito di punti).
4. Avendo dimostrato che $\sin x, \cos x$ sono continue in tutto $\mathbb{R}$, deduciamo che le funzioni $\operatorname{tg} x$, $\operatorname{cotg} x$ sono continue nel loro insieme di definizione.

Un'altra operazione che "conserva" i limiti e la continuità è la *composizione di funzioni*:

Teorema 3.23 (di cambio di variabile nel limite) *Siano f, g due funzioni per cui è ben definita la composizione $f \circ g$, almeno definitivamente per $x \to x_0$ ($x_0 \in \mathbb{R}^*$), e supponiamo che:*
i) $g(x) \to t_0$ per $x \to x_0$;
ii) esiste $\lim_{t \to t_0} f(t) = \ell \in \mathbb{R}^*$
iii) $g(x) \neq t_0$ definitivamente per $x \to x_0$.
Allora esiste

$$\lim_{x \to x_0} f(g(x)) = \lim_{t \to t_0} f(t).$$

I punti $x_0, t_0 \in \mathbb{R}^*$. L'ipotesi iii) non è necessaria nel caso in cui f sia continua in t_0, oppure (ovviamente) nel caso in cui $t_0 = \pm\infty$.

DIMOSTRAZIONE. Basta osservare che se x_n è una qualsiasi successione tale che $x_n \neq x_0 \forall n$ e $x_n \to x_0$ per $n \to \infty$, allora $g(x_n) \to t_0$ per $n \to \infty$ (per l'ipotesi (i)) e $g(x_n) \neq t_0$ definitivamente (per l'ipotesi (iii)). Perciò

$$f(g(x_n)) \to \ell$$

(per l'ipotesi (ii)), e questo prova la tesi. Se $t_0 = \pm\infty$ la condizione $g(x) \neq \pm\infty$ è ovviamente verificata, mentre se f è continua in t_0, $\ell = f(t_0)$, perciò nel caso risultasse $g(x_n) = t_0$ per qualche n, si avrebbe

$$f(g(x_n)) = f(t_0) = \ell$$

e quindi la convergenza $f(g(x_n)) \to \ell$ sarebbe comunque garantita. ◆

Dalla proprietà precedente, segue subito il prossimo:

TEOREMA 3.24 (DI CONTINUITÀ DELLA FUNZIONE COMPOSTA) *Siano: g una funzione definita almeno in un intorno di x_0 e continua in x_0, f una funzione definita almeno in un intorno di $t_0 = g(x_0)$ e continua in t_0, allora $f \circ g$ è definita almeno in un intorno di x_0 ed è continua in x_0.*

DIMOSTRAZIONE. Poiché g è continua in x_0,

$$\lim_{x \to x_0} g(x) = g(x_0) = t_0,$$

allora per il teorema del cambio di variabile nel limite, si ha:

$$\begin{aligned} \lim_{x \to x_0} f(g(x)) &= \lim_{t \to t_0} f(t) = \text{ poiché } f \text{ è continua in } t_0 \\ &= f(t_0) = f(g(x_0)), \end{aligned}$$

e la tesi è dimostrata. ◆

Vogliamo ora trarre qualche conclusione dai vari teoremi sulle funzioni continue visti finora. Sappiamo che

- le funzioni elementari dell'Analisi matematica sono continue nel loro insieme di definizione;
- somma, prodotto, quoziente di funzioni continue danno funzioni continue (dove il denominatore non si annulla);
- composizione di funzioni continue dà una funzione continua.

Ne segue che tutte le funzioni che si possono ottenere con somma, prodotto, quoziente e composizioni da funzioni elementari sono continue nel loro insieme di definizione. Risultano pertanto continue funzioni come

$$e^{-x^2} \qquad \sqrt{\sin x} \qquad \log_a\left(1 + (\operatorname{tg} x)^2\right), \ldots$$

Quindi, per esempio:

$$\lim_{x \to \pi} \sqrt{\sin x} = \sqrt{\sin \pi} = 0 \qquad \lim_{x \to 1} e^{-x^2} = e^{-1} = \frac{1}{e}$$

$$\lim_{x \to 0} \log_a\left(1 + (\operatorname{tg} x)^2\right) = \log_a\left(1 + (\operatorname{tg} 0)^2\right) = \log_a 1 = 0$$

poiché, per le funzioni continue, se x_0 è un punto del dominio, il limite si calcola semplicemente sostituendo x_0 nell'espressione analitica della funzione.

Mostriamo ora vari esempi di applicazioni dei teoremi di questo paragrafo.

Esempio

3.2 **Limiti di polinomi** *Per calcolare il limite di un polinomio per $x \to \pm\infty$ basta calcolare il limite del termine di grado massimo.* Infatti se il polinomio è $P(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n$, $a_0 \neq 0$, si ha:

$$P(x) = a_0x^n\left\{1 + \frac{a_1}{a_0x} + \frac{a_2}{a_0x^2} + \ldots + \frac{a_n}{x^n}\right\}$$

Tutti i termini dopo 1, in parentesi graffa, tendono a 0 per $x \to \pm\infty$ e quindi $\lim_{x\to\pm\infty} P(x) = \lim_{x\to\pm\infty} a_0x^n$.

Per esempio:

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

$$\lim_{x\to-\infty} (-2x^3 + x + 120) = \lim_{x\to-\infty} -2x^3 = +\infty$$

3.3 **Limiti di funzioni razionali** Sia ora $f(x) = P_n(x)/Q_m(x)$ dove P_n e Q_m sono, rispettivamente, polinomi di grado n e m dati da:

$$P_n(x) = a_0x^n + a_1x^{n-1} + \ldots + a_n$$

$$Q_m(x) = b_0x^m + b_1x^{m-1} + \ldots + b_m$$

Per calcolare il limite di $P_n(x)/Q_m(x)$ per $x \to \pm\infty$ basta calcolare il limite del rapporto dei monomi di grado massimo, cioè:

$$\lim_{x\to\pm\infty} \frac{P_n(x)}{Q_m(x)} = \lim_{x\to\pm\infty} \frac{a_0x^n}{b_0x^m}.$$

Infatti,

$$\lim_{x\to\pm\infty} \frac{P_n(x)}{Q_m(x)} = \lim_{x\to\pm\infty} \frac{a_0x^n\{\ 1 + \text{termini che tendono a } 0\}}{b_0x^m\{1 + \text{ termini che tendono a } 0\}} = \lim_{x\to\pm\infty} \frac{a_0x^n}{b_0x^m}.$$

Per esempio:

$$\lim_{x\to-\infty} \frac{3x^3+1}{x^2+2} = -\infty \qquad \lim_{x\to+\infty} \frac{3x^3+1}{x^4-10x+1} = 0 \qquad \lim_{x\to-\infty} \frac{6x^5+7x^2}{7x^5-4} = \frac{6}{7}$$

3.4 Una situazione simile analoga si ha nel calcolo del limite per $x \to 0$ di certi quozienti:

$$\lim_{x\to 0} \frac{x^{2/3}}{x^2 + 3\sqrt[3]{x} + x}$$

È una sorta di "quoziente di polinomi" (per la verità, sono potenze a esponente razionale, non intero): per $x \to 0$, tutte le potenze scritte tendono a 0, e si ha una forma di indeterminazione $\left[\frac{0}{0}\right]$. Questa volta, i termini preponderanti sono quelli di esponente minimo, perché $x \to 0$:

$$\frac{x^{2/3}}{x^2 + 3\sqrt[3]{x} + x} = \frac{x^{2/3}}{3x^{1/3}(1 + \frac{x^{2/3}}{3} + \frac{x^{5/3}}{3})} = \frac{x^{1/3}}{3(1 + \frac{x^{2/3}}{3} + \frac{x^{5/3}}{3})} \to 0$$

perché $x^{1/3} \to 0$ e $3\left(1 + \frac{x^{2/3}}{3} + \frac{x^{5/3}}{3}\right) \to 3$.

3.5
$$\lim_{x\to 0} x\sin\frac{1}{x} = 0$$

Infatti, poiché $\left|\sin\frac{1}{x}\right| \leq 1$, si ha

$$\left|x\sin\frac{1}{x}\right| \leq |x|$$

da cui, per il corollario del teorema del confronto, $x\sin\frac{1}{x} \to 0$. La conclusione è interessante perché il fattore $\sin\frac{1}{x}$, da solo, non ammette limite, quindi non si sarebbe potuto applicare il teorema sul limite del prodotto.

3.6
$$\lim_{x\to 0^{\pm}} e^{\frac{1}{x}} = \begin{cases} +\infty \\ 0 \end{cases}$$

Infatti, per $x \to 0^{\pm}$, $\frac{1}{x} \to \pm\infty$, e per $y \to \pm\infty, e^y \to \begin{cases} +\infty \\ 0 \end{cases}$ (rispettivamente). Abbiamo applicato, implicitamente, il teorema sul cambio di variabile nel limite.

3.7
$$\lim_{x\to 0^+} \frac{x-1}{\sin x} = -\infty$$

perché $(x-1) \to -1$ e $\sin x \to 0^+$. Abbiamo applicato il teorema sull'aritmetizzazione parziale di ∞.

3.8
$$\lim_{x\to -\infty} \left(\frac{1}{x} - 2\right) x^3 = +\infty$$

perché $\left(\frac{1}{x} - 2\right) \to -2$ e $x^3 \to -\infty$. Abbiamo applicato il teorema sull'aritmetizzazione parziale di ∞.

3.9
$$\lim_{x\to +\infty} \left(\frac{x+\sin x}{2x+\cos x}\right) = \frac{1}{2}.$$

Infatti:

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

Ora per il corollario del teorema del confronto,

$$\frac{\sin x}{x} \to 0 \text{ e } \frac{\cos x}{2x} \to 0, \text{ per } x \to +\infty.$$

Allora per il teorema sull'algebra dei limiti,

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

3.2 Limiti notevoli

Vediamo ora alcune comuni tecniche di calcolo dei limiti che, combinando i teoremi generali sui limiti con l'uso di alcuni *limiti notevoli* che riguardano certe funzioni elementari, risultano piuttosto potenti. I limiti notevoli, inoltre, saranno utili anche nel calcolo differenziale.

Limiti notevoli di seno e coseno

Nel calcolare $\lim_{x\to 0} \frac{\sin x}{x}$ si presenta la forma di indecisione $\frac{0}{0}$. Vogliamo dimostrare che

$$\boxed{\lim_{x\to 0} \frac{\sin x}{x} = 1} \tag{3.1}$$

Osserviamo subito che, essendo $\sin x$ e x funzioni dispari, $\frac{\sin x}{x}$ è funzione pari e quindi è sufficiente calcolare $\lim_{x\to 0^+} \frac{\sin x}{x}$. A tale scopo, osservando la figura 3.7 si vede che l'area del triangolo OPA è minore di quella del settore circolare OPA, a sua volta minore di quella del triangolo OTA. Ne segue

$$\frac{1}{2}\cdot 1\cdot \sin x \le \frac{1}{2}\cdot 1\cdot x \le \frac{1}{2}\cdot 1\cdot \operatorname{tg} x$$

ossia

$$\sin x < x < \operatorname{tg} x$$

Dividendo per $\sin x$, che è positivo perché $x \in (0, \frac{\pi}{2})$, si ha

$$1 < \frac{x}{\sin x} < \frac{1}{\cos x} \quad \forall x \in \left(0, \frac{\pi}{2}\right)$$

e infine, passando ai reciproci, si ottiene

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

Dal teorema del confronto, essendo $\lim_{x\to 0} \cos x = 1$, si deduce la (3.1).

$HP = \sin x$

$AT = \operatorname{tg} x$

$\widehat{AP} = x$

Figura 3.7.

Proviamo ora che:

$$\boxed{\lim_{x\to 0} \frac{1-\cos x}{x^2} = \frac{1}{2}} \tag{3.2}$$

Infatti

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

perché $\frac{\sin x}{x} \to 1$ e $(1+\cos x) \to 2$.

Prolungamento per continuità di una funzione

In base al limite (3.1) dimostrato, la funzione $f(x) = \frac{\sin x}{x}$, inizialmente non definita per $x = 0$, può essere *prolungata per continuità* anche in $x = 0$, ponendo

$$f(x) = \begin{cases} \frac{\sin x}{x} & \text{se } x \neq 0 \\ 1 & \text{se } x = 0 \end{cases}$$

La funzione f così definita risulta continua anche in $x = 0$.

Più in generale: se una funzione $f(x)$ non è definita in x_0 ma esiste finito

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

la funzione può essere prolungata per continuità anche in x_0, ponendo per definizione $f(x_0) = l$.

Lo studente si convinca del fatto che se invece la funzione f possiede in x_0 una discontinuità a salto, un asintoto verticale, o comunque non ammette limite finito, non è possibile renderla continua in x_0 alterandone la definizione *in un punto solo*.

Ogni volta che una funzione, inizialmente non definita in x_0, risulti essere prolungabile con continuità in x_0, si sottointende che la funzione sia stata effettivamente prolungata.

Limite di $\left(1+\frac{1}{x}\right)^x$ per $x \to \pm\infty$ e altri limiti notevoli collegati

Già sappiamo (v. Teorema 3.9) che per ogni successione $\{a_n\}$ divergente a $+\infty$ o $-\infty$ si ha

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

Questo fatto, per la definizione successionale di limite di funzione, implica immediatamente che:

$$\boxed{\left(1+\frac{1}{x}\right)^x \to e \quad \text{per} \quad x \to \pm\infty} \tag{3.3}$$

Dal limite (3.3) se ne possono dedurre alcuni altri.

Passando ai logaritmi nella (3.3), si ottiene

$$\log\left(1+\frac{1}{x}\right)^x = x\log\left(1+\frac{1}{x}\right) \to \log e = 1 \quad \text{per} \quad x \to \pm\infty$$

perché la funzione $\log x$ è continua.

Se si pone $y = \frac{1}{x}$, $x \to \pm\infty$ equivale a $y \to 0^{\pm}$ e l'ultimo limite si può scrivere nella forma seguente

$$\boxed{\frac{\log(1+y)}{y} \to 1 \quad \text{se} \quad y \to 0} \tag{3.4}$$

Se poi nella (3.4) poniamo $y = e^x - 1$, allora $y \to 0$ equivale a $x \to 0$, e sostituendo si ricava

$$\frac{\log(e^x)}{e^x - 1} = \frac{x}{e^x - 1} \to 1 \quad \text{se} \quad x \to 0$$

o, meglio:

$$\boxed{\frac{e^x - 1}{x} \to 1 \quad \text{se} \quad x \to 0} \tag{3.5}$$

Infine, se nella (3.4) poniamo $y = (1+x)^{\alpha} - 1$, con α esponente reale qualsiasi, allora $x \to 0$ equivale ad $y \to 0$, e si ha

$$\frac{\log(1+y)}{y} = \frac{\log[1 + (1+x)^{\alpha} - 1]}{(1+x)^{\alpha} - 1} = \frac{\alpha \log(1+x)}{(1+x)^{\alpha} - 1} =$$

$$= \frac{\alpha x}{(1+x)^{\alpha} - 1} \cdot \frac{\log(1+x)}{x} \to 1$$

e poiché anche $\frac{\log(1+x)}{x} \to 1$, se ne deduce:

$$\boxed{\lim_{x \to 0} \frac{(1+x)^{\alpha} - 1}{x} = \alpha} \tag{3.6}$$

3.3 Confronti e stime asintotiche

I limiti (3.1), (3.2), (3.4), (3.5) (3.6) sono espressi in modo forse più significativo usando il simbolo di equivalenza asintotica $\sim$, già introdotto per le successioni.

Definizione 3.16 Si dice che due funzioni f, g sono *asintotiche per* $x \to c$ se

$$\lim_{x \to c} \frac{f(x)}{g(x)} = 1$$

e si scrive $f \sim g$ per $x \to c$. ■

Il simbolo di asintotico gode di tutte le proprietà enunciate nel paragrafo 1.5. I limiti precedenti si possono allora riscrivere come segue:

$$\begin{gathered} \sin x \sim x \quad , \quad 1 - \cos x \sim \frac{1}{2}x^2 \quad , \quad e^x - 1 \sim x \\ \log(1+x) \sim x \quad , \quad (1+x)^{\alpha} - 1 \sim \alpha x \\ \text{per } x \to 0 \end{gathered} \tag{3.7}$$

Le (3.7) si possono esprimere dicendo ad esempio che $\sin x$, $\log(1+x)$ e $e^x - 1$, *in prima approssimazione* o *al primo ordine*, si comportano come x, per $x \to 0$.

Le (3.7) possono essere generalizzate. Se $\varepsilon(x)$ è una funzione che tende a zero [3] (cioè è un *infinitesimo*), si può scrivere

$$\boxed{\begin{array}{ll} \sin\varepsilon(x) \sim \varepsilon(x) & \\ 1-\cos\varepsilon(x) \sim \dfrac{1}{2}\varepsilon(x)^2 & \\ e^{\varepsilon(x)} - 1 \sim \varepsilon(x) & \varepsilon(x) \to 0 \\ \log(1+\varepsilon(x)) \sim \varepsilon(x) & \\ (1+\varepsilon(x))^\alpha - 1 \sim \alpha\varepsilon(x) & \end{array}} \tag{3.8}$$

Le (3.8) si deducono dalle (3.7) con il semplice cambio di variabile $y = \varepsilon(x)$.

Esempi

3.10
$$\lim_{x\to 0} \frac{\log(1+2x)}{\sin 3x} = \frac{2}{3}$$

Il limite dà una forma di indeterminazione $\left[\frac{0}{0}\right]$. Tuttavia, per $x \to 0$, $\log(1+2x) \sim 2x$, $\sin 3x \sim 3x$, e quindi

$$\frac{\log(1+2x)}{\sin 3x} \sim \frac{2x}{3x} = \frac{2}{3}$$

3.11
$$\lim_{x\to 1} \frac{(x-1)^2}{e^{3(x-1)^2} - 1} = \frac{1}{3}$$

Il limite dà una forma di indeterminazione $\left[\frac{0}{0}\right]$. Tuttavia, $e^{3(x-1)^2} - 1 \sim 3(x-1)^2$ per $x \to 1$ (la funzione $\varepsilon(x) = 3(x-1)^2$ è infinitesima per $x \to 1$), perciò

$$\frac{(x-1)^2}{e^{3(x-1)^2} - 1} \sim \frac{(x-1)^2}{3\,(x-1)^2} = \frac{1}{3}$$

3.12
$$\lim_{x\to +\infty} \left(\sqrt[3]{x^3+2x^2+1} - x\right) = \frac{2}{3}$$

È una forma di indeterminazione del tipo $[+\infty - \infty]$. Tuttavia:

$$\left(\sqrt[3]{x^3+2x^2+1} - x\right) = x\left(\sqrt[3]{1+\left(\frac{2}{x}+\frac{1}{x^3}\right)} - 1\right) \sim$$

$$\sim x\left[\frac{1}{3}\left(\frac{2}{x}+\frac{1}{x^3}\right)\right] \to \frac{2}{3}.$$

Abbiamo usato la stima $\sqrt[3]{1+\varepsilon(x)} - 1 \sim \frac{1}{3}\varepsilon(x)$, con $\varepsilon(x) = \left(\frac{2}{x} + \frac{1}{x^3}\right)$.

Altri importanti limiti riguardano il comportamento all'infinito di potenze, esponenziali e logaritmi, analoghi a quelli validi nel caso delle successioni:

[3] Non ha importanza a che cosa tende x.

Teorema 3.25 (Gerarchia degli infiniti) *Consideriamo le tre famiglie di funzioni:*

$$(\log_a x)^\alpha \qquad x^\beta \qquad b^x \quad \text{con } \alpha, \beta > 0, a, b > 1$$

Allora, per $x \to +\infty$, ognuna è un infinito di ordine inferiore rispetto a quella che le sta a destra. Esplicitamente:

$$\lim_{x\to+\infty} \frac{(\log_a x)^\alpha}{x^\beta} = 0 \qquad \lim_{x\to+\infty} \frac{x^\beta}{b^x} = 0 \tag{3.9}$$

il che si può esprimere a parole dicendo che qualunque potenza (positiva) di x prevale su qualunque potenza di $\log x$, e qualunque esponenziale (base > 1) di x prevale su qualunque potenza di x.

Inoltre ponendo $y = \frac{1}{x}$ nel primo limite in (3.9) si trova[4]

$$\boxed{\lim_{y\to 0^+} y^\beta(-\log y)^\alpha = 0} \qquad (\beta > 0\,,\ \forall\,\alpha) \tag{3.10}$$

Esempio 3.13

$$\lim_{x\to+\infty} \frac{x^3}{2^x} = 0$$

$$\lim_{x\to+\infty} \frac{(\log x)^2}{\sqrt{x}} = 0$$

$$\lim_{x\to 0^+} \sqrt{x}\log x = 0^-$$

$$\lim_{x\to 0^+} x^x = \lim_{x\to 0^+} e^{x\log x} = e^{0^-} = 1^-$$

$$\lim_{x\to 0^+} xe^{1/x} = \left[x = \frac{1}{t}\right] = \lim_{t\to+\infty} \frac{e^t}{t} = +\infty.$$

Il teorema 3.25 potrà essere dimostrato più avanti con gli strumenti del calcolo differenziale (cap. 4, par. 4.4).

3.4 Stime asintotiche e grafici

Le stime asintotiche non servono solo per calcolare limiti, ma anche per tracciare il grafico qualitativo di una funzione nell'intorno di un certo punto, oppure nell'intorno di $\pm\infty$.

Esempio 3.14 Sia

$$f(x) = \sqrt[3]{x} + x^2$$

[4] Perché ci vuole $(-\log y)^\alpha$ e sarebbe errato scrivere $(\log y)^\alpha$?

Si può ragionare così: la funzione è definita e continua su tutto $\mathbb{R}$; per $x \to \pm\infty$, $f(x) \sim x^2$. Dunque $f(x) \to +\infty$; inoltre il suo grafico sarà simile, per x grande in valore assoluto, a quello di x^2. La funzione inoltre si annulla in $x = 0$ e, per $x \to 0$, $f(x\ ) \sim \sqrt[3]{x}$; dunque il suo grafico sarà simile, in un intorno di $x = 0$, a quello di $\sqrt[3]{x}$: in particolare, avrà tangente vericale nell'origine. Il più semplice grafico compatibile con queste informazioni è:

Figura 3.8.

Spesso l'andamento di una funzione nell'intorno di un punto (ad esempio il fatto che abbia tangente verticale o orizzontale) è prevedibile in base a un'opportuna stima asintotica, che consente di tracciare il grafico qualitativo di f (nell'intorno del punto) per confronto con quello di una funzione nota (ad esempio, una potenza a esponente razionale). Analoghe stime sono utili per $x \to \pm\infty$.

3.15 Uno dei problemi fisici all'origine della teoria quantistica è quello della radiazione del corpo nero. La legge proposta nel 1900 da Planck per la distribuzione spettrale della radiazione emessa da un corpo nero è:

$$f(\lambda) = \frac{8\pi hc\lambda^{-5}}{e^{hc/\lambda kT} - 1}$$

Qui $f(\lambda)$ è la densità di energia raggiante corrispondente alla lunghezza d'onda λ, quando la temperatura del corpo è T; h è la costante di Planck, k la costante di Boltzmann e c la velocità della luce. Questa legge è in accordo con i dati sperimentali, mentre la legge che si può dedurre usando la fisica classica (legge di Rayleigh-Jeans) è

$$f(\lambda) = 8\pi kT\lambda^{-4}$$

che è in accordo coi dati sperimentali solo per λ grande, mentre per λ piccolo è in profondo disaccordo: tende a $+\infty$, mentre la curva sperimentale tende a zero!

Studiamo la curva di Planck con opportune stime asintotiche.

Per $\lambda \to +\infty$,

$$e^{hc/\lambda kT} - 1 \sim hc/\lambda kT$$

$$\text{e} \qquad \frac{8\pi hc\lambda^{-5}}{e^{hc/\lambda kT} - 1} \sim \frac{8\pi hc\lambda^{-5}}{hc/\lambda kT} = 8\pi kT\lambda^{-4}$$

ossia si ritrova la legge di Rayleigh-Jeans, in prima approssimazione; in particolare, la densità di energia raggiante tende a zero come λ^{-4}.

Per $\lambda \to 0^+$,

$$\frac{8\pi hc\lambda^{-5}}{e^{hc/\lambda kT}-1} \sim \frac{8\pi hc\lambda^{-5}}{e^{hc/\lambda kT}} \to 0^+$$

(l'esponenziale a denominatore tende a infinito più rapidamente di λ^{-5} a numeratore). Quindi la curva di Planck è in accordo coi dati sperimentali anche per λ piccolo. Poiché la funzione è sempre positiva, il suo grafico, in base a queste informazioni, sarà del tipo:

Figura 3.9.

Crescita di una funzione all'infinito

Supponiamo di voler tracciare il grafico di una funzione f che, per $x \to +\infty$ (o $-\infty$) tende a $+\infty$ (o $-\infty$). Per descrivere la velocità con cui la funzione tende all'infinito, sono utili le seguenti nozioni: diremo che, per $x \to +\infty$,

$$f \text{ ha crescita } \begin{cases} \text{sopralineare} \\ \text{lineare} \\ \text{sottolineare} \end{cases} \quad \text{se} \quad \lim_{x\to+\infty} \frac{f(x)}{x} = \begin{cases} \pm\infty \\ m \text{ (finito e diverso da 0)} \\ 0 \end{cases}$$

(Analoghe definizioni si danno per $x \to -\infty$). Solo nel caso in cui una funzione ha crescita lineare, è possibile che ammetta asintoto obliquo (v. par. 2 per il calcolo di tale asintoto).

Esempi di funzioni con crescita sopralineare per $x \to +\infty$ sono gli esponenziali a^x e le potenze x^a con $a > 1$; esempi di funzioni con crescita sottolineare per $x \to +\infty$ sono i logaritmi $\log_a x$ e le potenze x^a con $0 < a < 1$.

Esempio

3.14 La funzione

$$f(x) = 2x + e^x + e^{1/x}$$

per $x \to +\infty$ è asintotica a e^x; pertanto tende a $+\infty$ con crescita sopralineare; per $x \to -\infty$ è asintotica a $2x$; pertanto tende a $-\infty$ linearmente; poiché

$$\lim_{x\to-\infty} [f(x) - 2x] = 1,$$

la funzione ha asintoto obliquo $y = 2x + 1$ per $x \to -\infty$.

Esercizi

Sui limiti notevoli.

14 $\lim_{x\to+\infty} x\sin\frac{1}{x}$ 15 $\lim_{x\to 0}\frac{\cos x-1}{\sin^2 x}$ 16 $\lim_{x\to+\infty} x\log\left(\frac{x+3}{x+1}\right)$

Suggerimento: poiché $\frac{x+3}{x+1}\to 1$, si può porre $\frac{x+3}{x+1}=1+\varepsilon(x)$ con $\varepsilon(x)$ infinitesima; quindi...

17
$$\lim_{x\to+\infty}\left(\frac{x^2+3}{x^2+2}\right)^x$$

Suggerimento: dovendo calcolare il limite di una funzione del tipo $f(x)^{g(x)}$ che dà una forma di indeterminazione del tipo $[1^\infty]$, riscriverla nella forma

$$f(x)^{g(x)}=e^{g(x)\log f(x)}$$

e cominciare a calcolare il limite dell'esponente (che è ora una forma di indeterminazione del tipo $[\infty\cdot 0]$). Per far questo, seguire il suggerimento dell'esercizio precedente.

18 Fornire una stima asintotica, per $x\to 0$, delle funzioni: $\operatorname{tg}x$, $\operatorname{cotg}x$, $\operatorname{arctg}x$, $\arcsin x$.

19 $\lim_{x\to\pm\infty}\frac{2^{x+\frac{1}{x}}+\log|x|}{x^2+1}$

20 $\lim_{x\to 0}\frac{e^{x^2}-1}{\log(1+2x^2)}$ 21 $\lim_{x\to 0}\frac{x(\sin 2x)^2}{\sin(x^3)}$

22 $\lim_{x\to\pm\infty}\left(\sqrt{x^2+3x+2}-|x|\right)$ 23 $\lim_{x\to+\infty}\frac{\log(\log x)}{1+\log x}$

24 $\lim_{x\to 1}\frac{(\log x)^2}{(2x-2)^2}$ 25 $\lim_{x\to 0} xe^{\sin\frac{1}{x}}$

26 $\lim_{x\to+\infty} x\log\left(\frac{x+5}{x-1}\right)$ 27 $\lim_{x\to+\infty} xe^{\sin x}$

Per ciascuna delle seguenti funzioni: dire dove è definita e dove è continua; calcolare i limiti alla frontiera dell'insieme di definizione (punti in cui non è definita ed eventualmente $\pm\infty$); determinare tutti gli eventuali asintoti (verticali, orizzontali, obliqui); utilizzando opportune stime asintotiche nei punti in cui la funzione si annulla e all'infinito, prevedere l'andamento della funzione nell'intorno di tali punti.

28 $\frac{x^2+3x-1}{x+2}$ 29 $\frac{x^3+2x-1}{x+2}$ 30 $xe^{\frac{2x+1}{x+3}}$

31 $\frac{\log x}{\sqrt[3]{x-1}}$ 32 $\frac{\cos x-1}{x}$ 33 $\log\left(\frac{2x+1}{x^2+x-2}\right)$

34 $\operatorname{arctg}\left(\frac{x+2}{x-1}\right)$ 35 $x\operatorname{arctg}\frac{1}{x}$ 36 $x\log\left(\frac{2x^2+3}{x^2+x+1}\right)$

37 $2xe^{-1/x}$ 38 $x^2e^{-|x|}$ 39 $\frac{\sqrt[3]{x}\log|x|}{\log|x+1|}$

40 $\frac{x^2+2x-1}{x^2-1}$ 41 $xe^{-1/|x|}$

42 L'impedenza Z di un circuito elettrico RLC: R (resistenza), L (induttanza), C (capacità), è data da

$$Z(\omega)=\sqrt{R^2+\left(L\omega-\frac{1}{C\omega}\right)^2}\qquad(\omega>0).$$

Studiare, con l'aiuto del computer, come varia Z al variare dei parametri R, L, C.

4 PROPRIETÀ GLOBALI DELLE FUNZIONI CONTINUE O MONOTONE SU UN INTERVALLO

4.1 Funzioni continue su un intervallo

Le proprietà delle funzioni continue che abbiamo incontrato finora si basano sostanzialmente sul concetto di continuità *in un punto*. Come vedremo ora, invece, è soprattutto il fatto che una funzione sia continua su *un intervallo* ad avere delle conseguenze interessanti. La proprietà di continuità su un intervallo ha una semplice interpretazione geometrica: il grafico della funzione, su quell'intervallo, si può tracciare "senza staccare la penna dal foglio". Il concetto di funzione continua si presta quindi a tradurre in maniera rigorosa l'idea geometrica di *linea continua*, e quella fisica di *grandezza che varia con continuità*. Questo è vero, però, pur di utilizzare funzioni definite su un *intervallo di* $\mathbb{R}$. Come vedremo, infatti, le proprietà notevoli delle funzioni continue su un intervallo dipendono in maniera cruciale dall'assioma di continuità di $\mathbb{R}$, e non varrebbero, ad esempio, se il nostro ambiente di lavoro fosse il campo $\mathbb{Q}$ dei razionali.

La prima proprietà di cui ci occupiamo, che chiamiamo *teorema degli zeri*, riguarda la risoluzione di un'equazione del tipo

$$f(x) = 0 \tag{4.1}$$

Quando f è un polinomio di grado ≤ 4 esistono formule che forniscono le soluzioni della (4.1) mediante radicali. Se però f è un polinomio di grado > 4 o una funzione più complicata, salvo casi particolarmente fortunati, non esistono formule per le soluzioni. Geometricamente, risolvere la (4.1) significa determinare le ascisse dei punti di intersezione tra il grafico di $y = f(x)$ e l'asse delle ascisse. Naturalmente possono esserci infinite soluzioni, un numero finito di soluzioni, nessuna soluzione. Ogni soluzione si chiama *zero* di f.

Figura 3.10. L'equazione $f(x) = 0$ ha 4 soluzioni, ossia f ha 4 zeri.

Il teorema degli zeri dà alcune semplici condizioni sotto le quali esiste uno zero di f e anche un modo per calcolarlo.

Teorema 3.26 (degli zeri) *Sia:*

i) f *continua in* $[a, b]$

ii) $f(a) \cdot f(b) < 0$

Allora esiste $c \in (a, b)$ *tale che* $f(c) = 0$.

Se f *è anche strettamente monotona, lo zero è unico.*

DIMOSTRAZIONE. Costruiamo una successione che tende a uno zero di f. Poniamo $c_1 = \frac{a+b}{2}$, punto medio dell'intervallo $[a, b]$.
Se $f(c_1) = 0$ siamo fortunati e il teorema è dimostrato.
Se $f(c_1) \neq 0$ guardiamo il segno di $f(a) \cdot f(c_1)$ e procediamo come segue:

$$\text{se} \quad \begin{array}{lcl} f(a) \cdot f(c_1) < 0 & & \text{dove} \quad a_1 = a, \quad b_1 = c_1 \\ \searrow & & \nearrow \\ & \text{consideriamo l'intervallo } [a_1, b_1] & \\ \nearrow & & \searrow \\ f(a) \cdot f(c_1) > 0 & & \text{dove} \quad a_1 = c_1, \quad b_1 = b \end{array}$$

Poniamo ora $c_2 = \frac{b_1+a_1}{2}$, punto medio dell'intervallo $[a_1, b_1]$, calcoliamo $f(c_2)$ e procediamo come prima:
se $f(c_2) = 0$ lo zero cercato è c_2 e il teorema è dimostrato; se $f(c_2) \neq 0$ guardiamo il segno di $f(a_1) \cdot f(c_2)$:

$$\text{se} \quad \begin{array}{lcl} f(a_1) \cdot f(c_2) < 0 & & \text{dove} \quad a_2 = a_1, \quad b_2 = c_2 \\ \searrow & & \nearrow \\ & \text{consideriamo l'intervallo } [a_2, b_2] & \\ \nearrow & & \searrow \\ f(a_1) \cdot f(c_2) > 0 & & \text{dove} \quad a_2 = c_2, \quad b_2 = b_1 \end{array}$$

Continuando in questo modo troviamo una sequenza di intervalli $[a_n, b_n]$ con le seguenti proprietà:

1) $a_n \leq a_{n+1}$ e $b_n \geq b_{n+1}$ ($\{a_n\}$ cresce e $\{b_n\}$ decresce)

2) $b_n - a_n = \dfrac{b-a}{2^n}$ (ciascun intervallo è lungo la metà del precedente)

3) $f(a_n) \cdot f(b_n) < 0$ (per come sono stati scelti a_n e b_n a ogni passo)

Figura 3.11.

Per la 1), possiamo allora dedurre che le successioni $\{a_n\}$ e $\{b_n\}$ hanno limite finito, in base al teorema di esistenza del limite delle successioni monotone; cioè:

$$a_n \to \ell_1 \quad \text{e} \quad b_n \to \ell_2, \quad n \to +\infty$$

(Si noti che le successioni sono anche limitate perché contenute in $[a, b]$.)

Dalla 2) deduciamo che $b_n - a_n = \frac{b-a}{2^n} \to 0$ se $n \to +\infty$ e perciò $\ell_1 = \ell_2 = \ell$.

Per la continuità di f, abbiamo allora che

$$f(a_n) \cdot f(b_n) \to f(\ell)^2$$

mentre dalla 3) e dal teorema della permanenza del segno deduciamo $f(\ell)^2 \leq 0$. Deve perciò essere $f(\ell) = 0$ e così ℓ è lo zero cercato.

Arrestando il procedimento indicato nella dimostrazione dopo n passi, si può assumere a_n o b_n come valore approssimato dello zero ℓ di f, rispettivamente, per difetto o per eccesso. L'errore commesso è non superiore a $\frac{b-a}{2^n}$.

Si osservi che, se in $[a, b]$ vi sono più zeri della funzione f, il procedimento illustrato non indica quale di essi venga determinato.

Una seconda importante proprietà delle funzioni continue è espressa dal seguente:

Teorema 3.27 (di Weierstrass) *Sia $f : [a, b] \to \mathbb{R}$ una funzione continua. Allora f assume massimo e minimo in $[a, b]$, ossia: esistono $x_m, x_M \in [a, b]$ tali che $f(x_m) \leq f(x) \leq f(x_M)$ per ogni $x \in [a, b]$.*

Si dice che x_m è *punto di minimo* per f, e $m = f(x_m)$ è il *minimo* di f; analogamente, x_M è *punto di massimo*, e $M = f(x_M)$ è il *massimo* di f.

Dimostrazione. Proveremo che f ammette massimo in $[a, b]$. Analogamente si prova che ha minimo.

Cominciamo con la seguente osservazione sul concetto di estremo superiore: se E_1, E_2 sono due sottoinsiemi non vuoti di $\mathbb{R}$,

$$\sup(E_1 \cup E_2) = \max(\sup E_1, \sup E_2).$$

Questa proprietà è vera per insiemi sia limitati che illimitati (se E_i è superiormente illimitato si pone $\sup E_i = +\infty$; si intende che $\max(a, +\infty) = +\infty$). Sia ora

$$f : I \to \mathbb{R} \text{ con } I = I_1 \cup I_2$$

(I, I_1, I_2 intervalli). Applicando l'osservazione precedente agli insiemi $E_i = \{f(x) : x \in I_i\}$ si ha:

$$\sup_I f = \max\left(\sup_{I_1} f, \sup_{I_2} f\right).$$

In particolare quindi, è vera almeno una delle due relazioni:

$$\sup_I f = \sup_{I_1} f; \quad \sup_I f = \sup_{I_2} f.$$

Veniamo alla dimostrazione del teorema. Useremo una costruzione iterativa simile a quella vista nella dimostrazione del teorema degli zeri. Consideriamo la funzione $f : [a, b] \to \mathbb{R}$, e poniamo

$$\Lambda = \sup_{[a,b]} f$$

(con Λ eventualmente uguale a $+\infty$). Suddividiamo $[a,b]$ in due intervalli uguali; per quanto osservato sopra, per uno di essi, che chiamiamo $[a_1,b_1]$, sarà vero che

$$\Lambda = \sup_{[a_1,b_1]} f.$$

Suddividiamo $[a_1,b_1]$ in due intervalli uguali; per uno di questi, che chiamiamo $[a_2,b_2]$, sarà vero che

$$\Lambda = \sup_{[a_2,b_2]} f.$$

Procedendo per dicotomia, così facendo costruiamo una successione di intervalli $[a_n,b_n]$, ciascuno contenuto nei precedenti, con le proprietà:

1. a_n monotona crescente e limitata; b_n monotona decrescente e limitata;
2. $b_n - a_n = \frac{b-a}{2^n} \to 0$;
3. $\Lambda = \sup_{[a_n,b_n]} f$.

Per lo stesso ragionamento fatto nella dimostrazione del teorema degli zeri (che utilizza il teorema di monotonia per le successioni), dai punti 1) e 2) segue che le successioni a_n e b_n convergono ad uno stesso limite $x_0 \in [a,b]$. Distinguiamo ora due casi: $\Lambda < +\infty, \Lambda = +\infty$ (proveremo che questo secondo caso in realtà non può verificarsi). Se $\Lambda < +\infty$, per ogni n esiste un punto $t_n \in [a_n,b_n]$ tale che

$$\Lambda - \frac{1}{n} < f(t_n) \leq \Lambda. \tag{4.2}$$

Infatti: poiché $\Lambda - \frac{1}{n}$ è minore di Λ che è il minimo dei maggioranti dei valori di $f(x)$ in $[a_n,b_n]$, $\Lambda - \frac{1}{n}$ non è un maggiorante, quindi esiste $t_n \in [a_n,b_n]$ con la proprietà (4.2).

Poiché $t_n \in [a_n,b_n]$, $a_n \to x_0$ e $b_n \to x_0$, per il teorema del confronto anche $t_n \to x_0$.

Ancora per il teorema del confronto, la (4.2) dà allora

$$\lim_{n\to\infty} f(t_n) = \Lambda.$$

D'altro canto, poiché f è continua e $t_n \to x_0$, si ha che $\lim_{n\to\infty} f(t_n) = f(x_0)$, perciò $f(x_0) = \Lambda$. Allora, poiché Λ è il sup dei valori di f in $[a,b]$, Λ è il *massimo* di f in $[a,b]$, ed è assunto nel punto x_0. In questo caso perciò il teorema è dimostrato.

Supponiamo ora che sia $\Lambda = +\infty$. Allora per ogni n esiste un punto $t_n \in [a_n,b_n]$ tale che

$$f(t_n) \geq n. \tag{4.3}$$

Ragionando come sopra si prova che $t_n \to x_0$ per un certo $x_0 \in [a,b]$. Poiché f è continua in x_0,

$$\lim_{n\to\infty} f(t_n) = f(x_0),$$

ma per la (4.3) $\lim_{n\to\infty} f(t_n) = +\infty$, assurdo perché in x_0 la funzione deve avere un valore finito. ◆

Mostriamo ora che le ipotesi del teorema di Weierstrass sono tutte necessarie, con i seguenti esempi

Esempi

3.17 $f(x) = x$ in $(0,1)$.

La funzione è continua su un intervallo limitato, ma non chiuso. La funzione non ha né massimo né minimo (il suo estremo superiore, 1, e il suo estremo inferiore, 0, non sono assunti).

3.18 $f(x) = x$ in $\mathbb{R}$. La funzione è continua su un intervallo non limitato. La funzione non ha né massimo né minimo (non è nemmeno limitata).

3.19
$$f(x) = \begin{cases} x & \text{per } x \in (0,1) \\ \frac{1}{2} & \text{per } x = 0, x = 1 \end{cases}$$

La funzione è definita su un intervallo chiuso e limitato $[0,1]$, ma non è continua. La funzione non ha né massimo né minimo (il suo estremo superiore, 1, e il suo estremo inferiore, 0, non sono assunti).

Teorema 3.28 (dei valori intermedi) *Se f è continua su $[a,b]$, allora per ogni valore λ compreso tra m e M (minimo e massimo di f in $[a,b]$), esiste un ingresso x in $[a,b]$ che ha il valore λ come uscita (proprietà dei valori intermedi).*

Dimostrazione. Quest'ultimo teorema è una semplice conseguenza del teorema di Weierstrass e del teorema degli zeri: sia $m < \lambda < M$, $f(x_2) = m$, $f(x_1) = M$. Allora la funzione $g(x) = f(x) - \lambda$ è continua in $[x_1, x_2]$ (supponendo $x_1 < x_2$) e $g(x_1) = f(x_1) - \lambda = M - \lambda > 0$, $g(x_2) = f(x_2) - \lambda = m - \lambda < 0$. Dal teorema degli zeri, esiste ℓ tale che $g(\ell) = 0$ e cioè $f(\ell) = \lambda$. ♦

Figura 3.12. a) L'immagine di $[a,b]$ è l'intervallo $[m,M]$. b) Una funzione che non ha la proprietà dei valori intermedi. Ogni valore λ tra y_1 e y_2 non è un'uscita di f.

Le proprietà dei due teoremi precedenti si possono sintetizzare nell'unico enunciato seguente:

Teorema 3.29 *Se $f : [a,b] \to \mathbb{R}$ è continua, allora $f([a,b]) = [m,M]$; cioè l'immagine di un intervallo $[a,b]$ è l'intervallo di estremi $m = \min_{[a,b]} f$ e $M = \max_{[a,b]} f$ (vedi fig. 3.12a).*

Riguardo al teorema dei valori intermedi, si consideri il prossimo semplice

Esempio

3.20 Sia $f(x) = x^2$ e consideriamo f come funzione definita dall'insieme $\mathbb{Q}$ dei razionali a $\mathbb{Q}$ stesso (questo è lecito perché il quadrato di un razionale è un razionale). Allora f *non ha la proprietà dei valori intermedi.* Infatti, ad esempio, $f(1) = 1, f(2) = 4$, ma f non assume tutti i valori razionali compresi tra 1 e 4: per esempio non assume mai il valore 2, o 3. In altre parole, la proprietà dei valori intermedi è vera per le funzioni continue *grazie alle proprietà dell'insieme dei numeri reali.* Questi, anzi, sono i motivi più interessanti per l'Analisi per cui è utile avere come ambiente di lavoro l'insieme dei reali e non l'insieme dei razionali.

Mostriamo come il teorema degli zeri permetta di dimostrare direttamente il teorema sull'esistenza della radice n-esima (enunciato nel cap.1, par. 5.1):

Teorema 3.30 *Per ogni $y \in \mathbb{R}, y > 0$ e $n \in \mathbb{N}, n \geq 1$, esiste uno e un solo $x \in \mathbb{R}, x > 0$, tale che $x^n = y$. (Questo numero x si definisce radice n-esima di y).*

Dimostrazione. Consideriamo la funzione $f(x) = x^n$. Questa funzione è continua su tutto $\mathbb{R}$, ad esempio perché è il prodotto di n funzioni continue $g(x) = x$. Mostriamo che esistono $x_2 > x_1 > 0$ tali che $x_1^n < y < x_2^n$. Infatti:

se $y \geq 1$ basta scegliere $x_1 = 1, x_2 = y$ e si ha $x_1^n \leq y < x_2^n$ $(1 \leq y < y^n)$;

se $y < 1$, basta scegliere $x_1 = y, x_2 = 1$ e si ha $x_1^n < y \leq x_2^n$ $(y^n < y \leq 1)$.

Allora possiamo applicare il teorema dei valori intermedi alla funzione continua $f(x) = x^n$ sull'intervallo $[x_1, x_2]$: poiché f è strettamente crescente, il suo minimo è x_1^n e il suo massimo è x_2^n; poiché $x_1^n < y < x_2^n$, esiste $x_0 \in [x_1, x_2]$ tale che $x_0^n = y$. ♦

4.2 Funzioni monotone su un intervallo

Ci occupiamo ora di funzioni monotone su un intervallo (e non necessariamente continue). Il prossimo teorema si può vedere come una estensione alle funzioni del teorema di monotonia per le successioni, visto nel par. 1.2. Questo teorema condivide con i teoremi sulle funzioni continue su un intervallo, che abbiamo discusso nel paragrafo precedente, due caratteristiche importanti:

la sua dimostrazione si basa sull'*assioma di continuità* di $\mathbb{R}$;

il teorema assume come ipotesi una *proprietà globale* della funzione (ossia la sua monotonia su un intervallo).

Teorema 3.31 (di monotonia) *Sia $f : (a, b) \to \mathbb{R}$ una funzione monotona. Allora per ogni $c \in (a, b)$ esistono finiti i limiti destro e sinistro, per $x \to c$; ai due estremi a, b esistono i limiti destro (in a) e sinistro (in b), eventualmente infiniti.*

Dimostrazione. Questo teorema non sarà dimostrato utilizzando il teorema di monotonia per le successioni, ma piuttosto ripetendo una dimostrazione analoga.

Per fissare le idee, supponiamo f *crescente* in (a, b). Sia $c \in (a, b)$ e sia

$$\Lambda = \sup \{f(x) : x \in (a, c)\}.$$

Notiamo che Λ esiste finito per la proprietà dell'estremo superiore, in quanto $f(c)$ è un maggiorante dell'insieme $\{f(x) : x \in (a,c)\}$. Proviamo che

$$\lim_{x \to c^-} f(x) = \Lambda.$$

Sia dunque x_n una qualsiasi successione in (a,c) tale che $x_n \to c$, e proviamo che $f(x_n) \to \Lambda$, ossia che per ogni $\varepsilon > 0$ risulta definitivamente

$$\Lambda - \varepsilon < f(x_n) < \Lambda + \varepsilon.$$

La seconda disuguaglianza è ovvia in quanto $f(x_n) \leq \Lambda$ per definizione di Λ, perché $x_n \in (a,c)$. Per provare la prima, osserviamo che essendo $\Lambda - \varepsilon$ minore di Λ, cioè del minimo maggiorante di $\{f(x) : x \in (a,c)\}$, non è un maggiorante di tale insieme, perciò esiste un punto $\overline{x} \in (a,c)$ tale che $f(\overline{x}) \geq \Lambda - \varepsilon$. Poiché f è crescente, ne segue che

$$f(x) \geq \Lambda - \varepsilon \ \forall x \in (\overline{x}, c).$$

D'altro canto per ipotesi $x_n < c \ \forall n$ e $x_n \to c$ per $n \to \infty$, perciò $x_n \in (\overline{x}, c)$ definitivamente. Ma allora

$$f(x_n) \geq \Lambda - \varepsilon \text{ definitivamente,}$$

che è quanto occorreva provare.

Analogamente si può dimostrare che esiste

$$\lim_{x \to c^+} f(x) = \inf\{f(x) : x \in (c,b)\}.$$

Per quanto riguarda i limiti ai due estremi dell'intervallo, se si vuole provare ad esempio che esiste

$$\lim_{x \to b^-} f(x),$$

si definirà ancora

$$\Lambda = \sup\{f(x) : x \in (a,b)\};$$

tuttavia in questo caso Λ potrebbe anche essere $+\infty$ (non possiamo affermare che $f(b)$ sia un maggiorante dell'insieme, perché in b la funzione non è definita). Nel caso in cui $\Lambda < \infty$ si può ripetere la dimostrazione precedente; se invece $\Lambda = \infty$, il ragionamento si modifica come segue. Dall'ipotesi

$$\sup\{f(x) : x \in (a,b)\} = +\infty$$

segue che per ogni $K > 0$ esiste $\overline{x} \in (a,b)$ tale che $f(\overline{x}) > K$; per la monotonia di f, allora,

$$f(x) > K \ \forall x \in (\overline{x}, b),$$

allora presa una successione $x_n \to b$, si ha che $x_n \in (\overline{x}, b)$ definitivamente, e quindi

$$f(x_n) > K \text{ definitivamente,}$$

perciò

$$\lim_{x \to b^-} f(x_n) = +\infty.$$

Una conseguenza del teorema di monotonia è che se una funzione è monotona in un intervallo (a,b), i suoi eventuali punti di discontinuità in (a,b) sono necessariamente *discontinuità a salto*, ad eccezione degli estremi a, b, in cui può aversi anche un asintoto verticale.

4.3 Continuità e invertibilità

Sia $f : I \to \mathbb{R}$, con I intervallo. Occupiamoci ancora del problema dell'invertibilità di f, vedendo ora il problema dal punto di vista della continuità.

Sappiamo già che una funzione definita su un intervallo e strettamente monotona è invertibile (con inversa monotona; v. cap.2, par. 4.2). Sappiamo anche che il viceversa non è vero in generale: esistono funzioni invertibili su un intervallo, e non monotone (si veda la figura seguente).

Figura 3.13.

Tuttavia, se aggiungiamo l'ipotesi della continuità il teorema si inverte:

Teorema 3.32 *Sia $f : I \to \mathbb{R}$, con I intervallo, una funzione continua in I. Allora f è invertibile in I se e solo se è strettamente monotona. In tal caso la sua inversa è ancora strettamente monotona e continua.*

Dimostrazione. Sappiamo già che se f è strettamente monotona è invertibile (indipendentemente dalle ipotesi che f sia continua, e che sia definita su un intervallo). Mostriamo che vale il viceversa, ossia che se è continua e invertibile, allora è strettamente monotona. Per assurdo, non sia strettamente monotona, allora esistono tre punti $x_1 < x_2 < x_3$ in I tali che

$$f(x_1) < f(x_2) \text{ e } f(x_2) > f(x_3),$$

oppure tali che

$$f(x_1) > f(x_2) \text{ e } f(x_2) < f(x_3).$$

Per fissare le idee, supponiamo che valga la prima delle due alternative (nell'altro caso si ragionerà analogamente). Confrontiamo i valori di $f(x_1)$ e $f(x_3)$; non possono essere uguali perché f per ipotesi è invertibile, dunque $f(x_1) < f(x_3)$ oppure $f(x_1) > f(x_3)$. Di nuovo, per fissare le idee supponiamo che valga la prima delle due alternative (nell'altro caso si ragionerà analogamente). Dunque sappiamo che:

$$x_1 < x_2 < x_3 \text{ e } f(x_1) < f(x_3) < f(x_2).$$

Poiché f è continua, per il teorema dei valori intermedi esiste $x_0 \in (x_1, x_2)$ tale che $f(x_0) = f(x_3)$. Poiché $x_0 \neq x_3$ (perché $x_0 < x_2 < x_3$), ne segue che f non può essere invertibile, assurdo.

Questo dimostra la prima parte del teorema. Sia ora f continua, strettamente monotona e quindi invertibile in I, e sia g la sua funzione inversa, ancora strettamente monotona e

invertibile. Proviamo che g è continua. Per quanto osservato dopo il teorema di monotonia, la funzione g, strettamente monotona, o è continua, oppure ha dei punti di discontinuità a salto. In tal caso l'immagine di g non è un intervallo (ma è l'unione di almeno due intervalli disgiunti), il che è assurdo perché tale immagine è I. Dunque g è continua. ◆

Si noti che nella dimostrazione del teorema precedente si sono utilizzati sia il teorema dei valori intermedi, sia il teorema di monotonia per le funzioni. Di nuovo, quindi, abbiamo fatto implicitamente uso, in modo essenziale, dell'assioma di continuità di $\mathbb{R}$. Il teorema appena dimostrato significa in particolare che:

Una funzione continua e invertibile su un intervallo, ha inversa continua.

Questo fatto completa la dimostrazione del teorema di continuità delle funzioni elementari; infatti:

la continuità di a^x implica la continuità di $\log_a x$;
la continuità di $\sin x$, $\cos x$, $\operatorname{tg} x$ implica la continuità di $\arcsin x$, $\arccos x$, $\operatorname{arctg} x$.

COMPLEMENTI

7 Dimostrare che se per ogni successione $x_n \to c$, con $x_n \neq c\ \forall n$, la proprietà $p(x_n)$ è vera definitivamente, allora la proprietà $p(x)$ è vera definitivamente per $x \to c$.

Questo fatto è stato implicitamente utilizzato nella dimostrazione del teorema di permanenza del segno per le funzioni.
Suggerimento: ragionare per assurdo, scrivendo in maniera esplicita e quantitativa cosa significa la falsità della tesi, e costruire una successione che contraddice l'ipotesi.

8 Dimostrare la seguente variante del Teorema di Weierstrass su intervalli illimitati:

"Sia $f : \mathbb{R} \to \mathbb{R}$ continua, $f(x) \geq 0$ per ogni $x \in \mathbb{R}$ e supponiamo che $f(x) \to 0$ per $x \to \pm\infty$. Allora f ha massimo in $\mathbb{R}$".
Suggerimento: sia x_0 un punto in cui $f(x_0) > 0$; poiché $\lim_{x\to+\infty} f(x) = 0$, esiste $K > 0$ tale che $f(x) < f(x_0)\ \forall |x| > K$. Applicando il teorema di Weierstrass sull'intervallo $[-K, K]$ ....

Mostrare poi con esempi che:

- il teorema è falso senza l'ipotesi $f(x) \geq 0$;
- il teorema è falso senza l'ipotesi $\lim_{x\to\pm\infty} f(x) = 0$;
- nelle ipotesi del teorema, il *minimo* di f in $\mathbb{R}$ può non esistere.

9 Dimostrare le seguenti varianti del Teorema degli zeri su intervalli illimitati e su intervalli aperti:

(a) "Sia $f : \mathbb{R} \to \mathbb{R}$ continua e supponiamo che

$$\lim_{x\to+\infty} f(x) \cdot \lim_{x\to-\infty} f(x) < 0$$

(supponendo anche che i due limiti esistano, finiti o infiniti). Allora esiste $c \in \mathbb{R}$ tale che $f(c) = 0$".

(b) "Sia $f : (a, b) \to \mathbb{R}$ continua e supponiamo che

$$\lim_{x \to a^+} f(x) \cdot \lim_{x \to b^-} f(x) < 0$$

(supponendo anche che i due limiti esistano, finiti o infiniti). Allora esiste $c \in (a, b)$ tale che $f(c) = 0$".

Suggerimento: usando l'ipotesi sui limiti, mostrare che in entrambi i casi esiste un intervallo $[x_1, x_2]$ contenuto nel dominio della funzione, tale che $f(x_1) \cdot f(x_2) < 0$, e quindi...

10 Utilizzando opportunamente i teoremi enunciati nell'esercizio 9, dimostrare che l'equazione

$$\operatorname{tg} x + e^x - 5 = 0$$

ha almeno una soluzione in $\left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$.

11 Enunciando e dimostrando un'opportuna variante del teorema dimostrato nell'esercizio 8, provare che la funzione

$$f(x) = x^3 e^{-x^2}$$

ha massimo e minimo in $\mathbb{R}$.

4 Calcolo differenziale per funzioni di una variabile

1 INTRODUZIONE AL CALCOLO DIFFERENZIALE

In questa introduzione vogliamo far riflettere l'allievo su alcuni *problemi* da cui il calcolo differenziale è nato. Porremo questi problemi sotto forma di domande, a cui non risponderemo subito, ma nel seguito del capitolo. Come si vedrà, la risposta a questi problemi coinvolge la nozione di *limite*, discussa in precedenza. Anzi, si può dire che storicamente la nozione di limite sia stata introdotta principalmente per sviluppare le idee del calcolo differenziale, che introdurremo in questo capitolo, e del calcolo integrale, di cui ci occuperemo nel capitolo 6.

Che cos'è la retta tangente in un punto a una curva?

Anche se geometricamente il concetto sembra intuitivo, non è così ovvio quale possa essere una definizione corretta. Consideriamo la più semplice figura curvilinea, ossia la circonferenza. Cos'è la tangente a una circonferenza in un suo punto P? In base alla geometria elementare possiamo rispondere:

1. È la retta passante per P e perpendicolare al raggio passante per P.

Oppure:

2. È l'unica retta passante per P che non interseca la circonferenza in altri punti.

Figura 4.1. La retta tangente alla curva $y = x^3$ nel punto $(1, 1)$ è $y = 3x - 2$, che taglia la curva anche in $(-2, -8)$.

Queste definizioni si possono estendere a curve più generali? La definizione 1 certamente no: cos'è il raggio di una parabola, ad esempio? La definizione 2 sembra più incoraggiante perché "funziona" anche per le coniche (parabole, ellissi, iperboli). Ma già per la curva $y = x^3$ non funziona: la retta tangente a questa curva in un punto generico (escluso l'origine) taglia la curva anche in un altro punto (fig. 4.1); viceversa ci sono curve, come $y = |x|$ che in un punto (l'origine) hanno infinite rette soddisfacenti questa definizione, ma intuitivamente nessuna di esse si può chiamare tangente (fig. 4.2).

Figura 4.2. La funzione $y = |x|$ ha infinite rette che tagliano il suo grafico solo nell'origine: nessuna di esse è "tangente".

Ci accorgiamo dunque di parlare comunemente di un oggetto geometrico ("retta tangente") di cui non sappiamo (ancora) dare una definizione precisa e generale.

In realtà non c'è speranza di definire la tangente come la retta che ha "meno intersezioni delle altre" con la curva. Ci vuole un'idea diversa. Geometricamente, la retta tangente può essere approssimata dalla retta che passa per due punti molto vicini posti sulla curva, il primo fissato (il punto di tangenza) e l'altro "mobile" (e scelto via via più vicino al punto di tangenza). Due punti individuano una retta; più i due punti si avvicinano, più questa retta si avvicinerà alla tangente; se però partissimo subito dal considerare "due punti coincidenti", avremmo un solo punto, per il quale passano infinite rette, e la retta che cerchiamo resterebbe indeterminata. Si tratta quindi di capire cosa accade della retta per i due punti quando il secondo punto, mobile, si avvicina sempre più al primo, senza però coincidere con esso. La "retta limite" – se esiste – si chiamerà retta tangente.

Notiamo che il problema della determinazione della tangente è importante non solo da un punto di vista geometrico, ma anche per un'importante applicazione: è facile convincersi (vedi fig. 4.3) che una curva il cui grafico sia "regolare" ha tangente orizzontale nei punti di massimo e di minimo (ed eventualmente anche in altri punti).

Perciò, per cercare i punti di massimo e minimo di una funzione, è utile saper scrivere analiticamente l'equazione della retta tangente alla curva in un punto gene-

Figura 4.3.

rico, per vedere poi in quali punti essa è orizzontale. Questa idea si deve per primo a Fermat, che la elaborò intorno al 1630. Ad esempio, la meccanica insegna che un sasso lanciato verso l'alto (nel vuoto) con una velocità iniziale v_0, dopo un tempo t si trova a un'altezza

$$h(t) = v_0 t - \frac{1}{2} g t^2,$$

dove g è l'accelerazione di gravità, pari a 9,8 m/s^2. Problema: qual è il punto più alto raggiunto dal sasso? È il valore massimo assunto dalla funzione $h(t)$; cerchiamo dunque il valore di t per cui la curva $h(t)$ ha retta tangente orizzontale; in realtà *questa* curva è una parabola, e senza scomodare il calcolo infinitesimale sappiamo che ha valore massimo per t =ascissa del vertice, cioè $t = v_0/g$; perciò la quota massima raggiunta è $h(v_0/g) = \frac{1}{2} v_0^2/g$. È chiaro però che in problemi più generali potremmo trovare una funzione $h(t)$ "qualsiasi", e si pone il problema di avere un *metodo generale* per determinare il punto in cui la tangente è orizzontale, e prima ancora per calcolare il coefficiente angolare della retta tangente in un punto qualsiasi.

Un punto istruttivo di questa discussione è il seguente: senza il calcolo infinitesimale non si potrebbe non solo *risolvere* certi problemi, ma neppure *dar senso* ad essi (ad esempio, definire rigorosamente il concetto di retta tangente).

Che cos'è la velocità di un oggetto in moto?

Questo è un altro concetto che sembra ovvio ma non lo è. Consideriamo un punto materiale che si muove di moto vario, e proponiamoci di calcolare la sua velocità, nota la legge oraria del moto (ossia lo spazio percorso in funzione del tempo). Nella fisica galileiana l'unica velocità di cui si parla esplicitamente è la *velocità media*, ossia il rapporto tra lo spazio percorso in un certo intervallo di tempo e l'intervallo di tempo stesso; questo concetto è sufficiente per descrivere il moto uniforme. Ma in fenomeni come la caduta di un grave, l'oggetto cambia velocità a ogni istante. Che cos'è la velocità istantanea? Intuitivamente, è la velocità media in un intervallo di tempo brevissimo. Ma come si può definire rigorosamente? E come si calcola effettivamente?

Si noti che la parola "brevissimo" non ha alcun significato rigoroso, in matematica! Quello che si può pensare di fare, è calcolare la velocità media in un intervallo di tempo *variabile* Δt, ossia il rapporto tra lo spazio percorso nell'intervallo di tempo Δt e l'intervallo Δt stesso, e cercare di capire a quale valore si avvicina questa quantità quando Δt diventa sempre più piccolo (ma diverso da zero: se $\Delta t = 0$ il procedimento perde significato, dandoci un rapporto $0/0$). L'introduzione consapevole di questa nozione di velocità istantanea si deve a Newton (inizio 1700), ed è uno degli elementi di novità assoluta della fisica newtoniana rispetto a quella galileiana e pregalileiana. Per Newton, come per noi, la velocità è la *derivata rispetto al tempo* della funzione $s(t)$ ="spazio percorso nel tempo t", ovvero il limite a cui tende il rapporto tra lo spazio Δs percorso in un intervallo di tempo Δt e l'intervallo Δt, quando Δt diventa sempre più piccolo. Procediamo ancora di un passo. Galileo aveva già osservato che il concetto di velocità è sempre *relativo* ad un sistema di riferimento, e che "quiete" o "moto rettilineo uniforme" possono essere due descrizioni dello stesso fenomeno osservato da due sistemi di riferimento diversi (entrambi ugualmente "leciti": i cosiddetti *sistemi di riferimento inerziali*). Ne segue che l'effetto di una *forza* (o se vogliamo di una "causa fisica") non può essere la velocità. Qual è allora tale effetto? Newton è il primo ad affermarlo esplicitamente: tale effetto è l'*accelerazione*, ovvero la *velocità di variazione della velocità*. Questo concetto presuppone ovviamente quello di velocità istantanea, e introduce l'idea che il concetto matematico fisicamente interessante sia *la derivata della derivata della funzione* $s(t)$, ossia la derivata seconda di $s(t)$.

Il *calcolo differenziale* (ovvero lo studio della nozione di derivata), che è parte del calcolo infinitesimale, è quindi inscindibilmente legato alla nascita della scienza moderna.

2 DERIVATA DI UNA FUNZIONE

2.1 Derivata e retta tangente

Riprendiamo ora da un punto di vista quantitativo la discussione fatta nel paragrafo precedente sul concetto di retta tangente; arriveremo così alla definizione di *derivata*.

Cartelli stradali del tipo

indicano la "pendenza media" del percorso. Nel caso indicato la pendenza media è del 10%. Che cosa vuol dire? Significa che ad ogni avanzamento di 1 km corrisponde un innalzamento (o un abbassamento) di circa 100 m = 0,1 km (fig. 4.4).
Il 10% che indica la pendenza è il rapporto

$$\frac{\text{variazione quota}}{\text{variazione percorso}}$$

Figura 4.4.

che rappresenta il *tasso di variazione* della quota rispetto al percorso, relativamente al tratto di 1 km. Si noti che tasso di variazione positivo indica (in media!) innalzamento, mentre tasso negativo indica abbassamento.

Figura 4.5.

Se immaginiamo il nostro percorso idealmente descritto dal grafico di una funzione $y = f(x)$ possiamo pensare che le coordinate dei punti A e B (in fig. 4.5) siano rispettivamente $(x, f(x))$ e $(x + h, f(x + h))$. Allora avremo

$$\overline{AC} = h \quad \text{e} \quad \overline{CB} = f(x+h) - f(x)$$

e quindi

$$\frac{\text{variazione quota}}{\text{variazione percorso}} = \frac{f(x+h) - f(x)}{h}$$

L'ultimo rapporto prende il nome di *rapporto incrementale di f relativo all'intervallo* $[x, x + h]$. Geometricamente, essendo il triangolo ABC rettangolo, si ha che:

$$\frac{f(x+h) - f(x)}{h} = \operatorname{tg}\alpha = \text{ coefficiente angolare della retta } AB$$

D'altra parte, osservando la figura 4.5, si vede che la pendenza media da A a B non è un dato preciso in quanto, durante il percorso, nonostante la pendenza media sia positiva, vi sono tratti in discesa, a pendenza perciò negativa.

Per ovviare all'inconveniente e raggiungere un'adeguata precisione occorre considerare tratti di percorso sempre più piccoli, ossia far tendere la lunghezza del percorso a zero. Naturalmente, nel caso della pendenza stradale, un dato medio relativo a lunghezze dell'ordine del km può essere un'informazione sufficiente all'automobilista, ma vedremo più avanti in altri esempi che il passaggio al limite citato risulterà molto significativo. Ritornando al nostro esempio, consideriamo il rapporto:

$$\frac{f(x+h)-f(x)}{h} = \operatorname{tg}\alpha$$

e passiamo al limite (supponendo che esista) per $h \to 0$. Che cosa succede geometricamente? Il punto A, di coordinate $(x, f(x))$ rimane fisso, mentre il punto B, di coordinate $(x+h, f(x+h))$ si muove verso A, mantenendosi sul grafico di f (fig. 4.6).

Figura 4.6.

Contemporaneamente la retta AB varia la sua pendenza assestandosi su una posizione limite. Ne scaturiscono i seguenti fondamentali concetti:

1. la retta limite prende il nome di *retta tangente* al grafico di f nel punto di ascissa x;
2. la sua *pendenza* (o il suo *coefficiente angolare*) è data da $\operatorname{tg}\omega$ (fig. 4.6) e prende il nome di *derivata prima* di f nel punto x.

In generale, abbiamo la seguente definizione.

Definizione 4.1 Sia $f : (a,b) \longrightarrow \mathbb{R}$; f si dice *derivabile* in $x_0 \in (a,b)$ se esiste finito $\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h}$. Tale limite prende il nome di *derivata prima* (o semplicemente *derivata*) di f in x_0 e si indica con uno dei simboli seguenti:

$$f'(x_0) \qquad \left.\frac{df}{dx}\right|_{x=x_0} \qquad Df(x_0) \qquad \dot{f}(x_0)$$

$$\boxed{\lim_{h\to 0} \frac{f(x_0+h)-f(x_0)}{h} = f'(x_0)}$$

La retta di equazione:

$$\boxed{y = f(x_0) + f'(x_0)(x - x_0)} \tag{2.1}$$

si chiama *retta tangente* al grafico di f nel punto $(x_0, f(x_0))$.

Se f è derivabile in ogni punto di (a, b) è definita la funzione $f' : (a, b) \longrightarrow \mathbb{R}$, *derivata di* f, data da

$$x \mapsto f'(x)$$

In questo caso possiamo anche chiederci se la funzione $f'(x)$ sia a sua volta derivabile (in un punto o in tutto l'intervallo); in caso affermativo, chiameremo *derivata seconda di* f la derivata di f' e la indicheremo con uno dei simboli

$$f''(x_0) \qquad \left.\frac{d^2 f}{dx^2}\right|_{x=x_0} \qquad D^2 f(x_0) \qquad \ddot{f}(x_0)$$

Vedremo più avanti qual è il significato geometrico e l'utilizzo della derivata seconda. In modo del tutto analogo si definirà la derivata di ordine n, o *derivata n-esima*, indicata con i simboli:

$$f^{(n)}(x_0) \qquad \left.\frac{d^n f}{dx^n}\right|_{x=x_0} \qquad D^n f(x_0)$$

2.2 Altre interpretazioni della derivata

La definizione di derivata come limite del rapporto incrementale indica che essa deve considerarsi come *tasso di variazione puntuale o istantaneo*, anziché medio.

Questo modo di vedere è particolarmente fruttuoso in molte situazioni concrete come indicato qui di seguito.

- Un oggetto è in moto lungo una traiettoria rettilinea con legge oraria $s = s(t)$, $t \geq 0$. La funzione $t \mapsto s(t)$ indica lo spazio percorso al tempo t. Allora

$$\frac{ds}{dt} = \text{ velocità dell'oggetto}$$

- Analogamente, la velocità istantanea di variazione della velocità $v(t)$ dell'oggetto, rappresenta (per definizione!) l'accelerazione istantanea dell'oggetto stesso:

$$\frac{dv}{dt} = \textit{accelerazione} \text{ dell'oggetto}$$

Poiché a sua volta v era la derivata della funzione $s(t)$, si ottiene che

$$\text{accelerazione} = \frac{d}{dt}\left(\frac{ds}{dt}\right) = \frac{d^2 s}{dt^2}$$

Ecco un primo esempio di interpretazione della derivata seconda di una funzione. Se sull'oggetto agisce una forza $f = f(s, s')$, la seconda legge della dinamica si può scrivere nella forma

$$ms'' = f(s, s')$$

dove m è la massa dell'oggetto.

- Un pendolo oscilla in un piano verticale attorno a un asse, formando con questo un angolo $\theta = \theta(t)$, come in figura. In questo caso

$$\frac{d\theta}{dt} = \textit{velocità angolare} \text{ del pendolo}$$

Figura 4.7.

- In un filo percorso da corrente sia $q = q(t)$ la quantità di carica che attraversa una sezione del filo nell'istante t. Si ha:

$$\frac{dq}{dt} = \textit{intensità di corrente}$$

- Se $c = c(p)$ indica il costo di produzione di una quantità p (di un determinato bene), il rapporto tra $c(p+\Delta p)-c(p)$ e Δp indica la variazione del costo proveniente da una variazione (media) unitaria della produzione e prende il nome di *costo marginale medio*. Passando al limite per $\Delta p \to 0$ si ottiene il costo marginale; cioè, in questo caso

$$\frac{dc}{dp} = \text{costo marginale di produzione}$$

Invitiamo il lettore ad allungare la lista delle possibili interpretazioni della derivata con altri esempi (accelerazione angolare, densità di massa, ...).

2.3 Derivate di funzioni elementari

La tabella seguente contiene le derivate delle principali funzioni elementari.

f	f'	f	f'
1. c (costante)	0	11. e^x	e^x
2. x	1	12. $\log x$	$\frac{1}{x}$
3. x^2	$2x$	13. a^x $(a > 0)$	$a^x \log a$
4. $\frac{1}{x}$	$-\frac{1}{x^2}$	14. $\log_a x$ $(a > 0, a \neq 1)$	$\frac{1}{x \log a}$
5. $\sqrt{x}$	$\frac{1}{2\sqrt{x}}$ (per $x > 0$)	15. $\mathrm{Sh}x$	$\mathrm{Ch}x$
6. x^α $(\alpha \in \mathbb{R}, x > 0)$	$\alpha x^{\alpha-1}$	16. $\mathrm{Ch}x$	$\mathrm{Sh}x$
7. $\sin x$	$\cos x$	17. $\arcsin x$	$\frac{1}{\sqrt{1-x^2}}$
8. $\cos x$	$-\sin x$	18. $\arccos x$	$-\frac{1}{\sqrt{1-x^2}}$
9. $\mathrm{tg}\, x$	$1 + \mathrm{tg}^2 x = \frac{1}{(\cos x)^2}$	19. $\mathrm{arctg}\, x$	$\frac{1}{1+x^2}$
10. $\mathrm{cotg}\, x$	$-(1 + \mathrm{cotg}^2 x) = -\frac{1}{\sin^2 x}$		

La 1 è immediata; $2, 3, 4, 5$ sono casi particolari di 6, che abbiamo scritto esplicitamente perché si incontrano così frequentemente che meritano di essere ricordate. Proviamo ad esempio la:

3. $$\boxed{f(x) = x^2,\ f'(x) = 2x}$$

Si ha:

$$\frac{f(x+h) - f(x)}{h} = \frac{(x+h)^2 - x^2}{h} = \frac{x^2 + 2hx + h^2 - x^2}{h} = 2x + h$$

che tende a $2x$ per $h \to 0$; da qui, la formula.

6. $$\boxed{f(x) = x^\alpha;\ f'(x) = \alpha x^{\alpha-1}}$$

Sia $x > 0$. Scriviamo:

$$\frac{(x+h)^\alpha - x^\alpha}{h} = x^\alpha \frac{(1+\frac{h}{x})^\alpha - 1}{h} \sim x^\alpha \frac{\alpha \frac{h}{x}}{h} = \alpha x^{\alpha-1},$$

dove abbiamo usato il limite notevole $(1+\varepsilon(h)\ )^\alpha - 1 \sim \alpha\varepsilon(h)$ per $h \to 0$, con $\varepsilon(h) = h/x$ (cfr. la (3.8) del cap. 3).

7, 8. $$\boxed{f(x) = \sin x\,,\ f'(x) = \cos x} \qquad \boxed{f(x) = \cos x\,,\ f'(x) = -\sin x}$$

Si ha, utilizzando le formule di addizione

$$\frac{\sin(x+h) - \sin x}{h} = \frac{\sin x \cos h + \sin h \cos x - \sin x}{h} =$$
$$= \sin x \frac{\cos h - 1}{h} + \frac{\sin h}{h} \cos x \xrightarrow[h\to 0]{} \cos x$$

dove abbiamo usato i limiti notevoli $\frac{\sin h}{h} \to 1$ e

$$\frac{\cos h - 1}{h} \sim -\frac{1}{2}\frac{h^2}{h} = -\frac{1}{2}h \to 0$$

Analogamente si mostra la seconda formula. Omettiamo i dettagli.

11. $$\boxed{f(x) = e^x\,,\ f'(x) = e^x}$$

Si ha:

$$\frac{f(x+h) - f(x)}{h} = \frac{e^{x+h} - e^x}{h} = e^x \frac{e^h - 1}{h} \xrightarrow[h\to 0]{} e^x$$

essendo $\frac{e^h - 1}{h} \to 1$ per $h \to 0$ (limite notevole).

Dalla formula si ricava che la funzione $x \longmapsto e^x$ soddisfa l'equazione $y' = y$.

12. $$\boxed{f(x) = \log x\,,\ f'(x) = \frac{1}{x}}$$

Abbiamo:

$$\frac{f(x+h) - f(x)}{h} = \frac{\log(x+h) - \log x}{h} = \frac{\log(1+h/x)}{h} \sim \frac{h}{x} \cdot \frac{1}{h} = \frac{1}{x}$$

dove abbiamo usato la (3.8) del capitolo 3.

Le formule 9, 10 e 13-19 potranno essere dimostrate utilizzando le regole che vedremo nel prossimo paragrafo.

Le equazioni differenziali soddisfatte dalle funzioni esponenziali e trigonometriche
Osserviamo un fatto notevole che riguarda le funzioni esponenziali e trigonometriche. La 13 può essere riletta dicendo che le funzioni $f(x) = a^x$ soddisfano l'*equazione differenziale*

$$f'(x) = kf(x)$$

con k costante opportuna. La legge esponenziale governa quindi i fenomeni in cui *la velocità di crescita* (o *diminuzione*) *di una grandezza è proporzionale alla grandezza stessa.* Questa semplice legge si ritrova in molte leggi fisiche e questo è certamente uno dei motivi a cui le funzioni esponenziali devono la loro importanza.

La 7 e la 8 ci dicono invece che, posto $f(x) = \sin x$,

$$\frac{d^2 f}{dx^2}(x) = \frac{d}{dx}\left(\frac{df}{dx}\right)(x) = \frac{d}{dx}(\cos x) = -\sin x = -f(x)$$

Perciò la funzione $\sin x$ (e, come si vede con passaggi analoghi, anche la funzione $\cos x$) soddisfa l'equazione differenziale

$$f''(x) = -f(x)$$

Pertanto le funzioni sinusoidali governano i fenomeni in cui l'*accelerazione con cui varia una grandezza è uguale alla grandezza stessa, cambiata di segno.* Tali sono ad esempio molti fenomeni di tipo vibratorio.

Non è esagerato dire che *questo* è il motivo fondamentale per cui le funzioni trigonometriche sono così importanti nel calcolo infinitesimale e nella matematica applicata: un motivo che apparentemente non ha nulla a che vedere con la trigonometria, da cui queste funzioni sono nate! Approfondiremo questi fatti nel vol.2, trattando le equazioni differenziali.

Esempio

2.1 Calcoliamo l'equazione della retta tangente al grafico delle seguenti funzioni

$$f_1(x) = e^x\,, \quad f_2(x) = x^3$$

nei punti di ascissa $x = 2$.

Abbiamo $f_1(2) = e^2$ e

$$f_1'(x) = e^x \qquad f_1'(2) = e^2$$

e quindi l'equazione della retta è

$$y = f_1(2) + f_1'(2)(x-2) = e^2 + e^2(x-2)$$

Per l'altra funzione:

$$f_2(2) = 8\,, \quad f_2'(x) = 3x^2 \quad \text{(formula 6 della tabella con } \alpha = 3)$$

e quindi $f_2'(2) = 12$; pertanto si ottiene per la retta tangente, l'equazione

$$y = f_2(2) + f_2'(2)(x-2) = 8 + 12(x-2)$$

Riportiamo in figura 4.8 i grafici delle due funzioni, con le rispettive rette tangenti per $x = 2$:

Figura 4.8.

2.4 Punti angolosi, cuspidi, flessi a tangente verticale

Se una funzione f è derivabile in un punto x_0, nel punto di coordinate $(x_0, f(x_0))$ il grafico ha una retta tangente ben definita. Che cosa succede quando f non è derivabile in un punto? Vediamo alcuni esempi.

Punti angolosi

Sia $f(x) = |x|$. Essendo $f(x) = x$ per $x > 0$ e $f(x) = -x$ per $x < 0$, si ha $f'(x) = +1$ se $x > 0$ e $f'(x) = -1$ per $x < 0$, avendo f' il significato di coefficiente angolare. Nell'origine $x = 0$, occorre usare la definizione. Ora

$$\frac{f(h) - f(0)}{h} = \frac{|h|}{h}$$

e quindi, se $h \to 0^+$, $|h| = h$ e il limite del rapporto incrementale è 1, mentre se $h \to 0^-$, $|h| = -h$ e il limite è -1.

Si conclude che, non esistendo il limite del rapporto incrementale, *f non è derivabile in $x = 0$*. D'altra parte, ricordando il grafico di $f(x) = |x|$ si vede che la tangente nell'origine non è ben definita.

Figura 4.9. La funzione $|x|$ *non* è derivabile in $x = 0$.

Tuttavia i limiti destro e sinistro del rapporto incrementale di $|x|$ esistono finiti e in $(0, 0)$ il grafico presenta "un angolo".

La circostanza merita una definizione.

Definizione 4.2 Siano $f : (a, b) \longrightarrow \mathbb{R}$, $x_0 \in (a, b)$. Se esiste finito $\lim_{h \to 0^+} \frac{f(x_0+h)-f(x_0)}{h}$ $\left(\text{oppure } \lim_{h \to 0^-} \frac{f(x+h)-f(x_0)}{h}\right)$ allora f si dice *derivabile dalla destra* (oppure *dalla sinistra*); il limite si chiama *derivata destra* (oppure *sinistra*) e si indica con il simbolo $f'_+(x_0)$ (oppure $f'_-(x_0)$). ■

Nel caso in cui f sia continua e derivabile *da destra e da sinistra* (ma non derivabile) in x_0 si dice che f ha un *punto angoloso in* $x = x_0$. Dunque, $|x|$ ha un punto angoloso in $x = 0$.

Vale la pena ricordare la formula che esprime sinteticamente la derivata della funzione valore assoluto (fuori dall'origine):

$$|x|' = \operatorname{sgn}(x) = \begin{cases} 1 & \text{per } x > 0 \\ -1 & \text{per } x < 0 \end{cases}$$

Punti a tangente verticale. Cuspidi

Se f è continua in un punto x_0 e

$$\lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h} = +\infty \quad \text{oppure} \quad -\infty$$

f non è derivabile in x_0 ma, geometricamente, il grafico di f ha una retta tangente ben definita e parallela all'asse delle ordinate. Ammetteremo in tal caso la scrittura $f'(x_0) = +\infty$, $f'(x_0) = -\infty$ e parleremo di *flesso a tangente verticale* (fig. 4.10). Il concetto di flesso sarà definito più in generale in seguito (par. 5.2).

Figura 4.10.

Ad esempio, la funzione $f(x) = \sqrt[3]{x}$ ha un punto a tangente verticale in $x = 0$.

Consideriamo ora la funzione $f(x) = \sqrt[3]{|x|}$, il cui grafico è riportato in figura 4.11. In questo caso si ha $f'_+(0) = +\infty$, $f'_-(0) = -\infty$ e si dice che in $x = 0$, f ha una *cuspide*.

Figura 4.11. $f(x) = \sqrt[3]{|x|}$.

Definizione 4.3 Se f è continua in x_0 e $f'_+(x_0) = \pm\infty$, $f'_-(x_0) = \mp\infty$ si dice che f ha in x_0 una *cuspide*.

Nel caso misto in cui una delle due derivate è finita e l'altra infinita (con f continua) si parla ancora di punto angoloso.

Infine, se la funzione è definita solo per $x \geq x_0$ e in tal punto ha derivata (destra) infinita, diremo semplicemente che in tal punto *ha tangente verticale*, senza parlare né di cuspide né di flesso. Ad esempio, la funzione $\sqrt{x}$ ha un punto a tangente verticale in $x = 0$.

Continuità e derivabilità

Vale il seguente semplice ma importante

Teorema 4.1 *Se f è derivabile in un punto x_0 allora f è continua in x_0.*

Dimostrazione. Scriviamo

$$f(x_0 + h) - f(x_0) = \frac{f(x_0 + h) - f(x_0)}{h} \cdot h \sim f'(x_0)h \qquad \text{per } h \to 0\,.$$

Perciò $\lim_{h\to 0} [f(x_0 + h) - f(x_0)] = 0$ da cui $\lim_{h\to 0} f(x_0 + h) = f(x_0)$, che è la continuità di f in x_0.

Come conseguenza, se una funzione è *discontinua* in x_0, *non può* essere derivabile in x_0.

Viceversa, se f è continua in x_0, non necessariamente f è derivabile in x_0 come mostra $f(x) = |x|$ che è continua in $x = 0$ ma non ivi derivabile.

Consideriamo la funzione f periodica di periodo T che coincide con $\frac{2A}{T}|x|$ nell'intervallo $[-T/2, T/2]$ (fig. 4.12).

Figura 4.12.

Una funzione di questo genere modellizza *un'onda* o *segnale triangolare* di ampiezza A. Questa funzione presenta punti angolosi in $x = \pm\frac{T}{2}, \pm T, \pm\frac{3}{2}T$ ecc.

La sua derivata non è definita in questi punti ed è *costante a tratti*; precisamente, vale $\frac{2A}{T}$ per $0 < x < \frac{T}{2}$, vale $-\frac{2A}{T}$ per $\frac{T}{2} < x < T$ e si ripete con periodicità T, con un grafico illustrato in figura 4.13.

Figura 4.13.

La derivata del segnale triangolare di ampiezza $\frac{2A}{T}$ rappresenta quella che si chiama *onda quadra*.

Esercizi

1 Utilizzando la definizione di derivata, determinare il comportamento nell'origine delle seguenti funzioni (tangente orizzontale, cuspide, flesso a tangente verticale...):

$$x^{1/3} \qquad x^{4/3} \qquad x^{2/3} \qquad x^{5/3} \qquad x^{1/2} \qquad x^{3/2}$$

(*Questo argomento permette di completare la giustificazione del grafico delle funzioni potenza a esponente razionale o reale, che abbiamo descritto nel cap. 2, par. 3.1*).

2 Sia

$$f(x) = x \log x, \text{ per } x > 0$$

Dopo aver prolungato per continuità f anche in $x = 0$, calcolare la sua derivata destra in 0.

3 Sia

$$f(x) = e^{-1/x}, \text{ per } x > 0$$

Dopo aver prolungato per continuità (da destra) f anche in $x = 0$, calcolare la sua derivata destra in 0.

3 REGOLE DI CALCOLO DELLE DERIVATE

Vediamo ora la relazione tra l'operazione di derivata e le principali operazioni già note sulle funzioni; in particolare mostreremo la relazione tra:

$$\text{derivazione e} \begin{cases} \text{operazioni algebriche } (\pm, \cdot, /) \\ \text{composizione} \\ \text{inversione} \end{cases}$$

3.1 Algebra delle derivate

TEOREMA 4.2 *Siano $f, g : (a,b) \longrightarrow \mathbb{R}$, derivabili in (a,b); allora $f \pm g$, $f \cdot g$, f/g $(g \neq 0)$ sono derivabili in (a,b) e valgono le seguenti formule*

$$\boxed{(f \pm g)' = f' \pm g'} \tag{3.1}$$

$$\boxed{(f \cdot g)' = f' \cdot g + f \cdot g'} \tag{3.2}$$

$$\boxed{(f/g)' = (f' \cdot g - f \cdot g')/g^2} \tag{3.3}$$

In particolare, dalla (3.2) *si deduce*

$$\boxed{(k \cdot f)' = k \cdot f' \qquad k \text{ costante}} \tag{3.4}$$

essendo la derivata di una costante uguale a zero, e dalla (3.3) *si deduce, per $f = 1$,*

$$\boxed{\left(\frac{1}{g}\right)' = -\frac{g'}{g^2}} \tag{3.5}$$

La (3.2) si dice *regola di Leibnitz* e si estende al prodotto di n fattori:

$$(f_1 f_2 \cdots f_n)' = f_1' f_2 \cdots f_n + f_1 f_2' \cdots f_n + \ldots + f_1 f_2 \cdots f_n' .$$

Come si vedrà dalla dimostrazione, il teorema ha in realtà un carattere *puntuale*, ossia: se f e g sono derivabili in un punto $x_0 \in (a,b)$, allora in quel punto sono derivabili anche $f + g$, $f \cdot g, \ldots$, e valgono le formule scritte. Solitamente comunque il teorema si applica nella forma in cui l'abbiamo enunciato (cioè in situazioni in cui f e g sono derivabili in tutti i punti di un intervallo).

DIMOSTRAZIONE. Lasciamo per esercizio la dimostrazione di (3.1) e (3.4), molto semplici. Dimostriamo la (3.2): si ha, fissato $x \in (a,b)$

$$f(x+h)g(x+h) - f(x)g(x) = f(x+h)g(x+h) - f(x+h)g(x) + f(x+h)g(x) - f(x)g(x)$$

e quindi

$$\frac{f(x+h)g(x+h) - f(x)g(x)}{h} = f(x+h) \cdot \frac{g(x+h) - g(x)}{h} + g(x)\frac{f(x+h) - f(x)}{h} \longrightarrow$$
$$\longrightarrow f(x)g'(x) + f'(x)g(x)$$

poiché $f(x+h) \longrightarrow f(x)$ quando $h \to 0$, essendo f continua in quanto derivabile.

Proviamo ora la (3.5):

$$\frac{1}{h}\left[\frac{1}{g(x+h)} - \frac{1}{g(x)}\right] = \frac{g(x) - g(x+h)}{hg(x)g(x+h)} = -\frac{g(x+h) - g(x)}{h} \cdot \frac{1}{g(x)g(x+h)} \to -\frac{g'(x)}{g(x)^2}$$

dove si è sfruttato ancora il fatto che, essendo derivabile, g è continua, perciò $g(x+h) \to g(x)$ per $h \to 0$.

Notiamo infine che dalle (3.2) e (3.5) si deduce anche la (3.3), infatti:

$$\begin{aligned}\left(\frac{f}{g}\right)' &= \left(f\cdot\frac{1}{g}\right)' = \text{ per la (3.2) applicata a } f \text{ e } \frac{1}{g}\\ &= f'\cdot\frac{1}{g}+f\cdot\left(\frac{1}{g}\right)' = \text{ per la (3.5)}\\ &= f'\cdot\frac{1}{g}+f\cdot\left(-\frac{g'}{g^2}\right) = \frac{f'g-fg'}{g^2}, \text{ cioè la (3.3).}\end{aligned}$$

Esempi

3.1 Calcoliamo la velocità di un oggetto in moto rettilineo con legge oraria data da

$$s(t) = v_0 t + \frac{1}{2}gt^2$$

La derivata della funzione $v_0 t$ è v_0 mentre quella di $\frac{1}{2}gt^2$ è $\frac{1}{2}g\cdot 2t = gt$ (usando 2 volte la (3.4)). Usando ora la (3.1) si trova

$$v(t) = s'(t) = v_0 + gt$$

3.2 Calcoliamo la velocità di variazione del volume di una sfera rispetto al raggio. Essendo $V(r) = \frac{4}{3}\pi r^3$, la velocità richiesta è (regola (3.4))

$$V'(r) = \frac{4}{3}\pi(3r^2) = 4\pi r^2 \qquad (= \text{ superficie della sfera di raggio } r!)$$

3.3 Derivata della funzione tangente. Poiché $\operatorname{tg} x = \frac{\sin x}{\cos x}$, usando la formula (3.3) si trova

$$\frac{d}{dx}\operatorname{tg} x = \frac{(\cos x)^2+(\sin x)^2}{(\cos x)^2} = \frac{1}{(\cos x)^2} = 1+(\operatorname{tg} x)^2$$

Analogamente si calcola la derivata della funzione cotangente.

3.4 Se $c = c(p)$ è il costo di produzione di una quantità p di prodotto, il rapporto

$$\frac{c(p)}{p}$$

si chiama *costo medio* e rappresenta il *costo di una unità di prodotto.*

Per calcolare la derivata del costo medio, usiamo la formula (3.3); si trova

$$\frac{d}{dp}\frac{c(p)}{p} = \frac{c'(p)\cdot p - c(p)}{p^2}$$

3.2 Derivata di una funzione composta

Teorema 4.3 (Regola della catena) *Sia $g\circ f$ la composta di due funzioni f e g. Se f è derivabile in un punto x e g è derivabile in $y = f(x)$ allora $g\circ f$ è derivabile in x e vale la formula:*

$$\boxed{(g\circ f)'(x) = g'(f(x))\cdot f'(x)} \tag{3.6}$$

DIMOSTRAZIONE. Si ha $(g \circ f)(x+h) - (g \circ f)(x) = g(f(x+h)) - g(f(x))$.

Se poniamo $k = f(x+h) - f(x)$, $y = f(x)$, allora $f(x+h) = y + k$, e per la continuità di f, $h \to 0$ implica $k \to 0$. Con le nuove notazioni,

$$g(f(x+h)) - g(f(x)) = g(y+k) - g(y)$$

Osserviamo ora che la definizione di derivata

$$g'(y) = \lim_{k \to 0} \frac{g(y+k) - g(y)}{k}$$

si può riscrivere, per $k \neq 0$,

$$\frac{g(y+k) - g(y)}{k} = g'(y) + \varepsilon(k)$$

dove $\varepsilon(k)$ indica una quantità che tende a zero per $k \to 0$. Moltiplicando ambo i membri dell'equazione precedente per k si trova

$$g(y+k) - g(y) = g'(y) \cdot k + \varepsilon(k) \cdot k$$

relazione valida anche per $k = 0$. Dunque:

$$g\big(f(x+h)\big) - g\big(f(x)\big) = g'(y) \cdot k + \varepsilon(k) \cdot k$$

Dividendo per h, e osservando che $k/h \to f'(x)$ si ottiene la (3.6). ♦

La (3.6) si chiama *regola della catena*; usando le notazioni (di Leibniz) $\frac{df}{dx}$ e $\frac{dg}{dx}$ per le derivate di f e g e posto $w = g(y)$, la (3.6) acquista una forma più significativa:

$$\boxed{\frac{dw}{dx} = \frac{dw}{dy} \cdot \frac{dy}{dx}} \qquad \text{(come se } dy \text{ si semplificasse)} \tag{3.7}$$

La (3.7) esprime il fatto che il tasso di variazione di w rispetto a x è il prodotto dei tassi di variazione "intermedi", di w rispetto a y e di y rispetto a x.

Come suggerisce il nome di "regola della catena", la (3.7) può essere generalizzata alla composizione di un numero qualsiasi di funzioni, composte una con l'altra. Ad esempio per tre funzioni si ha

$$[f(g(h(x)))]' = f'(g(h(x))) \cdot g'(h(x)) \cdot h'(x)$$

Per usare questa regola, insieme alle altre dell'algebra delle derivate, occorre imparare a *vedere* una funzione complicata come composizione successiva di funzioni più semplici. Per individuare le componenti può essere utile immaginare come si calcola la funzione composta mediante una calcolatrice tascabile:

Esempi

3.5 Si voglia derivare

$$w(x) = (\sin x)^3$$

Per calcolarla, occorre inserire il valore di x, calcolare $\sin x$ e poi elevare tutto al cubo:

$$x \overset{\sin}{\to} \mapsto \sin x \overset{(\cdot)^3}{\to} \mapsto (\sin x)^3$$

Posto

$$f(x) = \sin x,\ g(y) = y^3$$

si ha allora $w(x) = g(f(x))$. Pertanto:

$$w'(x) = 3(\sin x)^2 \cdot \cos x$$

Usando la (3.7) si scriverebbe $y = \sin x$, $w = y^3$ e quindi

$$\frac{dw}{dx} = \frac{dw}{dy} \cdot \frac{dy}{dx} = 3y^2 \cdot \cos x = 3(\sin x)^2 \cdot \cos x$$

3.6 Derivata di $f(t) = A\sin(\omega t + \varphi)$ e di $g(t) = Ae^{-\alpha t}\cos(\omega t + \varphi)$

$$f'(t) \underset{\substack{\downarrow \\ \text{formula (3.4)}}}{=} A\frac{d}{dt}\sin(\omega t + \varphi) \underset{\substack{\downarrow \\ \text{regola della catena}}}{=} A\cos(\omega t + \varphi) \cdot \omega = A\omega\cos(\omega t + \varphi)$$

$$g'(t) \underset{\substack{\downarrow \\ \text{formula (3.4)}}}{=} A\frac{d}{dt}\left[e^{-\alpha t}\cos(\omega t + \varphi)\right] \underset{\substack{\downarrow \\ \text{formula (3.2)}}}{=} A\Big\{\frac{d}{dt}e^{-\alpha t} \cdot \cos(\omega t + \varphi) + e^{-\alpha t}\frac{d}{dt}\cos(\omega t + \varphi)\Big\} =$$

$$\underset{\text{(catena)}}{=} A\big\{e^{-\alpha t}(-\alpha) \cdot \cos(\omega t + \varphi) + e^{-\alpha t}(-\sin(\omega t + \varphi) \cdot \omega)\big\} =$$

$$= -Ae^{-\alpha t}\big\{-\alpha\cos(\omega t + \varphi) + \sin(\omega t + \varphi)\big\}$$

3.7 Possiamo ora completare la dimostrazione delle formule contenute nella tabella delle derivate di funzioni elementari.

$$(a^x)' = (e^{x\log a})' =$$

(usando le formule per la derivata di e^t e per la derivata della funzione composta)

$$= e^{x\log a} \cdot \log a = a^x \log a$$

$$(\log_a x)' = \left(\frac{\log x}{\log a}\right)' =$$

(usando le formule per la derivata di $\log x$ e per la derivata di $kf(x)$)

$$= \frac{1}{x\log a}$$

$$(\operatorname{Sh} x)' = \left(\frac{e^x - e^{-x}}{2}\right)' = \frac{1}{2}(e^x - (-e^{-x})) = \frac{e^x + e^{-x}}{2} = \operatorname{Ch} x$$

$$(\operatorname{Ch} x)' = \left(\frac{e^x + e^{-x}}{2}\right)' = \frac{1}{2}(e^x + (-e^{-x})) = \frac{e^x - e^{-x}}{2} = \operatorname{Sh} x$$

3.8 (*Derivate di funzioni del tipo* $f(x)^{g(x)}$). I passaggi per il calcolo di $(a^x)'$ si basano su un "trucco" di uso comune: riscrivere una funzione $f(x)^{g(x)}$ con $f(x) > 0$ nella forma seguente:

$$f(x)^{g(x)} = e^{g(x)\log f(x)}$$

Si è sfruttata la definizione di logaritmo $e^{\log A} = A$ e la proprietà dei logaritmi $\log[f(x)^{g(x)}] = g(x)\log f(x)$. A questo modo si può calcolare la derivata di una funzione di questo tipo:

$$[f(x)^{g(x)}]' = [e^{g(x)\log f(x)}]' = e^{g(x)\log f(x)} \cdot [g(x)\log f(x)]' =$$

$$= f(x)^{g(x)}\left[g'(x)\log f(x) + g(x)\frac{f'(x)}{f(x)}\right]$$

Ad esempio:

$$(x^x)' = (e^{x\log x})' = x^x[x\log x]' = x^x[\log x + 1]$$

3.9 (*Valore assoluto di una funzione*). Consideriamo una funzione del tipo:

$$|f(x)|$$

Sappiamo che il valore assoluto non è derivabile là dove il suo argomento si annulla. Tuttavia, nei punti in cui $f(x) \neq 0$, la derivazione di funzione composta dà:

$$|f(x)|' = \operatorname{sgn}(f(x)) \cdot f'(x) = \begin{cases} f'(x) & \text{per } f(x) > 0 \\ -f'(x) & \text{per } f(x) < 0 \end{cases}$$

In generale, ci aspettiamo che la funzione $|f(x)|$ presenti punti angolosi nei punti in cui $f(x)$ si annulla. Ad esempio, la funzione

$$e^{|x+1|}$$

ha un punto angoloso in $x = -1$.

3.10 (*Derivata di alcune funzioni logaritmiche*). Le seguenti situazioni si presentano di frequente:

$$(\log|x|)' = \frac{1}{|x|} \cdot \operatorname{sgn}(x) = \frac{1}{x} \quad \text{(senza modulo!)}$$

$$(\log|f(x)|)' = \frac{f'(x)}{f(x)}$$

(per la formula precedente e la derivazione di funzioni composte);

$$(\log(cx))' = (\log c + \log x)' = \frac{1}{x} \quad \text{(come la derivata di } \log x\text{!)};$$

$$\left(\log\left(\frac{ax+b}{cx+d}\right)\right)' = (\log(ax+b) - \log(cx+d))' = \frac{a}{ax+b} - \frac{c}{cx+d}$$

(ossia, talvolta conviene usare le proprietà dei logaritmi per trasformare una funzione logaritmica *prima* di calcolarne la derivata).

Il seguente esempio mette in luce le possibilità di calcolo connesse con la formula della catena.

Esempio

3.11 Un contenitore cilindrico con raggio di base $R = 1$ m e altezza 3 m è pieno d'acqua. Da un rubinetto posto in prossimità del fondo vengono prelevati 10 litri al minuto, (fig. 4.14).

Con quale velocità l'altezza dell'acqua decresce?

Sia h l'altezza (in decimetri) della colonna d'acqua e V il suo volume (in dm^3).

Vogliamo trovare dh/dt, sapendo che $dV/dt = -10\ \mathrm{dm}^3/\mathrm{min}$.

$$\frac{dV}{dt} = \frac{dV}{dh} \cdot \frac{dh}{dt}$$

essendo $V = 100\pi h\ \mathrm{dm}^3$ otteniamo $dV/dh = 100\pi\ \mathrm{dm}^2$ e infine

$$\frac{dh}{dt} = -\frac{10}{100\pi}\mathrm{dm/min} = -\frac{1}{10\pi}\ \mathrm{dm/min}$$

Figura 4.14.

Derivata logaritmica. Elasticità

Data $f > 0$, si chiama *derivata logaritmica* di f la derivata di $\log f$. In formule:

$$\boxed{\text{derivata logaritmica di } f = \frac{d}{dx}\log f(x) = \frac{f'(x)}{f(x)}} \tag{3.8}$$

La derivata logaritmica ha il significato di *tasso di incremento relativo* di f rispetto a x.

In casi concreti tale tasso è spesso più significativo del tasso assoluto f'. Ad esempio, un incremento annuo di capitale di 1 miliardo su un totale di 10 ha un effetto ben diverso di un incremento di 1 miliardo su 1000 miliardi. Nel primo caso il tasso di variazione relativo è $1/10$ nel secondo $1/1000$, mentre uguali sono gli incrementi assoluti.

Per i motivi sopraesposti, quando interessa visualizzare gli incrementi relativi si ricorre a grafici in scala *semilogaritmica*: sull'asse delle ascisse si collocano i valori di x, mentre su quello delle ordinate quelli di $\log f(x)$. A questo tipo di rappresentazione si ricorre anche quando $f(x)$ cresce così rapidamente da richiedere una compressione troppo elevata della scala (unità di misura) sull'asse delle ordinate o in quella delle ascisse. Basti pensare al grafico di $f(x) = e^x$ che in scala semilogaritmica coincide con ... la bisettrice del 1° quadrante:

Figura 4.15. a) Grafico di $f(x) = e^x$ in scala normale; b) grafico di e^x in una scala semilogaritmica.

Altra variante possibile è quella dei grafici in scala logaritmica, in cui invece di x sull'asse delle ascisse si collocano i valori di $\log x$; lo stesso si fa con f sull'asse delle ordinate.

La pendenza della retta tangente a un grafico in scala logaritmica (cioè la derivata di $\log f$ rispetto a $\log x$) rappresenta il tasso di variazione relativa di f rispetto a variazioni relative di x; questa quantità prende il nome di *elasticità* di f e si indica con $E(x)$.

Figura 4.16.

Per trovare l'espressione analitica di $E(x)$, osserviamo che, per il teorema di derivazione delle funzioni composte, si ha, posto $u = \log x$:

$$\frac{d\log f}{dx} = \frac{d\log f}{du} \cdot \frac{du}{dx}$$

ovvero

$$\frac{f'(x)}{f(x)} = E(x)\frac{1}{x}$$

da cui si ricava

$$E(x) = x\frac{f'(x)}{f(x)}$$

Esempio

3.12 Se $f(x) = x^\alpha$, $x > 0$,

$$E\,(x) = x\frac{\alpha x^{\alpha-1}}{x^\alpha} = \alpha$$

ossia: le potenze hanno elasticità costante. In realtà, si può mostrare che sono le uniche funzioni ad avere questa proprietà.

3.3 Derivata di funzione inversa

Teorema 4.4 *$f : (a, b) \to \mathbb{R}$ continua e invertibile in (a, b) e $g = f^{-1}$ la sua inversa, definita in $f\,(a, b)$. Supponiamo inoltre che esista $f'\,(x_0) \neq 0$ per un certo $x_0 \in (a, b)$. Allora g è derivabile in $y_0 = f\,(x_0)$ e*

$$g'\,(y_0) = \frac{1}{f'\,(x_0)}. \tag{3.9}$$

Altra variante possibile è quella dei grafici in scala logaritmica, in cui invece di x sull'asse delle ascisse si collocano i valori di $\log x$; lo stesso si fa con f sull'asse delle ordinate.

La pendenza della retta tangente a un grafico in scala logaritmica (cioè la derivata di $\log f$ rispetto a $\log x$) rappresenta il tasso di variazione relativa di f rispetto a variazioni relative di x; questa quantità prende il nome di *elasticità* di f e si indica con $E(x)$.

Figura 4.16.

Per trovare l'espressione analitica di $E(x)$, osserviamo che, per il teorema di derivazione delle funzioni composte, si ha, posto $u = \log x$:

$$\frac{d \log f}{dx} = \frac{d \log f}{du} \cdot \frac{du}{dx}$$

ovvero

$$\frac{f'(x)}{f(x)} = E(x)\frac{1}{x}$$

da cui si ricava

$$E(x) = x\frac{f'(x)}{f(x)}$$

Esempio

3.12 Se $f(x) = x^\alpha$, $x > 0$,

$$E\,(x) = x\frac{\alpha x^{\alpha-1}}{x^\alpha} = \alpha$$

ossia: le potenze hanno elasticità costante. In realtà, si può mostrare che sono le uniche funzioni ad avere questa proprietà.

3.3 Derivata di funzione inversa

Teorema 4.4 *$f : (a, b) \to \mathbb{R}$ continua e invertibile in (a, b) e $g = f^{-1}$ la sua inversa, definita in $f\,(a, b)$. Supponiamo inoltre che esista $f'\,(x_0) \neq 0$ per un certo $x_0 \in (a, b)$. Allora g è derivabile in $y_0 = f\,(x_0)$ e*

$$g'\,(y_0) = \frac{1}{f'\,(x_0)}. \tag{3.9}$$

DIMOSTRAZIONE. Sia:

$$f(x_0) = y_0; \qquad g(y_0) = x_0;$$

$$f(x_0 + h) = y_0 + k; \qquad g(y_0 + k) = x_0 + h$$

Consideriamo il rapporto incrementale di g in y_0:

$$\frac{g(y_0 + k) - g(y_0)}{k} = \frac{h}{f(x_0 + h) - f(x_0)}.$$

Se $k \neq 0$, $f(x_0 + h) - f(x_0) \neq 0$ e quindi anche $h \neq 0$; dunque l'ultimo quoziente si può riscrivere anche nella forma

$$\frac{1}{\dfrac{f(x_0 + h) - f(x_0)}{h}}.$$

Inoltre, per $k \to 0$ si ha $g(y_0 + k) \to g(y_0)$ perché g è continua, essendo l'inversa di una funzione continua su un intervallo (si veda il Cap.3, par. 4.3); d'altro canto $h = g(y_0 + k) - g(y_0)$, quindi per $k \to 0$ anche $h \to 0$, e per ipotesi

$$\frac{1}{\dfrac{f(x_0 + h) - f(x_0)}{h}} \to \frac{1}{f'(x_0)}$$

quindi esiste

$$\lim_{k \to 0} \frac{g(y_0 + k) - g(y_0)}{k} = \frac{1}{f'(x_0)}.$$

Osserviamo che, assumendo la derivabilità di f^{-1}, la (3.9) segue subito dall'identità $g(f(x)) = x$ e dalla regola della catena:

$$g'(f(x)) \cdot f'(x) = 1$$

da cui, se $f'(x) \neq 0$, la (3.9). La dimostrazione che abbiamo dato, tuttavia, è necessaria per *dedurre* la derivabilità di g dalle nostre ipotesi su f. ♦

La (3.9) ha un semplice significato geometrico, ricordando che i grafici di f e $g = f^{-1}$ sono simmetrici rispetto alla bisettrice $y = x$ (fig. 4.17).

Figura 4.17. Gli angoli α e β sono complementari ($\alpha + \beta = \pi/2$) e quindi $f'(x) = \operatorname{tg}\alpha = \operatorname{tg}\left(\frac{\pi}{2} - \beta\right) = \frac{1}{\operatorname{tg}\beta} = \frac{1}{g'(y)}$.

Con la notazione di Leibniz, posto $y = f(x)$, $x = g(y)$, (3.9) si scrive nella forma

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} \quad (!)$$

Si faccia attenzione al fatto che nella formula di derivazione della funzione inversa, le derivate f' e g' sono calcolate in *due punti diversi*: è questa la principale attenzione da avere nell'applicazione di questo teorema, come mostreranno i prossimi esempi.

Esempi

3.13 Derivata dell'arcotangente, dell'arcoseno e dell'arcocoseno.
Poniamo $y = \operatorname{tg} x$, $x = \operatorname{arctg} y$ con $x \in \left(-\frac{\pi}{2}, \frac{\pi}{2}\right)$, $y \in \mathbb{R}$.

$$\frac{dx}{dy} = \frac{1}{\frac{dy}{dx}} = \frac{1}{1 + (\operatorname{tg} x)^2} = \frac{1}{1 + y^2}$$

3.14 Poniamo ora $y = \sin x$, $x = \arcsin y$, con $x \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$, $y \in [-1, 1]$.
Poiché per quei valori di x si ha $\cos x = \sqrt{1 - (\sin x)^2} = \sqrt{1 - y^2}$:

$$\frac{dx}{dy} = \frac{1}{\cos x} = \frac{1}{\sqrt{1 - y^2}}$$

3.15 Analogamente, se $y = \cos x$, $x = \arccos y$ con $x \in [0, \pi]$, $y \in [-1, 1]$,

$$\frac{dx}{dy} = \frac{1}{-\sin x} = \frac{1}{-\sqrt{1 - y^2}}$$

Si osservi che le funzioni $\arcsin x, \arccos x$, pur essendo definite e continue in $[-1, 1]$, non sono derivabili agli estremi dell'intervallo: precisamente, presentano in questi punti tangente verticale.

L'utilità del teorema di derivazione della funzione inversa consiste nel fatto che permette di calcolare la derivata di g anche in situazioni in cui g non si sa scrivere esplicitamente:

Esempio

3.16 Sia $f(x) = x + e^x$. La funzione è strettamente crescente in tutto $\mathbb{R}$, dunque invertibile; sia g la sua inversa. Calcoliamo, ad esempio, $g'(y_0)$ per $y_0 = f(0) = 1$. Si ha:

$$f'(x) = 1 + e^x; f'(0) = 2 \neq 0; g'(1) = \frac{1}{f'(0)} = \frac{1}{2}.$$

Esercizi

In base alle regole di calcolo delle derivate e alla tabella delle derivate delle funzioni elementari, calcolare la derivata delle seguenti funzioni:

4. $3x^4 + 5x + x^{3/2} - 2x^{-3}$
5. $\log|x|$, $\log 3x$, $\log\left|\frac{x+2}{3-x}\right|$
6. $e^{-3x}\left(x^2 + 2x - 1\right)$
7. $\frac{1}{a}\operatorname{arctg}\frac{x}{a}$, $a > 0$

8 $x \log x$

9 $\operatorname{arctg} \frac{1+x}{1-x}$

10 $e^{2x}(2\sin 3x - 4\cos 3x)$

11 $e^{\frac{x+2}{x-3}}$

12 $\operatorname{cotg} x$, $\operatorname{Th} x$, $\operatorname{Coth} x$

13 2^{x^2+3x}

14 $\operatorname{SettSh} x$, $\operatorname{SettCh} x$ (*Suggerimento*: usare il teorema sulla derivata della funzione inversa)

15 $\log_2 |3x|$

16 $\frac{x^2+3x-2}{2x+1}$

17 $x^{x \log x}$

18 $\frac{ax+b}{cx+d}$

19 $\log |\log x|$

Scrivere l'equazione della retta tangente al grafico di $y = f(x)$ nel punto $(x_0, f(x_0))$:

20 $f(x) = \sin x$, $x_0 = \frac{\pi}{3}$

21 $f(x) = (x \log |x|)^3$, $x_0 = -1$

22 $f(x) = 3x^2 + 2x + 1$, $x_0 = 2$

23 $f(x) = \cos \log x$, $x_0 = e^{\frac{\pi}{2}}$

24 $f(x) = \log x$, $x_0 = 1$

25 $f(x) = e^{x^2}$, $x_0 = \log 2$

26 $f(x) = a^x$, $x_0 = 2$

27 $f(x) = e^{-|x|}$, $x_0 = -1$

28 Qual è il tasso di variazione del volume di una sfera rispetto al suo raggio? E rispetto all'area della sua superficie?

29 In un triangolo isoscele ABC (fig. 4.18) il vertice C si muove perpendicolarmente alla base AB in modo che l'area del triangolo cresca ad una velocità di 4 cm^2/s. La base AB è lunga 3 cm. A quale velocità cresce l'altezza CH? E il lato CB?

Figura 4.18.

30 Un parallelepipedo di base 1 m × 2 m e altezza 5 m è pieno d'acqua. Da un rubinetto posto in prossimità del fondo vengono prelevati 20 litri al minuto. Con quale velocità l'altezza dell'acqua decresce?

31 Una funzione è nota dal suo grafico (fig. 4.19); tracciare con buona approssimazione il grafico della sua derivata.

Figura 4.19.

Calcolare la derivata delle seguenti funzioni, dove esiste. Studiare i punti di non derivabilità, stabilendo se si tratta di punti angolosi (in questo caso, calcolare la derivata destra o sinistra), punti di cuspide, flessi a tangente verticale, punti a tangente verticale, tracciando un grafico locale della funzione in quei punti. Prima di eseguire il calcolo della derivata, cercare di prevedere quali saranno i punti di non derivabilità, in base alla forma della funzione.

32 $\left|x^2+3x-4\right|$

33 $\left|x^3\right|$

34 $\sqrt[3]{2x^2+x-1}$

35 $e^{-|x|}$

36 $e^{\sqrt{x}}$

37 $e^x\sqrt[3]{\frac{x+2}{x-3}}$

38 $\log^2\left(1+\sqrt[3]{x}\right)$

Calcolare $f'(0)$ in base alla definizione. Calcolare poi $f'(x)$ per $x \neq 0$, e stabilire se la derivata prima è continua in $x=0$:

39
$$f(x)=\begin{cases} x^2\sin\frac{1}{x} & \text{per } x>0 \\ 0 & \text{per } x<0 \end{cases}$$

40
$$f(x)=\begin{cases} e^{-1/x^2} & \text{per } x>0 \\ x^2 & \text{per } x\leq 0 \end{cases}$$

4 IL TEOREMA DEL VALOR MEDIO E LE SUE CONSEGUENZE

4.1 Punti stazionari. Massimi e minimi locali

Uno degli usi più proficui del calcolo differenziale consiste nella ricerca dei massimi e minimi (in breve, nell'*ottimizzazione*) di una funzione. Ecco alcuni problemi di ottimizzazione in ambiti diversi.

- *Ottica geometrica.* Le leggi della riflessione e della rifrazione si possono trovare utilizzando il "principio di Fermat": "il percorso scelto" da un raggio luminoso per collegare due punti A e B è quello che richiede il *minor* tempo di percorrenza.

- *Geometria elementare.* Inscrivere in un cono circolare retto di altezza h e raggio di base R, un cilindro di volume massimo.

- *Econometria elementare.* Sia $c(p)$ il costo di produzione di una quantità p di un certo prodotto. L'efficienza della produzione si può misurare col rapporto $\frac{c(p)}{p}$ = costo medio per unità di prodotto. Per raggiungere il massimo di efficienza occorre dunque minimizzare la funzione $\frac{c(p)}{p}$.

- *Dalla vita comune!.* Un tubo di lunghezza 4 m (sezione piccola) deve essere trasportato attraverso il cunicolo raffigurato in figura 4.20. È possibile il trasporto?

Figura 4.20. Passerà il tubo lungo 4 m?

Torniamo alla teoria. Richiamiamo le nozioni di massimo e minimo assoluti (o globali), per una funzione $f : [a, b] \longrightarrow \mathbb{R}$:

Definizione 4.4 Si dice che M è *massimo* di f in $[a, b]$ e $x_0 \in [a, b]$ è *punto di massimo* se

$$f(x_0) = M \geq f(x)\,, \quad \text{per ogni } x \in [a, b]$$

Analoga definizione per il *minimo*. Vi sono altri tipi di estremo (massimo o minimo) e cioè gli *estremi locali*.

Definizione 4.5 Si dice che M è *massimo locale* (o relativo) per f e che x_0 è *punto di massimo locale* se:

$$\begin{aligned}&\textit{esiste un intervallo } (x_0 - \delta, x_0 + \delta) \textit{ tale che } M = f(x_0) \geq f(x)\\ &\textit{per ogni } x \in (x_0 - \delta, x_0 + \delta) \cap [a, b]\end{aligned}$$

Analogamente per un *minimo locale*.

Notiamo espressamente che:

- il minimo e il massimo globale di f (se esistono) sono *unici* (naturalmente i punti di max e min possono essere più di uno);
- massimi e minimi locali possono essere più di uno. Evidentemente ogni estremo globale è anche locale.

Le seguenti figure illustrano varie situazioni.

- La funzione in figura 4.21 a) presenta:
  - massimo globale $M = f(x_2)$; x_2 unico punto di massimo globale.
  - minimo globale $m = f(a)$; a unico punto di minimo globale.
  - un massimo locale (non globale) in $x = x_0$.
  - due minimi locali (non globali) in $x = x_1$ (che è anche un punto angoloso) e $x = b$ (estremo destro di $[a, b]$).

Figura 4.21.

- La funzione in figura 4.21 b) presenta:
  - un massimo locale in $x = x_0$ (punto di discontinuità a salto).
  - minimo globale $m = f(b) = f(x_1)$; b e x_1 sono punti di minimo globale.

  Il massimo globale di f non esiste (si ha $\lim_{x \to a^+} f(x) = +\infty$).

Le figure mostrano che in un punto di estremo (locale o globale) f può non essere derivabile ed essere perfino discontinua. Se però $f : [a, b] \to \mathbb{R}$ è derivabile in un punto x_0 che sia di massimo o minimo locale e che sia diverso da a e da b, allora in x_0 la derivata si annulla, ossia la tangente al grafico in $(x_0, f(x_0))$ è orizzontale (fig. 4.21). Precisamente:

Teorema 4.5 (di Fermat) *Sia $f : [a, b] \longrightarrow \mathbb{R}$, derivabile in $x \in (a, b)$. Se x è punto di estremo locale allora*

$$f'(x) = 0$$

Dimostrazione. Sia, ad esempio, x punto di max locale. Allora, per z abbastanza vicino a x, si ha $f(z) \leq f(x)$. Perciò:

$$z < x \implies \frac{f(z) - f(x)}{z - x} \geq 0 \quad \text{e quindi} \quad f'_-(x) = \lim_{z \to x^-} \frac{f(z) - f(x)}{z - x} \geq 0$$

(abbiamo applicato il teorema della permanenza del segno, cap. 3, par. 3.1). D'altra parte

$$z > x \implies \frac{f(z) - f(x)}{z - x} \leq 0 \quad \text{e quindi} \quad f'_+(x) = \lim_{z \to x^+} \frac{f(z) - f(x)}{z - x} \leq 0$$

Essendo f derivabile in x, si ha $f'(x) = f'_-(x) = f'_+(x) = 0$. ◆

Punti stazionari

I punti in cui f' si annulla, si dicono *punti stazionari* per f. Abbiamo appena visto che, se x non si trova agli estremi dell'intervallo nel quale f è definita, allora

$$\boxed{x \quad \text{di estremo locale} \quad \Longrightarrow \quad x \quad \text{stazionario}}$$

Vi possono però essere punti stazionari che *non* sono di estremo. Ad esempio la funzione $f(x) = x^3$ ha $f'(x) = 3x^2$ che si annulla nell'origine, ma $x = 0$ non è punto di estremo (vedi fig. 4.22).
Si tratta di un punto di *flesso* (o di inflessione) a tangente orizzontale come vedremo più avanti.

Figura 4.22. $y = x^3$ ha un punto di flesso in $x = 0$, stazionario.

4.2 Teorema del valor medio. Test di monotonia

Supponiamo che per raggiungere una località B, partendo da una località A distante 160 km da B, si impieghino 2 h. La velocità media durante il percorso è dunque di 80 km/h. Se il viaggio non ha subìto interruzioni o altre irregolarità, certamente in qualche istante la velocità è stata esattamente di 80 km/h.

Cioè, in qualche istante:

$$\text{velocità media} = \text{velocità istantanea}$$

È concettualmente questo il contenuto del teorema del valor medio, che come vedremo, è denso di conseguenze.

Teorema 4.6 (del valor medio o di Lagrange) *Sia f derivabile in (a, b) e continua in $[a, b]$ (cioè continua fino agli estremi dell'intervallo). Allora*

$$\boxed{\textit{esiste} \quad c \in (a, b) \quad \textit{tale che} \quad \frac{f(b) - f(a)}{b - a} = f'(c)} \tag{4.1}$$

Il significato geometrico del teorema è illustrato nella figura seguente.

Figura 4.23.

Si ha:

$$\frac{f(b)-f(a)}{b-a} = \text{pendenza della retta } AB$$

$$f'(c) = \text{pendenza della retta tangente al grafico di } f \text{ nel punto } (c, f(c)).$$

La (4.1) esprime dunque il fatto che nel punto $(c, f(c))$ la tangente al grafico di f è parallela alla retta AB. In figura 4.23 esistono due di tali punti, di ascissa c_1, c_2.

DIMOSTRAZIONE DEL TEOREMA DEL VALOR MEDIO. Osserviamo che la retta AB ha equazione

$$y = f(a) + \frac{f(b)-f(a)}{b-a}(x-a)$$

e consideriamo la funzione

$$w(x) = f(x) - \left[f(a) + \frac{f(b)-f(a)}{b-a}(x-a)\right]$$

È facile verificare che: $w(a) = w(b) = 0$, w è continua in $[a,b]$ e w è derivabile in (a,b).

Poiché

$$w'(x) = f'(x) - \frac{f(b)-f(a)}{b-a}$$

la (4.1) equivale a dimostrare che esiste $c \in (a,b)$ tale che $w'(c) = 0$.

Essendo w continua in $[a,b]$, per il teorema di Weierstrass esistono due punti x_1 e x_2 in $[a,b]$ tali che

$$\begin{aligned} f(x_1) &= \text{ massimo di } f \text{ in } [a,b] = M \\ f(x_2) &= \text{ minimo } \quad \text{di } f \text{ in } [a,b] = m \end{aligned}$$

Se $M = m$, allora $w(x)$ è costante in $[a,b]$, e quindi $w'(x) = 0 \ \forall x \in [a,b]$.
Se $M > m$, almeno uno dei due punti x_1, x_2 non si trova agli estremi dell'intervallo, essendo $w(a) = w(b) = 0$.

Il teorema di Fermat implica allora che nel punto di massimo o minimo che risulta interno (eventualmente entrambi) la derivata di w si annulla e il teorema è così dimostrato.

Esempi

4.1 Sia $f(x) = x^2$. Allora $f'(x) = 2x$ e il teorema afferma che in ogni intervallo $[a, b]$ esiste un numero c tale che

$$\frac{b^2 - a^2}{b - a} = 2c \quad \text{da cui}$$

$$\boxed{c = \frac{a+b}{2} = \textit{media aritmetica} \text{ di } a \text{ e } b}$$

Cioè: ogni corda AB della parabola $y = x^2$ è parallela alla tangente nel punto di ascissa uguale alla media aritmetica delle ascisse di A e B.

Figura 4.24.

4.2 Sia $f(x) = \frac{1}{x}$. Allora $f'(x) = \frac{-1}{x^2}$ e dal teorema si deduce l'esistenza di $c \in [a, b]$ tale che

$$\frac{\frac{1}{b} - \frac{1}{a}}{b - a} = -\frac{1}{c^2}, \quad \text{ossia} \quad \frac{1}{ab} = \frac{1}{c^2}$$

da cui

$$\boxed{c = \sqrt{ab} = \textit{media geometrica} \text{ di } a \text{ e } b}$$

Cioè: ogni corda AB dell'iperbole $y = \frac{1}{x}$ è parallela alla tangente nel punto di ascissa uguale alla *media geometrica* delle ascisse di A e B.

Si noti che $\sqrt{ab} \leq \frac{a+b}{2}$ e cioè che

$$\boxed{\text{media geometrica} \leq \text{media aritmetica}}$$

Figura 4.25.

Ricordando il significato geometrico di derivata, si deduce subito che se una funzione derivabile è crescente oppure decrescente in un intervallo (a, b) la sua derivata è ≥ 0 oppure ≤ 0 rispettivamente. Il teorema di Lagrange permette di dimostrare anche il viceversa. Le due implicazioni sono contenute nel prossimo:

Teorema 4.7 (Test di monotonia) *Sia $f : (a, b) \longrightarrow \mathbb{R}$, derivabile. Allora*

$$\begin{aligned} f \ \textit{crescente} &\iff f'(x) \geq 0 \\ & \qquad\qquad\qquad\qquad \forall\, x \in (a, b) \\ f \ \textit{decrescente} &\iff f'(x) \leq 0 \end{aligned}$$

Dimostrazione. Infatti, considerata $f : (a, b) \longrightarrow \mathbb{R}$ e due punti qualunque $x, z \in (a, b)$

$$f \text{ è } \begin{matrix} \text{crescente} \\ \text{decrescente} \end{matrix} \iff \frac{f(z) - f(x)}{z - x} \text{ è } \begin{matrix} \geq 0 \\ \leq 0 \end{matrix}$$

Passando al limite per $z \to x$, per il teorema della permanenza del segno, dalle due precedenti relazioni si ottiene

$$\begin{array}{lll} f \text{ crescente} & \Longrightarrow & f'(x) \geq 0 \\ & & \qquad\qquad \forall x \in (a,b) \\ f \text{ decrescente} & \Longrightarrow & f'(x) \leq 0 \end{array}$$

Viceversa, sia, ad esempio, $f'(x) \geq 0$ per ogni $x \in (a,b)$, e proviamo che allora f è crescente in (a,b). Prendiamo dunque due punti qualsiasi $x_1, x_2 \in (a,b)$, $x_1 < x_2$, e mostriamo che $f(x_1) \leq f(x_2)$. Infatti, applicando il teorema di Lagrange ad f sull'intervallo $[x_1, x_2]$ abbiamo che esiste $c \in (x_1, x_2)$ tale che

$$\frac{f(x_2) - f(x_1)}{x_2 - x_1} = f'(c)$$

Poiché $f'(c) \geq 0$ e $x_2 - x_1 > 0$, ne segue $f(x_2) - f(x_1) \geq 0$, cioè la tesi. ◆

Con la stessa dimostrazione si vede anche che: se $f'(x) = 0$ per ogni $x \in (a,b)$, allora f è costante in (a,b). Poiché l'implicazione inversa (se f è costante in (a,b) allora ha derivata nulla) è ovvia, risulta dimostrata la seguente

Proposizione 4.1 (Caratterizzazione delle funzioni a derivata nulla) *Sia $f : (a,b) \to \mathbb{R}$. Allora*

$$f' = 0 \text{ in } (a,b) \Leftrightarrow f \text{ è costante in } (a,b)$$

Un errore da evitare è usare la proposizione 4.1 su insiemi più generali degli intervalli.

Esempio

4.3 Consideriamo la funzione

$$f(x) = \operatorname{arctg} x + \operatorname{arctg} \frac{1}{x}, \quad \text{per } x \neq 0$$

e calcoliamo

$$f'(x) = \frac{1}{1+x^2} + \frac{1}{1+\frac{1}{x^2}}\left(-\frac{1}{x^2}\right) = 0 \text{ per ogni } x \neq 0$$

Si può applicare la proposizione 4.1 a f e concludere che f è costante? La risposta è no: si può solo affermare che è costante su ciascuno dei due intervalli $(0, +\infty)$ e $(-\infty, 0)$. Per sapere quanto vale, è sufficiente calcolare f in un punto "comodo" di ciascun intervallo, per esempio:

$$f(1) = \operatorname{arctg} 1 + \operatorname{arctg} 1 = 2 \cdot \frac{\pi}{4} = \frac{\pi}{2}; \text{ dunque } f(x) = \frac{\pi}{2} \text{ per ogni } x > 0$$

$$f(-1) = \operatorname{arctg}(-1) + \operatorname{arctg}(-1) = -\frac{\pi}{2}; \text{ dunque } f(x) = -\frac{\pi}{2} \text{ per ogni } x < 0$$

Abbiamo dunque dimostrato l'identità:

$$\operatorname{arctg} x + \operatorname{arctg} \frac{1}{x} = \begin{cases} \frac{\pi}{2} & \text{se } x > 0 \\ -\frac{\pi}{2} & \text{se } x < 0 \end{cases}$$

Ricerca di massimi e minimi

Supponiamo di avere la solita funzione $f : [a, b] \longrightarrow \mathbb{R}$ e di volerne cercare i massimi e i minimi.

Se f è derivabile si può procedere nel modo seguente.

Passo 1. Si calcolano $f(a)$ e $f(b)$.

Passo 2. Si calcola $f'(x)$ e si risolve l'equazione

$$f'(x) = 0$$

In tal modo si trovano i punti stazionari, tra i quali vi sono gli eventuali punti di estremo locale, interni a (a, b).

Passo 3. Se non vi sono punti stazionari, $f(a)$ o $f(b)$ sono punti di estremo globale. Se viceversa $x = x_0$ è punto stazionario, occorre stabilirne la natura. A tale scopo si può studiare il segno di f' in un intorno di x_0, ricordando che $f' \geq 0$ implica f crescente, $f' \leq 0$ implica f decrescente. I casi che si presentano più comunemente sono illustrati qui di seguito:

a) x_0 è punto di max

b) x_0 è punto di min

$f'(x_0) = 0$

a) x_0 è punto di flesso

b) x_0 è punto di flesso

Passo 4. Trovati gli eventuali punti di estremo locale, si calcola il valore di f in questi punti e lo si confronta con $f(a)$ e $f(b)$.

Vediamo degli esempi numerici

Esempi

4.4 Sia $f(x) = xe^{-x^2}$ con $x \in [0, 2]$.

1. $f(0) = 0$, $f(2) = 2e^{-4}$
2. $f'(x) = e^{-x^2} + xe^{-x^2}(-2x) = e^{-x^2}(1 - 2x^2)$

$$f'(x) = 0 \iff 1 - 2x^2 = 0 \implies x = \pm\frac{1}{\sqrt{2}}$$

Solo $x_0 = \frac{1}{\sqrt{2}} \in [0, 2]$ e perciò è questo l'unico punto stazionario.

3. Studiamo ora il segno di f', vicino a $x_0 = \frac{1}{\sqrt{2}}$:

$$f'(x) \geq 0 \text{ per } 2x^2 \leq 1 \qquad -\frac{1}{\sqrt{2}} \leq x \leq \frac{1}{\sqrt{2}}$$

Si deduce quindi che $x_0 = \frac{1}{\sqrt{2}}$ è punto di massimo locale.

4. $f(\frac{1}{\sqrt{2}}) = \frac{1}{\sqrt{2}}e^{-1/2} = \frac{1}{\sqrt{2e}}$ è maggiore sia di $f(0) = 0$ che di $f(2) = \frac{2}{e^4}$. Si conclude quindi che:

$$f(0) = 0 \quad \text{è minimo globale}$$
$$f\left(\frac{1}{\sqrt{2}}\right) \quad \text{è massimo globale}$$

Si noti che in questo caso, il fatto che esista un solo punto stazionario con ordinata maggiore di quella agli estremi dell'intervallo, permette di trarre le stesse conclusioni anche senza lo studio del segno della derivata prima.

Un grafico qualitativo di f è il seguente

Figura 4.26. Grafico in $[0, 2]$ di $f(x) = xe^{-x^2}$; si noti che $f(x) \sim x$ per $x \to 0$.

Se volessimo studiare i massimi e minimi locali di $f(x)$ su tutto $\mathbb{R}$, il segno della derivata prima ci direbbe che

$$x = \frac{1}{\sqrt{2}} \text{ è punto di massimo locale;} \quad x = -\frac{1}{\sqrt{2}} \text{ è punto di minimo locale.}$$

4.5 (*Figura di diffrazione della luce attraverso una fenditura*). Un fascio di luce che attraversa una piccola fenditura la cui larghezza è dello stesso ordine di grandezza della lunghezza d'onda della luce, produce, su uno schermo su cui incide, una figura di interferenza. L'intensità luminosa in un punto dello schermo è data dalla funzione:

$$I = I_0 \left(\frac{\sin \frac{\varphi}{2}}{\frac{\varphi}{2}} \right)^2$$

dove I_0 è l'intensità massima (che si ottiene nel punto centrale) e l'angolo φ è collegato alla differenza di fase tra le due onde luminose che passano ai due estremi della fenditura. Cerchiamo i massimi e minimi di I in funzione di φ. Anzitutto, conviene porre $t = \varphi/2$ e cercare massimi e minimi della funzione

$$I = I_0 \left(\frac{\sin t}{t} \right)^2$$

(questo significa semplicemente cambiare scala sull'asse φ; basterà ricordare poi che $\varphi = 2t$). Il grafico qualitativo di questa funzione può essere tracciato facilmente, senza calcoli: la funzione $\frac{\sin t}{t}$ è pari, vale 1 in 0, ha infinite oscillazioni smorzate per $t \to \infty$; di conseguenza la funzione $I = I_0 \left(\frac{\sin t}{t}\right)^2$ è pari, non negativa, vale I_0 in 0, ed ha infinite oscillazioni smorzate:

Figura 4.27.

Vediamo quindi che I ha infiniti punti di minimo, nei punti in cui si annulla, $t = k\pi$ (cioè $\varphi = 2k\pi$), e infiniti punti di massimo, di cui il massimo assoluto è per $t = 0$ ($\varphi = 0$), e vale I_0. Il problema è determinare gli altri punti di massimo relativo. Questi si trovano risolvendo l'equazione

$$\frac{dI}{dt} = 0 \quad \text{ossia} \quad I_0 2\frac{\sin t}{t}\left(\frac{t\cos t - \sin t}{t^2}\right) = 0$$

Il fattore $\sin t$ si annulla nei punti di minimo che già conosciamo; i punti di massimo sono dunque le soluzioni di

$$t\cos t - \sin t = 0$$

equazione che non può essere risolta in modo esatto, ma che si studia facilmente con un *confronto grafico*: riscritta nella forma

$$t = \operatorname{tg} t$$

possiamo osservare dal grafico delle funzioni $t, \operatorname{tg} t$, dove sono collocate le soluzioni (v. fig. 4.28).

Figura 4.28.

Il primo punto positivo di massimo si trova poco prima di $\frac{3}{2}\pi$; il secondo poco prima di $\frac{5}{2}\pi$, e così via; più ci si allontana dall'origine, più questi punti tendono a coincidere esattamente con i valori $t = \frac{3}{2}\pi + k\pi$ (quindi $\varphi = 3\pi + 2k\pi$). Dunque ogni punto di massimo si trova approssimativamente a metà strada tra due punti di minimo successivi. I valori dei massimi relativi saranno all'incirca

$$I\left(\frac{3}{2}\pi + k\pi\right) = \frac{I_0}{(\frac{3}{2}\pi + k\pi)^2}$$

Una localizzazione più precisa delle soluzioni dell'equazione $t = \operatorname{tg} t$ si potrà dare con i metodi esposti nel paragrafo 7.5.

Successioni monotone

Come abbiamo visto nel capitolo 3, le successioni monotone hanno importanti proprietà. D'altro canto, non sempre è facile dimostrare la monotonia di una successione per via algebrica. Il calcolo differenziale ci offre un metodo utile.

Esempio

4.6 Sia

$$a_n = \frac{\log n}{n}$$

Sappiamo che $a_n \geq 0$ e $a_n \to 0$ (per confronto tra infiniti). Chiediamoci se la successione è monotona decrescente. Per definizione, ciò significa che

$$a_{n+1} \leq a_n \quad \text{ossia} \quad \frac{\log(n+1)}{n+1} \leq \frac{\log n}{n}$$

Poiché al crescere di n sia il numeratore che il denominatore crescono, non è facile provare questa disuguaglianza. D'altro canto, sia

$$f(x) = \frac{\log x}{x} \qquad \text{per } x > 0$$

(Ora la variabile x non è un intero, ma un numero reale!). Calcoliamo:

$$f'(x) = \frac{1 - \log x}{x^2} \leq 0 \qquad \text{per } x \geq e$$

Ne segue che f è decrescente per $x \geq e$; di conseguenza la successione $a_n = f(n)$ è decrescente almeno per $n \geq 3$ (il primo intero $> e$). Nel capitolo 5, parlando di *serie a segni alterni*, vedremo un'applicazione di questa osservazione.

Il passaggio "dal discreto al continuo" (cioè dagli interi ai reali) è un modo per avere a disposizione gli strumenti del calcolo differenziale. Attenzione a non usare indiscriminatamente questo espediente. Si ragioni sul fatto che in questo modo *non si potrebbe* studiare la monotonia di successioni come

$$\frac{n!}{2^n} \qquad \frac{n^2 + (-1)^n n}{n^3 + 1} \qquad \frac{n + (-2)^n}{3^n} \cdots$$

4.3 Soluzione di alcuni problemi di massimo e minimo

Ritorniamo in questo paragrafo ai problemi posti all'inizio del paragrafo 4.1, presentandone la soluzione.

Legge della riflessione

Vogliamo ricavare la legge sugli angoli di riflessione dal principio di Fermat. A tale scopo consideriamo, in un mezzo omogeneo, un raggio luminoso riflesso da uno specchio curvo. Immaginiamo che la sezione dello specchio coincida con il grafico di una funzione $y = f(x)$, derivabile. Possiamo sempre scegliere il sistema di riferimento in modo che il punto di riflessione sia l'origine $(0,0)$ e che l'asse x sia diretto come la tangente allo specchio nell'origine. Di conseguenza abbiamo

$$\boxed{f(0) = 0 \qquad f'(0) = 0}$$

Figura 4.29.

Con riferimento alla figura 4.29, P è un punto "sul raggio incidente" e Q è un punto "sul raggio riflesso". Il punto A è un altro punto sullo specchio. Quello che dobbiamo mostrare è che, poiché il tempo impiegato dal raggio per andare da P a Q è minimo (principio di Fermat), deve risultare $\hat{\imath} = \hat{r}$.

Ora, la velocità della luce è costante in un mezzo omogeneo e non varia dopo la riflessione e, pertanto, minimizzare il tempo di percorrenza da P a Q equivale a minimizzare il percorso $S = \overline{PA} + \overline{AQ}$.

Siano (x_0, y_0) le coordinate di P, (x_1, y_1) quelle di Q e $(x, f(x))$ quelle di A. Si ha:

$$S(x) = \sqrt{(x - x_0)^2 + (f(x) - y_0)^2} + \sqrt{(x - x_1)^2 + (f(x) - y_1)^2}$$

La funzione $S = S(x)$ deve avere un minimo in $x = 0$, ascissa del punto di riflessione e perciò deve essere $S'(0) = 0$.

Essendo

$$S'(x) = \frac{(x - x_0) + [f(x) - y_0] \cdot f'(x)}{\sqrt{(x - x_0)^2 + (f(x) - y_0)^2}} + \frac{(x - x_1) + [f(x) - y_1] \cdot f'(x)}{\sqrt{(x - x_1)^2 + (f(x) - y_1)^2}}$$

$S'(0) = 0$ equivale a (ricordando che $f(0) = f'(0) = 0$)

$$\frac{-x_0}{\sqrt{x_0^2 + y_0^2}} + \frac{-x_1}{\sqrt{x_1^2 + y_1^2}} = 0 \tag{4.2}$$

Ma

$$\sqrt{x_0^2 + y_0^2} = \overline{PO} \quad \text{e} \quad \frac{-x_0}{\sqrt{x_0^2 + y_0^2}} = \sin \hat{\imath}$$

mentre

$$\sqrt{x_1^2+y_1^2}=\overline{QO} \quad \text{e} \quad \frac{x_1}{\sqrt{x_1^2+y_1^2}}=\sin\hat{r}\,.$$

Da (4.2) si deduce

$$\sin\hat{\imath}=\sin\hat{r}$$

che implica (essendo $0<\hat{\imath},\hat{r}<\frac{\pi}{2}$), $\hat{\imath}=\hat{r}$.

Legge della rifrazione

Quando un raggio luminoso passa da un mezzo omogeneo a un altro (aria $\to$ acqua, ad esempio) la sua velocità cambia. Supponiamo che i due mezzi siano separati da una superficie piana e che la velocità (scalare) nei due mezzi sia v_1 nel primo e v_2 nel secondo.

Nel passare da un mezzo all'altro, un raggio luminoso cambia direzione di propagazione (viene cioè *rifratto*) ma raggio incidente e raggio rifratto giacciono in un piano verticale alla superficie di separazione.

Fissiamo il sistema di riferimento in modo che l'asse x sia sulla superficie di separazione e che il punto di rifrazione sia l'origine (fig. 4.30).

Sia $P=(x_0,y_0)$ un punto sul raggio incidente, $Q=(x_1,y_1)$ un punto sul raggio rifratto e $A=(x,0)$ un altro punto sull'asse x.

In base al principio di Fermat, il percorso POQ, scelto dal raggio luminoso, è quello che minimizza il tempo di percorrenza da P a Q. Ora, per un qualunque percorso PAQ il tempo $T=T(x)$ di percorrenza è dato da (tempo = spazio/velocità)

$$T(x)=\frac{\overline{PA}}{v_1}+\frac{\overline{AQ}}{v_2}=\frac{\sqrt{(x-x_0)^2+y_0^2}}{v_1}+\frac{\sqrt{(x-x_1)^2+y_1^2}}{v_2}$$

e quindi il principio di Fermat implica $T'(0)=0$.

Figura 4.30.

Poiché

$$T'(x)=\frac{x-x_0}{v_1\sqrt{(x-x_0)^2+y_0^2}}+\frac{x-x_1}{v_2\sqrt{(x-x_1)^2+y_1^2}}$$

l'equazione $T'(0) = 0$ equivale a

$$\frac{1}{v_1}\frac{-x_0}{\sqrt{x_0^2+y_0^2}} - \frac{1}{v_2}\frac{x_1}{\sqrt{x_1^2+y_1^2}} = 0 \tag{4.3}$$

Osservando che $\frac{-x_0}{\sqrt{x_0^2+y_0^2}} = \sin\hat{\imath}$ e $\frac{x_1}{\sqrt{x_1^2+y_1^2}} = \sin\hat{r}$, dalla (4.3) si deduce

$$\boxed{\frac{\sin\hat{\imath}}{\sin\hat{r}} = \frac{v_1}{v_2}}$$

che si chiama *legge di Snell* della rifrazione.
Inscrivere in un cono circolare retto di altezza h e raggio di base R, un cilindro di volume massimo.
Con riferimento alla figura 4.31, occorre scegliere $\overline{CH} = x$ in modo che il volume $\pi\overline{CH}\cdot\overline{CD}^2$ sia massimo. Poiché i triangoli CAB e DAK sono simili, si ha $\overline{DA} = x\cdot\frac{R}{h}$ e quindi $\overline{CD} = R - \overline{DA} = R - \frac{R}{h}x$.

Figura 4.31. $\overline{AC} = R$, $\overline{CB} = h$.

Il volume da massimizzare è quindi dato da

$$V(x) = \pi x\cdot\Big(R - \frac{R}{h}x\Big)^2 = \\ = \pi\Big(\frac{R}{h}\Big)^2 x\cdot(h-x)^2 \qquad 0 < x < h$$

Si ha $V'(x) = \pi\big(\frac{R}{h}\big)^2\big[(h-x)^2 - 2x(h-x)\big]$. Annullando $V'(x)$ si trova

$$0 = (h-x)^2 - 2x(h-x) = (h-x)[h-x-2x] = \\ = (h-x)(h-3x)$$

da cui $x = \frac{h}{3}$, essendo $x = h$ corrispondente a un cono degenere di volume 0. Poiché $V'(x) > 0$ per $x < \frac{h}{3}$, $x = \frac{h}{3}$ è effettivamente punto di massimo, corrispondente al volume

$$V\Big(\frac{h}{3}\Big) = \frac{4}{27}\pi h R^2$$

Consideriamo ora il problema dell'efficienza della produzione di un dato bene, il cui costo medio (per unità di prodotto) sia $c(p)/p$. Per ottenere il massimo di efficienza occorre minimizzare il costo medio e perciò, supponendo la funzione $c = c(p)$ derivabile, dovrà essere

$$\frac{d}{dp}\frac{c(p)}{p} = \frac{c'(p)\cdot p - c(p)}{p^2} = 0$$

Se p_0 è punto di minimo dovrà quindi verificarsi che

$$c'(p_0) = \frac{c(p_0)}{p_0} \tag{4.4}$$

ossia: $\boxed{\text{costo marginale} = \text{costo medio}}$

Figura 4.32. $\operatorname{tg}\omega = c'(p_0) = c(p_0)/p_0$.

La (4.4) si può riscrivere come $p_0 \dfrac{c'(p_0)}{c(p_0)} = 1$ e significa che l'*elasticità di c nel punto* p_0 è uguale a 1.

La (4.4) ha un'interessante interpretazione geometrica, illustrata nella figura 4.32: Se p_0 verifica (4.4), la retta OQ, di coefficiente angolare $\frac{c(p_0)}{p_0}$, coincide con la retta tangente al grafico di $c = c(p)$ nel punto Q, di coefficiente angolare $c'(p_0)$.

Queste considerazioni indicano che non sempre esiste un punto p_0 che soddisfa (4.4) e che, se esiste, può non essere unico (vedere figure 4.33 a), b)).

Figura 4.33. a) Non esistono punti ad elasticità 1; b) p_0 e p_1 sono punti ad elasticità 1.

Consideriamo il problema del passaggio di un tubo di lunghezza 4 m, con sezione trascurabile, attraverso il cunicolo in figura 4.34 e scegliamo il riferimento cartesiano come in figura 4.34

Figura 4.34.

Un minimo di riflessione indica che il passaggio è possibile se la lunghezza del tubo è minore della minima lunghezza dei possibili segmenti APB.

L'equazione della retta AP è

$$y = \frac{1}{2-a}(x-a)$$

Se $a > 2$, la retta interseca il semiasse positivo delle y nel punto $B = \left(0, \frac{a}{a-2}\right)$. La lunghezza del segmento AB, da minimizzare, è data allora da:

$$L(a) = \sqrt{a^2 + \left(\frac{a}{a-2}\right)^2}$$

Poiché la radice quadrata è una funzione crescente basterà minimizzare il suo argomento

$$L_0(a) = a^2 + \left(\frac{a}{a-2}\right)^2 \qquad a > 2$$

Si ha

$$L'(a) = 2a\left(1 - \frac{2}{(a-2)^3}\right)$$

$L'(a) = 0$ per $(a-2)^3 = 2$ cioè $a = \sqrt[3]{2} + 2$, che è punto di minimo come facilmente si verifica. La lunghezza minima è quindi

$$L(\sqrt[3]{2} + 2) = \sqrt{5 + 3\sqrt[3]{4} + 3\sqrt[3]{16}} \simeq 4.16$$

Essendo $4.16 > 4$, il tubo può passare.

Trovare il massimo o il minimo di una funzione è anche un modo per *dimostrare disuguaglianze*, altra tipica applicazione del calcolo differenziale.

Trovare la più piccola costante $c_n > 0$ per cui è vera la seguente disuguaglianza:

$$(a+b)^n \leq c_n\,(a^n + b^n) \text{ per ogni } a, b > 0$$

(dove n è un intero positivo fissato).

Anzitutto osserviamo che, per $a, b > 0$, è sempre

$$(a^n + b^n) < (a+b)^n$$

(basta sviluppare la potenza di $(a+b)$); si chiede di dimostrare che vale anche una disuguaglianza di segno opposto, per una opportuna costante c_n. Si osservi che la costante *più piccola* per cui questo è vero è la costante *migliore*, cioè quella per cui la disuguaglianza dà l'informazione più precisa. Anzitutto riduciamo il problema ad un problema in una sola variabile anziché due: basta raccogliere b^n ad ambo i membri e semplificare:

$$\left(\frac{a}{b}+1\right)^n \leq c_n\left(\left(\frac{a}{b}\right)^n+1\right)$$

e quindi porre $t = a/b$:

$$(t+1)^n \leq c_n\left(t^n+1\right).$$

Sia ora

$$f(t) = \frac{(t+1)^n}{t^n+1}.$$

La disuguaglianza da dimostrare è $f(t) \leq c_n$ per ogni $t > 0$, e la minima costante per cui questo è vero è *il massimo di f in* $(0, \infty)$. Ci siamo così ricondotti ad un consueto problema di massimizzazione di una funzione di una variabile. Calcoliamo

$$f'(t) = \frac{n(t+1)^{n-1}(t^n+1) - nt^{n-1}(t+1)^n}{(t^n+1)^2} = \frac{n(t+1)^{n-1}}{(t^n+1)^2}\left(1-t^{n-1}\right) \geq 0 \text{ per } t \leq 1.$$

Quindi $t = 1$ è il punto di massimo, e il massimo è

$$c_n = f(1) = 2^{n-1}.$$

4.4 Il teorema di de l'Hospital

Una notevole applicazione del calcolo differenziale si ha nel calcolo dei limiti che si presentano nelle forme di indecisione $\left[\frac{0}{0}\right]$ e $\left[\frac{\infty}{\infty}\right]$. Precisamente si ha:

Teorema 4.8 (di de l'Hospital) *Siano f, g funzioni derivabili in un intervallo (a, b) con $g, g' \neq 0$ in (a, b). Se*

i) $\lim\limits_{x\to a^+} f(x) = \lim\limits_{x\to a^-} g(x) = 0$ *oppure* $\pm\infty$

ii) $\lim\limits_{x\to a^+} f'(x)/g'(x) = L \in \mathbb{R}^*$

Allora

$$\lim_{x\to a^+} \frac{f(x)}{g(x)} = L$$

Il teorema continua a valere se $a = -\infty$ oppure se si considera il limite per $x \to b^-$ (anziché per $x \to a^+$), con $b \leq +\infty$.

DIMOSTRAZIONE. Nel caso $f(x), g(x) \to 0$. Daremo prima un'idea intuitiva (ma non concludente) della dimostrazione, e poi mostreremo come la si possa rendere rigorosa. Sia x_n una successione tendente ad a^+; prolunghiamo per continuità f e g in a ponendo $f(a) = g(a) = 0$. Allora

$$\frac{f(x_n)}{g(x_n)} = \frac{f(x_n) - f(a)}{g(x_n) - g(a)}. \tag{4.5}$$

Se applichiamo a f, g separatamente il teorema di Lagrange sull'intervallo $[a, x_n]$, otteniamo che l'ultimo quoziente scritto è uguale a:

$$\frac{f'(t_n)(x_n - a)}{g'(t_n^*)(x_n - a)} = \frac{f'(t_n)}{g'(t_n^*)}$$

dove t_n, t_n^* sono due punti opportuni che cadono nell'intervallo (a, x_n). Poiché quando $x_n \to 0$ anche t_n e $t_n^* \to 0$, sembra "ragionevole" che il limite del quoziente di f'/g' sia uguale al limite del quoziente f/g. Tuttavia questo non si può affermare rigorosamente, perché le successioni t_n, t_n^* sono a priori diverse tra loro. Per aggirare il problema occorre modificare leggermente l'argomentazione seguita. Riprendiamo dunque la dimostrazione dalla (4.5), e definiamo

$$h(x) = f(x_n)g(x) - g(x_n)f(x)$$

Notiamo che $h(a) = h(x_n) = 0$. La funzione h soddisfa le ipotesi del teorema di Lagrange sull'intervallo $[a, x_n]$, dunque esiste $t_n \in (a, x_n)$ tale che

$$h'(t_n) = \frac{h(x_n) - h(a)}{x_n - a} = 0$$

ovvero, calcolando

$$h'(x) = f(x_n)g'(x) - g(x_n)f'(x), \qquad f(x_n)g'(t_n) - g(x_n)f'(t_n) = 0$$

Dunque per ogni x_n esiste un punto $t_n \in (a, x_n)$ tale che

$$\frac{f(x_n)}{g(x_n)} = \frac{f'(t_n)}{g'(t_n)}$$

Per $n \to \infty$, $t_n \to a^+$, perciò $\frac{f'(t_n)}{g'(t_n)} \to L$, e di conseguenza anche $\frac{f(x_n)}{g(x_n)} \to L$, che è quanto volevamo dimostrare. ◆

Come prima applicazione del teorema di de L'Hospital, dimostriamo il teorema 3.25 del capitolo 3, relativo alla "gerarchia degli infiniti" nel calcolo dei limiti.

Sia da calcolare

$$\lim_{x \to +\infty} \frac{x^\alpha}{e^{\beta x}} \qquad \alpha > 0, \beta > 0$$

Posto $f(x) = x$, $g(x) = e^{\gamma x}$ $(\gamma > 0)$ si ha $f'(x) = 1$, $g'(x) = \gamma e^{\gamma x}$ e poiché

$$\lim_{x \to +\infty} \frac{f'(x)}{g'(x)} = \lim_{x \to +\infty} \frac{1}{\gamma e^{\gamma x}} = 0$$

si ha che $\lim_{x \to +\infty} \frac{x}{e^{\gamma x}} = 0$. Osservando ora che

$$\frac{x^\alpha}{e^{\beta x}} = \Big(\frac{x}{e^{(\beta/\alpha)x}}\Big)^\alpha = \Big(\frac{x}{e^{\gamma x}}\Big)^\alpha \text{ con } \gamma = \frac{\beta}{\alpha}$$

si ha $\lim_{x \to +\infty} \frac{x^\alpha}{e^{\beta x}} = \Big(\lim_{x \to +\infty} \frac{x}{e^{\gamma x}}\Big)^\alpha = 0$.

Per provare ora che

$$\lim_{x\to+\infty} \frac{(\log_a x)^\alpha}{x^\beta} = 0$$

è sufficiente eseguire il cambio di variabile $t = \log_a x, x = a^t$, e ricondursi al limite appena dimostrato.

Un'altra applicazione tipica del teorema di de L'Hospital è il calcolo di limiti che coinvolgono forme di indeterminazione non risolubili mediante i limiti notevoli (ossia, al prim'ordine).

Esempio

4.7

$$\lim_{x\to 0} \frac{x - \sin x}{x^3} = \left[\frac{0}{0}\right]$$

La stima al prim'ordine $\sin x \sim x$ non è sufficiente a risolvere tale forma di indeterminazione, in quanto porta a

$$\frac{x - \sin x}{x^3} = \frac{1}{x^2}\left(1 - \frac{\sin x}{x}\right) \to ? \quad (\infty \cdot 0)$$

(infatti la quantità tra parentesi tende a zero, per la stima asintotica).
Il teorema di de L'Hospital dà:

$$\lim_{x\to 0} \frac{(x - \sin x)'}{(x^3)'} = \lim_{x\to 0} \frac{1 - \cos x}{3x^2} = \frac{1}{6}$$

(per il limite notevole del coseno). Dunque anche il limite di partenza vale 1/6.

Per usare efficacemente il teorema, è utile talvolta fare qualche passaggio preliminare (come una stima asintotica, oppure un cambio di variabile), in modo che la successiva applicazione del teorema semplifichi l'espressione, anziché complicarla. Talvolta, inoltre, il teorema va applicato più volte consecutivamente, per sciogliere la forma di indeterminazione: in questo caso, comunque, già dopo la prima applicazione ci si dovrebbe accorgere che l'ordine di infinitesimo (o di infinito) a numeratore o a denominatore si è abbassato. Se così non è, significa che non si sta seguendo la strada giusta per calcolare il limite!

Esempio

4.8

$$\lim_{x\to 0} \frac{e^{-1/x^2}}{x} = \left[\frac{0}{0}\right]$$

Se applichiamo direttamente il teorema di de L'Hospital, l'espressione si complica (provare!). Invece, col cambio di variabile $x^2 = 1/t$ ci si riconduce a

$$\lim_{t\to+\infty} \frac{\sqrt{t}}{e^t}$$

che per confronto di infiniti (cioè per il teorema di de L'Hospital!) tende a zero.

Teorema di de L'Hospital: precauzioni per l'uso

– Il teorema si usa per quozienti, *non per prodotti*!
– Il teorema si usa per quozienti che siano effettive *forme di indeterminazione*!
– Il teorema prescrive di calcolare il quoziente delle derivate, *non la derivata del quoziente*!
– Se il limite di f'/g' non esiste, nulla si può affermare sul limite di f/g. In particolare, *non è lecito concludere che non esiste nemmeno il limite di* f/g.

Esempio

4.9 Si voglia calcolare:

$$\lim_{x\to+\infty} \frac{x-\sin x}{x+\sin x} = \left[\frac{+\infty}{+\infty}\right]$$

Il valore del limite si può stabilire immediatamente:

$$\frac{x-\sin x}{x+\sin x} \sim \frac{x}{x} = 1$$

Supponiamo di volerlo calcolare col teorema di de L'Hospital:

$$\lim_{x\to+\infty} \frac{(x-\sin x)'}{(x+\sin x)'} = \lim_{x\to+\infty} \frac{1-\cos x}{1+\cos x} \quad \text{non esiste}$$

Se concludessimo che il limite di partenza non esiste, diremmo il falso (quel limite vale 1). Il punto è che se $\lim\limits_{x\to x_0} \frac{f'(x)}{g'(x)}$ non esiste (né finito né infinito), cade una delle ipotesi del teorema, e in base ad esso non si può semplicemente concludere nulla.

4.5 Limite della derivata e derivabilità

Esempio

4.10 Supponiamo di voler studiare la derivabilità o meno di una funzione del tipo:

$$f(x) = x\,|\log x|\,.$$

La funzione è definita e continua per $x > 0$, ed è certamente derivabile per $x \neq 1$, mentre in $x = 1$, annullandosi l'argomento del modulo, sospettiamo la presenza di un punto angoloso. Poiché

$$f(x) = \begin{cases} x\log x & \text{per } x \geq 1 \\ -x\log x & \text{per } 0 < x \leq 1 \end{cases}$$

possiamo calcolare

$$f'(x) = \begin{cases} \log x + 1 & \text{per } x > 1 \\ -(\log x + 1) & \text{per } 0 < x < 1. \end{cases}$$

Rimane da studiare la derivabilità di f in $x = 1$. È facile osservare che

$$\lim_{x\to1^+} f'(x) = \lim_{x\to1^+} (\log x + 1) = 1 \text{ mentre}$$
$$\lim_{x\to1^-} f'(x) = \lim_{x\to1^-} (-\log x - 1) = -1.$$

Possiamo usare queste informazioni per affermare che la derivata destra e sinistra di f, in $x = 1$, valgono 1, -1, rispettivamente? (Se è così, $x = 1$ è un punto angoloso). In altre parole, è vero ad esempio che

$$\lim_{x \to 1^+} f'(x) = f'_+(1) \text{ ossia:}$$

il limite destro della derivata per $x \to 1$ è uguale alla derivata destra in $x = 1$?

Si osservi che la cosa non è ovvia perché, per definizione, la derivata destra è il limite destro del *rapporto incrementale*:

$$f'_+(1) = \lim_{h \to 0^+} \frac{f(1+h) - f(1)}{h}$$

che però è molto più scomodo da calcolare, in pratica, rispetto al limite della derivata. Il teorema di Lagrange permette di rispondere affermativamente a questa domanda:

Teorema 4.9 *Sia $f : [a, b) \to \mathbb{R}$, continua in a, derivabile in (a, b), ed esista (finito o infinito)*

$$\lim_{x \to a^+} f'(x) = m \in \mathbb{R}^*.$$

Allora esiste

$$f'_+(a) = m.$$

Detto in parole: se la funzione è continua in a ed esiste il limite destro della derivata, allora esiste la derivata destra, e coincide con quel limite. Un analogo enunciato vale per la derivata sinistra, e quindi per la derivata.

Dimostrazione. Sia $h > 0$. Applicando il teorema di Lagrange a f sull'intervallo $[a, a+h]$, otteniamo che esiste $t_h \in [a, a+h]$ tale che

$$\frac{f(a+h) - f(a)}{h} = f'(t_h).$$

Per $h \to 0^+$ si ha che $t_h \to a^+$, quindi per ipotesi $f'(t_h) \to m$. Di conseguenza esiste

$$f'_+(a) = \lim_{h \to 0^+} \frac{f(a+h) - f(a)}{h} = m.$$

Nell'esempio precedente, f è continua in $x = 1$, quindi il teorema è applicabile e permette di concludere che

$$f'_+(1) = 1, f'_-(1) = -1, \text{ per cui } x = 1 \text{ è un punto angoloso.}$$

Si considerino anche i prossimi esempi.

Esempi

4.11 Sia

$$f(x) = x^2 \log|x|,$$

definita per $x \neq 0$. Sappiamo però che esiste

$$\lim_{x \to 0} x^2 \log|x| = 0,$$

perciò possiamo prolungare f per continuità definendo $f(0) = 0$, ed f così risulta continua anche in 0. Chiediamoci se f è anche derivabile in 0. Cominciamo a calcolare, per $x \neq 0$,

$$f'(x) = 2x \log|x| + x.$$

Ora calcoliamo

$$\lim_{x \to 0} f'(x) = \lim_{x \to 0} (2x \log|x| + x) = 0.$$

Perciò per il teorema precedente esiste $f'(0) = 0$, in particolare f è derivabile in $x = 0$.

4.12 Sia

$$f(x) = \operatorname{arctg} \frac{1}{x}.$$

Calcoliamo

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

Osserviamo che

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

Si può concludere allora che esiste $f'(x) = -1$? No! In questo caso viene meno l'ipotesi di continuità di f in 0: si osservi che f non è definita per $x = 0$, e

$$\lim_{x \to 0^{\pm}} f(x) = \pm \frac{\pi}{2}$$

dunque $x = 0$ è un punto di discontinuità a salto, ed f non è derivabile in $x = 0$. Si osservi dunque che l'ipotesi di continuità, nel teorema precedente, è essenziale.

4.13 Sia

$$f(x) = x^2 \sin\left(\frac{1}{x}\right).$$

La funzione non è definita in $x = 0$, ma si può prolungarla per continuità ponendo $f(0) = 0$ (infatti, $f(x) \to 0$ per $x \to 0$). Dunque ora f è continua in 0; per $x \neq 0$, inoltre, f è certamente derivabile e si ha:

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

Osserviamo ora che

$$\lim_{x \to 0} f'(x) \text{ non esiste}$$

(né finito né infinito). Possiamo concludere che non esiste $f'(0)$? No! Si noti che il teorema non permette questa conclusione, ed in effetti, in questo caso, calcolando $f'(0)$ mediante la definizione si ha:

$$f'(0) = \lim_{h \to 0} \frac{f(h) - 0}{h} = \lim_{h \to 0} h \sin\left(\frac{1}{h}\right) = 0,$$

perciò in realtà f è derivabile e $f'(0) = 0$. Quando il limite della derivata non esiste (né finito né infinito), il teorema non è applicabile; in particolare non è detto che non esista la derivata nel punto considerato: occorre calcolarla dalla definizione.

Esercizi

Dopo aver stabilito l'insieme di definizione delle seguenti funzioni, determinarne i punti di massimo e minimo e tracciarne sommariamente il grafico.

41 $x^2 e^x$

42 $x^3 + x^2 - x + 1$

43 $\sqrt[3]{x} e^x$

44 $x^4 - 8x^3 + 22x^2 - 24x + 12$

45 $x^2 \log x$

46 $x^5 - 5x^4 + 5x^3 + 1$

47 $\sqrt{x} \log x$

48 $\dfrac{x+2}{x^2+1}$

Problemi di massimo e minimo.

Impostare e risolvere i seguenti problemi di massimo o minimo, cercando di imitare le tecniche illustrate negli esempi del paragrafo 4.3. In particolare, impostare il problema in modo da ridurlo ad un problema di massimizzazione o minimizzazione di una funzione di una variabile.

49 Una ditta produttrice di birra desidera minimizzare il costo della lattina. Essendo di materiale omogeneo e volume fissato (33 cl) occorre minimizzare la superficie totale del cilindro di volume pari a 33 cl. Quali sono le dimensioni (altezza e diametro) della lattina?

50 Un uomo deve raggiungere un punto che si trova sull'altra sponda di un fiume, 100 metri più a valle; il fiume è rettilineo e largo 10 metri; l'uomo può correre sulla sponda del fiume con velocità v, quindi tuffarsi e attraversare a nuoto il fiume, con velocità inferiore, pari a δv $(0 < \delta < 1)$.

Determinare dopo quanti metri di corsa l'uomo si deve tuffare, affinché sia minimo il tempo impiegato a raggiungere la meta.

Se l'uomo è un nuotatore provetto, δ sarà quasi uguale a 1: determinare il valore esatto di δ per il quale all'uomo conviene tuffarsi immediatamente, senza percorrere neanche un metro sulla terraferma.

51 Si vuole costruire una scatola, senza coperchio, col vincolo che la base sia quadrata e la superficie totale della scatola misuri 108 cm. Di quali dimensioni (lato della base e altezza) dev'essere affinché il volume sia massimo possibile? E quanto sarà il volume?

52 Si determini l'altezza e il raggio di base del cono di volume minimo circoscritto ad una sfera di raggio r. Si dimostri poi che il suddetto cono è anche quello di superficie minima totale.

53 Sia n un intero ≥ 2. Determinare (in dipendenza da n) le migliori costanti c_1, c_2 per cui si abbia:

$$c_1 (a+b) \leq \left(\sqrt[n]{a} + \sqrt[n]{b} \right)^n \leq c_2 (a+b) \quad \forall a, b \in \mathbb{R}.$$

(Dire che le costanti sono le migliori significa che c_1 è la massima e c_2 la minima costante che rendono vere le disuguaglianze precedenti). Suggerimento: imitare il procedimento seguito nell'ultimo esempio del paragrafo 4.3.

Monotonia di successioni.

54 Si stabilisca se le seguenti successioni a_n sono monotone (crescenti o decrescenti), almeno per $n \geq n_0$ opportuno:

$$\frac{n+3}{n^2-2n+4}; \qquad \frac{n^2}{\log n}; \qquad \frac{n!}{2^n};$$
$$n-\sqrt{n^2+1}; \qquad ne^{-n}; \qquad n\sin n.$$

Calcolare i seguenti limiti utilizzando opportunamente, se occorre, il teorema di de L'Hospital.

55 $\lim_{x\to 1} \frac{\cos(\pi/2x)}{\sin(\pi x)}$

56 $\lim_{x\to +\infty} \frac{x^x}{2^{x^2}}$

57 $\lim_{x\to 0} \frac{e^x - x - 1}{x^2}$

58 $\lim_{x\to 2} \frac{\sqrt{x^2-4}}{\sin(x-2)}$

59 $\lim_{x\to +\infty} x^{1/x}$

60 $\lim_{x\to 2} \frac{x^2-4}{\sqrt{\sin(x-2)^2}}$

61 $\lim_{x\to 0} \frac{x^2\cos x + 1 - e^{x^2}}{x^2\sin^2 x}$

62 $\lim_{x\to 0} \frac{x-\sin x}{x^2\sin x}$

63 $\lim_{x\to 0} \frac{x\sin x + \log(1-x^2)}{x^2(2x+x^2)^2}$

64 $\lim_{x\to 1} \frac{x\log x}{x^2-1}$

65 $\lim_{x\to +\infty} x\left(\operatorname{arctg} x - \frac{\pi x+2}{2x+1}\right)$

66 $\lim_{x\to 1} \frac{x^2-3x+3}{\cos\frac{\pi x}{2}}$

67 Sia $f(x) = \frac{\sin x}{x}$. Dopo aver prolungato per continuità f in $x=0$, calcolare $f'(0)$ e $f''(0)$. (*Suggerimento*: Nei due limiti che si chiede di calcolare, è utile il teorema di de L'Hospital).

68 Come l'esercizio precedente, per la funzione $\left(\frac{\sin x}{x}\right)^2$.

69 Il teorema di de L'Hospital a volte può produrre...il moto perpetuo. Si provi a calcolare il seguente limite (elementare!) applicando più volte il teorema:

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

70 Stabilire se le seguenti funzioni sono continue e se sono derivabili, nel punto $x=0$:

$$f_1(x) = \begin{cases} x^2\log x & \text{per } x>0 \\ e^x - 1 & \text{per } x\leq 0 \end{cases} \qquad f_2(x) = x|x| \qquad f_3(x) = \begin{cases} x^2\sin\frac{1}{x} & \text{per } x>0 \\ 0 & \text{per } x=0 \\ e^{1/x} & \text{per } x<0 \end{cases}$$

Suggerimento: utilizzare la definizione di derivata oppure il criterio visto nel par. 4.5, scegliendo caso per caso il metodo opportuno.

5 DERIVATA SECONDA

La derivata seconda di una funzione ha vari significati geometrici, che ci permetteranno di meglio studiare le proprietà del grafico di una funzione.

5.1 Significato geometrico della derivata seconda

Se la derivata prima ha come significato geometrico quello di pendenza del grafico, la derivata seconda rappresenta la velocità di variazione di tale pendenza e pertanto costituisce una misura del grado di scostamento del grafico dall'andamento rettilineo.

Per precisare meglio il concetto, consideriamo la situazione illustrata in figura 4.35.

Figura 4.35.

La funzione f in figura, soddisfa le condizioni $f(0) = f'(0) = 0$, $f''(0) \geq 0$.

Consideriamo ora la famiglia di semicirconferenze con centro sull'asse y tangente al grafico di f nell'origine.

L'equazione della famiglia è

$$y = R - \sqrt{R^2 - x^2}$$

come facilmente si verifica e, per ogni R, si ha $y'(0) = y(0) = 0$.

Tra queste semicirconferenze selezioniamo quella che non soltanto ha la stessa tangente in $x = 0$ ma anche la stessa velocità di variazione della pendenza in $x = 0$.

In altri termini, scegliamo R in modo che

$$y''(0) = f''(0)\,. \tag{5.1}$$

Si ha,

$$y'(x) = \frac{x}{\sqrt{R^2 - x^2}}\,, \quad y''(x) = \frac{R^2}{(R^2 - x^2)^{3/2}}$$

e quindi $y''(0) = \frac{1}{R}$. Se dunque vogliamo che sia soddisfatta la (5.1) occorre scegliere R in modo che

$$\boxed{f''(0) = \frac{1}{R}.} \tag{5.2}$$

La (5.2) esprime il significato geometrico della derivata seconda: nella situazione descritta, $f''(0)$ rappresenta il reciproco del raggio della circonferenza che "meglio approssima" f in $x = 0$; $\frac{1}{R}$ prende il nome di *curvatura* (del grafico) *di* f in $x = 0$ e

R è il *raggio di curvatura*. Ripetendo un'analoga costruzione in un punto generico x del grafico di f (senza supporre a priori $f(x) = f'(x) = 0$) si trova che il cerchio che meglio approssima il grafico nel punto di ascissa x ha raggio $R(x)$ dato da:

$$\boxed{\frac{1}{R(x)} = \frac{|f''(x)|}{(1 + f'(x)^2)^{3/2}}}$$

Esempio

5.1 Per una parabola del tipo $y = ax^2$ si ha $y'(x) = 2ax$, $y''(x) = 2a$ e quindi

$$\frac{1}{R(x)} = \frac{2|a|}{(1 + 4a^2x^2)^{3/2}}.$$

Si noti che la curvatura è massima per $x = 0$, cioè nel vertice.

5.2 Derivata seconda, concavità e convessità

Abbiamo visto che il significato geometrico della derivata seconda è legato al concetto di *curvatura* della curva grafico della funzione. Vedremo ora come rendere più preciso, qualitativamente e quantitativamente, questa idea di curvatura del grafico attraverso il concetto di *convessità*. Si tratta di una nozione che in matematica ha una portata molto vasta, di cui qui potremo solo toccare i primi aspetti.

Convessità e corde

Cominciamo col ricordare che, in geometria, si dice che una *figura* F (pensiamo ad esempio a un sottoinsieme del piano) è *convessa* se, per ogni coppia di punti $P_1, P_2 \in F$ avviene che il segmento che congiunge P_1 a P_2 è interamente contenuto in F:

Figura 4.36. a) un insieme convesso; b) un insieme non convesso.

Coerentemente a questa idea, diamo allora la seguente

Definizione 4.6 Consideriamo una funzione $f : I \to \mathbb{R}$ (dove I è un intervallo). Si chiama *epigrafico* (o sopragrafico) di f l'insieme

$$\text{epi} f = \left\{(x, y) \in \mathbb{R}^2 : x \in I \text{ e } y \geq f(x)\right\}.$$

Si dice allora che f è convessa in I se il suo epigrafico è un insieme convesso. Si dice che f è concava in I se $-f$ è convessa in I.

Figura 4.37. a) epigrafico; b) funzione convessa; c) funzione concava.

Segnaliamo subito che esiste una seconda terminologia, alternativa, che si usa talvolta per esprimere questi concetti: anziché f convessa in (a, b) si può dire f *concava verso l'alto* in (a, b), e invece di concava in (a, b) si può dire *concava verso il basso* in (a, b). Per non creare confusione, in questo paragrafo non useremo mai questa seconda terminologia.
Si dimostra facilmente che la definizione precedente è equivalente alla prossima, che fa riferimento al grafico anziché all'epigrafico della funzione:

Definizione 4.7 Se funzione $f : I \to \mathbb{R}$ (dove I è un intervallo), si dice che f è convessa (concava) in I se per ogni coppia di punti $x_1, x_2 \in I$ si ha che il segmento ("corda") di estremi $(x_1, f(x_1)), (x_2, f(x_2))$ non ha punti sotto (sopra) il grafico di f. ■

Questa definizione può essere riformulata anche mediante una disuguglianza analitica. Con riferimento alla figura 4.38, esprimiamo il generico punto (z_t, f_t) del segmento congiungente i punti $(x_1, f(x_1)), (x_2, f(x_2))$ come *combinazione lineare convessa* dei suoi estremi:

$$z_t = (1-t)\,x_1 + tx_2; \qquad f_t = (1-t)\,f(x_1) + tf(x_2)\,, \quad t \in [0,1]\,.$$

Si osservi che al variare di $t \in [0, 1]$ il punto z_1 percorre il segmento $[x_1, x_2]$ sull'asse x, e il punto del piano (z_t, f_t) percorre il segmento di estremi $(x_1, f(x_1)), (x_2, f(x_2))$.

Figura 4.38.

La seconda definizione di convessità che abbiamo dato si esprime dunque così:

$$f(z_t) \leq f_t \ \forall x_1, x_2 \in I, \forall t \in [0,1]$$

o, più esplicitamente

$$f((1-t)x_1 + tx_2) \leq (1-t)f(x_1) + tf(x_2) \qquad \forall x_1, x_2 \in I, \forall t \in [0,1].$$

Questa disuguaglianza esprime il modo in cui viene spesso sfruttata l'informazione che una funzione è convessa. Se nella precedente disuguglianza vale sempre il $<$ la funzione si dice *strettamente convessa*. Per le funzioni concave (o strettamente concave) varrà una analoga disuguaglianza col segno $\geq$ (o $>$).

Si noti che la definizione di funzione convessa non richiede a priori che f sia una funzione continua o derivabile, quindi è molto generale:

Figura 4.39. a) funzione convessa con due punti angolosi; b) funzione convessa con due discontinuità agli estremi di un intervallo.

Si può dimostrare che le situazioni esemplificate dalla figura precedente sono effettivamente "le peggiori possibili", nel senso che:

Teorema 4.10 *Una funzione convessa (o concava) su un intervallo I è continua in I, salvo al più negli estremi dell'intervallo. Inoltre possiede derivata destra e sinistra in ogni punto interno dell'intervallo.*

Dunque, punti angolosi all'interno e punti di discontinuità agli estremi sono i soli comportamenti irregolari permessi ad una funzione convessa o concava.

Convessità e derivate

Abbiamo visto che la definizione di convessità ha senso anche per funzioni non derivabili. Ad ogni modo, se sappiamo a priori che la funzione è derivabile una volta o due volte nell'intervallo considerato, allora la convessità risulta avere relazioni interessanti con la derivata prima e seconda della funzione:

Teorema 4.11 *Sia $f : (a,b) \to \mathbb{R}$.*
a) *Se f è derivabile in (a,b), allora f è convessa (concava) in (a,b) se e solo se f' è crescente (decrescente) in (a,b).*
b) *Se f è derivabile due volte in (a,b), allora f è convessa (concava) in (a,b) se e solo se $f''(x) \geq 0$ (≤ 0) per ogni $x \in (a,b)$.*

Il teorema si modifica in maniera ovvia per le funzioni strettamente convesse o concave: nella (a) "crescente" (decrescente) è sostituito da "strettamente crescente" (strettamente decrescente); nella (b) "$f''(x) \geq 0$" ($f''(x) \leq 0$) è sostituito con "$f''(x) > 0$" ($f''(x) < 0$).

Non dimostriamo il punto (a); il punto (b) segue da (a) per il test di monotonia, applicato ad f'.

Come conseguenza di questo teorema, lo studio del segno della derivata seconda ci permette di decidere della convessità o concavità di una funzione, aggiungendo quindi un elemento utile allo studio del grafico della funzione.

Esempi

5.2 Le funzioni esponenziali $f(x) = a^x$ sono convesse in $\mathbb{R}$, per qualunque base $a \lessgtr 1$. Infatti:

$$f'(x) = a^x \log a; \quad f''(x) = a^x (\log a)^2 > 0 \quad \forall x \in \mathbb{R}, \quad \forall a > 0,\ a \neq 1.$$

5.3 Le funzioni logaritmiche $f(x) = \log_a x$ sono concave in $(0, \infty)$ se $a > 1$, convesse in $(0, \infty)$ se $0 < a < 1$. Infatti:

$$f'(x) = \frac{1}{x \log a}; \quad f''(x) = -\frac{1}{x^2 \log a} \begin{cases} < 0 & \forall x > 0, \quad \text{se } a > 1 \\ > 0 & \forall x > 0, \quad \text{se } 0 < a < 1. \end{cases}$$

5.4 Le funzioni potenza a esponente reale α, $f(x) = x^\alpha$ sono convesse in $(0, \infty)$ se $\alpha > 1$ oppure $\alpha < 0$, sono concave in $(0, \infty)$ se $0 < \alpha < 1$. Infatti:

$$f'(x) = \alpha x^{\alpha - 1}; \quad f''(x) = \alpha(\alpha - 1) x^{\alpha - 2}$$

che ha, per ogni $x > 0$, il segno di $\alpha(\alpha - 1)$.

Lo studente si renda conto del significato geometrico di queste affermazioni osservando i grafici delle funzioni ora citate.

Convessità e rette tangenti

Un'utile caratterizzazione geometrica della convessità, su cui ritorneremo in seguito (par. 7.4) coinvolge le *rette tangenti* al grafico della funzione:

Teorema 4.12 *Una funzione $f : (a, b) \to \mathbb{R}$, derivabile in (a, b), è convessa (concava) in (a, b) se e solo se comunque si scelga un punto $x_0 \in (a, b)$ si ha che il grafico di f si mantiene in tutto (a, b)* sopra *(sotto) il grafico della sua retta tangente in $(x_0, f(x_0))$.*

Esempio

5.5 Sappiamo che $f(x) = e^x$ è convessa in tutto $\mathbb{R}$. Ne segue che se, ad esempio, costruiamo la retta tangente al grafico di f in $x = 0$, ossia $y = x + 1$, risulta $e^x \geq 1 + x\ \forall x \in \mathbb{R}$. La retta tangente al grafico di f in $x = 1$, invece, è $y = ex$. Di conseguenza risulta anche $e^x \geq ex\ \forall x \in \mathbb{R}$, e così via.

È utile talvolta sapere che una funzione (derivabile) convessa sta *sopra* le proprie rette tangenti e *contemporaneamente* sta *sotto* le proprie corde (definizione di convessità). Questo permette di concludere che, presi due punti qualsiasi sul grafico di una funzione convessa, il grafico tra quei due punti cade tutto nel triangolo che ha per lati la corda che li unisce e le rette tangenti al grafico nei due punti (fig. 4.40).

Figura 4.40.

Questa limitazione può essere utile, ad esempio, quando si cerca di localizzare l'intersezione tra due curve (si veda l'esercizio 106).

Punti di flesso

Ovviamente, il *verso della concavità* di una funzione (ossia il fatto che sia convessa o concava) può cambiare, nel suo insieme di definizione; questo ci conduce al concetto di *punto di flesso*:

Definizione 4.8 Sia $f : (a, b) \to \mathbb{R}$ una funzione e $x_0 \in (a, b)$ sia un punto di derivabilità per f, oppure sia $f'(x_0) = \pm\infty$. Il punto x_0 si dice di *flesso* per f se esiste un intorno destro $(x_0, x_0 + h)$, $h > 0$, in cui f è convessa (concava), e un intorno sinistro $(x_0 - h, x_0)$ in cui f è concava. ■

Attraversando un punto di flesso, la derivata seconda di f (se esiste) cambia segno. Ci aspettiamo allora che in questo punto f'' si annulli. Si può infatti dimostrare che:

Teorema 4.13 *Sia x_0 un punto di flesso per f; se esiste $f''(x_0)$, allora $f''(x_0) = 0$.*

Notiamo che, se sapessimo che f'' esiste in un intorno di x_0 ed è continua in x_0, allora la tesi del teorema seguirebbe dal teorema dei valori intermedi per le funzioni continue (applicato ad f''). Ad ogni modo, si può dimostrare il teorema appena enunciato anche senza queste ipotesi ulteriori.

L'implicazione opposta a quella enunciata dal teorema non è vera, come mostra l'esempio 5.8: un punto in cui la derivata seconda si annulla può non essere di flesso.

Il significato geometrico dei punti di flesso è ulteriormente chiarito dal prossimo:

Teorema 4.14 *Se $f : (a, b) \to \mathbb{R}$ è derivabile in (a, b) e $x_0 \in (a, b)$ è un punto di flesso, allora il grafico di $f(x)$ attraversa la propria retta tangente in $(x_0, f(x_0))$ (si veda la fig. 4.41):*

Figura 4.41.

DIMOSTRAZIONE. Tracciamo la retta tangente al grafico di $f(x)$ nel punto di ascissa x_0. Se f è (ad esempio) concava in (a, x_0), poiché f è derivabile, il grafico di f sta sotto la retta in (a, x_0) (si veda il teorema su convessità e rette tangenti); d'altra parte f è convessa in (x_0, b), perciò il suo grafico sta sopra la retta in (x_0, b). Di conseguenza in x_0 il grafico attraversa la retta tangente. ♦

Esempi

5.6 Sia $f(x) = x^3$. Poiché $f'(x) = 3x^2 \geq 0$, la funzione è crescente su tutto $\mathbb{R}$. Ha un punto stazionario per $x = 0$, che non sarà però punto di massimo o minimo, perché la funzione è crescente. $f''(x) = 6x \geq 0$ per $x \geq 0$. Dunque la funzione è convessa per $x > 0$, concava per $x < 0$ e ha un punto di flesso in $x = 0$, a tangente orizzontale.

5.7 Sia $f(x) = e^{-x^2}$. Poiché $f'(x) = -2xe^{-x^2} \geq 0$ per $x \leq 0$, la funzione cresce per $x < 0$, decresce per $x > 0$ perciò ha un punto di massimo relativo in $x = 0$.

$f''(x) = e^{-x^2}\left(4x^2 - 2\right) \geq 0$ per $x^2 \geq \frac{1}{2}$; $x \geq \frac{1}{\sqrt{2}}$; $x \leq -\frac{1}{\sqrt{2}}$. La funzione è convessa per questi valori, concava per $-\frac{1}{\sqrt{2}} \leq x \leq \frac{1}{\sqrt{2}}$, ha punti di flesso in $x = \pm\frac{1}{\sqrt{2}}$, con retta tangente di pendenza $f'\left(\pm\frac{1}{\sqrt{2}}\right) = \mp\sqrt{\frac{2}{e}}$ (v. fig. 4.42).

Figura 4.42.

5.8 Sia $f(x) = x^4$. Poiché $f'(x) = 4x^3 \geq 0$ per $x > 0$, la funzione è crescente per $x > 0$, decrescente per $x < 0$ e ha un punto di minimo in $x = 0$. $f''(x) = 12x^2 \geq 0$ per ogni x, quindi la funzione è convessa in tutto $\mathbb{R}$. Perciò il punto $x = 0$, in cui f'' si annulla, *non* è un punto di flesso.

5.9 Sia $f(x) = \sqrt[3]{x}$. Per $x > 0$ la funzione è concava, per $x < 0$ la funzione è convessa (come si deduce dal segno della derivata seconda); in $x = 0$ è $f'(0) = +\infty$. In base alla nostra definizione, $x = 0$ è un punto di flesso (a tangente verticale). Questo è anche coerente alla terminologia che abbiamo introdotto discutendo i punti di non derivabilità, nel par. 2.4. Si vede che anche in questo caso il grafico attraversa la tangente nel punto di flesso; si noti che $f''(0)$ non esiste (v. fig. 4.43).

5.10 Sia $f(x) = x|x|$. Per $x > 0$ è $f(x) = x^2$, convessa; per $x < 0$ è $f(x) = -x^2$, concava; si noti che esiste $f'(0) = 0$; dunque $x = 0$ è un punto di flesso. Tuttavia, $f''(0)$ non esiste.

Figura 4.43.

6 STUDIO DEL GRAFICO DI UNA FUNZIONE

Il calcolo differenziale sviluppato fin qui ci permette ora di affrontare in modo completo un tipico problema, quello di tracciare il grafico di una funzione f di una variabile. Anche se l'avvento dei mezzi elettronici di calcolo ha reso possibile la costruzione automatica dei grafici di funzioni, riteniamo ugualmente utile che lo studente sappia destreggiarsi manualmente in questo compito, armato unicamente di carta e matita. A questo proposito, daremo ora qualche regola generale su come procedere, e illustreremo le più comuni situazioni che si possono verificare, attraverso esempi ed esercizi.

Alcuni passi obbligati:

1. Determinare l'insieme su cui f è definita, ossia il dominio di f, $\mathrm{Dom}(f)$.
2. Quindi, calcolare i limiti (o limiti destri, sinistri) alla frontiera di tale insieme (compresi, eventualmente, $\pm\infty$). In questa fase si determinano gli eventuali asintoti orizzontali e verticali e alcuni eventuali punti di discontinuità.
3. Se per $x \to \pm\infty$ la funzione tende a $\pm\infty$, occorre chiedersi se la funzione presenta un asintoto obliquo. Più in generale, può essere utile una *stima asintotica* che ci dica se la funzione tende a ∞ in modo sopra o sottolineare (e in tal caso certamente *non ha* asintoto obliquo) o in modo lineare (e in tal caso può avere o non avere asintoto obliquo). La stima asintotica all'infinito dà normalmente anche un'informazione sulla *concavità* di f all'infinito.
4. Calcolare $f'(x)$, nei punti in cui esiste. Studiare accuratamente i punti in cui f è continua ma non derivabile e stabilirne la natura (punti angolosi, di flesso a tangente verticale, di cuspide). In punti angolosi o agli estremi del dominio è utile il calcolo di derivate destre o sinistre che determinano la pendenza del grafico in quei punti.
5. Studiare quindi il segno di f' (sul suo insieme di definizione), per ottenere le informazioni sulla monotonia di f e sui suoi punti di massimo e minimo relativi.

Vi sono altre indagini che *in certi casi* possono essere utili:

6. Osservare (quando c'è!) un'eventuale simmetria di f (pari o dispari) e restringere quindi lo studio a $x \geq 0$.

7. Osservare (quando c'è!) l'eventuale periodicità di f e restringere quindi lo studio a un periodo.

8. Determinare gli zeri della funzione e studiarne il segno. Si tenga presente che *solo nei casi particolarmente semplici* è possibile fare questo a priori; in generale, queste informazioni si possono dedurre solo al termine dello studio, come conseguenza delle altre informazioni.

9. Calcolare la derivata seconda e studiarne il segno, per dedurne informazioni sulla concavità e i flessi di f. Si tenga presente che in molti casi l'espressione di f'' è troppo complessa perché se ne riescano a determinare esattamente gli zeri. In questo caso, il calcolo di f'' non aggiunge niente a ciò che un utilizzo attento di limiti, stime asintotiche e derivata prima permette già di affermare.

Un'ultima avvertenza: è utile tracciare il grafico gradualmente, inserendo le informazioni via via raccolte (eventualmente correggendosi), anziché raccogliere tutte le informazioni e poi fare il grafico: i progressi graduali aiutano a controllare la coerenza del procedimento e a capire quali informazioni è ancora utile raccogliere.

Esempi

6.1

$$f(x) = e^{-|x|}\sqrt{x^2 - 5x + 6}$$

Ci aspettiamo: un punto angoloso in $x = 0$, per la presenza di $|x|$; punti a tangente verticale, dove si annulla il radicando.

Insieme di definizione: $x^2 - 5x + 6 \geq 0$; $(x-3)(x-2) \geq 0$; $x \leq 2$; $x \geq 3$.

$\lim\limits_{x \to \pm\infty} f(x) = 0$

Perciò $y = 0$ è asintoto orizzontale per $x \to \pm\infty$.

$f(x) \geq 0$ per ogni x. $f(x) = 0$ per $x = 2, 3$.

Per $x \to 3^+$, $f(x) \sim e^{-3}\sqrt{x-3} \to 0$ e ha tangente verticale in $x - 3$.

Per $x \to 2^-$, $f(x) \sim e^{-2}\sqrt{(2-x)} \to 0$ e ha tangente verticale in $x = 2$.

$$\begin{aligned} f'(x) &= e^{-|x|}\left(-\operatorname{sgn}(x)\sqrt{x^2-5x+6} + \frac{1}{2\sqrt{x^2-5x+6}}(2x-5)\right) = \\ &= \frac{e^{-|x|}}{2\sqrt{x^2-5x+6}}\left(-2\operatorname{sgn}(x)\left(x^2-5x+6\right) + 2x - 5\right) = \\ &= \begin{cases} \dfrac{e^{-x}}{2\sqrt{x^2-5x+6}}\left(-2x^2+12x-17\right) & \text{se } x > 0 \\ \dfrac{e^{x}}{2\sqrt{x^2-5x+6}}\left(2x^2-3x+1\right) & \text{se } x < 0 \end{cases} \end{aligned}$$

Per derivare la funzione contenente il modulo, abbiamo usato la scrittura compatta $|x|' = \operatorname{sgn}(x)$ per $x \neq 0$; solo al termine del calcolo, per rendere più leggibile l'espressione trovata, l'abbiamo riscritta esplicitamente per $x > 0$ e $x < 0$.

Calcoliamo:

$$\lim_{x \to 0^\pm} f'(x) = \begin{cases} -\dfrac{17}{2\sqrt{6}} \\ \dfrac{1}{2\sqrt{6}} \end{cases}$$

$x = 0$ è punto angoloso per f (f non è derivabile e le tangenti, da destra e sinistra, hanno pendenze diverse).

Per $x > 0$ $\quad f'(x) \geq 0$ se $2x^2 - 12x + 17 \leq 0 \quad \dfrac{6-\sqrt{2}}{2} \leq x \leq \dfrac{6+\sqrt{2}}{2}$

$\dfrac{6-\sqrt{2}}{2} \simeq 2.29$; $\dfrac{6+\sqrt{2}}{2} \simeq 3.7$, quindi

per $0 \leq x \leq 2$, $f(x)$ è decrescente;

per $2 \leq x \leq \dfrac{6+\sqrt{2}}{2}$, $f(x)$ è sempre crescente;

per $x \geq \dfrac{6+\sqrt{2}}{2}$, $f(x)$ è sempre decrescente.

$x = \dfrac{6+\sqrt{2}}{2}$ punto di massimo relativo. Il massimo è circa $M = f(3.7) = 0.027$.

Per $x < 0$, $f'(x) \geq 0$ se $2x^2 - 3x + 1 \geq 0, x \leq \dfrac{1}{2}$, $x \geq 1$, quindi:

per $x < 0$, $f(x)$ è sempre crescente.

Si vede anche che $x = 0$ (in cui f non è derivabile) è punto di massimo; il massimo (assoluto) è $M = \sqrt{6}$. Il grafico qualitativo è in figura 4.44. Ci sono anche due punti di flesso, uno in $(0, 2)$ e uno in $(3, +\infty)$. Si omette il calcolo di f''.

Figura 4.44.

6.2 $$f(x) = xe^{\frac{x+2}{x-1}}$$

Insieme di definizione: $x \neq 1$. $f(x) \geq 0$ per $x \geq 0$

Per $x \to 1^{\pm}$, $\dfrac{x+2}{x-1} \to \pm\infty$, $f(x) \to \begin{cases} +\infty \\ 0^+ \end{cases}$

$x = 1$ asintoto verticale per $x \to 1^+$. Per $x \to \pm\infty$, $\dfrac{x+2}{x-1} \to 1$, $f(x) \sim xe$, perciò $f(x) \to \pm\infty$ linearmente. Vediamo se c'è asintoto obliquo:

$$\lim_{x\to\pm\infty} [f(x) - xe] = \lim_{x\to\pm\infty} xe\left[e^{\frac{x+2}{x-1}-1} - 1\right]$$

Ora, $\dfrac{x+2}{x-1} - 1 = \dfrac{3}{x-1} \to 0$, perciò

$$e^{\frac{x+2}{x-1}-1} - 1 \sim \frac{3}{x-1} \qquad xe\left[e^{\frac{x+2}{x-1}-1} - 1\right] \sim \frac{3xe}{x-1} \to 3e$$

Quindi la funzione ha asintoto obliquo $y = xe + 3e$ per $x \to \pm\infty$.

(Si presti attenzione al calcolo del limite precedente: se $\frac{f(x)}{x} \to m$, il limite di $[f(x) - mx]$ è sempre una forma di indeterminazione $[\infty - \infty]$; raccogliendo mx, si ha $mx\left[\frac{f(x)}{mx} - 1\right]$, che è ora una forma di indeterminazione del tipo $[\infty \cdot 0]$; questa va risolta generalmente applicando qualche limite notevole all'espressione tra $[\cdot]$, per darne una stima asintotica).

$$f'(x) = e^{\frac{x+2}{x-1}}\left(1 + x\frac{(x-1)-(x+2)}{(x-1)^2}\right) = \frac{e^{\frac{x+2}{x-1}}}{(x-1)^2}[(x-1)^2 - 3x] =$$

$$= \frac{e^{\frac{x+2}{x-1}}}{(x-1)^2}(x^2 - 5x + 1)$$

Per ogni $x \neq 1$ $f'(x)$ è definita; per $x \to 1^-$, $f'(x) \to 0^-$ (l'esponenziale va a zero più rapidamente di $(x-1)^2$). Il grafico quindi arriva in $x = 1$ con tangente orizzontale, da sinistra.

$$f'(x) \geq 0 \text{ per } x^2 - 5x + 1 \geq 0$$

$$x \geq \frac{5+\sqrt{21}}{2} \simeq 4{,}8 \qquad x \leq \frac{5-\sqrt{21}}{2} \simeq 0{,}2$$

$$x = \frac{5+\sqrt{21}}{2} \text{ punto di min rel.} \qquad x = \frac{5-\sqrt{21}}{2} \text{ punto di max rel.}$$

Grafico:

Figura 4.45.

Esercizi

Studiare le seguenti funzioni, utilizzando i suggerimenti forniti, e tracciarne il grafico.

71 $$e^x \sqrt[3]{\frac{x+3}{x-2}}$$

Suggerimento. Ci aspettiamo: flesso a tangente verticale dove si annulla il radicando, asintoto verticale dove si annulla il denominatore.

72 $$\left|\frac{x^2+x-2}{2x+3}\right|$$

Suggerimento. Ci aspettiamo: punti angolosi dove si annulla il numeratore; asintoto verticale dove si annulla il denominatore. Conviene studiare la funzione senza modulo e poi...

73 $$e^{3/\log x}$$

Nei punti in cui la funzione non è definita ma ha limite (destro o sinistro) finito, studiare anche il limite della derivata prima, per sapere con quale pendenza la curva arriva in tali punti.

74 $$\operatorname{arctg}\left(\frac{x+1}{x-3}\right)+\frac{x}{4}$$

Suggerimento. Questo è un tipico esempio in cui l'equazione $f(x) = 0$ non è risolubile con passaggi algebrici: inutile quindi perder tempo cercando di studiare il segno di $f(x)$ e le sue intersezioni con l'asse x: sarà lo studio complessivo a darci informazioni al riguardo. Studiare in particolare il comportamento di $f(x)$ all'infinito e il punto di discontinuità che si trova: calcolare anche i limiti di $f'(x)$ in tale punto, per tracciare un grafico accurato.

Studiare le seguenti funzioni e tracciarne il grafico.

75 $$\frac{x^2+3x+2}{x+3}$$

76 $$x\,e^{1/(1-|x|)}$$

77 $$\frac{x^3+1}{x+2}$$

78 $$\frac{x}{\log(1+x)}$$

79 $$\frac{x^2+2x-3}{5x^2+x}$$

80 $$\sqrt{\frac{x^2-1}{x+3}}$$

81 $$x\log^{1/3}x$$

82 $$\log\left(\frac{x^2+3x-4}{x+1}\right)$$

83 $$\sqrt{x}\,|\log x|$$

84 $$e^{-\left|\frac{2x+1}{x-2}\right|}$$

85 $$\sqrt{x}\log^2 x$$

86 $$e^{\frac{x+2}{x^2-1}}$$

87 $$x^3\log|x|$$

88 $$e^{\frac{1}{x-1}}\log x$$

89 $$1-|x|e^{-x}$$

90 $$\frac{\log x}{\sqrt[3]{x-1}}$$

Le seguenti funzioni, tratte da modelli reali, contengono qualche costante o parametro. Tracciarne il grafico qualitativo.

91 (*Curva logistica*). Si tracci, al variare del parametro k, il grafico della curva

$$N(t) = \frac{kN_0e^t}{k - N_0 + N_0e^t}$$

dove N_0 è una costante positiva. (Questa funzione rappresenta, sotto opportune ipotesi e in opportune unità di misura, il numero di individui di una data specie all'istante t, se all'istante 0 tale numero è N_0; la costante k è detta *capacità dell'ambiente*; questo esempio sarà ripreso nel vol. 2°, trattando le equazioni differenziali).

92 (*Funzione di Langevin*). Si tracci per $y > 0$ il grafico della funzione:

$$L(y) = m_0\left(\frac{\operatorname{Ch} y}{\operatorname{Sh} y} - \frac{1}{y}\right)$$

dove m_0 è una costante positiva. Questa funzione entra nella descrizione dei fenomeni di *polarizzazione magnetica per orientamento*.

93 Siano γ, ω_0 costanti positive. Tracciare il grafico delle due funzioni:

$$n_1(\omega) = 1 + \frac{\omega_0^2 - \omega^2}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2} \qquad n_2(\omega) = \frac{\gamma\omega}{(\omega_0^2 - \omega^2)^2 + \gamma^2\omega^2}$$

(queste funzioni entrano nella descrizione degli indici di rifrazione delle onde elettromagnetiche nei dielettrici).

94 (*Effetto Doppler relativistico*). Se la sorgente e il ricevitore di un'onda si allontanano con velocità relativa v e l'onda è emessa con frequenza ν_0, la frequenza ricevuta sarà

$$\nu = \nu_0\frac{\sqrt{1 - v^2/c^2}}{1 + v/c}$$

dove c è la velocità della luce (costante positiva). Si tracci il grafico della funzione $\nu = \nu(v)$ nel dominio fisicamente ammissibile, $|v| < c$.

95 (*Angolo di rifrazione*). Dalla legge di Snell della rifrazione (cfr. par. 4.3) si ricava che nel passaggio di un raggio di luce da un mezzo a un altro, l'angolo θ_r di rifrazione dipende dall'angolo θ_i di incidenza in base alla legge:

$$\theta_r = \arcsin\left(\frac{\sin\theta_i}{n}\right)$$

dove n =indice di rifrazione relativo del mezzo 2 rispetto al mezzo 1, può essere maggiore o minore di 1 (ma è positivo). Studiare questa funzione di θ_i e tracciarne il grafico, distinguendo i casi $n > 1, n < 1$. L'angolo θ_i è sempre compreso in $\left(0, \frac{\pi}{2}\right)$. (In uno dei due casi è necessaria una limitazione più stretta e, per angoli di incidenza non ammissibili, *non si ha rifrazione*).

96 (*Distribuzione spettrale della radiazione emessa da un corpo nero*). Nell'esempio 3.40 del capitolo 3 si è studiata mediante stime asintotiche la funzione

$$f(\lambda) = \frac{8\pi hc\lambda^{-5}}{e^{hc/\lambda kT} - 1}$$

(h, c, k costanti positive; T =temperatura assoluta, parametro positivo; λ =lunghezza d'onda). Se ne completi ora lo studio calcolando la derivata prima e dimostrando che il punto di massimo di f cade (ad esempio) nell'intervallo $\left(\frac{hc}{5kT}, \frac{hc}{4kT}\right)$. Ciò significa che la lunghezza d'onda per cui è massima la densità di energia raggiante è inversamente proporzionale alla temperatura assoluta T (legge dello spostamento di Wien).

Infine, si pensi ora λ fissato e si tracci il grafico di f come funzione di T.

7 CALCOLO DIFFERENZIALE E APPROSSIMAZIONI

7.1 Differenziale e approssimazione lineare. Il simbolo di "o piccolo"

Un'operazione molto frequente sia in matematica che nelle sue applicazioni è quella di *linearizzazione*.

Sostanzialmente si tratta di approssimare una data quantità, che dipende in modo non lineare da una o più variabili, con una che dipende da queste linearmente.

Naturalmente, come in tutti i tipi di approssimazione, occorre fornire informazioni (qualitative o quantitative) sull'errore commesso.

Un esempio elementare di linearizzazione consiste nell'approssimare l'incremento subìto da una data funzione f, in conseguenza di una variazione del suo argomento da x_0 a $x_0 + dx$, sostituendo alla funzione stessa la retta tangente nel punto x_0.

Precisamente, sia $f : (a, b) \longrightarrow \mathbb{R}$ derivabile in un punto x_0 e diamo a x_0 un incremento dx (che pensiamo molto piccolo in valore assoluto, cioè $|dx| \ll 1$). In conseguenza, f subisce un incremento

$$\Delta f = f(x_0 + dx) - f(x_0)$$

che, in generale, pensando x_0 fissato, non è proporzionale a dx ossia non è lineare rispetto a dx.

Come indicato in figura 4.46, sostituiamo al grafico di f, vicino a x, la retta tangente nel punto $P = (x_0, f(x_0))$.

Figura 4.46. $df = \operatorname{tg}\alpha\, dx$.

L'incremento valutato lungo la retta tangente è pari alla lunghezza (con segno!) del segmento QR e cioè uguale a $\operatorname{tg}\alpha \cdot dx$, ovvero a $f'(x_0)dx$, ricordando che $f'(x_0) = \operatorname{tg}\alpha$. Tale incremento, proporzionale a dx, prende il nome di *differenziale* di f nel punto x_0 e si indica col simbolo $df(x_0)$:

$$\boxed{df(x_0) = f'(x_0)dx}$$

Qual è l'errore commesso nell'approssimazione? In altri termini, che cosa si può dire sulla differenza tra Δf e $df(x_0)$?

Per rispondere, è sufficiente ricorrere alla definizione di derivata; la formula

$$\lim_{dx \to 0} \frac{f(x_0 + dx) - f(x_0)}{dx} = f'(x_0)$$

si può scrivere

$$\frac{f(x_0+dx)-f(x_0)}{dx}=f'(x_0)+\varepsilon(dx) \tag{7.1}$$

dove $\varepsilon(dx)$ indica una quantità che tende a zero se $dx \to 0$; è cioè un *infinitesimo* per $dx \to 0$.

Moltiplicando per dx, dalla (7.1) si ottiene infine

$$f(x_0+dx)-f(x_0)=f'(x_0)dx+dx\cdot\varepsilon(dx)\,. \tag{7.2}$$

Dalla (7.2), si ricava subito

$$\Delta f-df(x_0)=dx\cdot\varepsilon(dx)$$

La quantità $dx\cdot\varepsilon\,(dx)$ è una quantità che, divisa per dx, tende a zero: questo significa che $dx\cdot\varepsilon(dx)$ è *infinitesimo di ordine superiore rispetto a* dx. Introduciamo un simbolo utile a indicare sinteticamente questa circostanza:

Definizione 4.9 Date due funzioni $f\,(x)\,, g\,(x)$, definite in un intorno di x_0, si dice che

$$f(x)=o\,(g(x)) \text{ per } x\to x_0$$

(si legge "$f(x)$ è o piccolo di $g(x)$") se

$$\frac{f(x)}{g(x)}\to 0 \text{ per } x\to x_0$$

Se $g(x)$ è un infinitesimo, dire che $f(x)=o\,(g(x))$ significa che $f(x)$ è infinitesimo di ordine superiore rispetto a $g(x)$. Ad esempio, per $x\to 0$, $x^2=o(x)$, $e^{-1/x^2}=o(x^4)$, ecc. In particolare, il simbolo $o(1)$ denota semplicemente una quantità infinitesima. ■

Con questa simbologia, si scrive:

$$f\,(x_0+dx)-f\,(x_0)=f'\,(x_0)\,dx+o\,(dx) \text{ per } dx\to 0 \tag{7.3}$$

o, più sinteticamente:

$$\Delta f\,(x_0)=df\,(x_0)+o\,(dx) \text{ per } dx\to 0$$

La (7.3) esprime quella che si chiama un'approssimazione di Δf al prim'ordine. Si scrive anche, talvolta,

$$\Delta f\approx df \text{ al prim'ordine, vicino a } x_0$$

Le regole di calcolo dei differenziali sono analoghe a quelle di derivazione:

$$\boxed{\begin{aligned} &d(f\pm g)=df\pm dg\\ &d(f\cdot g)=gdf+fdg\\ &d\Big(\frac{f}{g}\Big)=\frac{gdf-fdg}{g^2}\end{aligned}}$$

Si noti poi che $\frac{df}{f} = d\log|f| =$ differenziale logaritmico di f.

Esempi

7.1 Consideriamo

$$f(x) = \sqrt{1+x}$$

e approssimiamo al primo ordine $f(x) - f(0) = \sqrt{1+x} - 1$.

Si noti che qui $x_0 = 0$ e quindi x coincide con l'incremento dx. Essendo

$$f'(x) = \frac{1}{2\sqrt{1+x}} \qquad f'(0) = \frac{1}{2}$$

si ha

$$\boxed{\sqrt{1+x} - 1 \approx \frac{1}{2}x} \quad \textit{al primo ordine} \tag{7.4}$$

o anche

$$\sqrt{1+x} = 1 + \frac{1}{2}x + o(x)\,, \text{ per } x \to 0$$

Più in generale, considerando $f(x) = (1+x)^\alpha$, si ha

$$f'(x) = \alpha(1+x)^{\alpha-1} \qquad f'(0) = \alpha$$

e quindi ritroviamo (v. cap. 3, par. 3.3, formula (3.8))

$$\boxed{(1+x)^\alpha - 1 \approx \alpha x} \quad \textit{al primo ordine} \tag{7.5}$$

7.2 *Potenziale di un dipolo elettrico.* Un'interessante applicazione della (7.4) si trova nel calcolo dell'approssimazione al primo ordine del potenziale creato da un dipolo elettrico. Con riferimento alla figura 4.47, consideriamo un dipolo costituito da due cariche $+q$ e $-q$ concentrate in due punti R e S rispettivamente, a distanza $2a$ l'una dall'altra.

Figura 4.47.

Se P è un punto a distanza r dal punto medio O e a distanza r_1 da R, r_2 da S, il potenziale in P è assegnato dalla formula (in opportune unità di misura)

$$V(P) = \frac{q}{4\pi\varepsilon_0}\left(\frac{1}{r_1} - \frac{1}{r_2}\right)$$

Vogliamo valutare il potenziale quando $a \ll r$ (a molto minore di r), mantenendo costante il momento di dipolo $\mu = a \cdot q$.

Dal teorema di Carnot in trigonometria, si ha:

$$r_1^2 = r^2 + a^2 - 2ar\cos\theta = r^2\Big[1 + \Big(-2\frac{a}{r}\cos\theta + \frac{a^2}{r^2}\Big)\Big]$$

$$r_2^2 = r^2 + a^2 + 2ar\cos\theta = r^2\Big[1 + \Big(2\frac{a}{r}\cos\theta + \frac{a^2}{r^2}\Big)\Big]$$

Ricordiamo ora che, se x è vicino a zero, $\sqrt{1+x} \approx 1 + \frac{1}{2}x$ al primo ordine e quindi, essendo $\frac{a}{r}$ vicino a zero e $\frac{a}{r} \ll \frac{a^2}{r^2}$, si può scrivere, al primo ordine:

$$r_1 \approx r\Big(1 - \frac{a}{r}\cos\theta\Big) \qquad r_2 \approx r\Big(1 + \frac{a}{r}\cos\theta\Big)$$

Ne segue che, al primo ordine, $r_1 - r_2 \approx 2a\cos\theta$, $r_1 r_2 \approx r^2$ e infine

$$V(P) = \frac{q}{4\pi\varepsilon_0}\Big(\frac{1}{r_1} - \frac{1}{r_2}\Big) = \frac{q}{4\pi\varepsilon_0}\frac{r_2 - r_1}{r_1 r_2} \approx \frac{2a\cos\theta}{r^2}\frac{q}{4\pi\varepsilon_0} = \frac{2\mu\cos\theta}{4\pi\varepsilon_0 r^2}$$

che è l'approssimazione cercata. Si noti la dipendenza da r e dall'angolo θ.

7.3 *Oscillazioni del pendolo.* Una massa puntiforme m oscilla, appesa a un filo di lunghezza l. Detto $\theta(t)$ l'angolo formato dal filo con la verticale all'istante t, la forza agente sul punto nella direzione del moto è $-mg\sin\theta$ (v. fig. 4.48)

D'altro canto, questa forza deve uguagliare, per la seconda legge della dinamica, il prodotto della massa per l'accelerazione, pari a $ml\theta''$. Si ottiene quindi l'equazione

$$ml\theta'' = -mg\sin\theta$$

ossia

$$\theta'' = -\frac{g}{l}\sin\theta$$

Figura 4.48.

Tale equazione è difficile da studiare direttamente. Se però vogliamo studiare solo le piccole oscillazioni del punto attorno all'equilibrio, potremo assumere che θ sia piccolo; si può allora studiare in prima approssimazione l'equazione linearizzata

$$\theta'' = -\frac{g}{l}\theta$$

(Abbiamo sfruttato $\sin\theta \approx \theta$, al prim'ordine, per θ piccolo). Questa può essere risolta completamente. Si trova (ponendo $\omega^2 = g/l$)

$$\theta(t) = A\cos(\omega t + \theta_0)$$

con θ_0, A costanti da determinarsi mediante le condizioni iniziali del sistema. (Lo studente verifichi che le funzioni $\theta(t)$ trovate risolvono effettivamente l'equazione. Nelsecondo volume, trattando le equazioni differenziali si proverà che effettivamente non vi sono altre soluzioni).

7.4 *Energia cinetica classica e relativistica.* Nella meccanica classica, l'energia cinetica di un punto materiale di massa m in movimento, con velocità scalare v, è data da

$$E_c = \frac{1}{2}mv^2$$

La teoria della relatività modifica varie definizioni e leggi della meccanica classica. Anzitutto, la massa m non è più una costante, ma dipende dalla velocità, secondo la legge:

$$m = \frac{m_0}{\sqrt{1 - \frac{v^2}{c^2}}}$$

dove c è la velocità della luce nel vuoto e m_0 è la "massa a riposo" della particella (questa è costante); in secondo luogo, vale la famosa equazione dell'energia:

$$E = mc^2$$

L'energia della particella a riposo è invece

$$E_0 = m_0 c^2$$

e l'energia cinetica relativistica è per definizione pari a $(E - E_0)$. Sostituiamo ad m la sua espressione e troviamo E come funzione della velocità della particella:

$$E = \frac{m_0 c^2}{\sqrt{1 - \frac{v^2}{c^2}}}$$

Nell'ipotesi che la velocità della particella sia piccola rispetto a quella, elevatissima, della luce, v/c sarà molto piccolo. Linearizziamo allora l'espressione trovata come funzione di $(v/c)^2$ (usando la (7.5) con $\alpha = -1/2$):

$$E = m_0 c^2 \cdot \left(1 - \left(\frac{v}{c}\right)^2\right)^{-1/2} \approx m_0 c^2 \left(1 + \frac{1}{2}\frac{v^2}{c^2}\right) =$$
$$= m_0 c^2 + \frac{1}{2} m_0 v^2 = E_0 + E_c$$

dove E_0 è l'energia a riposo della particella, ed E_c l'energia cinetica classica. L'energia cinetica relativistica, che per definizione è pari a $E - E_0$, nell'approssimazione valida per velocità piccole rispetto a quella della luce coincide dunque con l'energia cinetica classica.

7.2 Limiti notevoli e sviluppi

Abbiamo visto che il processo di linearizzazione permette di approssimare una funzione derivabile, localmente, mediante la sua retta tangente, ossia una funzione lineare. D'altro canto, anche i "limiti notevoli" a suo tempo incontrati permettevano di scrivere simili risultati di approssimazione, per certe funzioni particolari. Ad esempio, per la funzione $y = \sin x$, il processo di linearizzazione, in $x = 0$, porta a scrivere

$$\sin x = x + o(x)$$

mentre a suo tempo abbiamo visto che

$$\sin x \sim x \text{ per } x \to 0$$

Che relazione c'è tra queste due affermazioni? Sono in effetti equivalenti, come mostra il prossimo fatto generale:

Teorema 4.15 (Relazione tra "o piccolo" e "asintotico") *Vale la seguente equivalenza:*

$$\textit{Per } x \to x_0,\ f(x) \sim g(x) \textit{ se e solo se } f(x) = g(x) + o(g(x))$$

Il lettore per esercizio dimostri le due implicazioni contenute nell'affermazione precedente: basta usare le definizioni.

Questo permette di rileggere i limiti notevoli che abbiamo finora espresso con stime asintotiche, mediante uguaglianze che coinvolgono "o piccolo". Ad esempio,

$$\text{per } x \to 0:$$
$$\sin x \sim x, \text{ quindi } \sin x = x + o(x);$$
$$1 - \cos x \sim \frac{1}{2}x^2, \text{ quindi } 1 - \cos x = \frac{1}{2}x^2 + o\left(x^2\right).$$

Riflettiamo sull'ultima: l'uguaglianza $1 - \cos x = \frac{1}{2}x^2 + o\left(x^2\right)$ si può riscrivere nella forma:

$$\cos x = 1 - \frac{1}{2}x^2 + o\left(x^2\right) \tag{7.6}$$

(questo è ovvio: abbiamo spostato dei termini da un membro a un altro di una uguaglianza). Notiamo, invece, che la relazione asintotica $1 - \cos x \sim \frac{1}{2}x^2$ non dice la stessa cosa della relazione $\cos x \sim 1 - \frac{1}{2}x^2$. Infatti, per $x \to 0$, $1 - \frac{1}{2}x^2 \sim 1$, perciò la stima $\cos x \sim 1 - \frac{1}{2}x^2$ contiene la stessa informazione che dire semplicemente "$\cos x \to 1$", ed è altrettanto vera che la stima $\cos x \sim 1 + 40x^{13}$! Questa osservazione mostra un vantaggio del simbolo di "o piccolo" rispetto a quello di asintotico: un'uguaglianza si può riscrivere in vari modi, è più facile da usare senza errori, rispetto ad una stima asintotica.

Per esercizio, il lettore riscriva mediante il simbolo di "o piccolo" i limiti notevoli che riguardano le funzioni $\sqrt{1+x}, e^x, \log(1+x), x^\alpha$, per $x \to 0$.

7.3 Formula di Taylor-MacLaurin con resto secondo Peano

Vogliamo ora generalizzare il procedimento di "approssimazione per linearizzazione" a quello di "approssimazione polinomiale". In altre parole, ci chiediamo: data una funzione, derivabile tutte le volte che sarà necessario, esiste un polinomio che, nell'intorno di un punto fissato, approssima la funzione meglio della sua retta tangente?

L'esempio (7.6) della funzione coseno si può rileggere in tal senso: la funzione $\cos x$ è approssimata dalla parabola $y = 1 - \frac{1}{2}x^2$ meglio che dalla retta tangente $y = 1$, per

Figura 4.49. La funzione $\cos x$ è approssimata dal polinomio $y = 1 - \frac{1}{2}x^2$ meglio che dalla retta $y = 1$, vicino ad $x = 0$.

$x \to 0$: infatti, lo scarto tra la funzione e questo polinomio di secondo grado è $o\left(x^2\right)$, cioè tende a zero più rapidamente di x^2 (e non solo più rapidamente di x).
Per semplicità, cominciamo a ragionare nell'intorno del punto $x_0 = 0$.

Procediamo in 2 passi:

a. Individuiamo un polinomio "candidato" ad approssimare bene la funzione, cercando un polinomio che abbia tutte le derivate fino all'ordine n uguali a quelle di $f(x)$, nel punto $x = 0$. Affinché questo sia sempre possibile, il grado del polinomio dev'essere almeno n. (Infatti, la derivata n-esima di un polinomio di grado minore di n è identicamente nulla, quindi non potrebbe essere uguale a $f^{(n)}(0)$, in generale). Facendo i calcoli, si trova che:

Teorema 4.16 (Polinomio di MacLaurin) *Data una funzione f derivabile n volte in $x = 0$, esiste uno e un sol polinomio di grado $\leq n$, chiamiamolo T_n, con la proprietà che:*

$$T_n(0) = f(0), T_n'(0) = f'(0), \ldots, T_n^{(n)}(0) = f^{(n)}(0)$$

e questo polinomio, detto polinomio di MacLaurin di $f(x)$ di grado n, è:

$$T_n(x) = f(0) + xf'(0) + \frac{1}{2}x^2 f''(0) + \frac{1}{3!}x^3 f'''(0) + \ldots + \frac{1}{n!}x^n f^{(n)}(0) =$$
$$= \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k$$

(avendo posto $f^{(0)} = f$).

Notiamo che il polinomio T_n assegnato dal teorema precedente, solitamente è proprio di grado n, ma può avere grado minore se $f^{(n)}(0) = 0$.

È interessante osservare la coerenza dimensionale della formula che assegna il polinomio di MacLaurin: supponiamo che f abbia le dimensioni di una lunghezza $[L]$ e x abbia le dimensioni di un tempo $[T]$. Allora $[f'] = [L] \cdot [T]^{-1}, [f''] = [L] \cdot [T]^{-2}$, e in generale $[f^{(k)}] = [L] \cdot [T]^{-k}$. Questo significa che ogni addendo $\frac{f^{(k)}(0)}{k!} x^k$ del polinomio $T_n(x)$ ha la struttura

(costante adimensionale) $\cdot$ (fattore di dimens. $[L] \cdot [T]^{-k}$) $\cdot$ (fattore di dimens. $[T]^k$)

Perciò il polinomio di MacLaurin ha le dimensioni di una lunghezza, esattamente come la funzione f che si vuole approssimare.

b. Proviamo ora che il polinomio trovato approssima bene $f(x)$, in un intorno di $x = 0$. Precisamente, vale il:

Teorema 4.17 (Formula di MacLaurin all'ordine n, con resto secondo Peano) *Sia $f : (a, b) \to \mathbb{R}$ derivabile n volte in $0 \in (a, b)$. Allora*

$$f(x) = T_n(x) + o\left(x^n\right) \text{ per } x \to 0$$

La formula precedente si dice "formula di MacLaurin di ordine n, con resto secondo Peano".

La formula ha la struttura:

funzione da approssimare = polinomio approssimante + errore di approssimazione

dove l'errore di approssimazione è il termine $o(x^n)$, detto *resto secondo Peano*.

Per $x \to 0$, il resto secondo Peano è tanto più piccolo quanto maggiore è n. Lo spirito della formula è dunque il seguente: conoscendo un numero abbastanza alto di derivate di f nel punto $x = 0$, si può approssimare sempre meglio f, in un intorno di $x = 0$.

DIMOSTRAZIONE. Proviamo per semplicità il teorema nel caso $n = 2$, ossia:

$$f(x) = f(0) + xf'(0) + \frac{1}{2}x^2 f''(0) + o(x^2) \text{ per } x \to 0$$

Occorre provare che

$$f(x) - \left[f(0) + xf'(0) + \frac{1}{2}x^2 f''(0)\right] = o(x^2) \text{ per } x \to 0$$

ossia (per definizione di "o piccolo") che:

$$\lim_{x\to 0} \frac{f(x) - \left[f(0) + xf'(0) + \frac{1}{2}x^2 f''(0)\right]}{x^2} = 0.$$

Questo limite dà una forma di indeterminazione [0/0], che calcoliamo con De L'Hospital:

$$\lim_{x\to 0} \frac{f'(x) - [f'(0) + xf''(0)]}{2x}$$

dà ancora una forma [0/0]. Osserviamo ora che l'ultimo limite trovato è zero, per il seguente motivo. Per ipotesi, f è derivabile due volte, quindi f' è derivabile; se applichiamo il teorema di linearizzazione alla funzione $f'(x)$ otteniamo

$$f'(x) = f'(0) + xf''(0) + o(x)$$

da cui segue che

$$\frac{f'(x) - [f'(0) + xf''(0)]}{x} \to 0.$$

Il caso generale (n qualsiasi) si può dimostrare *per induzione su* n (v, capitolo 1, par. 7). Abbiamo giò provato che la tesi vale per $n = 2$. Supponiamo ora di sapere che per qualsiasi funzione derivabile $n - 1$ volte in $x = 0$ vale la formula di Mac Laurin all'ordine n, e proponiamoci di dimostrare che

$$f(x) = T_n(x) + o(x^n) \text{ per } x \to 0, \text{ ossia che}$$

$$\lim_{x\to 0} \frac{f(x) - T_n(x)}{x^n} = 0. \tag{7.7}$$

Applicando il teorema di De L'Hospital, ci si trova a calcolare si ha

$$\lim_{x\to 0} \frac{f'(x) - T'_{n,f}(x)}{nx^{n-1}}. \tag{7.8}$$

Osserviamo ora che la derivata del polinomio $T_n(x)$ di f non è altro che il polinomio $T_{n-1}(x)$ della funzione f':

$$T'_{n,f}(x) = T_{n-1,f'}(x)$$

come si verifica direttamente dalla definizione di T_n. D'altro canto, per l'ipotesi induttiva, applicata alla funzione f', sappiamo che

$$f'(x) = T_{n-1,f'}(x) + o\left(x^{n-1}\right), \text{ ossia}$$

$$\lim_{x\to 0} \frac{f'(x) - T_{n-1,f'}(x)}{x^{n-1}} = 0,$$

quindi il limite (7.8) è zero e, per il teorema di De L'Hospital, anche il limite (7.7) è zero, come volevamo dimostrare.

Formula di Taylor all'ordine n

Tutto questo discorso si può generalizzare ad un punto $x_0 \neq 0$: data una funzione derivabile n volte in x_0, si può costruire il suo polinomio di Taylor

$$T_{n,x_0}(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

e si può dimostrare che vale un analogo risultato di approssimazione locale:

Teorema 4.18 (Formula di Taylor all'ordine n, con resto secondo Peano) *Sia $f : (a, b) \to \mathbb{R}$ derivabile n volte in $x_0 \in (a, b)$. Allora*

$$f(x) = T_{n,x_0}(x) + o\left((x - x_0)^n\right) \text{ per } x \to x_0.$$

Esempi

Approssimazione all'ordine n

7.5 Sia $f(x) = e^x$. Essendo $f^{(n)}(x) = e^x$, $f^{(n)}(0) = 1$ per ogni n, si ha:

$$e^x = T_{n,0}(x) + o(x^n) = 1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \ldots + \frac{x^n}{n!} + o(x^n), \text{ per } x \to 0$$

7.6 Sia $f(x) = \sin x$. Si ha:

$$f'(x) = \cos x \qquad f''(x) = -\sin x \qquad f'''(x) = -\cos x$$

e poi

$$f^{(iv)}(x) = \sin x = f(x) \qquad f^{(v)}(x) = \cos x = f'(x) \quad \text{ecc.}$$

Ne segue che per $n = 2k$, pari, le derivate di ordine n sono nulle per $x = 0$, mentre quelle di ordine $n = 2k + 1$, dispari, sono alternativamente $+1$ e -1. In sintesi:

$$f^{(2k)}(0) = 0 \qquad f^{(2k+1)}(0) = (-1)^k \qquad (k \geq 0)$$

I polinomi di MacLaurin della funzione seno contengono perciò solo termini di grado dispari:

$$\sin x = T_1(x) + o(x) = x + o(x)$$

$$\sin x = T_3(x) + o(x^3) = x - \frac{x^3}{6} + o(x^3)$$

$$\sin x = T_5(x) + o(x^5) = x - \frac{x^3}{6} + \frac{x^5}{120} + o(x^5)$$

$$\vdots$$

$$\sin x = T_{2n+1}(x) + o(x^{2n+1}) = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+1})$$

La figura 4.50 illustra il miglioramento dell'approssimazione di $\sin x$ mediante i suoi polinomi di MacLaurin, all'aumentare del grado del polinomio.

Figura 4.50.

Nel quadro che segue abbiamo raggruppato gli sviluppi di MacLaurin di alcune funzioni elementari, con il resto di Peano.

$$e^x = 1 + x + \frac{x^2}{2!} + \ldots + \frac{x^n}{n!} + o(x^n)$$

$$\sin x = x - \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+1})$$

$$\cos x = 1 - \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots + (-1)^n \frac{x^{2n}}{(2n)!} + o(x^{2n})$$

$$\operatorname{Sh} x = x + \frac{x^3}{3!} + \frac{x^5}{5!} + \ldots + \frac{x^{2n+1}}{(2n+1)!} + o(x^{2n+1})$$

$$\operatorname{Ch} x = 1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \ldots + \frac{x^{2n}}{(2n)!} + o(x^{2n})$$

$$\ln(1+x) = x - \frac{x^2}{2} + \frac{x^3}{3} + \ldots + (-1)^{n-1} \frac{x^n}{n} + o(x^n)$$

$$(1+x)^\alpha = 1 + \alpha x + \frac{\alpha(\alpha-1)}{2} x^2 + \ldots + \frac{\alpha(\alpha-1)\ldots(\alpha-n+1)}{n!} x^n + o(x^n)$$

Proprietà del simbolo di "o piccolo"

Quando si opera su sviluppi di Taylor-MacLaurin con resto secondo Peano, ad esempio per calcolare limiti, ci si trova ad eseguire calcoli che coinvolgono $o(x^n)$. Per procedere correttamente è utile comprendere alcune proprietà del simbolo di "o piccolo". Ricordiamo che

$$f(x) = o(g(x)) \text{ per } x \to x_0 \text{ se } \lim_{x \to x_0} \frac{f(x)}{g(x)} = 0.$$

Ad esempio, per $x \to 0$, $x^2 = o(x)$; $x^3 = o(x)$; $x^3 = o(x^2)$.

Il simbolo di $o(\ldots)$ non denota quindi una particolare funzione, ma qualsiasi funzione abbia la proprietà espressa dalla definizione. Da questo fatto seguono alcune proprietà della relazione di "o piccolo" che possono apparire "strane". Ne illustriamo alcune su esempi che si incontrano frequentemente:

$$o(x) - o(x) = o(x) \qquad \text{(non fa zero!)}$$

$$o\left(-3x^2\right) = o\left(x^2\right), \qquad -o(x) = o(x) \qquad \text{(coefficienti numerici moltiplicativi sono irrilevanti)}$$

Per $x \to 0, o(x) + o\left(x^2\right) = o(x)$ (l'errore più grossolano, cioè $o(x)$, ingloba quello più fine).

Analogamente si avrebbe:

$$\text{per } x \to +\infty, o(x) + o\left(x^2\right) = o\left(x^2\right)$$

Il simbolo di "o piccolo" si comporta al seguente modo con somme e prodotti (verificarlo in base alla definizione):

$$f \cdot o(g) = o(f \cdot g)\text{, ad esempio:}$$

$$xo\left(x^2\right) = o\left(x^3\right) \qquad \frac{o\left(x^3\right)}{x} = o\left(x^2\right) \qquad \frac{o\left(x^2\right)}{x^2} = o(1)$$

(ricordiamo che $o(1)$ indica semplicemente una funzione che tende a zero per $x \to x_0$, il punto considerato). Ed anche:

$$o(f) \cdot o(g) = o(f \cdot g)\text{, ad esempio:}$$

$$o(x)\, o\left(x^2\right) = o\left(x^3\right)$$

7.4 Formula di Taylor-MacLaurin con resto secondo Lagrange

Nella formula di Taylor con resto secondo Peano, l'informazione che abbiamo sull'errore commesso nell'approssimare f col suo polinomio di Taylor-MacLaurin è di tipo dinamico: al tendere a zero dell'incremento $(x - x_0)$, sappiamo che il resto tende a zero più rapidamente di $(x - x_0)^n$. Questa informazione, come abbiamo visto, è utile ad esempio nel calcolo dei limiti. Per un valore fissato dell'incremento $(x - x_0)$, però, la formula di Taylor con resto secondo Peano non dice nulla sull'entità dell'errore commesso. In varie questioni di calcolo approssimato è essenziale invece stimare l'errore che si commette approssimando una funzione col suo polinomio di Taylor-MacLaurin, quando l'incremento $(x - x_0)$ ha un valore fissato, o un valore che non supera una soglia fissata.

Ad esempio, sappiamo che il polinomio di MacLaurin al 3° ordine di e^x è

$$T_3(x) = 1 + x + \frac{1}{2}x^2 + \frac{1}{6}x^3$$

Se volessimo utilizzare questo fatto per calcolare un valore approssimato di $\sqrt{e}$, ossia pensassimo di approssimare

$$e^{1/2} \text{ con } T_3\left(\frac{1}{2}\right) = \frac{79}{48}$$

come potremmo stimare a priori (ossia senza conoscere già il valore vero di $\sqrt{e}$) di quanto stiamo sbagliando?
A questo tipo di problema risponde il prossimo risultato, che dà un modo alternativo di quantificare l'errore di approssimazione commesso:

Teorema 4.19 (Formula di Taylor all'ordine n, con resto secondo Lagrange) *Sia $f : [a, b] \to \mathbb{R}$ derivabile $n+1$ volte in $[a, b]$, e sia $x_0 \in [a, b]$. Allora esiste un punto c compreso tra x_0 e x tale che*

$$f(x) = T_{n,x_0}(x) + \frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}. \tag{7.9}$$

La formula precedente si dice "formula di Taylor (o di MacLaurin se $x_0 = 0$) di ordine n, con resto secondo Lagrange".

La formula ha la struttura:

funzione da approssimare = polinomio approssimante + errore di approssimazione

dove l'errore di approssimazione è il termine $\frac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}$, detto *resto secondo Lagrange*.

Per $n = 0$ la formula (7.9) coincide col Teorema di Lagrange.

Si noti che il punto c dipende da x_0, x e n, ed è compreso tra x_0 e x. Se si riesce a dimostrare che $\left|f^{(n+1)}(t)\right| \leq M$ per ogni t compreso tra x_0 e x, allora la formula di Taylor con resto secondo Lagrange dice che

$$|f(x) - T_{n,x_0}(x)| \leq \frac{M}{(n+1)!}|x - x_0|^{n+1}$$

che è appunto una stima dell'errore di approssimazione commesso.

Esempio

7.7 Torniamo al problema di approssimare $\sqrt{e}$ mediante la formula di Taylor. Ponendo, nella (7.9), $f(x) = e^x, x_0 = 0, x = \frac{1}{2}, n = 3$, otteniamo:

$$\sqrt{e} = T_3\left(\frac{1}{2}\right) + \frac{e^c}{4!}\left(\frac{1}{2}\right)^4.$$

Poiché per $0 \leq c \leq \frac{1}{2}$ è $e^c \leq \sqrt{e} < \sqrt{3}$ [1], si ottiene

$$\left|\sqrt{e} - T_3\left(\frac{1}{2}\right)\right| \leq \frac{\sqrt{3}}{2^4 \cdot 4!} \simeq 0.0045$$

Questo significa che approssimando $\sqrt{e}$ con il valore $\frac{79}{48}$ si commette un errore non superiore a 0.0045.

[1] Abbiamo usato la disuguaglianza grossolana $e < 3$ per non cadere in un circolo vizioso: per stimare il valore di $\sqrt{e}$ sembrerebbe necessario sapere già quanto vale $\sqrt{e}$! Invece, basta conoscere un valore che maggiora $\sqrt{e}$, e questo valore è, ad esempio, $\sqrt{3}$.

DIMOSTRAZIONE DEL TEOREMA 4.19 Proviamo la (7.9) nel caso $n = 1$. Ponendo per comodità $x_0 = a, x = b$ l'enunciato diviene: se $f : [a, b] \to \mathbb{R}$ è derivabile 2 volte in $[a, b]$, allora esiste un punto $c \in [a, b]$ tale che

$$f(b) = f(a) + f'(a)(b-a) + \frac{1}{2}f''(c)(b-a)^2. \tag{7.10}$$

Poniamo $f(b) - \{f(a) + f'(a)(b-a)\} = k(b-a)^2$ e cerchiamo di determinare la forma di k. Consideriamo la funzione

$$g(x) = f(b) - f(x) - f'(x)(b-x) - k(b-x)^2$$

e applichiamo ad essa il teorema del valor medio. Poiché $g(b) = g(a) = 0$ (la seconda uguaglianza segue dalla definizione di k) si trova che esiste $c \in (a, b)$ tale che

$$0 = \frac{g(b) - g(a)}{b-a} = g'(c)$$

Ma $g'(x) = -f'(x) - f''(x)(b-x) + f'(x) + 2k(b-x) = (b-x)[2k - f''(x)]$ e quindi $g'(c) = 0$ implica, essendo $c \neq b$:

$$k = \frac{1}{2}f''(c)$$

Con lo stesso metodo si può dimostrare la (7.9) per n qualsiasi. Definiamo:

$$g(x) = f(b) - \sum_{j=0}^{n} \frac{f^{(j)}(x)(b-x)^j}{j!} - k(b-x)^{n+1}$$

con k definito implicitamente dall'identità

$$f(b) - T_{n,a}(b) = k(b-a)^{n+1}. \tag{7.11}$$

Ora si procede così:

a) Si verifica che $g(b) = g(a) = 0$:

$$g(b) = f(b) - f(b) = 0$$
$$g(a) = f(b) - T_{n,a}(b) - k(b-a)^{n+1} = 0 \text{ per definizione di } k.$$

b) Si applica il teorema di Lagrange a g in $[a, b]$, e si mostra che l'affermazione "esiste $c \in (a, b)$ tale che $g'(c) = 0$" è esattamente la (7.9). Cominciamo a calcolare g':

$$\begin{aligned} g'(x) &= -\sum_{j=0}^{n} \frac{f^{(j+1)}(x)(b-x)^j}{j!} + \sum_{j=1}^{n} \frac{f^{(j)}(x)(b-x)^{j-1}}{(j-1)!} + k(n+1)(b-x)^n \\ &= -\sum_{j=0}^{n} \frac{f^{(j+1)}(x)(b-x)^j}{j!} + \sum_{j=0}^{n-1} \frac{f^{(j+1)}(x)(b-x)^j}{j!} + k(n+1)(b-x)^n \\ &= -\frac{f^{(n+1)}(x)(b-x)^n}{n!} + k(n+1)(b-x)^n \end{aligned}$$

(dove nel secondo passaggio abbiamo eseguito una traslazione di indici nella sommatoria).

Allora $g'(c) = 0$ significa:

$$-\frac{f^{(n+1)}(c)(b-c)^n}{n!} + k(n+1)(b-c)^n = 0$$

ossia

$$k = \frac{f^{(n+1)}(c)}{(n+1)!}$$

che, inserita nella (7.11), dà la tesi. ♦

Formula di Taylor e convessità

Consideriamo la formula (7.9) con $n = 1$:

$$f(x) = f(x_0) + f'(x_0)(x - x_0) + \frac{1}{2}f''(c)(x - x_0)^2$$

e supponiamo che in ogni punto di (a, b) si abbia $f''(x) \geq 0$, ovvero f sia convessa in (a, b). Allora si ha $\frac{1}{2}f''(c)(x - x_0)^2 \geq 0$ e perciò si può scrivere, per ogni coppia di punti $x_0, x \in (a, b)$:

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0) \tag{7.12}$$

La (7.12) significa geometricamente che il grafico di $f(x)$ si mantiene in tutto (a, b) sopra il grafico della sua retta tangente in x_0 (e questo vale per ogni scelta del punto $x_0 \in (a, b)$). Abbiamo dunque dimostrato che se una funzione (due volte derivabile) è convessa, è anche "convessa per tangenti" (vedi Teorema 4.12 par. 5.2).

Se invece f è concava, ossia $f''(x) \leq 0$ in tutto (a, b), la disuguaglianza (7.12) vale con "$\leq$" anziché $\geq$: il grafico di $f(x)$ si mantiene in tutto (a, b) sotto il grafico della sua retta tangente in x_0.

7.5 Risoluzione approssimata di equazioni: il metodo di Newton

Supponiamo di voler risolvere l'equazione

$$f(x) = 0$$

che equivale a cercare le intersezioni del grafico di f con l'asse x, e supponiamo di aver accertato (mediante confronto grafico, ad esempio) l'esistenza di un'unica soluzione $x = \alpha$, all'interno di un intervallo $[a, b]$. Si vuole dare una valutazione approssimata di α. Il metodo di Newton, che ora illustriamo, serve proprio a questo scopo. L'idea è quella di costruire una successione convergente ad α; questa successione, come si vedrà, è definita in modo ricorsivo, ossia:

- si assegna esplicitamente il primo termine x_0;
- si assegna esplicitamente la legge con cui calcolare x_{n+1} a partire da x_n, qualunque sia n.

In questo modo, il termine x_n può essere costruito a partire da x_0 con n iterazioni del medesimo algoritmo. Il metodo, perciò, si presta molto bene al calcolo automatico.

Illustriamo il metodo con riferimento alla figura 4.51. Partiamo da $x_0 = a$ e "linearizziamo" l'equazione $f(x) = 0$ sostituendo a f la retta tangente al suo grafico nel punto $(a, f(a))$. Tale retta ha equazione

$$y = f(x_0) + f'(x_0)(x - x_0)$$

Invece di risolvere $f(x) = 0$ risolviamo

$$f(x_0) + f'(x_0)(x - x_0) = 0$$

ottenendo

$$\boxed{x_1 = x_0 - \frac{f(x_0)}{f'(x_0)}} \rightarrow \text{ prima approssimazione di } \alpha$$

Figura 4.51.

Procediamo ora con x_1 al posto di $x_0 = a$, sostituendo f con la retta tangente in $(x_1, f(x_1))$. Anziché $f(x) = 0$ si risolve l'equazione

$$f(x_1) + f'(x_1)(x - x_1) = 0$$

ottenendo

$$\boxed{x_2 = x_1 - \frac{f(x_1)}{f'(x_1)}} \longrightarrow \text{seconda approssimazione di } \alpha$$

Proseguendo con x_2 come per x_1 e continuando si perviene alla legge di ricorrenza

$$\begin{cases} x_0 = a \\ x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)} \end{cases} \tag{7.13}$$

Nella situazione descritta in figura 4.51, si vede che $x_n \uparrow \alpha$.

Si tratta ora di dimostrare rigorosamente che, sotto opportune ipotesi, la successione x_n effettivamente converge alla soluzione cercata. Vale in proposito il seguente

Teorema 4.20 *Sia $f : [a, b] \to \mathbb{R}$ due volte derivabile in $[a, b]$, e supponiamo che:*

1. $f(a) \cdot f(b) < 0$;
2. $f'(x), f''(x)$ *hanno segno costante in* $[a, b]$.
3. $f(a) \cdot f''(a) > 0$.

Definiamo per ricorrenza la successione

$$\begin{cases} x_0 = a \\ x_{n+1} = x_n - \dfrac{f(x_n)}{f'(x_n)} \end{cases}$$

Allora esiste uno e un sol punto $c \in (a, b)$ tale che $f(c) = 0$, e la successione x_n converge ad c per difetto.

Se invece di valere la 3 vale la

3'. $f(b) \cdot f''(b) > 0$, *allora la successione*

$$\begin{cases} x_0 = b \\ x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \end{cases}$$

converge ad c per eccesso.

Osservazione Notiamo che, sotto le ipotesi 1 e 2, è senz'altro verificata la 3 o la 3'. Infatti f'' ha lo stesso segno in a e b per la 2, mentre f ha segni opposti in a e b, per la 1, perciò o vale la 3 o vale la 3'. Questo significa che il metodo è applicabile ogni volta che le ipotesi 1 e 2 sono verificate: si tratta solo, a quel punto, di scegliere opportunamente se porre $x_0 = a$ oppure b. A sua volta, la condizione 2, apparentemente restrittiva, sarà in generale soddisfatta pur di scegliere un intervallo $[a, b]$ abbastanza piccolo.

Dimostrazione. Poiché f è continua in $[a, b]$ e vale la 1, per il teorema degli zeri esiste almeno un $c \in (a, b)$ tale che $f(c) = 0$. Inoltre, poiché $f'(x)$ ha segno costante in (a, b), la f è strettamente monotona, quindi tale punto c è unico (una funzione strettamente monotona non può annullarsi in due punti distinti). Questo prova esistenza e unicità del punto c in cui f si annulla. Proveremo che la successione x_n è *monotona*. Da questo seguirà (teorema di monotonia) che la successione è convergente. Mostriamo per intanto che, se x_n converge, certamente converge ad c.

Infatti, se $x_n \to \beta \in [a, b]$, passando al limite nell'uguaglianza

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)}$$

si ha

$$\beta = \beta - \frac{f(\beta)}{f'(\beta)}$$

(abbiamo sfruttato il fatto che f è continua, e anche f' è continua, in quanto f' è derivabile, perché esiste per ipotesi f''). Dall'ultima uguaglianza segue $f(\beta) = 0$, per cui $\beta = c$ (unico punto in cui la funzione si annulla).

Proviamo dunque che x_n è monotona, sotto le ipotesi (1), (2), (3). Per fissare le idee, supponiamo che sia $f(a) < 0$. Allora per (1) è $f(b) > 0$. Poiché per la (2) $f'(x)$ ha segno costante in $[a, b]$, dovrà essere $f'(x) > 0$ in tutto $[a, b]$ (se valesse l'altra disuguaglianza, f sarebbe decrescente, e non potrebbe essere $f(a) < 0 < f(b)$). Poiché $f(a) < 0$ e vale la (3), $f''(a) < 0$, dunque, poiché per la (2) $f''(x)$ ha segno costante, $f''(x) < 0$ in tutto $[a, b]$. Siamo quindi nella situazione della seguente figura:

Figura 4.52.

Consideriamo ora la funzione

$$g(x) = x - \frac{f(x)}{f'(x)}.$$

Notiamo che è:

$$x_{n+1} = g(x_n);$$
$$g(c) = c$$

(perché $f(c) = 0$); inoltre

$$g'(x) = 1 - \frac{f'(x)^2 - f(x) f''(x)}{f'(x)^2} = \frac{f(x) f''(x)}{f'(x)^2} > 0 \Longleftrightarrow f(x) < 0$$

(perché $f''(x) < 0$ in tutto $[a, b]$), quindi $g'(x) > 0$ per $x \in [a, c]$ (vedi grafico di f), ossia: in $[a, c]$ la funzione g è strettamente crescente. Tenendo presenti queste cose, proviamo ora che:

$$\forall n \text{ è: } x_n < x_{n+1} < c. \tag{7.14}$$

Infatti: per $n = 0$ si ha:

$$x_1 = a - \frac{f(a)}{f'(a)} > a$$

(perché $f(a) < 0$ e $f'(a) > 0$), ossia $x_1 > x_0$. Inoltre $a < c$, e poiché g è crescente, questo implica

$$g(a) < g(c), \text{ ossia}$$
$$x_1 < c.$$

Quindi abbiamo provato che

$$x_0 < x_1 < c.$$

Applicando ora la g alle disuguaglianze precedenti (g è strettamente crescente in $[a, c]$, e i tre punti stanno in questo intervallo, quindi g conserva le disuguaglianze), si ha

$$g(x_0) < g(x_1) < g(c), \text{ ossia}$$
$$x_1 < x_2 < c.$$

Applicando ancora la g, troveremo $x_2 < x_3 < c$, e così via. Perciò la (7.14) è vera, e in particolare x_n è monotona. Questo conclude la dimostrazione. ♦

Esempio

7.8 Consideriamo ancora l'equazione

$$f(x) = e^{-x} - x = 0$$

Abbiamo:

$$f'(x) = -e^{-x} - 1 < 0 \quad \text{in} \quad [0, 1]$$

$$f''(x) = e^{-x} > 0 \quad \text{in} \quad [0, 1]$$

Inoltre $f(0) = 1 > 0$ e $f(1) < 0$.

La successione (7.13), è in questo caso

$$x_0 = 0\,, \quad x_{n+1} = x_n + \frac{e^{-x_n} - x_n}{e^{-x_n} + 1} = \frac{x_n + 1}{1 + e^{x_n}}\,.$$

Usando una calcolatrice si trova

$$x_1 = 0{,}500$$

$$x_2 = 0{,}5663\ldots$$

$$x_3 = 0{,}5671\ldots$$

$$x_4 = 0{,}5671\ldots$$

Si vede che già dopo 4 iterazioni si può dire $c = 0{,}5671\ldots$ con $e^{-x_4} - x_4 \approx 4{,}55 \cdot 10^{-7}$.

Esercizi

Approssimazioni al primo e al secondo ordine.

97 La velocità $v(t)$ di un corpo di massa m che cade nell'aria è regolata dalla legge

$$v(t) = \frac{mg}{h}\left(1 - e^{-\frac{h}{m}t}\right)$$

dove g è l'accelerazione di gravità e h un coefficiente che esprime l'attrito dell'aria. (Questo è vero se si assume la forza di attrito proporzionale alla velocità). Linearizzando l'espressione, si trovi come varia la velocità nei primi istanti di caduta. (Cioè si approssimi $v(t)$ per $t \to 0$). L'approssimazione così ottenuta è per eccesso o per difetto? Per rispondere, si scriva ora l'approssimazione al second'ordine (mediante la formula di Taylor), e si confrontino le due espressioni.

98 (Cfr. l'esercizio precedente). Se si assume invece la forza di attrito proporzionale al quadrato della velocità, la velocità del corpo in caduta risulta:

$$v(t) = \sqrt{\frac{mg}{h}} \cdot \frac{e^{2\sqrt{\frac{gh}{m}}t} - 1}{e^{2\sqrt{\frac{gh}{m}}t} + 1}$$

con lo stesso significato dei simboli (il coefficiente h ha ora dimensioni diverse). Linearizzando l'espressione, si trovi come varia la velocità nei primi istanti di caduta.

99 L'oscillazione di una pallina fissata all'estremità di una molla e libera di muoversi su una retta è descritta dalla funzione:

$$x = Ae^{-kt}\cos(\omega t)$$

con A, k, ω costanti positive. Linearizzando l'espressione, si approssimi $x(t)$ per $t \to 0$. L'approssimazione così ottenuta è per eccesso o per difetto? Per rispondere, si scriva ora l'approssimazione al second'ordine, e si confrontino le due espressioni.

100 Usando la formula di Taylor al second'ordine col resto secondo Lagrange, si dia una maggiorazione del massimo errore che si commette approssimando $\sin x$ con x, quando $0 \le x \le \frac{\pi}{6}$, e si precisi se l'errore commesso è per eccesso o per difetto.

Il massimo errore commesso è in realtà per $x = \frac{\pi}{6}$: si calcoli questo errore. Il risultato è coerente con la stima effettuata?

101 Usando la formula di Taylor al second'ordine col resto secondo Lagrange, si dia una maggiorazione del massimo errore che si commette approssimando $\sqrt{1+x}$ con $1+\frac{1}{2}x$, quando $0 \leq x \leq \frac{1}{2}$, e si precisi se l'errore commesso è per eccesso o per difetto.

Il massimo errore commesso è in realtà per $x = \frac{1}{2}$: si calcoli questo errore. Il risultato è coerente con la stima effettuata?

102 Dimostrare, sfruttando la definizione dei simboli di "o piccolo" e di "asintotico", tutte le proprietà enunciate nel paragrafo "Proprietà del simbolo di o piccolo".

Metodo di Newton

103 Nell'esempio 4.5 (diffrazione della luce attraverso una fenditura) abbiamo visto che la ricerca dei massimi dell'intensità luminosa nella figura di interferenza per la luce che attraversa una fenditura porta a risolvere l'equazione $\operatorname{tg} t = t$. Calcolare col metodo di Newton un valore approssimato della prima soluzione positiva di tale equazione, che si trova (come mostra il grafico) poco prima di $\frac{3}{2}\pi$.
(Suggerimento: applicare il metodo alla funzione $f(t) = \operatorname{tg} t - t$ sull'intervallo $[\frac{5}{4}\pi, \frac{3}{2}\pi - 0.1]$, usando come valore iniziale $a = \frac{3}{2}\pi - 0.1$. Calcolare le iterate con l'aiuto di un computer).

104 Calcolare con approssimazione l'unica soluzione reale dell'equazione $x^2 + 1 = \frac{1}{x}$. A tal fine:

a. Mostrare che il metodo di Newton è applicabile sull'intervallo $[\frac{1}{2}, 1]$;.

b. Scrivere esplicitamente l'algoritmo iterativo.

c. Calcolare esplicitamente le prime iterazioni, finché le prime due cifre decimali si stabilizzano.

105 Calcolare con approssimazione l'unica soluzione reale dell'equazione $\log x = \frac{1}{x}$. A tal fine:

a. Mostrare che il metodo di Newton è applicabile sull'intervallo $[1, 2]$.

b. Scrivere esplicitamente l'algoritmo iterativo.

c. Calcolare esplicitamente le prime iterazioni, finché le prime due cifre decimali si stabilizzano.

Convessità

106 Siano f, g due funzioni definite su tutto $\mathbb{R}$, derivabili due volte. Supponiamo che f sia convessa e g concava, su tutto $\mathbb{R}$. Utilizzando le proprietà geometriche enunciate nel paragrafo 5.2 per le funzioni concave e convesse, dimostrare che f e g non possono avere più di due intersezioni.

Calcolare i seguenti limiti utilizzando opportunamente, se necessario, gli sviluppi di Taylor.

107 $\lim_{x \to +\infty} \frac{\log \sqrt{1+x^2}}{\sqrt[4]{x}}$

108 $\lim_{x \to +\infty} x^\alpha \left(\sqrt{x^2 + 2x + 3} - x - 1\right)$

109 $\lim_{x \to +\infty} \left(\sqrt{x^2 + x + 1} - x\right)$

110 Sviluppare in serie di Taylor in vicinanza di $x = 0$ all'ordine 5 le seguenti funzioni:

$$\frac{\sin x}{x} \qquad \log(1+\cos x) \qquad \frac{1}{\cos x} \qquad \frac{1}{\sqrt{3+x}}$$

111 Se per $x \to 0$ (o $x \to 0^{\pm}$ se il caso lo richiede) si verifica che $f(x) \sim kx^{\alpha}$ $(k \neq 0,\ \alpha > 0)$, si dice che α è l'ordine di infinitesimo di f rispetto all'infinitesimo campione x.

Utilizzando opportunamente gli sviluppi di Taylor, determinare l'ordine di infinitesimo per $x \to 0$ delle seguenti funzioni

$$i)\ \ \sin x - x \qquad ii)\ \ (\sin x)^2 - x^2 \qquad iii)\ \ \cos(2\sqrt{x}) - e^{-2x}$$

112 Gli sviluppi di Taylor e il teorema di de L'Hospital sono due diversi strumenti per il calcolo dei limiti; spesso entrambi i metodi possono essere utilizzati. Provare a rifare gli esercizi 57, 58, 60, 61, 62, 63, 64 utilizzando opportunamente gli sviluppi di Taylor.

COMPLEMENTI

1 Si considerino le funzioni iperboliche inverse SettShx, SettChx (v. cap. 2, par. 4.4). Calcolare la derivata di queste funzioni inverse:

a) applicando il teorema della derivata della funzione inversa, sfruttando perciò le formule per la derivata di Shx e Chx, e la relazione tra le due funzioni;

b) derivando direttamente l'espressione esplicita della funzione inversa, calcolata nel cap. 2, par. 4.4. (Naturalmente, si deve trovare lo stesso risultato).

2 Fare un esempio di funzione derivabile n volte ma non $n-1$ volte, per n intero positivo qualsiasi.

3 Si consideri la funzione

$$f_\alpha(x) = \begin{cases} x^\alpha \sin\dfrac{1}{x} & \text{per } x > 0 \\ 0 & \text{per } x \leq 0 \end{cases}$$

Stabilire per quali valori del parametro reale α, la funzione f_α: è limitata; è continua; è derivabile in 0; ha derivata prima continua in 0; è derivabile due volte.

4 Dimostrare il Teorema enunciato nel par. 7.2 sulla relazione tra "asintotico" e "o piccolo".

5 Dimostrare la seguente relazione tra i polinomi di MacLaurin di f e f' (relazione che abbiamo utilizzato nella dimostrazione della formula di Taylor con resto secondo Peano):

$$T'_{n,f}(x) = T_{n-1,f'}(x)\,.$$

6 Dimostrare che dato un qualsiasi polinomio di grado n,

$$p_n(x) = a_n x^n + a_{n-1}x^{n-1} + \ldots + a_0$$

e assegnato un punto $x_0 \in \mathbb{R}$, si può esprimere p_n anche nella forma di polinomio di grado n centrato in x_0:

$$p_n(x) = b_n (x - x_0)^n + b_{n-1} (x - x_0)^{n-1} + ... + b_0$$

per opportuni coefficienti b_k determinati dai coefficienti a_k. (Questo fatto è implicitamente utilizzato nel provare il Teorema che caratterizza il polinomio di Taylor).

7 (a) Sia f una funzione invertibile in (a, b), e g la sua funzione inversa. Nell'ipotesi che f sia derivabile 2 volte in $x_0 \in (a, b)$ con $f'(x_0) \neq 0$, dimostrare la formula:

$$g''(y_0) = -\frac{f''(x_0)}{f'(x_0)^3} \quad (\text{con } y_0 = f(x_0)).$$

(b) Sapendo che

$$f(x) = a + b(x - x_0) + c(x - x_0)^2 + o(x - x_0)^2,$$

scrivere lo sviluppo di Taylor al second'ordine per g centrato in y_0, in funzione dei coefficienti a, b, c.

(c) Supponendo che f sia 3 volte derivabile in x_0, calcolare anche $g'''(y_0)$.

8 La formula di Taylor fornisce un ulteriore criterio per lo studio della natura dei punti stazionari. Sia f una funzione derivabile tutte le volte che ci occorrerà in x_0, e sia $f'(x_0) = 0$. Dimostrare che:

se il minimo intero n per cui risulta $f^{(n)}(x_0) \neq 0$ è pari, allora x_0 è un punto di massimo o minimo relativo, e precisamente: è di massimo (minimo) se $f^{(n)}(x_0) < 0$ (> 0); se invece il minimo intero n per cui risulta $f^{(n)}(x_0) \neq 0$ è dispari, allora x_0 non è un estremante.

5 Serie

1 SERIE NUMERICHE

1.1 Definizione e primi esempi

Nel capitolo 3 abbiamo introdotto il concetto di *limite di successione*, come punto di partenza per poter definire quello di *limite di funzione*, che è il fondamento concettuale del calcolo infinitesimale nel *continuo*, ossia il calcolo differenziale, trattato nel cap.4, e integrale, che tratteremo nel cap.6. In questo capitolo vedremo come il concetto di limite di successione permetta un altro importante sviluppo del calcolo infinitesimale, questa volta nell'ambito del *discreto*. Si tratta del concetto di *serie*, che come vedremo consente di estendere l'elementare operazione algebrica di somma a un numero infinito di addendi.

La possibilità di sommare infiniti numeri, magari tutti positivi, e ottenere un risultato finito, era qualcosa di paradossale per gli antichi filosofi greci: Zenone di Elea (intorno al 400 a.C.) formulò il celebre paradosso di Achille e la tartaruga.

In una gara di corsa Achille (A) e la tartaruga (T) partono inizialmente (tempo t_0) dalle posizioni x_0 e $x_1 > x_0$ rispettivamente (fig. 5.1). In un tempo t_1, A raggiungerà la posizione x_1 ma, nello stesso tempo, T avrà raggiunto una posizione $x_2 > x_1$. Per

tempo	posizione di A	posizione di T
t_0	x_0	x_1
$t_0 + t_1$	x_1	x_2
$t_0 + t_1 + t_2$	x_2	x_3
$\vdots$	$\vdots$	$\vdots$

Figura 5.1.

arrivare a x_2, A impiegherà un tempo t_2, ma in questo tempo T si sarà spostata in x_3 e così via...

O si ammette che la somma di infinite quantità positive: $t_0+t_1+t_2+\ldots$ possa dare un risultato finito, o si deve concludere che Achille non riuscirà mai a raggiungere la tartaruga!

D'altra parte non è così paradossale che una somma di infiniti addendi positivi, una volta che sia stata ben definita, possa dare un risultato finito; convinciamocene con questo semplice esempio. Immaginiamo di misurare un'asta lunga 2 m usando questo bizzarro procedimento: dividiamo l'asta a metà e misuriamo il primo pezzo: otterremo 1; poi dividiamo a metà il secondo pezzo e misuriamone la prima parte: otterremo $\frac{1}{2}$; il rimanente pezzo sia diviso ancora a metà... e così via indefinitamente.

Otteniamo una somma infinita

$$1+\frac{1}{2}+\frac{1}{4}+\frac{1}{8}+\ldots+\frac{1}{2^n}+\ldots$$

la quale, secondo le aspettative, dovrebbe avere come risultato 2.

Vediamo ora di precisare il discorso fatto con opportune definizioni.

Data una successione di numeri reali $\{a_n\}$, chiamiamo *serie dei termini* a_n la *scrittura formale*

$$\sum_{n=0}^{\infty} a_n \tag{1.1}$$

che si legge "serie (ma anche "somma") per n da 0 a ∞ di a_n". Per dare significato a questo simbolo, che intuitivamente rappresenta l'operazione di somma degli infiniti addendi a_n, costruiamo anzitutto un'altra successione, $\{s_n\}$, i cui termini sono così definiti:

$$\begin{aligned} s_0 &= a_0 \\ s_1 &= a_0 + a_1 \\ s_2 &= a_0 + a_1 + a_2 \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ s_n &= a_0 + a_1 + \ldots + a_n \\ &\ldots\ldots\ldots\ldots\ldots\ldots \end{aligned} \tag{1.2}$$

Il numero s_n viene detto *somma parziale* (o ridotta) n-esima della serie (1.1), e la successione $\{s_n\}$ si dice *successione delle somme parziali* della serie (1.1). Notare che le (1.2) si possono condensare nell'unica scrittura sintetica[1]:

$$s_n = \sum_{k=0}^{n} a_k \quad \text{per } n = 0, 1, 2, \ldots \tag{1.3}$$

[1]Si presti attenzione, nelle formule precedenti, al diverso ruolo giocato dai due indici (k e n): l'indice rispetto a cui si somma (n nella (1.1); k nella (1.3)) è sempre un indice muto: ad esempio la (1.3) si potrebbe scrivere anche così:

$$s_n = \sum_{j=0}^{n} a_j \quad \text{per } n = 0, 1, 2, \ldots$$

Definizione 5.1 Diremo che la serie (1.1) è *convergente, divergente, irregolare*, se la successione $\{s_n\}$ delle sue somme parziali è *convergente, divergente* o *irregolare*, rispettivamente. In particolare, se $\{s_n\}$ è convergente, $s_n \to s$, diremo che s è la *somma della serie*, e scriveremo:

$$\sum_{n=0}^{\infty} a_n = s$$

In questo caso dunque vale la relazione:

$$\sum_{n=0}^{\infty} a_n = \lim_{n\to+\infty} \sum_{k=0}^{n} a_k = \lim_{n\to+\infty} s_n$$

L'ultima formula scritta è interessante perché spiega in che modo il concetto di serie traduca con precisione l'idea di "somma di infiniti addendi": si calcola il *limite*, per $n \to \infty$, della *somma finita* dei primi n addendi.

L'espressione "studiare il carattere della serie" significa stabilire se la serie è *convergente* o *divergente*, o *irregolare*.

Talvolta, invece di sommare a partire da 0, si parte da un indice $N > 0$; scriveremo allora $\sum_{n=N}^{\infty} a_n$. Per indicare una serie useremo talvolta anche il simbolo sintetico $\sum a_n$. In questo caso sarà chiaro dal contesto a partire da quale indice N occorra sommare.

Osservazione Parlare di una serie numerica coinvolge sempre *due* diverse successioni: la successione $\{a_n\}$ dei *termini della serie* (a_n si dice anche "termine generale della serie") e la successione $\{s_n\}$ delle sue *somme parziali*. Lo studente presti bene attenzione, volta per volta, a quale delle due successioni si riferiscono le affermazioni fatte. Per esempio, abbiamo visto che la serie $\sum a_n$ si dice convergente se la successione $\{s_n\}$ è convergente (non se la successione $\{a_n\}$ è convergente!).

Esempi

1.1 (*Serie geometrica*) Sia $a_n = q^n$, $q \in \mathbb{R}$. Se $q \neq 1$, utilizzando la formula (1.1) del capitolo 1, abbiamo

$$s_n = 1 + q + q^2 + \ldots + q^n = \frac{1 - q^{n+1}}{1 - q}.$$

Se invece $q = 1$ abbiamo $s_n = n + 1$. Prendendo il limite, per $n \to +\infty$, otteniamo:

$$\lim_{n\to\infty} s_n = \begin{cases} \dfrac{1}{1-q} & \text{se } |q| < 1 \\ +\infty & \text{se } q \geq 1 \\ \text{non esiste} & \text{se } q \leq -1 \end{cases}$$

$$\text{e pertanto la serie } \sum_{n=0}^{\infty} q^n \quad \text{è} \quad \begin{cases} \text{convergente } \left(\text{con somma } \dfrac{1}{(1-q)}\right) & \text{se } |q| < 1 \\ \text{divergente a } +\infty & \text{se } q \geq 1 \\ \text{irregolare} & \text{se } q \leq -1 \end{cases}$$

1.2 (*Serie armonica*) È la serie

$$\sum_{n=1}^{\infty} \frac{1}{n} = 1 + \frac{1}{2} + \frac{1}{3} + \frac{1}{4} + \ldots + \frac{1}{n} + \ldots$$

Dimostreremo (nel cap. 6 par. 8.3) che la successione s_n delle somme parziali è $\geq \log(n+1)$ (e perciò, in particolare, tende a $+\infty$).

1.3 (*Serie di Mengoli*) È la serie:

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

Osservando che $\frac{1}{n(n+1)} = \frac{1}{n} - \frac{1}{n+1}$, si riesce a dare un'espressione semplice alla successione s_n:

$$s_n = \sum_{k=1}^{n} \left[\frac{1}{k} - \frac{1}{k+1}\right] = \left[1 - \frac{1}{2}\right] + \left[\frac{1}{2} - \frac{1}{3}\right] + \ldots + \left[\frac{1}{n} - \frac{1}{n+1}\right] =$$

(grazie al fatto che i termini si semplificano 2 a 2)

$$= 1 - \frac{1}{n+1}$$

Dunque $s_n \to 1$, ossia la serie converge e ha somma 1.

La serie di Mengoli è il più semplice esempio di *serie telescopica*, che significa quanto segue. Il termine generale a_k ha la forma $(b_k - b_{k+1})$ (dove b_k è un'altra opportuna successione) e di conseguenza, grazie alle cancellazioni, si ha

$$s_n = b_1 - b_{n+1}$$

Se il termine $b_n \to 0$, la serie è convergente e ha somma b_1.

Supponiamo che la serie $\sum_{n=1}^{\infty} a_n$ sia convergente e che s sia la sua somma. Ciò significa che $s_n \to s$ e quindi

$$a_n = (s_n - s_{n-1}) \to (s - s) = 0$$

Ne segue che:

Teorema 5.1 *Condizione necessaria affinché una serie $\sum_{n=0}^{\infty} a_n$ converga è che il termine generale a_n tenda a zero.*

Come mostra l'esempio della *serie armonica*, la condizione non è sufficiente. Comunque, se il termine generale non tende a zero, certamente la serie non converge. Ad esempio $\sum_{n=1}^{\infty} \cos\frac{1}{n}$ non converge, perché $\cos\frac{1}{n} \to 1$.

Talvolta è utile anche il seguente fatto:

Teorema 5.2 *Supponiamo che*

$$\sum_{k=0}^{\infty} a_k$$

sia convergente. Allora per ogni n risulta convergente anche la serie (che si dice resto *della serie di partenza)*

$$R_n = \sum_{k=n}^{\infty} a_k; \text{ inoltre } R_n \to 0 \text{ per } n \to \infty.$$

Detto in parole: il resto di una serie convergente è infinitesimo.

DIMOSTRAZIONE. Sia $s_n = \sum_{k=0}^{n} a_k$ e sia

$$S_h^{R_n} = \sum_{k=n}^{h} a_k \text{ (per } h \geq n)$$

la somma parziale h-esima della serie R_n. Dall'identità, valida per ogni $h > n$,

$$S_h^{R_n} + s_{n-1} = s_h$$

si ha, passando al limite per $h \to \infty$,

$$R_n + s_{n-1} = s.$$

Per $n \to \infty$, $s_{n-1} \to s$ e quindi $R_n = s - s_{n-1} \to 0$. ◆

Nei prossimi paragrafi studieremo vari *criteri di convergenza*, che forniscono alcune condizioni sufficienti affinché una serie converga; il problema di calcolarne esplicitamente la somma, invece, esula per il momento dai nostri obiettivi.

1.2 Serie a termini non negativi

Per una serie a termini non negativi si hanno alcuni semplici criteri sufficienti di convergenza. Cominciamo con l'osservare che la successione delle somme parziali di una tale serie sarà crescente, poiché

$$s_{n+1} = s_n + a_{n+1} \geq s_n$$

Perciò, per il teorema sull'esistenza del limite di una *successione monotona* (v. cap. 3, par. 1.2), esiste $\lim_{n\to+\infty} s_n$: tale limite è finito oppure $+\infty$ a seconda che la successione $\{s_n\}$ sia limitata oppure no. In ogni caso risulta: $\lim_{n\to+\infty} s_n = \sup_{n\in\mathbb{N}}\{s_n\}$.

Possiamo perciò affermare che:

una serie $\sum_{n=0}^{\infty} a_n$ a termini non negativi è convergente o è divergente a $+\infty$. Essa converge se e solo se la successione delle somme parziali n-esime è limitata.

Criterio del confronto

Siano $\sum a_n$ e $\sum b_n$ due serie a termini non negativi e tali che

$$a_n \leq b_n \qquad \text{definitivamente}$$

Allora valgono le seguenti implicazioni

i) $\sum b_n$ convergente $\implies \sum a_n$ convergente

ii) $\sum a_n$ divergente $\implies \sum b_n$ divergente.

La serie $\sum b_n$ viene detta *maggiorante*, la $\sum a_n$ *minorante*.

Questo e i prossimi criteri saranno dimostrati alla fine del paragrafo.

Esempio

1.4 La serie $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$, con $\alpha \leq 1$ è divergente.

Infatti l'affermazione è vera per $\alpha = 1$ (serie armonica); per $\alpha < 1$ la serie data è maggiorante della serie armonica, perciò diverge.

Criterio del confronto asintotico

Se le due successioni (a termini positivi) $\{a_n\}$ e $\{b_n\}$ sono asintotiche,

$$a_n \sim b_n$$

allora le corrispondenti serie $\sum a_n$ e $\sum b_n$ hanno lo stesso carattere, cioè o sono entrambe convergenti o sono entrambe divergenti.

Esempi

1.5 La serie $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converge perché $\frac{1}{n^2} \sim \frac{1}{n(n+1)}$ e la serie $\sum_{n=1}^{\infty} \frac{1}{n(n+1)}$ (di Mengoli) converge.

1.6 La serie $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ con $\alpha \geq 2$ converge per il criterio del confronto: $\frac{1}{n^\alpha} \leq \frac{1}{n^2}$ e $\sum_{n=1}^{\infty} \frac{1}{n^2}$ converge.

Abbiamo quindi stabilito il carattere della *serie armonica generalizzata* $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ per $\alpha \geq 2$ (convergente) e per $\alpha \leq 1$ (divergente). Si può dimostrare (una prova di questo fatto sarà fornita nel cap. 6 par. 8.3) che nei casi intermedi ($1 < \alpha < 2$) c'è ancora convergenza. Ricapitolando:

$$\boxed{\sum_{n=1}^{\infty} \frac{1}{n^\alpha} \quad \text{converge se } \alpha > 1; \text{ diverge se } \alpha \leq 1}$$

1.7 La serie $\sum_{n=1}^{\infty} \frac{5n+\cos n}{3+2n^3}$ converge, perché $\frac{5n+\cos n}{3+2n^3} \sim \frac{5}{2} \cdot \frac{1}{n^2}$. (Ovviamente la costante moltiplicativa è ininfluente sul carattere della serie).

Un'osservazione importante riguarda la differenza tra *stabilire il carattere della serie* (ad esempio: converge oppure no) e *calcolare la somma della serie* (nel caso converga). In generale, quando affermiamo in base al criterio del confronto asintotico che la serie $\sum a_n$ converge *perché* la serie $\sum b_n$ converge, ciò non significa affatto che le due serie abbiano la stessa somma. Intuitivamente, ciò è dovuto al fatto che l'affermazione $a_n \sim b_n$ ci dice che gli addendi delle due somme sono "simili" quando n è grande;

ma il valore della somma della serie dipende da *tutti* gli addendi, anche i primi. Ad esempio, abbiamo visto che

$$\sum_{n=1}^{\infty} \frac{1}{n(n+1)} = 1, \quad \text{e} \quad \frac{1}{n(n+1)} \sim \frac{1}{n^2}$$

Tuttavia si può dimostrare (lo vedremo nel vol.2°, trattando le *serie di Fourier*) che

$$\sum_{n=1}^{\infty} \frac{1}{n^2} = \frac{\pi^2}{6}$$

Ciò vale a maggior ragione quando si applica il criterio del confronto per affermare la convergenza di una serie.

L'ultimo esempio fatto è interessante anche per un altro motivo. La serie $\sum_{n=1}^{\infty} \frac{1}{n^2}$, convergente, ha per somma un numero irrazionale, nonostante il fatto che ogni termine della successione delle somme parziali sia razionale. Ciò significa che se il nostro ambiente di lavoro fosse $\mathbb{Q}$, questa serie non sarebbe convergente.

I criteri di convergenza per le serie a termini positivi si basano tutti sulla proprietà dell'estremo superiore, di cui gode l'insieme dei numeri reali.

Facciamo ora un'osservazione importante sull'applicazione pratica del criterio del confronto asintotico. Tutti gli strumenti che abbiamo incontrato nei capitoli 3 e 4 per stabilire stime asintotiche per *funzioni* (limiti notevoli, sviluppi di MacLaurin...) si possono applicare anche per ottenere stime asintotiche per successioni, e quindi forniscono strumenti potenti per lo studio del carattere di una serie a termini positivi:

Esempi

1.8 Si studi il carattere delle seguenti serie

$$\text{a)} \sum_{n=1}^{\infty} \left(\frac{e^{1/n} - 1}{n} \right); \quad \text{b)} \sum_{n=1}^{\infty} \left(\frac{1}{n} - \sin \frac{1}{n} \right); \quad \text{c)} \sum_{n=1}^{\infty} \log \left(\frac{n+1}{n+2} \right)$$

a) $\frac{e^{1/n}-1}{n} \sim \frac{1}{n^2}$ (sfruttando il limite notevole $e^{\varepsilon(x)} - 1 \sim \varepsilon(x)$ per $\varepsilon(x) \to 0$), perciò la serie, a termini positivi, per confronto asintotico con la serie di $1/n^2$, converge.

b) Dallo sviluppo di MacLaurin della funzione $\sin x$, si ha (ponendo $x = \frac{1}{n}$):

$$\frac{1}{n} - \sin \frac{1}{n} = \frac{1}{n} - \left(\frac{1}{n} - \frac{1}{6n^3} + o\left(\frac{1}{n^3} \right) \right) \sim \frac{1}{6n^3}.$$

Dunque si tratta di una serie a termini positivi, e per confronto asintotico con la serie di $1/n^3$, questa serie converge.

c) $\log\left(\frac{n+1}{n+2}\right) \sim \frac{n+1}{n+2} - 1 = \frac{-1}{n+2} \sim -\frac{1}{n}$. (Si è sfruttato il limite notevole $\log(1 + \varepsilon(x)) \sim \varepsilon(x)$ per $\varepsilon(x) \to 0$). Dunque si tratta una serie a termini negativi, il cui termine generale è asintotico a $-1/n$. Per confronto con la serie armonica, questa serie diverge (a $-\infty$).

Criterio della radice
Sia $\sum a_n$ una serie a termini non negativi. Se esiste il limite

$$\lim_{n\to+\infty} \sqrt[n]{a_n} = l$$

e $l > 1$ la serie diverge, se $l < 1$ la serie converge; se $l = 1$ nulla si può concludere.

Esempi

1.9 Sia data $\sum_{n=1}^{\infty} \frac{a^n}{n^n}$ con $a \geq 0$, risulta, per $n \to +\infty$

$$\sqrt[n]{\frac{a^n}{n^n}} = \frac{a}{n} \to 0$$

e perciò la serie data converge.

1.10 Sia data $\sum_{n=1}^{\infty} n^\alpha a^n$, con $\alpha \in \mathbb{R}$, $a > 0$; risulta

$$\sqrt[n]{n^\alpha a^n} = a n^{\alpha/n}$$

Calcoliamo $\lim_{n\to+\infty} n^{\alpha/n}$; basterà calcolare (cfr. il confronto di infiniti, cap.3, par.1.5)

$$\lim_{n\to+\infty} \log\left(n^{\alpha/n}\right) = \lim_{n\to+\infty} \frac{\alpha}{n} \log n = 0 \quad \forall \alpha \in \mathbb{R}$$

e perciò $\lim_{n\to+\infty} n^{\alpha/n} = 1$. Dunque, se $a < 1$ la serie converge. Se $a > 1$ la serie diverge. Se $a = 1$ la serie diventa $\sum_{n=1}^{\infty} n^\alpha$ e già sappiamo che per $\alpha < -1$ questa serie è convergente, per $\alpha \geq -1$ divergente.

Criterio del rapporto
Sia $\sum a_n$ una serie a termini positivi. Se esiste il limite

$$\lim_{n\to+\infty} \frac{a_{n+1}}{a_n} = l$$

e $l > 1$ la serie diverge, se $l < 1$ la serie converge; se $l = 1$ nulla si può concludere.

Esempio

1.11 La serie $\sum_{n=0}^{\infty} \frac{1}{n!}$ è convergente. Infatti:

$$\lim_{n\to+\infty} \frac{1/(n+1)!}{1/n!} = \lim_{n\to+\infty} \frac{1}{n+1} = 0$$

Più avanti mostreremo che la somma di questa serie è e.

Dimostrazione dei criteri di convergenza per le serie a termini positivi

Proviamo ora i criteri di convergenza che abbiamo enunciato. Come si vedrà, questi sono una conseguenza della regolarità delle serie a termini positivi, che a sua volta dipende dal teorema di esistenza del limite per successioni monotone (v. cap. 3, par. 1.2).

DIMOSTRAZIONE DEL CRITERIO DEL CONFRONTO

Con le stesse notazioni impiegate nell'enunciato, sia:

$$s_n = \sum_{k=1}^{n} a_k; \qquad s_n^* = \sum_{k=1}^{n} b_k$$

Poiché $0 \leq a_k \leq b_k$ per ogni k, sommando membro a membro le disuguaglianze per k da 1 a n, si ottiene $0 \leq s_n \leq s_n^*$. Sappiamo già che una serie a termini positivi è regolare, ossia o converge o diverge (non può oscillare). Dunque le affermazioni *i*) e *ii*) sono logicamente equivalenti, perciò basta dimostrare la seconda, che è immediata perché: dire che $\sum_{n=1}^{\infty} a_n$ diverge significa, per definizione di serie divergente, che $s_n \to +\infty$.

D'altro canto è $s_n \leq s_n^*$, perciò, per il criterio del confronto per le successioni, anche $s_n^* \to +\infty$, ossia $\sum_{n=1}^{\infty} b_n$ diverge. ♦

DIMOSTRAZIONE DEL CRITERIO DEL CONFRONTO ASINTOTICO

Dire che $a_n \sim b_n$ per $n \to \infty$ significa che $\frac{a_n}{b_n} \to 1$ per $n \to \infty$; questo implica che per ogni $\varepsilon > 0$ sia definitivamente $1 - \varepsilon < \frac{a_n}{b_n} < 1 + \varepsilon$, ad esempio (scegliendo $\varepsilon = \frac{1}{2}$) che sia

$$\frac{1}{2} < \frac{a_n}{b_n} < \frac{3}{2}$$

ossia (poiché per ipotesi $b_n > 0$)

$$\frac{1}{2} b_n < a_n < \frac{3}{2} b_n \quad \text{definitivamente.}$$

Per il teorema del confronto, la prima delle due disuguaglianze implica che se $\sum a_n$ converge, anche $\sum b_n$ converge, mentre la seconda implica che se $\sum a_n$ diverge, anche $\sum b_n$ diverge. Questo significa appunto che le serie hanno lo stesso carattere. ♦

DIMOSTRAZIONE DEL CRITERIO DELLA RADICE

Supponiamo prima che sia:

$$\lim_{n \to +\infty} \sqrt[n]{a_n} = l < 1$$

Poiché $\sqrt[n]{a_n} \to l$, allora fissato comunque un $\varepsilon > 0$, definitivamente $\sqrt[n]{a_n} \leq l + \frac{\varepsilon}{2}$.

D'altro canto $l < 1$, perciò è anche $l < 1 - \varepsilon$ per un $\varepsilon > 0$ opportuno.

Per questo ε si ha dunque che, definitivamente

$$\sqrt[n]{a_n} \leq l + \frac{\varepsilon}{2} < (1 - \varepsilon) + \frac{\varepsilon}{2} = 1 - \frac{\varepsilon}{2}$$

 © 978-88-08-06485-1

e quindi

$$a_n < \left(1 - \frac{\varepsilon}{2}\right)^n.$$

Per confronto con la serie geometrica convergente $\sum_{n=1}^{\infty} \left(1 - \frac{\varepsilon}{2}\right)^n$, la serie di partenza converge.
Se ora, invece, è

$$\lim_{n \to +\infty} \sqrt[n]{a_n} = l > 1$$

con un ragionamento simile si deduce che

$$a_n > \left(1 + \frac{\varepsilon}{2}\right)^n$$

definitivamente, per un certo $\varepsilon > 0$. Dunque $a_n \to +\infty$, e la serie diverge. ♦

Dimostrazione del criterio del rapporto

È simile alla precedente. Supponiamo prima che sia:

$$\lim_{n \to +\infty} \frac{a_{n+1}}{a_n} = l < 1$$

Ragionando come nella dimostrazione precedente, si ha che:

$$\frac{a_{n+1}}{a_n} < \left(1 - \frac{\varepsilon}{2}\right)$$

definitivamente, per qualche $\varepsilon > 0$. Ciò implica, ragionando iterativamente, che:

$$a_{n+1} < \left(1 - \frac{\varepsilon}{2}\right) a_n < \left(1 - \frac{\varepsilon}{2}\right) \cdot \left(1 - \frac{\varepsilon}{2}\right) a_{n-1} < \ldots < \left(1 - \frac{\varepsilon}{2}\right)^n a_1$$

Per confronto con la serie geometrica convergente $\sum_{n=1}^{\infty} \left(1 - \frac{\varepsilon}{2}\right)^n a_1$, la serie di partenza converge.

Se ora, invece, è

$$\lim_{n \to +\infty} \frac{a_{n+1}}{a_n} = l > 1$$

con un ragionamento simile si deduce che

$$a_{n+1} > \left(1 + \frac{\varepsilon}{2}\right)^n a_1$$

definitivamente, per un certo $\varepsilon > 0$. Dunque $a_n \to +\infty$, e la serie diverge. ♦

Vale la pena osservare il seguente fatto, che talvolta ci sarà utile: come si vede dalle dimostrazioni appena svolte, sia nel criterio della radice che in quello del rapporto, se $l > 1$ *non solo la serie non converge ma, addirittura, il termine generale della serie tende a* $+\infty$.

Osservazione Abbiamo già osservato che il carattere di una serie non cambia se noi ne alteriamo un numero finito di termini. Di conseguenza, i criteri visti per le serie a termini non negativi, si applicano anche alle serie con termini *definitivamente* non negativi. Raccogliendo un segno meno dall'intera serie, poi, si vede che questi si

possono applicare anche alle serie a termini non positivi, e quindi alle serie a termini *definitivamente non positivi*. In sintesi, quindi, questi criteri si applicano alle serie che hanno tutti i termini (tranne un numero finito) dello stesso segno. Nel prossimo paragrafo esamineremo invece le serie che possiedono infiniti termini positivi e infiniti termini negativi.

1.3 Serie a termini di segno variabile

Veniamo ora allo studio delle serie a termini di segno qualsiasi. Diamo anzitutto una definizione:

Definizione 5.2 Una serie $\sum a_n$ si dirà assolutamente convergente se converge la serie (a termini non negativi) $\sum |a_n|$.

Vale il seguente teorema:

Teorema 5.3 *Se la serie $\sum a_k$ converge assolutamente, allora converge.*

Dimostrazione. Lo proveremo come conseguenza del teorema di monotonia, decomponendo il termine generale della serie nella sua parte positive e negativa. Per ogni k, definiamo:

$$\begin{aligned} s_n^+ &= \text{somma dei soli termini } a_k \geq 0, \text{ per } k = 0, 1, ..., n \\ s_n^- &= -\,(\text{somma dei soli termini } a_k < 0, \text{ per } k = 0, 1, ..., n) \\ s_n &= \sum_{k=0}^{n} a_k. \end{aligned}$$

Poiché $s_n = s_n^+ - s_n^-$, sarà sufficiente mostrare che le successioni s_n^+ e s_n^- sono convergenti per concludere che s_n converge, e quindi che la serie $\sum_{k=0}^{\infty} a_k$ converge.

Osserviamo che s_n^+ e s_n^- sono successioni monotone non decrescenti (aumentando n di un'unità si somma ad s_n^+, s_n^- una quantità positiva, oppure non si somma nulla). Inoltre, per la disuguaglianza triangolare:

$$s_n^+ \leq \sum_{k=0}^{n} |a_k|\,; \qquad s_n^- \leq \sum_{k=0}^{n} |a_k|\,.$$

D'altro canto, per ipotesi la serie $\sum a_k$ converge assolutamente, ossia $\sum_{k=0}^{\infty} |a_k|$ converge, e quindi la quantità $\sum_{k=0}^{n} |a_k|$ è limitata; perciò le successioni s_n^+, s_n^- sono superiormente limitate e non decrescenti, e pertanto convergono, per il teorema di monotonia. Il teorema è così dimostrato.

Dunque la convergenza assoluta implica la convergenza (ordinaria), detta anche *convergenza semplice*; il viceversa non è vero.

Esempio

1.12 La serie $\sum_{n=1}^{\infty} (-1)^n/n^\alpha$, con $\alpha > 1$ è convergente, anzi assolutamente convergente. Infatti si ha

$$\left|\frac{(-1)^n}{n^\alpha}\right| = \frac{1}{n^\alpha}$$

e la serie $\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$ è convergente per $\alpha > 1$. Se $\alpha = 1$ nulla, per il momento, possiamo concludere circa la convergenza della serie data, poiché la serie dei valori assoluti corrispondente è la serie armonica, che è divergente.

Mostreremo ora che $\sum_{n=1}^{\infty} (-1)^n/n$ è convergente, fornendo così un esempio di serie convergente, ma non assolutamente convergente.

Serie a termini di segno alternato. Criterio di Leibniz

Tra le serie a termini di segno variabile, un caso particolarmente semplice è costituito dalle serie *a segni alterni*, per le quali vale il seguente criterio di convergenza:

TEOREMA 5.4 (CRITERIO DI LEIBNIZ) *Sia data la serie*

$$\sum_{n=0}^{\infty} (-1)^n a_n \text{ con } a_n \geq 0 \ \forall n.$$

Se:

(i) *la successione* $\{a_n\}$ *è decrescente;*

(ii) $a_n \to 0$ *per* $n \to \infty$

allora la serie è convergente. Inoltre, le somme parziali di indice pari approssimano la somma per eccesso, quelle di indice dispari per difetto; il resto della serie è maggiorato, in valore assoluto, dal primo termine trascurato. In formule, detta s la somma della serie, si ha:

(i)
$$s_{2n} = \sum_{k=0}^{2n} (-1)^k a_k \downarrow s;$$
$$s_{2n+1} = \sum_{k=0}^{2n+1} (-1)^k a_k \uparrow s;$$

(ii)
$$|R_n| = \left| \sum_{k=n}^{\infty} (-1)^k a_k \right| \leq a_n.$$

Osserviamo che, analogamente a quanto osservato riguardo ai criteri di convergenza per le serie a termini positivi, il criterio di Leibniz può essere applicato anche se i termini sono *definitivamente* di segno alterno e la successione a_n è definitivamente decrescente.

Si noti che la successione delle somme parziali di una serie a segni alterni ha la forma:

$$\begin{aligned}
s_1 &= -a_1 \\
s_2 &= -a_1 + a_2 \\
s_3 &= -a_1 + a_2 - a_3 \\
&\dots\dots\dots\dots\dots\dots \\
s_n &= -a_1 + a_2 - a_3 + \ldots + (-1)^n a_n
\end{aligned}$$

dove gli a_i sono tutti ≥ 0. Se allora la successione a_n tende a zero monotonamente, l'andamento di a_n e di s_n è del tipo descritto dalla figura seguente, che sostanzialmente contiene la dimostrazione del criterio di Leibniz:

Figura 5.2.

DIMOSTRAZIONE. Dimostriamo formalmente il teorema. Consideriamo la successione delle somme parziali, s_n, e le due successioni, estratte da questa, s_{2n} e s_{2n+1}. Consideriamo dapprima la successione s_{2n}, così formata:

$$
\begin{aligned}
&s_0 = a_0 \\
&s_2 = s_0 - a_1 + a_2 \leq s_0 \text{ perché } a_2 \leq a_1 \\
&s_4 = s_2 - a_3 + a_4 \leq s_2 \text{ perché } a_4 \leq a_3 \\
&\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots
\end{aligned}
$$

Come si vede, la successione $\{s_{2n}\}$ è monotona decrescente. In modo analogo si dimostra che la successione $\{s_{2n+1}\}$ è monotona crescente. Inoltre,

$$s_1 \leq s_{2n+1} = s_{2n} - a_{2n+1} \leq s_{2n} \leq s_0$$

perciò $\{s_{2n+1}\}$ è superiormente limitata e $\{s_{2n}\}$ è inferiormente limitata. Le due successioni sono quindi convergenti, per il teorema di monotonia; possiamo anche dire che convergono allo stesso limite, perché

$$0 \leq s_{2n} - s_{2n+1} \leq a_{2n+1} \to 0.$$

Detto s questo limite comune, abbiamo allora provato le (i). Inoltre, dal fatto che $s_{2n+1} \to s$ e $s_{2n} \to s$ si deduce che $s_n \to s$.

Proviamo ora (ii). Poiché $s_{2n+1} \uparrow s$ e $s_{2n} \downarrow s$, possiamo scrivere

$$
\begin{aligned}
s_{2n-1} &\leq s \leq s_{2n} \\
s_{2n+1} &\leq s \leq s_{2n}
\end{aligned}
$$

da cui si deduce

$$
\begin{aligned}
0 &\leq s - s_{2n-1} \leq s_{2n} - s_{2n-1} = a_{2n} \\
0 &\leq s_{2n} - s \leq s_{2n} - s_{2n+1} = a_{2n+1}.
\end{aligned}
$$

Perciò (per n sia pari che dispari) si ha

$$|R_n| = |s - s_{n-1}| \leq a_n$$

che è la (ii). ◆

Esempi

1.13 Applicando il criterio di Leibniz osserviamo che le serie

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

sono entrambe convergenti (la prima è assolutamente convergente, la seconda no). La prima serie è rapidamente convergente. Sommando infatti i primi 6 termini della serie

$$1 - 1 + \frac{1}{2} - \frac{1}{6} + \frac{1}{24} - \frac{1}{120} = \frac{11}{30} = 0{,}3\overline{6}$$

si ottiene un valore per difetto della somma con un errore che non supera $1/6! = 1/720$. Mostreremo nel par. 2.1 che tale somma è $\frac{1}{e}$.

La seconda serie è lentamente convergente; per avere un valore della somma approssimato (per eccesso) a meno di $1/100$ bisogna sommare 99 termini! Come vedremo (par. 2.1), la somma di questa serie è $\log 2$; ma la serie non è adatta al calcolo approssimato di questo numero.

1.14

$$\sum_{n=1}^{\infty} (-1)^n \frac{n-1}{n^2+n}$$

È una serie a segni alterni, $a_n = \frac{n-1}{n(n+1)}$ è positivo e tende a zero. Poiché $a_n \sim \frac{1}{n}$, la serie non converge assolutamente; vediamo dunque se è applicabile il criterio di Leibniz. Dobbiamo verificare se è decrescente, ovvero:

$$a_{n+1} \leq a_n$$

$$\frac{n}{(n+1)(n+2)} \leq \frac{n-1}{n(n+1)}$$

ossia

$$n^2 \leq n^2 + n - 2$$

vero per $n \geq 2$. Dunque si può applicare il criterio di Leibniz e concludere che la serie converge.

1.15

$$\sum_{n=2}^{\infty} (-1)^n \frac{\log n}{n}$$

Si tratta di una serie a segni alterni, poiché $a_n = \frac{\log n}{n}$ è positiva; inoltre $a_n \to 0$ (gerarchia degli infiniti). La serie non converge assolutamente, perché $\frac{\log n}{n} > \frac{1}{n}$ (e $\sum \frac{1}{n}$ diverge); vediamo se è applicabile il criterio di Leibniz. In questo caso non è facile provare la monotonia della successione per via puramente algebrica, cioè studiando la disequazione

$$a_{n+1} \leq a_n$$

$$\frac{\log(n+1)}{n+1} \leq \frac{\log n}{n}.$$

E' più comodo invece studiare la monotonia della successione a_n mediante il calcolo differenziale (si veda quanto osservato a riguardo nel cap. 4, par. 4.2): posto

$$f(x) = \frac{\log x}{x} \qquad \text{per } x > 0$$

(con x variabile reale), calcoliamo:

$$f'(x) = \frac{1 - \log x}{x^2} \leq 0 \quad \text{per } x \geq e$$

Ne segue che f è decrescente per $x \geq e$; di conseguenza la successione $a_n = f(n)$ è decrescente almeno per $n \geq 3$ (il primo intero $> e$). Pertanto, per il criterio di Leibnitz la serie converge.

1.16
$$\sum_{n=1}^{\infty} \frac{n + 1 + (-1)^n n^2}{n^3}$$

Scriviamo:

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

La prima serie converge perché è a termini positivi e $a_n \sim \frac{1}{n^2}$; la seconda converge per il criterio di Leibniz; quindi la serie di partenza converge. Abbiamo qui usato un fatto generale:

$$\text{se } \sum_{k=1}^{\infty} a_k \text{ converge e } \sum_{k=1}^{\infty} b_k \text{ converge, allora } \sum_{k=1}^{\infty} (a_k + b_k) \text{ converge,}$$

come si verifica vedendo la serie come limite della successione delle somme parziali, e applicando il teorema sul limite della somma.

1.17
$$\sum_{n=1}^{\infty} (-1)^n \left[\frac{\sqrt{n} + (-1)^n}{n} \right]$$

Scriviamo:

$$\sum_{n=1}^{\infty} (-1)^n \left[\frac{\sqrt{n} + (-1)^n}{n} \right] = \sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}} + \sum_{n=1}^{\infty} \frac{1}{n}$$

La prima serie converge per il criterio di Leibniz; la seconda diverge (serie armonica); quindi la serie di partenza diverge. Più in generale:

$$\text{se } \sum_{k=1}^{\infty} a_k \text{ converge e } \sum_{k=1}^{\infty} b_k \text{ diverge, allora } \sum_{k=1}^{\infty} (a_k + b_k) \text{ diverge,}$$

per lo stesso argomento usato nell'esempio precedente.

Questo esempio è interessante anche per il seguente motivo. Si noti che:

$$a_n = \frac{\sqrt{n} + (-1)^n}{n} \geq 0 \text{ e } a_n \to 0$$

Dunque la serie di partenza è una serie a segni alterni, e il suo termine generale tende a zero. Se affermassimo che essa converge per il criterio di Leibniz, diremmo il falso (abbiamo visto che la serie diverge). Il punto è che non abbiamo verificato l'altra ipotesi del criterio di Leibniz, ossia che $\frac{\sqrt{n}+(-1)^n}{n}$ sia *monotona decrescente*. Evidentemente, quest'ipotesi non è verificata. (Per esercizio lo studente tabuli i primi 10 valori della successione, e osservi cosa accade).

Esercizi

Dare una stima asintotica del termine generale delle seguenti serie a termini positivi. Dedurne il carattere della serie (converge o diverge?).

1. $\displaystyle\sum_{n=1}^{\infty} \frac{n+3}{2n^3+2n+2}$

2. $\displaystyle\sum_{n=0}^{\infty} \frac{\sqrt[3]{n}}{\sqrt{n^2+n+1}}$

Stabilire se le seguenti serie a termini positivi convergono o divergono, precisando il criterio utilizzato.

3. $\displaystyle\sum_{n=1}^{\infty} \frac{n!}{n^n}$

4. $\displaystyle\sum_{n=1}^{\infty} \frac{n^{1/n}}{n!}$

5. $\displaystyle\sum_{n=0}^{\infty} \frac{2^n}{e^{2n}}$

6. $\displaystyle\sum_{n=1}^{\infty} \log\left(\frac{n+2}{n+4}\right)$

7. $\displaystyle\sum_{n=0}^{\infty} \cos\left(\frac{n+2}{n^2+4}\right)$

8. $\displaystyle\sum_{n=2}^{\infty} \log\left(\frac{n^2+2}{n^2-2}\right)$

9. $\displaystyle\sum_{n=0}^{\infty} \sin\left(\frac{n+2}{n^3+4}\right)$

Stabilire se le seguenti serie a termini di segno variabile convergono assolutamente e/o semplicemente, precisando il criterio utilizzato.

10. $\displaystyle\sum_{n=1}^{\infty} \frac{(-1)^n}{\sqrt{n}}$

11. $\displaystyle\sum_{n=1}^{\infty} \frac{\cos\left(\frac{\pi}{2}n\right)}{n}$

12. $\displaystyle\sum_{n=2}^{\infty} \frac{\sin(\log n)}{n^2 \log n}$

Studiare il carattere delle seguenti serie, precisando il criterio utilizzato e giustificandone l'applicazione.

13. $\displaystyle\sum_{n=1}^{\infty} \frac{\sin n + (-1)^n n}{n^2}$

14. $\displaystyle\sum_{n=1}^{\infty} \frac{n \cdot 2^n}{e^{n/2}}$

15. $\displaystyle\sum_{n=1}^{\infty} \frac{1}{\sqrt{n^3-n}}$

16. $\displaystyle\sum_{n=1}^{\infty} \frac{n+\log n}{(n+\cos n)^3}$

17. $\displaystyle\sum_{n=1}^{\infty} \frac{\log n}{n^2}$

18. $\displaystyle\sum_{n=1}^{\infty} \frac{2^n+1}{3^n+n}$

19. Sapendo che $\sum_{n=0}^{\infty} q^n = \frac{1}{1-q}$ se $|q| < 1$, scrivere un algoritmo per calcolare $\frac{1}{0{,}815}$ con un errore inferiore a 10^{-3}.

20. Calcolare la lunghezza della spezzata poligonale inscritta nel triangolo ABC ottenuta mandando successivamente, a partire da A, le perpendicolari ai lati CB e AB. Sia $a = \overline{AC}$ e γ l'angolo $\widehat{ACB}$ (v. fig. 5.3).

Figura 5.3.

21. Data una circonferenza di raggio R, si consideri un triangolo equilatero inscritto, quindi una seconda circonferenza inscritta in questo triangolo, poi ancora un triangolo equilatero inscritto nella seconda circonferenza, e così via. Calcolare la somma delle aree degli infiniti triangoli così ottenuti (v. fig. 5.4).

Figura 5.4.

2 SERIE DI TAYLOR. ESPONENZIALE COMPLESSO

Il concetto di serie numerica, discusso nel paragrafo precedente, è il primo passo per introdurre serie di tipo più generale, ossia *serie di funzioni*, che si possono vedere come serie numeriche dipendenti da un parametro x: quando queste serie convergono per ogni valore x appartenente a un certo intervallo, rappresentano delle funzioni di nuovo tipo. Nel 2° volume di questo corso incontreremo due importanti classi di serie di funzioni, ovvero le serie trigonometriche e le serie di potenze. Qui (par. 2.1) daremo solo un primo cenno alle serie di potenze, dette anche *serie di Taylor*, che si possono vedere come una estensione naturale degli sviluppi di Taylor studiati nel cap.4. Un'altra estensione naturale dei concetti visti finora sulle serie è quello che consiste nello studiare serie *nel campo complesso*: come vedremo, questo permetterà di gettare nuova luce sulle funzioni esponenziali e trigonometriche, scoprendo la loro sorprendente parentela. Di questo ci occuperemo nel par. 2.2.

Il lettore noterà come, mentre nel par.1 ci siamo perlopiù concentrati sul problema di discutere il carattere di una serie, e solo in pochissimi casi ne abbiamo saputo calcolare la somma, in questo paragrafo incontreremo diverse situazioni interessanti in cui si riesce a calcolare la somma: in generale, infatti, il modo naturale per saper calcolare la somma di una serie numerica è vederla (quando questo è possibile) come serie di funzioni, calcolata per un particolare valore della variabile. Questa osservazione potrà essere meglio compresa nello studio delle serie di funzioni che si svilupperà nel 2° vol.

2.1 Serie di Taylor delle trascendenti elementari

Riprendiamo ora, per darne una nuova generalizzazione, la formula di Taylor con resto secondo Lagrange (v. cap.4, par. 7.4), che riscriviamo nella forma:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k + E_n(x)$$

dove $E_n(x) = \dfrac{f^{(n+1)}(c)}{(n+1)!}(x - x_0)^{n+1}$, e c è un opportuno punto tra x_0 e x.

Se la funzione f ha derivate di ogni ordine, si può scrivere la serie

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!}(x - x_0)^k$$

detta *serie di Taylor per la funzione* f (centrata in x_0). Non abbiamo, per il momento, alcuna garanzia che la serie sia convergente e che la sua somma sia $f(x)$. Ciò sarà vero, in un certo punto $x \in (a, b)$, se l'errore $E_n(x) \to 0$ per $n \to +\infty$; infatti ciò equivale a dire che la somma parziale n-esima della serie, $T_{n,x_0}(x)$, ammette limite finito e tale limite è precisamente $f(x)$.

Se accade che, per ogni x in un intervallo I, $E_n(x) \to 0$, diremo che $f(x)$ è *sviluppabile in serie di Taylor* nell'intervallo I.

Ci sono funzioni (infinitamente differenziabili) che sono sviluppabili in serie di Taylor su tutta la retta, altre che lo sono solo su un intervallo limitato, altre per le quali l'intervallo si riduce a un punto solo (il punto x_0).

Ci occupiamo ora principalmente delle tre funzioni elementari

$$e^x \qquad \cos x \qquad \sin x$$

che mostreremo essere sviluppabili in serie di Taylor su tutto $\mathbb{R}$.

La serie esponenziale

Scriviamo lo sviluppo di MacLaurin all'ordine n per e^x, con resto secondo Lagrange: per ogni intero n ed $x \in \mathbb{R}$ esiste un punto c, compreso tra 0 e x, tale che

$$e^x = \sum_{k=0}^{n} \frac{x^k}{k!} + \frac{x^{n+1}}{(n+1)!} e^c$$

Fissiamo ora x e facciamo tendere n a $+\infty$. Il punto c può variare con n ma, essendo sempre compreso tra 0 e x, si può comunque affermare che

$$e^c \leq \begin{cases} e^x & \text{se } x > 0 \\ 1 & \text{se } x < 0 \end{cases}$$

In ogni caso, per x fissato e $n \to +\infty$, e^c si mantiene limitato, mentre

$$\frac{x^{n+1}}{(n+1)!} \to 0$$

(confronto di infiniti che si dimostra, ad esempio, col criterio del rapporto, v. cap. 3, Teorema 3.11).

Ne segue quindi:

$$\boxed{e^x = \lim_{n\to+\infty} \sum_{k=0}^{n} \frac{x^k}{k!} = \sum_{k=0}^{\infty} \frac{x^k}{k!} \qquad \text{per ogni } x \in \mathbb{R}}$$

Abbiamo quindi dimostrato che la funzione e^x si può scrivere come somma di una serie di potenze, la sua serie di Taylor, convergente per ogni $x \in \mathbb{R}$. In particolare, per $x = 1$ otteniamo

$$e = \sum_{k=0}^{\infty} \frac{1}{k!}. \tag{2.1}$$

A questo punto conosciamo *due diversi algoritmi* per il calcolo del numero e: il primo (la definizione stessa di e) consiste nel calcolare

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

ovvero, dal punto di vista del calcolo numerico approssimato, calcolare il valore di $\left(1 + \frac{1}{n}\right)^n$ per un n molto grande; il secondo consiste nel calcolare la somma della serie (2.1) ovvero, dal punto di vista del calcolo numerico approssimato, calcolare la somma

dei primi n termini della serie per un n molto grande. Si verifica che il secondo metodo è molto più efficiente del primo. A titolo di esempio, si confrontino i seguenti valori:

n	$\left(1+\frac{1}{n}\right)^n$	$\sum_{k=0}^{n} \frac{1}{k!}$	e
1	2,	2	2,71828...
2	2,25	2,5	2,71828...
3	2,37037	2,66667	2,71828...
4	2,44141	2,70833	2,71828...
5	2,48832	2,71667	2,71828...
6	2,52163	2,71806	2,71828...
7	2,5465	2,71825	2,71828...
8	2,56578	2,71828	2,71828...
9	2,58117	2,71828	2,71828...
10	2,59374	2,71828	2,71828...

Le serie delle funzioni trigonometriche elementari

Un discorso completamente analogo si può ripetere per le funzioni seno e coseno. Il termine $f^{(n+1)}(c)$ che compare nelle rispettive formule di MacLaurin, ha la forma $\pm\cos c$ o $\pm\sin c$, pertanto ha valore assoluto ≤ 1. Questa limitazione è quel che basta a ripetere la dimostrazione precedente e concludere che

$$\boxed{\begin{aligned} \sin x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k+1}}{(2k+1)!} \\ \cos x &= \sum_{k=0}^{\infty} (-1)^k \frac{x^{2k}}{(2k)!} \end{aligned}}$$

per ogni $x \in \mathbb{R}$.

Le formule trovate sono molto interessanti, per almeno due motivi. Anzitutto, forniscono un algoritmo efficace per il calcolo numerico approssimato delle funzioni trigonometriche elementari, per valori arbitrari delle variabili: il fatto che le serie siano a segni alterni, unito al carattere rapidamente decrescente (in valore assoluto) del termine generale, rendono queste serie particolarmente adatte al calcolo approssimato.

Esempio

2.1 Si voglia calcolare $\sin(0{,}8)$. Posto

$$s_n = \sum_{k=0}^{n} (-1)^k \frac{(0{,}8)^{2k+1}}{(2k+1)!}$$

si ha

$$s_{2n-1} \leq \sin(0{,}8) \leq s_{2n} \text{ per ogni } n.$$

Ad esempio, per $n = 2$ valgono le disuguaglianze:

$$s_3 \leq \sin(0{,}8) \leq s_4$$
$$\text{ma } s_3 \simeq s_4 \simeq 0{,}717356$$

il che significa che anche $\sin(0{,}8) \simeq 0{,}717356$.

In questo caso la convergenza è molto veloce e basta un n abbastanza piccolo; concettualmente, comunque, il fatto di avere a disposizione una *serie* che converge a $\sin x$ garantisce di poter eseguire il calcolo con precisione arbitraria.

Approfondendo questa osservazione, notiamo che, ad eccezione dei (pochi) casi in cui x è un angolo notevole per cui i valori di $\sin x$ e $\cos x$ sono noti dalla geometria elementare (es. $\sin \frac{\pi}{6}$), oppure sono ricavabili da questi mediante opportune identità trigonometriche (es. $\sin \frac{\pi}{12} = \sin\left(\frac{1}{2} \cdot \frac{\pi}{6}\right) = \dots$), la "definizione" trigonometrica di $\sin x, \cos x$ non dà alcun criterio *operativo* per il calcolo effettivo del valore della funzione; in un certo senso, quindi, le serie che abbiamo appena scritto si possono considerare come la prima definizione completamente rigorosa di queste funzioni "elementari"!

Serie di potenze

In generale, si dice *serie di potenze* una serie del tipo

$$\sum_{k=0}^{\infty} a_k x^k$$

con a_k costanti reali (o complessi, come vedremo nel prossimo paragrafo), e x variabile reale (o complessa). La più semplice serie di potenze è la serie geometrica, che converge per $|x| < 1$:

$$\sum_{k=0}^{\infty} x^k = \frac{1}{1-x}$$

Altri sviluppi in serie di potenze per funzioni notevoli sono i seguenti:

$$\operatorname{Sh} x = \sum_{k=0}^{\infty} \frac{x^{2k+1}}{(2k+1)!} \qquad \text{per ogni } x \in \mathbb{R}$$

$$\operatorname{Ch} x = \sum_{k=0}^{\infty} \frac{x^{2k}}{(2k)!} \qquad \text{per ogni } x \in \mathbb{R}$$

$$\log(1+x) = \sum_{k=1}^{\infty} (-1)^{k+1} \frac{x^k}{k} \qquad \text{per } |x| < 1$$

$$\text{per } \alpha \in \mathbb{R} \qquad (1+x)^\alpha = \sum_{k=0}^{\infty} \binom{\alpha}{k} x^k \qquad \text{per } |x| < 1$$

dove si è indicato con

$$\binom{\alpha}{k} = \frac{\alpha(\alpha-1)(\alpha-2)\dots(\alpha-k+1)}{k!}$$

il coefficiente binomiale k-esimo.

Nel secondo volume riprenderemo lo studio delle serie di potenze, inquadrandolo nel punto di vista più generale dello studio delle *serie di funzioni*.

2.2 Serie nel campo complesso. Esponenziale complesso

Cominciamo a precisare il concetto di serie a termini complessi. Nel paragrafo 5.1, parlando di serie numeriche, ci siamo concentrati sulle serie a termini reali; tuttavia, le nozioni viste si estendono in modo naturale al caso di una serie

$$\sum_{n=0}^{\infty} a_n \qquad \text{con} \quad a_n \in \mathbb{C}$$

Infatti, una serie è definita mediante la successione delle somme parziali $s_n = \sum_{k=0}^{n} a_k$; per dar senso al concetto di serie a termini complessi è sufficiente dunque dire cosa si intenda per limite di una successione s_n a valori complessi. Diremo che

$$s_n \to l \quad (l \in \mathbb{C}) \quad \text{se} \quad |s_n - l| \to 0.$$

Poiché il modulo di un numero complesso è un numero reale, la nozione di limite nel campo complesso è ricondotta in questo modo a quella di limite nel campo reale.

Esempio

2.2 *Serie geometrica nel campo complesso.* Consideriamo la serie geometrica

$$\sum_{n=0}^{\infty} z^n \text{ con } z \in \mathbb{C}.$$

Si può ripetere il ragionamento visto nel par.1.1 a proposito della serie geometrica nel campo reale:

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

Se ora z è un qualsiasi numero complesso di modulo $|z| < 1$, si ha

$$|z^{n+1}| = |z|^{n+1} \to 0, \text{ e } s_n \to \frac{1}{1-z}.$$

Pertanto

$$\sum_{n=0}^{\infty} z^n = \frac{1}{1-z} \text{ per ogni } z \in \mathbb{C}, |z| < 1.$$

Per le serie a termini complessi si dimostra che continua a valere il *criterio della convergenza assoluta*:

Teorema 5.5 *Sia $\sum_{n=0}^{\infty} a_n$ una serie a termini a_n complessi. Se la serie (a termini reali non negativi) $\sum_{n=0}^{\infty} |a_n|$ converge, allora converge anche $\sum_{n=0}^{\infty} a_n$.*

Una dimostrazione di questo teorema è suggerita nei Complementi. Questa semplice condizione permette di provare la convergenza di una serie a termini complessi applicando alla serie dei moduli i criteri che conosciamo per le serie a termini positivi. Applicheremo ora questi fatti alla serie esponenziale.

Esponenziale complesso

La serie $\sum_{n=0}^{\infty} \frac{x^n}{n!}$, la cui somma è e^x per ogni $x \in \mathbb{R}$, si chiama *serie esponenziale*; per mezzo di essa è possibile definire l'espressione e^z quando $z = x + iy$ è un numero complesso. Si osserva infatti che la serie

$$\sum_{n=0}^{\infty} \frac{z^n}{n!}$$

è assolutamente convergente per ogni $z \in \mathbb{C}$ (infatti la serie a termini positivi $\sum_{n=0}^{\infty} |z|^n/n!$ è convergente, come si vede ad esempio applicando il criterio del rapporto) e perciò la sua somma è un numero complesso che indicheremo con e^z. Si dimostra che questa definizione è coerente con le proprietà formali dell'esponenziale:

Teorema 5.6 *Definendo*

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

si ha:

$$e^{z_1+z_2} = e^{z_1} \cdot e^{z_2} \ \forall z_1, z_2 \in \mathbb{C}.$$

Una traccia della dimostrazione di questo teorema sarà fornita nei Complementi. Risulta allora, posto $z = x + iy$,

$$e^z = e^x \cdot e^{iy}$$

dove

$$e^{iy} = \sum_{n=0}^{\infty} \frac{(iy)^n}{n!} = 1 + iy - \frac{y^2}{2} - i\frac{y^3}{3!} + \frac{y^4}{4!} + i\frac{y^5}{5!} - \frac{y^6}{6!} - i\frac{y^7}{7!} + \ldots =$$
$$= \sum_{k=0}^{\infty} (-1)^k \frac{y^{2k}}{(2k)!} + i\sum_{k=0}^{\infty} (-1)^k \frac{y^{2k+1}}{(2k+1)!} = \cos y + i \sin y$$

Analogamente si mostra che

$$e^{-iy} = \cos y - i \sin y.$$

Abbiamo così trovato un legame (nel campo complesso) tra le funzioni trigonometriche (sin, cos) e la funzione esponenziale. Le formule seguenti, dette *formule di Eulero*, sono valide per ogni $x \in \mathbb{R}$:

$$\boxed{\begin{array}{ll} e^{ix} = \cos x + i \sin x & \cos x = \dfrac{e^{ix} + e^{-ix}}{2} \\ e^{-ix} = \cos x - i \sin x & \sin x = \dfrac{e^{ix} - e^{-ix}}{2i} \end{array}}$$

La formula di Eulero $e^{iy} = \cos y + i \sin y$ è estremamente utile nel calcolo coi numeri complessi. Essa permette di trasformare la forma trigonometrica dei numeri complessi nella forma esponenziale, al modo seguente:

$$z = \rho(\cos\theta + i\sin\theta) = \rho e^{i\theta}$$

con $\rho \geq 0$ modulo e $\theta \in \mathbb{R}$ argomento del numero complesso z. La forma esponenziale dei numeri complessi serve agli stessi scopi della forma trigonometrica, ma è più compatta da scrivere e porta a formule più facili da ricordare.

Ad esempio il teorema di de Moivre sul prodotto di numeri complessi in forma trigonometrica:

$$\rho_1(\cos\theta_1 + i\sin\theta_1) \cdot \rho_2(\cos\theta_2 + i\sin\theta_2) = \rho_1\rho_2(\cos(\theta_1+\theta_2) + i\sin(\theta_1+\theta_2))$$

è estremamente naturale in forma esponenziale, perché formalmente è un'applicazione ovvia delle proprietà delle potenze:

$$\rho_1 e^{i\theta_1} \cdot \rho_2 e^{i\theta_2} = \rho_1\rho_2 e^{i(\theta_1+\theta_2)}$$

Il logaritmo nel campo complesso

La funzione logaritmo può essere definita nel campo complesso usando, formalmente, la "solita" definizione: diciamo che

$$w = \log z \iff z = e^w$$

Dato $z = \rho e^{i\theta}$, cerchiamo il suo logaritmo $w = x + iy$. Dovrà essere:

$$\rho e^{i\theta} = e^{x+iy} = e^x e^{iy}$$

da cui deduciamo

$$\begin{cases} \rho = e^x, & \text{cioè } x = \ln\rho \\ e^{i\theta} = e^{iy}, & \text{cioè } y = \theta + 2k\pi \qquad k \in \mathbb{Z} \end{cases}$$

(si ricordi la periodicità di $e^{i\theta}$!) In definitiva, si ha:

$$\log(\rho e^{i\theta}) = \ln\rho + i(\theta + 2k\pi) \qquad k \in \mathbb{Z}$$

Ciò significa che, nel campo complesso, il logaritmo di un numero non è univocamente determinato, ma ogni numero (diverso da zero) ha infiniti logaritmi (una successione). Si chiama *valore principale del logaritmo* quello che si ottiene per $k = 0$ e $\theta \in [0, 2\pi)$.

Ad esempio, calcoliamo

$$\log(1+i)$$

Poiché $|1+i| = \sqrt{2}$ e $\arg(1+i) = \frac{\pi}{4}$, si ha:

$$\log(1+i) = \frac{1}{2}\ln 2 + i\left(\frac{\pi}{4} + 2k\pi\right)$$

Il logaritmo principale di $(1+i)$ è $\frac{1}{2}\ln 2 + i\frac{\pi}{4}$.

 © 978-88-08-06485-1

L'elevamento a potenza complessa

Si può ora dare senso, nel campo complesso, anche alle potenze a esponente qualsiasi, z^α, con $\alpha \in \mathbb{C}$ (o reale). Si noti che finora sapevamo calcolare z^α solo con α intero (e allora z^α era univocamente determinato) o $\alpha = m/n$ razionale (mediante estrazione di radice n-esima e allora z^α indicava n numeri). Vedremo ora che in generale se α è complesso (o reale ma irrazionale), z^α indicherà infiniti numeri. Poniamo, per definizione:

$$z^\alpha = e^{\alpha \log z}$$

dove $\log z$ indica il logaritmo nel campo complesso, definito in precedenza.

Chiameremo *valore principale* di z^α quello che si ottiene da $e^{\alpha \log z}$ scegliendo il logaritmo principale di z. Ad esempio, calcoliamo:

$$i^{\sqrt{2}} = e^{\sqrt{2} \log i}$$

Poiché $i = e^{i\frac{\pi}{2}}$, $\log i = i\left(\frac{\pi}{2} + 2k\pi\right)$, e

$$i^{\sqrt{2}} = e^{\sqrt{2}i(\frac{\pi}{2}+2k\pi)} = \cos\left(\frac{\pi}{\sqrt{2}} + 2\sqrt{2}k\pi\right) + i\sin\left(\frac{\pi}{\sqrt{2}} + 2\sqrt{2}k\pi\right)$$

Si osservi che, per l'irrazionalità di $\sqrt{2}$, i numeri $2\sqrt{2}k\pi$ non sono mai multipli di 2π, perciò i numeri scritti sono tutti diversi, al variare di $k \in \mathbb{Z}$. Il valore principale di $i^{\sqrt{2}}$ è

$$\cos\frac{\pi}{\sqrt{2}} + i\sin\frac{\pi}{\sqrt{2}}$$

Criterio di Dirichlet e somma di alcune serie trigonometriche

Presentiamo ora un criterio per lo studio della convergenza di serie a termini complessi, che permette di trattare utilmente anche alcuni casi in cui la serie non converge assolutamente. Come vedremo, questo criterio avrà delle ricadute interessanti anche su certe serie a termini reali, ma di segno variabile, che non sarebbe possibile studiare con i criteri fin qui incontrati.

Teorema 5.7 (Criterio di Dirichlet) *Siano $\{a_n\}, \{b_n\}$ due successioni tali che:*

(1) *$\{a_n\}$ è a valori complessi e la sua successione delle somme parziali è limitata; in altre parole, esiste una costante $M > 0$ tale che, detta $A_n = \sum_{k=0}^{n} a_k$ risulta*

$$|A_n| \leq M \ \forall n.$$

(2) *$\{b_n\}$ è a valori reali positivi e tende monotonamente a zero:*

$$b_n > 0 \ \forall n; \quad b_{n+1} \leq b_n \ \forall n; \quad \lim_{n\to\infty} b_n = 0.$$

Allora la serie $\sum a_n b_n$ è convergente.

Si presti attenzione all'ipotesi (1): non si richiede semplicemente che la successione $\{a_n\}$ sia limitata, ma che la successione $\{A_n\}$ delle sue somme parziali sia limitata (il che è più impegnativo, in generale). Prima di occuparci della dimostrazione del teorema, mostriamone un'interessante applicazione:

Esempio

2.3 Le due serie (a termini reali di segno variabile):

$$\sum_{n=1}^{\infty} \frac{\cos n}{n}; \qquad \sum_{n=1}^{\infty} \frac{\sin n}{n}$$

convergono. Infatti: applichiamo il criterio di Dirichlet alle successioni:

$$a_n = e^{in}; \qquad b_n = \frac{1}{n}.$$

Ovviamente b_n è reale, positiva, e tende monotonamente a zero. Quanto ad a_n, scrivendo:

$$A_n = \sum_{k=0}^{n} a_k = \sum_{k=0}^{n} \left(e^i\right)^k = \frac{1 - e^{i(n+1)}}{1 - e^i}$$

si ha

$$|A_n| = \left| \frac{1 - e^{i(n+1)}}{1 - e^i} \right| \leq \frac{2}{|1 - e^i|}.$$

Pertanto A_n è limitata, e possiamo affermare che converge la serie a termini complessi

$$\sum_{n=1}^{\infty} a_n b_n = \sum_{n=1}^{\infty} \frac{e^{in}}{n} = \sum_{n=1}^{\infty} \frac{\cos n + i \sin n}{n}.$$

Separando parte reale e parte immaginaria, si conclude che convergono le due serie a termini reali $\sum_{n=1}^{\infty} \frac{\cos n}{n}; \sum_{n=1}^{\infty} \frac{\sin n}{n}$.

Si rifletta sul fatto che la convergenza di queste serie non si sarebbe potuta ottenere in base ai criteri studiati nel par.1. Infatti queste sono serie a termini di segno variabile, ma non a segni alterni; l'unico criterio visto nel par. 1 in quesi casi è quello della convergenza assoluta, ma non si vede perché le serie $\sum \frac{|\cos n|}{n}; \sum \frac{|\sin n|}{n}$ dovrebbero essere convergenti.

Il risultato appena discusso sarà molto utile nello studio della convergenza delle *serie di Fourier*, di cui ci occuperemo nel secondo volume.

Esercizi

Scrivere in forma algebrica $(a + ib)$ i seguenti numeri complessi:

22 e^{z^2}, con $z = x + iy$, $x, y \in \mathbb{R}$

23 $e^{\overline{z}}$, con $z = x + iy$, $x, y \in \mathbb{R}$

24 ie^{z}, con $z = x + iy$, $x, y \in \mathbb{R}$

25 e^{2z+3i}, con $z = x + iy$, $x, y \in \mathbb{R}$

26 $$e^{-iz}, \text{ con } z = x + iy,\ x, y \in \mathbb{R}$$

27 Sia $x \in \mathbb{R}$. Calcolare la parte reale del numero complesso:

$$\frac{1}{2+i} e^{(3-i)x}$$

28 Sia $x \in \mathbb{R}$. Calcolare la parte immaginaria del numero complesso:

$$(5+2i)e^{-3x+2ix}$$

Risolvere le seguenti equazioni nel campo complesso, ponendo $z = x + iy$ e riscrivendo in forma esponenziale ambo i membri dell'equazione:

29 $e^z = -1$

30 $e^z = 2i$

31 Calcolare nel campo complesso i logaritmi dei seguenti numeri:

$$-1 \qquad 2 \qquad 1 + i\sqrt{3}$$

32 Calcolare nel campo complesso le seguenti potenze, riscrivendo in forma algebrica il risultato:

$$2^i \qquad 2^{\sqrt{2}} \qquad (-2)^{\sqrt{3}} \qquad (1-i)^{3/5} \qquad (1+i)^{2+3i} \qquad i^i$$

33 Scrivere parte reale e parte immaginaria del valore principale della funzione

$$f(x) = x^{a-ib}, \text{ per } x, a, b \in \mathbb{R}$$

34 Utilizzando opportunamente lo sviluppo in serie di $\log(1+x)$, calcolare la somma della serie

$$\sum_{n=1}^{\infty} \frac{1}{n2^n}$$

35 Utilizzando opportunamente lo sviluppo in serie di e^x, calcolare la somma della serie

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{n!2^n}.$$

COMPLEMENTI

1 Sfruttando la somma della serie geometrica nel campo complesso, dimostrare che:

$$\sum_{n=0}^{\infty} \rho^n \cos(n\theta) = \frac{1 - \rho\cos\theta}{1 + \rho^2 - 2\rho\cos\theta}$$

$$\sum_{n=1}^{\infty} \rho^n \sin(n\theta) = \frac{\rho\sin\theta}{1 + \rho^2 - 2\rho\cos\theta}$$

per ogni $\rho < 1$, ogni θ.
Suggerimento: scrivere $z = \rho(\cos\theta + i\sin\theta)$, e separare parte reale e immaginaria nella somma della serie geometrica $\sum z^n$).

Calcolare quindi la somma delle serie

$$\sum_{n=0}^{\infty} \frac{\cos(n\theta)}{2^n}; \sum_{n=1}^{\infty} \frac{\sin(2n\theta)}{3^n}.$$

2 Dimostrare il criterio della convergenza assoluta per le serie a termini complessi (v. par. 2.2).
Suggerimento: ragionare in modo analogo alla dimostrazione data nel par. 1.3 per le serie a termini reali, introducendo le *quattro* successioni di somme parziali $s_n^{\mathrm{Re}^+}, s_n^{\mathrm{Re}^-}, s_n^{\mathrm{Im}^+}, s_n^{\mathrm{Im}^-}$ (con l'ovvio significato dei simboli).

3 Vogliamo dimostrare il seguente teorema, enunciato nel par. 2.2: ponendo $e^z = \sum_{n=0}^{\infty} \frac{z^n}{n!}$, risulta:

$$e^{z_1+z_2} = e^{z_1} \cdot e^{z_2} \ \forall z_1, z_2 \in \mathbb{C}.$$

(a) Consideriamo la successione delle somme parziali della serie che definisce $e^{z_1+z_2}$:

$$s_n = \sum_{k=0}^{n} \frac{(z_1+z_2)^k}{k!}.$$

Sviluppando la potenza $(z_1+z_2)^k$ con la formula del binomio di Newton e applicando opportunamente le proprietà delle sommatorie, provare che:

$$\sum_{k=0}^{n} \frac{(z_1+z_2)^k}{k!} = \sum_{k=0}^{n} \frac{z_1^k}{k!} \sum_{h=0}^{n-k} \frac{z_2^h}{h!}. \tag{2.2}$$

(b) Se al posto del secondo membro dell'ultima identità avessimo

$$\sum_{k=0}^{n} \frac{z_1^k}{k!} \sum_{h=0}^{n} \frac{z_2^h}{h!}$$

sarebbe ora sufficiente passare al limite per $n \to \infty$. Poiché

$$\sum_{k=0}^{n} \frac{z^k}{k!} \to e^z \text{ per } n \to \infty,$$

otterremmo proprio $e^{z_1+z_2} = e^{z_1} \cdot e^{z_2}$. La complicazione nasce dal fatto che, invece, nel secondo membro della (2.2), la sommatoria interna dipende anche dall'indice k. Si può allora spezzare la sommatoria al modo seguente (supponiamo per semplicità che n sia pari):

$$\begin{aligned} \sum_{k=0}^{n} \frac{z_1^k}{k!} \sum_{h=0}^{n-k} \frac{z_2^h}{h!} &= \sum_{k=0}^{n/2} \frac{z_1^k}{k!} \sum_{h=0}^{n/2} \frac{z_2^h}{h!} + \sum_{k=0}^{n/2} \frac{z_1^k}{k!} \sum_{h=n/2+1}^{n-k} \frac{z_2^h}{h!} + \sum_{k=n/2+1}^{n} \frac{z_1^k}{k!} \sum_{h=0}^{n-k} \frac{z_2^h}{h!} \\ &\equiv A_n + B_n + C_n. \end{aligned} \tag{2.3}$$

Ora il termine A_n ha la forma giusta per poter affermare che $A_n \to e^{z_1} \cdot e^{z_2}$. Si tratta di dimostrare che B_n e C_n tendono a zero. Per far questo, si ragiona così:

$$|B_n| \le \sum_{k=0}^{n/2} \frac{|z_1|^k}{k!} \sum_{h=n/2+1}^{n-k} \frac{|z_2|^h}{h!} \le \sum_{k=0}^{\infty} \frac{|z_1|^k}{k!} \sum_{h=n/2+1}^{\infty} \frac{|z_2|^h}{h!} = e^{|z_1|} \cdot \sum_{h=n/2+1}^{\infty} \frac{|z_2|^h}{h!} \to 0$$

perché l'ultima serie scritta è il resto di una serie convergente. In modo analogo si ragiona con C_n. Questo conclude la dimostrazione.

Per capire l'idea geometrica con cui si è eseguito il passaggio (2.3) si rappresenti nel piano l'insieme delle coppie (k, h) su cui si somma l'espressione $\sum_{k=0}^{n} \frac{z_1^k}{k!} \sum_{h=0}^{n-k} \frac{z_2^h}{h!}$. L'idea è di vedere questo triangolo di punti a coordinate intere come unione di un quadrato (che dà l'espressione A_n che vorremmo avere) e due triangoli più piccoli, che dovranno dare somme infinitesime con n. Si ha:

Figura 5.5.

4 Un modo alternativo per introdurre l'esponenziale complesso senza usare le serie è il seguente. Per $z = x + iy$, $x, y \in \mathbb{R}$, poniamo per definizione:

$$\exp(x + iy) = e^x (\cos y + i \sin y). \tag{2.4}$$

Sfruttando le formule di De Moivre per i numeri complessi scritti in forma trigonometrica, dimostrare in base alla (2.4) che:

$$\begin{aligned} e^{z_1+z_2} &= e^{z_1} \cdot e^{z_2} \\ (e^{z_1})^n &= e^{nz_1} \end{aligned} \tag{2.5}$$

$\forall z_1, z_2 \in \mathbb{C}, n \in \mathbb{N}$. Inoltre, ovviamente, se $z \in \mathbb{R}$ (cioè se $z = x$ e $y = 0$) si riottiene che $\exp(z) = e^x$. Perciò la funzione di variabile complessa appena definita può chiamarsi "esponenziale complesso". Si noti che le formule di Eulero sono tutte conseguenze delle proprietà (2.5) dell'esponenziale.

5 Discutere, al variare del parametro reale x, la convergenza delle serie

$$\sum \frac{\sin(nx)}{n}; \sum \frac{\cos(nx)}{n}.$$

Suggerimento: applicare il criterio di Dirichlet, adattando l'esempio che abbiamo presentato subito dopo il criterio.

6 Dare condizioni sulla successione $\{a_n\}$ affinché le serie

$$\sum a_n \cos(nx); \sum a_n \sin(nx)$$

risultino convergenti.
Suggerimento: applicare il criterio di Dirichlet.

7 Il criterio di Leibniz per le serie a segni alterni si può dedurre anche dal criterio di Dirichlet. Come?

6 Calcolo integrale per funzioni di una variabile

1 INTRODUZIONE AL CALCOLO INTEGRALE

Come abbiamo fatto nel capitolo precedente per il calcolo differenziale, introduciamo anche il calcolo integrale partendo da alcuni dei problemi che storicamente hanno portato alla sua nascita.

Che cosa vuol dire misurare l'area di una figura piana a contorno curvilineo?
Questo è il più antico problema di calcolo infinitesimale che sia stato affrontato e risolto, almeno in qualche caso particolare.

L'idea elementare di misura di un'area è: fissare un quadratino unità di misura e contare quante volte questo può essere riportato nella figura da misurare. Per questa via, si riescono però a misurare solo rettangoli e figure scomponibili in rettangoli. Andando oltre, la geometria elementare postula che se due figure sono equiscomponibili (cioè possono essere divise in un numero *finito* di parti rispettivamente sovrapponibili), allora sono equivalenti (cioè hanno la stessa area). Inoltre, si conviene che l'area sia *additiva*, cioè se una figura T è unione di due (o più) figure T_1, T_2 (disgiunte), l'area di T è la somma dell'area di T_1 e di T_2. L'utilizzo di questi princìpi permette di calcolare le aree di *poligoni* qualsiasi.

Comunque, per questa via non si arriverà mai a dar senso al concetto di *area del cerchio*, per esempio (né a calcolarla): infatti il cerchio non è scomponibile in un numero finito di triangoli e/o rettangoli.

Che cosa vuol dire misurare la lunghezza di una curva?
Il problema è analogo al precedente: non è possibile riportare un segmento unitario su una curva, e dunque "contare quante volte esso è contenuto". Cos'è dunque la lunghezza di una curva?

Entrambi questi problemi vengono risolti, nella geometria elementare di Euclide (attorno al 300 a.C.), nel caso particolare del cerchio (area del cerchio, lunghezza della circonferenza), con un ragionamento che fa ricorso all'idea di approssimazioni successive, sempre più accurate: si calcola l'area, o il perimetro, dei poligoni regolari inscritti e circoscritti, e ci si chiede cosa accade quando il numero dei lati cresce indefinitamente. Ecco dunque un'idea di *procedimento infinito* e di *limite*, che si innesta nella geometria elementare come un'idea nuova.

Archimede di Siracusa (287-212 a.C.) applicò ingegnosamente idee simili (il suo *metodo di esaustione*) al calcolo dell'area del segmento di parabola e ad alcuni altri casi specifici. Nel 1600, vari matematici trovarono altri metodi per calcolare l'area sottesa al grafico di semplici funzioni: x^α ($\alpha > -1$) (Fermat, 1636), $1/x$ (Mercator, 1668). Tuttavia, fu solo con l'invenzione del calcolo infinitesimale (fine 1600-inizio 1700) che mediante il concetto di *integrale*, inteso come opportuno *limite di somme*, venne data una definizione abbastanza generale di area di una figura piana e di lunghezza di una curva, e contemporaneamente venne fornito un algoritmo per il calcolo effettivo utilizzabile non solo in casi molto particolari: il famoso "Teorema fondamentale del calcolo integrale" (v. par. 4) di Newton, Leibniz, Giovanni Bernoulli, che ricondusse tale problema, almeno in molti casi, a un esercizio di routine: la ricerca di una primitiva (o antiderivata) di una funzione.

2 L'INTEGRALE COME LIMITE DI SOMME

2.1 La definizione di integrale

Figura 6.1.

Vediamo, ad esempio, come si potrebbe calcolare l'area compresa tra l'asse x e l'arco di parabola $y = x^2$ per $x \in [0, 1]$, usando idee simili a quelle di Fermat, uno dei "pionieri" del calcolo integrale.

Si divide l'intervallo [0,1] in n segmenti uguali di estremi $x_i = \frac{i}{n}$, $x_{i+1} = \frac{i+1}{n}$, con $i = 0, \ldots, n-1$.

Si approssima l'area cercata con la somma delle aree dei rettangoli ombreggiati in figura 6.1:

$$\text{Area} \approx S_n = \sum_{i=1}^{n-1} \underbrace{\frac{1}{n}}_{\text{base}} \cdot \underbrace{\left(\frac{i}{n}\right)^2}_{\text{altezza}} = \frac{1}{n^3} \sum_{i=1}^{n-1} i^2 = \frac{(n-1)n(2n-1)}{6n^3}$$

(per l'ultima uguaglianza, v. esercizio 5, capitolo 1, par.2)

A questo punto si può far tendere n a $+\infty$, ottenendo

$$\frac{(n-1)n(2n-1)}{6n^3} \longrightarrow \frac{1}{3}$$

e concludere così che

$$\text{Area} = \frac{1}{3}$$

La definizione di integrale che daremo segue questi stessi passi.

Sia $f : [a, b] \longrightarrow \mathbb{R}$, *limitata*. Notiamo esplicitamente che f potrebbe non essere continua; ad ogni modo, almento per il momento, assumiamo che sia almeno *definita* in ogni punto di $[a, b]$ (e limitata).

Consideriamo la suddivisione di $[a, b]$, individuata dai punti

$$a = x_0\,, \qquad x_1, x_2, \ldots, x_{n-1}\,, \qquad x_n = b$$

con

$$x_j = a + jh\,, \quad h = \frac{b-a}{n}\,, \quad j = 0, \ldots, n$$

e scegliamo in ciascuno degli n intervalli $[x_{j-1}, x_j]$, un punto arbitrario ξ_j $(j = 1, 2, \ldots, n)$.

Costruiamo la somma (detta *somma di Cauchy-Riemann*):

$$S_n = \sum_{j=1}^{n} f(\xi_j) \cdot (x_j - x_{j-1}) = \frac{b-a}{n} \sum_{j=1}^{n} f(\xi_j) \tag{2.1}$$

Figura 6.2. Una particolare scelta dei punti ξ_j per $n = 4$ e $n = 6$.

Si rifletta su questa costruzione. Ad ogni passo della costruzione, l'intervallo $[a, b]$ viene diviso in n intervallini, e in ciascuno di questi viene scelto un punto ξ_j. Si noti che, ad esempio, i punti $\xi_1, \xi_2, \ldots, \xi_6$ scelti al passo $n = 6$ (v. figura 6.2) sono in generale *tutti diversi* dai punti $\xi_1, \ldots, \xi_4$ scelti al passo $n = 4$. Ad ogni passo della costruzione, in generale, *tutti* gli addendi della somma S_n cambiano, diventando via via più numerosi e più piccoli in valore assoluto.

A questo punto si passa al limite per $n \to +\infty$ e... si spera di ottenere un risultato significativo; che cosa questo significhi esattamente, è espresso dalla prossima definizione:

Definizione 6.1 Diciamo che la funzione $f : [a, b] \to \mathbb{R}$, limitata, è *integrabile* in $[a, b]$, se, detta S_n una sua qualsiasi successione di Cauchy-Riemann, esiste finito il limite di S_n, e tale limite non dipende da come abbiamo scelto i punti ξ_j ad ogni passo della costruzione iterativa. In tal caso, si pone

$$\lim_{n\to\infty} S_n = \int_a^b f(x)\, dx.$$

Il simbolo

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

si legge "integrale da a a b di $f(x)$ in dx". In esso, il segno $\int$ (una esse allungata) è una deformazione del simbolo di somma; la scrittura $f(x)\,dx$ ricorda il prodotto del valore di $f(x)$ per la lunghezza di un piccolo intervallo sull'asse x. Tutta la scrittura, quindi, ricorda il procedimento con cui l'integrale è stato definito. La variabile x si dice *variabile d'integrazione* ed è una *variabile muta*; infatti la scrittura $\int_a^b f(t)\,dt$ ha lo stesso significato di $\int_a^b f(x)\,dx$, esattamente come, ad esempio, $\sum_{j=1}^{10} a_j = \sum_{n=1}^{10} a_n$. Ancora, notiamo che l'integrale di una funzione su un intervallo fissato è un *numero*, non una funzione.

Possiamo esprimere sinteticamente la definizione data dicendo che *l'integrale è un (particolare) limite di somme*, che può avere molte interpretazioni diverse (geometriche, fisiche...). Vediamone qualcuna.

Interpretazione geometrica (cfr. figura 6.2)
Sia $f \geq 0$, e sia f continua. Ogni addendo della (2.1) rappresenta l'area del rettangolo avente come base il segmento $[x_{j-1}, x_j]$ e come altezza $f(\xi_j)$. La somma S_n rappresenta dunque un'approssimazione dell'area della parte di piano compreso tra l'asse x, $a \leq x \leq b$, e il grafico di f (*trapezioide* individuato da f).

Passando al limite per $n \to +\infty$, si ha

$$\boxed{S_n \longrightarrow \int_a^b f(x)dx = \text{ area del trapezioide}}$$

Più precisamente, è l'integrale che costituisce una definizione precisa dell'area del trapezioide individuato da una curva $y = f(x)$, e non viceversa. Perciò, è il calcolo infinitesimale che permette di dar senso all'idea di *area di una figura piana* in generale.

Figura 6.3.

Si noti che, nel caso in cui f cambia segno, l'integrale rappresenta una somma di *aree con segno*. Nel caso in figura 6.3 si ha:

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

$$= \text{area } (T_1) - \text{area } (T_2) + \text{area } (T_3)$$

Ad esempio, per simmetria si ha:

$$\int_0^{2\pi} \sin x\,dx = 0$$

Figura 6.4.

Si noti anche il caso elementare: se c è una costante, $\int_a^b c\,dx = c\,(b-a)$ (area del rettangolo di altezza c e base $(b-a)$).

Interpretazione cinematica

Supponiamo che un punto materiale si muova lungo una traiettoria fissata, con velocità $v(t)$ variabile nel tempo, e chiediamoci quanto spazio è percorso dal punto nell'intervallo di tempo $[0, T]$. Suddividiamo $[0, T]$ in n intervallini $[t_{j-1}, t_j]$ sufficientemente piccoli da poter pensare che in ciascuno di essi la velocità vari di poco. Dunque, se ξ_j è un qualsiasi punto di $[t_{j-1}, t_j]$, $v(\xi_j)$ sarà circa uguale alla velocità media del punto nell'intervallo di tempo $[t_{j-1}, t_j]$; perciò lo spazio percorso in questo intervallino di tempo sarà $v(\xi_j)(t_j - t_{j-1})$, e lo spazio totale percorso sarà

$$\sum_{j=1}^{n} v(\xi_j)(t_j - t_{j-1})$$

Al crescere di n questa approssimazione dello spazio percorso si fa sempre migliore. Perciò il valore vero dello spazio percorso è il limite per $n \to \infty$ di questa espressione, ossia:

$$\boxed{\text{Spazio percorso in } [0, T] = \int_0^T v(t)\, dt}$$

Lo spazio percorso (o più precisamente lo spostamento netto) è l'integrale della velocità rispetto al tempo.

Interpretazione meccanica

Supponiamo che un sistema fisico caratterizzato da pressione ($= p$) e volume ($= V$) evolva da uno stato (p_A, V_A) ad uno stato (p_B, V_B), a temperatura costante. Graficamente, nel piano V, p si ha:

Figura 6.5.

In questo caso

$$\boxed{\int_{V_A}^{V_B} p(V)\, dV = \text{ area } (T) = \text{lavoro effettuato nella trasformazione da } (p_A, V_A) \text{ a } (p_B, V_B)}$$

Infatti, ogni addendo della somma $S_n = \sum_{j=1}^{n} p(\xi_j)(V_j - V_{j-1})$ costituisce il lavoro effettuato in corrispondenza di una variazione di volume $\Delta V = V_j - V_{j-1}$ (dimensionalmente: $[p] = [\text{Forza}] \cdot [\text{lunghezza}]^{-2}$, $[V] = [\text{lunghezza}]^3$ e perciò $[p \cdot \Delta V] =$

[Forza] · [lunghezza]). Passando al limite per $n \to +\infty$ si ottiene l'interpretazione richiesta.

Alla stessa conclusione si può arrivare con un procedimento, spesso usato nelle applicazioni dell'integrale, come il seguente.
In corrispondenza alla variazione di volume da V a $V + dV$, con dV molto piccolo, si può considerare la pressione costante ($= p(V)$) e quindi il lavoro elementare dL effettuato durante la trasformazione è dato da

$$dL = p(V)dV$$

Sommando tutti i contributi da V_A a V_B" e cioè integrando da V_A a V_B si ottiene

$$L = \int_{V_A}^{V_B} dL = \int_{V_A}^{V_B} p(V)dV$$

2.2 Classi di funzioni integrabili

La definizione di integrale che abbiamo dato lascia aperta finora una domanda naturale: come possiamo stabilire, in pratica, se una funzione è integrabile? Valgono in proposito i seguenti teoremi:

Teorema 6.1 *Se $f : [a, b] \to \mathbb{R}$ è continua, allora è integrabile.*

Teorema 6.2 *Se $f : [a, b] \to \mathbb{R}$ è monotona e limitata, allora è integrabile.*

In base a questi teoremi, conosciamo dunque alcune classi significative di funzioni che risultano certamente integrabili. Inoltre, a partire da funzioni integrabili se ne possono costruire altre integrabili con varie procedure. Una di queste è espressa dal prossimo teorema:

Teorema 6.3 *Se $f_1 : [a, b] \to \mathbb{R}$ e $f_2 : [b, c] \to \mathbb{R}$ sono integrabili, allora la funzione*

$$f(x) = \begin{cases} f_1(x) & se\ x \in [a, b) \\ f_2(x) & se\ x \in (b, c] \end{cases}$$

(f definita in qualsiasi modo in b) è integrabile in $[a, c]$.

Applicando più volte questo teorema, il risultato si estende a funzioni ottenute "incollando tra loro" un numero finito di funzioni integrabili su intervalli adiacenti. Ad esempio, una funzione continua e limitata in un intervallo $[a, c]$, ad eccezione di un numero finito di punti di discontinuità a salto, risulta integrabile. Riprenderemo nel seguito questa affermazione, con opportuni esempi (si veda il par.5.6).

I tre teoremi appena enunciati sono piuttosto delicati da dimostrare; proveremo solo il primo di questi (integrabilità delle funzioni continue), ma per non spezzare il discorso, rimandiamo questa dimostrazione all'ultimo paragrafo di questo capitolo.

È bene notare che *non tutte* le funzioni limitate sono integrabili, come mostra il prossimo esempio:

Esempio

2.1 Sia $f : [0,1] \to \mathbb{R}$ così definita:

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

La funzione precedente (nota come "funzione di Dirichlet") vale 1 nei punti razionali e 0 nei punti irrazionali dell'intervallo $[0,1]$. Un attimo di riflessione mostra che questa funzione è discontinua in ogni punto! Proviamo a calcolare S_n. Qualunque sia l'intero n, in ogni intervallino $[x_{j-1}, x_j]$ possiamo scegliere ξ_j razionale o irrazionale, a nostro piacere. Supponiamo di scegliere ξ_j razionale in tutti gli intervallini e per tutti gli n. Allora sarà

$$S_n = \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1}) = \sum_{j=1}^{n} 1 \cdot (x_j - x_{j-1}) = x_n - x_0 = 1.$$

Per questa costruzione di S_n, dunque, risulta $\lim S_n = 1$. Se invece scegliamo, in tutti gli intervallini e per tutti gli n, dei punti ξ_j irrazionali, si avrà

$$S_n = \sum_{j=1}^{n} f(\xi_j)(x_j - x_{j-1}) = \sum_{j=1}^{n} 0 \cdot (x_j - x_{j-1}) = 0$$

e in questo caso $\lim S_n = 0$. Poiché il limite delle successioni di Cauchy-Riemann dipende dalle scelte dei punti ξ_j, la funzione non è integrabile.

3 PROPRIETÀ DELL'INTEGRALE

Direttamente dalla definizione si possono dimostrare le seguenti proprietà dell'integrale.

Teorema 6.4 *Siano f, g integrabili in $[a,b]$. Valgono allora le seguenti proprietà*

1. ***Linearità dell'integrale.*** *Se α, β sono costanti, allora anche la funzione $\alpha f(x) + \beta g(x)$ è integrabile, e vale l'identità*

$$\int_a^b [\alpha f(x) + \beta g(x)]dx = \alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx$$

2. ***Additività dell'integrale rispetto all'intervallo di integrazione.*** *Se $a \le r \le b$ allora f è integrabile anche su $[a,r]$ e $[r,b]$, e vale l'identità:*

$$\int_a^b f(x)dx = \int_a^r f(x)dx + \int_r^b f(x)dx \tag{3.1}$$

Convenzione: se $a < b$, si pone

$$\int_b^a f(x)dx = -\int_a^b f(x)dx$$

allora la (3.1) vale qualunque sia l'ordinamento dei punti a, b, r.

3. ***Positività e monotonia dell'integrale.***

$$f \geq 0 \; in \; [a,b] \implies \int_a^b f(x)dx \geq 0$$

$$f \geq g \; in \; [a,b] \implies \int_a^b f(x)dx \geq \int_a^b g(x)dx \qquad (a < b)$$

In particolare

$$\boxed{\left| \int_a^b f(x)dx \right| \leq \int_a^b |f(x)|\, dx} \tag{3.2}$$

DIMOSTRAZIONE.

1. Dimostriamo la linearità dell'integrale. Sia S_n una somma di Cauchy-Riemann n-esima relativa ad $(\alpha f + \beta g)$:

$$S_n = \frac{b-a}{n} \sum_{j=1}^{n} [\alpha f(\xi_j) + \beta g(\xi_j)] =$$

$$= \alpha \left[\frac{b-a}{n} \sum_{j=1}^{n} f(\xi_j) \right] + \beta \left[\frac{b-a}{n} \sum_{j=1}^{n} g(\xi_j) \right]$$

Facciamo ora tendere n a $+\infty$. Il secondo membro tende a

$$\alpha \int_a^b f(x)dx + \beta \int_a^b g(x)dx;$$

inoltre, questo limite non dipende da come sono scelti i punti ξ_j (perché per ipotesi f e g sono integrabili); dunque anche il primo membro avrà un limite indipendente dalla scelta dei punti ξ_j, e questo limite per definizione è

$$\int_a^b [\alpha f(x) + \beta g(x)]dx.$$

Questo mostra che $\alpha f(x) + \beta g(x)$ è effettivamente integrabile e che vale l'uguaglianza delle due espressioni, ossia la tesi.
Si rifletta sull'idea elementare della dimostrazione precedente: l'integrale è un limite di somme; ma la sommatoria è lineare e l'operazione di limite è lineare: perciò l'integrale è un'operazione lineare.

2. Non dimostriamo l'additività rispetto all'intervallo di integrazione, che è un po' delicata. Ci limitiamo a segnalare che, nel caso di una funzione continua e non negativa, il significato geometrico dell'additività dell'integrale è semplicemente quello di additività dell'area, come mostra la figura 6.6.

3. Sia $f(x) \geq 0$ in $[a,b]$, e consideriamo una sua successione di Cauchy-Riemann:

$$S_n = \sum_{k=1}^{n} f(\xi_k) \left(\frac{b-a}{n} \right) \geq 0$$

Figura 6.6. Additività dell'area.

perché $f(\xi_k) \geq 0$ per ogni k. Passando al limite per $n \to \infty$, il teorema di permanenza del segno per le successioni garantisce che

$$\int_a^b f(x)\,dx = \lim_{n\to\infty} S_n \geq 0.$$

La proprietà di positività appena dimostrata implica quella di monotonia: se ora $f \geq g$ in $[a,b]$, allora $f - g \geq 0$ in $[a,b]$, che per quanto appena dimostrato implica

$$\int_a^b [f(x) - g(x)]\,dx \geq 0.$$

Per la linearità dell'integrale, questo implica a sua volta

$$\int_a^b f(x)\,dx \geq \int_a^b g(x)\,dx$$

che è la proprietà di monotonia. Infine, essendo $f \leq |f|$ e $-f \leq |f|$ in $[a,b]$ si ha:

$$\int_a^b f(x)dx \leq \int_a^b |f(x)|dx \quad \text{e} \quad \int_a^b -f(x)dx \leq \int_a^b |f(x)|dx$$

che insieme equivalgono alla (3.2). ◆

La prossima proprietà non vale per tutte le funzioni integrabili, ma specificamente per quelle *continue*:

Teorema 6.5 (Teorema della media) *Sia $f : [a,b] \to \mathbb{R}$ continua. Allora esiste $c \in [a,b]$ tale che*

$$\frac{1}{b-a}\int_a^b f(x)\,dx = f(c). \tag{3.3}$$

Dimostrazione. Essendo f continua in $[a,b]$, per il Teorema di Weierstrass (cap. 3, par. 4.1), essa è dotata di massimo ($= M$) e minimo ($= m$). Dalla proprietà di monotonia si ha

$$m = \frac{1}{b-a}\int_a^b m\,dx \leq \frac{1}{b-a}\int_a^b f(x)dx \leq \frac{1}{b-a}\int_a^b M\,dx = M$$

Quindi il valore $\frac{1}{b-a}\int_a^b f(x)\,dx$ è compreso tra il minimo ed il massimo di f. Per la proprietà dei valori intermedi delle funzioni continue (ancora cap. 3, par. 4.1) tale valore è uguale a $f(c)$ per qualche $c \in [a,b]$. ◆

Valor medio, valor efficace
Chiamiamo:

$$\frac{1}{b-a}\int_a^b f(x)dx = \text{ valor medio di } f \text{ su } [a,b] = f_M$$

$$\left(\frac{1}{b-a}\int_a^b [f(x)]^2 dx\right)^{1/2} = \text{valor efficace di } f \text{ su } [a,b] = f_E$$

Essendo

$$\frac{1}{b-a}\int_a^b f(x)dx = \lim_{n\to+\infty}\frac{1}{b-a}S_n = \lim_{n\to+\infty}\frac{1}{n}\sum_{j=1}^{n} f(\xi_j)$$

si vede che f_M è una generalizzazione dell'ordinaria media aritmetica tra n numeri.

Il valore efficace f_E interviene, ad esempio, nella teoria dei circuiti elettrici. Sia $I(t) = I\sin\omega t$, $\omega = \frac{2\pi}{T}$, l'intensità di una corrente sinusoidale. Abbiamo, prendendone il valore efficace su un periodo:

$$I_E^2 = \frac{1}{T}\int_0^T I^2(\sin\omega t)^2 dt = \left(\text{essendo } (\sin x)^2 = \frac{1-\cos 2x}{2}\right)$$

$$= \frac{I^2}{T}\int_0^T \left[\frac{1}{2} - \frac{1}{2}\cos 2\omega t\right] dt = \frac{I^2}{T}\left\{\frac{T}{2} - \int_0^T \frac{1}{2}\cos 2\omega t\, dt\right\}$$

Poiché, per simmetria, $\int_0^T \cos 2\omega t\, dt = 0$ (vedi figura 6.7) si ha:

$$I_E^2 = \frac{I^2}{2} \quad \text{e cioè} \quad \boxed{I_E = \frac{I}{\sqrt{2}}}$$

Figura 6.7. $\cos 2\omega t$ è periodico di periodo $\frac{T}{2}$ se $\omega = \frac{2\pi}{T}$.

4 IL TEOREMA FONDAMENTALE DEL CALCOLO INTEGRALE

La definizione di integrale non si presta in generale al suo calcolo effettivo. Uno dei metodi più usati per il calcolo è quello per *variazione di una primitiva*, che qui illustriamo.

Definizione 6.2 Si dice che una funzione G, derivabile in $[a,b]$ è una *primitiva* di f in $[a,b]$ se

$$G'(x) = f(x) \quad \text{per ogni } x \in [a,b].$$

Ad esempio, $G(x) = x^2$ è una primitiva di $f(x) = 2x$ e $G(x) = \sin x$ è una primitiva di $f(x) = \cos x$, su tutto l'asse reale.

Evidentemente se G è una primitiva di f, lo è anche $G + c$ dove c è una costante; d'altra parte se G_1 e G_2 sono due primitive di f in $[a, b]$ allora $G_1' - G_2' = 0$ in $[a, b]$ ossia $(G_1 - G_2)' = 0$ e quindi $G_1 - G_2 =$ costante (Il fatto che, se una funzione ha derivata nulla su un intervallo, allora è costante, è stato dimostrato come conseguenza del teorema di Lagrange, si veda il teorema 4.6 del capitolo 4).

Ne segue che *se si conosce una primitiva G di f, tutte le altre sono della forma $G + c$, $c \in \mathbb{R}$.*

Tutte le funzioni possiedono primitiva? Non tutte; si può mostrare che se una funzione presenta punti di discontinuità a salto in un intervallo $[a, b]$ allora *non* può avere primitiva.

Se però f è continua in $[a, b]$, proveremo in seguito (par. 9) che essa possiede certamente una primitiva (e quindi infinite, per quanto detto precedentemente).
Fatte queste importanti premesse, enunciamo ora il risultato che ci permetterà di calcolare effettivamente gli integrali.

Teorema 6.6 (Teorema fondamentale del calcolo integrale)
Se $f : [a, b] \to \mathbb{R}$ è continua, e G è una sua primitiva su $[a, b]$, allora

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

Per indicare $G(b) - G(a)$ si usa comunemente il simbolo

$$[G(x)]_a^b$$

Ad esempio, si voglia calcolare

$$\int_0^1 x^2 dx$$

Poiché $(\frac{x^3}{3})' = x^2$, $G(x) = \frac{x^3}{3}$ è una primitiva di x^2 in $[0, 1]$, perciò

$$\int_0^1 x^2 dx = \left[\frac{x^3}{3}\right]_0^1 = \frac{1}{3}$$

Si confronti la semplicità di questo calcolo con il procedimento illustrato in apertura del capitolo, per calcolare questo integrale in base alla definizione.

Si esprime talvolta questo teorema dicendo che *integrazione e derivazione sono operazioni inverse*. È questa un'affermazione che non va equivocata: l'integrale è stato definito mediante un opportuno limite di somme, che (apparentemente!) non ha niente a che vedere col concetto di derivata (limite del rapporto incrementale). Tuttavia, grazie al teorema precedente, il calcolo effettivo di un integrale è ricondotto alla ricerca di una primitiva, che è l'operazione inversa della derivazione. (Un sinonimo di *primitiva* talvolta usato è *antiderivata*).

Dimostrazione. Siano $a = x_0, x_1, \ldots, x_n = b$ punti che suddividono l'intervallo $[a, b]$ in n intervallini di ugual ampiezza. Allora, aggiungendo e togliendo $G(x_j)$ per $j = 1, 2, \ldots, n-1$ si

ha:

$$G(b) - G(a) = G(x_n) - G(x_0) = [G(x_n) - G(x_{n-1})] + [G(x_{n-1}) - G(x_{n-2})] + \cdots$$
$$\cdots + [G(x_2) - G(x_1)] + [G(x_1) - G(x_0)] =$$
$$= \sum_{j=1}^{n} [G(x_j) - G(x_{j-1})]$$

Applichiamo ora il teorema di Lagrange alla funzione $G(x)$ su ciascuno degli intervalli $[x_{j-1}, x_j]$. Esiste allora $\xi_j \in (x_{j-1}, x_j)$ tale che

$$G(x_j) - G(x_{j-1}) = (x_j - x_{j-1})G'(\xi_j) = (x_j - x_{j-1})f(\xi_j)$$

perché per ipotesi G è una primitiva di f e perciò $G'(\xi_j) = f(\xi_j)$. Ne segue che

$$G(b) - G(a) = \sum_{j=1}^{n} (x_j - x_{j-1})f(\xi_j) = S_n$$

dove S_n è una somma n-esima di Cauchy-Riemann di f. L'identità scritta vale per ogni n; possiamo allora far tendere n a $+\infty$, trovando

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

Si osservi che questa dimostrazione prova che *una certa* successione di Cauchy-Riemann di f tende a $G(b) - G(a)$. Poiché però sappiamo già che f è integrabile, in quanto continua (per il Teorema 6.1, che proveremo in seguito), questo è sufficiente a concludere che ogni altra successione di Cauchy-Riemann converge allo stesso limite, e quindi vale la formula enunciata. ◆

Nel prossimo paragrafo affronteremo il problema di come si determina una primitiva di $f(x)$, almeno per certe classi di funzioni $f(x)$.

Va tenuto presente, comunque, che:

1. A dispetto del fatto che ogni funzione continua su $[a, b]$ ha una primitiva, anche per funzioni continue abbastanza semplici succede che tale primitiva non sia esprimibile come composizione di funzioni elementari. Ciò significa che in molti casi non si sa scrivere un'espressione analitica della primitiva di $f(x)$ utilizzabile per il calcolo effettivo dell'integrale.
2. Anche quando una primitiva di $f(x)$ non è nota, è possibile calcolare un valore approssimato dell'integrale di una funzione continua su un intervallo, con i metodi dell'analisi numerica, che si basano, in sostanza, sulla definizione di integrale come limite di somme, e non sul teorema fondamentale del calcolo integrale. Ecco quindi un caso in cui la "scomoda" definizione di integrale fornisce il modo migliore per calcolarlo. Alla discussione di uno di questi metodi dedicheremo il successivo par. 7.

5 METODI ELEMENTARI PER LA RICERCA DI UNA PRIMITIVA. CALCOLO DI INTEGRALI INDEFINITI E DEFINITI

5.1 Integrali immediati, per scomposizione, per sostituzione

Come abbiamo visto, se f è una funzione continua e G è una sua primitiva, tutte le primitive di f avranno la forma

$$G(x) + c \qquad (c \in \mathbb{R})$$

L'insieme di tutte le primitive di f prende il nome di *integrale indefinito* di f, e si indica col simbolo

$$\int f(x)\,dx$$

Lo stesso simbolo si usa talvolta per indicare una particolare primitiva di f. Ricordiamo ancora che questo simbolo ha un significato ben diverso da quello di

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

che indica l'integrale di f sull'intervallo $[a, b]$, detto anche *integrale definito*. L'integrale definito è un numero, l'integrale indefinito è un insieme di funzioni (o una qualunque di esse).

Ci occuperemo ora dei *metodi di integrazione*, ossia dei metodi per trovare una primitiva di una funzione data (integrazione indefinita) e quindi per calcolarne il suo integrale definito, mediante il teorema fondamentale del calcolo integrale, almeno per funzioni di tipo abbastanza semplice.

Queste tecniche di integrazione saranno usate in molti argomenti trattati nel secondo volume (equazioni differenziali ordinarie, integrali di linea, integrali doppi e di superficie, serie di Fourier...).

Leggendo la tabella delle derivate delle funzioni elementari "in senso inverso", si ottiene la prima tabella di primitive:

Funzione	Primitiva	Funzione	Primitiva
k	kx	e^x	e^x
x^α	$\frac{x^{\alpha+1}}{\alpha+1}$ $(\alpha \neq -1)$	a^x	$\frac{a^x}{\log a}$ $(a > 0, a \neq 1)$
$\frac{1}{x}$	$\log\lvert x\rvert$	$\frac{1}{1+x^2}$	$\operatorname{arctg} x$
$\sin x$	$-\cos x$	$\frac{1}{\sqrt{1-x^2}}$	$\arcsin x$
$\cos x$	$\sin x$	$\operatorname{Sh} x$	$\operatorname{Ch} x$
$\frac{1}{(\cos x)^2} = 1 + (\operatorname{tg} x)^2$	$\operatorname{tg} x$	$\operatorname{Ch} x$	$\operatorname{Sh} x$
$\frac{1}{(\sin x)^2}$	$-\operatorname{cotg} x$	$\frac{1}{(\operatorname{Ch} x)^2}$	$\operatorname{Th} x$
		$\frac{1}{(\operatorname{Sh} x)^2}$	$-\operatorname{Coth} x$

Nella tabella precedente, quella fornita è *una* delle primitive. Per scrivere l'integrale indefinito, come insieme di *tutte* le primitive, occorre aggiungere sempre una costante arbitraria, ad esempio:

$$\int x^2 dx = \frac{x^3}{3} + c$$

Le regole elementari con cui, a partire da certi integrali, se ne calcolano altri, sono:

Integrazione per scomposizione

$$\int [\alpha f(x) + \beta g(x)]dx = \alpha \int f(x)dx + \beta \int g(x)dx \tag{5.1}$$

(segue dalla linearità della derivata);

Integrazione per sostituzione

Sia G una primitiva di f in un intervallo I, cioè $G'(t) = f(t)$ per ogni $t \in I$. Sia ora $t = \varphi(x)$ una funzione derivabile con continuità su un intervallo $[a, b]$ tale che $\varphi([a, b]) \subset I$. Dal teorema di derivazione delle funzioni composte abbiamo allora:

$$\frac{d}{dt}G(\varphi(x)) = G'(\varphi(x)) \cdot \varphi'(x) = f(\varphi(x))\varphi'(x) \quad \forall x \in [a, b]$$

e cioè:

$$\begin{array}{ccc} G(t) & & \Phi(x) = G(\varphi(x)) \\ \text{primitiva di} & \Longleftrightarrow & \text{primitiva di} \\ f(t) & & f(\varphi(x)) \cdot \varphi'(x). \end{array}$$

Ne segue la formula (di integrazione per sostituzione)

$$\int f(\varphi(x))\,\varphi'(x)\,dx = \int f(t)\,dt \quad (\varphi(x) = t) \tag{5.2}$$

È facile ricordare la (5.2) col seguente procedimento formale: si sostituisce $t = \varphi(x)$ nell'integrale a primo membro, e si calcola $dt = \varphi'(x)\,dx$. Così facendo si ottiene proprio la (5.2) (si tenga presente che questo passaggio è puramente formale: il motivo per cui la (5.2) è vera è stato spiegato in precedenza).

La versione (5.2) per l'integrale definito è la seguente

$$\int_a^b f(\varphi(t))\,\varphi'(t)\,dt = \int_{\varphi(a)}^{\varphi(b)} f(x)\,dx \quad (\varphi(t) = x) \tag{5.3}$$

Di nuovo, è facile ricordare la formula precedente con i passaggi formali già visti: si pone $t = \varphi(x)$ nell'integrale a primo membro, e si calcola $dt = \varphi'(x)\,dx$; inoltre, si modificano gli estremi di integrazione tenendo conto che

$$x \in [a, b] \Longrightarrow t = \varphi(x) \in [\varphi(a), \varphi(b)].$$

dove la scrittura $[\varphi(a), \varphi(b)]$ indica che $\varphi(a), \varphi(b)$ sono, in quest'ordine, gli estremi di integrazione del nuovo integrale, anche se, eventualmente, $\varphi(a) > \varphi(b)$.

Esempi

5.1 $\int (4x^2 + 5x + 1)dx =$ (per la (5.1)) =

$$= 4\int x^2 dx + 5\int x dx + \int 1 dx = \frac{4}{3}x^3 + \frac{5}{2}x^2 + x + c$$

5.2 $$\int_0^1 \frac{1}{e^x + e^{-x}}\, dx$$

Poniamo $e^x = t$ ossia $x = \log t$, $dx = \frac{1}{t}\, dt$.
Si ha inoltre: $x = 0 \Rightarrow t = 1; x = 1 \Rightarrow t = e$; perciò dalla (5.3) abbiamo:

$$\int_0^1 \frac{1}{e^x + e^{-x}}\, dx = \int_1^e \frac{1}{t + \frac{1}{t}} \frac{1}{t}\, dt = \int_1^e \frac{1}{1+t^2}\, dt = [\operatorname{arctg} t]_1^e = \operatorname{arctg} e - \frac{\pi}{4}.$$

5.3 $$\int \operatorname{tg} x dx = \int \frac{\sin x}{\cos x} dx = [\text{per la (5.3): } \cos x = t; -\sin x dx = dt] =$$
$$= \int -\frac{dt}{t} = -\log|t| + c = -\log|\cos x| + c$$

L'esempio precedente è anche un caso particolare della seguente formula:

$$\boxed{\int \frac{f'(x)}{f(x)} dx = \log|f(x)| + c}$$

come si verifica con la sostituzione $f(x) = t \quad f'(x)dx = dt$.

5.4 $$\boxed{\int [f(x)]^\alpha f'(x)dx = \frac{[f(x)]^{\alpha+1}}{\alpha+1} + c \qquad \alpha \neq -1}$$

Basta porre $t = f(x)$, $dt = df(x) = f'(x)dx$; si ha:

$$\int [f(x)]^\alpha f'(x)dx = \int t^\alpha\, dt = \frac{t^{\alpha+1}}{\alpha+1} + c = \frac{[f(x)]^{\alpha+1}}{\alpha+1} + c$$

Ad esempio:

$$\int x^2 \sqrt[3]{2+x^3}\, dx = \frac{1}{3}\int \sqrt[3]{2+x^3} 3x^2 dx = \frac{1}{3}\int (\underbrace{2+x^3}_{f(x)})^{1/3} \underbrace{3x^2}_{f'(x)}\, dx =$$
$$= \frac{1}{3}(2+x^3)^{4/3} \cdot \frac{3}{4} + c = \frac{1}{4}(2+x^3)^{4/3} + c$$

Simmetrie

Facciamo ora un'osservazione sull'*integrale di una funzione simmetrica* (*pari* o *dispari*) *su un intervallo simmetrico* $[-k, k]$. Si osservino le figure 6.8*a* e 6.8*b*.

Pensando l'integrale come area con segno sottesa al grafico di f, si vede subito che se f è una funzione continua su $[-k, k]$ e

$$\boxed{\begin{array}{l} f \text{ è pari, allora } \displaystyle\int_{-k}^{k} f(x)dx = 2\int_0^k f(x)dx \\[2ex] f \text{ è dispari, allora } \displaystyle\int_{-k}^{k} f(x)dx = 0 \end{array}}$$

Figura 6.8.

Ad esempio:

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \cos x\, dx = (\text{integranda pari}) \ = 2\int_0^{\frac{\pi}{2}} \cos x\, dx = 2[\sin x]_0^{\pi/2} = 2$$

$$\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} \sin x\, dx = (\text{integranda dispari}) \ = 0$$

Si osservi anche il prossimo esempio:

$$\int_{-1}^{1} e^{-|x|} dx = 2\int_0^1 e^{-|x|} dx = 2\int_0^1 e^{-x} dx = 2\left[-e^{-x}\right]_0^1 = 2 - \frac{2}{e}.$$

In questo caso la simmetria ha consentito di ridursi a un integrale sull'intervallo $[0,1]$; a sua volta, questo ha consentito di togliere il valore assoluto (per $x \in [0,1]$, $e^{-|x|} = e^{-x}$), permettendoci quindi di trovare la primitiva.

Valori assoluti

L'ultimo esempio suggerisce un'idea generale con cui si può calcolare l'integrale definito di una funzione contenente un valore assoluto. Si consideri il prossimo esempio:

$$\int_0^2 |(x-1)(x+3)|\, dx.$$

Per $x \in [0,2]$, abbiamo

$$|(x-1)(x+3)| = \begin{cases} (x-1)(x+3) & \text{se } x \in [1,2] \\ (1-x)(x+3) & \text{se } x \in [0,1] \end{cases}$$

(si osservi che non ci interessano i casi in cui $x \notin [0,2]$). Quindi, per l'additività dell'integrale abbiamo:

$$\begin{aligned} \int_0^2 |(x-1)(x+3)|\, dx &= \int_0^1 (1-x)(x+3)\, dx + \int_1^2 (x-1)(x+3)\, dx = \\ &= \left[-\frac{x^3}{3} - x^2 + 3x\right]_0^1 + \left[\frac{x^3}{3} + x^2 - 3x\right]_1^2 = \\ &= \left(\frac{5}{3} - 0\right) + \left(\frac{2}{3} + \frac{5}{3}\right) = \frac{10}{3}. \end{aligned}$$

In generale, volendo calcolare l'integrale definito di una funzione contenente uno o più valori assoluti, si spezzerà l'integrale nella somma di integrali su sottointervalli, in modo che in ciascun sottointervallo l'argomento di ciascun valore assoluto abbia segno costante, e quindi si possa esprimere l'integranda senza usare valori assoluti.

5.2 Integrazione delle funzioni razionali

Diamo ora un'idea schematica di come si calcoli in generale la primitiva di una funzione razionale:

$$\int \frac{P_n(x)}{Q_m(x)}\,dx$$

(abbiamo indicato con P_n, Q_m polinomi di grado n, m, rispettivamente).

Anzitutto: *se il grado del numeratore è $\geq$ del grado del denominatore, per prima cosa si esegue la divisione di polinomi*; questo porta a riscrivere l'integranda come somma di un polinomio (che si integra immediatamente), più una funzione razionale con lo stesso denominatore di quella di partenza, e il numeratore di grado inferiore. (Questo non conclude il calcolo, ma è un primo passo che semplifica le cose).

Esempio

5.5

$$\int \frac{x^3+x}{x^2+x+1}\,dx$$

La divisione di polinomi dà:

$$x^3+x=(x^2+x+1)(x-1)+(x+1)$$

quindi:

$$\int \frac{x^3+x}{x^2+x+1}\,dx=\int (x-1)dx+\int \frac{x+1}{x^2+x+1}\,dx$$

In base a questa osservazione, possiamo d'ora in poi supporre che il grado del numeratore sia già minore del grado del denominatore (altrimenti ci si riduce a questo caso mediante divisione).

Se il denominatore è di primo grado il suo integrale si calcola immediatamente, mediante logaritmo:

5.6

$$\int \frac{2}{3x+5}\,dx=\frac{2}{3}\log|3x+5|+c$$

La formula generale è la seguente:

$$\int \frac{dx}{ax+b}=\frac{1}{a}\log|ax+b|+c.$$

Se il denominatore è di secondo grado (e il numeratore di grado ≤ 1), occorre distinguere 3 casi, a seconda del segno del discriminante del denominatore.

A. *Il denominatore ha due radici distinte.* La frazione si scompone in *fratti semplici* e si integra poi mediante somma di logaritmi:

Esempi

5.7
$$\int \frac{x+2}{(x-2)(x+3)}dx$$

Scriviamo:
$$\frac{x+2}{(x-2)(x+3)} = \frac{a}{x-2} + \frac{b}{x+3}$$
per opportuni coefficienti a, b da determinarsi. Per determinarli, si esegue la somma a secondo membro facendo denominatore comune, e si impone che la frazione trovata sia uguale a quella a primo membro. Si trova:
$$\frac{x(a+b)+(3a-2b)}{(x-2)(x+3)} = \frac{x+2}{(x-2)(x+3)}$$
che è vero se e solo se:
$$\begin{cases} a+b=1 \\ 3a-2b=2 \end{cases} \qquad \begin{cases} a=\frac{4}{5} \\ b=\frac{1}{5} \end{cases}$$
Quindi l'integrale di partenza è uguale a:
$$\int (\frac{\frac{4}{5}}{x-2} + \frac{\frac{1}{5}}{x+3})dx = \frac{4}{5}\log|x-2| + \frac{1}{5}\log|x+3| + c$$

5.8
$$\int \frac{dx}{x^2-4} = \int \frac{1}{(x-2)(x+2)}dx = \int \frac{1}{4}[\frac{1}{x-2} - \frac{1}{x+2}]dx =$$
$$= \frac{1}{4}[\log|x-2| - \log|x+2|] + c$$

B. *Il denominatore è un quadrato perfetto.* Mediante sostituzione ci si riconduce alla somma di potenze (con esponente positivo o negativo), che si integra immediatamente:

Esempio

5.9
$$\int \frac{x+1}{(3x+2)^2}dx = [3x+2=t, 3dx=dt, x=\frac{t-2}{3}]$$
$$= \int \frac{\frac{t-2}{3}+1}{t^2}\frac{dt}{3} = \frac{1}{9}\int \frac{t+1}{t^2}dt = \frac{1}{9}\int (\frac{1}{t} + \frac{1}{t^2})dt =$$
$$= \frac{1}{9}(\log|t| - \frac{1}{t}) + c = \frac{1}{9}(\log|3x+2| - \frac{1}{3x+2}) + c$$

C. *Il denominatore non si annulla mai.* Vediamo su un paio di esempi come si procede in generale in questo caso:

Esempi

5.10
$$\int \frac{dx}{x^2+3} = \frac{1}{3}\int \frac{dx}{(x/\sqrt{3})^2+1} = \frac{\sqrt{3}}{3}\operatorname{arctg}\left(\frac{x}{\sqrt{3}}\right) + c$$

In generale:

$$\boxed{\int \frac{1}{x^2+a^2}dx = \frac{1}{a}\operatorname{arctg}\frac{x}{a} + c} \tag{5.4}$$

come si verifica immediatamente calcolando la derivata del secondo membro. Questo esempio illustra il caso più semplice in cui il denominatore non si annulla mai. La situazione più generale è descritta invece nel prossimo esempio.

5.11

$$\int \frac{x}{x^2+2x+4}dx = \frac{1}{2}\int \frac{2x+2}{x^2+2x+4}dx - \int \frac{1}{x^2+2x+4}dx$$

Si osservi com'è costruita l'identità: il primo addendo deve avere a numeratore la derivata del denominatore, in questo caso $2x+2$; d'altro canto il secondo addendo non deve contenere x a numeratore; il "trucco" è perciò scrivere:

$$x = \frac{1}{2}(2x+2) - 1$$

e *non*, ad esempio, $x = (2x+2) - (x+2)$, che sarebbe vero ma inconcludente, in quanto il secondo integrale avrebbe ancora la x a numeratore. Ora l'espressione trovata è uguale a:

$$\frac{1}{2}\log(x^2+2x+4) - \int \frac{1}{(x+1)^2+3}dx = \frac{1}{2}\log(x^2+2x+4) - \frac{1}{\sqrt{3}}\operatorname{arctg}\frac{x+1}{\sqrt{3}} + c$$

Nell'ultimo integrale si è usata la formula vista per la primitiva di $\frac{1}{x^2+a^2}$, più la sostituzione $x+1=t$, $dx=dt$.

In generale, se il denominatore ha discriminante negativo, ossia è irriducibile, si spezza la frazione in due addendi, di cui una a numeratore presenta la derivata del denominatore, e porta ad un integrale logaritmico; l'altra ha numeratore costante, e si riporta all'integrale di una funzione arcotangente. Il procedimento generale è sintetizzato dalla formula seguente:

$$\begin{aligned}\int \frac{x+d}{(x+b)^2+a^2}dx &= \frac{1}{2}\int \frac{2x+2b}{x^2+2bx+b^2+a^2}dx + \int \frac{d-b}{(x+b)^2+a^2}dx = \\ &= \frac{1}{2}\log\left(x^2+2bx+b^2+a^2\right) + \frac{d-b}{a}\operatorname{arctg}\left(\frac{x+b}{a}\right) + c.\end{aligned}$$

Se il denominatore è un polinomio di grado maggiore di due è sempre possibile (teoricamente!) scomporlo in un prodotto di (potenze di) fattori di primo grado, oppure di secondo grado irriducibili. Fatto questo, si scompone la frazione in fratti semplici, a cui si applicano i discorsi precedenti. I prossimi esempi illustrano solo alcuni casi semplici in cui questo accade. Se il grado del denominatore è molto maggiore di due... procurarsi un buon software!!

Esempi

5.12

$$\int \frac{2x^2+3x-1}{x^3+2x^2+x+2}dx$$

Osserviamo il denominatore:

$$x^3+2x^2+x+2 = x^2(x+2)+(x+2) = (x+2)(x^2+1)$$

Scomposto il denominatore, si scompone il quoziente in fratti semplici:

$$\frac{2x^2+3x-1}{(x+2)(x^2+1)} = \frac{a}{x+2} + \frac{bx+c}{x^2+1}$$

con a, b, c da determinarsi (il criterio con cui abbiamo scritto i numeratori incogniti è *che il grado del numeratore sia minore di uno del grado del denominatore*). Mettendo a denominatore comune, si trova:

$$\frac{a}{x+2} + \frac{bx+c}{x^2+1} = \frac{a(x^2+1)+(bx+c)(x+2)}{(x+2)(x^2+1)} = \frac{2x^2+3x-1}{(x+2)(x^2+1)}$$

se e solo se:

$$\begin{cases} a+b=2 \\ 2b+c=3 \\ a+2c=-1 \end{cases} \qquad \text{sistema che risolto dà:} \quad \begin{cases} a=\frac{1}{5} \\ b=\frac{9}{5} \\ c=-\frac{3}{5} \end{cases}$$

Perciò:

$$\begin{aligned} \int \frac{2x^2+3x-1}{x^3+2x^2+x+2}dx &= \frac{1}{5}\int \frac{1}{x+2}dx + \frac{1}{5}\int \frac{9x-3}{x^2+1}dx = \\ &= \frac{1}{5}\log|x+2| + \frac{9}{10}\int \frac{2x}{x^2+1}dx - \frac{3}{5}\int \frac{dx}{x^2+1} = \\ &= \frac{1}{5}\log|x+2| + \frac{9}{10}\log(x^2+1) - \frac{3}{5}\operatorname{arctg} x + c \end{aligned}$$

5.13

$$\int \frac{x+1}{x^2(x+3)}dx$$

Si scompone:

$$\frac{x+1}{x^2(x+3)} = \frac{ax+b}{x^2} + \frac{c}{x+3}$$

e si procede al solito modo. Si noti che

$$\int \frac{ax+b}{x^2}dx = \int (\frac{a}{x} + \frac{b}{x^2})dx$$

che è immediato.

5.14

$$\int \frac{x^2+1}{(x+1)^2(x^2-x+1)}dx$$

Il fattore (x^2-x+1) è irriducibile. Si scompone:

$$\frac{x^2+1}{(x+1)^2(x^2-x+1)} = \frac{ax+b}{(x+1)^2} + \frac{cx+d}{x^2-x+1}$$

e si imposta un sistema di quattro equazioni in quattro incognite che risolto dà la scomposizione in fratti semplici. Si trovano così due integrali di tipi che abbiamo già trattato.

Funzioni razionali di e^x

Dovendo integrare una funzione razionale di e^x, si pone $e^x = t; x = \log t; dx = \frac{dt}{t}$, e ci si riconduce a una funzione razionale di t. Occorre, alla fine, tornare alla variabile x.

Esempio

5.15

$$\int \frac{1}{\mathrm{Ch}x} dx = 2\int \frac{dx}{e^x + e^{-x}} = \left[e^x = t; x = \log t; dx = \frac{dt}{t}\right] =$$
$$= 2\int \frac{1}{t + \frac{1}{t}} \frac{dt}{t} = 2\int \frac{dt}{1+t^2} = 2\,\mathrm{arctg}\, t + c = 2\,\mathrm{arctg}(e^x) + c$$

Esercizi

1. $\displaystyle\int \frac{x^3}{1+x^2} dx$
2. $\displaystyle\int \frac{2x-1}{x^2+2x+4} dx$
3. $\displaystyle\int \frac{x+2}{x^2+4x+3} dx$
4. $\displaystyle\int \frac{x+5}{x^2(x+1)} dx$
5. $\displaystyle\int \frac{1+2x^2}{x^4-1} dx$
6. $\displaystyle\int \frac{1-2x}{(x+1)^2(x^2+1)} dx$
7. $\displaystyle\int_0^1 \frac{1}{x^3+1} dx$
8. $\displaystyle\int \frac{1+e^{-x}}{\mathrm{Ch}x} dx$
9. $\displaystyle\int \frac{x^3+2}{x^2+1} dx$
10. $\displaystyle\int_0^{\sqrt{2}} \frac{x+2}{2x^2+1} dx$
11. $\displaystyle\int \frac{2x+3}{(3x+2)^2} dx$
12. $\displaystyle\int_2^5 \frac{2x+3}{x^2+x-2} dx$
13. $\displaystyle\int \frac{1-x^6}{1-x} dx$
14. $\displaystyle\int \frac{3x+1}{x^2+4x+4} dx$

5.3 Integrazione per parti

Se f e g sono derivabili in $[a, b]$ si ha

$$(fg)' = f'g + fg'$$

ossia

$$fg' = (fg)' - f'g\,.$$

Prendendo l'integrale indefinito di entrambi i membri ed osservando che $\int (fg)'dx = fg$ si trova la formula di *integrazione per parti*

$$\boxed{\int f(x)g'(x)dx = f(x)g(x) - \int f'(x)g(x)dx\,.} \tag{5.5}$$

Usando la notazione differenziale, la (5.5) si può scrivere nella forma seguente (più sintetica)

$$\int \underbrace{f}_{\substack{\text{fattore}\\ \text{finito}}} \underbrace{dg}_{\substack{\text{fattore}\\ \text{differenziale}}} = fg - \int g\,df \tag{5.6}$$

Le (5.5) e (5.6) hanno una versione per l'*integrale definito*:

$$\boxed{\int_a^b f(x)g'(x)\,dx = \big[f(x)g(x)\big]_a^b - \int_a^b f'(x)g(x)dx}$$

Nei casi concreti occorre saper scegliere correttamente il fattore finito e quello differenziale.

Esempi

5.16

$$\int x \sin x\, dx$$

Sappiamo calcolare sia la derivata che la primitiva di entrambe le funzioni $x, \sin x$; tuttavia la scelta migliore è quella di derivare x, il che semplifica l'integrale:

$$\int \underset{\substack{\downarrow \\ f}}{x}\ \underset{\substack{\downarrow \\ g'}}{\sin x}\, dx = -x\cos x - \int(-\cos x)dx = -x\cos x + \sin x + c$$

La scelta inversa avrebbe portato

$$\int \underset{\substack{\downarrow \\ g'}}{x}\ \underset{\substack{\downarrow \\ f}}{\sin x}\, dx = \frac{x^2}{2}\sin x - \int \frac{x^2}{2}\cos x\, dx = \text{più complicato dell'integrale di partenza}$$

Questo semplice esempio contiene un'idea molto generale: dovendo calcolare un integrale

$$\int x^n f(x)\, dx \quad \text{con} \quad f(x) = \cos x, \sin x, e^x, \operatorname{Sh} x, \operatorname{Ch} x$$

si integra per parti, derivando x^n (che si abbassa di grado) e integrando $f(x)$ (che dà una funzione "simile": $\sin x$, $-\cos x$, e^x, $\operatorname{Ch} x$, $\operatorname{Sh} x$, rispettivamente).

Per $n = 1$ questo porta un integrale immediato; per $n > 1$ si itera il procedimento finché la derivata della potenza scompare del tutto. Per la linearità dell'integrale indefinito, il procedimento permette di calcolare un integrale del tipo

$$\int P_n(x)f(x)\, dx$$

con $P_n(x)$ polinomio di grado n, e $f(x)$ come sopra.

5.17

$$\int e^x \sin x\, dx$$

Entrambe le funzioni $e^x, \sin x$ hanno derivata e primitiva altrettanto semplici: la scelta di quale fattore integrare e quale derivare sembra essere indifferente (e lo è). Procediamo ad esempio così:

$$\int \underset{\substack{\downarrow \\ g'}}{e^x}\ \underset{\substack{\downarrow \\ f}}{\sin x}\, dx = e^x \sin x - \int e^x \cos x dx =$$

Applichiamo un'altra integrazione per parti, scegliendo f, g' coerentemente alla prima scelta fatta, ossia $g'(x) = e^x, f(x) = \cos x$:

$$= e^x \sin x - \{e^x \cos x - \int e^x(-\sin x)\, dx\}$$

Se chiamiamo I l'integrale di partenza, abbiamo trovato:

$$I = e^x \sin x - e^x \cos x - I \tag{5.7}$$

Apparentemente, la doppia integrazione per parti non ha semplificato nulla; in realtà, però, l'integrale I si ottiene ora risolvendo l'equazione elementare (5.7):

$$2I = e^x \sin x - e^x \cos x; \implies I = \frac{1}{2}e^x(\sin x - \cos x) + c$$

5.18 Con lo stesso metodo dell'esempio precedente si calcolano le seguenti primitive (si lasciano i calcoli per esercizio):

$$\int (\mathrm{Ch}\, x)^2 dx = \frac{1}{2}(\mathrm{Sh}x\mathrm{Ch}x + x) + c; \quad \int (\mathrm{Sh}\, x)^2 dx = \frac{1}{2}(\mathrm{Sh}\, x\mathrm{Ch}x - x) + c$$

In generale: dovendo calcolare un integrale del tipo:

$$\int f(x)g(x)dx \quad \text{con } f(x) = \begin{cases} e^{ax} \\ \mathrm{Sh}(ax) \\ \mathrm{Ch}(ax) \end{cases} \qquad g(x) = \begin{cases} \cos(bx) \\ \sin(bx) \end{cases}$$

si eseguono due integrazioni per parti successive; nella prima, la scelta della funzione da derivare o da integrare è indifferente; nella seconda, però, la scelta dev'essere coerente alla prima. Dopo due integrazioni per parti si trova un'espressione del tipo $I = h(x) - \left(\frac{b}{a}\right)^2 I$, da cui si ricava I.

5.19
$$\int \log x\, dx$$

Il logaritmo è una funzione elementare di cui ancora non conosciamo la primitiva. Integriamo per parti, ponendo $f(x) = \log x, g'(x) = 1$:

$$\int \log x\, dx = x \log x - \int x \cdot \frac{1}{x} dx = x \log x - \int 1 dx = x \log x - x + c$$

Più in generale, con questo metodo si può calcolare

$$\int \underset{\substack{\downarrow \\ g'}}{x^m} \underset{\substack{\downarrow \\ f}}{(\log x)^n}\, dx = \frac{x^{m+1}}{m+1}(\log x)^n - \int \frac{x^{m+1}}{m+1} n \frac{(\log x)^{n-1}}{x} dx =$$

$$= \frac{x^{m+1}}{m+1}(\log x)^n - \frac{n}{m+1}\int x^m (\log x)^{n-1}\, dx.$$

L'ultimo integrale ottenuto è simile a quello di partenza, con il logaritmo elevato ad $(n-1)$ anziché a n. Questo significa che, iterativamente, con n integrazioni per parti si può calcolare l'integrale. Più in generale, questo metodo consente allora di calcolare l'integrale di funzioni del tipo

$$P_m(x)\, Q_n(\log x)$$

dove P_m, Q_n sono polinomi di grado m, n, rispettivamente.

5.20.
$$\int \operatorname{arctg} x dx$$

L'arcotangente è una funzione elementare di cui ancora non conosciamo la primitiva. Integriamo per parti, ponendo $f(x) = \operatorname{arctg} x, g'(x) = 1$:

$$\int \operatorname{arctg} x dx = x \operatorname{arctg} x - \int \frac{x}{1+x^2} dx =$$
$$= x \operatorname{arctg} x - \frac{1}{2} \log\left(1+x^2\right) + c,$$

dove si è integrata coi metodi noti la funzione razionale $\frac{x}{1+x^2}$.

Più in generale, si può calcolare:

$$\int \underset{g'}{\underset{\downarrow}{x^n}} \underset{f}{\underset{\downarrow}{\operatorname{arctg} x}} dx = \frac{x^{n+1}}{n+1} \operatorname{arctg} x - \int \frac{x^{n+1}}{n+1} \cdot \frac{dx}{1+x^2}.$$

L'ultimo integrale ottenuto è quello di una funzione razionale: se $n \geq 1$ occorrerà per prima cosa eseguire la divisione di polinomi $x^{n+1} : 1 + x^2$ per riscrivere

$$\frac{x^{n+1}}{1+x^2} = P_{n-1}(x) + \frac{ax+b}{1+x^2}$$

con P_{n-1} polinomio di grado $n-1$, che si integra direttamente, mentre la funzione razionale $\frac{ax+b}{1+x^2}$ si integra coi metodi visti, riportandosi a funzioni logaritmo e arcotangente.

Esercizi

15 Si presti attenzione ai prossimi integrali: in quali casi è utile effettuare una sostituzione? (se non è utile, si può comunque sviluppare il cubo e integrare il polinomio termine a termine)

$$\int (3x+1)^3 dx \qquad \int (3x^2+1)^3 dx \qquad \int x(3x^2+1)^3 dx$$

16 $\int (x^{5/6} + 2x^{-2} - 3x^{-1} + 2) dx$

17 $\int_0^{\frac{3\pi}{2}} (\sin x)^2 \, dx$

18 $\int \operatorname{cotg} gx \, dx \qquad \int \operatorname{Th} x \, dx$

19 $\int_0^{\pi/4} \cos 2x \sin x \, dx$

20 $\int \left(\sqrt{x} + \sqrt[3]{x+2} + \frac{1}{x^2}\right) dx$

21 $\int_0^1 (3x^3+1)^2 x^2 dx$

22 $\int (1 + (\sin x)^2)^3 \sin 2x \, dx$

23 $\int_e^{e^2} \frac{1}{x(\log x)^2} dx$

24 $\int (\operatorname{tg} x)^2 \, dx \qquad \int \operatorname{Th}^2 x \, dx$

25 $\int_{-1}^1 \frac{x}{1+x^2} dx$

26 $\int \left(\frac{e^{2x} - e^{-x}}{3}\right) dx$

27 $\int_{-1}^1 e^{-|x|} dx$

28 $\int e^{-x^2} x\,dx$

29 $\int (x^2+5x+4)e^x\,dx$

30 $\int \frac{dx}{x\log x}$

31 $\int_0^1 xe^{3x}dx$

32 $\int \sqrt{2+\sin x}\cos x\,dx$

33 $\int_1^3 x^2\log x\,dx$

34 $\int \frac{1}{x(1+(\log x)^2)}dx$

35 $\int_0^\pi x\cos x\,dx$

36 $\int \frac{1}{x(x+1)}dx$

37 $\int e^x(\sin x)^2\,dx$

38 $\int \frac{1}{(x+1)^2}dx$

39 $\int e^{-x^2}x^3dx$

40 $\int_0^{\frac{\pi}{4}} (\cos x)^2\,dx$

41 $\int (\mathrm{Ch}\,x)^2(\mathrm{Sh}\,x)^2\,dx$

5.4 Integrazione delle funzioni trigonometriche

In questo paragrafo faremo una breve panoramica di metodi per integrare una funzione trigonometrica.

Integrali del tipo

$$\int f(\sin x)\cos x\,dx: \quad \text{si pone} \quad \sin x = t, \cos x\,dx = dt$$

$$\int f(\cos x)\sin x\,dx: \quad \text{si pone} \quad \cos x = t, -\sin x\,dx = dt$$

Esempi

5.21

$$\int_0^{\pi/3} (\sin x)^3(\cos x)^2\,dx$$

Si osserva che:

$$(\sin x)^3(\cos x)^2 = \sin x\cdot(\sin x)^2\cdot(\cos x)^2 = \sin x(1-(\cos x)^2)(\cos x)^2 = f(\cos x)\sin x$$

$$\int_0^{\pi/3} \sin x(1-(\cos x)^2)(\cos x)^2\,dx = \left[\cos x = t;\ -\sin x\,dx = dt; t\in\left(1,\frac{1}{2}\right)\right] =$$

$$= \int_1^{1/2} -(1-t^2)t^2dt = \int_{1/2}^1 (t^2-t^4)dt =$$

$$= [\frac{t^3}{3}-\frac{t^5}{5}]_{1/2}^1 = \frac{1}{3}-\frac{1}{5}-\frac{1}{24}+\frac{1}{160}$$

In generale: dovendo calcolare

$$\int (\sin x)^n(\cos x)^m\,dx$$

con *almeno uno degli esponenti dispari*, è sempre possibile riscrivere l'integranda in una delle forme $f(\sin x)\cos x$ o $f(\cos x)\sin x$, sfruttando la relazione $(\sin x)^2+(\cos x)^2=1$. A questo punto l'integrale si calcola con la sostituzione $\sin x = t$, o $\cos x = t$, rispettivamente.

Se invece entrambi gli esponenti sono pari, si possono usare le formule trigonometriche per l'abbassamento del grado:

$$\boxed{(\cos x)^2 = \frac{1}{2}(1+\cos 2x) \qquad (\sin x)^2 = \frac{1}{2}(1-\cos 2x)}$$

5.22
$$\int (\sin x)^2 (\cos x)^4 \, dx$$

Si verifichi che dalle formule precedenti segue:

$$\int (\sin x)^2 (\cos x)^4 \, dx = \frac{1}{8}\int (-(\cos 2x)^3 - (\cos 2x)^2 + \cos 2x + 1) dx$$

Ora: gli addendi $\cos 2x + 1$ hanno integrale immediato; $(\cos 2x)^2$ si integra come visto in precedenza (attenzione alla presenza di $2x$ anziché x); $(\cos 2x)^3$ rientra nel caso in cui uno degli esponenti è dispari: si riscrive $(\cos 2x)^3 = \cos 2x(1-(\sin 2x)^2)$, e si integra ponendo $\sin 2x = t$. Si trova (verificare i calcoli per esercizio):

$$\int (\sin x)^2 (\cos x)^4 \, dx = \frac{12x + 3\sin 2x - 3\sin 4x - \sin 6x}{192} + c$$

(o una delle molte espressioni trigonometriche equivalenti a questa!)

5.23 Calcoliamo, in particolare, due integrali notevoli di funzioni trigonometriche:

$$\int (\sin x)^2 \, dx, \int (\cos x)^2 \, dx$$

Sfruttando le le formule trigonometriche per l'abbassamento del grado (v. sopra) si ha:

$$\boxed{\begin{aligned} \int (\sin x)^2 \, dx &= \frac{1}{2}\int (1-\cos 2x)\, dx = \frac{1}{2}\left(x - \frac{1}{2}\sin 2x\right) + c = \frac{x - \sin x \cos x}{2} + c \\ \int (\cos x)^2 \, dx &= \frac{1}{2}\int (1+\cos 2x)\, dx = \frac{1}{2}\left(x + \frac{1}{2}\sin 2x\right) + c = \frac{x + \sin x \cos x}{2} + c \end{aligned}}$$

5.24 Sfruttando il risultato dell'esempio precedente si possono calcolare, ad esempio, i seguenti integrali definiti:

$$\boxed{\int_0^{k\frac{\pi}{2}} (\sin x)^2 dx = \int_0^{k\frac{\pi}{2}} (\cos x)^2 dx = k\frac{\pi}{4} \quad \text{con} \quad k \in \mathbb{Z}} \tag{5.8}$$

La formula precedente, che esprime due *integrali notevoli* che si presentano spesso, è facile da ricordare: l'integrale di $(\cos x)^2$, $(\sin x)^2$ su un intervallo del tipo $[0, k\frac{\pi}{2}]$ vale la metà dell'area del rettangolo di base $[0, k\frac{\pi}{2}]$ e altezza 1. Questa "coincidenza" ha un significato geometrico; si osservi la figura 6.9:

Figura 6.9.

Le due regioni disegnate hanno la stessa area; quella inferiore è l'integrale di $(\sin x)^2$ su $[0, \frac{\pi}{2}]$, quella superiore è l'integrale di $(\cos x)^2$, perché $(\cos x)^2 = 1 - (\sin x)^2$; ognuno dei due integrali vale dunque $\frac{1}{2} \cdot \frac{\pi}{2}$. Per le simmetrie e periodicità dei grafici di $(\cos x)^2, (\sin x)^2$, ne segue la formula generale.

Integrali del tipo

$$\int \cos \alpha x \sin \beta x \, dx, \int \cos \alpha x \cos \beta x \, dx, \int \sin \alpha x \sin \beta x \, dx$$

Si possono utilizzare le formule di prostaferesi (v. appendice A), che riconducono alla somma di due integrali immediati.

Esempio

5.25 Applicando il metodo visto, si verifica che:

$$\int_0^{2\pi} \cos nx \sin mx \, dx = 0 \text{ per ogni coppia di interi } n, m$$

$$\int_0^{2\pi} \sin nx \sin mx \, dx = 0 \text{ per ogni coppia di interi } n, m, \ n \neq m$$

$$\int_0^{2\pi} \cos nx \cos mx \, dx = 0 \text{ per ogni coppia di interi } n, m, \ n \neq m$$

Questi integrali sono importanti nello studio delle serie di Fourier (v. secondo volume), assieme ai seguenti (che non si calcolano con le formule di prostaferesi, ma sostituendo $nx = t$ e sfruttando la (5.8)):

$$\int_0^{2\pi} (\cos nx)^2 \, dx = \int_0^{2\pi} (\sin nx)^2 \, dx = \pi \quad \text{per ogni intero positivo } n$$

Funzioni razionali di $\sin x$, $\cos x$

L'integrale di una funzione razionale di $\sin x$, $\cos x$, può sempre essere ricondotto all'integrale di una funzione razionale, mediante la seguente sostituzione:

$$t = \operatorname{tg} \frac{x}{2} \qquad x = 2 \operatorname{arctg} t \qquad dx = \frac{2}{1+t^2} dt$$

La sostituzione è utile in forza delle seguenti identità trigonometriche:

$$\begin{cases} \cos x = \dfrac{1-t^2}{1+t^2} \\ \sin x = \dfrac{2t}{1+t^2} \end{cases} \qquad \text{se} \quad t = \operatorname{tg} \frac{x}{2}$$

Questo metodo porta, in genere, a calcoli laboriosi, e va utilizzato perciò solo quando non sembra esservi una via più semplice o un buon software a portata di mano.

Esempio

5.26

$$\int \frac{\sin x - 5\cos x}{3+\sin x}\,dx$$

L'integrale si scompone nella somma di due:

$$I = \int \frac{\sin x}{3+\sin x}dx - 5\int \frac{\cos x}{3+\sin x}dx = I_1 - 5I_2$$

Il secondo è immediato:

$$I_2 = \log(3+\sin x) + c$$

Il primo non è immediato, ma può essere trattato con la sostituzione appena vista. Può essere utile un'ultima semplificazione preliminare:

$$I_1 = \int \frac{\sin x}{3+\sin x}dx = \int \left(1 - \frac{3}{3+\sin x}\right)dx = x - 3\int \frac{1}{3+\sin x}dx = x - 3I_3$$

Calcoliamo ora

$$I_3 = \left[t = \operatorname{tg}\frac{x}{2}; \sin x = \frac{2t}{1+t^2};\ dx = \frac{2dt}{1+t^2}\right] =$$
$$= \int \frac{1}{3+\frac{2t}{1+t^2}}\frac{2dt}{1+t^2} = 2\int \frac{1}{3t^2+2t+3}dt$$

che è l'integrale di una funzione razionale di t. Scritta la primitiva di questa, occorre ritornare alla variabile x. I calcoli (noiosi ma ormai di routine) danno:

$$I_3 = \frac{1}{\sqrt{2}}\operatorname{arctg}\left(\frac{3}{2\sqrt{2}}\left(\operatorname{tg}\frac{x}{2} + \frac{1}{3}\right)\right) + c$$

e in definitiva si ha:

$$I = I_1 - 5I_2 = x - 3I_3 - 5\log(3+\sin x) =$$
$$= x - \frac{3}{\sqrt{2}}\operatorname{arctg}\left(\frac{3}{2\sqrt{2}}\left(\operatorname{tg}\frac{x}{2} + \frac{1}{3}\right)\right) - 5\log(3+\sin x) + c$$

Esercizi

42 $\displaystyle\int (\sin x)^3\,dx$

43 $\displaystyle\int (\sin x)^4\,dx$

45 $\displaystyle\int_0^{\pi/8} (\sin x)^2(\cos x)^2\,dx$

65 $\displaystyle\int_0^{\pi/2} \frac{\sin x + 2\cos x}{2+\sin x}\,dx$

44 $\displaystyle\int (\sin x)^2 \sin 2x\,dx$

46 $\displaystyle\int \sin 2x \cos 3x\,dx$

5.5 Integrazione delle funzioni irrazionali

Vediamo alcuni metodi di integrazioni applicabili a funzioni irrazionali di tipo molto semplice.

Sostituzioni tipiche per trattare funzioni razionali di x e una delle seguenti funzioni

$$\sqrt{a^2 - x^2};\ \sqrt{a^2 + x^2};\ \sqrt{x^2 - a^2}$$

Se l'integranda è una funzione razionale di x e di una sola delle precedenti radici quadrate, si effettua una sostituzione standard:

A. $\sqrt{a^2 - x^2}$. Si pone: $x = a \sin t, dx = a \cos t dt$;

$$\sqrt{a^2 - x^2} = \sqrt{a^2(1 - (\sin t)^2)} = |a \cos t|$$

(Se si sta calcolando un integrale definito, si sa dove varia x, e quindi t; perciò si può togliere il modulo mettendo il segno opportuno). L'integrale è quindi ricondotto a quello di una funzione razionale di $\sin t, \cos t$. A sua volta, questo è sempre riconducibile all'integrale di una funzione razionale, mediante la sostituzione $\operatorname{tg} \frac{t}{2} = u$ (o un'altra via più semplice, se possibile).

B. $\sqrt{a^2 + x^2}$. Si pone: $x = a \operatorname{Sh} t$, $dx = a \operatorname{Ch} t\, dt$;

$$\sqrt{a^2 + x^2} = \sqrt{a^2(1 + (\operatorname{Sh} t)^2)} = a\operatorname{Ch} t$$

(Si osservi che $\operatorname{Ch} t$ è sempre positivo). L'integrale è quindi ricondotto a quello di una funzione razionale di $\operatorname{Sh} t$, $\operatorname{Ch} t$. A sua volta, questo è sempre riconducibile all'integrale di una funzione razionale, mediante la sostituzione $e^t = u$; $t = \log u$; $dt = \frac{du}{u}$ (oppure, l'integrale si calcola per un'altra via più semplice).

C. $\sqrt{x^2 - a^2}$. Si pone: $x = a \operatorname{Ch} t$, $dx = a \operatorname{Sh} t\, dt$;

$$\sqrt{x^2 - a^2} = \sqrt{a^2((\operatorname{Ch} t)^2 - 1)} = |a\operatorname{Sh} t|$$

(Se si sta calcolando un integrale definito, si sa dove varia x, e quindi t; perciò si può togliere il modulo mettendo il segno opportuno). L'integrale è quindi ricondotto a quello di una funzione razionale di $\operatorname{Sh} t$, $\operatorname{Ch} t$ (v. caso B).

Osserviamo che quando si effettuano le sostituzioni B e C, nel ritornare alla variabile x occorre usare le *funzioni iperboliche inverse*:

$$x = a\operatorname{Ch} t \Rightarrow t = \operatorname{SettCh}\left(\frac{x}{a}\right) = \log\left(\frac{x}{a} + \sqrt{\frac{x^2}{a^2} - 1}\right)$$

$$x = a\operatorname{Sh} t \Rightarrow t = \operatorname{SettSh}\left(\frac{x}{a}\right) = \log\left(\frac{x}{a} + \sqrt{\frac{x^2}{a^2} + 1}\right)$$

A questo proposito, è utile anche notare che:

$$\operatorname{Sh}(\operatorname{SettCh} \alpha) = \sqrt{\alpha^2 - 1} \qquad \operatorname{Ch}(\operatorname{SettSh} \alpha) = \sqrt{\alpha^2 + 1}$$

Esempi

5.27 Calcoliamo l'area dell'ellisse:

$$\frac{x^2}{a^2}+\frac{y^2}{b^2}=1$$

Per simmetria, quest'area è quattro volte quella compresa nel primo quadrante, che può calcolarsi come integrale su $[0,a]$ della funzione:

$$y=b\sqrt{1-\frac{x^2}{a^2}}$$

$$\text{Area ellisse}=4\int_0^a b\sqrt{1-\frac{x^2}{a^2}}dx=\left[x=a\sin t;\, dx=a\cos t\,dt;\; t=\arcsin\frac{x}{a}\in\left(0,\frac{\pi}{2}\right)\right]$$

$$=4ab\int_0^{\frac{\pi}{2}}(\cos t)^2\,dt=4ab\cdot\frac{\pi}{4}=\pi ab$$

risultato che generalizza in modo naturale la formula per l'area del cerchio, πr^2.

5.28 $$\int_{\sqrt{2}}^2\frac{1}{x^2\sqrt{x^2-1}}\,dx=$$

$$[x=\mathrm{Ch}t;\, dx=\mathrm{Sh}t\,dt;\, t=\mathrm{SettCh}x\in(\mathrm{SettCh}\sqrt{2},\mathrm{SettCh}2)]$$

$$=\int_{\mathrm{SettCh}\sqrt{2}}^{\mathrm{SettCh}2}\frac{\mathrm{Sh}t}{(\mathrm{Ch}\,t)^2\mathrm{Sh}t}dt=\int_{\mathrm{SettCh}\sqrt{2}}^{\mathrm{SettCh}2}\frac{1}{(\mathrm{Ch}\,t)^2}\,dt=\left[\frac{\mathrm{Sh}t}{\mathrm{Ch}t}\right]_{\mathrm{SettCh}\sqrt{2}}^{\mathrm{SettCh}2}=\frac{\sqrt{3}}{2}-\frac{1}{\sqrt{2}}$$

Integrale di una funzione razionale di $x, x^{n_1/m_1}, x^{n_2/m_2},\ldots$
Si pone $x=t^n$ con n =minimo comune multiplo di $m_1, m_2,\ldots$. Si ha quindi $dx=nt^{n-1}dt$, e si ottiene una funzione razionale di t. Integrata questa, si torna alla variabile x.

Esempio

5.29
$$\int\frac{\sqrt{x}}{2x^{1/3}+3}\,\frac{dx}{x}$$

La funzione integranda dipende da due diverse potenze a esponente razionale di x: $x^{1/2}, x^{1/3}$. Se poniamo:

$$x=t^6,\; dx=6t^5dt$$

otteniamo una funzione razionale di t, perché $x^{1/2}=t^3, x^{1/3}=t^2$:

$$\int\frac{\sqrt{x}}{2x^{1/3}+3}\,\frac{dx}{x}=\int\frac{t^3}{2t^2+3}\cdot\frac{6t^5}{t^6}dt=$$

$$=\int\frac{6t^2}{2t^2+3}dt=\int(3-\frac{9}{2t^2+3})dt=$$

$$=3t-3\sqrt{\frac{3}{2}}\,\mathrm{arctg}\sqrt{\frac{2}{3}}t+c=3x^{1/6}-3\sqrt{\frac{3}{2}}\,\mathrm{arctg}(\sqrt{\frac{2}{3}}x^{1/6})+c$$

5.6 Integrazione di funzioni discontinue

Se una funzione $f : [a, b] \to \mathbb{R}$ è limitata in $[a, b]$ e continua salvo un numero finito di punti di discontinuità a salto nei punti $r_1, \ldots, r_k$, la funzione f è continua (o meglio prolungabile con continuità fino agli estremi) in ogni intervallino $[r_{i-1}, r_i]$; pertanto è integrabile su ogni intervallino $[r_{i-1}, r_i]$ e dunque (per il Teorema 6.3 visto nel par. 2.2) anche in tutto $[a, b]$. Inoltre, per la proprietà di additività dell'integrale, si può scrivere:

$$\int_a^b f(x)dx = \int_a^{r_1} f(x)dx + \int_{r_1}^{r_2} f(x)dx + \ldots + \int_{r_k}^b f(x)dx \tag{5.9}$$

Questo è utile perché ciascuno degli integrali scritti a secondo membro è ora l'integrale di una funzione continua, di cui possiamo cercare la primitiva (o calcolare il valore numerico approssimato dell'integrale). In ogni caso, il problema viene ricondotto a un numero finito di problemi più semplici.

Figura 6.10. Integrazione di una funzione con due discontinuità a salto.

Si noti che, nel ragionamento precedente, ciò che conta è che su ciascun intervallino (r_k, r_{k+1}) la funzione sia continua, e i limiti ai due estremi esistano finiti; non ha importanza, invece, il valore effettivo di f nei due estremi. In altre parole, alterare il valore della funzione integranda in un numero finito di punti è del tutto irrilevante ai fini del calcolo dell'integrale. (Si pensi al significato geometrico: se un rettangolo viene alterato togliendogli un segmento e sostituendolo con un altro di diversa lunghezza, la sua area non cambia).

Esempio

5.30 Le seguenti funzioni sono integrabili (lo studente si renda conto in dettaglio del motivo):

$$\frac{\sin x}{x}, \text{ su } [-1, 1] \qquad \frac{x}{|x|}, \text{ su } [-1, 2]$$

$$\operatorname{arctg} \frac{1}{x}, \text{ su } [-2, 3] \qquad e^{-1/x}, \text{ su } [0, 1] \text{ (ma non su } [-1, 0]!)$$

Esercizi

47 $\displaystyle\int \frac{x^3}{\sqrt{2 - x^2}}\, dx$

48 $\displaystyle\int_0^1 \frac{x^2}{\sqrt{1 + x^2}} dx$

49 $\displaystyle\int \frac{\sqrt{3+x^2}}{x^2}\,dx$

50 $\displaystyle\int_2^3 \frac{\sqrt{x^2-2}}{x^2}\,dx$

51 $\displaystyle\int \frac{x^2}{\sqrt{x^2-1}}\,dx$

52 $\displaystyle\int_1^2 \frac{\sqrt{x^2-2}}{x}\,dx$

53 $\displaystyle\int \frac{1-x}{\sqrt{1+x}}\,dx$

54 $\displaystyle\int \frac{\sqrt{x}+\sqrt[3]{x}+1}{\sqrt[4]{x}}\,dx$

55 $\displaystyle\int_0^1 (1-x^2)^{3/2}\,dx$

56 $\displaystyle\int \frac{\sqrt{x}-\sqrt[3]{x}}{x(1+\sqrt[4]{x})}\,dx$

57 $\displaystyle\int x\sqrt{4-x^2}\,dx$

58 $\displaystyle\int_0^{\frac{\pi}{2}} \sqrt{1+(\sin x)^2}\cos x\,dx$

I prossimi esercizi mostrano esempi di situazioni in cui, con opportune integrazioni per parti, ci si riconduce all'integrale di una funzione razionale o una funzione irrazionale dei tipi trattati in precedenza.

59 $\displaystyle\int x\arcsin x\,dx$

60 $\displaystyle\int_0^1 \operatorname{arctg} x\,dx$

61 $\displaystyle\int_1^2 x(\log x)^2\,dx$

62 $\displaystyle\int x\operatorname{arctg}(x+1)\,dx$

Esercizi vari

63 a) Verificare che

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

è una somma di Cauchy-Riemann per la funzione $f(x) = \frac{1}{1+x^2}$, relativa all'intervallo $[0,1]$.
b) Calcolare

$$\lim_{n\to+\infty} S_n$$

64 Sia $f(t) = a_0 + a_1\cos\omega t + b_1\sin\omega t$.
Calcolare il valor medio e il valore efficace di f. Generalizzare a

$$f(t) = a_0 + \sum_{k=1}^{n}\{a_k\cos k\,\omega t + b_k\sin k\,\omega t\}$$

65 Calcolare i seguenti integrali indefiniti:

a) $\displaystyle\int \frac{x}{x^2-4}\,dx$ *b)* $\displaystyle\int e^{-x}(x^2+x)\,dx$ *c)* $\displaystyle\int e^x\sin 2x\,dx$

d) $\displaystyle\int \frac{1}{\sqrt{x}(1+x)}\,dx$ *e)* $\displaystyle\int \sqrt{2x+1}\,\frac{1}{x}\,dx$

6.6 Calcolare i seguenti integrali definiti:

$$a)\ \int_0^{\pi} x \sin 2x\, dx \qquad b)\ \int_1^2 \frac{1}{x(\log x + 1)} dx$$

$$c)\ \int_1^2 \frac{\sqrt{x+1}}{x} dx \qquad d)\ \int_0^9 \sqrt{1+\sqrt{x}}\, dx \qquad (\text{porre } \sqrt{1+\sqrt{x}} = t)$$

6.7 Calcolare per ogni n intero positivo

$$a_n = \int_0^n \frac{x-1}{(x+1)^3} dx$$

Discutere il carattere della serie

$$\sum_{n=1}^{\infty} a_n$$

6 ALCUNE APPLICAZIONI FISICHE E GEOMETRICHE

Oltre a quelle già incontrate nel par. 2.1, l'integrale ha numerose altre importanti applicazioni.

Energia potenziale di una forza elastica di richiamo. Lavoro di una forza

Una particella di massa M è soggetta ad una forza elastica di richiamo, in direzione dell'asse x, cioè ad una forza direttamente proporzionale allo spostamento da un punto fisso determinato (posizione a riposo) con verso tale da ridurre lo spostamento. Il caso tipico è quello prodotto da una molla soggetta a spostamenti sufficientemente piccoli dalla posizione a riposo.

Se si scelgono le coordinate sull'asse in modo che la posizione di riposo per la particella corrisponda ad $x = 0$ la forza avrà un'intensità data dalla formula

$$F = -kx \tag{6.1}$$

dove $k > 0$ è una costante (detta costante elastica) che dipende dalla molla. La (6.1) si chiama *legge di Hooke*. Per calcolare l'energia potenziale associata alla forza F, calcoliamo il lavoro L che una forza applicata, F_{app}, deve compiere per spostare la massa M, inizialmente a riposo, in un punto di ascissa x_f. L'energia potenziale sarà data da $E_{\text{pot}} = -L$. In base alla (6.1), la forza applicata è funzione della posizione ed è in ogni punto uguale ed opposta a F:

$$F_{\text{app}} = -F = kx$$

Figura 6.11. a) Molla a riposo. b) Lo spostamento x provoca una forza $F = -kx$.

Il lavoro elementare dL, relativo a uno spostamento da x a $x+dx$ ($|dx|$ molto piccolo) nel quale si può considerare $F_{\text{app}} = \text{costante} = kx$, è dato da $dL = kx\,dx$. Integrando tra 0 e x_f si ottiene

$$L = \int_0^{x_f} kx\,dx = k\left[\frac{1}{2}x^2\right]_0^{x_f} = \frac{1}{2}kx_f^2$$

e infine $E_{\text{pot}} = -\frac{1}{2}kx_f^2$.

Più in generale, se una forza agisce lungo una retta, che scegliamo come asse x, spostando un punto materiale da a a b, il lavoro della forza $F(x)$ è dato per definizione dall'integrale:

$$L = \int_a^b F(x)\,dx$$

Naturalmente, da un punto di vista fisico, è molto restrittivo assumere che il moto si svolga lungo una retta. Questa è una delle motivazioni che porteranno a sviluppare ulteriormente la teoria dell'integrazione, definendo gli *integrali di linea* (si veda il secondo volume), mediante i quali si darà senso al concetto di lavoro lungo una curva qualsiasi dello spazio.

Lunghezza di un grafico

Sia $y = f(x)$ una funzione continua con la sua derivata prima $f'(x)$, in un intervallo $[a,b]$. Vogliamo trovare una formula per la lunghezza del grafico di f. A tale scopo suddividiamo l'intervallo $[a,b]$ in n parti uguali, mediante i "soliti" punti $x_0 = a, x_1, x_2, \ldots, x_n = b$; consideriamo ora la linea spezzata che ha per estremi i punti $P_i \equiv (x_i, f(x_i))$, che stanno sul grafico di f.

Figura 6.12.

In base alla nostra idea intuitiva di lunghezza di una curva, la lunghezza ℓ_n di tale spezzata dovrebbe essere una *approssimazione per difetto* della lunghezza del grafico. Se esiste finito il limite di ℓ_n per $n \to \infty$, sarà naturale assumere questo limite come definizione della lunghezza del grafico. Vediamo ora di mettere in pratica questa idea.

Anzitutto, è possibile calcolare ℓ_n al modo seguente:

$$\ell_n = \sum_{i=1}^{n} |P_i - P_{i-1}| = \sum_{i=1}^{n} \sqrt{(x_i - x_{i-1})^2 + (f(x_i) - f(x_{i-1}))^2}.$$

D'altro canto, per il teorema di Lagrange applicato ad f su ciascun intervallino $[x_{i-1}, x_i]$ si ha che esiste $\xi_i \in [x_{i-1}, x_i]$ per cui risulta

$$f(x_i) - f(x_{i-1}) = f'(\xi_i)(x_i - x_{i-1})$$

e quindi

$$\ell_n = \sum_{i=1}^{n} \sqrt{(x_i - x_{i-1})^2 + (f'(\xi_i)(x_i - x_{i-1}))^2} =$$
$$= \sum_{i=1}^{n} (x_i - x_{i-1}) \sqrt{1 + f'(\xi_i)^2}.$$

Si riconosce quindi che la lunghezza della n-esima poligonale che approssima il grafico non è altro che una somma di Cauchy-Riemann n-esima della funzione (continua, nelle nostre ipotesi) $\sqrt{1 + f'(x)^2}$. perciò si ha:

$$\text{lunghezza del grafico} = \lim_{n\to\infty} \ell_n = \int_a^b \sqrt{1 + f'(x)^2}\,dx.$$

Ecco come il calcolo infinitesimale permette di risolvere un altro problema "geometrico" che con metodi elementari non è affrontabile. (Si confronti con quanto discusso nel par. 1).

Esempi

6.1 Calcoliamo la lunghezza dell'arco di curva $y = x^{3/2}$ per $x \in [0, a]$. Essendo $y' = \frac{3}{2}\sqrt{x}$, si ha:

$$L = \int_0^a \sqrt{1 + \frac{9}{4}x}\,dx = \frac{8}{27}[(1 + \frac{9}{4}x)^{3/2}]_0^a = \frac{8}{27}\left[\left(1 + \frac{9}{4}a\right)^{3/2} - 1\right]$$

6.2 Calcoliamo la lunghezza dell'arco di parabola $y = x^2$ per $x \in [0, a]$. Essendo $y' = 2x$, si ha:

$$L = \int_0^a \sqrt{1 + 4x^2}dx = \left[2x = \mathrm{Sh}t; dx = \frac{1}{2}\mathrm{Ch}t dt; t \in (0, \mathrm{SettSh}2a)\right]$$

$$= \int_0^{\mathrm{SettSh}2a} \tfrac{1}{2}(\mathrm{Ch}\,t)^2 dt = \text{(ricordando la primitiva di } (\mathrm{Ch}\,t)^2\text{, calcolata nell'esempio 5.18)}$$

$$= \frac{1}{2}\left[\frac{1}{2}(\mathrm{Sh}t\mathrm{Ch}t + t)\right]_0^{\mathrm{SettSh}2a} = \frac{1}{4}(2a\sqrt{1 + 4a^2} + \log(2a + \sqrt{1 + 4a^2}))$$

Come si vede, anche in un caso elementare come quello della parabola, l'espressione trovata per la lunghezza dell'arco ha una certa complessità. Questo è un fatto abbastanza generale. Lo studente provi, a titolo d'esempio, a impostare il calcolo della lunghezza dell'arco di ellisse:

$$y = b\sqrt{1 - \frac{x^2}{a^2}} \qquad x \in [0, a]$$

(cioè un quarto dell'ellisse). Troverà che, salvo il caso banale $a = b$ (circonferenza!) non è possibile determinare una primitiva della funzione integranda con i metodi noti.

OSSERVAZIONE (GRAFICI DI LUNGHEZZA INFINITA) Notiamo ora che per ogni funzione $f : [a, b] \to \mathbb{R}$, continua, è possibile costruire la successione ℓ_n delle lunghezze delle poligonali di vertici $(x_i, f(x_i))$. Abbiamo inoltre dimostrato che, se f è derivabile con continuità, esiste certamente $\lim_{n\to\infty} \ell_n$, che è calcolabile mediante un certo integrale. Se f è continua ma non è derivabile con continuità, può accadere però che il $\lim_{n\to\infty} \ell_n$ non esista finito: in questo caso la lunghezza della curva risulterà infinita.

Esempio

6.3 La funzione

$$f(x) = \begin{cases} x \sin\left(\dfrac{1}{x}\right) & \text{per } x \in (0, 1] \\ 0 & \text{per } x = 0 \end{cases}$$

è continua in $[0, 1]$ (ma la sua derivata non è continua in 0 e non è limitata in $[0, 1]$). Si può dimostrare che il suo grafico ha lunghezza infinita:

Figura 6.13.

Volume di un solido di rotazione attorno a un asse

Consideriamo un solido ottenuto per rotazione attorno all'asse x del trapezioide individuato dalla funzione continua $y = f(x)$, $a \le x \le b$ (cfr. fig. 6.14). Il volume si ottiene sommando i contributi dV dei volumi dei cilindretti aventi altezza dx e raggio

Figura 6.14.

di base $f(x)$. Essendo $dV = \pi f(x)^2 dx$, integrando tra a e b si trova la formula per il volume:

$$\boxed{\text{vol} = \pi \int_a^b f(x)^2 dx} \tag{6.2}$$

Esempio

6.4 Calcoliamo il volume dell'ellissoide di rotazione, cioè del solido che si ottiene facendo ruotare attorno all'asse x la parte superiore dell'ellisse

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} = 1$$

(L'equazione cartesiana della superficie dell'ellissoide trovato è

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{b^2} = 1$$

ovvero un ellissoide con due semiassi su tre uguali tra loro). In questo caso

$$f(x) = b\sqrt{1 - \frac{x^2}{a^2}}, \quad \text{con } x \in [-a, a]$$

e il volume del solido è

$$\begin{aligned} \text{vol} &= \pi \int_{-a}^{a} (b\sqrt{1 - \frac{x^2}{a^2}})^2 dx = 2\pi b^2 \int_0^a (1 - \frac{x^2}{a^2}) dx = \\ &= 2\pi b^2 [x - \frac{x^3}{3a^2}]_0^a = 2\pi b^2 \cdot \frac{2}{3} a = \frac{4}{3}\pi a b^2 \end{aligned}$$

formula che generalizza in modo naturale quella del volume della sfera, $\frac{4}{3}\pi r^3$, e che a sua volta suggerisce che il volume di un generico ellissoide (non di rotazione), di semiassi a, b, c, sia $\text{vol} = \frac{4}{3}\pi abc$ (come dimostreremo nel secondo volume).

Esercizi

68 Calcolare il lavoro della forza $F(x) = \frac{k}{x}$ nello spostare un punto da a a b $(0 < a < b)$.

69 Calcolare il volume del solido che si ottiene facendo ruotare intorno all'asse y l'arco di parabola $y = x^2, x \in [0, r]$, e chiudendo la figura ottenuta col piano $y = r^2$. (Attenzione: poiché l'asse di rotazione è l'asse y, occorre vedere la curva come funzione $x = f(y)$).

70 Calcolare il volume dei solidi che si ottengono dalla rotazione della curva $\sqrt[3]{x}$ attorno ai due assi cartesiani, per $0 \leq x \leq 1$.

7 CALCOLO NUMERICO APPROSSIMATO DI UN INTEGRALE

Come abbiamo già osservato, anche per funzioni continue abbastanza semplici può capitare di non essere in grado di scrivere esplicitamente una primitiva, in termini di funzioni elementari. Ad esempio, la semplice funzione e^{-x^2}, che come vedremo gioca un ruolo importante in molte questioni applicate, ha una primitiva non esprimibile mediante composizione di funzioni elementari. Quindi, per una funzione di questo tipo, il calcolo esatto dell'integrale definito su un certo intervallo, mediante il teorema fondamentale del calcolo integrale, risulta in pratica impossibile. In questi e in molti altri casi può essere utile avere comunque una valutazione approssimata del valore numerico di un integrale definito. Il modo più semplice di procedere è usare la definizione stessa:

sia data una funzione continua $f : [a, b] \to \mathbb{R}$, e costruiamo una successione s_n di Cauchy-Riemann per f. Poiché sappiamo che

$$s_n \to \int_a^b f(x)\, dx \text{ per } n \to \infty,$$

ci aspettiamo che il valore numerico di s_n per n "molto grande" sia una buona approssimazione numerica dell'integrale; d'altro canto s_n può essere calcolato facilmente mediante un computer. Il problema di questo procedimento è che non sappiamo, così facendo, *di quanto sbagliamo*. Questa prima idea, ingenua, diventa però un metodo efficace non appena abbiamo un risultato che ci permetta di stimare a priori l'errore commesso in questa approssimazione. Questo è il significato del prossimo teorema:

Teorema 6.7 (Calcolo numerico di un integrale col metodo del punto medio) *Sia $f : [a, b] \to \mathbb{R}$ una funzione due volte derivabile con continuità, e sia K una costante tale che $|f''(x)| \leq K$ per ogni $x \in [a, b]$ (poiché f'' è continua su $[a, b]$, è anche limitata, perciò una tale costante esiste certamente). Consideriamo ora, per un qualsiasi intero positivo n, la somma di Cauchy-Riemann s_n per f, costruita scegliendo come punto $\xi_i \in [x_{i-1}, x_i]$ il punto medio di $[x_{i-1}, x_i]$, per ogni $i = 1, 2, , \ldots, n$:*

$$s_n = \sum_{i=1}^{n} f\left(\frac{x_{i-1} + x_i}{2}\right) \frac{(b-a)}{n}.$$

Allora vale la stima:

$$\left| \int_a^b f(x)\, dx - s_n \right| \leq \frac{K}{24} \cdot \frac{(b-a)^3}{n^4}.$$

Dimostrazione. Consideriamo un singolo intervallino $[x_{i-1}, x_i]$; chiamiamo ξ_i il suo punto medio e h la sua ampiezza. Dunque risulta

$$[x_{i-1}, x_i] = \left[\xi_i - \frac{h}{2}, \xi_i + \frac{h}{2}\right]; \quad h = \frac{b-a}{n};$$

$$s_n = h \sum_{i=1}^{n} f(\xi_i);$$

Scriviamo ora lo sviluppo di Taylor al prim'ordine con resto secondo Lagrange, per $f(x)$ con centro ξ_i e $x \in \left[\xi_i - \frac{h}{2}, \xi_i + \frac{h}{2}\right]$. Sarà:

$$f(x) = f(\xi_i) + f'(\xi_i)(x - \xi_i) + \frac{f''(c)}{2}(x - \xi_i)^2$$

per qualche $c \in \left[\xi_i - \frac{h}{2}, \xi_i + \frac{h}{2}\right]$. Poiché $|f''(x)| \leq K$ per ogni $x \in [a, b]$ possiamo scrivere allora:

$$f(\xi_i) + f'(\xi_i)(x - \xi_i) - \frac{K}{2}(x - \xi_i)^2 \leq f(x) \leq f(\xi_i) + f'(\xi_i)(x - \xi_i) + \frac{K}{2}(x - \xi_i)^2$$

per ogni $x \in \left[\xi_i - \frac{h}{2}, \xi_i + \frac{h}{2}\right]$. Integrando le disuguaglianze precedenti in $\left[\xi_i - \frac{h}{2}, \xi_i + \frac{h}{2}\right]$ otteniamo:

$$\int_{\xi_i - \frac{h}{2}}^{\xi_i + \frac{h}{2}} f(x)\,dx \leq \int_{\xi_i - \frac{h}{2}}^{\xi_i + \frac{h}{2}} \left[f(\xi_i) + f'(\xi_i)(x - \xi_i) + \frac{K}{2}(x - \xi_i)^2\right] dx =$$

$$= f(\xi_i)\,h + f'(\xi_i) \int_{\xi_i - \frac{h}{2}}^{\xi_i + \frac{h}{2}} (x - \xi_i)\,dx + \frac{K}{2} \int_{\xi_i - \frac{h}{2}}^{\xi_i + \frac{h}{2}} (x - \xi_i)^2\,dx.$$

Notiamo ora che per simmetria

$$\int_{\xi_i - \frac{h}{2}}^{\xi_i + \frac{h}{2}} (x - \xi_i)\,dx = \int_{-h/2}^{h/2} x dx = 0$$

mentre

$$\int_{\xi_i - \frac{h}{2}}^{\xi_i + \frac{h}{2}} (x - \xi_i)^2\,dx = \int_{-h/2}^{h/2} x^2 dx = 2\left[\frac{x^3}{3}\right]_0^{\frac{h}{2}} = \frac{h^3}{12}.$$

Quindi

$$\int_{\xi_i - \frac{h}{2}}^{\xi_i + \frac{h}{2}} f(x)\,dx \leq f(\xi_i)\,h + \frac{Kh^3}{24}.$$

Analogamente

$$\int_{\xi_i - \frac{h}{2}}^{\xi_i + \frac{h}{2}} f(x)\,dx \geq f(\xi_i)\,h - \frac{Kh^3}{24}.$$

quindi

$$\left|\int_{\xi_i - \frac{h}{2}}^{\xi_i + \frac{h}{2}} f(x)\,dx - f(\xi_i)\,h\right| \leq \frac{Kh^3}{24}.$$

Possiamo quindi stimare:

$$\left|\int_a^b f(x)\,dx - s_n\right| = \left|\int_a^b f(x)\,dx - \sum_{i=1}^n f(\xi_i)\,h\right| =$$

$$= \left|\sum_{i=1}^n \left(\int_{x_{i-1}}^{x_i} f(x)\,dx - f(\xi_i)\,h\right)\right| \leq \sum_{i=1}^n \left|\int_{x_{i-1}}^{x_i} f(x)\,dx - f(\xi_i)\,h\right| \leq$$

$$\leq \sum_{i=1}^n \frac{Kh^3}{24} = \frac{K}{24} \cdot \frac{h^3}{n} = \frac{K}{24} \cdot \frac{(b-a)^3}{n^4}.$$

8 INTEGRALI GENERALIZZATI

In molte situazioni concrete (calcolo delle probabilità e statistica ne sono esempi) si presentano integrali di funzioni illimitate e integrali estesi a intervalli illimitati. Ci occuperemo ora di queste situazioni, mostrando come la teoria precedente si possa estendere in maniera naturale per tenere conto anche di questi casi.

8.1 Integrazione di funzioni non limitate

Consideriamo il caso tipico illustrato in figura 6.15, in cui $f : [a, b) \to \mathbb{R}$ è continua e

$$\lim_{x \to b^-} f(x) = +\infty$$

(Del tutto analogo è il caso: $\lim_{x \to b^-} f(x) = -\infty$).

Figura 6.15. $\lim_{x \to b^-} f(x) = +\infty$.

Per definire l'integrale di f in $[a, b]$, l'idea è molto semplice: si integra tra a e $b - \varepsilon$ ($\varepsilon > 0$) e poi si passa al limite per $\varepsilon \to 0^+$. In simboli, si pone

$$\boxed{\int_a^b f(x)dx = \lim_{\varepsilon \to 0^+} \int_a^{b-\varepsilon} f(x)dx} \tag{8.1}$$

Definizione 6.3 Se il limite in (8.1) *esiste finito* allora f si dice *integrabile* in $[a, b]$ oppure che l'integrale $\int_a^b f(x)dx$ è *convergente*.
Se il limite in (8.1) è $+\infty$ oppure $-\infty$, l'integrale si dirà *divergente*.
Se il limite non esiste allora l'integrale *non esiste*.

Analoghe definizioni si hanno se $f : (a, b] \longrightarrow \mathbb{R}$, con f continua e

$$\lim_{x \to a^+} f(x) = \pm\infty$$

Si pone:

$$\boxed{\int_a^b f(x)dx = \lim_{\varepsilon\to 0^+}\int_{a+\varepsilon}^b f(x)dx} \tag{8.2}$$

Esempio importante

8.1 Calcolo dell'integrale

$$\int_a^b \frac{dx}{(b-x)^\alpha} \qquad (\alpha > 0)$$

Caso $\boxed{\alpha = 1}$. Si ha:

$$\int_a^{b-\varepsilon} \frac{dx}{(b-x)} = \big[-\log(b-x)\big]_a^{b-\varepsilon} = -\log\varepsilon + \log(b-a)$$

Dunque

$$\int_a^b \frac{dx}{(b-x)} = \lim_{\varepsilon\to 0^+}\big[-\log\varepsilon + \log(b-a)\big] = +\infty$$

Quindi l'integrale è *divergente.*

Caso $\boxed{\alpha \neq 1}$. Si ha:

$$\int_a^{b-\varepsilon} \frac{dx}{(b-x)^\alpha} = \frac{1}{1-\alpha}\big[-(b-x)^{1-\alpha}\big]_a^{b-\varepsilon} = \frac{1}{1-\alpha}\big[-\varepsilon^{1-\alpha} + (b-a)^{1-\alpha}\big]$$

Dunque

$$\int_a^b \frac{dx}{(b-x)^\alpha} = \lim_{\varepsilon\to 0^+}\frac{1}{1-\alpha}\big[-\varepsilon^{1-\alpha} + (b-a)^{1-\alpha}\big] = \begin{cases} +\infty & \text{se } \alpha > 1 \\ \dfrac{(b-a)^{1-\alpha}}{1-\alpha} & \text{se } \alpha < 1 \end{cases}$$

Riassumendo abbiamo:

$$\boxed{\int_a^b \frac{dx}{(b-x)^\alpha} \quad \text{è} \quad \begin{cases} \text{divergente a } +\infty & \text{se} \quad \alpha \geq 1 \\ \text{convergente } = \dfrac{(b-a)^{1-\alpha}}{1-\alpha} & \text{se} \quad \alpha < 1 \end{cases}} \tag{8.3}$$

Un risultato perfettamente analogo vale per l'integrale

$$\int_a^b \frac{dx}{(x-a)^\alpha}$$

8.2 Criteri di integrabilità al finito

Siano $f, g\colon [a,b) \longrightarrow \mathbb{R}$, continue, con

$$\lim_{x\to b^-} f(x) = \lim_{x\to b^-} g(x) = +\infty$$

I seguenti criteri permettono di decidere se un integrale è convergente o divergente, senza calcolarlo:

Confronto
Se $0 \leq f(x) \leq g(x)$ in $[a, b)$, allora

g	integrabile	$\Longrightarrow$	f	integrabile
f	non integrabile	$\Longrightarrow$	g	non integrabile

Infatti, per la proprietà di monotonia dell'integrale, si ha:

$$0 \leq \int_a^{b-\varepsilon} f(x)dx \leq \int_a^{b-\varepsilon} g(x)dx$$

e, passando al limite per $\varepsilon \to 0^+$, si prova la tesi.

Confronto asintotico
Se $f > 0$, $g > 0$ e $f \sim g$ per $x \to b^-$ allora

f	integrabile	$\Longleftrightarrow$	g	integrabile

Analoghi criteri valgono se $f, g \to +\infty$ per $x \to a^+$, o se $f, g \to -\infty$.

In quest'ultimo caso, le disuguaglianze del criterio del confronto devono valere tra i moduli di f e g.

Esempio

8.2 Consideriamo gli integrali

$$I_1 = \int_0^1 \frac{dx}{\sqrt[3]{1-x^2}}, \qquad I_2 = \int_1^3 \frac{dx}{x^2 - 5x + 4}$$

I_1: La funzione $f(x) = \frac{1}{\sqrt[3]{(1-x^2)}} = \frac{1}{(1-x)^{1/3}(1+x)^{1/3}}$ è continua e positiva in $[0, 1]$ e tende a $+\infty$ per $x \to 1^-$. Inoltre $f(x) \sim \frac{1}{\sqrt[3]{2}} \frac{1}{(1-x)^{1/3}}$ per $x \to 1^-$. D'altra parte la funzione $g(x) = \frac{1}{\sqrt[3]{2}(1-x)^{1/3}}$ è positiva e integrabile (dalla (8.3) con $\alpha = 1/3$) e quindi anche f risulta integrabile, in base al confronto asintotico.

L'integrale I_1 è perciò convergente.

I_2: La funzione $f(x) = \frac{1}{x^2-5x+4} = \frac{1}{(x-1)(x-4)}$ è continua e negativa in $[1, 3]$ e tende a $-\infty$ per $x \to 1^+$. Il segno non costituisce un problema in quanto possiamo riferirci a $-f(x)$ che è positiva.

Osserviamo che $-f(x) \sim \frac{1}{3(x-1)}$ e che $g(x) = \frac{1}{3(x-1)}$ è (positiva) e non integrabile (sempre dalla (8.3) con $\alpha = 1$). Ne segue che anche $-f$ ed f non sono integrabili.

L'integrale I_2 è perciò divergente a $-\infty$.

Analogamente a quanto accade con i criteri di convergenza per le serie numeriche (v. cap. 5), quando la convergenza di un integrale viene stabilita mediante confronto o confronto asintotico tra due funzioni, il *valore numerico* dei due integrali sarà in generale *diverso*.

Una funzione $f(x)$ potrebbe essere illimitata per $x \to a^+$ senza avere segno definitivamente costante. Ad esempio,

$$f(x) = \frac{1}{\sqrt{x}} \sin \frac{1}{x}$$

è illimitata per $x \to 0^+$, ma non tende a $\pm\infty$, e il suo segno è variabile in ogni intorno destro di 0. Per funzioni di questo tipo i criteri del confronto e del confronto asintotico non sono applicabili. Vale invece l'implicazione seguente:

Teorema 6.8

$$\int_a^b |f(x)|dx \quad \textit{convergente} \quad \Longrightarrow \int_a^b f(x)dx \quad \textit{convergente}$$

Se $|f|$ è integrabile in $[a, +\infty)$ si dice che f è *assolutamente integrabile* in $[a, +\infty)$. Perciò il teorema afferma che una funzione assolutamente integrabile è anche integrabile.

Dimostrazione. La dimostrazione è simile a quella dell'analoga proprietà delle serie numeriche (si veda il Teorema 5.3 cap. 5). Per definizione,

$$\int_a^b f(x)\,dx = \lim_{\varepsilon \to 0} \int_{a+\varepsilon}^b f(x)\,dx.$$

Separiamo ora la funzione f nella sua *parte positiva e negativa*:

$$f(x) = f^+(x) - f^-(x),$$

dove, per definizione:

$$f^+(x) = \begin{cases} f(x) & \text{se } f(x) \geq 0 \\ 0 & \text{se } f(x) < 0 \end{cases} \qquad f^-(x) = \begin{cases} -f(x) & \text{se } f(x) \leq 0 \\ 0 & \text{se } f(x) > 0 \end{cases}$$

Le funzioni f^+, f^- sono ≥ 0 e soddisfano le disuguaglianze:

$$\int_{a+\varepsilon}^b f^+(x)\,dx \leq \int_{a+\varepsilon}^b |f(x)|\,dx;$$
$$\int_{a+\varepsilon}^b f^-(x)\,dx \leq \int_{a+\varepsilon}^b |f(x)|\,dx.$$

Poiché per ipotesi esiste finito $\lim_{\varepsilon \to 0} \int_{a+\varepsilon}^b |f(x)|\,dx$ (in quanto $\int_a^b |f(x)|dx$ è convergente), ne segue che le funzioni

$$\varepsilon \longmapsto \int_{a+\varepsilon}^b f^+(x)\,dx; \quad \varepsilon \longmapsto \int_{a+\varepsilon}^b f^-(x)\,dx$$

sono limitate. Inoltre, queste funzioni sono anche monotone decrescenti, perché

$$\varepsilon_1 < \varepsilon_2 \Longrightarrow (a+\varepsilon_1, b) \supset (a+\varepsilon_2, b) \Longrightarrow \int_{a+\varepsilon_1}^b f^+(x)\,dx \geq \int_{a+\varepsilon_2}^b f^+(x)\,dx$$

perché $f^+ \geq 0$ (e lo stesso vale per f^-). Ricordiamo ora che per il Teorema 3.31 (cap. 3), una funzione limitata e monotona ammette limite finito, perciò esistono finiti

$$\lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^{b} f^+(x)\,dx, \quad \lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^{b} f^-(x)\,dx$$

e quindi anche

$$\lim_{\varepsilon \to 0} \int_{a+\varepsilon}^{b} f(x)\,dx = \lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^{b} f^+(x)\,dx - \lim_{\varepsilon \to 0^+} \int_{a+\varepsilon}^{b} f^-(x)\,dx.$$

Di conseguenza f è integrabile. ♦

Esempio

8.3 L'integrale

$$\int_0^1 \frac{1}{\sqrt{x}} \sin \frac{1}{x} dx$$

converge, perché

$$\left| \frac{1}{\sqrt{x}} \sin \frac{1}{x} \right| \leq \frac{1}{\sqrt{x}}$$

che ha integrale convergente.

8.3 Integrazione su intervalli illimitati

Sia $f : [a, +\infty) \longrightarrow \mathbb{R}$, continua. Poniamo

$$\boxed{\int_a^{+\infty} f(x)dx = \lim_{\omega \to +\infty} \int_a^{\omega} f(x)dx} \tag{8.4}$$

Figura 6.16. L'integrale su $[a, +\infty)$ è definito come $\lim_{\omega \to +\infty} \int_a^{\omega} f(x)dz$.

Definizione 6.4 Se il limite in (8.4) *esiste finito* allora f si dice *integrabile* in $[a, +\infty)$ oppure che l'integrale $\int_a^{+\infty} f(x)dx$ è *convergente*. Se il limite in (8.4) è $+\infty$ oppure $-\infty$, l'integrale si dirà *divergente*. Se infine il limite non esiste allora l'integrale *non esiste*.

Analogamente se $f : (-\infty, b] \longrightarrow \mathbb{R}$ è continua, si pone

$$\boxed{\int_{-\infty}^{b} f(x)dx = \lim_{\omega \to -\infty} \int_{\omega}^{b} f(x)dx}$$

ed infine, se $f : (-\infty, +\infty) \longrightarrow \mathbb{R}$ è continua, si pone

$$\boxed{\int_{-\infty}^{+\infty} f(x)dx = \int_{-\infty}^{c} f(x)dx + \int_{c}^{+\infty} f(x)dx}$$

dove c è un punto qualunque.

Nel seguito, per semplicità, ci riferiremo all'intervallo $[a, +\infty)$.

Esempio importante

8.4 Calcolo dell'integrale

$$\int_{1}^{+\infty} \frac{1}{x^{\alpha}} dx \qquad \alpha > 0$$

Caso $\boxed{\alpha = 1}$. Si ha

$$\int_{1}^{\omega} \frac{1}{x} dx = [\log x]_1^{\omega} = \log \omega$$

Poiché $\lim_{\omega \to +\infty} \log \omega = +\infty$, l'integrale è divergente.

Caso $\boxed{\alpha \neq 1}$. Si ha:

$$\int_{1}^{\omega} \frac{1}{x^{\alpha}} dx = \frac{1}{1-\alpha} [x^{1-\alpha}]_1^{\omega} = \frac{1}{1-\alpha} (\omega^{1-\alpha} - 1)$$

Dunque

$$\int_{1}^{+\infty} \frac{1}{x^{\alpha}} dx = \lim_{\omega \to +\infty} \frac{1}{1-\alpha} (\omega^{1-\alpha} - 1) = \begin{cases} +\infty & \text{se} \quad \alpha < 1 \\ \dfrac{1}{\alpha - 1} & \text{se} \quad \alpha > 1 \end{cases}$$

Riassumendo,

$$\boxed{\int_{1}^{+\infty} \frac{1}{x^{\alpha}} dx \quad \text{è} \quad \begin{cases} \text{divergente a} & +\infty & \text{se} \quad \alpha \leq 1 \\ \text{convergente} & = \dfrac{1}{\alpha - 1} & \text{se} \quad \alpha > 1 \end{cases}} \tag{8.5}$$

Divergenza della serie armonica

Nel capitolo 5, par. 1.1 abbiamo affermato che la serie armonica

$$\sum_{n=1}^{\infty} \frac{1}{n}$$

è divergente. L'uso degli integrali generalizzati permette una rapida giustificazione dell'affermazione.

Figura 6.17. L'area ombreggiata è uguale a $\int_1^6 \frac{1}{x} dx$ mentre la somma delle aree dei rettangoli è uguale a $\sum_{n=1}^{5} \frac{1}{n}$.

Dalla figura 6.17 è facile convincersi che fissato N, intero, vale sempre la disuguaglianza

$$\int_1^N \frac{1}{x} dx < \sum_{n=1}^{N-1} \frac{1}{n} \tag{8.6}$$

Poiché $\int_1^N \frac{1}{x} dx = \log N \to +\infty$ se $N \to +\infty$, anche $\sum_{n=1}^{N} \frac{1}{n} \to +\infty$ se $N \to +\infty$, che dimostra la divergenza della serie armonica.

Convergenza della serie armonica generalizzata per $\alpha > 1$

Abbiamo visto (cap. 5, par. 1.2) che la serie

$$\sum_{n=1}^{\infty} \frac{1}{n^\alpha}$$

converge per $\alpha > 1$ (e diverge per $\alpha \leq 1$). L'affermazione è stata dimostrata per $\alpha \geq 2$ (confronto con la serie di Mengoli). Siamo ora in grado di dimostrarla per qualunque $\alpha > 1$. Il ragionamento è analogo a quello sulla serie armonica, con le disuguaglianze in senso inverso. Dalla figura 6.18 si vede che per ogni intero N vale la disuguaglianza

$$\sum_{n=2}^{N} \frac{1}{n^\alpha} \leq \int_1^N \frac{dx}{x^\alpha}$$

Poiché per $\alpha > 1$ e $N \to \infty$ l'integrale converge (per quanto già visto), anche la serie converge (la successione delle somme parziali è crescente e superiormente limitata).

Figura 6.18.

8.4 Criteri di integrabilità all'infinito

Siano $f, g : [a, +\infty) \longrightarrow \mathbb{R}$, continue.

Per decidere se un integrale è convergente o meno, valgono criteri analoghi a quelli per l'integrale di funzioni illimitate.

Confronto

Se $0 \leq f(x) \leq g(x)$ in $[a, +\infty)$ allora

	g integrabile	$\Longrightarrow$	f	integrabile
f	non integrabile	$\Longrightarrow$	g	non integrabile

Confronto asintotico

Se $f > 0$, $g > 0$ e $f \sim g$ per $x \to +\infty$, allora

f	integrabile	$\Longleftrightarrow$	g	integrabile

Esempi

8.5 L'integrale

$$\int_0^{+\infty} e^{-x^2}\,dx$$

è convergente.

Infatti, si può scrivere

$$\int_0^{+\infty} e^{-x^2}\,dx = \int_0^1 e^{-x^2}\,dx + \int_1^{+\infty} e^{-x^2}\,dx$$

Osserviamo ora che per $x > 1$ si ha $x^2 > x$ e quindi $e^{-x^2} < e^{-x}$.

D'altra parte

$$\int_1^{+\infty} e^{-x}dx = \lim_{\omega\to+\infty}\int_1^{\omega} e^{-x}dx = \lim_{\omega\to+\infty}\left[-e^{-\omega}+\frac{1}{e}\right] = \frac{1}{e}$$

Per confronto si deduce che anche $\int_1^{+\infty} e^{-x^2}dx$ è convergente.

86 $$I_1 = \int_0^{+\infty}\frac{x^2+1}{x^4+x+1}dx, \qquad I_2 = \int_0^{+\infty}\frac{1}{\sqrt{x^2+1}}dx$$

I_1: $f(x) \sim \frac{1}{x^2}$ per $x \to +\infty$ e quindi è integrabile ((8.5), con $\alpha = 2$). L'integrale I_1 è pertanto convergente.

I_2: $f(x) \sim \frac{1}{x}$ per $x \to +\infty$ e quindi non è integrabile ((8.5) con $\alpha = 1$).

L'integrale I_2 è pertanto divergente a $+\infty$.

Per funzioni di segno qualunque si ha ancora il:

Teorema 6.9

$$\int_a^{+\infty}|f(x)|dx \quad \textit{convergente} \quad \Longrightarrow \int_a^{+\infty} f(x)dx \quad \textit{convergente}$$

In altre parole, se f è *assolutamente integrabile* in $[a, +\infty)$, è anche integrabile. La dimostrazione è molto simile a quella fatta nel caso degli integrali generalizzati su intervalli limitati.

Esercizi

Stabilire quali dei seguenti integrali esistono (eventualmente in senso generalizzato). In caso affermativo, calcolarlo quando è possibile.

71 $\displaystyle\int_0^1 x\log x\,dx$

72 $\displaystyle\int_0^{+\infty}\frac{1}{\sqrt{x}}e^{-\sqrt{x}}dx$

73 $\displaystyle\int_0^1 \frac{\sin x}{x^2}dx$

74 $\displaystyle\int_0^{+\infty} e^{-2x}\sin(e^{-x})\,dx$

75 $\displaystyle\int_2^4 \operatorname{arctg}\left(\frac{x}{x-3}\right)dx$

76 $\displaystyle\int_2^{+\infty}\frac{1}{x\log x}dx$

77 $\displaystyle\int_0^1 \frac{1-e^{-x}}{x^{3/2}}dx$

78 $\displaystyle\int_2^{+\infty}\frac{1}{x(\log x)^2}dx$

79 $\displaystyle\int_0^{1/2}\frac{1}{x\log x}dx$

80 $\displaystyle\int_0^{+\infty}\frac{x}{(1+x^2)^2}dx$

81 $\displaystyle\int_0^{1/2}\frac{1}{x(\log x)^2}dx$

82 $\displaystyle\int_0^{+\infty}\sin(x^2)dx$

L'ultimo integrale è detto integrale di Fresnel, ed interviene in ottica.

Suggerimento: su $[0,1]$ la funzione $\sin(x^2)$ è integrabile. Per studiare l'integrale su $[1,+\infty]$, valutare l'integrale su $[1,a]$ eseguendo prima la sostituzione $x^2 = t$ e poi un'integrazione per parti. Cosa succede per $a \to +\infty$?

9 FUNZIONI INTEGRALI

Sia f una funzione integrabile (eventualmente in senso generalizzato) in un intervallo I, limitato o illimitato. Consideriamo l'integrale di f su un intervallo J contenuto in I. Ovviamente, al variare dell'intervallo J varia il valore di tale integrale. Se J è l'intervallo che ha un estremo x_0 fissato una volta per tutte, e l'altro estremo x variabile, l'integrale di f su tale intervallo $[x_0, x]$ diventa una funzione di x. Tale funzione si dice *funzione integrale di* f, e si indica con:

$$F(x) = \int_{x_0}^{x} f(t)\,dt$$

Si noti che la variabile d'integrazione t (variabile muta), ha un nome diverso dalla variabile x, estremo mobile dell'intervallo di integrazione: infatti t varia tra x_0 e x.

Non sarà inutile a questo punto riepilogare i diversi concetti che sono stati introdotti a proposito dell'integrale:

1. L'*integrale* (*definito*) di f su $[a, b]$,

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

è un *numero*, definito come opportuno limite di somme;

2. La *funzione integrale* di f,

$$F(x) = \int_{x_0}^{x} f(t)\,dt$$

è, appunto, una *funzione*: ad ogni x associa il numero ottenuto calcolando un integrale definito;

3. L'*integrale indefinito* di f,

$$\int f(x)\,dx$$

è invece un simbolo che denota un *insieme di funzioni*, precisamente l'insieme di tutte le primitive di f, ossia delle funzioni G tali che $G'(x) = f(x)$ per ogni x.

Come si vede, sono tre concetti ben diversi, che non devono essere confusi.

Le funzioni integrali, di cui qui ci occupiamo, compaiono in modo naturale in varie applicazioni:

1. *Meccanica*. Lo spazio percorso fino all'istante t da una particella che si muove di velocità variabile $v(t)$ lungo una traiettoria assegnata, è la funzione integrale della velocità:

$$s(t) = \int_0^t v(\tau)\,d\tau$$

2. *Probabilità*. La "funzione degli errori"

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt$$

descrive qual è la probabilità che una certa "variabile casuale" assuma un valore $\leq x$; la variabile in questione descrive, ad esempio, sotto certe ipotesi, l'errore

commesso nella misurazione di una grandezza fisica. Si noti che per ogni x quello che si calcola in questo caso è un integrale generalizzato, esteso all'intervallo illimitato $(-\infty, x)$.

Il grafico della funzione integranda e della funzione integrale, rispettivamente, sono riportati nella figura 6.19.

Figura 6.19.

3. *Ottica.* Nello studio di certi fenomeni legati all'interferenza luminosa, compaiono le *funzioni integrali di Fresnel*:

$$C(x) = c_1 \int_0^x \cos(t^2)dt \qquad S(x) = c_2 \int_0^x \sin(t^2)dt$$

dove c_1, c_2 sono opportune costanti di normalizzazione.

Figura 6.20.

Ma le funzioni integrali hanno anche una fondamentale importanza matematica, dovuta al:

Teorema 6.10 (Secondo teorema fondamentale del calcolo integrale) *Sia $f : [a, b] \to \mathbb{R}$ una funzione integrabile (in senso proprio o generalizzato), sia $x_0 \in [a, b]$ e sia*

$$F(x) = \int_{x_0}^{x} f(t)\, dt.$$

Allora:

1. *La funzione F è continua in $[a, b]$.*
2. *Se inoltre f è continua in $[a, b]$, allora F è derivabile in $[a, b]$, e vale*

$$F'(x) = f(x) \text{ per ogni } x \in [a, b].$$

Dimostrazione. Proviamo il punto 1 prima nell'ipotesi in cui f è integrabile in senso proprio, e quindi, in particolare, è limitata. Perciò esiste $k > 0$ per cui è

$$|f(t)| \leq k \ \forall t \in [a, b].$$

Per provare che F è continua in un generico punto $\overline{x} \in [a, b]$, consideriamo ora:

$$\begin{aligned} |F(\overline{x} + h) - F(\overline{x})| &= \left| \int_{x_0}^{\overline{x}+h} f(t)\, dt - \int_{x_0}^{\overline{x}} f(t)\, dt \right| = \\ &= \left| \int_{\overline{x}}^{\overline{x}+h} f(t)\, dt \right| \leq \left| \int_{\overline{x}}^{\overline{x}+h} |f(t)|\, dt \right| \leq \\ &\leq \left| \int_{\overline{x}}^{\overline{x}+h} k dt \right| = k\,|h|. \end{aligned}$$

(Abbiamo usato l'additività dell'integrale rispetto all'intervallo di integrazione, e la disuguaglianza del valore assoluto). Perciò per $h \to 0$ si ha $F(\overline{x} + h) \to F(\overline{x})$, e F è continua in $\overline{x}$.

Supponiamo ora che f sia integrabile solo in senso generalizzato, e che $\overline{x}$ sia proprio un punto nell'intorno del quale f è illimitata (altrimenti la precedente dimostrazione vale anche in questo caso). Abbiamo ancora:

$$F(\overline{x} + h) - F(\overline{x}) = \int_{\overline{x}}^{\overline{x}+h} f(t)\, dt.$$

Scriviamo ora:

$$\int_{\overline{x}}^{b} f(t)\, dt = \int_{\overline{x}}^{\overline{x}+h} f(t)\, dt + \int_{\overline{x}+h}^{b} f(t)\, dt$$

Poiché per ipotesi f è integrabile in senso generalizzato, si ha, per definizione:

$$\int_{\overline{x}+h}^{b} f(t)\, dt \to \int_{\overline{x}}^{b} f(t)\, dt \text{ per } h \to 0^+$$

e quindi per differenza

$$\int_{\overline{x}}^{\overline{x}+h} f(t)\, dt \to 0 \text{ per } h \to 0^+$$

ossia $F(\overline{x} + h) \to F(\overline{x})$ per $h \to 0^+$. Analogamente si ragiona per $h \to 0^-$.

Proviamo ora il punto 2, partendo ancora dall'uguaglianza, già ottenuta,

$$F(\overline{x}+h)-F(\overline{x})=\int_{\overline{x}}^{\overline{x}+h} f(t)\,dt.$$

Poiché ora per ipotesi f è continua, per il Teorema della media (Teorema 6.5 par. 3) si ha

$$\int_{\overline{x}}^{\overline{x}+h} f(t)\,dt = hf(x_h)$$

per qualche $x_h \in [\overline{x}, \overline{x}+h]$ (supponendo, per fissare le idee, $h>0$). Ne segue

$$\frac{F(\overline{x}+h)-F(\overline{x})}{h}=f(x_h).$$

Per $h \to 0$ si avrà $x_h \to \overline{x}$ e quindi, ancora per la continuità di f, $f(x_h) \to f(\overline{x})$. Pertanto esiste

$$\lim_{h\to 0}\frac{F(\overline{x}+h)-F(\overline{x})}{h}=f(\overline{x})$$

e il teorema è dimostrato. ♦

Il teorema ha varie conseguenze:

1. Anzitutto, se f è continua, allora F è derivabile *con continuità*, perché $F'=f$ e f è continua.
2. Se f non è solo continua, ma è derivabile con continuità, allora anche F' è derivabile con continuità (perché è $=f$), perciò F è due volte derivabile con continuità. Iterando: *la funzione integrale ha sempre un grado di regolarità in più rispetto alla funzione integranda.*
3. Dal teorema segue in particolare che *ogni funzione continua su I ha ivi una primitiva*: la sua funzione integrale. (Questa affermazione era stata fatta nel par. 4, ma non dimostrata). Possiamo scrivere quindi:

$$\int f(x)dx = \int_{x_0}^{x} f(t)dt + c$$

Esempio

9.1 Sia

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}}\,dt$$

(Si noti che non si sa scrivere esplicitamente una primitiva di $e^{-\frac{t^2}{2}}$, e quindi la funzione integrale non si può riscrivere in forma più semplice). Allora:

$$F'(x)=\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$$

Poiché $F'(x)>0$ per ogni x, $F(x)$ è sempre crescente. Inoltre

$$F''(x)=-\frac{2x}{\sqrt{2\pi}}e^{-\frac{x^2}{2}} \geq 0 \text{ per } x \leq 0$$

perciò $F(x)$ ha un punto di flesso per $x=0$, ed è concava verso l'alto per $x<0$. Per $x \to +\infty$, $F(x)$ tende all'integrale generalizzato $\int_{-\infty}^{+\infty}\frac{1}{\sqrt{2\pi}}e^{-\frac{t^2}{2}}\,dt$, che è finito. Perciò $F(x)$ ha un asintoto orizzontale per $x \to +\infty$ (si può dimostrare che questo asintoto è $y=1$). Infine, per come è definita, per $x \to -\infty$ $F(x) \to 0$. Tutti questi fatti si possono controllare sul grafico di $F(x)$ riportato in precedenza (v. fig. 6.19).

Il punto 2 del teorema si può anche raffinare, al modo seguente. Se la funzione integranda $f(t)$ non è continua su tutto I, ma è integrabile in senso generalizzato, in tutti i punti in cui l'integranda è continua, $F(x)$ è derivabile, e $F'(x) = f(x)$.
In sostanza, dove f è discontinua, F è ancora continua, ma non sarà derivabile.

Esempi

9.2 Sia $f(t) = \text{sgn}(t)$,e

$$F(x) = \int_0^x f(t)dt$$

Allora:

$$F(x) = \begin{cases} \text{se } x > 0 & \displaystyle\int_0^x 1dt = x \\ \text{se } x < 0 & \displaystyle\int_0^x -1dt = -x \end{cases}$$

dunque $F(x) = |x|$. Come si vede in figura 6.21, F è continua ovunque, ma presenta un punto angoloso (quindi di non derivabilità) esattamente dove la funzione integranda è discontinua.

Figura 6.21. a) Grafico di $f(t)$; b) grafico di $F(x)$.

9.3 Sia $f(t) = \frac{1}{\sqrt[3]{t}}$, e

$$F(x) = \int_0^x f(t)dt = \left[\frac{3}{2}t^{2/3}\right]_0^x = \frac{3}{2}x^{2/3}$$

Come si vede in figura 6.22, F è continua ovunque (perché f, anche se illimitata, è comunque integrabile in senso generalizzato), ma presenta un punto di cuspide (quindi di non derivabilità) esattamente dove la funzione integranda è discontinua.

Figura 6.22. a) Grafico di $f(t)$; b) grafico di $F(x)$.

Esercizi

83 Si considerino le funzioni di Fresnel $C(x), S(x)$ definite negli esempi precedenti. Calcolare le loro derivate prima e seconda, e determinare i loro punti di massimo, minimo e flesso.

84 Sia $f(t) = |t|$, $F(x) = \int_0^x f(t)\,dt$. Calcolare esplicitamente $F(x)$ e disegnarne il grafico. Quante volte la funzione F è derivabile su tutto $\mathbb{R}$?

85 Sia $f(t) = e^{-t}$ per $t > 0$, $f(t) = 0$ per $t \leq 0$, e sia $F(x) = \int_0^x f(t)dt$. Calcolare esplicitamente $F(x)$ e tracciarne il grafico. F è continua? È derivabile in tutto $\mathbb{R}$?

86 Sia f la funzione periodica di periodo 2 (cioè $f(x+2) = f(x)$, per ogni x) che assume il valore 1 per $0 \leq x < 1$ ed assume il valore zero per $1 \leq x < 2$.
a) Disegnare il grafico di f.
b) Calcolare $F(x) = \int_0^x f(t)dt$ per $x \geq 0$.
c) Disegnare il grafico di F per $x \geq 0$.

87 Sia $F(x) = \int_0^x f(t)dt$. Provare che se f è pari allora F è dispari, se f è dispari allora F è pari.

88 Studiare la seguente funzione integrale e tracciarne un grafico plausibile:

$$F(x) = \int_0^x \frac{\sin t}{1+t^2}dt$$

Suggerimento: non cercare una primitiva elementare dell'integranda: è inutile.

10 CONVOLUZIONE E SISTEMI FISICI LINEARI

Gli integrali estesi ad intervalli illimitati intervengono nel concetto di *convoluzione* fra due funzioni, che gioca un ruolo importante nella modellizzazione dei cosiddetti "sistemi lineari".

Precisamente, consideriamo un sistema fisico che operi nel modo descritto dal diagramma seguente:

$$e(t) \longrightarrow \boxed{\quad S \quad} \longrightarrow r(t)$$

cioè a un *ingresso* $e(t)$ corrisponde una *risposta* $r(t)$, dipendente dall'ingresso mediante una funzione S caratterizzante il sistema. Si abbia cioè

$$r(t) = S[e](t)$$

Qui ingresso e risposta sono due grandezze variabili nel tempo, cioè due funzioni $e(t), r(t)$; S è un operatore che trasforma la funzione $e(t)$ nella funzione $r(t)$. Ciò significa, in particolare, che il valore $r(t)$ a un istante particolare può dipendere da *tutta* la funzione e, e non solo dal valore $e(t)$ in quell'istante.

Supponiamo che il sistema sia *lineare* ovvero che soddisfi il *principio* di *sovrapposizione* seguente:

Definizione 6.5 La risposta a una somma di ingressi è uguale alla somma delle risposte a ciascun ingresso.

Precisamente, se $e_1(t), e_2(t), \ldots, e_n(t)$ indicano n ingressi e $\lambda_1, \ldots, \lambda_n$ indicano numeri reali, si ha:

$$S\left[\sum_{j=1}^{n} \lambda_j\, e_j\right](t) = \sum_{j=1}^{n} \lambda_j\, S[e_j](t) = \sum_{j=1}^{n} \lambda_j\, r_j(t) \tag{10.1}$$

Supponiamo inoltre che il sistema abbia la seguente proprietà di *invarianza temporale*: se all'ingresso $e(t)$ si sostituisce l'ingresso "ritardato" $e(t-h)$ (h fissato), si ottiene la stessa risposta dell'ingresso $e(t)$, con uguale ritardo, cioè $r(t-h)$. In simboli:

$$S[e(t-h)] = S[e](t-h) \tag{10.2}$$

Sotto queste ipotesi, la risposta a un qualunque ingresso $e(t)$ può essere determinata mediante una particolare operazione, detta *convoluzione*, conoscendo la risposta a un ingresso speciale, l'*impulso di Dirac*.

Proviamo questa affermazione rimanendo tuttavia a un livello intuitivo.

Ricordiamo che l'impulso di Dirac concentrato in $t = 0$, e denotato con $\delta(t)$ si può considerare come il "limite" per $\Delta\tau \to 0$ dell'impulso unitario di durata $\Delta\tau$, definito dalla funzione

$$I_{\Delta\tau}(t) = \begin{cases} 1/\Delta\tau & 0 < t < \Delta\tau \\ 0 & \text{altrove} \end{cases}$$

L'impulso di Dirac concentrato in $t = \tau$, denotato con $\delta(t-\tau)$, sarà il "limite" per $\Delta\tau \to 0$ dell'impulso unitario $I_{\Delta\tau}(t-\tau)$.

Figura 6.23. Si noti che $I_{\Delta\tau}(t-\tau) \cdot \Delta\tau = 1$ per $\tau < t < \tau + \Delta\tau$.

Sia ora $e = e(t)$ un ingresso, integrabile in $(-\infty, +\infty)$. Possiamo pensare di approssimare $e(t)$ con la somma (generalmente infinita) di impulsi di durata $\Delta\tau$, relativi ad un intervallo temporale $(\tau, \tau + \Delta\tau)$, nel modo seguente (v. fig. 6.24).
Per ogni t tale che $k\Delta\tau \le t \le (k+1)\Delta\tau$, si può scrivere $e(t) \approx e(k\Delta\tau) = e(k\Delta\tau) I_{\Delta\tau}(t - k\Delta\tau)\Delta\tau$. Dunque, per ogni t:

$$e(t) \approx \sum_{k=-\infty}^{+\infty} e(k\Delta\tau) I_{\Delta\tau}(t - k\Delta\tau) \cdot \Delta\tau \tag{10.3}$$

(per ogni valore di t uno solo degli addendi nella somma (10.3) è diverso da zero!)

Figura 6.24.

Per il principio di sovrapposizione (10.1) si avrà allora

$$
\begin{aligned}
S[e](t) &\approx \sum_{k=-\infty}^{+\infty} e(k\Delta\tau) S[I_{\Delta\tau}(t-k\Delta\tau)]\Delta\tau \\
&\underset{(\text{se } \Delta\tau\approx 0)}{\approx} \sum_{k=-\infty}^{+\infty} e(k\Delta\tau) S[\delta(t-k\Delta\tau)]\Delta\tau
\end{aligned}
\tag{10.4}
$$

Ponendo $h(t) = S[\delta](t)$ =risposta all'impulso di Dirac, e ricordando che, per la (10.2), si può anche scrivere $S[\delta(t-\tau)] = h(t-\tau)$, passando al limite nella (10.4) per $\Delta\tau \to 0$ si ha:

$$
S[e(t)] = \int_{-\infty}^{+\infty} e(\tau)h(t-\tau)d\tau \tag{10.5}
$$

L'integrale nella (10.5) prende il nome di *convoluzione* di e e h e si indica col simbolo $(e * h)(t)$.

Dunque:

$$
\boxed{S[e](t) = (e * h)(t)}
$$

Ossia:

La risposta di un sistema lineare ad un qualunque ingresso $e(t)$ si ottiene come convoluzione con la risposta all'impulso di Dirac.

In generale, date due funzioni f e g integrabili in $\mathbb{R}$ si chiama *prodotto di convoluzione*, o semplicemente *convoluzione* di f e g la funzione definita dalla formula seguente

$$
\boxed{(f * g)(x) = \int_{-\infty}^{+\infty} f(y)g(x-y)dy} \tag{10.6}
$$

Il simbolo $f * g$ si legge "f *convoluzione* g".

Frequentemente f e g sono *nulle* per $x < 0$; in questo caso particolare si ha:

$$(f * g)(x) = \int_0^x f(y)g(x-y)dy$$

Infatti:

$$\int_{-\infty}^{+\infty} f(y)g(x-y)dy = \int_0^x f(y)g(x-y)dy + \int_{-\infty}^0 f(y)g(x-g)dy + \int_x^{+\infty} f(y)g(x-y)dy$$

ma il secondo integrale è nullo, essendo $f = 0$ per $y < 0$, ed il terzo è nullo essendo $g(x-y) = 0$ per $x - y < 0$ cioè $y > x$.

Proprietà della convoluzione

$$f * g = g * f \qquad \text{(il prodotto "}*\text{" è commutativo)}$$

$$(f * g) * h = f * (g * h) \qquad \text{("}*\text{" è associativo)}$$

$$(f + g) * h = f * h + g * h \qquad \text{("}*\text{" è distributivo)}.$$

Esempi significativi

10.1 *Convoluzione con l'impulso unitario di durata ε.*

$$\int_{-\infty}^{+\infty} I_\varepsilon(t)f(x-t)dt = \frac{1}{\varepsilon}\int_0^\varepsilon f(x-t)dt \underset{(x-t=u)}{\longrightarrow} = \frac{1}{\varepsilon}\int_x^{x+\varepsilon} f(u)du$$

Perciò abbiamo:

$$(I_\varepsilon * f)(x) = \frac{1}{\varepsilon}\int_{x-\varepsilon}^x f(t)dt = \text{ valor medio di } f \text{ su } \quad [x-\varepsilon, x] \tag{10.7}$$

10.2 *Convoluzione con l'impulso di Dirac.*
Passando al limite per $\varepsilon \to 0$ nella (10.7) si ha:

$$(I_\varepsilon * f)(x) \longrightarrow (\delta * f)(x)$$

e, usando il teorema della media per gli integrali

$$\lim_{\varepsilon \to 0} \frac{1}{\varepsilon}\int_{x-\varepsilon}^x f(t)dt = \lim_{c \to x} f(c) = f(x)$$

essendo $x - \varepsilon < c < x$ ed f continua.

Abbiamo dunque l'importante formula:

$$\boxed{(\delta * f)(x) = f(x)} \tag{10.8}$$

La (10.8) indica che l'impulso di Dirac opera come l'elemento "unità" per il prodotto di convoluzione.

Esercizio

88 Sia $I_1 = I_1(t)$ un impulso unitario di durata 1:

$$I_1(t) = \begin{cases} 1 & \text{se } 0 \leq x \leq 1 \\ 0 & \text{se } x < 0 \quad \text{oppure } x > 1 \end{cases}$$

Calcolare

$$(I_1 * I_1)(x) = \int_{-\infty}^{+\infty} I_1(y) I_1(x-y) dy$$

e disegnarne il grafico.

COMPLEMENTI

11 TEOREMA DI BOLZANO-WEIERSTRASS, CONTINUITÀ UNIFORME E INTEGRABILITÀ DELLE FUNZIONI CONTINUE

In questo paragrafo dimostreremo il teorema, enunciato nel par. 2.2, che garantisce l'integrabilità delle funzioni continue su $[a, b]$. Questa dimostrazione costituisce la parte più delicata della teoria dell'integrazione, che come vedremo fa uso delle proprietà "fini" dell'insieme dei numeri reali, e richiederà l'introduzione di diverse nuove idee.

Anzitutto (par. 11.1) discuteremo certe *proprietà delle successioni di numeri reali*, che finora non erano state introdotte; quindi (par. 11.2) introdurremo una nozione simile alla continuità, ma più forte di questa, detta *uniforme continuità*, e proveremo un risultato che la riguarda; con questi strumenti, nel paragrafo 11.3 potremo provare l'integrabilità delle funzioni continue.

Segnaliamo che le nozioni introdotte nei paragrafi 11.1-11.2, pur molto importanti per i fondamenti dell'analisi matematica e per gli ulteriori sviluppi dell'analisi superiore, in questo corso saranno utilizzati esclusivamente in questo paragrafo.

11.1 Alcuni risultati fondamentali per le successioni di numeri reali

Iniziamo con la seguente

DEFINIZIONE 6.6 Data una successione $\{a_n\}_{n=n_0}^{\infty}$ di numeri reali, chiamiamo *sottosuccessione* di $\{a_n\}$ ogni successione estratta da questa, ossia ogni successione del tipo $\{a_{n_k}\}_{k=1}^{\infty}$ dove $n_1 < n_2 < n_3 < \ldots$ è una successione crescente di indici interi.

Ad esempio, dalla successione

$$a_1, a_2, a_3, \ldots, a_n \ldots$$

potremmo estrarre le sottosuccessioni

$$a_2, a_4, a_6, \ldots$$

oppure

$$a_1, a_3, a_5, \ldots$$

o qualsiasi altra

$$a_{n_1}, a_{n_2}, a_{n_3}, \ldots$$

pur di avere scelto una ben precisa successione (strettamente crescente) di indici interi $\{n_k\}_{k=1}^{\infty}$.

Si vede subito che se una successione $\{a_n\}$ converge o diverge, cioè se $a_n \to l$ per qualche $l \in \mathbb{R}^*$, allora ogni sottosuccessione estratta da $\{a_n\}$ tenderà allo stesso l. Invece, in generale, dal fatto che una particolare sottosuccessione $\{a_{n_k}\}$ tenda a un certo l non possiamo dedurre che lo stesso valga per la successione intera. Ad esempio, data la successione

$$a_n = (-1)^n ,$$

se estraiamo la sottosuccessione

$$a_{2n} = (-1)^{2n} = 1$$

otteniamo una successione costante (e quindi convergente), mentre la successione di partenza è oscillante.

La situazione descritta da quest'esempio è più frequente di quanto si potrebbe credere: sovente una successione non è convergente, ma possiede una sottosuccessione convergente, e sapere questo può essere estremamente utile per certi scopi. Il prossimo teorema dà una condizione sufficiente perché questo accada:

Teorema 6.11 (di Bolzano-Weierstrass) *Ogni successione di numeri reali contenuta in un intervallo limitato $[a, b]$ possiede una sottosuccessione convergente in $[a, b]$.*

Per esempio, la successione $\{\sin n\}_{n=1}^{\infty}$, limitata ma estremamente irregolare, in base a questo teorema certamente possiede una sottosuccessione convergente (anche se non è così facile costruirla esplicitamente).

Dimostrazione. Sia $\{x_n\}$ una successione contenuta in $[a, b]$. Eseguiamo una costruzione con la tecnica di dicotomia, già utilizzata nella dimostrazione del teorema di Weierstrass e del teorema degli zeri (v. cap. 3, par. 4.1): dividiamo $[a, b]$ in due sottointervalli di ugual ampiezza; almeno uno di essi dovrà contenere infiniti elementi x_n della successione (non necessariamente distinti: in altre parole, conterrà elementi x_n per un numero infinito di *indici* diversi). Chiamiamo $[a_1, b_1]$ questo sottointervallo e scegliamo un elemento $x_{n_1} \in [a_1, b_1]$. Ora dividiamo $[a_1, b_1]$ in due sottointervalli uguali; almeno uno di essi, che chiamiamo $[a_2, b_2]$, conterrà termini x_n per un numero infinito di indici diversi; in particolare, ne conterrà certamente uno del tipo x_{n_2} con $n_2 > n_1$. Procedendo iterativamente costruiamo una successione di intervalli $[a_k, b_k]$ e una sottosuccessione $\{x_{n_k}\}$ della successione di partenza, con le seguenti proprietà:

1. $\{a_k\}$ è monotona crescente e superiormente limitata da b;
2. $\{b_k\}$ è monotona decrescente e inferiormente limitata da a;
3. $b_k - a_k = (b - a)/2^k$
4. $x_{n_k} \in [a_k, b_k]$

Per la (1) e (2), a_k e b_k convergono (per il teorema di monotonia, v. cap. 3 par. 1.2); per la (3), il limite a cui tendono a_k e b_k è lo stesso, chiamiamolo c; per la (4) e il teorema del confronto, anche $x_{n_k} \to c$; inoltre, poiché $x_{n_k} \in [a_k, b_k] \subset [a, b]$, per il teorema di permanenza del segno anche $c \in [a, b]$. ◆

Il prossimo concetto che dobbiamo introdurre, riguardo alle successioni di numeri reali, è contenuto nella seguente

Definizione 6.7 Sia $\{a_n\}$ una successione di numeri reali. Si dice che la successione soddisfa la *condizione di Cauchy*, o semplicemente che è $\{a_n\}$ una *successione di Cauchy*, se avviene che

$$\forall \varepsilon > 0 \ \exists n_0 : \forall n, m \geq n_0 \text{ si ha } |a_n - a_m| < \varepsilon.$$

(Più brevemente: $\forall \varepsilon > 0$ è $|a_n - a_m| < \varepsilon$ definitivamente). ■

In sostanza, la condizione afferma che i termini della successione sono *sempre più vicini tra loro*. Si confronti con la definizione di successione convergente, che richiede invece che i termini della successione siano *sempre più vicini ad un certo numero reale* ℓ. È naturale pensare che una successione convergente soddisfi la condizione di Cauchy. Infatti, vale il prossimo:

TEOREMA 6.12
(i) Se una successione è convergente, allora è di Cauchy.
(ii) Se una successione è di Cauchy, allora è limitata.

Si lascia al lettore la facile dimostrazione. Più interessante, e più delicato da dimostrare, è il fatto che la condizione di Cauchy sia non solo necessaria ma anche sufficiente per la convergenza:

TEOREMA 6.13 *Sia $\{a_n\}$ una successione di numeri reali. Se la successione è di Cauchy, allora ammette limite finito in $\mathbb{R}$.*

L'interesse del precedente teorema sta nel fatto che per provare la convergenza di una successione è sufficiente verificare la condizione di Cauchy che, come si vede dalla definizione, non richiede la previa conoscenza del *limite* a cui tende la successione.

DIMOSTRAZIONE. Sia $\{a_n\}_{n=1}^{\infty}$ una successione di Cauchy. Per il teorema precedente, $\{a_n\}$ è limitata. Allora, per il teorema di Bolzano-Weierstrass, contiene una sottosuccessione convergente:

$$\exists \{a_{n_k}\}_{k=1}^{\infty} \text{ ed } \exists \ell \in \mathbb{R} \text{ tale che } a_{n_k} \to \ell.$$

Mostriamo che, grazie alla condizione di Cauchy, allora, l'intera successione converge al limite ℓ. Difatti:

$$|a_n - \ell| \leq |a_n - a_{n_k}| + |a_{n_k} - \ell|.$$

Fissato $\varepsilon > 0$, $|a_{n_k} - \ell| < \varepsilon$ defnitivamente, perché $a_{n_k} \to \ell$; d'altro canto, per la condizione di Cauchy è anche $|a_n - a_m| < \varepsilon$ definitivamente, e quindi in particolare $|a_n - a_{n_k}| < \varepsilon$ definitivamente. Perciò

$$|a_n - \ell| < 2\varepsilon$$

definitivamente, e il teorema è dimostrato. ◆

Il teorema appena dimostrato esprime un'importante proprietà dell'insieme dei numeri reali che, ad esempio, non vale nell'insieme dei razionali. Per convincersene, si consideri

$$a_n = \left(1 + \frac{1}{n}\right)^n, \text{ per } n = 1, 2, \ldots$$

Questa è una successione di numeri razionali, che converge (in $\mathbb{R}$) al numero e. Di conseguenza la successione è di Cauchy (sia in $\mathbb{R}$ che in $\mathbb{Q}$). Tuttavia, in $\mathbb{Q}$ la successione non è convergente (perché il numero e è irrazionale).

11.2 Continuità uniforme

La dimostrazione dell'integrabilità delle funzioni continue coinvolge in modo essenziale, come vedremo, una nozione simile alla continuità, ma più forte. Premettiamo quindi la seguente

DEFINIZIONE 6.8 Si dice che $f : I \to \mathbb{R}$ è *uniformemente continua* in I se

$$\begin{aligned} &\forall \varepsilon > 0 \quad \exists \delta > 0 \text{ tale che } \forall x_1, x_2 \in I, \\ &\text{se } |x_1 - x_2| < \delta \text{ allora } |f(x_1) - f(x_2)| < \varepsilon. \end{aligned} \tag{11.1}$$

■

Notiamo la sottile differenza esistente tra le affermazioni "f è uniformemente continua in I" e "f è continua in I". La seconda significa:

$$\forall x_1 \in I, \forall \varepsilon > 0, \quad \exists \delta > 0 \text{ tale che } \forall x_2 \in I,$$
$$\text{se } |x_1 - x_2| < \delta \text{ allora } |f(x_1) - f(x_2)| < \varepsilon.$$

Nella definizione di continuità, il numero δ dipende da ε *e anche da* x_1, mentre in quella di uniforme continuità dipende solo da ε. (Questo si capisce facendo attenzione all'*ordine* in cui compaiono i vari quantificatori $\forall, \exists$ nelle due affermazioni). Detto in modo meno formale: uniforme continuità significa che la quantità $|f(x_1) - f(x_2)|$ può essere controllata conoscendo quanto è piccola la differenza $|x_1 - x_2|$, e questo senza bisogno di sapere dove si trovano i punti x_1, x_2.

Esempio

11.1 La funzione $f(x) = x^2$ è uniformemente continua su $I = [0, 1]$, ma non lo è su $I = \mathbb{R}$. Per provare la prima affermazione, scriviamo:

$$|f(x_1) - f(x_2)| = |x_1^2 - x_2^2| = |(x_1 - x_2)(x_1 + x_2)| \leq$$
$$\leq |x_1 - x_2| (|x_1| + |x_2|) \leq 2|x_1 - x_2|$$

perché $x_1, x_2 \in [0, 1]$. Quindi fissato $\varepsilon > 0$, se $2|x_1 - x_2| < \varepsilon$, ossia se $|x_1 - x_2| < \varepsilon/2$, sarà anche $|f(x_1) - f(x_2)| < \varepsilon$. In questo caso quindi basta scegliere $\delta = \varepsilon/2$, e la definizione di uniforme continuità è soddisfatta.

Consideriamo ora f su tutto $\mathbb{R}$. In questo caso, fissato $\varepsilon > 0$ noi vorremmo trovare $\delta > 0$ tale che $|x_1 - x_2| < \delta$ implichi $|x_1^2 - x_2^2| < \varepsilon$. Questo non è possibile, infatti: comunque scegliamo ε e δ, sia $x_1 = n$ e $x_2 = n + \delta/2$ (con n intero che sceglieremo poi). Si ha:

$$|x_1 - x_2| = \delta/2 < \delta$$

ma

$$|x_1^2 - x_2^2| = |(x_1 - x_2)(x_1 + x_2)| = \left|\frac{\delta}{2}(n + \delta)\right| > n\frac{\delta}{2}.$$

Per quanto piccolo sia δ, questa quantità, per n opportuno, sarà sempre maggiore di $\varepsilon > 0$ fissato. Quindi f non è uniformemente continua su $\mathbb{R}$.

L'esempio precedente voleva illustrare come le nozioni di continuità e continuità uniforme siano effettivamente distinte. La continuità uniforme implica la continuità, ma è una condizione più forte, in generale.

Vale tuttavia il seguente importante

Teorema 6.14 (di Cantor-Heine) *Sia* $f : [a, b] \to \mathbb{R}$ *è continua, allora è* uniformemente continua *in* $[a, b]$.

In altre parole, una funzione continua può violare la condizione di uniforme continuità solo se l'intervallo su cui è definita non è limitato o non è chiuso.

Dimostrazione. Per assurdo, supponiamo che f sia continua in $[a, b]$ ma non uniformemente continua. Questo significa che:

$$\exists \varepsilon > 0 \text{ tale che } \forall \delta > 0 \ \exists x_1, x_2 \in I \text{ tali che } |x_1 - x_2| < \delta \text{ ma } |f(x_1) - f(x_2)| \geq \varepsilon.$$

(si noti come abbiamo costruito la *negazione logica* della definizione di uniforme continuità). Scegliendo allora $\delta = 1/n$ successivamente per $n = 1, 2, 3, \ldots$, possiamo costruire due successioni $\{x_n\}, \{y_n\} \subset [a, b]$ tali che

$$|x_n - y_n| < \frac{1}{n} \text{ ma } |f(x_n) - f(y_n)| \geq \varepsilon.$$

Poiché la successione $\{x_n\}$ è limitata (contenuta in $[a, b]$), per il teorema di Bolzano-Weierstrass ammette una sottosuccessione $\{x_{n_k}\}$ convergente a un certo $c \in [a, b]$. Consideriamo allora la sottosuccessione $\{y_{n_k}\}$ corrispondente agli stessi indici; anche questa, essendo limitata, avrà una sottosuccessione $\left\{y_{n_{k_h}}\right\}$ convergente. Possiamo allora concludere che le due sottosuccessioni

$$\left\{x_{n_{k_h}}\right\}, \left\{y_{n_{k_h}}\right\}$$

sono convergenti; anzi, poiché

$$\left|x_{n_{k_h}} - y_{n_{k_h}}\right| < \frac{1}{n_{k_h}} \to 0,$$

il limite è lo stesso c. D'altro canto la funzione f è continua, perciò

$$f\left(x_{n_{k_h}}\right) \to c \text{ e } f\left(y_{n_{k_h}}\right) \to c.$$

Ma questo è assurdo perché sappiamo che

$$\left|f\left(x_{n_{k_h}}\right) - f\left(y_{n_{k_h}}\right)\right| \geq \varepsilon.$$

Perciò il teorema è dimostrato. ◆

11.3 Integrabilità delle funzioni continue

Abbiamo ora gli strumenti per provare il seguente:

Teorema 6.15 *Se $f : [a, b] \to \mathbb{R}$ è continua, allora è integrabile.*

Dimostrazione. La dimostrazione è in due passi.

1. Mostriamo prima che una qualsiasi successione di Cauchy-Riemann s_n di f è di Cauchy, ossia che per ogni $\varepsilon > 0$ risulta $|s_n - s_m| < \varepsilon$ definitivamente. Consideriamo dunque, per due interi n, m fissati $(n < m)$, i due insiemi di punti, equispaziati in $[a, b]$:

$$x_0 = a, x_1, x_2, \ldots, x_n = b$$
$$y_0 = a, y_1, y_2, \ldots, y_m = b$$

e le corrispondenti somme di Cauchy-Riemann:

$$s_n = \sum_{i=1}^{n} f(\xi_i)(x_i - x_{i-1}); s_m = \sum_{j=1}^{m} f(\eta_j)(y_j - y_{j-1})$$

con

$$\xi_i \in [x_{i-1}, x_i], i = 1, 2, \ldots, n;$$
$$\eta_j \in [y_{j-1}, y_j], j = 1, 2, \ldots, m.$$

Per studiare la differenza $s_n - s_m$ è utile esprimere entrambe le somme rispetto alla medesima suddivisione di $[a, b]$. Siano quindi

$$z_0 = a < z_1 < z_2 < \cdots < z_k = b$$

(con $k = n + m$) tutti i punti x_i e y_j, disposti in ordine crescente. Consideriamo allora la somma s_n. Possiamo riscriverla come somma di k addendi, al modo seguente

$$s_n = \sum_{i=1}^{k} f(\xi_i^*)(z_i - z_{i-1})$$

a patto di definire i punti ξ_i^* a questo modo: per tutti gli intervallini $[z_{i-1}, z_i]$ che compongono l'intervallo (più grande) $[x_{j-1}, x_j]$, si pone $\xi_i^* = \xi_j$ (vedi figure 6.25-6.27).

Figura 6.25. La somma s_3.

Figura 6.26. La somma s_4.

Analogamente riscriviamo la somma s_m come:

$$s_m = \sum_{i=1}^{k} f(\eta_i^*)(z_i - z_{i-1})$$

Figura 6.27. La somma s_3 riscritta aggiungendo ai punti y_i anche i punti x_2, x_3: il punto ξ_2 che si era scelto in $[y_1, y_2]$ ora si usa sia per $[y_1, x_1]$ sia per $[x_1, y_2]$, e si chiamerà ξ_2^* e ξ_3^*; il punto ξ_3 che si era scelto in $[y_2, y_3]$ ora si usa sia per $[y_2, x_2]$ sia per $[x_2, y_3]$, e si chiamerà ξ_4^* e ξ_5^*.

e possiamo quindi valutare la differenza:

$$s_n - s_m = \sum_{i=1}^{k} \left[f\left(\xi_i^*\right) - f\left(\eta_i^*\right) \right] \left(z_i - z_{i-1}\right).$$

Ragioniamo ora sul fatto che i due punti ξ_i^*, η_i^*, anche se non appartengono allo stesso intervallino $[z_{i-1}, z_i]$, si trovano ad una distanza che non supera $\frac{b-a}{n}$. Poiché f è continua, e quindi uniformemente continua, fissato $\varepsilon > 0$ esiste un $\delta > 0$ tale che

$$\left| f\left(\xi_i^*\right) - f\left(\eta_i^*\right) \right| < \varepsilon \text{ purché } \left|\xi_i^* - \eta_i^*\right| < \delta,$$

e questo capita se $\frac{b-a}{n} < \delta$, ossia per $n > \delta/(b-a)$. Allora se $m > n > \delta/(b-a)$ si ha:

$$\left|s_n - s_m\right| \leq \sum_{i=1}^{k} \left| f\left(\xi_i^*\right) - f\left(\eta_i^*\right) \right| \left(z_i - z_{i-1}\right) \leq \varepsilon \sum_{i=1}^{k} \left(z_i - z_{i-1}\right) = \varepsilon\left(b-a\right).$$

Questo dimostra che *ogni successione di somme di Cauchy-Riemann di f è di Cauchy, e quindi convergente.*

2. Proviamo ora che, prese due diverse successioni di somme di Cauchy-Riemann di f, il loro limite (che già sappiamo esistere) è lo stesso. Per questo, considerate due successioni

$$s_n = \sum_{i=1}^{n} f\left(t_i\right)\left(x_i - x_{i-1}\right);$$

$$s_n^* = \sum_{i=1}^{n} f\left(t_i^*\right)\left(x_i - x_{i-1}\right)$$

con $t_i, t_i^* \in [x_{i-1}, x_i]$, sarà sufficiente provare che $s_n - s_n^* \to 0$. Il ragionamento è simile a quello del primo punto, ma più semplice, perché in questo caso confrontiamo due somme relative alla stessa suddivisione. Fissiamo $\varepsilon > 0$; poiché f è uniformemente continua e

$$\left|t_i - t_i^*\right| \leq \frac{b-a}{n},$$

per n abbastanza grande sarà $|f(t_i) - f(t_i^*)| < \varepsilon$ e quindi:

$$|s_n - s_n^*| = \left| \sum_{i=1}^{n} [f(t_i) - f(t_i^*)](x_i - x_{i-1}) \right| \leq$$
$$\leq \sum_{i=1}^{n} |f(t_i) - f(t_i^*)| \, |x_i - x_{i-1}| < \sum_{i=1}^{n} \varepsilon \, |x_i - x_{i-1}| = \varepsilon (b - a).$$

Questo prova che $s_n - s_n^* \to 0$. Pertanto tutte le successioni di somme di Cauchy-Riemann di f tendono allo stesso limite, ed f è integrabile. ◆

7 Modelli dinamici discreti

1 INTRODUZIONE ALLA MODELLISTICA

Da qualche decennio, la modellistica matematica, già tradizionalmente fondamentale in Fisica, è entrata pesantemente (forse anche inaspettatamente) in importanti rami dell'Ecologia, della Biologia, della Medicina e in modo ancor più rilevante in Economia e Finanza. Le ragioni di questo successo sono dovute alla speciale flessibilità e all'universalità degli strumenti matematici, spesso capaci (e a volte gli unici) di dare risposte soddisfacenti a problemi di crescente complessità, grazie alla potenza di calcolo di macchine sempre più sofisticate, oggi disponibili.

Che cos'è un *modello matematico*? Come si costruisce? Un modello matematico è un'interfaccia tra una parte del mondo reale e il mondo delle teorie matematiche. Supponiamo di aver isolato un fenomeno reale (per esempio il trasporto e/o la diffusione di un inquinante in un corso d'acqua) e di volerlo studiare allo scopo di fare previsioni. Se si vuole usare un modello matematico occorre anzitutto capire quali siano i fattori (tradotti in variabili matematiche) che riteniamo essenziali e caratteristici nell'evoluzione. Questa fase è probabilmente la più delicata. Per queste variabili si cercano poi le relazioni fondamentali che descrivono quantitativamente la dinamica. Solitamente si perviene a una o più equazioni di vario tipo che costituiscono il modello matematico. Si tratta poi di analizzare il modello, cercando di estrarne le informazioni desiderate.

Qual è la credibilità di un modello matematico? Dipende dalla sua bontà, determinata dalla dicotomia completezza/analizzabilità. La completezza, in generale, richiederebbe di tener conto di un numero troppo elevato di componenti e di relazioni tra esse il più realistiche possibile; d'altra parte, quanto maggiore è il numero di variabili e quanto più realistiche sono le relazioni tra esse, tanto più complicato è il modello, e minore la capacità di analisi. Un buon modello costituisce un efficace compromesso, ma proprio per questo motivo, occorrono sempre test a posteriori, usando casi già noti, per garantirne la validità predittiva. In ogni caso, si deve tener sempre presente che un modello matematico (ma questo vale per qualunque modello descrittivo di fenomeni reali) non ha nessuna pretesa di universalità e va sempre usato *cum grano salis*. Alcuni modelli permettono analisi e previsioni quantitative, come per esempio i modelli in Fisica, Chimica, Medicina, Finanza; in altri, tipicamente quelli

in Economia, le analisi sono forzatamente qualitative e vanno usate unicamente per testare l'influenza di alcuni fattori sull'evoluzione di un dato sistema.

Perché usare modelli matematici? Alcune caratteristiche li rendono particolarmente appetibili. Il loro basso costo, la loro flessibilità, legata alla possibilità praticamente illimitata di simulazioni ripetute al computer per descrivere l'andamento di fenomeni complessi, altrimenti incomprensibili. A volte sono necessari, essendo l'unico strumento utilizzabile. Un esempio notevole in questo senso è costituito dalla simulazione della circolazione sanguigna nel cosiddetto *circolo di Willis*, che assicura l'irrorazione nei due emisferi cerebrali, usato da medici oltreoceano per valutare l'effetto di un'eventuale occlusione in uno dei vasi. In questo caso, non è certo possibile fare esperimenti in loco!

Molta cautela e un atteggiamento sempre critico vanno impiegati nell'uso del computer, evitando il macroscopico errore, tipico dei principianti, di affermare "l'ha detto il computer e quindi deve essere così." Particolarmente vero è questo fatto per i modelli di tipo economico.

In generale, un modello matematico si costruisce traducendo in termini matematici *leggi generali* di evoluzione accoppiate con *leggi specifiche* (*costitutive*) del fenomeno da descrivere. I modelli di cui ci occuperemo in questo capitolo appaiono spesso nelle scienze applicate e sono costituiti da *equazioni alle differenze*, che traducono il monitoraggio dell'evoluzione ad intervalli di tempo regolari. Per questa ragione si chiamano anche *modelli dinamici a tempo discreto*. Nel secondo volume esamineremo modelli *a tempo continuo*, che conducono ad *equazioni differenziali*, in cui le incognite sono funzioni che appaiono insieme ad alcune loro derivate.

Vediamo ora come si procede nel caso di alcuni esempi classici che inoltre ci permetteranno di motivare lo sviluppo della teoria.

1.1 Modello di Malthus

È storicamente il primo modello per la dinamica delle *popolazioni*, dovuto a Malthus[1] nel 1798. Consideriamo una popolazione[2] che evolve isolata e i cui unici fattori di evoluzione sono *fertilità* e *mortalità*. Indichiamo con $x(t)$ il numero di individui presenti al tempo t e ci proponiamo di studiarne l'evoluzione a partire dall'istante (convenzionale) $t = 0$. Naturalmente $x(t)$ è un valore medio (per esempio, numero di individui per unità di area), quindi non necessariamente intero, che possiamo senz'altro ritenere reale. Siano λ il tasso di fertilità e μ il tasso di mortalità, cioè il numero di individui nati (rispettivamente, morti) nell'unità di tempo per ogni individuo già esistente. Ne segue che, in un intervallo di tempo di durata h, la percentuale di nuovi nati e di morti è, rispettivamente, λh e μh. Si noti che, a priori, λ e μ potrebbero dipendere da t e da $x(t)$.

La legge generale di evoluzione qui è semplice: *la variazione relativa del numero di individui in un intervallo di tempo di durata h è funzione di $(\lambda - \mu)h$.*

La legge costitutiva indica *che tipo di funzione si intende adottare.* Malthus assume che questa funzione sia esattamente $(\lambda - \mu)h$, con λ e μ costanti. Il tutto si traduce

[1] Thomas Robert Malthus (1766-1834), economista inglese.

[2] Non necessariamente di esseri umani: potrebbe, per esempio, trattarsi di una popolazione di virus, o della quantità di atomi radioattivi in una massa di materiale fissile.

allora facilmente in termini matematici:

$$\frac{x(t+h)-x(t)}{x(t)} = (\lambda - \mu)\, h. \tag{1.1}$$

Il numero $\varepsilon = \lambda - \mu$ si chiama *potenziale biologico*.

Pensiamo ora di monitorare l'evoluzione di x ad intervalli di tempo regolari. Questo modo di procedere corrisponde a considerare il tempo come variabile *discreta*. Prendendo allora come unità di misura temporale l'intervallo tra due osservazioni successive si ha $t \in \mathbb{N}$. Lo stato del sistema è quindi descritto dalla successione $x(0)$, $x(1)$, $x(2)$, ... Si usa anche scrivere x_0, x_1, x_2 e, in generale, x_n con n intero, al posto di $x(t)$. Sostituendo nella (1.1) e scegliendo $h = 1$ (l'intervallo di tempo più piccolo possibile, che potrebbe essere un mese, un anno, un giorno...) si ha

$$x_{n+1} - x_n = \varepsilon x_n. \tag{1.2}$$

La (1.2) connette l'incremento $x_{n+1} - x_n$ con lo stato al tempo n e per questa ragione si chiama *equazione (lineare) alle differenze*, a *un passo* o del *prim'ordine*. Spesso conviene scrivere la (1.2) nella forma

$$\boxed{x_{n+1} = (1+\varepsilon)\, x_n} \tag{1.3}$$

dove lo stato al tempo $n+1$ è espresso in funzione dello stato al tempo precedente n. La conoscenza della consistenza della popolazione all'istante iniziale $n = 0$ permette di determinare un'unica soluzione, che è

$$\boxed{x_n = M_0 (1+\varepsilon)^n} \tag{1.4}$$

dove $x_0 = M_0$ è la condizione iniziale.

Infatti, sostituendo $n = 0$ nella (1.3), si ottiene $x_1 = (1+\varepsilon) M_0$ e, successivamente, sostituendo $n = 1$, si ha

$$x_2 = (1+\varepsilon)\, x_1 = (1+\varepsilon)^2 M_0.$$

Procedendo in tal modo si ottiene

$$x_n = (1+\varepsilon)\, x_{n-1} = (1+\varepsilon)^2 x_{n-2} = (1+\varepsilon)^3 x_{n-3} = \cdots = (1+\varepsilon)^n M_0.$$

Alla condizione iniziale $M_0 = 0$ corrisponde la soluzione nulla. Essendo costante nel tempo si dice che è una *soluzione d'equilibrio (steady state)*. Se lasciamo variare la condizione iniziale M_0 su tutti i numeri reali, si ottiene la totalità delle soluzioni.

Le informazioni più interessanti per le applicazioni si ricavano dal comportamento asintotico delle soluzioni rispetto alla soluzione di equilibrio. Tecnicamente si tratta di calcolare il limite per $n \to +\infty$ di x_n.

Vediamo che cosa succede. Se $\varepsilon > 0$ allora $1 + \varepsilon > 1$ e $(1+\varepsilon)^n M_0 \to +\infty$ se $n \to +\infty$. In questo caso, dunque le soluzioni si allontanano dalla soluzione di equilibrio. Diremo allora che quest'ultima è *instabile*.

Se $-2 < \varepsilon < 0$ allora $-1 < 1+\varepsilon < 1$ e $(1+\varepsilon)^n M_0 \to 0$ se $n \to +\infty$ e la soluzione nulla si dice *asintoticamente stabile*; se $\varepsilon = -2$, allora $1+\varepsilon = -1$ e la soluzione oscilla tra i valori $-M_0$ e M_0 (soluzione nulla *stabile ma non asintoticamente*), mentre se $\varepsilon < -2$, allora $1+\varepsilon < -1$ e x_n è oscillante e illimitata (soluzione nulla *instabile*).

Preciseremo più avanti i concetti di instabilità, stabilità e stabilità asintotica.

1.2 Modello logistico

Il modello di Malthus è troppo semplicistico e irrealistico nella maggioranza dei casi. In tale modello si fa l'ipotesi che l'ambiente esterno non influenzi la crescita e, di conseguenza, che il tasso di crescita sia proporzionale alla popolazione presente, ossia che il tasso relativo di crescita sia costante. Una popolazione più numerosa comporta però minori risorse e ciò implica un minor tasso di crescita. Nel 1845 Verhulst[3] propose un modello nel quale la legge costitutiva di evoluzione prevede una *soglia di sopravvivenza* M che la popolazione non può superare, e prevede che la variazione relativa di $x(t)$ nell'intervallo di tempo h decresca linearmente come funzione di $x(t)$. Sotto queste condizioni la legge di evoluzione si può scrivere nel modo seguente:

$$x(t+h) - x(t) = \varepsilon h x(t) \left[1 - \frac{x(t)}{M}\right] \tag{1.5}$$

Scegliendo come prima $h = 1$, ponendo $t = n$, intero, e $x(n) = x_n$, dalla (1.5) si ottiene

$$x_{n+1} = x_n + \varepsilon x_n \left(1 - \frac{x_n}{M}\right) = (1+\varepsilon) x_n \left(1 - \frac{\varepsilon}{(1+\varepsilon) M} x_n\right)$$

che è un' *equazione alle differenze* del primo ordine, *non lineare*, che prende il nome di *equazione logistica discreta*. Rendiamo più leggibile l'equazione ponendo

$$x_n = \frac{(1+\varepsilon) M}{\varepsilon} y_n.$$

L'equazione per y_n è

$$\boxed{y_{n+1} = r y_n (1 - y_n)} \tag{1.6}$$

dove $r = 1 + \varepsilon$.

La successione $y_n = 0$ per ogni n (cioè la successione di zeri) è *soluzione di equilibrio*. Che cosa possiamo dedurre dalla (1.6) riguardo alle altre soluzioni? Anzitutto: vi sono altre soluzioni di equilibrio? Una soluzione di equilibrio è costante, per cui $y_{n+1} = y_n$. Dalla (1.6) otteniamo allora

$$y_{n+1} = r y_n (1 - y_n) = y_n$$

da cui

$$r(1 - y_n) = 1$$

e successivamente

$$y_n = 1 - \frac{1}{r}.$$

Questa soluzione è tra 0 e 1 se $r > 1$ e quindi, in questo caso vi sono due soluzioni di equilibrio. Ci si chiede se l'andamento delle altre soluzioni tenderà verso una delle due. Intanto, se il potenziale biologico ε è negativo ma non troppo, in modo che $0 < r = 1 + \varepsilon < 1$, e se y_n parte da y_0 tra 0 e 1, si ha subito

$$0 < y_{n+1} = r y_n (1 - y_n) < r y_n < r^2 y_{n-1} < \cdots < r^n y_0$$

e quindi $y_n \to 0$ per $n \to +\infty$. La soluzione nulla è asintoticamente stabile.

[3] Pierre Francois Verhulst, 1804-1849.

E se $\varepsilon > 0$? Allora $r > 1$ ed i due fattori ry_n e $(1 - y_n)$ entrano in competizione: la moltiplicazione per r fa aumentare rapidamente y_n, ma appena y_n si avvicina alla quota 1 il secondo fattore è molto piccolo e tende a diminuire il valore di y_n. Quale dei due fattori prevalga è molto difficile da stabilire *a priori*, senza contare il fatto che $1 - y_n$ potrebbe diventare negativo e la successione perderebbe il suo significato di numero (medio) di individui di una popolazione.

Proviamo a ricorrere al computer e a disegnare nel piano (n, y_n) i punti generati dalla logistica discreta per vari valori del parametro r. In figura 7.1 si vede quello che si ottiene con $y_1 = 0{,}2$, dando a r i valori 0,7 (circolo), 2,6 (rombo), 3,4 (crocetta), 4 (asterisco).

Figura 7.1. Sensibilità della logistica rispetto al parametro r.

Si osserva un'estrema sensibilità alle variazioni del parametro r ed una varietà, a priori insospettabile, di possibili comportamenti asintotici, non escluso quello che prende il nome di *comportamento caotico*. Per interpretare i grafici della figura 7.1 occorre un po' di teoria. Ce ne occuperemo più avanti.

1.3 Modello dell'acceleratore

Questo classico modello (P.A. Samuelson, 1938) costituisce un tentativo di spiegare elementarmente l'alternarsi di periodi d'espansione e di recessione nello sviluppo economico. Consideriamo un sistema economico osservato per un certo numero d'anni in sequenza. Indichiamo con Y_n il *prodotto nazionale* per l'anno n e con C_n il *consumo* nell'anno n. Nel modello vi sono tre leggi costitutive:

a) C_n è *funzione lineare* (*affine*[4]) di quanto prodotto nell'anno precedente:

$$C_n = aY_{n-1} + k, \tag{1.7}$$

dove a è un coefficiente di propensione al consumo, $0 < a < 1$ e k è un numero positivo che rappresenta un consumo fisso costante. Questa funzione è nota come *funzione del consumo*;

[4]Correttamente: $y = ax$ è *lineare*, $y = ax + b$ è *lineare affine*.

b) i produttori investono un ammontare I_n costituito da una parte autonoma $A > 0$ e da una parte proporzionale alla variazione dei consumi[5]:

$$I_n = A + b\,(C_n - C_{n-1})\,,$$

con $b > 0$.

c) Il livello di Y_n (ossia la decisione sulla produzione) è scelto in modo da soddisfare la domanda di beni (consumo + investimento). Si avrà allora

$$Y_n = C_n + I_n.$$

Sostituiamo in quest'ultima uguaglianza l'espressione precedente di I_n ed eliminiamo i consumi attraverso la (1.7); otteniamo:

$$Y_n - a\,(1+b)\,Y_{n-1} + abY_{n-2} = A + k. \tag{1.8}$$

La (1.8) è un'equazione alle differenze che connette il valore di Y_n con quelli di Y_{n-1} e Y_{n-2}; inoltre, il primo membro è un polinomio di primo grado in Y_n, Y_{n-1} e Y_{n-2}. Per questi motivi l'equazione si dice *lineare del second'ordine* (o *a due passi*).

Per le soluzioni di questo tipo di equazioni ricaveremo (v. par. 5) una formula generale da cui sarà facile dedurre il comportamento di Y_n per $n \to +\infty$.

2 GENERALITÀ SULLE EQUAZIONI ALLE DIFFERENZE

Abbiamo considerato modelli matematici che descrivono l'evoluzione di una variabile, detta *variabile di stato*, in funzione del tempo, misurato ad intervalli regolari. In tal caso, scegliendo opportunamente l'unità di misura temporale si può far in modo che la variabile temporale assuma i valori interi $0, 1, 2, \ldots$.

L'unità di misura temporale può essere un anno, un semestre, un mese, un giorno, ecc. In genere si sceglie come unità di misura l'intervallo di tempo tra due "osservazioni" successive.

La funzione che descrive l'evoluzione della variabile di stato è allora una successione $\{x_n\}$:

$$n \mapsto x_n \qquad n \in \mathbb{N}.$$

Gli ingredienti tipici di un modello dinamico discreto sono i seguenti. Dato $k \geq 1$, si assegna:

a) la *dinamica,* ossia *la legge d'evoluzione di* x , che esprime un legame tra il valore della variabile x al tempo $n+k$, in termini del valore di x ai tempi precedenti n, $n+1, \ldots, n+k-1$.

In formule, si ha un'equazione del tipo

$$x_{n+k} = f\,(n, x_{n+k-1}, \ldots, x_n)$$

che si chiama equazione *alle differenze di ordine* k o *a* k *passi.*

[5] L'idea sottostante è che se, per esempio, i consumi sono in crescita, per far fronte alla domanda che si manifesterà nel periodo successivo, è opportuno espandere la capacità produttiva. Viceversa nel caso di consumi in contrazione.

b) le *condizioni iniziali*, ossia i primi k valori di x, cio è $x_0, x_1, \ldots, x_{k-1}$.

Nei prossimi due paragrafi ci occuperemo delle equazioni alle differenze del prim'ordine o a un passo, del tipo

$$x_{n+1} = f(n, x_n), \quad n \in \mathbb{N} \tag{2.1}$$

con

$$x_0 \text{ assegnato.} \tag{2.2}$$

Se f non dipende esplicitamente da n, cioè

$$x_{n+1} = f(x_n), \quad n \in \mathbb{N}$$

l'equazione si dice *autonoma*.

Abbiamo assunto $n \in \mathbb{N}$, ma può succedere che l'equazione abbia senso da un certo intero n_0 in poi, nel qual caso si assegna x_{n_0} come condizione iniziale.

Per alcuni tipi di equazione, sono disponibili formule esplicite per la totalità delle soluzioni, che chiameremo *soluzione generale*. Un caso importante è quello delle equazioni lineari a coefficienti costanti, che analizzeremo nel paragrafo 3.

Osserviamo subito che se la funzione f risulta ben definita in (n, x_n) per ogni $n \geq 0$, *esistenza ed unicità* della soluzione del problema (2.1), (2.2) sono assicurate.

L'analisi degli equilibri ed il comportamento asintotico per $t \to +\infty$ possono invece presentare aspetti notevolmente complessi.

Il caso particolare delle equazioni *non lineari* autonome, ci servir à per introdurre gli importanti concetti di punto *fisso* o *di equilibrio*, di *orbite periodiche* e di *caos deterministico* (par. 4).

Infine, nel paragrafo 5 ci occuperemo invece delle equazioni alle differenze del 2° ordine, lineari a coefficienti costanti.

3 EQUAZIONI LINEARI DEL PRIM'ORDINE A COEFFICIENTI COSTANTI

Un'equazione *lineare del prim'ordine a coefficienti costanti* è del tipo

$$x_{n+1} = ax_n + b, \qquad n \in \mathbb{N} \tag{3.1}$$

dove $a, b \in \mathbb{R}$. Se $b = 0$ l'equazione si dice *omogenea*. Per queste equazioni è disponibile una formula esplicita per la soluzione generale.

Iniziamo con il caso dell'equazione *omogenea associata*

$$w_{n+1} = aw_n,$$

che corrisponde al modello di Malthus, con $a = 1 + \varepsilon$. Come abbiamo già visto, la soluzione generale è

$$\boxed{w_n = ca^n, \qquad c \in \mathbb{R}}$$

dove il parametro c coincide con il valore iniziale della variabile di stato: $c = w_0$.

Il caso non omogeneo non è molto più complicato. Infatti, supponiamo di conoscere una soluzione particolare $\{y_n\}$ della (3.1). Data un'altra soluzione qualsiasi $\{x_n\}$ e posto $w_n = x_n - y_n$, si ha:

$$\begin{aligned} w_{n+1} &= x_{n+1} - y_{n+1} = ax_n + b - (ay_n + b) = \\ &= a(x_n - y_n) = aw_n. \end{aligned}$$

Pertanto, w_n è soluzione dell'equazione omogenea associata e quindi $w_n = ca^n$. Deduciamo che

$$x_n = ca^n + y_n$$

ossia:

La soluzione generale della (3.1) si ottiene come somma della soluzione generale dell'omogenea associata e di una soluzione particolare della non omogenea.

Il problema è così ricondotto alla ricerca di una soluzione particolare della (3.1).

Esendo b costante, una soluzione particolare si trova facilmente. Infatti, provando con una successione costante $\overline{x}_n = k$, deve essere, dato che $\overline{x}_{n+1} = k$,

$$k = ak + b.$$

Se $\boxed{a \neq 1}$ si trova

$$\overline{x}_n = k = \frac{b}{1-a}$$

che, essendo costante, è una *soluzione di equilibrio*. Di conseguenza, la soluzione generale dell'equazione

$$x_{n+1} = ax_n + b, \qquad a, b \in \mathbb{R},\ a \neq 1$$

è

$$\boxed{x_n = ca^n + \frac{b}{1-a}, \qquad c \in \mathbb{R}} \tag{3.2}$$

e la soluzione che "parte" da x_0 è

$$\boxed{x_n = \left(x_0 - \frac{b}{1-a}\right)a^n + \frac{b}{1-a}.} \tag{3.3}$$

Infatti, sostituendo nella (3.2) $n = 0$, si trova

$$x_0 = c + \frac{b}{1-a}$$

e quindi

$$c = x_0 - \frac{b}{1-a}.$$

Se $\boxed{a = 1}$ da

$$x_{n+1} = x_n + b$$

sostituendo $n = 0, 1, 2, \ldots$ si ha

$$\begin{aligned} x_1 &= x_0 + b, \\ x_2 &= x_1 + b = x_0 + 2b, \\ x_3 &= x_2 + b = x_0 + 3b, \\ &\vdots \end{aligned}$$

e si ricava facilmente

$$\boxed{x_n = x_0 + nb}$$

Esempi

3.1 Cerchiamo la soluzione dell'equazione alle differenze

$$x_{n+1} = \frac{1}{2}x_n + 1,$$

che soddisfa la condizione iniziale $x_0 = 1$. La soluzione generale dell'equazione omogenea corrispondente è

$$z_n = c\left(\frac{1}{2}\right)^n$$

mentre la soluzione di equilibrio è (qui $a = 1/2 \neq 1$) $\overline{x}_n = 2$ per cui la soluzione generale è

$$x_n = c\left(\frac{1}{2}\right)^n + 2, \quad c \in \mathbb{R}. \tag{3.4}$$

Per determinare la soluzione particolare che verifica la condizione iniziale assegnata determiniamo la costante c, sostituendo il valore 0 nella (3.4); si ha

$$x_0 = 1 = c + 2 \quad \Rightarrow \quad c = -1.$$

La soluzione cercata è quindi

$$x_n = 2 - \left(\frac{1}{2}\right)^n.$$

Si controlla facilmente che $x_n \to 2$ per $n \to +\infty$.

3.2 *Capitalizzazione semplice e composta.* Un capitale C, investito in capitalizzazione semplice con tasso mensile d'interesse semplice r, ogni mese frutta interessi pari a rC. Indicando con C_n la somma disponibile dopo n mesi e con C_{n+1} la somma disponibile al mese successivo, si ha

$$C_{n+1} = C_n + rC. \tag{3.5}$$

Se la stessa somma fosse investita in capitalizzazione composta con tasso mensile d'interesse composto i, indicando sempre con C_n la somma disponibile dopo n mesi e con C_{n+1} la somma disponibile al mese successivo, si avrebbe, essendo gli interessi calcolati non sul capitale iniziale ma sul capitale maturato,

$$C_{n+1} = C_n + iC_n = (1 + i)\, C_n. \tag{3.6}$$

La (3.5) e la (3.6) sono equazioni lineari a coefficienti costanti, la seconda è anche omogenea. Per le due equazioni (3.5) e (3.6), tenendo conto della condizione iniziale $C_0 = C$ (il capitale in data 0 è C), si ottengono, rispettivamente, le soluzioni

$$C_n = C\,(1 + rn)\,, \qquad C_n = C\,(1 + i)^n\,.$$

3.2 *Modello della ragnatela.* La quantità domandata e la quantità offerta d'un bene dipendono dal prezzo dello stesso. In un semplice modello che descrive la formazione del prezzo di mercato, si assume che domanda e offerta siano funzioni lineari del prezzo.

Consideriamo prima il caso statico, in cui il prezzo p, e quindi anche domanda e offerta siano costanti. Se D è la quantità domandata, O la quantità offerta e a, b, c, d parametri positivi[6], si ha

$$D\,(p) = a - bp, \qquad O\,(p) = -c + dp.$$

Il mercato è in equilibrio se $D = O$ ovvero se

$$a - bp = -c + dp$$

da cui s'ottiene il *prezzo* p^* *d'equilibrio*

$$p^* = \frac{a + c}{b + d}.$$

Supponiamo ora che il prezzo e, quindi, la domanda e l'offerta siano funzioni del tempo. Mentre la domanda s'adegua immediatamente al prezzo, l'offerta richiede un po' di tempo. La quantità offerta dipende, in particolare, dal prezzo nel periodo temporale precedente. Scegliamo come unità di tempo il "ritardo" nell'adeguamento dell'offerta. Per molti prodotti, per esempio agricoli, che vengono seminati un anno prima dell'immissione sul mercato, l'offerta dipende dal prezzo dell'anno precedente. In tal caso l'unità di tempo è un anno. Siano, quindi,

$$\begin{cases} D_n = a - bp_n \\ O_n = -c + dp_{n-1} \end{cases}$$

rispettivamente, la funzione di domanda e la funzione di offerta (con a, b, c, d parametri positivi). All'equilibrio sarà $D_n = O_n$, ovvero

$$a - bp_n = -c + dp_{n-1}.$$

Risolvendo rispetto a p_n s'ottiene

$$p_n = -\frac{d}{b}p_{n-1} + \frac{a+c}{b},$$

equazione alle differenze del primo ordine lineare a coefficienti costanti, la cui soluzione è data (poiché il primo coefficiente è diverso da 1) dalla formula (3.3):

$$p_n = \left(-\frac{d}{b}\right)^n (p_0 - p^*) + p^*,$$

dove p_0 indica il prezzo iniziale e $p^* = \dfrac{a + c}{b + d}$ è il prezzo di equilibrio del mercato in condizioni statiche.

Osserviamo che se $p_0 = p^*$, la soluzione è costante e coincide con il prezzo di equilibrio: $p_n = p^*$ per ogni n. Se $p_0 \neq p^*$, abbiamo una dinamica oscillante: il prezzo p_n oscilla attorno a p^*.

[6] Il parametro a rappresenta il *mercato potenziale*. Il parametro c può anche essere nullo.

L'andamento per $n \to +\infty$ dipende dal rapporto d/b. Se $d > b$ le oscillazioni s'amplificano, se $d < b$ le oscillazioni si smorzano. Se $d = b$ l'ampiezza delle oscillazioni rimane costante. *La dinamica converge verso l'equilibrio se e solo se* $d < b$, cioè se la pendenza della curva d'offerta è minore del modulo della pendenza della curva di domanda:

$$\boxed{\lim_{n\to+\infty} p_n = p^* \qquad \text{se e solo se} \quad d < b}$$

La dinamica del sistema può essere illustrata graficamente: si tracciano le due rette d'equazione

$$D = a - bp \quad \text{e} \quad O = -c + dp,$$

che rappresentano, rispettivamente, la funzione di domanda e la funzione d'offerta. Si colloca p_0 sull'asse p e si determina l'ordinata O_1 sulla retta di offerta. Spostandosi orizzontalmente sulla retta di domanda si determina D_1 (essendo $O_1 = D_1$): l'ascissa corrispondente è p_1. Trovato p_1 si ripete il percorso precedente e si determinano O_2, poi $D_2 = O_2$ e quindi p_2. Procedendo in questo modo, s'ottengono le rappresentazioni sull'asse p dei termini della successione p_n.

Per visualizzare la dinamica del sistema, si può "automatizzare" la costruzione geometrica, procedendo nel modo seguente. Si parte da p_0 sull'asse p, ci si sposta in verticale sulla retta O, poi in orizzontale sulla retta D, poi in verticale su O, poi in orizzontale su D e così via. Nel caso di convergenza, si traccia un cammino a "ragnatela", che giustifica il nome del modello (v. fig. 7.2).

Figura 7.2. Modello della ragnatela.

4 EQUAZIONI AUTONOME NON LINEARI

4.1 Orbite, diagrammi a gradino e punti fissi o d'equilibrio

Tra i modelli discreti non lineari ci limitiamo a considerare equazioni alle differenze *autonome* del primo ordine del tipo:

$$x_{n+1} = f(x_n), \qquad n \geq 0. \tag{4.1}$$

Assegnando x_0, attraverso la relazione di ricorrenza (4.1) è possibile calcolare $x_1 = f(x_0)$, $x_2 = f(x_1)$,... e ottenere così la successione $\{x_n\}$, che prende il nome di *orbita uscente da* x_0.

La funzione f si chiama *generatrice* e, affinché la successione x_n sia ben definita, occorre che

$$f : I \to I$$

dove $I \subseteq \mathbb{R}$ è un intervallo. Infatti, la (4.1) richiede ogni volta di calcolare f su uno dei suoi *output*, secondo la sequenza di calcolo raffigurata sotto:

$$x_0 \to \boxed{f} \to x_1 \to \boxed{f} \to x_2 \to \cdots \to x_n \to \boxed{f} \to x_{n+1} \cdots$$

Assumeremo che f sia *continua* in I. Dato che

$$\begin{aligned} x_1 &= f(x_0), \\ x_2 &= f(x_1) = f(f(x_0)) = (f \circ f)(x_0), \\ x_3 &= f(x_2) = f(f(f(x_0))) = (f \circ f \circ f)(x_0), \\ &\cdots \end{aligned}$$

indicando con f^n l'iterata $n-$esima di f cioè $f^n = \underbrace{f \circ f \circ \cdots \circ f}_{\text{n volte}}$, si ha

$$\boxed{x_n = f^n(x_0)}$$

che evidenzia il punto di partenza dell'orbita.

Osservazione Sottolineiamo che il simbolo f^n indica l'iterata $n-$esima di f. Occorre evitare confusioni tra

$$f^n(x) \quad \text{e} \quad [f(x)]^n.$$

L'andamento di un'orbita si può efficacemente visualizzare tramite il cosiddetto *diagramma di fase* o *diagramma a gradino*. Vediamo come si costruisce tale diagramma per l'equazione

$$x_{n+1} = f(x_n).$$

Figura 7.3. Diagramma a gradino.

Disegnamo il grafico di f e la bisettrice $y = x$ e, in riferimento alla figura 7.3, cerchiamo di visualizzare l'orbita uscente da x_0.

Il punto sul grafico di f con ascissa x_0 ha ordinata $f(x_0) = x_1$, il successivo valore assunto dalla variabile di stato. Se sfruttiamo la bisettrice possiamo "riportare" x_1 sull'asse delle ascisse ed individuare $x_2 = f(x_1)$.

Procedendo in questo modo, connettendo ogni volta con un segmento i punti di coordinate

$$(x_0, x_1), (x_1, x_1), (x_1, x_2), \ldots, (x_n, x_{n+1}), (x_{n+1}, x_{n+1}), \ldots$$

si visualizza l'orbita mediante una spezzata.

Dopo un po' d'esercizio la costruzione diventa automatica: si prende un punto iniziale sull'asse x, si cerca la corrispondente ordinata sulla curva, da lì ci si sposta orizzontalmente verso la bisettrice, poi si va verticalmente alla curva, poi alla bisettrice e così via.

I punti come x^* in figura sono particolarmente importanti. Infatti se il punto iniziale x_0 coincide con x^*, *l'intera orbita coincide con* x^* e genera perciò una successione costante, ossia una soluzione di equilibrio. Si dice anche che x^* è *punto di equilibrio* o *punto fisso*. L'ultima locuzione è dovuta al fatto che

$$f(x^*) = x^*$$

e cioè il punto x^* rimane *fisso* sotto l'azione di f.

Il diagramma ottenuto in tal modo, evidenziando i punti fissi e tracciando le spezzate che visualizzano l'andamento delle orbite, prende il nome di *diagramma a gradino*.

Esempio

4.1 Nel modello logistico

$$x_{n+1} = rx_n(1 - x_n) \qquad (r > 0)$$

abbiamo

$$f(x) = rx(1 - x)$$

il cui grafico è una parabola. I punti fissi si trovano risolvendo l'equazione

$$rx(1 - x) = x$$

da cui

$$x_1^* = 0 \quad \text{e} \quad x_2^* = 1 - \frac{1}{r}.$$

4.2 Punti fissi e stabilità

Dai diagrammi a gradino si possono ricavare informazioni utili sul comportamento asintotico delle orbite. Sempre nel caso in figura 7.3, si vede che l'orbita uscente da x_0 converge al punto fisso x^*. Provando con altri punti di partenza, anche alla destra dell'equilibrio, s'osserva lo stesso fenomeno[7]: x^* è *asintoticamente stabile*. Precisamente, introduciamo i seguenti concetti.

[7] Invitiamo il lettore a controllare.

Definizione 7.1 (Stabilità, instabilità e stabilità asintotica) Il punto di equilibrio x^* si dice:

a) *stabile* se: per ogni $\varepsilon > 0$ esiste $\delta = \delta_\varepsilon$ tale che se $|x_0 - x^*| < \delta$ e x_n è l'orbita uscente da x_0, si ha

$$|x_n - x^*| < \varepsilon \qquad \text{per ogni } n \geq 0.$$

b) *asintoticamente stabile* se è stabile e, inoltre, esiste δ_1 tale che, se $|x_0 - x^*| < \delta_1$ e x_n è l'orbita uscente da x_0, si ha

$$x_n \to x^* \qquad \text{per } n \to +\infty.$$

c) *instabile* se non è stabile (se cioè non vale la condizione (a)). ■

Nel caso in cui x^* sia asintoticamente stabile, dalla definizione si deduce che esiste un intorno U di x^* tale che ogni orbita uscente da U converge a x^*. Il più grande insieme $B \supseteq U$ con la proprietà che ogni orbita uscente da B converga a x, si chiama *bacino di attrazione* di x^*.

Esempio

4.2 *Il caso lineare.* Nel caso lineare

$$x_{n+1} = ax_n + b, \qquad a, b \in \mathbb{R},\ a \neq 0,$$

se $a \neq 1$, l'unico punto di equilibrio è

$$x^* = \frac{b}{1-a}.$$

Dato che la soluzione generale è

$$x_n = ca^n + \frac{b}{1-a}, \qquad c \in \mathbb{R}$$

si ricava facilmente che per ogni altra soluzione (non costante), si ha:

$$\lim_{n\to+\infty} x_n = \begin{cases} \dfrac{b}{1-a} & \text{se } |a| < 1 \\ \pm\infty \text{ (col segno di } c) & \text{se } a > 1 \\ \text{non esiste} & \text{se } a \leq -1 \end{cases}$$

quindi il punto di equilibrio è (globalmente) asintoticamente stabile se se solo se $|a| < 1$.

Se $a = 1$ e $b \neq 0$ si ha

$$x_{n+1} = x_n + b$$

che non ha punti fissi.

Se $a = 1$ e $b = 0$ l'equazione è

$$x_{n+1} = x_n$$

ed ogni punto è d'equilibrio.

I diagrammi a gradino confermano i risultati trovati poco sopra (v. fig. 7.4).

Figura 7.4. Diagrammi a gradino nel caso lineare.

4.3 Stabilità per linearizzazione

In qualche caso, attraverso il diagramma a gradino si può capire l'andamento asintotico di x_n, anche senza conoscere l'espressione analitica di f. A volte bastano proprietà qualitative del grafico per convincerci che se il grafico di f non è "troppo ripido" vicino al punto di equilibrio, quest'ultimo è asintoticamente stabile. "Molto ripido" o "poco ripido" s'intendono rispetto alla bisettrice, che ha pendenza unitaria, quindi è decisivo il confronto del modulo della pendenza del grafico di f con la pendenza 1. Sostanzialmente l'idea è riportarsi al caso lineare sostituendo il grafico di f con la retta tangente ad esso nel punto fisso (x^*, x^*).

Questa tecnica prende perciò il nome di *linearizzazione*. Precisamente, vale il seguente criterio:

Teorema 7.1 *Se $f : I \longrightarrow \mathbb{R}$ è derivabile con continuità in I, $f(x^*) = x^*$ e $|f'(x^*)| \neq 1$, allora:*

se $|f'(x^*)| < 1$, x^* è punto di equilibrio asintoticamente stabile;
se $|f'(x^*)| > 1$, x^* è punto di equilibrio instabile.

Dimostrazione. Se $|f'(x^*)| < 1$, essendo f' continua, per il teorema della permanenza del segno, in un intorno $I_\delta = [x^* - \delta, x^* + \delta]$ di x^* si ha,

$$|f'(x)| \leq k < 1.$$

Poiché $f(I_\delta) \subseteq I_\delta$ (il lettore lo dimostri!), se $x_0 \in I_\delta$, anche $x_n \in I_\delta$. Valutiamo la distanza tra x_n e x^*; dall'equazione

$$x_{n+1} = f(x_n)$$

e dal fatto che $f(x^*) = x^*$ si ha

$$|x_{n+1} - x^*| = |f(x_n) - f(x^*)|.$$

Usando il teorema del valor medio di Lagrange e il fatto che $|f'| \leq k$, possiamo scrivere

$$|f(x_n) - f(x^*)| = |f'(c_n)(x_n - x^*)| \leq k|x_n - x^*|$$

da cui

$$|x_{n+1} - x^*| \leq k|x_n - x^*|.$$

Rifacendo gli stessi calcoli per valutare $|x_n - x^*|$ si trova

$$|x_n - x^*| \leq k|x_{n-1} - x^*| \leq k^2|x_{n-2} - x^*|.$$

Procedendo ancora in questo modo si arriva a

$$|x_n - x^*| \leq k|x_{n-1} - x^*| \leq k^2|x_{n-2} - x^*| \leq \cdots \leq k^n|x_0 - x^*|.$$

Essendo $k < 1$ si conclude che $|x_n - x^*| \to 0$, ossia $x_n \to x^*$, purché $x_0 \in I$.

Se $|f'(x^*)| > 1$, si ha, in un intorno I,

$$|f'(x)| \geq \rho > 1$$

e quindi se $x_0 \in I$, si ha, applicando il teorema del valor medio,

$$|x_1 - x^*| = |f(x_0) - f(x^*)| = |f'(c)||x_0 - x^*| \geq \rho|x_0 - x^*|;$$

se anche $x_1 \in I$, con lo stesso ragionamento, si trova

$$|x_2 - x^*| \geq \rho^2|x_0 - x^*|$$

e dopo un numero finito di passi si esce da I . Cioè x^* è instabile[8]

Esempio

7.5 Studiamo il carattere della successione per ricorrenza

$$\begin{cases} s_{n+1} = 3 - e^{-s_n} \\ s_0 = a. \end{cases}$$

La funzione generatrice della successione è

$$f(x) = 3 - e^{-x}.$$

Come mostra il grafico, f ha due punti fissi $\alpha < 0$ e β $(2 < \beta < 3)$. Dato che $f'(x) = e^{-x}$, si ha

$$f'(\alpha) > 1, \qquad 0 < f'(\beta) < 1.$$

Dunque α è punto fisso instabile (repulsore), β è punto fisso (localmente) asintoticamente stabile.

Se $a < \alpha$, la successione s_n è decrescente[9] e tende a $-\infty$, se $\alpha < a < \beta$ è crescente e tende a β, se $a > \beta$ è decrescente e tende a β. Ovviamente se $a = \alpha$ o $a = \beta$, s_n è costante. La figura 7.5 mostra che il bacino d'attrazione del punto β è $(\alpha, +\infty)$.

[8]Prima o poi, tutte le orbite uscenti da I escono da I, allontanandosi da x^*. In tal caso si dice che x^* è *repulsore*.

[9]Se la funzione generatrice è crescente, la succesione per ricorrenza è monotona (vedi l'esercizio 6 alla fine del par. 4).

4.4 Orbite periodiche e stabilità

Le *orbite periodiche* sono del tipo

$$x_0, x_1, \ldots x_{p-1}, \boxed{x_p = x_0}, x_1, \ldots x_{p-1}, x_0, \ldots$$

dove i primi p punti sono tutti diversi tra loro $(x_0 \neq x_1 \neq \cdots \neq x_{p-1})$, mentre $x_p = x_0$ e si ricomincia da capo. I loro punti si dicono *punti periodici* e il numero p si chiama *periodo* dell'orbita. Osserviamo che

$$x_0 = x_p = f^p(x_0)$$

quindi x_0 è *punto fisso per l'iterata* f^p. Nello stesso modo si verifica che anche $x_1, \ldots x_{p-1}$ sono punti fissi per f^p:

$$x_j = f^p(x_j) \quad \text{per ogni} \quad j = 1, \ldots, p-1.$$

Per esempio, un'*orbita di periodo* 2 è un insieme di due punti

$$\{x^*, y^*\}$$

entrambi fissi per l'iterata seconda f^2 e si ha

$$f(x^*) = y^*, \quad f(y^*) = x^*$$

Esempio

4.4 La successione $\left\{2, \frac{1}{2}, 2, \frac{1}{2}, 2, \frac{1}{2}, \ldots\right\}$ è un'orbita di periodo 2 per l'equazione

$$x_{n+1} = \frac{1}{x_n}.$$

Il diagramma a gradino di un'orbita di periodo 2 è particolarmente semplice (v. fig. 7.7).

Figura 7.7. Orbita di periodo 2.

Anche per le orbite periodiche si può parlare di stabilità, stavolta con l'idea che se si parte abbastanza vicino ad uno dei suoi punti si continua a rimanere vicino, non necessariamente a quel punto, ma all'orbita.

Precisamente, consideriamo l'equazione (4.1) (con f derivabile con continuità in un intervallo $I \subseteq \mathbb{R}$). Sia Γ_p un'orbita periodica di periodo p, costituita dai punti

$$w_0, w_1 = f(w_0), \ldots, w_{p-1} = f^{p-1}(w_0).$$

Per *distanza di un punto* z da Γ_p intendiamo la più piccola tra le distanze di z da un punto di Γ:

$$\operatorname{dist}(z, \Gamma_p) = \min\{|z - w_0|, |z - w_1|, \ldots, |z - w_{p-1}|\}.$$

Definizione 7.2 (Stabilità, instabilità e stabilità asintotica) L'orbita Γ_p si dice:

a) *stabile* se: per ogni $\varepsilon > 0$ esiste un insieme aperto A che contiene Γ_p e tale che, se $x_0 \in A$ e x_n è l'orbita uscente da x_0, si ha

$$\operatorname{dist}(x_n, \Gamma_p) < \varepsilon \qquad \text{per ogni } n \geq 0.$$

b) *asintoticamente stabile* se è stabile e, inoltre, esiste un insieme A_1, costituito da un unione di intervalli aperti, tale che $\Gamma_p \subset A_1$ e, se $x_0 \in A_1$ si ha

$$\operatorname{dist}(x_n, \Gamma_p) \to 0 \qquad \text{per } n \to +\infty.$$

c) *instabile* se non è stabile (se cioè non vale la condizione (a)). ■

Usare la definizione in generale non è semplice. C'è però un modo piuttosto efficace per assicurarsi che un'orbita periodica sia asintoticamente stabile.

Sappiamo infatti che ciascuno dei punti w_j $(j = 0, \ldots p-1)$ è punto fisso per l'iterata f^p. Notiamo che la *derivata prima dell'iterata f^p è uguale in tutti i punti w_j*, infatti si ha

$$(f^p)'(x) = f'(x) \cdot f'(f(x)) \cdot \cdots \cdot f'(f^{p-2}(x)) \cdot f'(f^{p-1}(x))$$

e in w_j:

$$(f^p)'(w_j) = f'(w_j) \cdot f'(w_{j+1}) \cdot \cdots \cdot f'(w_{j+p-1}) = \prod_{j=0}^{p-1} f'(w_j).$$

Si deduce che i punti dell'orbita sono *simultaneamente tutti iperbolici* oppure *tutti non iperbolici*. Pertanto se in uno dei punti w_j (basta controllarne uno qualsiasi!) si ha

$$\left|(f^p)'(w_j)\right| < 1$$

allora tutti i punti dell'orbita sono asintoticamente stabili per f^p e quindi l' *orbita stessa è asintoticamente stabile.* Analogamente, se in uno dei punti w_j si ha

$$\left|(f^p)'(w_j)\right| > 1$$

allora tutti i punti dell'orbita sono instabili e quindi l' *orbita stessa è instabile.*

 © 978-88-08-06485-1

4.5 Esistenza di orbite periodiche

La presenza di orbite periodiche rivela un andamento ciclico del sistema dinamico ed è piuttosto importante stabilire condizioni che ne garantiscano l'esistenza e, se possibile, la loro natura stabile/instabile. I due teoremi seguenti danno una risposta parziale al problema. Il primo si applica a funzioni generatrici cosiddette *unimodali*, che presentano cioè un unico punto stazionario nell'intervallo di definizione (come per la logistica).

TEOREMA 7.2 (DI SINGER) *Sia*

$$f : [0,1] \to [0,1]$$

continua con le sue derivate fino alla terza inclusa. Se

1. $$f(0) = f(1) = 0,$$

2. *f ha un solo punto $x_0 \in (0,1)$ tale che*

$$f'(x_0) = 0,$$

3. *la quantità*

$$\{f, x\} = \frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2$$

è negativa in $[0,1]$, allora l'equazione alle differenze

$$x_{n+1} = f(x_n)$$

ammette al massimo un'unica orbita periodica asintoticamente stabile, contenuta in $(0,1)$. Se tale orbita esiste, x_0 è nel suo bacino d'attrazione.

La quantità $\{f, x\}$ si chiama *derivata di Schwarz di f* (la definizione risale al 1869!). Il fatto che x_0 appartenga al bacino di attrazione dell'eventuale orbita periodica è un'informazione preziosa qualora si disponga di un computer per individuare tale orbita. In particolare, il teorema di Singer si applica al caso dell'equazione logistica, come vedremo più avanti.

Il secondo teorema, interessante se non sorprendente per la sua originalit à, richiede l'introduzione di uno speciale ordinamento dei numeri naturali, detto *ordinamento di Sharkowsky*. Prima si collocano i numeri dispari, poi il doppio di tutti i dispari, poi il quadruplo di tutti i dispari e così via. Infine si collocano le potenze di due in ordine decrescente. Se utilizziamo il simbolo $\blacktriangleleft$ per indicare che, nella formula

$$m \blacktriangleleft n$$

n segue m, l'ordinamento si può scrivere nel modo seguente:

$$3 \blacktriangleleft 5 \blacktriangleleft 7 \blacktriangleleft \cdots \blacktriangleleft 2\cdot 3 \blacktriangleleft 2\cdot 5 \blacktriangleleft 2\cdot 7 \blacktriangleleft \cdots \blacktriangleleft 2^2\cdot 3 \blacktriangleleft 2^2\cdot 5 \blacktriangleleft 2^2\cdot 7 \blacktriangleleft \cdots$$
$$\cdots \blacktriangleleft 2^3\cdot 3 \blacktriangleleft 2^3\cdot 5 \blacktriangleleft 2^3\cdot 7 \blacktriangleleft \cdots \blacktriangleleft 2^4 \blacktriangleleft 2^3 \blacktriangleleft 2^2 \blacktriangleleft 2 \blacktriangleleft 1$$

Teorema 7.3 (di Sharkowsky) *Sia $f : \mathbb{R} \to \mathbb{R}$ continua. Se l'equazione alle differenze*

$$x_{n+1} = f(x_n)$$

ammette un'orbita periodica di periodo m, allora, per ogni n che segue m nell'ordinamento di Sharkowsky esiste un'orbita periodica di periodo n.

Non si può evitare di sottolineare una (clamorosa!) conseguenza del teorema: *se esiste un'orbita di periodo 3 esistono orbite di qualunque periodo.* Il periodo 3 è sinonimo di ... caos, come vedremo.

Non solo: si può esibire un esempio di funzione generatrice continua il cui sistema dinamico corrispondente ammette un'orbita di periodo 5 ma non di periodo 3, a conferma della profondità del teorema. Si può usare il teorema anche per mostrare che *non* possono esistere orbite di un certo periodo. Per esempio, consideriamo la logistica con

$$f(x) = 3{,}2x(1-x)$$

e disegnamo i grafici di f, dell'iterata seconda f^2 e dell'iterata quarta f^4 (v. fig. 7.8).

Figura 7.8. Grafici di f, f^2, f^4.

Il primo grafico mostra i due punti fissi 0 e 0,7.... Nel secondo, osserviamo ancora i due punti fissi di f e altri due punti fissi di f^2 che costituiscono un'orbita periodica di periodo 2. Nel terzo i punti fissi coincidono con quelli di f^2. Ne segue che *non* vi sono orbite di periodo 4 e quindi, per il teorema di Sharkowsky, non possono esistere orbite periodiche di periodo maggiore di 2.

4.6 Comportamento caotico

Accenniamo brevemente al concetto di *caos deterministico*. La parola *caos* è collegata spesso a mancanza di regolarità, di ordine o di qualche correlazione. Da un punto di vista dinamico, caos significa impossibilità di misurare, osservare o prevedere le variazioni di un sistema in evoluzione, entrando, inevitabilmente, in una condizione di aleatorietà. In realtà vi sono vari tipi di caos e quello che tratteremo qui, brevemente, non si riferisce all'osservazione di fenomeni aleatori (come il lancio di un dado) bensì a sistemi che evolvono secondo *leggi deterministiche*, di cui le equazioni che stiamo considerando costituiscono un esempio.

Nel nostro caso parleremo di *caos deterministico* e della conseguente *dinamica caotica*. Uno dei problemi che si presenta è caratterizzare in modo rigoroso che cosa si debba intendere con ciò. La questione non è così semplice, tant'è vero che esistono varie definizioni più o meno soddisfacenti di comportamento caotico. Adottiamo quella di Li e Yorke.

Sia $f : I \to I$. L'idea è che lo stato caotico sia determinato dall'esistenza di un insieme non numerabile $S \subset I$ con la proprietà seguente: per ogni coppia di punti $p, q \in S$, esistono due successioni di iterate f^{n_j} e f^{m_k} tali che

$$\lim_{n_j \to +\infty} |f^{n_j}(p) - f^{n_j}(q)| = 0$$

e

$$\lim_{m_k \to +\infty} |f^{m_k}(p) - f^{m_k}(q)| = \delta > 0.$$

In altri termini, le orbite $f^n(p)$ e $f^n(q)$, uscenti da punti qualunque p e q in S, hanno un andamento erratico, nel senso che per qualche indice n sono molto vicine mentre per certi altri si mantengono a distanza positiva. Naturalmente ciò significa che la dinamica esibisce estrema sensibilità rispetto alle condizioni iniziali: una piccolissima variazione in queste ultime provoca, in generale, notevoli differenze nelle orbite corrispondenti. Se una dinamica è caotica, le previsioni sul suo andamento eseguite al computer non possono essere credibili!

Un risultato ormai classico è il seguente:

Teorema 7.4 (di Li e Yorke) *Periodo tre implica caos.*

L'equazione logistica

$$x_{n+1} = 4x_n(1 - x_n)$$

fornisce un importante e classico esempio di dinamica caotica. La esaminiamo in qualche dettaglio nella prossima sezione.

4.7 Equazione logistica discreta

In questa sezione analizziamo il comportamento asintotico delle orbite dell'equazione logistica

$$x_{n+1} = rx_n(1 - x_n), \qquad r > 0. \tag{4.2}$$

La funzione generatrice è

$$f(x) = rx(1 - x)$$

Se limitiamo, come faremo, i valori del parametro r nell'intervallo $(0, 4]$, si ha (il lettore controlli) $f([0, 1]) \subseteq [0, 1]$ ossia $f : [0, 1] \to [0, 1]$.

Osserviamo subito che f soddisfa le ipotesi del teorema di Singer; infatti, $f(0) = f(1) = 0$, ha un solo punto stazionario $x = 1/2$ e

$$\{f, x\} = \frac{f'''(x)}{f'(x)} - \frac{3}{2}\left(\frac{f''(x)}{f'(x)}\right)^2 = -\frac{3}{2}\left(\frac{-2}{1 - 2x}\right)^2 < 0$$

in $[0, 1]$. Deduciamo, in base al teorema di Singer, che per ogni valore di $r \in (0, 4]$ esiste al massimo un'orbita periodica in $(0, 1)$, asintoticamente stabile e $x_0 = 1/2$ è nel suo bacino d'attrazione.

La dinamica della (4.2) presenta una insospettata dipendenza dal parametro r e un'interessante varietà di fenomeni. Come abbiamo visto nel paragrafo 1.2, i punti fissi, soluzioni dell'equazione

$$rx(1 - x) = x$$

sono

$$x_1 = 0 \quad \text{e} \quad x_2 = 1 - \frac{1}{r}.$$

Distinguiamo ora vari casi.

$\boxed{0 < r \leq 1}$

Solo $x_1 = 0$ sta nell'intervallo $[0, 1]$ e risulta asintoticamente stabile con bacino d'attrazione $[0, 1]$, infatti per $0 < r < 1$ si ha

$$f'(x) = r - 2rx$$

e

$$f'(0) = r \in (0, 1).$$

Si ha stabilità asintotica anche per $r = 1$, come mostra il diagramma a gradino (v. fig. 7.9).

Figura 7.9. Mappa logistica per $r = 1$.

Si osservi che se $0 < r < 1$ il punto $x_1 = 0$ risulterebbe (localmente) asintoticamente stabile anche considerando $f : \mathbb{R} \to \mathbb{R}$, mentre per $r = 1$ è *semistabile* (le orbite che partono a sinistra di 0, si allontanano).

$\boxed{1 < r \leq 3}$

Anche $x_2 \in [0, 1]$ e si ha

$$f'\left(1 - \frac{1}{r}\right) = r - 2r\left(1 - \frac{1}{r}\right) = 2 - r$$

e

$$|2 - r| < 1 \quad \text{per} \quad 1 < r < 3.$$

Il punto $x_1 = 0$ è instabile, mentre x_2 è asintoticamente stabile con bacino di attrazione $(0, 1)$. Anche per $r = 3$, x_2 è asintoticamente stabile.

$\boxed{3 < r \leq 1 + \sqrt{6} = 3{,}449499\ldots}$

Per $r > 3$ i punti x_1 e x_2 sono instabili. Dal grafico di f non si vede di più.

Cerchiamo orbite periodiche di periodo 2 considerando l'iterata seconda di f :

$$f^2(x) = r^2 x(1 - x)(1 - rx(1 - x)).$$

I punti periodici di periodo 2 corrispondono ai punti fissi di f^2 e si trovano risolvendo l'equazione

$$r^2 x (1-x)(1-rx(1-x)) = x.$$

Oltre (ovviamente) a $x_1 = 0$ e $x_2 = 1 - \frac{1}{r}$, si trovano i punti

$$x_{3,4} = \frac{1 + r \pm \sqrt{r^2 - 2r - 3}}{2r}$$

soluzioni dell'equazione

$$r^2 x^2 - r(r+1)x + r + 1 = 0.$$

Per studiarne la stabilità calcoliamo

$$(f^2)'(x_3) = (f^2)'(x_4) = f'(x_3) f'(x_4) = -r^2 + 2r + 4.$$

Si ha che

$$\left| -r^2 + 2r + 4 \right| < 1$$

per $3 < r < 1 + \sqrt{6}$ e quindi, per tali valori del parametro, esiste un'orbita periodica asintoticamente stabile di periodo 2. Il bacino d'attrazione dell'orbita $\{x_3, x_4\}$ è l'intervallo $(0,1)$, tranne un'infinità numerabile di punti che arriva a x_2 dopo un numero finito di iterazioni. Dal grafico di f^2 si intuisce che lo stesso accade per $r = r_2 = 1+\sqrt{6}$.

In conclusione, quando r supera il valore $r_1 = 3$ il punto fisso x_2 diventa instabile e si crea un'orbita periodica asintoticamente stabile di periodo 2. Questo cambio di comportamento delle orbite prende il nome di *biforcazione*. La figura 7.10 mostra,

Figura 7.10. Mappa logistica per $r = 3{,}2$.

in corrispondenza del valore $r = 3{,}2$, l'orbita uscente dal punto $x_0 = 0{,}7$; la curva tratteggiata è il grafico dell'iterata seconda f^2.

$$\boxed{1+\sqrt{6} < r \leq r_\infty = 3{,}5699456\ldots}$$

Superato il valore $r_2 = 1+\sqrt{6}$, l'orbita di periodo 2 diventa instabile e si crea un'orbita stabile di periodo 4. Ciò si vede esaminando l'iterata quarta f^4. Per $r = r_2$ si verifica quindi un'altra biforcazione con raddoppio del periodo. La figura 7.11 mostra, in corrispondenza del valore $r = 3{,}52$, l'orbita uscente dal punto $x_0 = 0{,}6$; la curva tratteggiata è il grafico dell'iterata quarta f^4.

Figura 7.11. Mappa logistica per $r = 3{,}52$.

L'orbita di periodo 4 rimane stabile fino ad un valore $r_3 = 3{,}544090\ldots$, in corrispondenza al quale si verifica un'altra biforcazione con raddoppio del periodo; si crea, cioè, un'orbita asintoticamente stabile di periodo 8.

Successivamente, al crescere di r, si verifica una sequenza infinita di biforcazioni con raddoppio del periodo in corrispondenza ad infiniti valori r_k. I primi otto sono

$$\begin{array}{llll} r_1 = 3 & r_2 = 3{,}449499\ldots & r_3 = 3{,}544090\ldots & r_4 = 3{,}564407\ldots \\ r_5 = 3{,}568759\ldots & r_6 = 3{,}569692\ldots & r_7 = 3{,}569891\ldots & r_8 = 3{,}569934\ldots \end{array}$$

I valori r_k convergono ad un valore

$$r_\infty = 3{,}5699456\ldots$$

Sperimentalmente, si è osservato che $r_n = r_\infty -$ costante$\cdot\delta^{-n}$ $(n \gg 1)$ dove $\delta = 4{,}6692016091\ldots$ è detta *costante di Feigenbaum*.

In corrispondenza a r_∞ esiste un insieme infinito F che attrae i punti di $(0,1)$, tranne i punti periodici (instabili) e i punti che arrivano ai punti instabili dopo un numero finito di iterazioni. L'insieme F ha struttura *frattale*. L'insieme $1 < r < r_\infty$ definisce il *regime di biforcazione* per la mappa logistica.

$\boxed{r_\infty < r \leq 4}$

Per $r > r_\infty$ si entra nel *regime caotico* ed il comportamento è molto più complicato. Esistono orbite periodiche instabili di periodo 2^n (per ogni n ce n'è una), orbite stabili (di periodo $3 \cdot 2^n, 5 \cdot 2^n$ ecc.) e anche orbite aperiodiche.

Per $r > 3{,}67857\ldots$ cominciano a presentarsi orbite di periodo dispari. In particolare, per $r = 1 + \sqrt{8}$ esiste un'unica orbita di periodo 3 asintoticamente stabile e, per il teorema di Sharkowsky, orbite di periodo qualunque.

Figura 7.12. Orbita di periodo 3.

Come abbiamo già visto, il regime caotico presenta forte instabilità rispetto alle condizioni iniziali. La figura 7.13 mostra i grafici di due soluzioni dell'equazione logistica in corrispondenza del valore $r = 4$. Le due soluzioni, pur partendo da punti molto vicini ($x_1 = 0.,71$ e $x_1 = 0{,}72$) hanno un andamento assai differente.

Figura 7.13. Grafici di due soluzioni dell'equazione logistica per $r = 4$.

Esercizi

1 Scrivere la soluzione generale delle seguenti equazioni alle differenze lineari del prim'ordine.

$$s_{n+1} = -\frac{s_n}{3}, \qquad s_{n+1} = -2s_n + 1.$$

Nei due casi discutere il carattere di s_n.

2 Stabilire il carattere delle successioni per ricorrenza

a) $\begin{cases} s_{n+1} = 2s_n - 1 \\ s_0 = 3. \end{cases}$ b) $\begin{cases} s_{n+1} = s_n - 1 \\ s_0 = 4 \end{cases}$ c) $\begin{cases} s_{n+1} = -s_n \\ s_0 = 4 \end{cases}$

3 Determinare una soluzione particolare per l'equazione

$$x_{n+1} = \frac{1}{2}x_n + b_n,$$

nei casi

$$b_n = n \quad \text{e} \quad b_n = 2^n.$$

4 Sia $f : [a, b] \to [a, b]$ continua. Dimostrare che f ammette un *punto fisso* in $[a, b]$, ossia l'equazione

$$f(x) = x$$

ha almeno una soluzione $x^* \in [a, b]$.

5 Sia

$$\begin{cases} x_{n+1} = f(x_n) \\ x_0 \text{ (assegnato)}, \end{cases}$$

con $f : [a, b] \to [a, b]$ continua. Dimostrare che se $x_n \to l \in \mathbb{R}$, allora l è punto fisso per f.

6 Sia

$$\begin{cases} x_{n+1} = f(x_n) \\ x_0 \text{ (assegnato)}, \end{cases}$$

con $f : [a, b] \to [a, b]$ crescente. Dimostrare che se $x_1 > x_0$, allora x_n è crescente e che se $x_1 < x_0$, allora x_n è decrescente, quindi, in ogni caso, x_n è monotona e, quindi, regolare.

7 Studiare, al variare di $a > 0$, il comportamento asintotico della successione per ricorrenza

$$\begin{cases} s_0 = a \\ s_{n+1} = \sqrt{2 + s_n}. \end{cases}$$

8 Sia

$$\begin{cases} s_{n+1} = f(s_n) \\ s_0 \text{ (assegnato)}, \end{cases}$$

con $f : [a, b] \to [a, b]$ continua e decrescente. Dimostrare che

a) la successione[10]

$$\begin{cases} x_{n+1} = f^2(x_n) \\ x_0 = s_0 \end{cases}$$

coincide con s_{2n} e la successione

$$\begin{cases} y_{n+1} = f^2(y_n) \\ y_0 = s_1 \end{cases}$$

coincide con s_{2n+1};

[10] Ricordiamo che il simbolo f^2 indica $f \circ f$, ossia l'iterata seconda di f.

b) le successioni x_n e y_n sono monotone; in particolare $x_n = s_{2n}$ è crescente se e solo se $y_n = s_{2n+1}$ è decrescente e viceversa.

c) $x_n \to x^*$ se e solo se $y_n \to y^*$ con $x^* = f(y^*)$ e $y^* = f(x^*)$.

d) $s_n \to l \in \mathbb{R}$ se e solo se $l = x^* = y^*$.

9 Calcolare i punti di equilibrio stabile per l'equazione alle differenze

$$x_{n+1} = \frac{1}{2}\left(x_n + \frac{A}{x_n}\right), \quad A > 0.$$

La relazione indicata può servire per calcolare velocemente radici quadrate. Calcolare $\sqrt{7}$ con tre cifre decimali esatte.

10 Studiare, al variare di $a > 0$, il comportamento asintotico della successione per ricorrenza

$$\begin{cases} s_0 = a \\ s_{n+1} = \dfrac{2}{1+s_n}. \end{cases}$$

11 Nel cuore, l'eccitazione generata da un normale *pacemaker* negli atrii, passa ai ventricoli causando contrazione del cuore e l'invio di sangue agli altri organi. L'eccitazione deve passare attraverso il nodo atrio-ventricolare, che connette elettricamente atrii e ventricoli. Il problema seguente è basato su un modello matematico per la conduzione atrio-ventricolare nei mammiferi.[11] Assumiamo che i tempi (in millisecondi) di conduzione atrio-ventricolare soddisfino la seguente legge di ricorrenza

$$x_{n+1} = \frac{375}{x_n - 90} + 100, \quad x_n > 90.$$

a) Determinare i punti fissi e studiarne la stabilità.

b) Analizzare il comportamento asintotico della soluzione che verifica la condizione iniziale $x_0 = 150$.

12 La seguente equazione alle differenze è stata proposta come modello di evoluzione per la densità di popolazione di insetti in anni successivi

$$x_{n+1} = f(x_n) = \alpha x_n e^{-\beta x_n^3},$$

dove α, β sono parametri positivi. Per $\alpha = e$ e $\beta = 1/3$,

a) tracciare il grafico di f;

b) determinare i punti di equilibrio e studiarne la stabilità.

13 Nello studio delle reti neurali[12] si incontra la seguente equazione alle differenze

$$x_{n+1} = f(x_n) = \frac{\alpha x_n}{1 + \beta x_n},$$

con α e β parametri positivi e $x_n \geq 0$.

[11] Simson e altri, *American Journal of Phisiology*, 1981.
[12] Glass e Pasternack, *Journal of math. biology*, 1978.

a) Determinare al variare di α e β i punti di equilibrio, la loro stabilità e se la dinamica in un intorno dei punti fissi è monotona o oscillatoria.

b) Nel caso $\alpha = \beta = 1$, tracciare il diagramma di fase e studiare il comportamento asintotico di x_n.

c) Esprimere x_n in funzione di x_0 (> 0 generico) e verificare il risultato ottenuto in (*b*).

5 EQUAZIONI LINEARI A COEFFICIENTI COSTANTI DEL SECOND'ORDINE

In questo capitolo analizziamo le equazioni lineari a coefficienti costanti del second'ordine:

Definizione 7.3 Un'equazione alle differenze lineare del second'ordine o a due passi a coefficienti costanti è un'equazione del tipo[13]

$$ax_{n+2} + bx_{n+1} + cx_n = f_n, \qquad a, b, c \in \mathbb{R},\ a, c \neq 0.$$

Se la successione f_n è identicamente nulla (cioè $f_n = 0$ per ogni n) l'equazione si dice *omogenea*. Per esempio, è omogenea l'equazione $x_{n+2} + x_n = 0$, mentre non lo è l'equazione $x_{n+2} + x_n = 3n$.

5.1 Equazioni omogenee

Si può dimostrare che *la soluzione generale di un'equazione alle differenze del second'ordine lineare omogenea*

$$ax_{n+2} + bx_{n+1} + cx_n = 0 \tag{5.1}$$

è costituita da tutte le combinazioni lineari di due soluzioni particolari φ_n e ψ_n linearmente indipendenti (ossia che non siano una multipla dell'altra) può, cioè, essere scritta nelle forma

$$x_n = C_1\varphi_n + C_2\psi_n, \qquad C_1, C_2 \in \mathbb{R}.$$

La *soluzione generale* della (5.1) contiene due costanti arbitrarie. Per determinare una soluzione particolare è necessario assegnare una coppia di condizioni; di solito si assegnano due valori per la successione x_n in due tempi diversi, tipicamente per $n = 0$ e $n = 1$.

Per scrivere la soluzione generale della (5.1), occorre, dunque, determinare le due soluzioni particolari φ_n e ψ_n opportune. Ricordiamo che la soluzione generale dell'equazione del primo ordine:

$$x_{n+1} = \alpha x_n,$$

è $x_n = C\alpha^n$. Proviamo quindi a cercare funzioni del tipo $x_n = \lambda^n$, con $\lambda \neq 0$ da determinarsi in modo da ottenere una soluzione della (5.1). Se $x_n = \lambda^n$, si ha $x_{n+1} = \lambda^{n+1}$ e $x_{n+2} = \lambda^{n+2}$. Sostituendo nella (5.1) troviamo

$$a\lambda^2\lambda^n + b\lambda\lambda^n + c\lambda^n = 0,$$

[13] Chiediamo che i coefficienti a e c siano diversi da 0, altrimenti non sarebbe di second'ordine.

da cui, dividendo per λ^n, otteniamo l'equazione di secondo grado in λ:

$$\boxed{a\lambda^2 + b\lambda + c = 0} \tag{5.2}$$

La (5.2) si dice *equazione caratteristica* dell'equazione alle differenze (5.1). Dovremo allora distinguere tre casi.

$\boxed{b^2 - 4ac > 0}$

Se il discriminante dell'equazione è positivo la (5.2) ha le due soluzioni reali e distinte:

$$\lambda_{1,2} = \frac{-b \pm \sqrt{b^2 - 4ac}}{2a}.$$

Le due successioni λ_1^n, λ_2^n verificano la (5.1) e sono linearmente indipendenti (controllare). In questo caso la soluzione generale dell'equazione alle differenze (5.1) è

$$\boxed{x_n = C_1\lambda_1^n + C_2\lambda_2^n, \qquad C_1, C_2 \in \mathbb{R}}$$

Esempio

5.1 Per l'equazione alle differenze

$$x_{n+2} - 6x_{n+1} + 5x_n = 0$$

l'equazione caratteristica è

$$\lambda^2 - 6\lambda + 5 = 0 \tag{5.3}$$

che ha le due soluzioni $\lambda_1 = 1$ e $\lambda_2 = 5$. Pertanto, la soluzione generale dell'equazione è

$$x_n = C_1 + C_2 5^n, \qquad C_1, C_2 \in \mathbb{R}.$$

Supponiamo interessi trovare la soluzione particolare soddisfa le condizioni $x_0 = 2$ e $x_1 = 1$. Sostituendo nella soluzione generale $n = 0$ e successivamente $n = 1$ si trova il sistema

$$\begin{cases} C_1 + C_2 = 2 \\ C_1 + 5C_2 = 1. \end{cases}$$

che ha soluzione

$$C_1 = \frac{9}{4}, \quad C_2 = -\frac{1}{4}.$$

La soluzione particolare cercata per l'equazione alle differenze è, quindi,

$$x_n = \frac{9}{4} - \frac{1}{4}5^n.$$

$\boxed{b^2 - 4ac = 0}$

L'equazione caratteristica ha una sola radice (*doppia*) $\lambda = -\dfrac{b}{2a}$; essa genera la soluzione λ^n. Si può controllare che un'altra soluzione linearmente indipendente è data dalla successione $n\lambda^n$. Infatti da $x_n = n\lambda^n$ si ricava

$$x_{n+1} = (n+1)\,\lambda^{n+1} \quad \text{e} \quad x_{n+2} = (n+2)\,\lambda^{n+2}.$$

Sostituendo nell'equazione alle differenze

$$a(n+2)\lambda^{n+2} + b(n+1)\lambda^{n+1} + cn\lambda^n = 0,$$

dopo aver diviso per λ^n, ci si riduce a

$$\left(a\lambda^2 + b\lambda + c\right)n + 2a\lambda^2 + b\lambda = 0$$

che è un'identità poiché $\lambda = -\dfrac{b}{2a}$ è radice dell'equazione caratteristica. La soluzione generale per l'equazione alle differenze in questo caso è

$$\boxed{x_n = C_1\lambda^n + C_2 n\lambda^n = (C_1 + C_2 n)\lambda^n, \qquad C_1, C_2 \in \mathbb{R}}$$

Esempio

5.2 Sia

$$x_{n+2} + 4x_{n+1} + 4 = 0.$$

L'equazione caratteristica è

$$\lambda^2 + 4\lambda + 4 = 0,$$

che possiede la sola soluzione $\lambda = -2$. La sua soluzione generale è

$$x_n = (-2)^n (C_1 + C_2 n),$$

con C_1, C_2 arbitrari.

$\boxed{b^2 - 4ac < 0}$

L'equazione caratteristica ha due soluzioni complesse coniugate $\lambda_{1,2} = \alpha \pm i\beta$. Introduciamo il loro modulo

$$|\lambda_{1,2}| = \rho = \sqrt{\alpha^2 + \beta^2}$$

e (a meno del segno) l'argomento $\theta \neq 0$, definito da

$$\operatorname{tg}\theta = \frac{\beta}{\alpha}.$$

Si può controllare che le due successioni[14]

$$\rho^n \cos n\theta \quad \text{e} \quad \rho^n \sin n\theta$$

sono soluzioni indipendenti, per cui la soluzione generale in questo caso è

$$\boxed{x_n = C_1\rho^n \cos n\theta + C_2\rho^n \sin n\theta, \qquad C_1, C_2 \in \mathbb{R}}$$

[14] Scrivendo $\lambda_{1,2} = \rho(\cos\theta \pm i\sin\theta)$, si ha, per la formula di De Moivre

$$\lambda_1^n = \rho^n(\cos n\theta + i\sin n\theta), \quad \lambda_2^n = \rho^n(\cos n\theta - i\sin n\theta).$$

Due soluzioni (reali) dell'equazione sono quindi

$$\frac{\lambda_1^n + \lambda_2^n}{2} \quad \text{e} \quad \frac{\lambda_1^n - \lambda_2^n}{2i}.$$

 © 978-88-08-06485-1

Esempio

5.3 Sia

$$x_{n+2} + 4x_n = 0.$$

Il discriminante dell'equazione caratteristica è negativo: le due radici sono $\pm 2i$ e, conseguentemente

$$\rho = 2 \quad \text{e} \quad \theta = \frac{\pi}{2}.$$

La soluzione generale è

$$x_n = C_1 2^n \cos\left(\frac{\pi}{2}n\right) + C_2 2^n \sin\left(\frac{\pi}{2}n\right),$$

con C_1 e C_2 costanti reali.

Scriviamo ora la soluzione dell'equazione alle differenze $x_{n+2} + 4x_n = 0$, che verifica le condizioni iniziali

$$x_0 = 1, \quad x_1 = 2.$$

Per determinare tale soluzione particolare dobbiamo determinare le costanti C_1 e C_2. Per far questo sostituiamo i valori 0 e 1 nella soluzione generale; si ha

$$\begin{cases} x_0 = C_1 = 1 \\ x_1 = 2C_2 = 2 \end{cases} \Rightarrow \begin{cases} C_1 = 1 \\ C_2 = 1. \end{cases}$$

La soluzione cercata[15] è

$$x_n = 2^n \left(\cos\left(\frac{\pi}{2}n\right) + \sin\left(\frac{\pi}{2}n\right)\right).$$

5.2 I numeri di Fibonacci

Presentiamo un problema esaminato da Fibonacci[16] nel suo *Liber Abaci* (1202). Si tratta di un problema di ... allevamento di conigli. Ci chiediamo quante coppie di conigli si possono ottenere in un anno a partire da una sola coppia se formuliamo le seguenti ipotesi:

- la prima coppia è costituita da conigli appena nati;
- ogni coppia di conigli genera una nuova coppia ogni mese;
- ogni coppia di conigli diventa fertile dopo un mese di vita;
- i conigli non muoiono.

---

[15] Si noti che

$$\left\{\cos\frac{n\pi}{2}\right\} = \{1, 0, -1, 0, 1, 0, -1, 0, \ldots\}$$
$$\left\{\sin\frac{n\pi}{2}\right\} = \{0, 1, 0, -1, 0, 1, 0, -1, \ldots\}$$

per cui

$$\left\{\cos\frac{n\pi}{2} + \sin\frac{n\pi}{2}\right\} = \{1, 1, -1, -1, 1, 1, -1, -1, \ldots\}.$$

La soluzione è periodica di periodo 4.

[16] Soprannome di Leonardo Pisano (1175-1240).

Indichiamo con F_k il numero di coppie dopo k mesi ($k \geq 0$). All'inizio e anche dopo un mese avremo solo la prima coppia, quindi $F_0 = F_1 = 1$. Al secondo mese avremo un'altra coppia, quindi $F_2 = 1 + 1 = 2$, al terzo mese solo la prima coppia è in grado di generarne un'altra e quindi $F_3 = 2 + 1 = 3 \ldots$ e così via.

Alla fine del generico mese $k+2$ avremo oltre le coppie di conigli già presenti F_{k+1}, e le neonate, pari al numero di coppie fertili del mese precedente F_k, quindi

$$F_{k+2} = F_{k+1} + F_k. \tag{5.4}$$

La (5.4) è un'equazione alle differenze a due passi, lineare omogenea a coefficienti costanti.

Dalla (5.4), con la condizione (iniziale) $F_0 = F_1 = 1$ è facile calcolare F_k (specialmente se si ha un computer!). I primi 12 termini sono

$$1,\ 1,\ 2,\ 3,\ 5,\ 8,\ 13,\ 21,\ 34,\ 55,\ 89,\ 144.$$

Risolviamo la (5.4). L'equazione caratteristica è

$$\lambda^2 - \lambda - 1 = 0$$

che ha le soluzioni

$$\lambda_1 = \frac{1-\sqrt{5}}{2}, \quad \lambda_2 = \frac{1+\sqrt{5}}{2},$$

per cui la soluzione generale è

$$F_k = c_1 \left(\frac{1-\sqrt{5}}{2}\right)^k + c_2 \left(\frac{1+\sqrt{5}}{2}\right)^k, \qquad c_1, c_2 \in \mathbb{R}.$$

Determiniamo c_1 e c_2 imponendo le condizioni iniziali $F_0 = F_1 = 1$. Si ottiene il sistema

$$\begin{cases} c_1 + c_2 = 1 \\ c_1 \dfrac{1-\sqrt{5}}{2} + c_2 \dfrac{1+\sqrt{5}}{2} = 1 \end{cases}$$

che ha la soluzione

$$\begin{cases} c_1 = -\dfrac{1-\sqrt{5}}{2\sqrt{5}} \\ c_2 = \dfrac{1+\sqrt{5}}{2\sqrt{5}}. \end{cases}$$

Quindi

$$F_k = \frac{1}{\sqrt{5}} \left\{ \left(\frac{1+\sqrt{5}}{2}\right)^{k+1} - \left(\frac{1-\sqrt{5}}{2}\right)^{k+1} \right\}.$$

Si noti che F_k è un numero naturale per ogni k. Poiché

$$\frac{\sqrt{5}-1}{2} < 1$$

si ha

$$\left(\frac{1-\sqrt{5}}{2}\right)^{k+1} \to 0 \quad \text{per} k \to +\infty$$

quindi

$$F_k \sim \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^{k+1}$$

e

$$\lim_{k\to+\infty} \frac{F_{k+1}}{F_k} = \frac{1+\sqrt{5}}{2}.$$

Incidentalmente, osserviamo che il numero $\dfrac{\sqrt{5}-1}{2}$ è noto come *sezione aurea* (golden ratio).

Figura 7.14. Il problema dei conigli e i numeri di Fibonacci.

5.3 Equazioni non omogenee

Consideriamo ora un'equazione alle differenze del tipo

$$ax_{n+2} + bx_{n+1} + cx_n = f_n, \tag{5.5}$$

con secondo membro non identicamente nullo. Si può dimostrare che *la soluzione generale può scriversi come*

$$x_n = C_1\varphi_n + C_2\psi_n + y_n, \qquad C_1, C_2 \in \mathbb{R} \tag{5.6}$$

ove y_n è una soluzione particolare della (5.5) e $C_1\varphi_n + C_2\psi_n$ è la soluzione generale dell'equazione omogenea associata

$$az_{n+2} + bz_{n+1} + cz_n = 0.$$

Il problema è quindi determinare una soluzione particolare. In alcuni casi si può utilizzare il *metodo di somiglianza*: se f_n è d'un certo tipo, allora anche una soluzione particolare è dello stesso tipo. Per esempio, se f_n è un polinomio allora anche una soluzione particolare y_n è un polinomio, se il termine noto è esponenziale allora anche una soluzione particolare è un'esponenziale e così via.

Esempio

5.4 Scriviamo la soluzione generale dell'equazione alle differenze

$$6x_{n+2} + 5x_{n+1} + x_n = 3. \tag{5.7}$$

Come prima cosa determiniamo le soluzioni dell'equazione omogenea associata

$$6y_{n+2} + 5y_{n+1} + y_n = 0.$$

L'equazione caratteristica

$$6\lambda^2 + 5\lambda + 1 = 0$$

ha le radici

$$\lambda = -\frac{1}{3} \quad \text{e} \quad \lambda = -\frac{1}{2},$$

quindi la soluzione generale dell'equazione omogenea è

$$y_n = c_1\left(-\frac{1}{3}\right)^n + c_2\left(-\frac{1}{2}\right)^n \qquad c_1, c_2 \in \mathbb{R}.$$

Cerchiamo poi una soluzione dell'equazione non omogenea. Dato che il secondo membro dell'equazione completa è una costante, pensiamo come soluzione particolare dell'equazione completa a una successione costante $u_n = k$. Se $u_n = k$ anche $u_{n+1} = k$ e $u_{n+2} = k$ e sostituendo nella (5.7), si ricava

$$6k + 5k + k = 3 \quad \Rightarrow \quad k = \frac{1}{4}.$$

La soluzione generale della (5.7) è, dunque,

$$x_n = c_1\left(-\frac{1}{3}\right)^n + c_2\left(-\frac{1}{2}\right)^n + \frac{1}{4} \qquad c_1, c_2 \in \mathbb{R}.$$

 © 978-88-08-06485-1

5.4 Stabilità

Per l'equazione alle differenze omogenea la successione $x_n = 0$ per ogni n, rappresenta una *soluzione di equilibrio*. Se ogni altra soluzione dell'equazione converge a 0 per $n \to +\infty$ la soluzione nulla si dice *globalmente asintoticamente stabile*.

Per le equazioni alle differenze non omogenee è importante sapere se la soluzione nulla è asintoticamente stabile per l'omogenea associata. In tal caso infatti può succedere che sia la soluzione particolare u_n a determinare il comportamento a regime del sistema, mentre l'integrale generale dell'omogenea assume un significato di regime transitorio.

Avendo a disposizione una formula esplicita per la soluzione generale possiamo dedurre direttamente il comportamento all'infinito.

Proposizione 7.1 *La soluzione nulla è*

a) *asintoticamente stabile se e solo se il modulo di ogni radice dell'equazione caratteristica è minore di* 1;

b) *stabile se e solo se le radici sono complesse con modulo uguale a* 1 *oppure le radici sono reali e distinte, una con modulo minore di* 1 *e l'altra uguale a* ± 1.

c) *instabile se l'equazione caratteristica ha una radice di modulo maggiore di uno oppure se ha una radice doppia uguale a* ± 1.

Esistono criteri di stabilità particolarmente comodi che non richiedono di risolvere l'equazione caratteristica. Ci limitiamo al seguente.

Proposizione 7.2 *Criterio per equazioni di secondo grado.*
La soluzione nulla è asintoticamente stabile per l'equazione

$$ax_{n+2} + bx_{n+1} + cx_n = 0$$

se e solo se, dette λ_1 *e* λ_2 *le radici dell'equazione caratteristica, si ha*

$$\lambda_1\lambda_2 < 1 \quad e \quad |\lambda_1 + \lambda_2| < 1 + \lambda_1\lambda_2$$

ossia se e solo se

$$\frac{c}{a} < 1 \quad e \quad \left|\frac{b}{a}\right| < 1 + \frac{c}{a}$$

quindi se e solo se, posto $a > 0$,

$$c < a \quad e \quad |b| < a + c.$$

La dimostrazione non è difficile ed il lettore può cimentarsi (si veda l'esercizio 18).

Esempio

5.5 Riprendiamo il *modello di Samuelson* considerato nell'introduzione, a cui rimandiamo per la derivazione. Si tratta dell'equazione del second'ordine

$$Y_n - a(1+b)Y_{n-1} + abY_{n-2} = A + k. \tag{5.8}$$

Interessa sapere se esiste una soluzione d'equilibrio e, in caso affermativo, studiarne la stabilità. La soluzione di equilibrio, cioè una soluzione costante $\overline{Y}$, si trova subito dall'equazione

$$\overline{Y} - a(1+b)\overline{Y} + ab\overline{Y} = A + k.$$

ed è

$$\overline{Y} = \frac{A+k}{1-a}.$$

Questa soluzione risulta asintoticamente stabile se e solo se la soluzione nulla è asintoticamente stabile per l'equazione omogenea e cioè se e solo se le radici λ_1 e λ_2 dell'equazione

$$\lambda^2 - a(1+b)\lambda + ab = 0 \tag{5.9}$$

sono in modulo minori di 1. In base al criterio di stabilità, ciò accade se

$$ab < 1 \quad \text{e} \quad a(1+b) < 1 + ab.$$

Essendo $0 < a < 1$, la seconda condizione è sempre verificata per cui ci si riduce a

$$ab < 1. \tag{5.10}$$

L'andamento delle soluzioni è differente a seconda dei valori dei parametri. Il discriminante dell'equazione (5.9) risulta positivo o nullo se

$$a^2(1+b)^2 - 4ab \geq 0 \quad \text{ossia se} \quad a \geq \frac{4b}{(1+b)^2}, \tag{5.11}$$

quindi, se sono verificate entrambe le condizioni (5.10) e (5.11), la soluzione dell'equazione (5.8) tende all'equilibrio senza compiere oscillazioni; se è verificata la (5.10), ma non la (5.11), la soluzione dell'equazione (5.8) tende all'equilibrio oscillando. La figura 7.15 mostra il grafico una soluzione di questo tipo, ottenuta in corrispondenza dei valori dei parametri $a = 1/4$ e $b = 3$.

Figura 7.15. Una soluzione del modello di Samuelson.

Esercizi

14 Scrivere la soluzione dell'equazione alle differenze

$$x_{n+2} - 5x_{n+1} + 6x_n = 0$$

che soddisfa le condizioni $x_0 = 0$ e $x_1 = 1$.

15 Scrivere la soluzione generale delle seguenti equazioni alle differenze lineari del second'ordine.

$$(a)\ 2x_{n+2} - 3x_{n+1} + x_n = 0, \quad (b)\ x_{n+2} + x_n = 0, \quad (c)\ x_{n+2} - 4x_{n+1} + 4x_n = 1.$$

Dire se la soluzione costante è stabile.

16 Risolvere le equazioni alle differenze

$$(a)\ x_{n+2} - 3x_{n+1} + 2x_n = 3^n, \quad (b)\ x_{n+3} - x_{n+2} + x_{n+1} - x_n = -1.$$

17 Determinare la soluzione dell'equazione alle differenze

$$x_{n+3} + x_n = 0$$

che corrisponde ai dati iniziali $x_0 = 1$, $x_1 = 0$ e $x_2 = 2$. Controllare che tale soluzione è periodica di periodo 6.

18 Dimostrare la Proposizione 7.2 ossia che *se λ_1 e λ_2 sono radici di un trinomio di secondo grado*

$$|\lambda_1| < 1 \text{ e } |\lambda_2| < 1 \iff \lambda_1\lambda_2 < 1 \text{ e } |\lambda_1 + \lambda_2| < 1 + \lambda_1\lambda_2.$$

19 Sia

$$x_{n+2} = x_{n+1} + ax_n, \quad a \in \mathbb{R}.$$

a) Stabilire per quali valori del parametro la soluzione nulla è asintoticamente stabile.

b) Posto $a = \dfrac{1}{8}$, calcolare la soluzione che verifica la condizione iniziale $x_0 = 0$, $x_1 = 10$.

A Formule utili

1 COSTANTI MATEMATICHE

e	2,7182818285...
π	3,1415926536...
$\log_{10} 2$	0,3010299957...
$\log_{10} e$	0,4342944819...
$\log_{10} \pi$	0,4971498727...
$\log_e 2$	0,6931471806...
$\log_e \pi$	1,1447298858...
$\log_e 10$	2,3025850930...
$\sqrt{2}$	1,4142135624...
$\sqrt{e}$	1,6487212707...
$\sqrt{3}$	1,7320508076...
$\sqrt{\pi}$	1,7724538509...
$\sqrt{5}$	2,2360679775...
$\sqrt{10}$	3,1622776602...
1°	0,0174532925... radianti
1 radiante	$57^\circ 17' 44'',806...$

2 FUNZIONI TRIGONOMETRICHE

$$\sin x \qquad \cos x \qquad \operatorname{tg} x = \frac{\sin x}{\cos x} \qquad \operatorname{cotg} x = \frac{\cos x}{\sin x}$$

$$(\sin x)^2 + (\cos x)^2 = 1$$

Angoli notevoli

x	$\cos x$	$\sin x$	$\operatorname{tg} x$	$\operatorname{cotg} x$
0	1	0	0	$\pm\infty$
$\frac{\pi}{10} = 18°$	$\frac{1}{4}\sqrt{10+2\sqrt{5}}$	$\frac{\sqrt{5}-1}{4}$	$\sqrt{\frac{5-2\sqrt{5}}{5}}$	$\sqrt{5+2\sqrt{5}}$
$\frac{\pi}{6} = 30°$	$\frac{\sqrt{3}}{2}$	$\frac{1}{2}$	$\frac{\sqrt{3}}{3}$	$\sqrt{3}$
$\frac{\pi}{5} = 36°$	$\frac{\sqrt{5}+1}{4}$	$\frac{1}{4}\sqrt{10-2\sqrt{5}}$	$\sqrt{5-2\sqrt{5}}$	$\sqrt{\frac{5+2\sqrt{5}}{5}}$
$\frac{\pi}{4} = 45°$	$\frac{\sqrt{2}}{2}$	$\frac{\sqrt{2}}{2}$	1	1
$\frac{3\pi}{10} = 54°$	$\frac{1}{4}\sqrt{10-2\sqrt{5}}$	$\frac{\sqrt{5}+1}{4}$	$\sqrt{\frac{5+2\sqrt{5}}{5}}$	$\sqrt{5-2\sqrt{5}}$
$\frac{\pi}{3} = 60°$	$\frac{1}{2}$	$\frac{\sqrt{3}}{2}$	$\sqrt{3}$	$\frac{\sqrt{3}}{3}$
$\frac{2\pi}{5} = 72°$	$\frac{\sqrt{5}-1}{4}$	$\frac{1}{4}\sqrt{10-2\sqrt{5}}$	$\sqrt{5+2\sqrt{5}}$	$\sqrt{\frac{5-2\sqrt{5}}{5}}$
$\frac{\pi}{2} = 90°$	0	1	$\pm\infty$	0

Simmetrie, archi complementari e supplementari

$$\sin(-x) = -\sin x \qquad \cos(-x) = \cos x \qquad \operatorname{tg}(-x) = -\operatorname{tg} x$$

$$\sin\left(\frac{\pi}{2} \pm x\right) = \cos x \qquad \cos\left(\frac{\pi}{2} \pm x\right) = \mp \sin x \qquad \operatorname{tg}\left(\frac{\pi}{2} \pm x\right) = \mp \operatorname{cotg} x$$

$$\sin(\pi \pm x) = \mp \sin x \qquad \cos(\pi \pm x) = -\cos x \qquad \operatorname{tg}(\pi \pm x) = \pm \operatorname{tg} x$$

Formule di addizione

$$\sin(x \pm y) = \sin x \cos y \pm \cos x \sin y$$

$$\cos(x \pm y) = \cos x \cos y \mp \sin x \sin y$$

$$\operatorname{tg}(x \pm y) = \frac{\operatorname{tg} x \pm \operatorname{tg} y}{1 \mp \operatorname{tg} x \operatorname{tg} y}$$

Formule di duplicazione

$$\sin 2x = 2\sin x\cos x$$

$$\cos 2x = (\cos x)^2 - (\sin x)^2 = 2(\cos x)^2 - 1 = 1 - 2(\sin x)^2$$

$$\operatorname{tg} 2x = \frac{2\operatorname{tg} x}{1-(\operatorname{tg} x)^2}$$

Formule di bisezione (scegliere il segno corretto)

$$\sin\frac{x}{2} = \pm\sqrt{\frac{1-\cos x}{2}}$$

$$\cos\frac{x}{2} = \pm\sqrt{\frac{1+\cos x}{2}}$$

$$\operatorname{tg}\frac{x}{2} = \frac{1-\cos x}{\sin x} = \frac{\sin x}{1+\cos x}$$

Formule di prostaferesi

$$\sin u + \sin v = 2\sin\frac{u+v}{2}\cos\frac{u-v}{2}$$

$$\sin u - \sin v = 2\cos\frac{u+v}{2}\sin\frac{u-v}{2}$$

$$\cos u + \cos v = 2\cos\frac{u+v}{2}\cos\frac{u-v}{2}$$

$$\cos u - \cos v = -2\sin\frac{u+v}{2}\sin\frac{u-v}{2}$$

Formule parametriche Posto $t = \operatorname{tg}(x/2)$:

$$\sin x = \frac{2t}{1+t^2} \qquad \cos x = \frac{1-t^2}{1+t^2} \qquad \operatorname{tg} x = \frac{2t}{1-t^2}$$

Teorema di Carnot

$$a^2 = b^2 + c^2 - 2bc\cdot\cos\alpha$$

 © 978-88-08-06485-1

3 FUNZIONI IPERBOLICHE

$$\operatorname{Sh} x \qquad \operatorname{Ch} x \qquad \operatorname{Th} x = \frac{\operatorname{Sh} x}{\operatorname{Ch} x}$$

$$(\operatorname{Ch} x)^2 - (\operatorname{Sh} x)^2 = 1$$

$$\operatorname{Sh}(-x) = -\operatorname{Sh} x \qquad \operatorname{Ch}(-x) = \operatorname{Ch}(x) \qquad \operatorname{Th}(-x) = -\operatorname{Th} x$$

Formule di addizione

$$\operatorname{Sh}(x \pm y) = \operatorname{Sh} x \operatorname{Ch} y \pm \operatorname{Ch} x \operatorname{Sh} y$$

$$\operatorname{Ch}(x \pm y) = \operatorname{Ch} x \operatorname{Ch} y \pm \operatorname{Sh} x \operatorname{Sh} y$$

$$\operatorname{Th}(x \pm y) = \frac{\operatorname{Th} x \pm \operatorname{Th} y}{1 \pm \operatorname{Th} x \operatorname{Th} y}$$

Formule di duplicazione

$$\operatorname{Sh} 2x = 2 \operatorname{Sh} x \operatorname{Ch} x$$

$$\operatorname{Ch} 2x = (\operatorname{Ch} x)^2 + (\operatorname{Sh} x)^2 = 2(\operatorname{Ch} x)^2 - 1 = 1 + 2(\operatorname{Sh} x)^2$$

$$\operatorname{Th} 2x = \frac{2 \operatorname{Th} x}{1 + (\operatorname{Th} x)^2}$$

Formule di bisezione (scegliere il segno corretto)

$$\operatorname{Sh} \frac{x}{2} = \pm\sqrt{\frac{\operatorname{Ch} x - 1}{2}}$$

$$\operatorname{Ch} \frac{x}{2} = \sqrt{\frac{\operatorname{Ch} x + 1}{2}}$$

$$\operatorname{Th} \frac{x}{2} = \frac{\operatorname{Ch} x - 1}{\operatorname{Sh} x} = \frac{\operatorname{Sh} x}{\operatorname{Ch} x + 1}$$

Formule di prostaferesi

$$\operatorname{Sh} u + \operatorname{Sh} v = 2 \operatorname{Sh} \frac{u+v}{2} \operatorname{Ch} \frac{u-v}{2}$$

$$\operatorname{Sh} u - \operatorname{Sh} v = 2 \operatorname{Ch} \frac{u+v}{2} \operatorname{Sh} \frac{u-v}{2}$$

$$\operatorname{Ch} u + \operatorname{Ch} v = 2 \operatorname{Ch} \frac{u+v}{2} \operatorname{Ch} \frac{u-v}{2}$$

$$\operatorname{Ch} u - \operatorname{Ch} v = 2 \operatorname{Sh} \frac{u+v}{2} \operatorname{Sh} \frac{u-v}{2}$$

Formule parametriche Posto $t = \operatorname{Th}(x/2)$:

$$\operatorname{Sh} x = \frac{2t}{1-t^2} \qquad \operatorname{Ch} x = \frac{1+t^2}{1-t^2} \qquad \operatorname{Th} x = \frac{2t}{1+t^2}$$

4 DERIVATE ELEMENTARI

$f(x)$	$f'(x)$
x^α	$\alpha x^{\alpha-1}$
$\lvert x\rvert$	$\operatorname{sgn} x$
$\log\lvert x\rvert$	$1/x$
$\log_a\lvert x\rvert$	$1/(x\log a)$
e^x	e^x
a^x	$a^x\log a$
$\sin x$	$\cos x$
$\cos x$	$-\sin x$
$\operatorname{tg} x$	$1+(\operatorname{tg} x)^2 = 1/(\cos x)^2$
$\operatorname{ctg} x$	$-1-(\operatorname{ctg} x)^2 = -1/(\sin x)^2$
$\operatorname{Sh} x$	$\operatorname{Ch} x$
$\operatorname{Ch} x$	$\operatorname{Sh} x$
$\operatorname{Th} x$	$1-(\operatorname{Th} x)^2 = 1/(\operatorname{Ch} x)^2$
$\operatorname{Cth} x$	$1-(\operatorname{Cth} x)^2 = -1/(\operatorname{Sh} x)^2$
$\log\lvert\sin x\rvert$	$-\operatorname{ctg} x$
$\log\lvert\cos x\rvert$	$\operatorname{tg} x$
$\log\lvert\operatorname{Sh} x\rvert$	$\operatorname{Cth} x$
$\log\operatorname{Ch} x$	$\operatorname{Th} x$
$\arcsin x$	$1/\sqrt{1-x^2}$
$\arccos x$	$-1/\sqrt{1-x^2}$
$\operatorname{arctg} x$	$1/(1+x^2)$
$\operatorname{arccotg} x$	$-1/(1+x^2)$

5 REGOLE DI DERIVAZIONE

$$
\begin{aligned}
\mathrm{D}(\lambda f(x)+\mu g(x)) &= \lambda f'(x)+\mu g'(x)\\
\mathrm{D}(f(x)g(x)) &= f'(x)g(x)+f(x)g'(x)\\
\mathrm{D}\,\frac{f(x)}{g(x)} &= \frac{f'(x)g(x)-f(x)g'(x)}{g(x)^2}\\
\mathrm{D}\,f(g(x)) &= f'(g(x))g'(x)\\
\mathrm{D}\,f(g(h(x))) &= f'(g(h(x)))g'(h(x))h'(x)\\
\mathrm{D}\,e^{f(x)} &= e^{f(x)}f'(x)
\end{aligned}
$$

$$\mathrm{D}\log|f(x)| = \frac{f'(x)}{f(x)}$$

$$\mathrm{D}[f(x)]^{g(x)} = [f(x)]^{g(x)}\Big\{g'(x)\log f(x) + \frac{g(x)f'(x)}{f(x)}\Big\}$$

6 SVILUPPI DI MAC LAURIN DELLE PRINCIPALI FUNZIONI

$f(x)$	Sviluppo
e^x	$1 + x + \frac{x^2}{2!} + \frac{x^3}{3!} + \cdots + \frac{x^n}{n!} + \cdots$
$\operatorname{Sh} x$	$x + \frac{x^3}{3!} + \frac{x^5}{5!} + \cdots + \frac{x^{2n+1}}{(2n+1)!} + \cdots$
$\operatorname{Ch} x$	$1 + \frac{x^2}{2!} + \frac{x^4}{4!} + \cdots + \frac{x^{2n}}{(2n)!} + \cdots$
$\sin x$	$x - \frac{x^3}{3!} + \frac{x^5}{5!} - \cdots + (-1)^n \frac{x^{2n+1}}{(2n+1)!} + \cdots$
$\cos x$	$1 - \frac{x^2}{2!} + \frac{x^4}{4!} - \cdots + (-1)^n \frac{x^{2n}}{(2n)!} + \cdots$
$\log(1+x)$	$x - \frac{x^2}{2} + \frac{x^3}{3} - \cdots + (-1)^{n-1}\frac{x^n}{n} + \cdots$
$\operatorname{arctg} x$	$x - \frac{x^3}{3} + \cdots + (-1)^n \frac{x^{2n+1}}{2n+1} + \cdots$
$(1+x)^\alpha$	$1 + \alpha x + \binom{\alpha}{2}x^2 + \cdots + \binom{\alpha}{n}x^n + \cdots$ $\qquad \alpha \in \mathbb{R}$

In particolare:

$\frac{1}{1+x}$	$1 - x + x^2 - x^3 + \cdots + (-1)^n x^n + \cdots$	$(\alpha = -1)$
$\sqrt{1+x}$	$1 + \frac{x}{2} - \frac{x^2}{8} + \frac{x^3}{16} - \cdots + \frac{(-1)^{n+1}(2n-3)!!x^n}{(2n)!!}$	$(\alpha = \frac{1}{2})$
$\frac{1}{\sqrt{1+x}}$	$1 - \frac{x}{2} + \frac{3x^2}{8} - \frac{5x^3}{16} + \cdots + \frac{(-1)^n(2n-1)!!x^n}{(2n)!!}$	$(\alpha = -\frac{1}{2})$

ove

$$\binom{\alpha}{n} = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!} \qquad \text{(coefficiente binomiale generalizzato)}$$

$$k!! = k(k-2)(k-4)\cdots 2 \qquad (k \text{ semifattoriale})$$

7 TABELLA DI PRIMITIVE

$f(x)$	$F(x)$	
x^α	$\dfrac{x^{\alpha+1}}{(\alpha+1)}$	se $\alpha \neq -1$
$\dfrac{1}{x}$	$\log\lvert x\rvert$	
e^x	e^x	
a^x	$\dfrac{a^x}{\log a}$	$a>0\,,\quad a\neq 1$
$\sin x$	$-\cos x$	
$\cos x$	$\sin x$	
$\operatorname{tg} x$	$-\log\lvert\cos x\rvert$	
$\operatorname{cotg} x$	$-\log\lvert\sin x\rvert$	
$\operatorname{Sh} x$	$\operatorname{Ch} x$	
$\operatorname{Ch} x$	$\operatorname{Sh} x$	
$\operatorname{Th} x$	$\log \operatorname{Ch} x$	
$\operatorname{Cth} x$	$\log\lvert\operatorname{Sh} x\rvert$	
$\dfrac{1}{(\cos x)^2}$	$\operatorname{tg} x$	
$\dfrac{1}{(\sin x)^2}$	$-\operatorname{cotg} x$	
$\dfrac{1}{(\operatorname{Ch} x)^2}$	$\operatorname{Th} x$	
$\dfrac{1}{(\operatorname{Sh} x)^2}$	$-\operatorname{Cth} x$	
$\dfrac{1}{1+x^2}$	$\operatorname{arctg} x$	
$\dfrac{1}{1-x^2}$	$\dfrac{1}{2}\log\dfrac{1+x}{1-x}$	se $\lvert x\rvert<1$
$\dfrac{1}{1-x^2}$	$\dfrac{1}{2}\log\dfrac{x+1}{x-1}$	se $\lvert x\rvert>1$
$\dfrac{1}{\sqrt{1-x^2}}$	$\arcsin x$	
$\dfrac{1}{\sqrt{1+x^2}}$	$\log\lvert x+\sqrt{1+x^2}\rvert$	
$\dfrac{1}{\sqrt{x^2-1}}$	$\log\lvert x+\sqrt{x^2-1}\rvert$	se $\lvert x\rvert>1$

 © 978-88-08-06485-1

$f(x)$	$F(x)$
$f^\alpha f'$	$\dfrac{f^{\alpha+1}}{(\alpha+1)}$ se $\alpha \neq -1$
$\dfrac{f'}{f}$	$\log\lvert f\rvert$
$\dfrac{f'}{1+f^2}$	$\operatorname{arctg} f$
$\dfrac{f'}{\sqrt{1-f^2}}$	$\arcsin f$

B Grafici

1
$f(x) = H(x)$: gradino di Heaviside

2
$f(x) = \mathrm{sgn}(x)$: segno di x

3
$f(x) = |x|$

4
$f(x) = x_+ = \max\{0,x\}$

 © 978-88-08-06485-1

11
$f(x) = 1/(1-x^2)$

12
$f(x) = \sqrt{x^2-1}$

13
$f(x) = \sqrt{1-x^2}$

14
$\exp(-x)$
$\exp(x)$

15
0.6
0.5
$0.\overline{3}$
0.1
10
3
2
$1.\overline{6}$
0.8
1.25
1
$f(x) = a^x$, a reale > 0

16
Ch x
$\frac{1}{2}e^x$
Sh x

 © 978-88-08-06485-1

23
1/e
$f(x) = x^x = \exp(x \log x)$

24
$f(x) = \sin x$

25
$f(x) = \cos x$

26
$f(x) = \arcsin(x)$

27
$f(x) = \arccos(x)$

28
$f(x) = \mathrm{tg} x$

 © 978-88-08-06485-1

35
$f(x) = x \sin x$

36
$f(x) = \sin(1/x)$

37
$f(x) = x \sin(1/x)$

38
$n = 1$
$n = 2$
$n = 3$
$n = 4$
$f(x) = x^n \exp(-x)$

39
$n = 1$
$n = 2$
$n = 3$
$n = 4$
$f(x) = x^n \exp(-x^2)$

Indice analitico

A

accelerazione, 153
ampiezza, 64
angolo polare, 40
antiderivata, 267
antisimmetrica, proprietà, 19
appartenenza, 1
approssimazione, 28
– lineare, 208
– polinomiale, 213
arcocoseno, 80
arcoseno, 80
arcotangente, 81
area di una figura piana, 260
argomento
– di un numero complesso, 40
– principale, 41
aritmetizzazione parziale del simbolo di infinito, 100, 122
arrotondamento, 28
asintotico
– per funzioni, 130, 212
– per successioni, 104
asintoto
– obliquo, 114
– orizzontale, 112
– verticale, 116
asse
– immaginario, 37
– reale, 37
assioma
– di continuità, 22
– di Dedekind, 22

B

bacino di attrazione, 336
base naturale, 62, 102
Bernoulli, disuguaglianza di, 33
biforcazione, 346

C

calcolo numerico di un integrale, 294
campo, 20
– ordinato, 20
Cantor procedimento diagonale di, 30
caos deterministico, 343
carattere della serie, 231
cardinalità, 29
catenaria, 70
Cauchy
– condizione di, 315
– successione di, 315
codominio, 51
coefficiente
– angolare, 152
– binomiale, 16
complementare, 7
complesso coniugato, 37
condizione di Cauchy, 315
condizioni iniziali, 329
congiunzione logica, 8
coniugio, 38
connessione, 25
continuità uniforme, 316
controesempio, 10
controinversa, implicazione, 12
convergenza
– assoluta, 239
– semplice, 239
convoluzione, 310
coordinate polari, 40
corrispondenza biunivoca, 29
coseno, 63
– iperbolico, 69
costante di Feigenbaum, 347
cotangente, 63
crescita
– lineare, 134
– sopralineare, 134
– sottolineare, 134

 © 978-88-08-06485-1

curvatura, 195
cuspide, 158

D

De Moivre, formule di, 42
De Morgan, leggi di, 9
Dedekind, assioma di, 22
definitivamente, 89, 119
derivata, 152
– destra, 158
– di Schwarz, 342
– logaritmica, 166
– prima, 152
– sinistra, 158
diagramma
– a gradino, 334
– di fase, 334
differenza, tra due insiemi, 6
differenziale, 208
dimostrazione
– indiretta, 11
– per assurdo, 12
– per induzione, 32
dinamica, caotica, 343
dipolo, elettrico, 210
Dirac, impulso di, 311
discontinuità a salto, 117
disgiunzione logica, 8
disuguaglianza
– di Bernoulli, 33
– triangolare
– – in $\mathbb{C}$, 38
– – in $\mathbb{R}$, 24
dominio, 50

E

elasticità, 166
elemento, 1
energia cinetica, 211
enunciato, 9
epigrafico, 196
equazione
– alle differenze
– – autonoma, 333
– – del prim'ordine, 329
– – del second'ordine, 351
– – di ordine k, 328
– autonoma, 329, 333
– caratteristica, 352
– differenziale, 156
– logistica discreta, 326
– non omogenea, 356
– omogenea, 329, 351
esponenziale complesso, 250
estremo, 172
– inferiore, 22
– superiore, 22
– – proprietà dell', 22
Eulero, formule di, 250

F

fattoriale, 15
Feigenbaum, costante di, 347
Fibonacci, numeri di, 354
flesso, a tangente verticale, 158
forma
– algebrica dei numeri complessi, 36
– di indecisione, 100
– esponenziale dei numeri complessi, 251
– trigonometrica dei numeri complessi, 41
formula
– di De Moivre, 42
– di MacLaurin, 214
– di Newton, 16, 33
– di Taylor, 216
formule di Eulero, 250
frequenza, 64
Fresnel
– funzioni integrali di, 306
– integrale di, 304
funzione, 49
– composta, 75
– concava, 196
– – strettamente, 198
– continua, 117
– convessa, 196
– – strettamente, 198
– crescente, 55, 176
– – strettamente, 55
– decrescente, 55, 176
– – strettamente, 55
– definita a tratti, 74
– degli errori, 305
– derivabile, 152
– discontinua, 117
– dispari, 54
– elementare, 56
– esponenziale, 61
– generatrice, 334
– infinitesima, 111
– integrabile, 259, 296, 300
– – assolutamente, 299, 304

– integrale, 305
– – di Fresnel, 306
– inversa, 77, 143
– invertibile, 77, 143
– limitata, 53
– logaritmo, 61
– monotona, 55
– non crescente, 55
– non decrescente, 55
– pari, 54
– periodica, 55
– potenza, 56
– prolungamento per continuità di una, 129
– simmetrica, 54
– uniformemente continua, 316
funzioni
– asintotiche, 130
– iperboliche, 69
– – inverse, 82
– trigonometriche, 63
– – inverse, 80

G

Gauss, piano di, 37
gerarchia degli infiniti
– per funzioni, 132
– per successioni, 105

I

immagine, 51
implicazione, 3
– controinversa, 12
– universale, 9, 11
impulso, di Dirac, 311
inclusione, 3, 11
– stretta, 3
indice
– di sommatoria, 13
– muto, 14
infinitesimi dello stesso ordine, 104
infinitesimo di ordine
– inferiore, 104
– superiore, 104
infiniti dello stesso ordine, 104
infinito di ordine
– inferiore, 104
– superiore, 104
insieme, 1
– convesso, 196
– delle parti, 5
– infinito, 28
– limitato, 21
– – inferiormente, 21
– – superiormente, 21
– ordinato, 19
– totalmente ordinato, 19
– vuoto, 4
insiemi operazioni su, 6
instabilità
– di un punto d'equilibrio, 336
– di un'orbita periodica, 341
integrabilità delle funzioni continue, 318
integrale
– convergente, 296, 300
– definito, 258, 269, 305
– di Fresnel, 304
– divergente, 296, 300
– generalizzato, 296
– indefinito, 269, 305
integrazione
– delle funzioni irrazionali, 285
– delle funzioni razionali, 273
– delle funzioni trigonometriche, 281
– per parti, 277
– per scomposizione, 270
– per sostituzione, 270
intersezione di due insiemi, 6
intervallo, 24
– di periodicità, 55
intorno, 119

L

Lagrange, resto secondo, 219
lavoro di una forza, 289
legge delle controinverse, 12
leggi di De Morgan, 9
limite, 90
– destro, 115
– di funzione
– – definizione successionale, 111
– – definizione topologica, 119
– – per difetto, 112
– – per eccesso, 112
– di una successione, 90
– – per difetto, 93
– – per eccesso, 93
– sinistro, 115
limiti notevoli, 128
linearizzazione, 208, 337
logaritmo, 27
– nel campo complesso, 251

lunghezza di un grafico, 290

M

MacLaurin
– formula di, 214
– polinomio di, 214
maggiorante, 22
Malthus, modello di, 324
mantissa, 68
massimo
– di funzione, 138,172
– di un insieme, 21
– locale, 172
media
– aritmetica, 176
– geometrica, 176
Mengoli, serie di, 232
metodo
– del punto medio, 294
– di Newton, 221
– di somiglianza, 357
minimo
– di funzione, 138, 172
– di un insieme, 21
– locale, 172
minorante, 22
modello
– dell'acceleratore, 327
– della ragnatela, 332
– di Malthus, 324
– logistico, 326
– matematico, 323
modulo, 23
– di un numero complesso, 38
molteplicità di una radice di un'equazione, 46

N

negazione, 11
– logica, 8
Nepero, numero e di, 62, 101
Newton
– formula di, 16, 33
– metodo di, 221
numerabile, 29
numeri complessi, 6, 35
– – forma algebrica, 36
– di Fibonacci, 354
– immaginari puri, 37
– interi, 5
– naturali, 5
– razionali, 18
– reali, 5, 18
– – definizione assiomatica, 22
numero e di Nepero, 62, 101

O

o piccolo, 209, 212
orbita, 334
– instabile, 341
– periodica, 340
– stabile, 341
– – asintoticamente, 341
ordinamento di Sharkowsky, 342
oscillazione
– amplificata, 67
– smorzata, 67

P

parte
– decimale, 68
– immaginaria
– – di un numero complesso, 36
– intera, 68
– reale
– – di un numero complesso, 36
Peano, resto secondo, 215, 216
pendenza, 152
pendolo, 211
periodo, 55, 340
permutazione, 16
piano
– complesso, 37
– di Gauss, 37
polinomio
– di MacLaurin, 214
– di Taylor, 216
potenza
– a esponente complesso, 252
– a esponente razionale, 59
– a esponente reale, 61
– del continuo, 31
– di un insieme, 29
potenze a esponente razionale o reale, 26
predicato, 9
primitiva, 266
principio di induzione, 32
procedimento diagonale di Cantor, 30
prodotto cartesiano, 7
progressione geometrica, 15
prolungamento per continuità di una funzione, 129

proposizione, 9
proprietà, 9
– dell'estremo superiore, 22
pulsazione, 64
punto
– angoloso, 157
– di equilibrio, 335
– di flesso, 200
– – a tangente orizzontale, 174
– di massimo, 138, 172
– – locale, 172
– di minimo, 138, 172
– fisso, 335
– instabile, 336
– iperbolico, 339
– periodico, 340
– semi-stabile, 339
– stabile, 336
– – asintoticamente, 336
– stazionario, 174

Q

quantificatore
– esistenziale, 4
– universale, 3

R

radice n-esima
– aritmetica, 25
– complessa, 43
raggio
– di curvatura, 196
– polare, 40
rapporto incrementale, 151
relazione
– d'ordine, 19
– di equivalenza, 105
repulsore, 338
resto
– secondo Lagrange, 219
– secondo Peano, 215, 216
retta tangente, 152
riflessione di indici, 14
riflessiva, proprietà, 19, 105

S

scala semilogaritmica, 166
Schwarz, derivata di, 342
seno, 63
– iperbolico, 69
serie, 230
– a segni alterni, 240
– a termini di segno variabile, 239
– a termini non negativi, 233
– armonica, 232, 301
– – generalizzata, 234, 302
– carattere della, 231
– convergente, 231
– di funzioni, 245
– di Mengoli, 232
– di potenze, 248
– di Taylor, 245
– divergente, 231
– esponenziale, 246
– geometrica, 231
– – nel campo complesso, 249
– irregolare, 231
– maggiorante, 234
– minorante, 234
– somma della, 231
– telescopica, 232
– termine generale della, 231
settore
– coseno iperbolico, 82
– seno iperbolico, 82
sezione aurea, 356
Sharkowsky, ordinamento di, 342
simmetrica, proprietà, 105
sistema fisico lineare, 310
soluzione di equilibrio, 330
somma
– della serie, 231
– di Cauchy-Riemann, 259
– parziale, 230
sommatoria, 13
sottoinsieme, 3
sottosuccessione, 314
stabilità
– asintotica
– – di un punto d'equilibrio, 336
– – di un'orbita periodica, 341
– di un punto d'equilibrio, 336
– di un'orbita periodica, 341
successione, 87
– convergente, 89
– definita per ricorrenza, 222
– delle somme parziali, 230
– di Cauchy, 315
– divergente, 91
– geometrica, 96
– indeterminata, 92
– infinita, 92
– infinitesima, 92

© 978-88-08-06485-1

– irregolare, 92
– limitata, 89
– – inferiormente, 89
– – superiormente, 89
– monotona, 93, 181
successione limitata, 88
– inferiormente, 88
– superiormente, 88

T

tangente, 63
– iperbolica, 69
Tartaglia
– triangolo di, 17
tasso
– di incremento relativo, 166
– di variazione, 151
Taylor
– formula di, 216
– polinomio di, 216
– serie di, 245
termine
– generale
– – della serie, 231
transitiva, proprietà, 19, 105
traslazione di indici, 14
triangolo
– di Tartaglia, 17

U

uguaglianza
– tra insiemi, 3
unione
– di due insiemi, 6
unità
– immaginaria, 36

V

valore
– assoluto, 23
– – di un numero complesso, 38
– efficace, 266
– medio, 266
– principale
– – del logaritmo, 251
– – di una potenza, 252
variabile
– d'integrazione, 260
– di stato, 328
– dipendente, 50
– indipendente, 50
– logica, 9
– muta, 260
velocità, 153
vibrazione elementare, 64
volume di un solido di rotazione, 292

Marco Bramanti Carlo D. Pagani Sandro Salsa

Analisi matematica 1

Gli autori

Marco Bramanti è professore associato di Analisi matematica presso il Dipartimento di Matematica del Politecnico di Milano. I suoi interessi di ricerca vertono sulle equazioni alle derivate parziali e sull'analisi reale, in particolare sulla teoria degli integrali singolari.
Carlo Domenico Pagani è professore ordinario di Analisi matematica presso il Dipartimento di Matematica del Politecnico di Milano. Ha svolto attività didattica e di ricerca presso l'Università della California (Berkeley), l'Accademia delle Scienze (Russia), l'Accademia Sinica (Cina), il Tata Institute for Fundamental Research (India). Si interessa principalmente di equazioni alle derivate parziali e di problemi inversi.
Sandro Salsa è professore ordinario di Analisi matematica presso il Dipartimento di Matematica del Politecnico di Milano. Ha svolto attività di ricerca presso l'Università del Minnesota (Minneapolis), il Courant Institute (New York), l'Institute for Advanced Study (Princeton) e l'Università del Texas (Austin). Si occupa principalmente di equazioni a derivate parziali e problemi di frontiera libera.
Insieme sono autori anche di *Matematica – calcolo infinitesimale e algebra lineare* (Zanichelli, 2004).

L'opera

Analisi matematica di Bramanti, Pagani e Salsa è un corso per la formazione di base che riesce a conferire anche il giusto spazio all'approfondimento grazie ai rigorosi criteri didattici adottati:

- **Il minimo di astrazione necessaria** viene inserita per raggiungere l'obiettivo di conoscere, comprendere e saper utilizzare i contenuti fondamentali dell'analisi matematica.
- **Equilibrio tra sinteticità e chiarezza:** la giustificazione del risultato, quando non richieda un apparato formale troppo pesante, rende più consapevoli dei nessi logici.
- **Motivazione:** ogni nuovo concetto è introdotto attraverso esempi tratti dalle applicazioni più comuni e la teoria è accompagnata costantemente con riferimenti a problemi tratti da altre scienze, evidenziando il ruolo dello strumento matematico nella modellizzazione.
- **Nessuna separazione tra "teoria" e "pratica":** esempi, esercizi e applicazioni sono costantemente alternati alla presentazione teorica.
- **Modularità:** si è mantenuta la massima indipendenza possibile tra gli argomenti trattati, compatibilmente con la struttura logica del discorso matematico.

BRAMANTI"PAGANI"SALSA*ANALISI MAT

ISBN 978-88-08-06485-1

9 788808 064851

0 1 2 3 4 5 6 7 (60B)