Copyright © 2009 Zanichelli editore S.p.A., via Irnerio 34, 40126 Bologna [6485]

I diritti di elaborazione in qualsiasi forma o opera, di memorizzazione anche digitale, su supporti di qualsiasi tipo (inclusi magnetici e ottici), di riproduzione e di adattamento totale o parziale con qualsiasi mezzo (compresi i microfilm e le copie fotostatiche), i diritti di noleggio, di prestito e di traduzione sono riservati per tutti i paesi.
L'acquisto della presente copia dell'opera non implica il trasferimento dei suddetti diritti né li esaurisce.

Le fotocopie per uso personale (cioè privato e individuale) possono essere effettuate, nei limiti del 15% di ciascun volume, dietro pagamento alla S.I.A.E. del compenso previsto dall'art. 68, commi 4 e 5, della legge 22 aprile 1941 n. 633.
Tali fotocopie possono essere effettuate negli esercizi commerciali convenzionati S.I.A.E. o con altre modalità indicate da S.I.A.E.

Per le riproduzioni ad uso non personale (ad esempio: professionale, economico o commerciale) l'editore potrà concedere a pagamento l'autorizzazione a riprodurre un numero di pagine non superiore al 15% delle pagine del presente volume.
Le richieste per tale tipo di riproduzione vanno inoltrate a:

Associazione Italiana per i Diritti di Riproduzione
delle Opere dell'ingegno (AIDRO)
Corso di Porta Romana, 108
20122 Milano
e-mail segreteria@aidro.org e sito web www.aidro.org

L'editore, per quanto di propria spettanza, considera rare le opere fuori del proprio catalogo editoriale.
La riproduzione degli esemplari esistenti nelle biblioteche di tali opere è consentita, non essendo concorrenziale all'opera. Non possono considerarsi rare le opere di cui esiste, nel catalogo dell'editore, una successiva edizione, le opere presenti in cataloghi di altri editori o le opere antologiche.

Maggiori informazioni sul nostro sito: www.zanichelli.it/f_info_fotocopie.html

---

Rilettura critica e redazione: Beatrice Orlandini

Impaginazione e disegni: Compomat, Configni (RI)

Copertina:
– *Progetto grafico*: Miguel Sal & C, Bologna
– *Immagine di copertina*: © Mike Powell/Allsport Concepts /Getty Images.

Prima edizione: luglio 2009

Ristampa
5 4 3 2 1 2013 2012 2011 2010 2009

Realizzare un libro è un'operazione complessa, che richiede numerosi controlli: sul testo, sulle immagini e sulle relazioni che si stabiliscono tra essi.
L'esperienza suggerisce che è praticamente impossibile pubblicare un libro privo di errori. Saremo quindi grati ai lettori che vorranno segnalarceli.
Per segnalazioni o suggerimenti relativi a questo libro l'indirizzo a cui rivolgersi è:

Zanichelli editore S.p.A.
Via Irnerio 34
40126 Bologna
fax 051293322
e-mail: linea_universitaria@zanichelli.it
sito web: www.zanichelli.it

Prima di effettuare una segnalazione è possibile verificare se questa sia già stata inviata in precedenza, identificando il libro interessato all'interno del nostro catalogo on line (www.zanichelli.it/f_catalog.html) e selezionando il link ERRATA CORRIGE, dove sono disponibili le eventuali correzioni in formato PDF.

Per comunicazioni di tipo commerciale: universita@zanichelli.it

Stampa: Grafica Ragno
Via Lombardia 25, 40064 Tolara di Sotto, Ozzano Emilia (Bologna)
per conto di Zanichelli editore S.p.A.
Via Irnerio 34, 40126 Bologna

Marco Bramanti Carlo D. Pagani Sandro Salsa

Analisi matematica 2

ZANICHELLI

Indice

Prefazione vii

1 Equazioni differenziali 1
1 Modelli differenziali 2
2 Equazioni del primo ordine 3
2.1 Generalità 3
2.2 Equazioni a variabili separabili 5
2.3 Equazioni lineari del primo ordine 10
3 Equazioni lineari del secondo ordine 18
3.1 Spazi di funzioni 19
3.2 Generalità sulle equazioni lineari. Problema di Cauchy 20
3.3 La struttura dell'integrale generale 22
3.4 Equazioni omogenee a coefficienti costanti 26
3.5 Equazioni non omogenee 28
3.6 Vibrazioni meccaniche 35
4 Complementi 42
4.1 Teorema di esistenza e unicità per le equazioni a variabili separabili 42
4.2 Cenni alle equazioni lineari di ordine n 44

2 Calcolo infinitesimale per le curve 49
1 Richiami di calcolo vettoriale 49
2 Funzioni a valori vettoriali, limiti e continuità 52
3 Curve regolari e calcolo differenziale vettoriale 55
3.1 Esempi introduttivi 55
3.2 Arco di curva continua 58
3.3 Derivata di una funzione vettoriale. Arco di curva regolare 60
3.4 Integrale di una funzione a valori vettoriali 64
3.5 Alcune classi di curve piane 65
4 Lunghezza di un arco di curva 68
4.1 Curve rettificabili e lunghezza 68
4.2 Cambiamenti di parametrizzazione, curve equivalenti 72
4.3 Parametro arco o ascissa curvilinea 73
5 Integrali di linea (di prima specie) 74
6 Elementi di geometria differenziale delle curve 78
6.1 Curvatura e normale principale per una curva in $\mathbb{R}^m$ 78

© 978-88-08-06485-1

6.2 Calcolo della curvatura per curve nello spazio $\mathbb{R}^3$ o nel piano 82
6.3 Torsione e terna intrinseca per curve nello spazio $\mathbb{R}^3$ 85
7 Complementi 90
7.1 Lunghezza di una curva regolare 90
7.2 Alcune applicazioni fisiche notevoli 92

3 Calcolo differenziale per funzioni reali di più variabili 95
1 Grafici e insiemi di livello 95
2 Limiti e continuità per funzioni di più variabili 99
2.1 Definizioni e proprietà di limiti e funzioni continue 99
2.2 Calcolo dei limiti in più variabili: analisi delle forme di indeterminazione 102
3 Topologia in $\mathbb{R}^n$ e proprietà delle funzioni continue 107
3.1 Concetti fondamentali 108
3.2 Proprietà topologiche delle funzioni continue 115
4 Derivate parziali, piano tangente, differenziale 119
4.1 Derivate parziali 119
4.2 Piano tangente 122
4.3 Differenziabilità e approssimazione lineare 124
4.4 Derivate direzionali 130
4.5 Calcolo delle derivate 134
5 Derivate di ordine superiore e approssimazioni successive 143
5.1 Derivate di ordine superiore 143
5.2 Differenziale secondo, matrice hessiana, formula di Taylor al secondo ordine 148
6 Ottimizzazione. Estremi liberi 152
6.1 Generalità sui problemi di ottimizzazione 152
6.2 Estremi liberi. Condizioni necessarie del prim'ordine 156
6.3 Forme quadratiche. Classificazione 157
6.4 Forme quadratiche. Test degli autovalori 162
6.5 Studio della natura dei punti critici 165
7 Funzioni convesse di n variabili 178
7.1 Generalità sulle funzioni convesse 178
7.2 Ottimizzazione di funzioni convesse e concave 180
8 Funzioni definite implicitamente 183
8.1 Funzione implicita di una variabile 183
8.2 Funzione implicita di n variabili 188
9 Complementi 190
9.1 Topologia e funzioni continue 190
9.2 Funzioni omogenee 192
9.3 Differenziali e formula di Taylor di ordine superiore 197

4 Calcolo differenziale per funzioni di più variabili a valori vettoriali 201
1 Funzioni di più variabili a valori vettoriali: generalità 201
1.1 Superfici in forma parametrica 201
1.2 Trasformazioni di coordinate 203
1.3 Campi vettoriali 205

2 Limiti, continuità e differenziabilità per funzioni $f : \mathbb{R}^n \to \mathbb{R}^m$ 207
3 Superfici regolari in forma parametrica 210
4 Varietà k-dimensionali in $\mathbb{R}^n$ e funzioni definite implicitamente 218
4.1 Varietà k-dimensionali in $\mathbb{R}^n$ in forma parametrica 218
4.2 Funzioni implicite definite da sistemi di equazioni 220
4.3 Varietà k-dimensionali in $\mathbb{R}^n$ in forma implicita 222
5 Trasformazioni di coordinate e loro inversione 224
5.1 Il teorema della funzione inversa 224
5.2 Trasformazione di operatori differenziali 229
6 Ottimizzazione. Estremi vincolati 231
6.1 Vincoli di uguaglianza e moltiplicatori di Lagrange. Funzioni di due variabili 231
6.2 Moltiplicatori di Lagrange. Il caso generale 240
6.3 Vincoli di disuguaglianza e teorema di Kuhn-Tucker 246

5 Calcolo integrale per funzioni di più variabili 249
1 Integrali doppi 249
1.1 Integrale di una funzione limitata definita su un rettangolo 249
1.2 Funzioni integrabili su domini non rettangolari. Insiemi semplici, regolari, misurabili 256
1.3 Proprietà elementari dell'integrale doppio 261
1.4 Calcolo degli integrali doppi: metodo di riduzione 263
1.5 Calcolo degli integrali doppi: cambiamento di variabili 271
2 Integrali doppi generalizzati 276
3 Il calcolo degli integrali tripli 278
4 Derivazione sotto il segno di integrale 284
5 Complementi 286
5.1 La funzione Gamma di Eulero 286
5.2 Definizioni e proprietà elementari degli integrali in $\mathbb{R}^n$ 289

6 Campi vettoriali 293
1 Campi vettoriali e integrali di linea di seconda specie 293
1.1 Linee di campo 293
1.2 Gradiente, rotore e divergenza 295
1.3 Integrale di linea di un campo vettoriale. Lavoro e circuitazione. 299
1.4 Campi conservativi e potenziali 301
1.5 Campi irrotazionali. Insiemi semplicemente connessi 305
1.6 Campi solenoidali e potenziale vettore 310
1.7 Il linguaggio delle forme differenziali 314
2 Formula di Gauss-Green nel piano 316
3 Area e integrali di superficie 320
3.1 Area di una superficie 320
3.2 Integrale di superficie di una funzione continua 325
4 Integrale di superficie di un campo vettoriale. Flusso 327
4.1 Superfici orientate. Bordo di una superficie. Superfici regolari a pezzi 328
4.2 Flusso 331

5 Teorema della divergenza 334
6 Teorema del rotore 339

7 Serie di potenze e serie di Fourier 347
1 Serie di funzioni e convergenza totale 347
2 Serie di potenze 354
2.1 Proprietà fondamentali delle serie di potenze 354
2.2 Serie di Taylor e serie di potenze 361
3 Serie trigonometriche e serie di Fourier 367
3.1 Polinomi trigonometrici e serie trigonometriche 367
3.2 Richiami sugli spazi vettoriali con prodotto scalare 371
3.3 Coefficienti e serie di Fourier di una funzione. Approssimazione in media quadratica 374
3.4 Esempi e osservazioni sul calcolo dei coefficienti di Fourier 379
3.5 Forma esponenziale complessa delle serie di Fourier 383
3.6 Convergenza puntuale delle serie di Fourier 386
3.7 Alcune interpretazioni fisiche 394
3.8 Applicazioni alle equazioni differenziali della fisica matematica. Metodo di separazione delle variabili 397
4 Complementi 409
4.1 Il metodo di Frobenius per la soluzione delle equazioni differenziali 409
4.2 Criteri per la convergenza delle serie trigonometriche 412
4.3 Fenomeno di Gibbs 414

8 Teoria qualitativa di equazioni differenziali e sistemi 417
1 Equazioni del prim'ordine 417
1.1 Problema di Cauchy 417
1.2 Alcune classi di equazioni del prim'ordine 427
1.3 Equazioni autonome. Diagrammi di fase. Stabilità 430
2 Problema di Cauchy per sistemi o equazioni di ordine n 435
3 Complementi 441
3.1 Lemma di Gronwall e dipendenza continua 441
4 Sistemi autonomi bidimensionali 443
4.1 Generalità 443
4.2 Stabilità per sistemi autonomi lineari 454
4.3 Stabilità per sistemi autonomi non lineari 461

A Trasformata di Laplace e trasformata di Fourier 467
1 Trasformata di Laplace. Definizione ed esempi 467
2 Proprietà della trasformata di Laplace 470
3 Trasformazione inversa di Laplace 477
4 Funzione di trasferimento di un sistema 478
5 Trasformata di Fourier 481
6 Proprietà della trasformazione di Fourier 484
7 Una applicazione: studio di un circuito RC 488

Prefazione

Questo secondo volume di Analisi matematica, che segue a distanza di un anno il primo, ne mantiene l'impostazione generale e le scelte didattiche di fondo: una trattazione degli argomenti fondamentali dell'analisi che unisca il rigore a una certa sintesi e modularità, per favorire l'effettiva utilizzabilità del testo da parte di docenti e studenti.

Per quanto riguarda gli argomenti, si comincia qui con un'introduzione elementare alle equazioni differenziali ordinarie; si prosegue, gradualmente, con il calcolo differenziale in più variabili, prima per le curve, poi per le funzioni di più variabili a valori scalari e quindi vettoriali; si passa poi al calcolo integrale in due o tre variabili; un capitolo sui campi vettoriali completa il calcolo differenziale e integrale con gli importanti teoremi di Gauss-Green-Stokes e i necessari complementi sulle superfici. A questo punto si ritorna alle funzioni di una variabile con un capitolo su serie di potenze e di Fourier e un successivo approfondimento su equazioni differenziali e sistemi.

Alcune caratteristiche che abbiamo cercato di imprimere a questo secondo volume e che riteniamo utile segnalare sono le seguenti.

1. La trattazione del calcolo differenziale in più variabili comprende sempre l'enunciato di definizioni e teoremi in un numero qualunque di variabili, affinché lo studente possa impadronirsi di concetti e notazioni di portata generale; il caso bidimensionale rimane comunque una linea guida intuitiva, utilizzata sistematicamente in esempi, esercizi e spesso nella prima introduzione dei nuovi concetti, per favorire la gradualità ed anche un eventuale utilizzo del testo ad un livello meno approfondito. Agli argomenti più propriamente di base se ne affiancano altri, sempre tradizionali ma più impegnativi (elementi di geometria differenziale delle curve, funzioni convesse, funzioni implicite definite da sistemi di equazioni, ottimizzazione vincolata...) che ricevono qui un'ampiezza maggiore di quanto avessero nel precedente volume unico *Matematica*. Così pure è stato dato più spazio al calcolo differenziale e integrale vettoriale e alla trattazione delle serie di potenze e di Fourier. Di queste ultime si è voluto illustrare brevemente l'applicazione al metodo di separazione di variabili, che costituisce un esempio significativo di utilizzo di vari concetti introdotti nell'intero corso.
2. Coerentemente all'impostazione del volume 1, in tutto il testo è stata fornita una dimostrazione dei teoremi enunciati ovunque questo fosse possibile senza far uso di un eccessivo apparato fondazionale.
3. In molti capitoli sono presenti dei *Complementi* (la cui omissione non pregiudica la comprensione del testo) che contengono, caso per caso, dimostrazioni un po' delicate di alcuni risultati enunciati nel testo, generalizzazioni, oppure argomenti particolari ma significativi, spesso in vista di qualche applicazione.

© 978-88-08-06485-1

4. Un discorso a parte merita l'argomento delle equazioni differenziali, a cui si è deciso di dedicare due distinti capitoli (1 e 8). Nel primo, di carattere elementare e che utilizza solo gli strumenti e il linguaggio del volume 1, introdurremo i concetti fondamentali e i metodi di risoluzione esplicita per le equazioni del primo ordine lineari, in particolare a coefficienti costanti, ed esempi di equazioni lineari di ordine superiore. Nel capitolo 8, facendo uso degli strumenti del calcolo differenziale in più variabili, riprenderemo lo studio delle equazioni differenziali da un punto di vista meno elementare, più vicino a quello che oggi si chiama *teoria dei sistemi*. Accanto ai classici elementi della teoria generale di equazioni e sistemi (questioni di esistenza e unicità, prolungabilità, ecc.), si darà spazio allo studio qualitativo e in particolare al problema della stabilità degli equilibri. La scelta di svolgere questo argomento in due capitoli[1], di stile e contenuto diversi, può venire incontro anche all'esigenza di qualche docente che intenda svolgere soltanto la prima parte (elementare), tralasciando la seconda, per far posto ad altri argomenti pure importanti, tenendo presente anche la possibilità che alcuni allievi, nella continuazione dei loro studi, scelgano un corso specifico di Equazioni Differenziali. Chi invece decida di svolgere questo argomento in modo un po' più approfondito potrà facilmente integrare il materiale presente nel primo capitolo con tutto o parte del materiale del capitolo 8.

Un'ultima osservazione. Nel capitolo 7 del volume 1 abbiamo parlato di modelli matematici discreti e di equazioni alle differenze, mentre in questo e nel capitolo 8 parliamo di modelli differenziali e di equazioni differenziali. Spesso i modelli presi in esame sono gli stessi, presentati in versione "discreta" nel volume 1 e in versione "continua" nel volume 2. Tuttavia abbiamo preferito svolgere la trattazione in maniera del tutto indipendente, in particolare senza assumere, da parte del lettore, la conoscenza del capitolo 7 (volume 1) per lo studio delle equazioni differenziali. Ciò comporta qualche ripetizione (come pure esiste qualche inevitabile ripetizione tra i capitoli 1 e 8 di questo volume), compensata però da una maggiore praticità nell'utilizzo del testo.

Saremo grati ai colleghi e agli studenti che vorranno farci avere i loro commenti anche su questo secondo volume.

Giugno 2009

Gli Autori

[1] Segnaliamo per inciso che al materiale sulle equazioni differenziali ordinarie presente nei capitoli 1 e 8 va aggiunto il *metodo di Frobenius* (di integrazione per serie di potenze), trattato brevemente nel capitolo 7, complementi.

1 Equazioni differenziali

Come abbiamo visto nel volume 1, capitolo 4, il concetto di derivata di una funzione si presta a numerose interpretazioni fisiche. Ricordando solo uno degli esempi fatti, se $x(t)$ rappresenta la posizione all'istante t di una particella che si muove lungo una retta, le derivate $x'(t)$ e $x''(t)$ rappresentano, rispettivamente, la velocità e l'accelerazione della particella. Non c'è da stupirsi, allora, se la traduzione matematica di molte leggi che governano fenomeni naturali porti a scrivere equazioni che coinvolgono una funzione incognita e qualcuna delle sue derivate: è questo il concetto di *equazione differenziale*. Numerosissimi modelli matematici sviluppati negli ultimi trecento anni dalle scienze applicate (fisica, biologia, ecologia, economia...) si formulano per mezzo di equazioni differenziali. Si parla anche, perciò, di *modelli differenziali*. In matematica si distinguono le *equazioni differenziali ordinarie* e le *equazioni differenziali alle derivate parziali*. Pur limitando la nostra attenzione alle prime[1], l'argomento è comunque molto vasto. Noi tratteremo l'argomento in due capitoli; nel primo, cioè questo, tratteremo da un punto di vista elementare alcune classi di equazioni che lo studente incontra frequentemente in moltissime applicazioni, come certe equazioni del 1° ordine, equazioni del 2° ordine lineari, in particolare a coefficienti costanti, ed esempi di equazioni lineari di ordine superiore. Per queste equazioni ci concentreremo sui metodi elementari di risoluzione esplicita. Gli strumenti e i metodi saranno qui quelli utilizzati nel volume 1 per lo studio di funzioni di una sola variabile reale. Più avanti, nel capitolo 8, dopo aver acquisito gli strumenti del calcolo differenziale in più variabili, riprenderemo lo studio delle equazioni differenziali da un punto di vista meno elementare, occupandoci di dare un quadro teorico più preciso della teoria delle equazioni e dei sistemi di equazioni differenziali, per affrontare poi in particolare il tema della stabilità degli equilibri.

[1] alle equazioni alle derivate parziali dedicheremo solo qualche cenno nel capitolo 3, paragrafo 5 (esempi) e nel capitolo 7, paragrafo 3.8 (metodo di separazione delle variabili).

© 978-88-08-06485-1

1 MODELLI DIFFERENZIALI

Si dice *equazione differenziale*[2] *di ordine n* un'equazione del tipo:

$$F\left(t, y, y', y'', \ldots, y^{(n)}\right) = 0 \tag{1.1}$$

dove $y(t)$ è la funzione incognita e F è una funzione assegnata delle $n+2$ variabili $t, y, y', y'', \ldots, y^{(n)}$, a valori reali. L'*ordine* dell'equazione è l'ordine massimo di derivazione che vi compare. Si dirà *soluzione*, o *(curva) integrale*, della (1.1), nell'intervallo $I \subset \mathbb{R}$, una funzione $\varphi(t)$, definita almeno in I e a valori reali, per cui risulti[3]

$$F\left(t, \varphi(t), \varphi'(t), \varphi''(t), \ldots, \varphi^{(n)}(t)\right) = 0 \quad \forall t \in I. \tag{1.2}$$

Infine, si dirà *integrale generale* dell'equazione (1.1) una formula che rappresenti la famiglia di tutte le soluzioni dell'equazione (1.1), eventualmente al variare di uno o più parametri in essa contenuti.

Mostriamo un primo esempio elementare, ma storicamente significativo, di modello differenziale.

Esempio

Modello di Malthus per la dinamica delle popolazioni, 1798. Si considera una popolazione che evolve isolata e i cui unici fattori di evoluzione sono fertilità e mortalità. Indicheremo con $N(t)$ il numero di individui presenti al tempo t, con λ (rispettivamente, μ) il numero di nuovi nati (di morti) per individuo nell'unità di tempo, cosicché in un tempo di durata h il numero di nuovi nati e di morti sarà rispettivamente $\lambda h N(t)$ e $\mu h N(t)$. Perciò, la variazione del numero di individui in un tempo h sarà:

$$N(t+h) - N(t) = \lambda h N(t) - \mu h N(t).$$

Dividendo ambo i membri della precedente equazione per h, abbiamo:

$$\frac{N(t+h) - N(t)}{h} = (\lambda - \mu) N(t). \tag{1.3}$$

Assumiamo valida la (1.3) per ogni intervallo di tempo h; prendendo il limite di ambo i membri per $h \to 0$ abbiamo (ammesso che la funzione $N(t)$ sia derivabile)

$$\dot{N}(t) = (\lambda - \mu) N(t). \tag{1.4}$$

La (1.4) è un'equazione differenziale lineare del primo ordine[4]; il numero $\varepsilon = \lambda - \mu$ si chiama *potenziale biologico*. Scrivendo la (1.4) nella forma $\dot{N}/N = \varepsilon$, si osserva che il tasso relativo di crescita (o diminuzione) di N è costante.

[2] In tutto questo capitolo ci riferiremo alle *equazioni differenziali ordinarie* chiamandole semplicemente *equazioni differenziali*.

[3] Si noti la differenza tra le due scritture (1.1) e (1.2): mentre (1.1) indica l'equazione differenziale – che deve ancora essere risolta –, la (1.2) indica l'identità che si ottiene sostituendo ad $y, y', \ldots, y^{(n)}$ la soluzione $\varphi(t)$ trovata e le sue derivate calcolate effettivamente, che ora sono funzioni note di t.

[4] Nelle equazioni differenziali della fisica, le derivate vengono spesso indicate col punto, scrivendo $\dot{N}, \ddot{N}$ anziché N', N''. Questo uso risale a Newton.

© 978-88-08-06485-1

La (1.4) non ammette una sola soluzione, ma infinite; come vedremo, esse sono tutte della forma:

$$N(t) = c e^{\varepsilon t} \tag{1.5}$$

dove c è una qualsiasi costante reale. Notiamo esplicitamente cosa significa che la (1.5) è soluzione della (1.4): se nella (1.4) sostituiamo ad $N(t)$ la sua espressione $c e^{\varepsilon t}$ e a $\dot{N}(t)$ la sua espressione $c\varepsilon e^{\varepsilon t}$ (calcolata dalla (1.5)), otteniamo un'identità. Se si conosce il numero di individui presenti in un dato istante, per esempio, $N(0) = N_0$ (*condizione iniziale*) si può selezionare, tra le soluzioni (1.4), quella (unica) soddisfacente alla condizione data: risulta così

$$N(t) = N_0 e^{\varepsilon t}. \tag{1.6}$$

La (1.6) è una formula esplicita che ci consente di conoscere, in ogni istante $t \geq 0$ (ma anche $t < 0$), il numero di individui presenti nella popolazione; ci consente anche di sapere come sarà l'evoluzione finale ($t \to +\infty$) della popolazione: essa tenderà all'estinzione ($N(t) \to 0$) se $\varepsilon < 0$, mentre crescerà esponenzialmente se $\varepsilon > 0$.

Mostriamo come si arriva a determinare le soluzioni (1.5) dell'equazione (1.4). Anticiperemo così facendo, in un caso particolare, un ragionamento che vedremo più in generale in seguito.

L'equazione $\dot{N}/N = \varepsilon$ (equivalente alla (1.4) nell'ipotesi $N \neq 0$) si può riscrivere

$$\frac{d}{dt}(\log|N|) = \varepsilon$$

da cui

$$\log|N| = \varepsilon t + c_1$$

$$|N(t)| = e^{c_1} e^{\varepsilon t} = k^2 e^{\varepsilon t}$$

con $k^2 = e^{c_1}$ costante positiva arbitraria, e infine

$$N(t) = \pm k^2 e^{\varepsilon t}, \text{ cioè } N(t) = c e^{\varepsilon t}$$

con c costante arbitraria diversa da zero. Questa soluzione non si annulla mai, come richiesto dal fatto che per dedurla abbiamo diviso per N. D'altro canto se $N(t) \equiv 0$ abbiamo ancora una soluzione della (1.4); possiamo ritenere, allora, che nella (1.5) c sia una costante reale *qualsiasi* (anche nulla).

2 EQUAZIONI DEL PRIMO ORDINE

2.1 Generalità

Consideriamo ora equazioni del tipo:

$$F(t, y, y') = 0. \tag{2.1}$$

Ad esempio, la ricerca delle primitive di una funzione f continua su I equivale a risolvere l'equazione differenziale del primo ordine

$$y'(t) = f(t)$$

© 978-88-08-06485-1

che ha infinite soluzioni del tipo:

$$y(t) = \int f(t)dt + c \qquad c \in \mathbb{R}.$$

Anche l'equazione (1.4), come abbiamo detto, ha infinite soluzioni dipendenti da una costante arbitraria c. Si dimostra che l'insieme delle soluzioni di un'equazione differenziale del primo ordine è costituito da una famiglia di funzioni, dipendente da un parametro c: $t \mapsto \varphi(t; c)$; tale famiglia prende il nome di *integrale generale* dell'equazione. La condizione supplementare

$$y(t_0) = y_0 \tag{2.2}$$

permette, in generale, di selezionare una *soluzione particolare*. Il problema di risolvere le equazioni (2.1), (2.2) prende il nome di *problema di Cauchy*.

Quando l'equazione (2.1) si presenta nella forma

$$y'(t) = f(t, y(t)) \tag{2.3}$$

si dice che è in *forma normale*. Per equazioni di questo tipo si può assicurare, sotto larghe ipotesi, che il problema di Cauchy (2.3), (2.2) ammette un'unica soluzione, almeno *localmente*, cioè per valori di t in un intorno del punto t_0 in cui è assegnata la condizione iniziale (cfr. capitolo 8, paragrafo 1.1).

Si noti che le soluzioni dell'equazione differenziale, espresse dall'integrale generale, potrebbero essere definite su insiemi diversi a seconda del valore della costante, e talvolta su insiemi più complicati di un intervallo (ad esempio: per ogni $t \neq 0$). Tuttavia, quando si parla di soluzione di un problema di Cauchy, si intende sempre una funzione che:

a) è definita su un intervallo I, contenente il punto t_0 in cui è assegnata la condizione iniziale;

b) è derivabile in tutto I e soddisfa l'equazione in tutto I.

L'idea di soluzione di un problema di Cauchy corrisponde infatti a quella di evoluzione di un sistema, soggetto a una certa legge differenziale e a una certa condizione iniziale. Questa interpretazione ha senso solo se la variabile t varia con continuità, *cioè in un intervallo*[5], a partire dall'istante iniziale t_0.

Qui di seguito vedremo qualche classe di equazioni per cui l'integrale è ricavabile in forma esplicita. Per la maggior parte delle equazioni, tuttavia, ciò non è possibile ed allora si può ricorrere a metodi di integrazione approssimati, di cui si occupa il *calcolo numerico*.

Esempio

Il problema di Cauchy:

$$\begin{cases} N'(t) = 2N(t) \\ N(0) = 5 \end{cases}$$

[5] Nel capitolo 8, paragrafo 1.1, approfondiremo il concetto di intervallo massimale su cui è definita la soluzione del problema di Cauchy.

© 978-88-08-06485-1

ha un'unica soluzione, data per quanto visto in precedenza da:

$$N(t) = 5e^{2t} \quad \text{per ogni } t \in \mathbb{R}.$$

Si può dimostrare che l'integrale generale dell'equazione

$$xy' + y = 0$$

è dato dalla famiglia di iperboli

$$y = \frac{c}{x} \tag{2.4}$$

definite per $x \neq 0$ (che queste funzioni risolvano l'equazione si verifica immediatamente; il punto è provare che non ce ne sono altre). Supponiamo di voler risolvere il problema di Cauchy:

$$\begin{cases} xy' + y = 0 \\ y(1) = 2. \end{cases}$$

Imponendo la condizione $y(1) = 2$ nella (2.4) si trova

$$c = 2$$

da cui

$$y = \frac{2}{x}.$$

Ora, questa funzione è la soluzione del problema di Cauchy sull'intervallo $(0, +\infty)$ (che contiene il punto $x = 1$ in cui è assegnata la condizione iniziale). Sarebbe un errore affermare che la soluzione del problema di Cauchy è $y = 2/x$ per ogni $x \neq 0$.

2.2 Equazioni a variabili separabili

Le equazioni a variabili separabili sono equazioni del tipo:

$$y' = a(t)\, b(y) \tag{2.5}$$

con a continua in $I \subset \mathbb{R}$ e b continua in $J \subset \mathbb{R}$. Osserviamo anzitutto che se il numero $\overline{y}$ è una soluzione dell'equazione $b(y) = 0$, la funzione costante $y(t) = \overline{y}$ è una soluzione dell'equazione differenziale. Infatti in tal caso il secondo membro si annulla perché $b(\overline{y}) = 0$, e il primo membro si annulla, perché la derivata della funzione costante è zero. Supponendo invece $b(y) \neq 0$, la (2.5) si può riscrivere nella forma

$$\frac{y'}{b(y)} = a(t).$$

Un'ipotetica soluzione $y(t)$ soddisfa dunque l'identità

$$\frac{y'(t)}{b(y(t))} = a(t).$$

Prendendo gli integrali indefiniti di ambo i membri, si ottiene:

$$\int \frac{y'(t)}{b(y(t))} dt = \int a(t)\, dt + c.$$

© 978-88-08-06485-1

Ora nell'integrale a primo membro si può effettuare il cambio di variabile $y = y(t)$; $dy = y'(t)dt$ ottenendo

$$\int \frac{dy}{b(y)} = \int a(t)\,dt + c \tag{2.6}$$

con c costante arbitraria.

La (2.6) assegna l'integrale generale della (2.5). Ciò significa che se $B(y)$ è una primitiva di $\frac{1}{b(y)}$ e $A(t)$ è una primitiva di $a(t)$, l'integrale generale della (2.5) è assegnato *implicitamente* dall'equazione

$$B(y) = A(t) + c$$

con c costante arbitraria.

Infine, se si riesce a ricavare esplicitamente y dall'ultima equazione (cioè se si sa scrivere la funzione inversa di B, B^{-1}) si ottiene:

$$y = B^{-1}(A(t) + c)$$

cioè un'espressione del tipo

$$y = F(t, c)$$

(in cui la costante arbitraria c può comparire in qualsiasi forma, anche non additiva).

In pratica, data un'equazione a variabili separabili, non è detto che si sappiano determinare esplicitamente le primitive $B(y)$, $A(t)$, e anche in caso positivo non è detto che si sappia esprimere y nella forma $y = F(t, c)$. Ciò sarà possibile in casi particolarmente semplici.

Ricordare la (2.6) è facile: basta pensare di scrivere $y'(t) = \frac{dy}{dt}$ e, trattandolo formalmente come un quoziente, trasformare la (2.5) nella

$$\frac{dy}{b(y)} = a(t)\,dt$$

che poi si integra formalmente membro a membro, arrivando alla (2.6).

Per le equazioni a variabili separabili vale il seguente risultato di esistenza e unicità della soluzione del problema di Cauchy:

TEOREMA 1.1 (PROBLEMA DI CAUCHY PER UN'EQUAZIONE A VARIABILI SEPARABILI) *Si consideri il problema di Cauchy:*

$$\begin{cases} y' = a(t)\,b(y) \\ y(t_0) = y_0 \end{cases} \tag{2.7}$$

dove a è una funzione continua in un intorno I di t_0 e b è una funzione continua in un intorno J di y_0. Allora esiste un intorno di t_0 $I' \subset I$ e una funzione[6] *$y \in C^1(I')$ soluzione del problema (2.7).*

Se inoltre $b \in C^1(J)$ (o, più in generale, b è una funzione continua, eventualmente non derivabile ma con rapporto incrementale limitato in J), allora tale soluzione è unica.

[6] La scrittura $y \in C^1(I')$ significa che y è derivabile in I' e y' è continua in I'. Approfondiremo nel paragrafo 3.3 la discussione dello spazio di funzioni C^1.

© 978-88-08-06485-1

Una dimostrazione di questo teorema (che è un caso particolare del teorema enunciato nel capitolo 8, paragrafo 1.1) sarà fornita nei Complementi in fondo al capitolo.

Esempi

Nota bene. *Nello svolgimento della teoria, abbiamo sempre indicato con y la funzione incognita e con t la variabile indipendente. Negli esempi e negli esercizi saranno utilizzate anche notazioni diverse, che risulteranno chiare dal contesto.*

2.3 Risolvere il problema di Cauchy:

$$\begin{cases} yy' = 1 \\ y(0) = 2. \end{cases}$$

L'equazione si integra scrivendo:

$$\int y\,dy = \int dx$$

$$\frac{1}{2}y^2 = x + c$$

$$y = \pm\sqrt{2(x + c)}.$$

Si osservi che per ogni valore di $c \in \mathbb{R}$ esistono due soluzioni (corrispondenti ai due segni davanti alla radice), definite solo per $x \geq -c$.

Imponendo la condizione iniziale si ha:

$$2 = \pm\sqrt{2c}$$

perciò occorre scegliere il segno + davanti alla radice, e porre $c = 2$. Si ottiene quindi:

$$y = \sqrt{2x + 4},$$

unica soluzione del problema di Cauchy, definita per $x \geq -2$.

Osserviamo che in questo caso $b(y) = 1/y$ è derivabile con continuità per $y \neq 0$, mentre $a(t) = 1$ è continua ovunque. Il problema di Cauchy per questa equazione ha dunque una e una sola soluzione purché la condizione iniziale *non sia* del tipo

$$y(t_0) = 0.$$

Infatti, l'equazione $yy' = 1$ non può essere soddisfatta in tale punto t_0, perché darebbe $0 = 1$. Dunque il problema di Cauchy

$$\begin{cases} yy' = 1 \\ y(t_0) = 0 \end{cases}$$

non ammette alcuna soluzione. In altre situazioni in cui sono violate le ipotesi del teorema, potrebbe cadere l'*unicità* di tale soluzione (si vedano gli esercizi successivi).

2.4 Sia data l'equazione

$$y' = 2t\sqrt{1 - y^2}.$$

Le rette $y = 1$ e $y = -1$ sono soluzioni. Le altre soluzioni sono date dalla formula:

$$\int \frac{dy}{\sqrt{1 - y^2}} = 2\int t\,dt + c$$

cioè

$$\arcsin y = t^2 + c$$

© 978-88-08-06485-1

e quindi, ricordando che $\arcsin y$ è definito solo per $-1 \le y \le 1$ e che in questo intervallo è la funzione inversa di $y = \sin x$, abbiamo

$$y = \sin(t^2 + c) \text{ purché } -\frac{\pi}{2} \le t^2 + c \le \frac{\pi}{2}.$$

Se dovessimo risolvere, ad esempio, il problema di Cauchy

$$\begin{cases} y' = 2t\sqrt{1-y^2} \\ y(\sqrt{\pi}) = \dfrac{1}{2} \end{cases}$$

procederemmo così:

a) Si impone la condizione iniziale nell'integrale generale $y = \sin(t^2 + c)$:

$$\frac{1}{2} = \sin(\pi + c).$$

b) Si determina c: poiché dev'essere $-\frac{\pi}{2} \le \pi + c \le \frac{\pi}{2}$, si trova:

$$\pi + c = \frac{\pi}{6} \qquad c = -\frac{5\pi}{6}.$$

c) Si sostituisce il valore trovato di c nell'integrale generale e si trova così la soluzione del problema:

$$y = \sin(t^2 - \frac{5\pi}{6})$$

per $-\frac{\pi}{2} \le t^2 - \frac{5\pi}{6} \le \frac{\pi}{2}$, quindi $\frac{\pi}{3} \le t^2 \le \frac{4}{3}\pi$ ossia:

$$\sqrt{\frac{\pi}{3}} \le t \le \sqrt{\frac{4}{3}\pi} \quad \text{oppure} \quad -\sqrt{\frac{4\pi}{3}} \le t \le -\sqrt{\frac{\pi}{3}}.$$

Ma la soluzione del problema di Cauchy dev'essere definita in un intervallo contenente il punto iniziale $t_0 = \sqrt{\pi}$; in definitiva la soluzione del problema è:

$$y = \sin\left(t^2 - \frac{5\pi}{6}\right) \quad \text{per} \quad \sqrt{\frac{\pi}{3}} \le t \le \sqrt{\frac{4}{3}\pi}.$$

Equazione logistica. Consideriamo l'equazione

$$y' = ay(1 - by) \qquad a, b \text{ costanti} > 0.$$

Le rette $y = 0$ e $y = \frac{1}{b}$ sono soluzioni. Le altre soluzioni sono date dalla formula:

$$\int \frac{dy}{y(1-by)} = \int a\,dt + c.$$

Poiché

$$\frac{1}{y(1-by)} = \frac{1}{y} + \frac{b}{1-by} \quad \text{e} \quad \int \frac{dy}{y(1-by)} = \log\left|\frac{y}{1-by}\right|$$

le soluzioni sono date da:

$$\log\left|\frac{y}{1-by}\right| = at + c, \text{ con } c \text{ costante arbitraria}$$

$$\left|\frac{y}{1-by}\right| = c_0\, e^{at}, \text{ con } c_0 = e^c \text{ costante positiva}$$

$$y = \frac{c_1\, e^{at}}{1 + bc_1\, e^{at}}, \text{ con } c_1 \text{ costante arbitraria.}$$

© 978-88-08-06485-1

Nell'Esempio 1.1 abbiamo visto il modello di Malthus per la dinamica di una popolazione; tale modello è troppo semplicistico e chiaramente irrealistico nella maggioranza dei casi: una popolazione più numerosa comporta minori risorse e ciò implica un minore tasso di crescita. Nel 1845 Verhulst propose la seguente legge di evoluzione (al posto della (1.4)):

$$\dot{N}(t) = \varepsilon N(t)\left(1 - \frac{N(t)}{k}\right) \qquad (\varepsilon > 0,\ k > 0). \tag{2.8}$$

La (2.8) è un'equazione logistica (v. Esempio precedente), con $a = \varepsilon$ e $b = 1/k$. La soluzione (positiva) tale che $N(0) = N_0 > 0$ è data da

$$N(t) = \frac{kN_0\, e^{\varepsilon t}}{k - N_0 + N_0\, e^{\varepsilon t}}.$$

Il comportamento della soluzione al variare di t è illustrato in figura 1.1 distinguendo i casi $N_0 \gtrless k$; si vede che $N(t) \to k$ per $t \to +\infty$; la costante k si chiama *capacità dell'ambiente*. Il caso di Malthus corrisponde ad assumere $k = \infty$.

Le soluzioni $N(t) = 0$ e $N(t) = k$ si chiamano *soluzioni di equilibrio*. Poiché $N(t) \to k$ per ogni $N_0 > 0$ (quando $t \to +\infty$) la soluzione $N(t) = k$ si dice *asintoticamente stabile*.

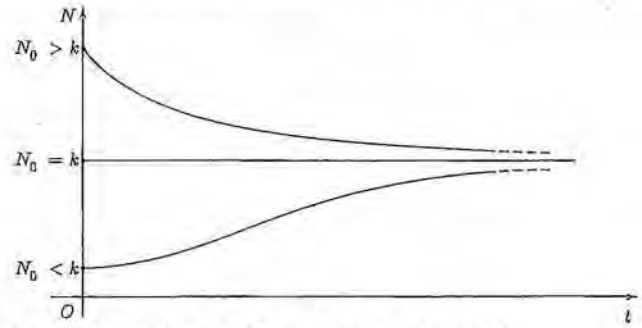


Figura 1.1. Soluzioni dell'equazione logistica per diverse condizioni iniziali.

Caduta di un grave nel vuoto e nell'aria. Lo spazio y percorso in verticale da un corpo di massa m in caduta libera nel vuoto soddisfa l'equazione:

$$my'' = mg$$

con g = accelerazione di gravità, costante. Si trova subito $y = \frac{1}{2}gt^2 + at + b$ con a, b costanti arbitrarie.

Se il corpo cade nell'aria (situazione più realistica) e la resistenza è proporzionale alla velocità, il modello si modifica così:

$$my'' = mg - hy'$$

(la costante positiva h tiene conto dell'attrito dell'aria). Si tratta di un'equazione del secondo ordine in y, che però si può riscrivere come equazione del prim'ordine per la velocità $v = y'$:

$$v' = g - \frac{h}{m}v.$$

Questa è un'equazione a variabili separabili; $v = \frac{mg}{h}$ è una soluzione costante. Se $v \neq \frac{mg}{h}$, si ha:

$$\int \frac{dv}{g - \frac{h}{m}v} = \int dt$$

$$-\frac{m}{h}\log\left|g - \frac{h}{m}v\right| = t + c.$$

La precedente equazione si può esplicitare in v:

$$v(t) = \frac{mg}{h} + ke^{-\frac{h}{m}t}$$

con k costante arbitraria (si noti che, per $k = 0$, si ritrova la soluzione costante $v = \frac{mg}{h}$). È interessante osservare che in questo caso la velocità si stabilizza rapidamente sulla *velocità limite* $v_{\lim} = \frac{mg}{h}$, mentre in assenza dell'aria la velocità teorica $y' = gt + a$ cresce indefinitamente; si osservi che, in assenza di attrito, la velocità del corpo non dipende dalla sua massa!

2.3 Equazioni lineari del primo ordine

Sono equazioni del tipo (2.1) con F lineare in y e y', ossia che si possono scrivere nella forma: $a_1(t)y'(t) + a_0(t)y(t) = g(t)$; se il coefficiente $a_1(t)$ non si annulla, dividendo per questo si può riscrivere l'equazione in *forma normale*

$$y'(t) + a(t)\,y(t) = f(t); \tag{2.9}$$

supporremo a e f continue sull'intervallo $I \subseteq \mathbb{R}$. Se $f = 0$ l'equazione si dice *omogenea*[7]

$$z'(t) + a(t)\,z(t) = 0. \tag{2.10}$$

Per contro, la (2.9) si dirà *completa*.

Dimostriamo anzitutto il seguente teorema fondamentale (che, come vedremo, si estenderà alle equazioni lineari di ordine qualsiasi):

TEOREMA 1.2 *L'integrale generale dell'equazione completa si ottiene aggiungendo all'integrale generale dell'omogenea una soluzione particolare della completa.*

DIMOSTRAZIONE. Sia infatti $y(t)$ una qualunque soluzione della (2.9) e $\bar{y}(t)$ una soluzione particolare, cioè:

$$y' + a(t)y = f(t)$$

$$\bar{y}' + a(t)\bar{y} = f(t).$$

[7] In tutto il seguito del capitolo, nello studio delle equazioni lineari del primo e del second'ordine, useremo la lettera z per denotare la soluzione di un'equazione omogenea e y per denotare quella di un'equazione completa.

Sottraendo membro a membro si ha:

$$(y - \bar{y})' + a(t)(y - \bar{y}) = 0$$

Perciò la funzione $z(t) = y(t) - \bar{y}(t)$ è soluzione della (2.10).

Viceversa, sia $z(t)$ una qualunque soluzione della (2.10) e $\bar{y}(t)$ una soluzione particolare della (2.9); per somma si ottiene che la funzione $y(t) = z(t) + \bar{y}(t)$ è soluzione della (2.9).

Abbiamo così dimostrato che la generica soluzione dell'equazione completa si ottiene sommando alla soluzione generica dell'equazione omogenea una soluzione particolare (fissata una volta per tutte) dell'equazione completa. ∎

Soluzione dell'equazione omogenea

Occupiamoci ora dell'equazione omogenea (2.10). Sia $A(t)$ una primitiva di $a(t)$ (cioè $A'(t) = a(t)$); moltiplichiamo ambo i membri della (2.10) per $e^{A(t)}$; abbiamo

$$z'(t)e^{A(t)} + a(t)z(t)e^{A(t)} = 0$$

che si può anche scrivere

$$\left[z(t)e^{A(t)}\right]' = 0$$

e cioè

$$z(t)e^{A(t)} = c$$

da cui, essendo $A(t) = \int a(t)dt$,

$$z(t) = ce^{-\int a(t)dt}. \tag{2.11}$$

La formula precedente dice, in particolare, che l'integrale generale dell'equazione omogenea ha la forma

$$z = cz_0$$

dove z_0 è una particolare soluzione dell'omogenea e c è una costante arbitraria. Ciò significa che l'insieme delle soluzioni di un'equazione differenziale lineare omogenea del primo ordine è uno spazio vettoriale di dimensione 1. Vedremo che l'insieme delle soluzioni delle equazioni lineari omogenee del secondo ordine è uno spazio vettoriale di dimensione 2.

Ricerca di una soluzione particolare dell'equazione completa

Spesso una soluzione particolare si riesce a scoprire facilmente; altrimenti il seguente *metodo*, detto *di variazione della costante*, consente comunque di trovarne una. Questa viene cercata nella forma seguente:

$$\bar{y}(t) = c(t)e^{-A(t)}$$

dove la funzione $c(t)$ (non più costante!) deve essere trovata in modo che $\bar{y}(t)$ sia soluzione della (2.9).

© 978-88-08-06485-1

Sostituendo l'espressione di $\bar{y}$ nella (2.9) si trova:

$$e^{-A(t)}(c'(t) - c(t)a(t)) + a(t)\,c(t)\,e^{-A(t)} = f(t)$$

cioè

$$e^{-A(t)}c'(t) = f(t)$$

$$c'(t) = f(t)\,e^{A(t)}$$

$$c(t) = \int f(t)\,e^{A(t)}dt.$$

Perciò

$$\bar{y}(t) = e^{-A(t)} \int f(t)\,e^{A(t)}dt.$$

L'integrale *generale* della (2.9) è quindi dato dalla formula:

$$\boxed{y(t) = ce^{-A(t)} + e^{-A(t)} \int f(t)\,e^{A(t)}dt} \tag{2.12}$$

ottenuta aggiungendo la (2.11), dove $A(t) = \int a(t)dt$.

OSSERVAZIONE Nelle formule precedenti, $A(t)$ è una primitiva qualsiasi di $a(t)$, scelta però una volta per tutte; in altre parole, non c'è bisogno di aggiungere ad $A(t)$ la costante arbitraria di integrazione, che figura già nell'addendo $ce^{-A(t)}$. Si noti che l'integrale generale non cambia se, anziché $A(t)$, scegliessimo un'altra primitiva di $a(t)$, ad esempio $A(t) + k$. Infatti:

$$ce^{-[A(t)+k]} + e^{-[A(t)+k]} \int f(t)\,e^{A(t)+k}dt = ce^{-k}e^{-A(t)} + e^{-A(t)}e^{-k}e^{k} \int f(t)\,e^{A(t)}dt =$$

$$= c_1 e^{-A(t)} + e^{-A(t)} \int f(t)\,e^{A(t)}dt$$

con c_1 costante arbitraria.

Analogamente, nel calcolare l'integrale indefinito $\int f(t)e^{A(t)}dt$ non c'è bisogno di aggiungere la costante d'integrazione, in quanto aggiungendola si ha:

$$ce^{-A(t)} + e^{-A(t)}\left[\int f(t)\,e^{A(t)}dt + k\right] = (c+k)e^{-A(t)} + e^{-A(t)} \int f(t)\,e^{A(t)}dt$$

ottenendo la stessa formula.

Soluzione del problema di Cauchy

La costante c nella (2.12) sarà determinata da una condizione iniziale

$$y(t_0) = y_0. \tag{2.13}$$

© 978-88-08-06485-1

Scegliendo la primitiva $A(t)$ tale che $A(t_0) = 0$ (cioè $A(t) = \int_{t_0}^{t} a(s)ds$), l'integrale della (2.9) soddisfacente la (2.13) sarà:

$$\boxed{y(t) = y_0 e^{-A(t)} + e^{-A(t)} \int_{t_0}^{t} f(s)\,e^{A(s)}ds} \tag{2.14}$$

Si osservi il diverso significato dell'integrale nella (2.12) e (2.14): nella (2.12) il simbolo $\int f(t)e^{A(t)}dt$ denota una primitiva qualsiasi di $f(t)e^{A(t)}$, nella (2.14) compare invece una primitiva ben precisa, la funzione integrale con estremo fisso t_0 ed estremo variabile t; s è la "variabile muta" d'integrazione.

In particolare, abbiamo quindi dimostrato il seguente

TEOREMA 1.3 (PROBLEMA DI CAUCHY PER UN'EQUAZIONE LINEARE DEL PRIM'ORDINE) *Siano a, f funzioni continue in un intervallo $I \ni t_0$. Allora, per ogni $y_0 \in \mathbb{R}$ il problema di Cauchy*

$$\begin{cases} y'(t) + a(t)\,y(t) = f(t) \\ y(t_0) = y_0 \end{cases}$$

ha una e una sola soluzione $y \in C^1(I)$. Tale soluzione è assegnata dalla (2.14).

Si noti che per le equazioni lineari del prim'ordine il dominio di definizione della soluzione coincide con quello di continuità dei coefficienti, mentre per le equazioni a variabili separabili poteva anche essere più piccolo (si veda il teorema 1.1).



 Si voglia risolvere il problema di Cauchy:

$$\begin{cases} y' + \dfrac{2y}{x} = \dfrac{1}{x^2} \\ y(-1) = 2. \end{cases}$$

L'equazione è del tipo (2.9) con $a(x) = \frac{2}{x}$ e $f(x) = \frac{1}{x^2}$; il coefficiente a è continuo in $(-\infty, 0) \cup (0, +\infty)$; essendo il punto iniziale $x = -1$, considereremo l'intervallo $I = (-\infty, 0)$. La primitiva $A(x)$ tale che $A(-1) = 0$ è

$$A(x) = \int_{-1}^{x} \frac{2}{t}dt = [2\log|t|]_{-1}^{x} = 2\log(-x).$$

Applicando la (2.14) abbiamo:

$$y(x) = 2e^{-2\log(-x)} + e^{-2\log(-x)} \int_{-1}^{x} \frac{1}{t^2} e^{2\log(-t)}dt = \frac{2}{x^2} + \frac{1}{x^2}(x+1) = \frac{1}{x} + \frac{3}{x^2},$$

per $x < 0$.

© 978-88-08-06485-1

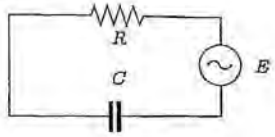


Figura 1.2.

Circuiti con resistenza e capacità. La differenza di potenziale u sulle facce opposte di un condensatore di capacità C inserito in un circuito dotato di una resistenza R ai cui capi sia applicata una forza elettromotrice costante E è descritta dall'equazione

$$CR\dot{u} + u = E. \tag{2.15}$$

La (2.15) è un'equazione lineare a coefficienti costanti; si osservi che una soluzione particolare è data dalla costante E, mentre l'integrale generale della omogenea: $CR\dot{u} + u = 0$ è dato da $ce^{-t/RC}$. Allora la soluzione della (2.15), tale che $u(0) = u_0$, è data da:

$$u(t) = E - (E - u_0)e^{-\frac{t}{\tau}} \tag{2.16}$$

avendo posto $\tau = RC$. Il termine esponenziale a secondo membro della (2.16) tende rapidamente a zero quando $t \to +\infty$ e perciò rappresenta un regime *transitorio* del fenomeno descritto dalla (2.15); il secondo termine (la costante E) rappresenta invece il regime *permanente*.

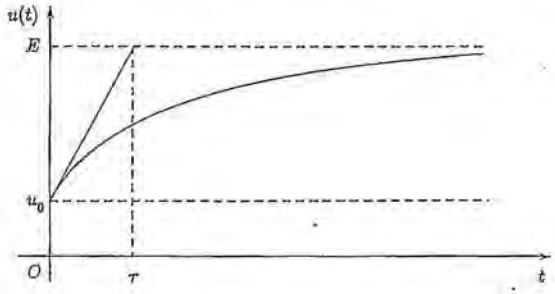


Figura 1.3.

Più in generale, sia S un sistema fisico che a una sollecitazione esterna (*entrata* o *ingresso*) $f(t)$ reagisce con una risposta (*uscita*) $y(t)$ legata a f dalla relazione:

$$\dot{y}(t) + ay(t) = f(t) \qquad a \text{ costante } > 0. \tag{2.17}$$

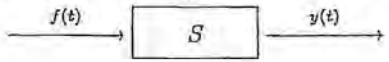


Figura 1.4.

Posto $y(0) = y_0$, l'uscita $y(t)$ del sistema è data dalla formula (2.14) (con $A(t) = at$) che si scrive:

$$y(t) = y_0e^{-at} + e^{-at}\int_0^t f(s)e^{as}ds = y_0e^{-at} + F(t),$$

dove $y_0e^{-at} \to 0$ per $t \to +\infty$, e perciò il *regime permanente* è dato da:

$$F(t) = \int_0^t f(s)e^{-a(t-s)}ds$$

cioè dalla *convoluzione*[8] $(h * f)(t)$ della sollecitazione esterna f con la funzione $h(t) = e^{-at}$ che caratterizza, attraverso la costante a, le proprietà fisiche del sistema S.

Se la sollecitazione esterna è di tipo impulsivo (un impulso di Dirac all'istante t_0: $f(t) = \delta(t - t_0)$), la risposta sarà:

$$F(t) = \begin{cases} 0 & \text{se } t < t_0 \\ e^{-a(t-t_0)} & \text{se } t \geq t_0 \end{cases}$$

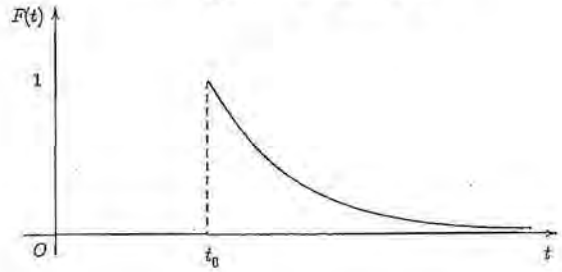


Figura 1.5.

ESERCIZI

1 Per ciascuna delle seguenti equazioni differenziali del primo ordine, dire se è a variabili separabili, se è lineare, se è entrambe le cose o nessuna delle due.

a) $x^2y' + xy = 1$ c) $y' = \frac{xy}{1+x^2}$
b) $y' = \frac{xy^2}{1+x^2}$ d) $y(y')^2 + xy' = 0$

2 Per ciascuna delle seguenti equazioni differenziali del prim'ordine, dire per quali valori di x_0, y_0 si può affermare che il problema di Cauchy con la condizione iniziale $y(x_0) = y_0$ ha certamente una e una sola soluzione (in base ai teoremi 1.1 e 1.3).

a) $y' \log x = y^2 + 1$ d) $xy' = y^2 - 1$
b) $y' = x^2\sqrt{y^2 - 1}$ e) $y' = \sqrt[3]{x}(y + 1)$
c) $y' + \frac{x^2}{x+1}y = e^x$ f) $xy' + e^xy = 2$

3 Determinare l'integrale generale della seguente equazione:

$$y' + xy - x = 0.$$

4 Determinare l'integrale generale della seguente equazione:

$$y' + y\operatorname{cotg} x - 2\cos x = 0.$$

[8] v. volume 1, capitolo 6, paragrafo 10.

© 978-88-08-06485-1

Determinare tutte le soluzioni della seguente equazione:

$$y' = \frac{2xy}{x^2 - 1}.$$

Determinare tutte le soluzioni della seguente equazione:

$$e^{x+y}y' + x = 0.$$

Risolvere il seguente problema di Cauchy:

$$\begin{cases} y' = -\dfrac{2x}{1+x^2}y + \dfrac{1}{x(1+x^2)} \\ y(-1) = 0 \end{cases}$$

e specificare qual è il più ampio intervallo su cui la soluzione è definita.

Risolvere il seguente problema di Cauchy:

$$\begin{cases} y' + 2xy^2 = 0 \\ y(0) = -1 \end{cases}$$

e specificare qual è il più ampio intervallo su cui la soluzione è definita.

L'equazione della caduta dei gravi con resistenza dell'aria:

$$v' = g - \frac{h}{m}v$$

è stata risolta (esempio 2.7) vedendola come equazione a variabili separabili; in realtà, quest'equazione è anche lineare. La si risolva ora applicando il procedimento visto per le equazioni lineari.

Un altro modello per descrivere la caduta di un grave nell'aria suppone la resistenza dell'aria proporzionale al quadrato della velocità:

$$v' = g - \frac{h}{m}v^2.$$

Anche in questo caso, la velocità si stabilizzerà rapidamente su una velocità limite. Dopo aver riconosciuto di che tipo di equazione si tratta, la si risolva e si determini la velocità limite.

Sia $T(t)$ la temperatura di un corpo ed $E =$ costante la temperatura dell'ambiente esterno. La temperatura del corpo si evolverà in base alla legge:

$$T' = k(E - T)$$

con $k > 0$ costante di conducibilità. Risolvere l'equazione differenziale sotto la condizione iniziale $T(0) = T_0$. Che cosa succede per tempi lunghi?

Consideriamo un modello di circuito elettrico con resistenza R e induttanza L costanti (positive) e una forza elettromotrice applicata, variabile nel tempo secondo la legge $V(t) = A\sin\omega t$ (A, ω costanti positive); l'intensità di corrente $I(t)$ soddisferà allora l'equazione:

$$LI' = -RI + A\sin\omega t.$$

© 978-88-08-06485-1

Risolvere l'equazione e determinare il regime permanente di $I(t)$ (cioè, nell'espressione di $I(t)$, l'addendo che non tende a zero per $t \to +\infty$).

Si consideri il problema di Cauchy:

$$\begin{cases} y' = \sqrt[3]{y} \\ y(t_0) = y_0. \end{cases}$$

a) In base al teorema di esistenza e unicità, si dica per quali valori di t_0, y_0 tale problema ammette certamente una e una sola soluzione.
b) Si determinino tutte le soluzioni dell'equazione differenziale.
c) Si determinino tutte le soluzioni del problema di Cauchy con condizione iniziale $y(0) = 0$.

Si consideri il problema di Cauchy:

$$\begin{cases} y' = t\sqrt{y} \\ y(t_0) = y_0. \end{cases}$$

a) In base al teorema di esistenza e unicità, si dica per quali valori di t_0, y_0 tale problema ammette certamente una e una sola soluzione.
b) Si determinino tutte le soluzioni dell'equazione differenziale.
c) Si mostri che il problema di Cauchy con condizione iniziale $y(1) = 0$ ha più di una soluzione, esibendo (almeno) due soluzioni distinte.

Risolvere il problema di Cauchy:

$$\begin{cases} y'(x) = \dfrac{(1+y)^2}{x} \\ y(1) = 2 \end{cases}$$

e specificare qual è il più ampio intervallo su cui la soluzione è definita.

Determinare l'integrale generale dell'equazione:

$$y' + y\sin x = \sin 2x.$$

Determinare l'integrale generale dell'equazione:

$$y' + \frac{2}{t}y = 2t + 1, \text{ per } t > 0.$$

Quindi:
a) Determinare l'unica soluzione che si mantiene limitata per $t \to 0^+$.
b) Determinare l'unica soluzione che soddisfa la condizione $y(1) = 0$.

In una reazione chimica tra due reagenti A e B si produce una certa sostanza C. Detta $x(t)$ la concentrazione della sostanza C all'istante t, se un'unità della sostanza C si forma dalla combinazione di m unità di A e n unità di B, si dimostra che $x(t)$ soddisfa l'equazione differenziale:

$$x'(t) = k(a - mx)(b - nx)$$

dove a, b sono rispettivamente il numero iniziale di unità di A, B, e k è una costante positiva opportuna. All'istante $t = 0$ (ossia prima che la reazione abbia luogo), è $x(0) = 0$. Determinare $x(t)$; calcolare il valore limite di $x(t)$, per $t \to +\infty$, distinguendo i due casi $a > b, a < b$.

© 978-88-08-06485-1

Calcolare l'istante t in cui la quantità $x(t)$ raggiunge la metà del suo valore finale. Tracciare infine il grafico di $x(t)$.

19 Si determinino le curve $y = y(x)$ tali che il segmento di tangente che unisce il punto di tangenza T al punto P di intersezione con l'asse x uguaglia il segmento che unisce il punto di tangenza all'origine (v. fig. 1.6)

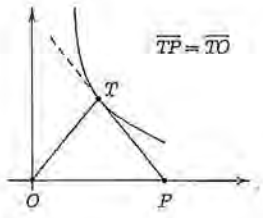


Figura 1.6.

20 Un serbatoio di superficie A ha un foro sul fondo, di superficie a. Si vuole sapere in quanto tempo il serbatoio, inizialmente pieno di un liquido, si svuota. Se $h(t)$ è l'altezza del liquido all'istante t, si dimostra che h soddisfa la seguente equazione differenziale:

$$\frac{dh}{dt} = \frac{a}{A}\sqrt{2\gamma h}$$

dove γ è il peso specifico del liquido. (vedi fig. 1.7)
a) Risolvere l'equazione differenziale.
b) Imponendo la condizione iniziale $h(0) = H$, determinare il tempo T necessario perché il serbatoio si svuoti (ossia risulti $h(T) = 0$).

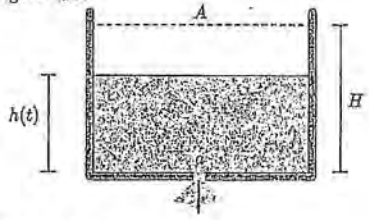


Figura 1.7.

21 Un razzo di massa M contiene carburante di massa iniziale m_0, che viene consumato a ritmo costante, in modo che la massa effettiva del razzo è una funzione del tempo:

$$m(t) = M + m_0 - kt$$

(quando $m_0 - kt = 0$ il carburante è esaurito; consideriamo quindi $t < m_0/k$). Il razzo si muove in verticale con velocità $v(t)$ secondo l'equazione differenziale:

$$m(t)\frac{dv}{dt} = -m(t)\,g - A\frac{dm}{dt}$$

dove g è l'accelerazione di gravità e A la velocità con cui il carburante è spinto fuori. Il termine $-A\frac{dm}{dt}$ rappresenta la forza propulsiva (ed è positiva, perché $\frac{dm}{dt} < 0$), mentre $-m(t)\,g$ è la forza di gravità. Calcolare la velocità in funzione del tempo, fino all'istante in cui il carburante è esaurito.

3 EQUAZIONI LINEARI DEL SECONDO ORDINE

Vogliamo ora occuparci di equazioni differenziali di ordine superiore al primo. Concentreremo la nostra attenzione su una classe particolare di equazioni, quelle lineari. Queste equazioni, che compaiono nella descrizione di molti modelli reali, sono più semplici da studiare di quelle non lineari e rappresentano, almeno talvolta, un'utile approssimazione anche di modelli non lineari. Più in particolare, in questa sezione tratteremo con un certo dettaglio le equazioni lineari del *second'ordine*. In realtà molte delle idee che illustreremo in questo contesto valgono, più in generale, per le equazioni di ordine n, come si vedrà brevemente nei Complementi. Per comprendere il significato e l'importanza dell'ipotesi di *linearità* di un'equazione differenziale, è utile premettere qualche osservazione sugli *spazi di funzioni* coinvolti.

© 978-88-08-06485-1

3.1 Spazi di funzioni

I concetti di *spazio vettoriale* e *trasformazione lineare*, introdotti nello studio dell'algebra lineare, permettono di rileggere alcuni fatti del calcolo differenziale da un punto di vista che si rivela molto utile nello studio delle equazioni differenziali lineari: il punto di vista degli *spazi di funzioni*.

Sia I un intervallo e consideriamo l'insieme $\mathcal{F}_I$ di tutte le funzioni definite in I, a valori reali. Risultano definite le operazioni naturali di somma di due funzioni e prodotto per uno scalare:

$$(f+g)(x) = f(x) + g(x)$$
$$(\lambda f)(x) = \lambda f(x).$$

Con queste operazioni, $\mathcal{F}_I$ risulta essere uno spazio vettoriale. La verifica di questo fatto è immediata.

Spesso è più interessante considerare, anziché l'insieme di tutte le funzioni possibili, l'insieme delle funzioni con qualche proprietà di regolarità. Vediamo alcuni esempi.

Si indica con $C(I)$ (o anche $C^0(I)$) l'insieme di tutte le funzioni continue in I. Ovviamente $C(I) \subset \mathcal{F}_I$. Di più, $C(I)$ è un sottospazio di $\mathcal{F}_I$, in quanto sappiamo che la combinazione lineare di funzioni continue è una funzione continua.

Si indica con $C^1(I)$ l'insieme di tutte le funzioni derivabili in I, con derivata continua[9] in I. Notiamo che $C^1(I) \subset C(I)$ (se una funzione è derivabile, allora è continua); di più, $C^1(I)$ è un sottospazio di $C(I)$:

se $f_1, f_2 \in C^1(I)$, ogni loro combinazione lineare $\lambda_1 f_1 + \lambda_2 f_2$ è derivabile; inoltre $(\lambda_1 f_1 + \lambda_2 f_2)' = \lambda_1 f_1' + \lambda_2 f_2'$ è continua (perché combinazione lineare di funzioni continue).

Per la linearità della derivata possiamo dire che l'operatore di derivazione è una trasformazione lineare tra gli spazi vettoriali $C^1(I)$ e $C(I)$.

Generalizzando queste considerazioni, definiamo $C^n(I)$ lo spazio delle funzioni dotate di derivata n-esima (e quindi di tutte le derivate di ordine inferiore), con derivata n-esima continua. Vale, ovviamente, la catena di inclusioni:

$$C(I) \supset C^1(I) \supset C^2(I) \supset \ldots \supset C^n(I) \supset \ldots$$

L'operatore di derivazione può vedersi come una trasformazione lineare di $C^{n+1}(I)$ in $C^n(I)$, per ogni $n \geq 0$.

Analogamente, l'operatore che associa a una funzione la sua derivata di ordine $k \geq 1$ può vedersi come operatore lineare da $C^{n+k}(I)$ a $C^n(I)$, per ogni $n \geq 0$.

[9] Attenzione a non equivocare: se f è derivabile, f è certamente continua; tuttavia ciò non significa necessariamente che f' sia a sua volta continua. Dunque chiedere che sia $f \in C^1$ è qualcosa di più rispetto a chiedere che sia continua e derivabile. Se $f \in C^1$, si dice anche che f è "derivabile con continuità".

© 978-88-08-06485-1

3.2 Generalità sulle equazioni lineari. Problema di Cauchy

Un'equazione differenziale del second'ordine si dice *lineare* se è del tipo:

$$a_2(t)\,y'' + a_1(t)\,y' + a_0(t)\,y = g(t) \tag{3.1}$$

dove i coefficienti a_i e il termine noto g sono funzioni definite in un certo intervallo I e ivi continue. Se il termine noto g è identicamente nullo, l'equazione si dice *omogenea*; altrimenti si dice *completa*. Se le funzioni a_i sono costanti, l'equazione si dirà *a coefficienti costanti* (notare che il termine noto può invece dipendere da t), in caso contrario si dirà *a coefficienti variabili*. Se $a_2(t) \equiv 1$ l'equazione si dirà *in forma normale*.

Il motivo per cui l'equazione è detta lineare è il seguente. Se indichiamo con $Ly(t)$ il primo membro dell'equazione (3.1), notiamo che l'operatore

$$L : C^2(I) \to C^0(I)$$

$$L : y \mapsto Ly$$

risulta proprio un operatore lineare tra questi spazi di funzioni. Infatti, nell'ipotesi di continuità dei coefficienti, se $y \in C^2(I)$ si ha che $Ly \in C^0(I)$; inoltre, se $y_1, y_2 \in C^2(I)$ e $\lambda_1, \lambda_2 \in \mathbb{R}$, si ha

$$L(\lambda_1 y_1 + \lambda_2 y_2) = \lambda_1 L y_1 + \lambda_2 L y_2.$$

La linearità di L è un fatto ricco di conseguenze, come vedremo nel prossimo paragrafo. Cominciamo ora a fare alcune considerazioni sul problema di Cauchy per queste equazioni.

Esempio

La più semplice equazione del secondo ordine è

$$y''(t) = 0.$$

Essa equivale a: $y'(t) = c_1$ e quindi: $y(t) = c_1 t + c_2$ con c_1 e c_2 costanti arbitrarie. Perciò le soluzioni di questa equazione sono tutti e soli i polinomi di primo grado.

Vedremo che l'integrale generale di una qualsiasi equazione lineare del second'ordine dipende da due parametri arbitrari. Per selezionare una soluzione particolare dovremo allora assegnare *due* condizioni; se queste sono del tipo:

$$\begin{cases} y(t_0) = y_0 \\ y'(t_0) = y_1 \end{cases} \tag{3.2}$$

il problema corrispondente (3.1)+(3.2) prende il nome di *problema di Cauchy*.

Nelle applicazioni, l'equazione (3.1) descrive l'evoluzione di un sistema fisico, il cui *stato* è individuato dalla funzione $y(t)$; il termine noto $g(t)$ rappresenta le forze esterne che agiscono sul sistema, i coefficienti $a_i(t)$ descrivono generalmente proprietà fisiche del sistema stesso.

© 978-88-08-06485-1

Esempio

Un punto materiale di massa m, libero di muoversi su una retta orizzontale, è attaccato a una molla, che esercita una forza di richiamo di tipo elastico. Detta $y(t)$ la posizione del punto sulla retta (rispetto alla posizione di riposo), y soddisfa l'equazione

$$my'' = -ky$$

con $k > 0$ costante di elasticità. L'equazione si può riscrivere dunque nella forma

$$y'' + \omega^2 y = 0$$

con $\omega^2 = k/m$, e prende il nome di *equazione dell'oscillatore armonico*. È un'equazione del secondo ordine lineare, omogenea, a coefficienti costanti.

Se sul punto agisse una forza esterna (dipendente solo dal tempo) che ne sollecita il moto, l'equazione si scriverebbe

$$y'' + \omega^2 y = f(t)$$

(dove f rappresenta la forza per unità di massa), cioè non sarebbe più omogenea.

Se il moto fosse smorzato da una forza d'attrito proporzionale alla velocità, l'equazione sarebbe

$$y'' + hy' + \omega^2 y = 0$$

con h costante positiva.

Come si vede, tutti questi esempi sono casi particolari dell'equazione (3.1).

Notiamo anche che, in questi esempi, assegnare le condizioni (3.2) significa assegnare la posizione e la velocità del punto materiale in un istante iniziale t_0. Per questo motivo ci si riferisce spesso alle (3.2) come alle "condizioni iniziali".

Se il coefficiente $a_2(t)$ nella (3.1) non si annulla mai, dividendo per questo si può riscrivere l'equazione in forma normale:

$$y'' + a(t)\,y' + b(t)\,y = f(t). \tag{3.3}$$

Vale il seguente risultato generale sul problema di Cauchy per le equazioni lineari del second'ordine in forma normale (che è un caso particolare del Teorema di Cauchy di esistenza e unicità locale, capitolo 8, paragrafo 2):

TEOREMA 1.4 *Se a, b, f sono funzioni continue in un intervallo $I \ni t_0$, per ogni $y_0, y_1 \in \mathbb{R}$ il problema di Cauchy*

$$\begin{cases} y'' + a(t)y' + b(t)y = f(t) \\ y(t_0) = y_0 \\ y'(t_0) = y_1 \end{cases}$$

ha una e una sola soluzione $y \in C^2(I)$.

Si noti l'analogia con il caso delle equazioni lineari del prim'ordine. Come al solito, tale soluzione sarà individuata imponendo le condizioni iniziali nell'espressione che assegna l'integrale generale dell'equazione (3.3). Il problema è quindi capire come si scrive tale integrale generale.

© 978-88-08-06485-1

3.3 La struttura dell'integrale generale

Abbiamo visto che un'equazione differenziale lineare del second'ordine si può scrivere nella forma

$$Ly = f$$

dove $L : C^2(I) \to C^0(I)$ è un operatore lineare tra i due spazi di funzioni. L'equazione $Ly = 0$ si dice equazione omogenea associata all'equazione completa $Ly = f$. La linearità di L permette di determinare facilmente la struttura dell'integrale generale dell'equazione, come afferma il seguente:

TEOREMA 1.5 (STRUTTURA DELL'INTEGRALE GENERALE DELL'EQUAZIONE LINEARE COMPLETA)

a. L'insieme delle soluzioni dell'equazione omogenea $Lz = 0$ in un dato intervallo I è uno spazio vettoriale (sottospazio di $C^2(I)$).

b. L'integrale generale dell'equazione completa si ottiene sommando l'integrale generale dell'equazione omogenea e una soluzione particolare dell'equazione completa.

DIMOSTRAZIONE.

a. Siano $z_1, z_2 \in C^2(I)$ soluzioni dell'equazione omogenea in I, e siano λ_1, λ_2 costanti. Allora:

$$L(\lambda_1 z_1 + \lambda_2 z_2) = \lambda_1 L z_1 + \lambda_2 L z_2 = 0$$

(il primo passaggio segue dalla linearità di L, il secondo dal fatto che per ipotesi $Lz_1 = 0$ e $Lz_2 = 0$). Perciò anche $\lambda_1 z_1 + \lambda_2 z_2$ risolve l'equazione omogenea, dunque l'insieme delle soluzioni dell'equazione omogenea è uno spazio vettoriale, in particolare è un sottospazio di $C^2(I)$.

b. Siano y_1 una soluzione particolare dell'equazione completa e z_0 una generica soluzione dell'equazione omogenea, ossia: $Ly_1 = f, Lz_0 = 0$. Allora per linearità

$$L(y_1 + z_0) = Ly_1 + Lz_0 = f + 0 = f$$

ossia $y_1 + z_0$ è soluzione dell'equazione completa.

Viceversa, se y_2 ora è una qualsiasi soluzione dell'equazione completa ($Ly_2 = f$) per linearità è:

$$L(y_2 - y_1) = Ly_2 - Ly_1 = f - f = 0$$

ossia $y_2 - y_1$ è soluzione dell'omogenea, ossia $y_2 - y_1 = z_0$ per una certa soluzione z_0 dell'omogenea. Dunque

$$y_2 = y_1 + z_0$$

ossia: la generica soluzione dell'equazione completa si può scrivere come somma di una particolare soluzione y_1 dell'equazione completa (fissata una volta per tutte) e di una soluzione dell'equazione omogenea. Questo prova *b*. ♦

Osservando la dimostrazione del teorema precedente, il lettore noti che questa non dipende in alcun modo dal fatto che l'equazione $Ly = f$ sia un'equazione differenziale (piuttosto che un'equazione algebrica, o di altro tipo), né dall'ordine 2 dell'equazione, ma solo dal fatto che sia un'equazione *lineare*. Questo risultato, infatti, si estenderà alle equazioni lineari di ordine n (si vedano i Complementi in fondo al capitolo).

Nel seguito del discorso supporremo sempre che l'equazione lineare del second'ordine sia scritta in forma normale, cioè (3.3). L'equazione omogenea associata è allora:

$$Lz \equiv z'' + a(t)z' + b(t)z = 0. \tag{3.4}$$

© 978-88-08-06485-1

Un'ulteriore importante proprietà, che riguarda l'equazione omogenea (3.4), è espressa dal seguente:

TEOREMA 1.6 *Lo spazio vettoriale delle soluzioni di un'equazione differenziale lineare omogenea del second'ordine ha dimensione* 2.

Esplicitamente, questo significa che esistono 2 soluzioni dell'equazione omogenea in un intervallo I, chiamiamole $z_1(t), z_2(t)$, tali che:

1. queste funzioni sono linearmente indipendenti, ossia non sono una multipla dell'altra (o anche: il quoziente $z_2(t)/z_1(t)$ non è costante nell'intervallo I);

2. ogni altra soluzione dell'equazione omogenea è combinazione lineare di $z_1(t)$, $z_2(t)$, il che significa che l'integrale generale dell'equazione omogenea è assegnato dalla formula

$$c_1 z_1(t) + c_2 z_2(t)$$

al variare in ogni modo dei coefficienti reali c_1, c_2.

DIMOSTRAZIONE. La dimostrazione si basa totalmente sul Teorema 1.4.

Siano z_1, z_2, rispettivamente, le soluzioni dei problemi di Cauchy nell'intervallo I, con $t_0 \in I$ fissato:

$$\begin{cases} Lz_1 = 0 \\ z_1(t_0) = 1 \\ z_1'(t_0) = 0 \end{cases} \qquad \begin{cases} Lz_2 = 0 \\ z_2(t_0) = 0 \\ z_2'(t_0) = 1. \end{cases}$$

Proviamo che:

1) Le funzioni z_1, z_2, soluzioni dell'equazione omogenea (3.4), sono linearmente indipendenti. Infatti il quoziente $z_2(t)/z_1(t)$ si annulla in $t = t_0$; se fosse costante in I, dovrebbe essere identicamente nullo; ma allora $z_2(t)$ sarebbe identicamente nulla, quindi $z_2'(t_0) = 0$, contro l'ipotesi $z_2'(t_0) = 1$.

2) Ogni altra soluzione dell'equazione omogenea è combinazione lineare di z_1, z_2 in I. Sia $z_0(t)$ una qualsiasi soluzione di $Lz = 0$ in I, e cerchiamo due costanti c_1, c_2 per cui sia

$$z_0(t) = c_1 z_1(t) + c_2 z_2(t) \text{ per ogni } t \in I.$$

Scegliendo $c_1 = z_0(t_0), c_2 = z_0'(t_0)$, si ottiene:

$$(c_1 z_1 + c_2 z_2)(t_0) = c_1 \cdot 1 + c_2 \cdot 0 = c_1 = z_0(t_0);$$
$$(c_1 z_1 + c_2 z_2)'(t_0) = c_1 \cdot 0 + c_2 \cdot 1 = c_2 = z_0'(t_0).$$

Pertanto la funzione $c_1 z_1(t) + c_2 z_2(t)$ è soluzione del problema di Cauchy

$$\begin{cases} Ly = 0 \\ y(t_0) = z_0(t_0) \\ y'(t_0) = z_0'(t_0). \end{cases}$$

Poiché, d'altro canto, anche $z_0(t)$ risolve, ovviamente, il medesimo problema, per l'unicità della soluzione del problema di Cauchy si ha che:

$$z_0(t) = (c_1 z_1 + c_2 z_2)(t) \text{ per ogni } t \in I,$$

che è quanto volevamo dimostrare. ♦

Esempio

L'equazione

$$t^2 z'' - 3t z' + 3z = 0$$

possiede (lo studente verifichi l'affermazione!) le due soluzioni particolari t e t^3 le quali sono indipendenti.

Allora l'integrale generale è dato da

$$z(t) = c_1 t + c_2 t^3.$$

Quando due soluzioni di un'equazione omogenea sono note esplicitamente, è facile verificare se queste siano linearmente indipendenti, semplicemente controllando che non sono una multipla dell'altra. Tuttavia, nello sviluppo della teoria, è utile avere un criterio generale per decidere dell'indipendenza o meno di due soluzioni, e per sfruttare tale informazione. Vale in proposito il seguente:

TEOREMA 1.7 (DETERMINANTE WRONSKIANO E INDIPENDENZA) *Siano z_1, z_2 due funzioni $C^2(I)$, soluzioni dell'equazione lineare omogenea (3.4) nell'intervallo I. Allora esse sono linearmente indipendenti in $C^2(I)$ se e solo se la seguente matrice*

$$\begin{bmatrix} z_1(t) & z_2(t) \\ z_1'(t) & z_2'(t) \end{bmatrix},$$

detta matrice Wronskiana, *ha determinante diverso da zero per ogni $t \in I$. Inoltre, affinché questo accada è sufficiente che il determinante sia diverso da zero in un punto $t_0 \in I$.*

L'ultima parte del teorema si può esprimere anche così: il determinante della matrice Wronskiana (detto brevemente *Wronskiano*) o si annulla in tutti i punti di I, oppure è diverso da zero in tutti i punti di I. Nel secondo caso, z_1 e z_2 sono indipendenti.

Si noti che questo teorema non caratterizza l'indipendenza di due qualsiasi funzioni in $C^2(I)$, ma quella di due soluzioni dell'equazione (3.4). Come vedremo dalla dimostrazione, anche questo teorema si basa essenzialmente sul risultato di esistenza e unicità per il problema di Cauchy (Teorema 1.4).

DIMOSTRAZIONE. Siano z_1, z_2 due funzioni linearmente *dipendenti* in $C^2(I)$; quindi esistono due costanti $(c_1, c_2) \neq (0,0)$ tali che

$$c_1 z_1(t) + c_2 z_2(t) = 0 \text{ per ogni } t \in I. \tag{3.5}$$

Derivando rispetto a t quest'identità si trova

$$c_1 z_1'(t) + c_2 z_2'(t) = 0 \text{ per ogni } t \in I$$

e quindi i numeri (c_1, c_2) sono una soluzione non banale del sistema lineare omogeneo

$$\begin{cases} c_1 z_1(t) + c_2 z_2(t) = 0 \\ c_1 z_1'(t) + c_2 z_2'(t) = 0, \end{cases} \tag{3.6}$$

che pertanto deve avere determinante nullo, per ogni $t \in I$. Ma il determinante di questo sistema è proprio il Wronskiano.

Ne deduciamo:

$$z_1, z_2 \text{ linearmente dipendenti} \Longrightarrow \text{Wronskiano identicamente nullo}$$

e perciò anche

$$\text{Wronskiano diverso da zero in un punto } t_0 \Longrightarrow z_1, z_2 \text{ linearmente indipendenti.}$$

Viceversa, supponiamo che z_1, z_2 siano soluzioni linearmente indipendenti di (3.4), e mostriamo che il Wronskiano non si annulla mai. Supponiamo, per assurdo, che si annulli in un punto t_0. Allora esiste una soluzione (c_1, c_2) non banale del sistema:

$$\begin{cases} c_1 z_1(t_0) + c_2 z_2(t_0) = 0 \\ c_1 z_1'(t_0) + c_2 z_2'(t_0) = 0, \end{cases}$$

Sia z la soluzione del problema di Cauchy:

$$\begin{cases} z'' + a(t) z' + b(t) z = 0 \\ z(t_0) = c_1 z_1(t_0) + c_2 z_2(t_0) = 0 \\ z'(t_0) = c_1 z_1'(t_0) + c_2 z_2'(t_0) = 0. \end{cases}$$

Sappiamo che la funzione identicamente nulla soddisfa queste condizioni; d'altro canto le soddisfa anche la funzione $c_1 z_1(t) + c_2 z_2(t)$; per l'unicità della soluzione del problema di Cauchy, ne segue che $c_1 z_1(t) + c_2 z_2(t)$ è identicamente nulla in I, ossia z_1, z_2 sono dipendenti, in contrasto con l'ipotesi di z_1, z_2 linearmente indipendenti. ♦

Ricapitoliamo ora il percorso di questo paragrafo. In base ai teoremi 1.5 e 1.6, il problema della determinazione dell'integrale generale di un'equazione differenziale lineare completa del second'ordine si riconduce ai due passi seguenti:

1) determinare l'integrale generale dell'equazione omogenea; questo a sua volta significa determinare 2 soluzioni $z_1(t), z_2(t)$ dell'equazione omogenea, linearmente indipendenti;

2) determinare una soluzione particolare $\overline{y}(t)$ dell'equazione completa.

A questo punto l'integrale generale dell'equazione completa sarà dato da:

$$\overline{y}(t) + c_1 z_1(t) + c_2 z_2(t)$$

al variare in ogni modo dei coefficienti reali c_i.

Diciamo subito che non esiste alcun metodo generale per risolvere effettivamente questi due problemi, quando i coefficienti a, b, f siano funzioni continue qualsiasi. Vedremo invece come risolvere facilmente il problema 1) quando a, b sono costanti (vedi paragrafo 3.4), e come risolvere il problema 2) quando f è una funzione di forma particolare, abbastanza semplice ("metodo di somiglianza"), oppure quando si è già risolto il problema 1) ("metodo di variazione delle costanti") (vedi paragrafo 3.5).

Infine, per risolvere un problema di Cauchy, una volta determinato l'integrale generale nella forma precedente, basterà imporre le 2 condizioni iniziali, ricavare di conseguenza il valore delle costanti c_1, c_2 e scrivere la soluzione corrispondente a questi valori di c_1, c_2.

 © 978-88-08-06485-1

3.4 Equazioni omogenee a coefficienti costanti

Consideriamo l'equazione omogenea

$$z''(t) + az'(t) + bz(t) = 0 \qquad (a, b \text{ costanti}) \tag{3.7}$$

e cerchiamo soluzioni di tipo esponenziale: $t \mapsto e^{rt}$, $r \in \mathbb{C}$. Ciò è suggerito dal fatto che, nel caso analogo del primo ordine, gli integrali sono proprio esponenziali.

Sostituendo $z(t) = e^{rt}$ nella (3.7), abbiamo:

$$e^{rt}(r^2 + ar + b) = 0.$$

Perciò, affinché l'esponenziale e^{rt} sia soluzione della (3.7), la costante r deve essere una radice dell'equazione

$$r^2 + ar + b = 0, \tag{3.8}$$

detta *equazione caratteristica* della (3.7).

Distinguiamo tre casi, secondo il valore di $\Delta = a^2 - 4b$, discriminante dell'equazione caratteristica.

1. $a^2 > 4b$, cioè $\Delta > 0$: la (3.8) possiede due radici reali e distinte r_1 e r_2. Le funzioni $z_1(t) = e^{r_1 t}$ e $z_2(t) = e^{r_2 t}$ sono due soluzioni distinte e indipendenti della (3.7) il cui integrale generale si scrive dunque:

$$z(t) = c_1 e^{r_1 t} + c_2 e^{r_2 t}. \tag{3.9}$$

2. $a^2 < 4b$, cioè $\Delta < 0$: la (3.8) ha due radici complesse coniugate $r_1 = \alpha + i\beta$, $r_2 = \alpha - i\beta$ (α, β reali). Soluzioni (indipendenti) della (3.7) sono perciò le funzioni:

$$z_1(t) = e^{(\alpha+i\beta)t} = e^{\alpha t}(\cos\beta t + i \sin\beta t)$$
$$z_2(t) = e^{(\alpha-i\beta)t} = e^{\alpha t}(\cos\beta t - i \sin\beta t).$$

Può essere desiderabile avere soluzioni reali; ricordando che ogni combinazione lineare di soluzioni dell'equazione omogenea è ancora soluzione della stessa equazione, scegliamo, in luogo di z_1 e z_2, le soluzioni $\frac{1}{2}(z_1 + z_2)$ e $\frac{1}{2i}(z_1 - z_2)$, cioè

$$e^{\alpha t}\cos\beta t \qquad e^{\alpha t}\sin\beta t.$$

L'integrale generale della (3.7) si può scrivere perciò nella forma:

$$z(t) = e^{\alpha t}(c_1 \cos\beta t + c_2 \sin\beta t). \tag{3.10}$$

Un altro modo, spesso utile, di scrivere la (3.10) è il seguente:

$$z(t) = e^{\alpha t} A \cos(\beta t + \varphi) \tag{3.11}$$

con A, φ costanti reali arbitrarie (cfr. volume 1, capitolo 2, paragrafo 3.4).

© 978-88-08-06485-1

3. $a^2 = 4b$, cioè $\Delta = 0$: la (3.7) possiede l'unica radice (doppia) $r = -a/2$. Perciò $e^{-\frac{a}{2}t}$ è una soluzione; per trovare una seconda soluzione usiamo ancora il metodo della variazione delle costanti (paragrafo 2.3); cerchiamola nella forma:

$$z(t) = e^{rt}c(t) \quad \text{con } r = -a/2$$
$$z'(t) = e^{rt}(rc(t) + c'(t))$$
$$z''(t) = e^{rt}(r^2 c(t) + 2rc'(t) + c''(t)).$$

Sostituendo nella (3.7), abbiamo:

$$e^{rt}[(r^2 + ar + b)c(t) + (2r + a)c'(t) + c''(t)] = 0$$

Poiché $r = -a/2$, risulta verificata l'equazione caratteristica $r^2 + ar + b = 0$ e anche $2r + a = 0$; perciò dovrà essere $c''(t) = 0$, cioè $c(t) = c_2 t + c_1$. La soluzione generale della (3.7) sarà allora:

$$z(t) = e^{-\frac{a}{2}t}(c_1 + c_2 t) \tag{3.12}$$

In ciascuno dei tre casi ora studiati abbiamo determinato due soluzioni dell'equazione omogenea, che risultano linearmente indipendenti (come si verifica immediatamente):

$$\begin{array}{llll} e^{r_1 t} & e^{r_2 t} & (\text{caso } \Delta > 0) \\ e^{\alpha t}\cos\beta t & e^{\alpha t}\sin\beta t & (\text{caso } \Delta < 0) \\ e^{rt} & te^{rt} & (\text{caso } \Delta = 0). \end{array} \tag{3.13}$$

A titolo di esercizio, lo studente è invitato a verificare l'indipendenza di ciascuna delle tre coppie di soluzioni trovate applicando il criterio del determinante Wronskiano (Teorema 1.7).

Esempi

3.24 Trovare l'integrale generale delle seguenti equazioni:

i) $z'' - \omega^2 z = 0$ ii) $z'' + \omega^2 z = 0$ iii) $z'' - 2\omega z' + \omega^2 z = 0$

L'equazione caratteristica è, rispettivamente:

i) $r^2 - \omega^2 = 0$ ii) $r^2 + \omega^2 = 0$ iii) $r^2 - 2\omega r + \omega^2 = 0$

e perciò abbiamo le soluzioni:

i) $z(t) = c_1 e^{\omega t} + c_2 e^{-\omega t}$ o anche $c_3 \mathrm{Sh}\omega t + c_4 \mathrm{Ch}\omega t$

avendo utilizzato, invece delle soluzioni $e^{\omega t}$, $e^{-\omega t}$, le combinazioni $\frac{1}{2}(e^{\omega t} \pm e^{-\omega t})$.

ii) $z(t) = c_1 \sin\omega t + c_2 \cos\omega t$ o anche $A\cos(\omega t - \varphi)$

iii) $z(t) = e^{\omega t}(c_1 + c_2 t)$

3.25 Risolvere il problema di Cauchy:

$$\begin{cases} z'' + 2z' + 3z = 0 \\ z(0) = 1 \\ z'(0) = 2 \end{cases}$$

© 978-88-08-06485-1

L'equazione caratteristica è:

$$r^2 + 2r + 3 = 0$$

che ha soluzioni:

$$r = -1 \pm i\sqrt{2}.$$

Due soluzioni indipendenti dell'equazione sono dunque:

$$z_1 = e^{-x}\cos(\sqrt{2}x) \qquad z_2 = e^{-x}\sin(\sqrt{2}x)$$

e l'integrale generale è:

$$z = e^{-x}(c_1\cos(\sqrt{2}x) + c_2\sin(\sqrt{2}x)).$$

Imponiamo le condizioni iniziali.

$$\begin{aligned} z(0) &= c_1 = 1 \\ z'(x) &= e^{-x}(-c_1\cos(\sqrt{2}x) - c_2\sin(\sqrt{2}x) - \sqrt{2}c_1\sin(\sqrt{2}x) + \sqrt{2}c_2\cos(\sqrt{2}x)) \\ z'(0) &= -c_1 + \sqrt{2}c_2 = 2 \end{aligned}$$

Dunque

$$c_1 = 1 \qquad c_2 = \frac{3}{\sqrt{2}}$$

e la soluzione del problema di Cauchy è:

$$z = e^{-x}(\cos(\sqrt{2}x) + \frac{3}{\sqrt{2}}\sin(\sqrt{2}x)).$$

3.5 Equazioni non omogenee

Ci occupiamo ora del problema di determinare un integrale particolare dell'equazione completa, associata a un'equazione lineare a coefficienti costanti:

$$y''(t) + ay'(t) + by(t) = f(t) \qquad (a, b \text{ costanti}) \tag{3.14}$$

Illustriamo prima il *metodo di somiglianza*, che per una certa classe di casi particolari permette di determinare tale integrale particolare con brevi calcoli; vedremo poi il *metodo di variazione delle costanti*, che può essere impiegato *sempre*, ma in compenso porta solitamente a calcoli più pesanti (perciò, quando possibile, è meglio applicare il primo metodo).

Metodo di somiglianza

Quando il termine noto $f(t)$ ha un aspetto particolarmente semplice, si può cercare una soluzione pure abbastanza semplice, e *simile* a f, in un senso che ora preciseremo.

- $f(t) = p_r(t)$ (polinomio di grado r). Si cerca una soluzione di tipo polinomiale:

$$\begin{aligned} &y(t) = q_r(t) && \text{se} \quad b \neq 0 \\ &y(t) = tq_r(t) && \text{se} \quad b = 0 \quad \text{e} \quad a \neq 0 \\ &y(t) = t^2 q_r(t) && \text{se} \quad b = 0 \quad \text{e} \quad a = 0. \end{aligned}$$

dove $q_r(t)$ è un generico polinomio di grado r, di cui occorre determinare i coefficienti.

© 978-88-08-06485-1

Esempio

Trovare una soluzione particolare delle equazioni:

$$\text{i)} \quad y'' - \omega^2 y = 1 + t^2 \qquad \text{ii)} \quad y'' + 2y' = t.$$

i) Cerchiamo una soluzione polinomiale di secondo grado:

$$y(t) = c_0 + c_1 t + c_2 t^2.$$

Inserendo l'espressione nell'equazione abbiamo:

$$2c_2 - \omega^2 c_0 - \omega^2 c_1 t - \omega^2 c_2 t^2 = 1 + t^2$$

da cui, uguagliando i coefficienti dei termini di pari grado in t, ricaviamo:

$$c_0 = -\frac{2+\omega^2}{\omega^4} \qquad c_1 = 0 \qquad c_2 = -\frac{1}{\omega^2},$$

e quindi

$$y = -\left(\frac{2+\omega^2}{\omega^4} + \frac{1}{\omega^2}t^2\right).$$

ii) Cerchiamo ancora una soluzione polinomiale di secondo grado (il termine noto è di primo grado, ma nell'equazione manca il termine in y):

$$y(t) = c_1 t + c_2 t^2.$$

Abbiamo: $2c_2 + 2c_1 + 4c_2 t = t$ da cui ricaviamo:

$$c_1 = -\frac{1}{4} \qquad c_2 = \frac{1}{4},$$

e quindi

$$y = \frac{1}{4}(t^2 - 1).$$

- $f(t) = Ae^{\lambda t}$, $\lambda \in \mathbb{C}$. Si cerca una soluzione del tipo $y(t) = e^{\lambda t}\gamma(t)$. Si trova (con calcoli simili a quelli svolti per l'equazione omogenea):

$$\gamma'' + \gamma'(2\lambda + a) + \gamma(\lambda^2 + a\lambda + b) = A. \tag{3.15}$$

È sufficiente trovare una *qualsiasi* $\gamma(t)$ che soddisfi la (3.15).

- Se $\lambda^2 + a\lambda + b \neq 0$ (cioè se λ *non* è radice dell'equazione caratteristica), basterà prendere:

$$\gamma(t) = \text{ costante } = \frac{A}{\lambda^2 + a\lambda + b}, \text{ e dunque } y(t) = \frac{Ae^{\lambda t}}{\lambda^2 + a\lambda + b}.$$

- Se $\lambda^2 + a\lambda + b = 0$, ma $2\lambda + a \neq 0$, basterà prendere

$$\gamma'(t) = \text{ costante } = \frac{A}{2\lambda + a}, \text{ da cui } \gamma(t) = \frac{At}{2\lambda + a} \text{ e } y(t) = \frac{Ate^{\lambda t}}{2\lambda + a}.$$

© 978-88-08-06485-1

– Se infine $\lambda^2 + a\lambda + b = 0$ e $2\lambda + a = 0$, la (3.15) si scrive semplicemente $\gamma'' = A$, da cui

$$\gamma(t) = \frac{A}{2}t^2, \quad \text{e } y(t) = \frac{A}{2}t^2 e^{\lambda t}.$$

Nella classe di termini noti del tipo $e^{\lambda t}$ con $\lambda \in \mathbb{C}$ rientrano anche i casi:

$$\cos \omega t, \quad \sin \omega t, \quad e^{\lambda t}\cos \omega t, \quad e^{\lambda t}\sin \omega t \quad (\text{con } \lambda \in \mathbb{R}).$$

Mostriamo attraverso esempi come si opera in queste varie situazioni.

Esempi

Troviamo una soluzione particolare delle seguenti equazioni:

1)

$$y'' + 2y' + 3y = 2e^{3x}.$$

Cerchiamo:

$$y = ce^{3x}$$
$$y'' + 2y' + 3y = ce^{3x}(9 + 2 \cdot 3 + 3) = 18ce^{3x} = 2e^{3x}$$

per $c = \frac{1}{9}$. Dunque:

$$y = \frac{1}{9}e^{3x}.$$

2)

$$y'' + 2y' - 3y = 2e^{-3x}.$$

Se cerchiamo

$$y = ce^{-3x}$$

abbiamo:

$$y'' + 2y' - 3y = ce^{-3x}(9 - 2 \cdot 3 - 3) = 0$$

e per ogni c la funzione non risolve l'equazione completa. Il punto è che l'esponente -3 annulla l'equazione caratteristica. Cerchiamo allora:

$$y = cxe^{-3x}$$

e abbiamo:

$$y' = ce^{-3x}(1 - 3x); y'' = ce^{-3x}(-6 + 9x)$$
$$y'' + 2y' - 3y = ce^{-3x}((-6 + 9x) + 2(1 - 3x) - 3x) = -4ce^{-3x} = 2e^{-3x}$$

per $c = -1/2$, dunque una soluzione è:

$$y = -\frac{x}{2}e^{-3x}.$$

Se anche la soluzione cxe^{-3x} avesse annullato identicamente il primo membro dell'equazione, avremmo dovuto cercare una soluzione del tipo cx^2e^{-3x}.

3)

$$y'' + 2y' - y = 2e^{x}\cos 3x.$$

Notiamo che:

$$2e^{x(1+3i)} = 2e^{x}\cos 3x + i2e^{x}\sin 3x.$$

© 978-88-08-06485-1

Se risolviamo allora l'equazione con termine noto $2e^{x(1+3i)}$ e poi prendiamo la parte reale di tale soluzione complessa, per linearità troveremo la soluzione dell'equazione con termine noto $2e^{x}\cos 3x$.

(Analogamente, per risolvere l'equazione con termine noto $2e^{x}\sin 3x$ dovremmo prendere la parte immaginaria della soluzione complessa dell'equazione con termine noto $2e^{x(1+3i)}$).

Cerchiamo dunque:

$$y = ce^{x(1+3i)}$$
$$y'' + 2y' - y = ce^{x(1+3i)}((1+3i)^2 + 2(1+3i) - 1) = 2e^{x(1+3i)}$$

purché

$$c((1+3i)^2 + 2(1+3i) - 1) = 2$$

ossia

$$c = \frac{2}{1 - 9 + 6i + 2 + 6i - 1} = \frac{2}{-7 + 12i} = \frac{2}{49 + 144}(-7 - 12i).$$

La soluzione complessa è:

$$y = \frac{2e^{x}}{193}(-7 - 12i)(\cos 3x + i \sin 3x).$$

La soluzione che cerchiamo è la parte reale di questa funzione, ossia:

$$y = \frac{2e^{x}}{193}(-7\cos 3x + 12 \sin 3x).$$

Per ottenere lo stesso risultato senza usare i numeri complessi, avremmo dovuto cercare una soluzione del tipo:

$$y = e^{x}(c_1 \cos 3x + c_2 \sin 3x).$$

Si tenga presente comunque che l'utilizzo dell'esponenziale complesso rende *meno laborioso il calcolo delle derivate* di y.

4)

$$2y'' + y' + 2y = 3 \sin 2x.$$

Si può ragionare come nel caso precedente, risolvendo prima l'equazione che ha termine noto complesso

$$3e^{2ix} = 3\cos 2x + 3i \sin 2x$$

e poi prendendo la parte immaginaria della soluzione complessa trovata. Cerchiamo dunque:

$$y = ce^{2ix}$$
$$2y'' + y' + 2y = ce^{2ix}(2(2i)^2 + 2i + 2) = 3e^{2ix}$$

per

$$c = \frac{3}{-6 + 2i} = \frac{3}{40}(-6 - 2i).$$

La soluzione complessa è quindi:

$$y = \frac{3}{40}(-6 - 2i)(\cos 2x + i \sin 2x)$$

e la soluzione che cerchiamo è la parte immaginaria di questa:

$$y = \frac{3}{40}(-2\cos 2x - 6 \sin 2x) = -\frac{3}{20}\cos 2x - \frac{9}{20}\sin 2x.$$

Per ottenere lo stesso risultato senza usare i numeri complessi, avremmo dovuto cercare una soluzione del tipo:

$$y = c_1 \cos 2x + c_2 \sin 2x.$$

Si osservi che, anche se il termine noto contiene solo la funzione $\sin 2x$, in generale la soluzione va cercata come combinazione lineare di funzioni seno e coseno.

Fa eccezione a questa regola solo il caso in cui nell'equazione differenziale manchi il termine in y', come mostra il prossimo esempio.

Metodo di sovrapposizione.

$$y'' + 3y = x + 2\cos x.$$

Questa è un'equazione il cui termine noto è somma di un polinomio e una funzione trigonometrica. Per linearità, se si trova una soluzione y_1 di

$$y'' + 3y = x$$

una soluzione y_2 di

$$y'' + 3y = 2\cos x$$

e si sommano le due funzioni, si trova una soluzione dell'equazione di partenza. Per i metodi visti in precedenza, cercheremo y_1 nella forma

$$y_1 = ax + b$$

(anzi, si vede subito che $y_1 = x/3$ è soluzione) e y_2 nella forma

$$y_2 = c\cos x.$$

Si trova

$$y_2 = \cos x$$

e dunque una soluzione dell'equazione di partenza è

$$y = \frac{x}{3} + \cos x.$$

Metodo di variazione delle costanti

Illustriamo ora un metodo generale che consente di determinare una soluzione particolare dell'equazione completa, *qualunque sia la forma del termine noto* $f(t)$, quindi anche quando questo non rientra nei casi speciali trattati nel paragrafo precedente. Il metodo è applicabile *purché si conoscano già due soluzioni indipendenti dell'equazione omogenea.* Se l'equazione omogenea è a coefficienti costanti, abbiamo visto un metodo standard per ottenere tali soluzioni. Segnaliamo, comunque, che il metodo di variazione delle costanti è applicabile anche alle equazioni lineari a coefficienti variabili. Il punto è che, in questo caso, non vi sono invece *metodi generali* per determinare due soluzioni indipendenti dell'equazione omogenea, come già osservato.

Siano dunque $z_1(t), z_2(t)$ due soluzioni indipendenti dell'equazione omogenea associata in un intervallo I. L'idea è cercare una soluzione particolare nella forma:

$$\overline{y}(t) = c_1(t)z_1(t) + c_2(t)z_2(t) \tag{3.16}$$

(si confronti col procedimento seguito per risolvere le equazioni lineari del prim'ordine non omogenee).

Ricordiamo che le funzioni z_1, z_2 sono note, mentre le funzioni c_1, c_2 sono incognite, e vanno determinate in modo tale che $\overline{y}$ soddisfi l'equazione

$$y'' + ay' + by = f \text{ in } I. \tag{3.17}$$

Quest'equazione fornirà una condizione sulle due funzioni c_1, c_2. Poiché le funzioni da determinare sono due, potremo imporre una seconda condizione su c_1, c_2, che sceglieremo in modo opportuno. Cominciamo a calcolare, dalla (3.16),

$$\overline{y}' = c_1'z_1 + c_2'z_2 + c_1z_1' + c_2z_2'.$$

Scegliamo di imporre come condizione su c_1, c_2 la seguente:

$$c_1'z_1 + c_2'z_2 = 0.$$

Questa fa sì che risulti

$$\overline{y}' = c_1z_1' + c_2z_2' \tag{3.18}$$

e di conseguenza

$$\overline{y}'' = c_1'z_1' + c_2'z_2' + c_1z_1'' + c_2z_2''. \tag{3.19}$$

Sostituendo (3.28) e (3.29) nella (3.17) otteniamo:

$$(c_1'z_1' + c_2'z_2' + c_1z_1'' + c_2z_2'') + a(c_1z_1' + c_2z_2') + b(c_1z_1 + c_2z_2) = f$$

che si può riscrivere come:

$$c_1(z_1'' + az_1' + bz_1) + c_2(z_2'' + az_2' + bz_2) + (c_1'z_1' + c_2'z_2') = f$$

che a sua volta, ricordando che per ipotesi z_1, z_2 soddisfano l'equazione omogenea, si riduce a:

$$c_1'z_1' + c_2'z_2' = f$$

In definitiva, siamo arrivati a scrivere il sistema lineare di due equazioni nelle due funzioni incognite c_1', c_2':

$$\begin{cases} c_1'z_1 + c_2'z_2 = 0 \\ c_1'z_1' + c_2'z_2' = f. \end{cases}$$

Si osserva ora che la matrice dei coefficienti di questo sistema è esattamente la matrice Wronskiana corrispondente a z_1, z_2; essendo queste soluzioni linearmente indipendenti per ipotesi, il determinante Wronskiano

$$\begin{vmatrix} z_1 & z_2 \\ z_1' & z_2' \end{vmatrix}$$

è diverso da zero per ogni $t \in I$ (si veda il Teorema 1.7). Perciò il sistema può essere risolto in c_1', c_2'. Si trova:

$$c_1' = \frac{-z_2 f}{z_2'z_1 - z_2z_1'} \qquad c_2' = \frac{z_1 f}{z_2'z_1 - z_2z_1'}.$$

(Si noti che il denominatore è proprio il determinante non nullo). Ne segue che, essendo f continua e z_1, z_2 derivabili con continuità, le funzioni c_1', c_2' sono continue; si possono quindi antiderivare, ottenendo $c_1(t), c_2(t)$ che, sostituite nella (3.16), forniscono, finalmente, un integrale particolare dell'equazione completa.

Più precisamente: se determiniamo due particolari primitive $c_1(t), c_2(t)$ otteniamo una soluzione particolare dell'equazione completa, del tipo

$$\overline{y}(t) = c_1(t)z_1(t) + c_2(t)z_2(t).$$

Se invece sostituiamo nella (3.16) le *generiche* primitive, sommando cioè costanti arbitrarie di integrazione, otteniamo

$$\begin{aligned} y(t) &= (c_1(t) + k_1)z_1(t) + (c_2(t) + k_2)z_2(t) = \\ &= \overline{y}(t) + k_1 z_1(t) + k_2 z_2(t) \end{aligned}$$

ovvero l'integrale generale dell'equazione completa.

Il problema di scrivere una soluzione dell'equazione completa è dunque ricondotto a quello del calcolo di due integrali indefiniti. (In pratica, solo in casi molto semplici tali integrazioni si riescono ad eseguire in forma esplicita).

Esempio

Scriviamo una formula che assegni l'integrale generale dell'equazione:

$$y'' + \omega^2 y = f(t)$$

qualunque sia il termine noto $f(t)$.

Soluzioni indipendenti dell'equazione omogenea:

$$z_1 = \cos\omega t \qquad z_2 = \sin\omega t.$$

Risolviamo il sistema:

$$\begin{cases} c_1' \cos\omega t + c_2' \sin\omega t = 0 \\ -\omega c_1' \sin\omega t + \omega c_2' \cos\omega t = f(t) \end{cases}$$

$$\begin{cases} c_2' = -c_1' \dfrac{\cos\omega t}{\sin\omega t} \\ \omega c_1' (-\sin\omega t - \dfrac{\cos\omega t}{\sin\omega t} \cdot \cos\omega t) = f(t) \end{cases}$$

$$\begin{cases} c_1' = -\dfrac{1}{\omega} f(t) \sin\omega t \\ c_2' = \dfrac{1}{\omega} f(t) \cos\omega t \end{cases}$$

$$c_1(t) = \int_0^t -\frac{1}{\omega} f(\tau)\sin\omega\tau\, d\tau \qquad c_2(t) = \int_0^t \frac{1}{\omega} f(\tau)\cos\omega\tau\, d\tau.$$

Una soluzione particolare dell'equazione completa è dunque:

$$\begin{aligned} \overline{y}(t) &= \cos\omega t \int_0^t -\frac{1}{\omega} f(\tau)\sin\omega\tau\, d\tau + \sin\omega t \int_0^t \frac{1}{\omega} f(\tau)\cos\omega\tau\, d\tau = \\ &= \int_0^t \frac{1}{\omega} f(\tau)[\sin\omega t\cos\omega\tau - \cos\omega t\sin\omega\tau]d\tau = \\ &= \int_0^t \frac{1}{\omega} f(\tau)\sin[\omega(t-\tau)]d\tau \end{aligned}$$

che è la convoluzione (vedi volume 1, capitolo 6, paragrafo 10) del termine noto $f(t)$ con la funzione $\frac{\sin\omega t}{\omega}$, pensando entrambe le funzioni uguali a zero per $t < 0$.

L'integrale generale dell'equazione completa ha ovviamente la forma:

$$y(t) = \overline{y}(t) + c_1\cos\omega t + c_2\sin\omega t.$$

3.6 Vibrazioni meccaniche

In questo paragrafo passeremo in rassegna certe equazioni lineari del second'ordine di particolare importanza modellistica. Alcune delle equazioni qui considerate sono già state studiate in esempi precedenti; qui si intende sottolineare in particolare il significato fisico dei parametri in gioco. Nel capitolo 8, paragrafo 3.2, riprenderemo questo discorso dal punto di vista della stabilità degli equilibri.

Vibrazioni libere

Un problema di fondamentale interesse in meccanica è quello di calcolare il movimento di un punto materiale P soggetto ad una forza elastica attrattiva (cioè ad una forza centrale proporzionale alla distanza dal centro O come, ad esempio, quella esercitata da una molla). Come è noto (o come è facile dimostrare), se la velocità iniziale del punto è radiale, cioè diretta come la retta OP (v. fig. 1.8), il movimento si svolgerà su questa retta; fissata su di essa l'origine (in O) e un verso, detta $y(t)$ l'ascissa di P all'istante t, l'equazione del moto del punto è:

$$m\ddot{y} = -ky$$

cioè

$$\ddot{y} + \omega^2 y = 0 \tag{3.20}$$

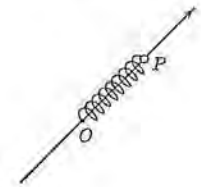


Figura 1.8.

avendo posto $\omega = \sqrt{k/m}$, dove $k > 0$ è la costante della molla, m la massa del punto.

L'integrale generale della (3.20) è: $\varphi(t) = c_1\cos\omega t + c_2\sin\omega t$ (c_1, c_2 costanti arbitrarie), che si può anche scrivere nella forma, più significativa

$$\varphi(t) = A\cos(\omega t + \alpha) \tag{3.21}$$

dove A e α sono costanti arbitrarie che possono essere determinate assegnando la posizione e la velocità iniziali del punto. Questo esegue infinite oscillazioni di *periodo* $2\pi/\omega$, di *ampiezza* A, mentre α rappresenta la *fase iniziale*. Il *moto* è detto *armonico*; la frequenza $2\pi/\omega$ è detta *frequenza caratteristica* dell'oscillatore. Si parla, in questo caso, di *oscillazioni libere*.

Vibrazioni smorzate

Se sul punto P agisce, oltre alla forza elastica, anche una resistenza di tipo viscoso, cioè proporzionale alla velocità (per esempio, la resistenza dell'aria), allora l'equazione del moto è: $m\ddot{y} = -ky - h\dot{y}$, che può scriversi:

$$\ddot{y} + 2\delta\dot{y} + \omega^2 y = 0 \tag{3.22}$$

 © 978-88-08-06485-1

avendo posto $\delta = h/2m$; questo parametro, positivo, rappresenta la resistenza del mezzo al moto del punto. L'equazione caratteristica della (3.22): $\lambda^2 + 2\delta\lambda + \omega^2 = 0$ possiede le due radici

$$\lambda_1 = -\delta + \sqrt{\delta^2 - \omega^2} \qquad \lambda_2 = -\delta - \sqrt{\delta^2 - \omega^2}.$$

Si presentano i seguenti tre casi:

i) $\delta > \omega$ (resistenza elevata). Le radici λ_1 e λ_2 sono reali negative, $\lambda_1 \neq \lambda_2$. L'integrale generale di (3.22) è:

$$\varphi(t) = c_1 e^{\lambda_1 t} + c_2 e^{\lambda_2 t}. \tag{3.23}$$

Il *moto non è oscillatorio.*

Lo studente verifichi le seguenti affermazioni.

Si assuma che il punto si trovi inizialmente in $\bar{y}$ e abbia una assegnata velocità iniziale; se tale velocità è centrifuga[10], la y raggiunge (in valore assoluto) un massimo e poi decresce tendendo asintoticamente a zero (fig. 1.9a); se la velocità iniziale è centripeta e grande (precisare quanto grande!) il punto passa una volta per l'origine, procede oltre fino ad un certo limite e poi inverte il moto, tendendo asintoticamente verso il centro (fig. 1.9b); se la velocità iniziale è centripeta e piccola, il punto tende asintoticamente verso il centro senza eseguire alcuna oscillazione (fig. 1.9c).

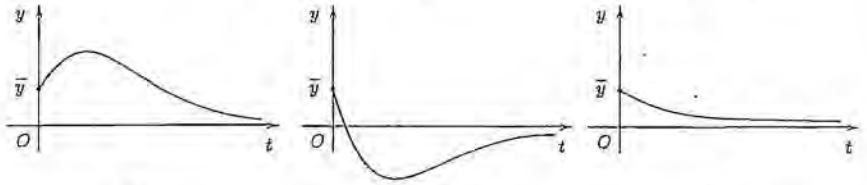


Figura 1.9. Vibrazioni smorzate.

ii) $\delta = \omega$. Allora $\lambda_1 = \lambda_2 = -\delta < 0$. L'integrale generale di (3.22) è:

$$\varphi(t) = (c_1 + c_2 t)e^{-\delta t}. \tag{3.24}$$

Ancora il *moto non* è *oscillatorio*; il comportamento è simile a quello discusso in i).

iii) $\delta < \omega$ (resistenza debole). Abbiamo $\lambda_1 = -\delta + i\nu, \lambda_2 = -\delta - i\nu, \nu = \sqrt{\omega^2 - \delta^2}$

L'integrale generale della (3.22) si può scrivere nella forma:

$$\varphi(t) = Ae^{-\delta t}\cos(\nu t + \alpha) \tag{3.25}$$

con A, α costanti arbitrarie. Il punto mobile esegue infinite oscillazioni attorno al centro, di frequenza $\nu = \sqrt{\omega^2 - \delta^2}$ e di ampiezza $Ae^{-\delta t}$; l'ampiezza delle oscillazioni diventa dunque sempre più piccola col passare del tempo e il punto tende asintoticamente all'origine (fig. 1.10). Si parla, in tal caso, di *oscillazioni smorzate.*

Riassumendo, l'effetto principale della viscosità è quello di smorzare o addirittura di eliminare le oscillazioni, e il punto P tende asintoticamente alla sua posizione di equilibrio naturale (il centro O).

[10]Con le nostre notazioni, la velocità è centrifuga se è positiva, centripeta se è negativa.

© 978-88-08-06485-1

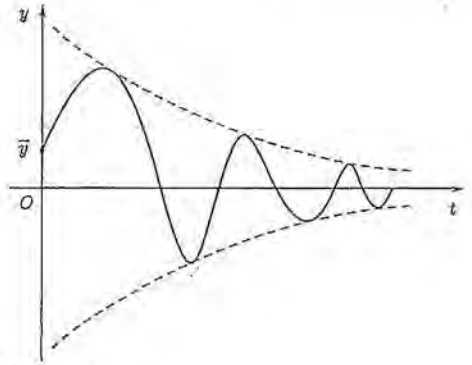


Figura 1.10. Oscillazioni smorzate.

Vibrazioni forzate. Risonanza

Supponiamo che sul punto P agisca, oltre alla forza elastica e alla resistenza viscosa, anche una forza impressa dall'esterno (diretta come l'asse y, cosicché il moto permane rettilineo). L'equazione del moto è ora non omogenea: $m\ddot{y} = -ky - h\dot{y} + F(t)$; ponendo $f(t) = F(t)/m$ (e adottando le stesse notazioni già introdotte in (3.20) e (3.22)) l'equazione prende la forma:

$$\ddot{y} + 2\delta\dot{y} + \omega^2 y = f(t). \tag{3.26}$$

La soluzione generale della (3.26) sarà del tipo:

$$y(t) = \varphi(t) + \psi(t)$$

dove $\varphi(t)$ è l'integrale generale dell'omogenea associata e $\psi(t)$ è una soluzione particolare della (3.26).

Consideriamo un caso tipico:

$$f(t) = B\cos\gamma t \tag{3.27}$$

(se $\gamma = 0$ la forza impressa è costante; altrimenti è armonica, cioè periodica semplice, con frequenza $\gamma/2\pi$).

i) Supponiamo dapprima che la resistenza sia nulla ($\delta = 0$). Allora un integrale particolare della (3.26) è (verificare!):

$$\psi(t) = \begin{cases} \dfrac{B}{\omega^2 - \gamma^2}\cos\gamma t & \text{se } \gamma \neq \omega \\ \dfrac{B}{2\omega} t\sin\omega t & \text{se } \gamma = \omega. \end{cases} \tag{3.28}$$

Pertanto, alle oscillazioni libere, rappresentate da $\varphi(t)$ in (3.21), si sovrappongono delle *oscillazioni forzate*, rappresentate da $\psi(t)$ in (3.28); se $\gamma \neq \omega$ queste sono di ampiezza costante pari a $B/(\omega^2 - \gamma^2)$ e la soluzione è sovrapposizione di due oscillazioni periodiche di pulsazioni γ e ω; essa può anche non essere periodica. Il caso

© 978-88-08-06485-1

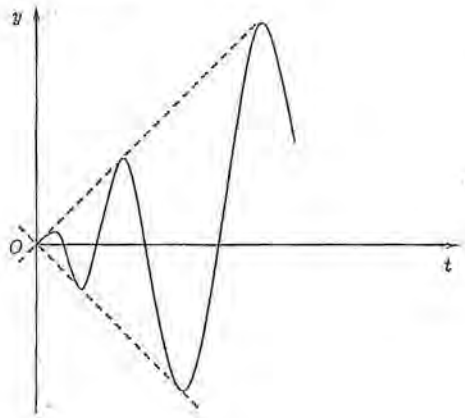


Figura 1.11. Risonanza.

interessante si ha quando la forza impressa oscilla con la stessa frequenza caratteristica dell'oscillatore ($\omega = \gamma$); allora le oscillazioni forzate hanno un'ampiezza ($tB/2\omega$) che si amplifica col tempo, crescendo oltre ogni limite; è questo il fenomeno della *risonanza* (fig. 1.11).

ii) Sia ora $\delta > 0$. Allora un integrale particolare della (3.26) è:

$$\psi(t) = B\rho\cos(\gamma(t-\tau)) \tag{3.29}$$

dove si è posto:

$$\rho = \frac{1}{\sqrt{(\omega^2-\gamma^2)^2+4\delta^2\gamma^2}} \qquad \operatorname{tg}\gamma\tau = \frac{2\gamma\delta}{\omega^2-\gamma^2} \tag{3.30}$$

(si osservi che la (3.29) è ora un integrale particolare della (3.26) qualunque sia il valore di γ).

Dalla discussione svolta in precedenza sappiamo che l'integrale generale della omogenea, $\varphi(t)$, tende esponenzialmente a zero per $t \to +\infty$ (si vedano le formule (3.23), (3.24), (3.25) che assegnano l'integrale generale); perciò, dopo un transitorio iniziale, la situazione di regime è descritta dall'integrale particolare $\psi(t)$. Dalla (3.29) ricaviamo perciò che il *moto* a regime è *armonico*, con lo stesso *periodo* $2\pi/\gamma$ della forza impressa, con *ampiezza* $B\rho$ costante (dipendente dai vari parametri del sistema) e con *ritardo di fase* τ rispetto alla forza impressa. È interessante osservare come varia l'ampiezza $B\rho$ di queste oscillazioni forzate al variare della frequenza di eccitazione γ (per valori fissati di ω, δ, B); bisogna studiare cioè la funzione $\rho = \rho(\gamma)$. Si osserva che, se $\delta < \omega/\sqrt{2}$ (resistenza debole), la funzione $\rho(\gamma)$ presenta un massimo per $\gamma = \sqrt{\omega^2 - 2\delta^2}$ e tale massimo è

$$\rho_{\max} = \frac{1}{2\delta\sqrt{\omega^2-\delta^2}}.$$

Appare cioè ancora il fenomeno della *risonanza*: l'ampiezza dell'oscillazione, per quanto in questo caso sia limitata, può essere molto grande se la resistenza δ è piccola. Rispetto alla situazione descritta in i) ($\delta = 0$) si osserva che ora la risonanza compare quando la frequenza di eccitazione è "vicina" (ma non uguale) a quella caratteristica ω e l'ampiezza delle oscillazioni forzate rimane limitata. Se invece $\delta \geq \omega/\sqrt{2}$, la funzione $\rho(\gamma)$ è decrescente e non si ha risonanza (fig. 1.12)

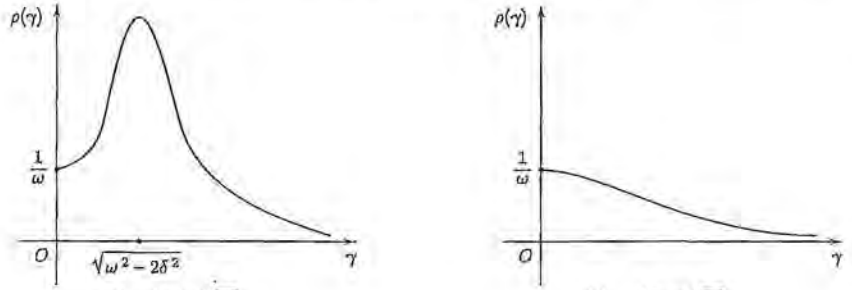


Figura 1.12.

Concludiamo con la seguente osservazione. Si consideri un circuito elettrico costituito da una resistenza R, un'induttanza L, un condensatore di capacità C, a cui sia applicata una forza elettromotrice variabile $V(t)$ (fig. 1.13).

La differenza di potenziale ai capi del condensatore è soluzione della seguente equazione:

$$CL\ddot{y} + CR\dot{y} + y = -V(t)$$

formalmente identica a quella delle vibrazioni meccaniche. Lo studente interessato può rivedere tutta la discussione svolta sopra e reinterpretare i parametri nei termini adatti alla nuova situazione.

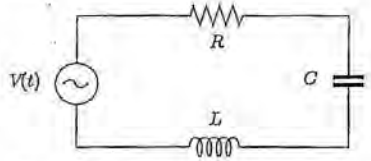


Figura 1.13.

Esercizi

Come suggerito alla fine del paragrafo 3.4, si verifichi, calcolando la matrice Wronskiana, che per l'equazione lineare omogenea a coefficienti costanti le tre coppie di soluzioni (3.13) calcolate in ciascuno dei tre casi sono effettivamente indipendenti.

Siano z_1, z_2 due soluzioni indipendenti dell'equazione omogenea

$$z'' + a(t)z' + b(t)z = 0$$

in un intervallo I e sia $W(t) = z_1z_2' - z_2z_1'$ il Wronskiano di questa coppia di soluzioni (v. paragrafo 3.3). Dimostrare che W risolve in I l'equazione differenziale:

$$W' = -a(t)W.$$

© 978-88-08-06485-1

Di conseguenza è

$$W(t) = W(t_0)\, e^{-\int_{t_0}^{t} a(\tau)d\tau}, \text{ per ogni } t_0, t \in I.$$

L'ultima relazione scritta fornisce un altro modo per dimostrare che il Wronskiano o si annulla identicamente oppure non si annulla mai (v. Teorema 1.7).

24 Rifare gli esercizi contenuti negli esempi 3.7, 3.8, 3.9, 3.10, 3.11 utilizzando il metodo di variazione delle costanti, anziché il metodo di somiglianza.

25 Risolvere il seguente problema di Cauchy:

$$\begin{cases} u'' + 3u' - 4u = 0 \\ u(0) = 1 \\ u'(0) = 2. \end{cases}$$

26 Risolvere il seguente problema di Cauchy:

$$\begin{cases} u'' + 3u' + 4u = 0 \\ u(0) = 0 \\ u'(0) = 1. \end{cases}$$

27 Risolvere il seguente problema di Cauchy:

$$\begin{cases} u'' + 4u' + 4u = 0 \\ u(0) = 3 \\ u'(0) = 2. \end{cases}$$

28 Risolvere il seguente problema di Cauchy:

$$\begin{cases} u'' + 13u' - 104u = 0 \\ u(0) = 0 \\ u'(0) = 0. \end{cases}$$

È necessario fare calcoli?

29 Risolvere il seguente problema di Cauchy:

$$\begin{cases} y'' - 2y' + 4y = 0 \\ y(0) = 1 \\ y'(0) = 4. \end{cases}$$

30 Si consideri l'equazione:

$$y'' + 3y' + 2y = t^2 + 2t.$$

a) Determinare una soluzione particolare di tale equazione.
b) Scrivere l'integrale generale dell'equazione.
c) Per tempi lunghi, qual è il regime permanente?

31 Determinare una soluzione particolare dell'equazione:

$$y'' + 3y' = x^2 + 1.$$

© 978-88-08-06485-1

32 Si consideri l'equazione:

$$y'' + 3y' + 2y = 2\sin 4x.$$

a) Determinare una soluzione particolare dell'equazione.
b) Determinare l'integrale generale dell'equazione.
c) Risolvere il problema di Cauchy con le condizioni iniziali:

$$y(0) = 0,\ y'(0) = 0.$$

33 Si consideri l'equazione:

$$y'' - 2y' + y = 2e^x \cos 3x.$$

a) Determinare una soluzione particolare dell'equazione.
b) Determinare l'integrale generale dell'equazione.
c) Risolvere il problema di Cauchy con le condizioni iniziali:

$$y(0) = 1,\ y'(0) = 0.$$

34 Determinare una soluzione particolare dell'equazione:

$$y'' + 2y' + 3y = 3e^{-x}\cos\sqrt{2}x.$$

35 Determinare una soluzione particolare dell'equazione:

$$y'' + y = e^x(x^2 - 1).$$

36 Determinare una soluzione particolare dell'equazione:

$$y'' - 2y' + y = e^x.$$

37 Determinare l'integrale generale dell'equazione:

$$y'' + y' - 2y = e^x + x.$$

38 Determinare l'integrale generale dell'equazione:

$$y'' + 2y' + 3y = 2x^2 - 1.$$

39 Determinare l'integrale generale dell'equazione:

$$y'' + 9y = t - e^t.$$

40 Determinare l'integrale generale dell'equazione:

$$y'' + y = \cos 3t.$$

41 L'equazione

$$y'' + \frac{a}{x}y' + \frac{b}{x^2}y = 0,$$

 © 978-88-08-06485-1

con a, b costanti, è uno dei pochi esempi di equazione lineare omogenea del secondo ordine a coefficienti variabili che si sappia risolvere in modo elementare: si chiama *equazione di Eulero*. Cercare due soluzioni indipendenti, del tipo

$$y = x^\alpha \quad \text{per} \quad x > 0$$

e $\alpha \in \mathbb{R}$ da determinarsi. Con questo metodo, scrivere l'integrale generale dell'equazione:

$$y'' - \frac{4}{x}y' + \frac{1}{x^2}y = 0.$$

Sia I un intervallo. Di ciascuno dei seguenti sottoinsiemi di $C^2(I)$, si dica se costituisce uno spazio vettoriale oppure no, giustificando la risposta, e in caso affermativo specificando la dimensione dello spazio:

(a) L'insieme delle soluzioni in I di un'equazione differenziale lineare del second'ordine completa (a coefficienti continui);

(b) L'insieme delle soluzioni in I di un'equazione differenziale lineare del second'ordine omogenea (a coefficienti continui);

(c) L'insieme delle soluzioni in I di un'equazione differenziale lineare del second'ordine omogenea (a coefficienti continui) che soddisfano l'ulteriore condizione $y'(t_0) = 0$ (per un certo $t_0 \in I$ fissato).

4 COMPLEMENTI

4.1 Teorema di esistenza e unicità per le equazioni a variabili separabili

Vogliamo dimostrare qui, nel caso particolare delle equazioni del prim'ordine a variabili separabili, il teorema di esistenza e unicità della soluzione di un problema di Cauchy. Questo risultato, pur essendo un caso molto particolare del Teorema generale di esistenza e unicità che enunceremo (ma non dimostreremo) nel capitolo 8, è istruttivo per le seguenti ragioni:

- si applica ad una classe di problemi che non sempre sappiamo risolvere esplicitamente "a mano";
- riguarda una classe di equazioni non lineari;
- evidenzia il fatto che le ipotesi sotto cui si può stabilire l'esistenza di una soluzione sono più deboli di quelle sotto cui si può stabilire l'unicità della soluzione stessa.

Riportiamo per comodità l'enunciato del teorema già presentato nel paragrafo 2.2.

TEOREMA 1.8 *Consideriamo il problema di Cauchy:*

$$\begin{cases} y' = a(t)\, b(y) \\ y(t_0) = y_0 \end{cases} \tag{4.1}$$

dove a è una funzione continua in un intorno I di t_0 e b è una funzione continua in un intorno J di y_0. Allora esiste un intorno di t_0 $I' \subset I$ e una funzione $y \in C^1(I')$ soluzione del problema (4.1).

Se inoltre $b \in C^1(J)$ (o, più in generale, b è una funzione continua, eventualmente non derivabile ma con rapporto incrementale limitato in J), allora tale soluzione è unica.

DIMOSTRAZIONE. Cominciamo dalla prima parte del teorema, che stabilisce l'*esistenza* di una soluzione. La dimostrazione ricalca la tecnica risolutiva che abbiamo illustrato nel paragrafo 2.2.

Se $b(y_0) = 0$, la funzione costante $y = y_0$ risolve evidentemente il problema, in tutto I, e l'esistenza di una soluzione è in tal caso dimostrata.

Sia dunque $b(y_0) \neq 0$. Poiché b è continua, per il teorema di permanenza del segno questo implica che $b(y)$ è diversa da zero e ha segno costante in un certo intorno di y_0 $J' \subset J$. In tale intorno $1/b(y)$ è continua, perciò ammette una primitiva $B(y)$. Anche a, essendo continua, ammette una primitiva $A(t)$ in I. Vogliamo determinare una funzione $y(t)$ che soddisfi l'identità:

$$B(y(t)) - B(y_0) = A(t) - A(t_0) \text{ per ogni } t \in I',$$

dove I' è un opportuno intorno di t_0 da determinarsi, $I' \subset I$. Poiché $B'(y) = 1/b(y)$ e b ha segno costante in J', $B(y)$ è una funzione strettamente monotona in J', e quindi invertibile in J'. Detta B^{-1} la sua inversa, che è definita nell'intervallo $B(J')$, possiamo scrivere quindi:

$$y(t) = B^{-1}(B(y_0) + A(t) - A(t_0)) \text{ per ogni } t \in I', \tag{4.2}$$

dove I' è stato scelto in modo tale che

$$t \in I' \Longrightarrow B(y_0) + A(t) - A(t_0) \in B(J').$$

Questo è possibile perché A è continua e per $t = t_0$ è $B(y_0) + A(t) - A(t_0) = B(y_0) \in B(J')$.

La funzione $y(t)$ definita da (4.2) soddisfa il problema di Cauchy (4.1). Anzitutto y è di classe C^1, perché composta di funzioni C^1; infatti, poiché A e B sono primitive di funzioni continue, esse sono di classe C^1 nei rispettivi domini di definizione, e B^{-1} è anch'essa C^1.

Inoltre:

$$y(t_0) = B^{-1}(B(y_0) + A(t_0) - A(t_0)) = B^{-1}(B(y_0)) = y_0,$$

mentre, per il teorema di derivazione della funzione inversa e della funzione composta:

$$y'(t) = \frac{1}{B'(B^{-1}(B(y_0) + A(t) - A(t_0)))} \cdot A'(t) = \frac{1}{B'(y(t))} \cdot A'(t) =$$

(ricordando che A è primitiva di a e B è primitiva di $1/b$)

$$= b(y(t))\, a(t).$$

Questo completa la dimostrazione della prima parte del teorema.

Proviamo ora l'*unicità* della soluzione, sotto l'ulteriore ipotesi che b abbia rapporto incrementale limitato, ossia che valga

$$|b(y_1) - b(y_2)| \leq c|y_1 - y_2|$$

per qualche costante c, per ogni $y_1, y_2 \in J$. (Questo in particolare sarà vero se $b \in C^1$, pur di restringere eventualmente l'intervallo J).

Supponiamo dunque che $y_1(t)$ e $y_2(t)$ siano due soluzioni del problema (4.1), e mostriamo che in realtà coincidono. Sia

$$z(t) = y_1(t) - y_2(t).$$

Si ha:

$$\begin{aligned} z(t_0) &= y_1(t_0) - y_2(t_0) = y_0 - y_0 = 0; \\ z'(t) &= y_1'(t) - y_2'(t) = a(t)\,[b(y_1(t)) - b(y_2(t))]. \end{aligned}$$

Poiché a è continua in I, e quindi limitata in qualunque intervallo $I' \subset I$, scegliendo I' sufficientemente piccolo in modo che risulti anche $y_1(t), y_2(t) \in J$ per ogni $t \in I'$, possiamo scrivere:

$$|z'(t)| \leq c|a(t)|\,|y_1(t) - y_2(t)| \leq M|y_1(t) - y_2(t)| = M|z(t)|$$

© 978-88-08-06485-1

per ogni $t \in I'$ e un'opportuna costante M. Ricordando che $z(t_0) = 0$, possiamo scrivere:

$$z(t) = \int_{t_0}^{t} z'(s)\,ds$$

e quindi, per $t \in I' \equiv [t_0 - \delta, t_0 + \delta]$,

$$|z(t)| \leq \left| \int_{t_0}^{t} |z'(s)|\,ds \right| \leq M \left| \int_{t_0}^{t} |z(s)|\,ds \right| \leq M\delta \cdot \max_{s \in I'} |z(s)|$$

da cui

$$\max_{s \in I'} |z(s)| \leq M\delta \cdot \max_{s \in I'} |z(s)|.$$

Riducendo eventualmente δ in modo che risulti $M\delta < 1$, si deduce che $\max_{s \in I'} |z(s)| = 0$, ossia $z(t)$ è identicamente nulla in I'. Dunque in questo intorno di t_0 le due soluzioni $y_1(t), y_2(t)$ coincidono. Ripetendo il ragionamento a partire da $t_0 + \delta$ e $t_0 - \delta$ si prolunga l'intervallo di coincidenza a $[t_0 - 2\delta, t_0 + 2\delta]$. Dopo un numero finito di passi si mostra infine che le due soluzioni coincidono in tutto l'intervallo in cui sono entrambe definite. ☆

4.2 Cenni alle equazioni lineari di ordine n

Le proprietà delle equazioni lineari del second'ordine studiate nel paragrafo 3 si generalizzano in modo molto naturale alle equazioni di ordine n. Le esporremo molto schematicamente.

Consideriamo l'equazione lineare completa in forma normale:

$$y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y' + a_0 y = f(t) \tag{4.3}$$

con a_i $(i = 0, 1, 2, \ldots, n-1)$ e f funzioni continue in un intervallo I, e l'equazione omogenea associata:

$$z^{(n)} + a_{n-1}z^{(n-1)} + \ldots + a_1 z' + a_0 z = 0. \tag{4.4}$$

Struttura dell'integrale generale

Valgono le seguenti proprietà:

- L'integrale generale della (4.3) si ottiene come somma dell'integrale generale della (4.4) e di un integrale particolare (qualsiasi) della (4.3).
- L'integrale generale della (4.4) si ottiene prendendo tutte le combinazioni lineari di n soluzioni linearmente indipendenti della (4.4). Si può quindi affermare che l'insieme delle soluzioni della (4.4) costituisce uno spazio vettoriale di dimensione n.
- Per riconoscere se n funzioni $z_1(t), z_2(t), \ldots, z_n(t)$, soluzioni della (4.4) nell'intervallo I, sono linearmente indipendenti, si può utilizzare ancora la *matrice wronskiana* relativa a questo sistema di soluzioni, che in questa situazione più generale è definita come segue:

$$W(t) = \begin{bmatrix} z_1(t) & z_2(t) & \ldots & z_n(t) \\ z_1'(t) & z_2'(t) & \ldots & z_n'(t) \\ \vdots & \vdots & & \vdots \\ z_1^{(n-1)}(t) & z_2^{(n-1)}(t) & \ldots & z_n^{(n-1)}(t) \end{bmatrix}.$$

Il determinante di questa matrice si chiama *wronskiano* dell'equazione. Vale la seguente

PROPOSIZIONE 1.1 *Condizione necessaria e sufficiente perché n soluzioni dell'equazione (4.4) siano linearmente indipendenti in $C^n(I)$ è che il determinante wronskiano sia diverso da zero $\forall t \in I$. Inoltre, il wronskiano di un sistema di soluzioni dell'equazione (4.4) o è identicamente nullo su I oppure è diverso da zero in ogni punto di I.*

© 978-88-08-06485-1

Dunque anche in questo caso è sufficiente verificare il criterio in un solo punto. L'ultima affermazione si può vedere anche come conseguenza del prossimo risultato[11]:

TEOREMA 1.9 (DI LIOUVILLE) *Il Wronskiano di una n-upla di soluzioni dell'equazione (4.4) in I soddisfa l'equazione differenziale*

$$W' = -a_{n-1}(t)\,W;$$

in particolare, per ogni $t_0, t \in I$ si può scrivere:

$$W(t) = W(t_0)\,e^{-\int_{t_0}^{t} a_{n-1}(\tau)\,d\tau}.$$

Problema di Cauchy

Vale il seguente risultato, perfettamente analogo al Teorema 1.4:

TEOREMA 1.10 *Siano f e a_i $(i = 0, 1, \ldots, n-1)$ funzioni continue in un intervallo $I \ni t_0$; allora per ogni $b_0, b_1, \ldots, b_{n-1} \in \mathbb{R}$ assegnati, il problema di Cauchy*

$$\begin{cases} y^{(n)} + a_{n-1}y^{(n-1)} + \ldots + a_1 y' + a_0 y = f(t) \\ y(t_0) = b_0 \\ y'(t_0) = b_1 \\ \ldots \\ y^{(n-1)}(t_0) = b_{n-1} \end{cases} \tag{4.5}$$

ammette una e una sola soluzione $y \in C^n(I)$.

Equazioni a coefficienti costanti

Nel caso particolare in cui i coefficienti dell'equazione (4.4) siano *costanti*, esiste un metodo standard per determinare n soluzioni indipendenti della (4.4): cercare, come per le equazioni del secondo ordine, soluzioni di tipo esponenziale e^{rt} $(r \in \mathbb{C})$. In pratica, si procede al modo seguente:

1. Si considera l'equazione caratteristica associata alla (4.4):

$$r^n + a_{n-1}r^{n-1} + \ldots + a_1 r + a_0 = 0 \tag{4.6}$$

Questa è un'equazione polinomiale di grado n a coefficienti reali, che può avere radici reali o coppie di radici complesse coniugate, ciascuna delle quali può essere semplice o avere molteplicità maggiore di 1.

2. Ogni radice r reale della (4.6), di molteplicità s, fornisce le s soluzioni della (4.4):

$$e^{rt}, te^{rt}, t^2e^{rt}, \ldots, t^{s-1}e^{rt}$$

3. Ogni coppia di radici complesse coniugate della (4.6), $\alpha \pm i\beta$, di molteplicità s, fornisce s coppie di soluzioni della (4.4):

$$\begin{array}{ll} e^{\alpha t}\cos\beta t & e^{\alpha t}\sin\beta t \\ te^{\alpha t}\cos\beta t & te^{\alpha t}\sin\beta t \\ \ldots\ldots\ldots\ldots & \ldots \\ t^{s-1}e^{\alpha t}\cos\beta t & t^{s-1}e^{\alpha t}\sin\beta t \end{array}$$

[11] Si confronti con quanto provato, per le equazioni del second'ordine, nell'Esercizio 23.

© 978-88-08-06485-1

Il procedimento appena illustrato fornisce sempre n soluzioni indipendenti della (4.4) (pur di saper risolvere l'equazione polinomiale (4.6)!) e permette quindi di scriverne l'integrale generale. L'indipendenza delle soluzioni sopra trovate si può verificare applicando la Proposizione 1.1.

– Una volta scritto l'integrale generale della (4.4), una soluzione particolare della (4.3) può essere determinata o con il *metodo di somiglianza*, adattando al caso presente le considerazioni svolte nel paragrafo 3.5, o col *metodo di variazione delle costanti*, che nel caso delle equazioni di ordine n acquista la forma seguente:

1. Si considerano le n soluzioni indipendenti dell'equazione omogenea (4.4), determinate nel modo descritto in precedenza,

$$z_1(t), z_2(t), \dots, z_n(t),$$

definite per $t \in \mathbb{R}$, e si cerca una soluzione particolare della (4.3), del tipo seguente:

$$\bar{y}(t) = c_1(t)z_1(t) + c_2(t)z_2(t) + \dots + c_n(t)z_n(t) \tag{4.7}$$

dove $c_i(t)$ sono n funzioni incognite, da determinarsi imponendo su di esse n condizioni opportune.

2. Si risolve ora il seguente sistema lineare di n equazioni nelle n incognite $c_i'(t)$ (le derivate delle funzioni incognite):

$$\begin{cases} c_1'z_1 + c_2'z_2 + \dots + c_n'z_n = 0 \\ c_1'z_1' + c_2'z_2' + \dots + c_n'z_n' = 0 \\ \dots \\ c_1'z_1^{(n-2)} + c_2'z_2^{(n-2)} + \dots + c_n'z_n^{(n-2)} = 0 \\ c_1'z_1^{(n-1)} + c_2'z_2^{(n-1)} + \dots + c_n'z_n^{(n-1)} = f(t) \end{cases}$$

Il determinante di questo sistema è esattamente il determinante wronskiano associato alle n soluzioni $z_i(t)$; poiché per ipotesi queste soluzioni sono indipendenti, per la Proposizione 1.1 il determinante è diverso da zero per ogni $t \in \mathbb{R}$; di conseguenza il sistema lineare ha sempre un'unica soluzione $(c_1'(t), c_2'(t), \dots, c_n'(t))$ definita almeno sull'intervallo I in cui $f(t)$ è definita e continua.

3. Si calcola ora una primitiva di ciascuna delle n funzioni $c_n'(t)$ soluzione del sistema precedente, determinando così n funzioni $c_i(t)$ che, sostituite nella (4.7), forniscono una soluzione particolare della (4.3).

– Noto l'integrale generale della (4.4) e un integrale particolare della (4.3), si può ora scrivere l'integrale generale della (4.3).

– Per risolvere il *problema di Cauchy* (4.5) per l'equazione (4.3), una volta determinato l'integrale generale della (4.3) è sufficiente calcolare le derivate fino all'ordine $n-1$ della soluzione generica (dipendente da n costanti arbitrarie $c_1, c_2, \dots, c_n$), imporre le condizioni iniziali e risolvere il sistema lineare di n equazioni nelle n incognite $c_1, c_2, \dots, c_n$ che si ottiene in questo modo. (Il sistema risulta avere determinante diverso da zero per l'ipotesi di indipendenza delle n soluzioni dell'equazione omogenea; pertanto ha un'unica soluzione). I valori delle costanti così ottenuti determinano la soluzione del problema di Cauchy.

ESEMPIO

Equazione della trave. Consideriamo una trave a sezione costante, composta da un materiale omogeneo, isotropo e linearmente elastico; rappresentiamo per semplicità la

© 978-88-08-06485-1

trave mediante un segmento AB sull'asse delle x (che supponiamo essere l'asse baricentrico). La trave (vedi fig. 1.14) è appoggiata alle estremità e soggetta a un carico distribuito $q(x)$ agente nel piano xy. A causa di questo carico, l'asse baricentrico AB inizialmente rettilineo, cambia forma e diventa una curva $y = y(x)$. Si può dimostrare che questa curva è governata dall'equazione:

$$\frac{d^4y}{dx^4} EI = q(x)$$

dove E, detto *modulo di Young*, è una caratteristica del materiale, e I è una caratteristica della sezione.

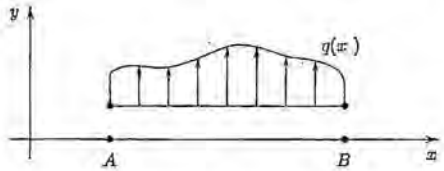


Figura 1.14.

Nel caso elementare in cui il carico è costante, $q(x) = q$, l'equazione si integra immediatamente:

$$y(x) = \frac{q}{24EI}x^4 + c_1x^3 + c_2x^2 + c_3x + c_4$$

dove c_i $(i = 1, 2, 3, 4)$ sono costanti arbitrarie, da determinarsi imponendo quattro condizioni opportune.

Supponiamo ora che la trave sia appoggiata su un suolo elastico; l'equazione differenziale si modifica allora nel modo seguente:

$$\frac{d^4y}{dx^4} EI + ky = q(x)$$

A titolo d'esempio, integriamo l'equazione precedente nel caso particolare $q(x) = q$ (costante). Una soluzione particolare dell'equazione completa è la costante

$$\bar{y} = \frac{q}{k}$$

mentre l'equazione omogenea ha equazione caratteristica

$$r^4 + \frac{k}{EI} = 0$$

che dà

$$r = \left(-\frac{k}{EI}\right)^{1/4} = \begin{cases} \sqrt[4]{\frac{k}{EI}}\left(\frac{1 \pm i}{\sqrt{2}}\right) = \omega(1 \pm i) \\ \sqrt[4]{\frac{k}{EI}}\left(\frac{-1 \pm i}{\sqrt{2}}\right) = \omega(-1 \pm i) \end{cases} \tag{4.8}$$

dove $\omega = \sqrt[4]{\frac{k}{4EI}}$. Nella (4.8), il simbolo $\left(-\frac{k}{EI}\right)^{1/4}$ rappresenta una qualsiasi delle 4 radici quarte complesse di $-k/EI$, mentre $\sqrt[4]{\frac{k}{EI}}$ rappresenta l'unica radice quarta aritmetica (cioè positiva) di k/EI.

L'equazione omogenea ha quindi le quattro soluzioni indipendenti

$$e^{\omega x}\cos\omega x \qquad e^{\omega x}\sin\omega x \qquad e^{-\omega x}\cos\omega x \qquad e^{-\omega x}\sin\omega x$$

e l'integrale generale dell'equazione completa è

$$y(x) = c_1e^{\omega x}\cos\omega x + c_2e^{\omega x}\sin\omega x + c_3e^{-\omega x}\cos\omega x + c_4e^{-\omega x}\sin\omega x + \frac{q}{k}.$$

2 Calcolo infinitesimale per le curve

1 RICHIAMI DI CALCOLO VETTORIALE

Lo studio del calcolo infinitesimale in più variabili, di cui ci occuperemo a partire da questo capitolo, richiede un certo uso del calcolo vettoriale e, più in generale, dell'algebra lineare. Qui assumiamo che lo studente abbia già seguito un corso di geometria e algebra lineare, in cui avrà sicuramente incontrato i concetti che ci serviranno.

È utile comunque richiamare fin d'ora le notazioni vettoriali che useremo in tutto il seguito del libro, per evitare equivoci nell'interpretare certe formule.

Lo spazio $\mathbb{R}^n$ sarà il nostro ambiente naturale. I suoi *elementi* (*vettori*) si indicheranno in grassetto, con una scrittura del tipo:

$$\mathbf{x} = (x_1, x_2, \dots, x_n)$$

dove $x_i \in \mathbb{R}$ sono le *componenti* del vettore.[1]

Il *modulo* di un vettore (chiamato anche *norma*) è

$$|\mathbf{x}| = \sqrt{\sum_{i=1}^{n} x_i^2}.$$

Si dice *versore* un vettore di modulo unitario e, dato un vettore $\mathbf{x}$ non nullo, di modulo qualsiasi, si pone:

$$\operatorname{vers}(\mathbf{x}) = \frac{\mathbf{x}}{|\mathbf{x}|}.$$

La *base canonica* di $\mathbb{R}^n$ è costituita dai versori:

$$\begin{aligned} \mathbf{e}_1 &= (1, 0, 0, \dots, 0) \\ \mathbf{e}_2 &= (0, 1, 0, \dots, 0) \\ &\dots \\ \mathbf{e}_n &= (0, 0, \dots, 0, 1). \end{aligned}$$

[1] Nella scrittura a mano, il vettore si può indicare con $\underline{x}$ o con $\vec{x}$ (anziché col grassetto). Quale che sia la convenzione usata, invitiamo il lettore a *non indicare* i punti di $\mathbb{R}^n$ semplicemente con $x, y, \dots$ (minuscole corsive), ma riservare l'utilizzo di queste lettere per indicare degli scalari (per esempio, ciascuna *coordinata* dei punti di $\mathbb{R}^n$).

 © 978-88-08-06485-1

Perciò si può anche scrivere:

$$(x_1, x_2, \ldots, x_n) = x_1\mathrm{e}_1 + x_2\mathrm{e}_2 + \ldots + x_n\mathrm{e}_n = \sum_{j=1}^{n} x_j\mathrm{e}_j. \tag{1.1}$$

Due vettori x, y si dicono *paralleli* se risulta:

$$\lambda\mathrm{x} = \mu\mathrm{y} \text{ per qualche } \lambda, \mu \in \mathbb{R}$$

(in particolare, uno dei due vettori potrebbe essere nullo); se inoltre $\mathrm{y} = \lambda\mathrm{x}$ con $\lambda > 0$ li diremo paralleli e *concordi*.

Nei casi particolari $n = 2, n = 3$ (il piano e lo spazio tridimensionale della geometria analitica) si usano solitamente notazioni particolari che non fanno uso di indici, come

$$(x, y), (x, y, z)$$

per denotare il generico punto di $\mathbb{R}^2, \mathbb{R}^3$.

I versori della base canonica sono indicati con

$$\mathrm{i}, \mathrm{j}, \mathrm{k} \text{ (in } \mathbb{R}^3\text{) o } \mathrm{i}, \mathrm{j} \text{ in } \mathbb{R}^2,$$

perciò, anziché (x, y, z), si può scrivere, per esempio:

$$x\mathrm{i} + y\mathrm{j} + z\mathrm{k}.$$

Oltre all'operazione di somma tra vettori e prodotto di un vettore per uno scalare, che abbiamo già utilizzato, per esempio nella (1.1), vi sono altre due operazioni comunemente utilizzate. La prima è il *prodotto scalare*:

$$\mathrm{u} \cdot \mathrm{v} = (u_1, u_2, \ldots, u_n) \cdot (v_1, v_2, \ldots, v_n) = u_1v_1 + u_2v_2 + \ldots + u_nv_n.$$

Anziché $\mathrm{u} \cdot \mathrm{v}$, si usa talvolta la notazione $\langle \mathrm{u}, \mathrm{v} \rangle$. Il prodotto scalare è ben definito per due vettori di $\mathbb{R}^n$ e ha per risultato un numero reale. Nel caso particolare $n = 2$, $n = 3$, quest'operazione ha un significato geometrico elementare, essendo in tal caso:

$$\mathrm{u} \cdot \mathrm{v} = |\mathrm{u}|\,|\mathrm{v}| \cos\theta, \tag{1.2}$$

dove θ è l'angolo tra i due vettori. Di conseguenza, $\mathrm{u} \cdot \mathrm{v} = 0$ se e solo se i due vettori sono ortogonali. La formula (1.2), che nel piano e nello spazio tridimensionale si dimostra, nel caso generale di $\mathbb{R}^n$ permette la *definizione* di angolo tra due vettori, se scritta nella forma:

$$\cos\theta = \frac{\mathrm{u} \cdot \mathrm{v}}{|\mathrm{u}|\,|\mathrm{v}|}.$$

In particolare, quindi, in $\mathbb{R}^n$ si dirà che due vettori sono *ortogonali* se $\mathrm{u} \cdot \mathrm{v} = 0$.

Un'altra operazione tra vettori, che introduciamo esclusivamente in $\mathbb{R}^3$, è quella di *prodotto vettoriale*:

$$\mathrm{u} \times \mathrm{v} = \begin{vmatrix} \mathrm{i} & \mathrm{j} & \mathrm{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \mathrm{i}\,(u_2v_3 - u_3v_2) + \mathrm{j}\,(u_3v_1 - u_1v_3) + \mathrm{k}\,(u_1v_2 - u_2v_1).$$

© 978-88-08-06485-1

Il prodotto vettoriale di due vettori di $\mathbb{R}^3$ è un vettore di $\mathbb{R}^3$. Si annulla se e solo se i due vettori sono paralleli.

Infine, il prodotto misto di tre vettori $\mathrm{u}, \mathrm{v}, \mathrm{w} \in \mathbb{R}^3$ è definito come:

$$\mathrm{u} \cdot (\mathrm{v} \times \mathrm{w}) = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

(che si può scrivere anche senza parentesi $\mathrm{u} \cdot \mathrm{v} \times \mathrm{w}$ in quanto c'è solo un modo possibile di interpretare l'ordine in cui vanno eseguite le operazioni: $(\mathrm{u} \cdot \mathrm{v}) \times \mathrm{w}$ non avrebbe senso, essendo il prodotto vettoriale di uno scalare per un vettore). Il prodotto misto si annulla se e solo se i tre vettori sono linearmente dipendenti, ossia se e solo se, spiccando i vettori dall'origine, i tre punti terminali dei vettori e l'origine stessa risultano quattro punti complanari; il modulo di questo vettore ha il significato di volume del parallelepipedo che ha i tre vettori (sempre pensati spiccati da un'unica origine) come spigoli.

Concludiamo con una sintesi schematica delle proprietà delle operazioni tra vettori in $\mathbb{R}^n$, che abbiamo richiamato:

Proprietà della somma di vettori:
per ogni $\mathrm{u} = (u_1, u_2, \ldots, u_n)$, $\mathrm{v} = (v_1, v_2, \ldots, v_n)$, $\mathrm{w} = (w_1, w_2, \ldots, w_n) \in \mathbb{R}^n$ si ha:

$$(\mathrm{u} + \mathrm{v}) + \mathrm{w} = \mathrm{u} + (\mathrm{v} + \mathrm{w})$$
$$\mathrm{u} + \mathrm{v} = \mathrm{v} + \mathrm{u}$$
$$\mathrm{v} + \mathbf{0} = \mathrm{v}, \text{ dove } \mathbf{0} = (0, 0, \ldots, 0)$$
$$\mathrm{v} + (-\mathrm{v}) = \mathbf{0}, \text{ dove } -\mathrm{v} = (-v_1, -v_2, \ldots, -v_n)$$

Proprietà del prodotto per uno scalare:
per ogni $\mathrm{u}, \mathrm{v} \in \mathbb{R}^n$, $s, t \in \mathbb{R}$ si ha:

$$1 \cdot \mathrm{v} = \mathrm{v}$$
$$s(t\,\mathrm{v}) = (st)\mathrm{v}$$
$$t(\mathrm{u} + \mathrm{v}) = t\,\mathrm{u} + t\,\mathrm{v}$$
$$(s + t)\mathrm{v} = s\mathrm{v} + t\mathrm{v}$$

Proprietà del prodotto scalare:
per ogni $\mathrm{u}, \mathrm{v}, \mathrm{w} \in \mathbb{R}^n$, $t \in \mathbb{R}$ si ha:

$$\mathrm{v} \cdot \mathrm{w} = \mathrm{w} \cdot \mathrm{v}$$
$$\mathrm{u} \cdot (\mathrm{v} + \mathrm{w}) = \mathrm{u} \cdot \mathrm{v} + \mathrm{u} \cdot \mathrm{w}$$
$$(t\mathrm{v}) \cdot \mathrm{w} = t(\mathrm{v} \cdot \mathrm{w})$$
$$\mathrm{v} \cdot \mathrm{v} = |\mathrm{v}|^2$$
$$|\mathrm{u} \cdot \mathrm{v}| \leq |\mathrm{u}|\,|\mathrm{v}| \quad \text{(disuguaglianza di Cauchy-Schwarz)}$$

Proprietà del modulo:
per ogni $\mathrm{u}, \mathrm{v} \in \mathbb{R}^n$, $t \in \mathbb{R}$ si ha:

$$|\mathrm{v}| \geq 0 \text{ e } (|\mathrm{v}| = 0 \Leftrightarrow \mathrm{v} = 0)$$
$$|t\mathrm{v}| = |t| \cdot |\mathrm{v}|$$
$$|\mathrm{u} + \mathrm{v}| \leq |\mathrm{u}| + |\mathrm{v}| \quad \text{(disuguaglianza triangolare)}$$

Proprietà del prodotto vettoriale (in $\mathbb{R}^3$):
per ogni $\mathrm{u}, \mathrm{v}, \mathrm{w} \in \mathbb{R}^n$, $t \in \mathbb{R}$ si ha:

$$\mathrm{v} \times \mathrm{w} = -\mathrm{w} \times \mathrm{v}$$
$$\mathrm{u} \times (\mathrm{v} + \mathrm{w}) = \mathrm{u} \times \mathrm{v} + \mathrm{u} \times \mathrm{w}$$
$$(t\mathrm{v}) \times \mathrm{w} = t(\mathrm{v} \times \mathrm{w})$$
$$\mathrm{v} \times \mathrm{v} = 0$$
$$\mathrm{i} \times \mathrm{j} = \mathrm{k}; \quad \mathrm{j} \times \mathrm{k} = \mathrm{i}; \quad \mathrm{k} \times \mathrm{i} = \mathrm{j}.$$

Il modulo del prodotto vettoriale, $|\mathrm{v} \times \mathrm{w}|$, uguaglia l'area del parallelogramma individuato da v e w.

Proprietà del prodotto misto (in $\mathbb{R}^3$):
per ogni $\mathrm{u}, \mathrm{v}, \mathrm{w} \in \mathbb{R}^n$ si ha:

$$\mathrm{u} \cdot (\mathrm{v} \times \mathrm{w}) = \mathrm{v} \cdot (\mathrm{w} \times \mathrm{u}) = \mathrm{w} \cdot (\mathrm{u} \times \mathrm{v})$$

Il modulo del prodotto misto, $|\mathrm{u} \cdot (\mathrm{v} \times \mathrm{w})|$, uguaglia il volume del parallelepipedo di cui u, v e w sono i tre spigoli.

2 FUNZIONI A VALORI VETTORIALI, LIMITI E CONTINUITÀ

Nel volume 1, capitolo 2, è stata data la definizione generale di *funzione fra due insiemi qualsiasi*: ricordiamo che se A, B sono insiemi, una funzione $f : A \to B$ è una legge che associa a ogni elemento di A uno e un solo elemento di B. Nello studio del calcolo infinitesimale, finora ci siamo concentrati sulle *funzioni reali di variabile reale*, $f : A \subseteq \mathbb{R} \to \mathbb{R}$, per le quali "ingresso" e "uscita" erano numeri reali.

Possiamo ora immaginare situazioni in cui i dati d'ingresso siano più d'uno: a un gruppo di $2, 3, \ldots, n$ dati viene associato *univocamente* un numero reale. Parliamo allora di funzioni di $2, 3, \ldots, n$ variabili.

Per esempio: il volume V occupato da un gas perfetto è funzione della sua pressione P e della sua temperatura T secondo la nota legge

$$V = k\frac{T}{P} \ (k \text{ costante positiva})$$
$$\underset{\text{ingresso}}{(T, P)} \longmapsto \underset{\text{uscita}}{V}$$

Analogamente, ci sono situazioni in cui a un ingresso (numero reale o gruppo di numeri reali) corrisponde univocamente una *coppia* (o *terna*, ecc.) di numeri reali. La funzione avrà dunque valori vettoriali.

Per esempio: la posizione di una particella in movimento è univocamente individuata in ogni istante t; tale posizione si può vedere perciò come una funzione che associa al numero t la terna ordinata delle coordinate del punto:

$$\underset{\text{ingresso}}{t} \longmapsto \underset{\text{uscita}}{(x, y, z)}$$

Si noti il linguaggio usato:

si dice *funzione di una variabile* o *funzione di più variabili* una funzione *definita* in $A \subseteq \mathbb{R}$ o in $A \subseteq \mathbb{R}^n$ $(n > 1)$, rispettivamente;

si dice *funzione a valori reali* o *funzione a valori vettoriali* (o più semplicemente, *funzione reale* o *funzione vettoriale*) una funzione che ha *codominio* $\mathbb{R}$ o $\mathbb{R}^m$ $(m > 1)$, rispettivamente.

Per esempio, una funzione $\mathrm{f} : \mathbb{R} \to \mathbb{R}^3$ è una funzione vettoriale di una variabile; una funzione $f : \mathbb{R}^2 \to \mathbb{R}$ è una funzione reale di due variabili. Le funzioni a valori vettoriali si indicano, come i vettori, in grassetto[2].

Il calcolo infinitesimale per funzioni $\mathrm{f} : A \subseteq \mathbb{R} \to \mathbb{R}^m$ è più facile di quello per funzioni $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$. Per questo motivo partiremo proprio dall'illustrare, in questo capitolo, le idee fondamentali del calcolo differenziale e integrale per le funzioni vettoriali di variabile reale, $\mathrm{f} : A \subseteq \mathbb{R} \to \mathbb{R}^m$. Nei capitoli successivi tratteremo invece, in modo graduale, il calcolo differenziale e integrale per le funzioni $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ e infine per le funzioni $\mathrm{f} : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$.

Come vedremo, nel caso $m = 2, 3$, le funzioni $\mathrm{f} : A \subseteq \mathbb{R} \to \mathbb{R}^m$ hanno il significato geometrico di curve nel piano o nello spazio tridimensionale. Spesso i nostri esempi faranno riferimento a queste situazioni, più intuitive; tuttavia la teoria verrà sviluppata per lo più in dimensione qualunque, come effettivamente è richiesto nelle applicazioni di questi concetti.

Per introdurre le nozioni del calcolo infinitesimale per funzioni $\mathrm{f} : A \subseteq \mathbb{R} \to \mathbb{R}^m$, il primo passo è naturalmente definire e studiare la nozione di *limite* in questo contesto. Consideriamo quindi una funzione

$$\mathrm{r} : I \to \mathbb{R}^m$$

con I intervallo in $\mathbb{R}$; sia $t_0 \in I$ (oppure sia t_0 un estremo di I non appartenente a I) e sia $\mathrm{l} \in \mathbb{R}^m$.

DEFINIZIONE 2.1 Si dice che

$$\lim_{t \to t_0} \mathrm{r}(t) = \mathrm{l}$$

se

$$\lim_{t \to t_0} |\mathrm{r}(t) - \mathrm{l}| = 0. \tag{2.1}$$

[2] Vale in proposito la stessa avvertenza che abbiamo fatto nella nota a pag. 49.

Poiché la funzione $t \mapsto |\mathrm{r}(t) - \mathrm{l}|$ è una funzione reale di variabile reale (il modulo di un vettore è un numero reale), la definizione precedente riconduce la nozione di limite per funzioni a valori vettoriali a quella, già nota, di limite per funzioni a valori reali[3]. Geometricamente, la (2.1) significa che la *distanza* tra il punto $\mathrm{r}(t)$ e il punto l, nello spazio $\mathbb{R}^m$, tende a zero per $t \to t_0$.

È facile vedere che il limite per $t \to t_0$ di $\mathrm{r}(t)$ si calcola semplicemente calcolando i limiti delle *componenti scalari* di $\mathrm{r}(t)$. Vale infatti la seguente:

PROPOSIZIONE 2.1 *Siano:*

$$\mathrm{r}(t) = (r_1(t), r_2(t), \ldots, r_m(t)) \quad \text{con } r_i : I \to \mathbb{R},\ i = 1, 2, \ldots, m$$

e sia

$$\mathrm{l} = (l_1, l_2, \ldots, l_m) \in \mathbb{R}^m.$$

Allora, per $t \to t_0$,

$$\mathrm{r}(t) \to \mathrm{l} \text{ se e solo se } r_i(t) \to l_i \text{ per ogni } i = 1, 2, \ldots, m.$$

In altre parole: il limite della funzione a valori vettoriali si calcola componente per componente; in simboli:

$$\lim_{t \to t_0} (r_1(t), r_2(t), \ldots, r_m(t)) = \left(\lim_{t \to t_0} r_1(t), \lim_{t \to t_0} r_2(t), \ldots, \lim_{t \to t_0} r_m(t) \right). \tag{2.2}$$

DIMOSTRAZIONE. Basta osservare che

$$|r_i(t) - l_i|^2 \le |\mathrm{r}(t) - \mathrm{l}|^2 = \sum_{j=1}^{n} |r_j(t) - l_j|^2. \tag{2.3}$$

Allora se $\mathrm{r}(t) \to \mathrm{l}$, la prima disuguaglianza implica che $r_i(t) \to l_i$ per ogni $i = 1, 2, .., m$ (teorema del confronto); viceversa se $r_j(t) \to l_j$ per ogni $j = 1, 2, \ldots, m$, allora la sommatoria nella (2.3) tende a zero, quindi $\mathrm{r}(t) \to \mathrm{l}$. ❖

Il calcolo dei limiti per funzioni a valori vettoriali non introduce quindi nuove difficoltà rispetto al calcolo dei limiti per funzioni a valori reali.

Inoltre, molte definizioni e proprietà riguardanti i limiti di funzioni vettoriali si possono enunciare e dimostrare in modo perfettamente analogo al caso unidimensionale, per esempio:

- il *teorema di unicità del limite* vale per funzioni $\mathrm{r} : I \subseteq \mathbb{R} \to \mathbb{R}^m$;
- il *teorema sul limite della somma* o del *prodotto per una costante* vale per funzioni $\mathrm{r} : I \subseteq \mathbb{R} \to \mathbb{R}^m$;
- la *definizione di funzione continua*, in un punto o in un insieme, è analoga al caso unidimensionale: si dice che $\mathrm{r} : I \subseteq \mathbb{R} \to \mathbb{R}^m$ è continua in t_0 se $\lim_{t \to t_0} \mathrm{r}(t) = \mathrm{r}(t_0)$. Per la (2.2), si ha che *una funzione a valori vettoriali è continua se e solo se lo sono tutte le sue componenti.*

[3] v. volume 1, capitolo 3, paragrafo 2.

3 CURVE REGOLARI E CALCOLO DIFFERENZIALE VETTORIALE

3.1 Esempi introduttivi

Come già anticipato nel paragrafo 2, nel caso $m = 2$ o 3 le funzioni $\mathrm{r} : I \subseteq \mathbb{R} \to \mathbb{R}^m$ hanno un significato geometrico e fisico notevole, che illustriamo ora meglio con alcuni esempi. Questi ci condurranno a una definizione precisa di *curva*.

ESEMPIO Consideriamo un punto materiale che si muove nello spazio tridimensionale lungo una certa traiettoria. Le sue coordinate (x, y, z) saranno funzione del tempo t, perciò la sua posizione è specificata da una funzione del tipo:

$$\mathrm{r} : [t_1, t_2] \to \mathbb{R}^3$$

con $\mathrm{r} = (x, y, z)$ e

$$\begin{cases} x = x(t) \\ y = y(t) \\ z = z(t) \end{cases} \qquad t \in [t_1, t_2].$$

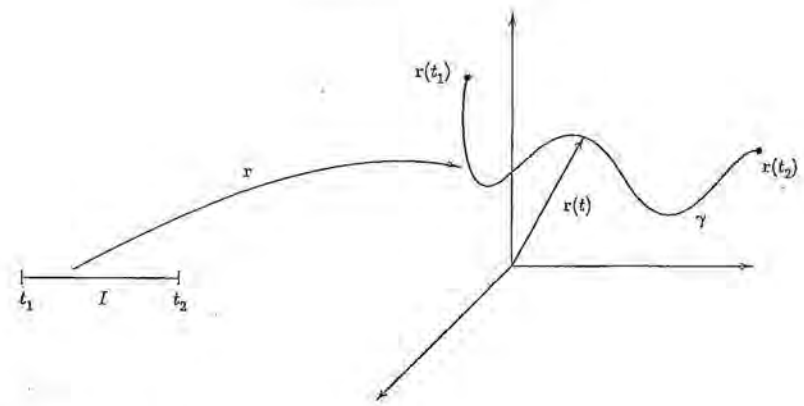


Figura 2.1. Curva in $\mathbb{R}^3$.

Si può usare anche la notazione vettoriale:

$$\mathrm{r}(t) = x(t)\,\mathrm{i} + y(t)\,\mathrm{j} + z(t)\,\mathrm{k}$$

o anche

$$\mathrm{r}(t) = (x(t), y(t), z(t)).$$

Se il punto si muovesse nel piano, per rappresentarne il moto basterebbe una funzione:

$$\mathrm{r} : [t_1, t_2] \to \mathbb{R}^2$$

con $\mathrm{r} = (x, y)$ e

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases} \qquad t \in [t_1, t_2].$$

 © 978-88-08-06485-1

3.2 Arco di curva continua

Possiamo ora dare una definizione generale di curva, precisando le idee descritte dagli esempi precedenti con qualche richiesta di regolarità.

DEFINIZIONE 2.2 Sia I un intervallo in $\mathbb{R}$. Si dice *arco di curva continua*, o cammino, in $\mathbb{R}^m$ una funzione $\mathrm{r} : I \to \mathbb{R}^m$ continua (ovvero tale che le sue componenti sono funzioni continue).

Se la variabile t si pensa come tempo, un arco di curva è la legge oraria di un punto mobile: assegna la traiettoria e il punto in cui si trova in ogni istante il punto mobile.

Il *sostegno della curva* è l'immagine della funzione, cioè l'insieme dei punti di $\mathbb{R}^m$ percorsi dal punto mobile (ovvero la linea geometrica, a prescindere dalla legge con cui è percorsa).

La curva si dice *chiusa* se $\mathrm{r}(a) = \mathrm{r}(b)$ con $I = [a, b]$ (il punto di partenza e di arrivo nel moto del punto coincidono); si dice *semplice* se non ripassa mai dallo stesso punto cioè se $t_1 \neq t_2 \Rightarrow \mathrm{r}(t_1) \neq \mathrm{r}(t_2)$, con l'unica possibile eccezione $(t_1, t_2) = (a, b)$ (in altre parole una curva semplice può essere chiusa, ma non deve "intrecciarsi").

Una curva si dice *piana* se esiste un piano che contiene il suo sostegno.

Più rigorosamente, si potrebbe dire che un arco di curva continuo γ è la *coppia* costituita da una funzione (continua) $\mathrm{r} : I \to \mathbb{R}^m$, che chiamiamo *parametrizzazione della curva*, e un insieme di punti di $\mathbb{R}^m$ (l'immagine di r), che chiamiamo *sostegno della curva*. Questa precisazione fa sì che, per esempio, si possa parlare di "due parametrizzazioni diverse della stessa curva", per intendere due diverse funzioni $\mathrm{r} : I \to \mathbb{R}^m$ che però individuano lo stesso sostegno. Per non appesantire il discorso, tuttavia, useremo questa terminologia con molta flessibilità.

Esempi

L'arco di *ellisse*

$$\begin{cases} x = 2\cos t \\ y = 3\sin t \end{cases} \qquad t \in [0, \pi]$$

è un arco di curva continua, non chiusa, semplice. Se t variasse in $[0, 2\pi]$ avremmo una curva chiusa, semplice, il cui sostegno è l'intera ellisse (fig. 2.3.*b*).

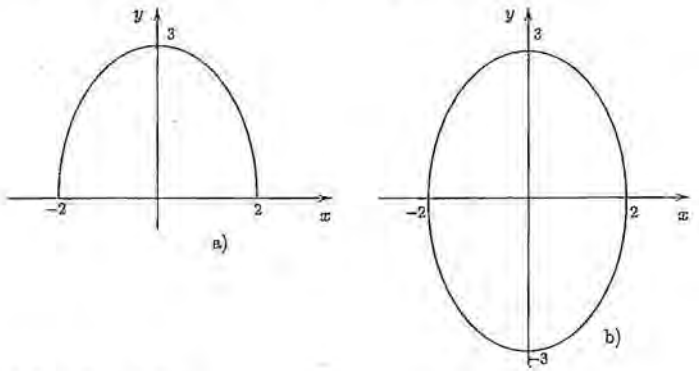


Figura 2.3. Semiellisse ed ellisse.

Il *Folium* di Cartesio

$$\begin{cases} x = t(t-1) \\ y = t(t-1)(2t-1) \end{cases} \qquad t \in (-\infty, +\infty)$$

è una curva continua, non chiusa (fig. 2.4*a*), non semplice. Se considerassimo solo l'arco per $t \in [0, 1]$ avremmo una curva chiusa e semplice (fig. 2.4*b*): infatti si vede che per $t = 0$ e per $t = 1$ passa dallo stesso punto $(0, 0)$.

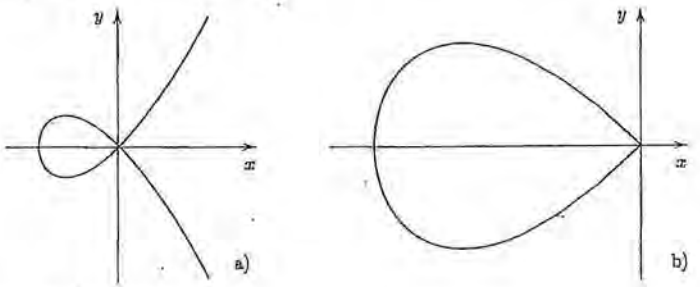


Figura 2.4. *Folium* di Cartesio.

Le equazioni

$$\begin{cases} x = a\mathrm{Cht} \\ y = b\mathrm{Sht} \end{cases} \qquad t \in \mathbb{R}$$

con $a, b > 0$ descrivono il *ramo di iperbole*

$$\frac{x^2}{a^2} - \frac{y^2}{b^2} = 1 \qquad x > 0$$

(da cui l'origine del nome di *funzioni iperboliche* per Cht, Sht). L'altro ramo (per $x < 0$) è descritto da

$$\begin{cases} x = -a\mathrm{Cht} \\ y = b\mathrm{Sht} \end{cases} \qquad t \in \mathbb{R}.$$

Ognuno dei due è un arco di curva continuo, non chiuso (fig. 2.5).

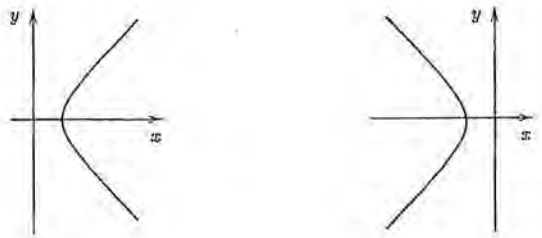


Grafico di $\begin{cases} x = \mathrm{Ch}\, t \\ y = \mathrm{Sh}\, t \end{cases}$ $t \in \mathbb{R}$ Grafico di $\begin{cases} x = -\mathrm{Ch}\, t \\ y = \mathrm{Sh}\, t \end{cases}$ $t \in \mathbb{R}$

Figura 2.5.

Gli esempi precedenti sono curve piane, cioè in $\mathbb{R}^2$. Un esempio di curva in $\mathbb{R}^3$ è l'*elica cilindrica*, già incontrata nel paragrafo 3.1 (Esempio 3.4):

$$\begin{cases} x = R\cos t \\ y = R\sin t \\ z = pt \end{cases} \qquad t \in [a, b]$$

con a, b, R, p parametri fissati. Si chiama *cilindrica* perché giace sul cilindro di equazione $x^2 + y^2 = R^2$. È un arco di curva continua, non chiusa.

La curva

$$\begin{cases} x = t \\ y = 2t \\ z = t^2 \end{cases} \qquad t \in [0, 1]$$

è una curva piana in $\mathbb{R}^3$; infatti giace sul piano $y = 2x$.

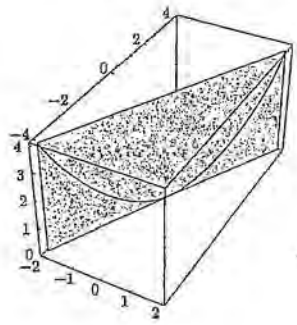


Figura 2.6. Curva piana.

Su un arco di curva parametrizzato esiste un *orientamento* o *orientazione*, cioè un verso di percorrenza, dettato dalla parametrizzazione. Per esempio, l'ellisse nell'Esempio 3.7 è percorsa in senso antiorario (per t crescente, il punto mobile gira in tal senso). La curva di ugual sostegno percorsa in senso opposto (orario) è parametrizzata da

$$\begin{cases} x = 2\cos(-t) = 2\cos t \\ y = 3\sin(-t) = -3\sin t \end{cases} \qquad t \in [0, 2\pi].$$

3.3 Derivata di una funzione vettoriale. Arco di curva regolare

Vogliamo ora generalizzare il concetto di *derivata* a una funzione $\mathbf{r} : I \to \mathbb{R}^m$ (con I intervallo in $\mathbb{R}$). La definizione usuale, come limite del rapporto incrementale, ha perfettamente senso:

DEFINIZIONE 2.3 Sia $\mathbf{r} : I \to \mathbb{R}^m$ e $t_0 \in I$; si dice che $\mathbf{r}$ è derivabile in t_0 se esiste finito

$$\mathbf{r}'(t_0) = \lim_{h \to 0} \frac{\mathbf{r}(t_0 + h) - \mathbf{r}(t_0)}{h}. \tag{3.1}$$

Se $\mathbf{r}$ è derivabile in tutto I e inoltre la funzione $\mathbf{r}'$ è continua in I, si dice che $\mathbf{r}$ è *di classe* $C^1(I)$, e si scrive $\mathbf{r} \in C^1(I)$. Analogamente si definiscono le derivate di ordine successivo e le funzioni di classe $C^k(I)$ per $k > 1$.

Si noti che il rapporto incrementale è il quoziente tra un vettore e uno scalare (il che ha senso), quindi è un vettore; il limite va inteso rispetto alla distanza in $\mathbb{R}^m$.

Ricordando poi che i limiti di funzioni a valori vettoriali si fanno componente per componente, vediamo che

$$\mathbf{r}'(t_0) = \left(\lim_{h \to 0} \frac{r_1(t_0 + h) - r_1(t_0)}{h}, \ \ldots, \ \lim_{h \to 0} \frac{r_m(t_0 + h) - r_m(t_0)}{h} \right) = \tag{3.2}$$
$$= (r_1'(t_0), r_2'(t_0), \ \ldots, \ r_m'(t_0)).$$

Ossia, anche la derivata di una funzione a valori vettoriali si fa componente per componente: il *vettore derivato è il vettore delle derivate delle componenti*.

Esempio

Per l'ellisse $\mathbf{r}(t) = (2\cos t, 3\sin t)$, abbiamo:

$$\mathbf{r}'(t) = (-2\sin t, 3\cos t)$$

Mentre la (3.2) è utile per il *calcolo effettivo* del vettore derivato, per capire il *significato cinematico* del vettore derivato è meglio ragionare sulla (3.1), cioè sul limite del rapporto incrementale dei vettori, anziché sulle singole componenti. Il vettore $\mathbf{r}(t_0 + h) - \mathbf{r}(t_0)$ rappresenta lo spostamento del punto mobile dall'istante t_0 all'istante $t_0 + h$. Ne segue che il limite del rapporto incrementale rappresenta cinematicamente il *vettore velocità istantanea* del moto, che risulta tangente alla curva (fig. 2.7).

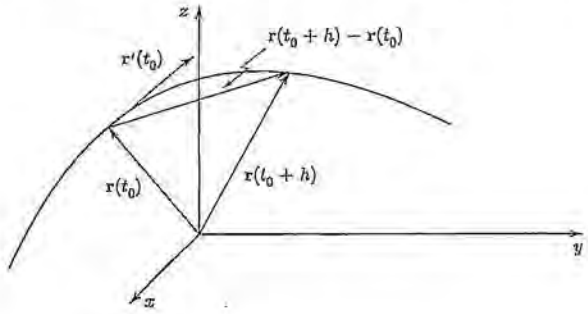


Figura 2.7. Vettore velocità.

Analogamente, la derivata seconda $\mathbf{r}''(t)$, se esiste, avrà il significato di *vettore accelerazione istantanea* del moto.

Si può definire anche la *velocità scalare*:

$$v(t) = |\mathbf{r}'(t)|.$$

 © 978-88-08-06485-1

Il vettore velocità istantanea individua effettivamente una direzione (e quindi ha il significato di vettore tangente) solo quando le sue componenti non sono tutte nulle. Questo motiva la prossima definizione:

Definizione 2.4 Sia $I \subseteq \mathbb{R}$ un intervallo. Si dice *arco di curva regolare* un arco di curva $\mathrm{r} : I \to \mathbb{R}^m$ tale che $\mathrm{r} \in C^1(I)$ e $\mathrm{r}'(t) \neq 0$ per ogni $t \in I$.

In altre parole: il vettore derivato, o vettore velocità, esiste in ogni punto, varia con continuità e non si annulla mai. Di conseguenza, per le curve regolari è ben definito il *versore tangente*,

$$\mathrm{T} = \frac{\mathrm{r}'(t)}{|\mathrm{r}'(t)|},$$

che inoltre dipende con continuità da t.

Esempi

L'ellisse $\mathrm{r}(t) = (2\cos t, 3\sin t)$, $t \in [0, 2\pi]$ è una curva regolare, infatti il vettore derivato $r'(t) = (-2\sin t, 3\cos t) \neq (0,0)$ per ogni t (si noti che quando una componente si annulla, l'altra è diversa da zero). Il versore tangente è:

$$\mathrm{T} = \frac{(-2\sin t, 3\cos t)}{\sqrt{4(\sin t)^2 + 9(\cos t)^2}}$$

L'astroide

$$\begin{cases} x = (\cos\theta)^3 \\ y = (\sin\theta)^3 \end{cases} \qquad \theta \in [0, 2\pi]$$

non è una curva regolare in quanto $\mathrm{r}'(\theta) = (-3(\cos\theta)^2 \sin\theta, 3(\sin\theta)^2 \cos\theta)$ si annulla per $\theta = k\frac{\pi}{2}$ $(k = 0, 1, 2, 3)$. Infatti la curva in figura 2.8 presenta delle cuspidi nei punti corrispondenti:

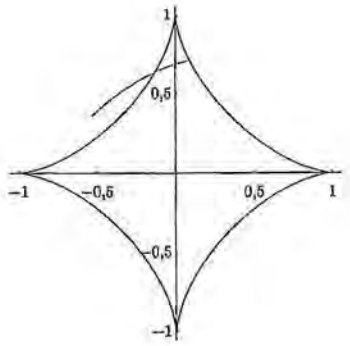


Figura 2.8. Astroide.

L'ultimo esempio visto presenta una situazione abbastanza tipica: la curva non è regolare, tuttavia le condizioni di regolarità sono violate in un numero finito di punti.

Questa situazione merita una definizione:

Definizione 2.5 Si dice *arco di curva regolare a tratti* un arco di curva $\mathrm{r} : I \to \mathbb{R}^m$ (con I intervallo in $\mathbb{R}$) tale che: r è continua e l'intervallo I può essere suddiviso in un numero finito di sottointervalli, su ciascuno dei quali r è un arco di curva regolare.

L'astroide dell'Esempio 3.14 è un arco di curva regolare a tratti, anche se non è regolare; si può dire anche che l'astroide è unione di quattro archi di curva regolare, pur di dare, però, una definizione precisa di unione di due archi di curve:

Definizione 2.6 Siano γ_1, γ_2 due curve in $\mathbb{R}^m$, di equazione, rispettivamente, $\mathrm{r}_1 = \mathrm{r}_1(t)$ per $t \in [a, b]$, $\mathrm{r}_2 = \mathrm{r}_2(t)$ per $t \in [b, c]$, soddisfacenti la condizione di raccordo $\mathrm{r}_1(b) = \mathrm{r}_2(b)$. Allora si definisce *unione di* γ_1 *e* γ_2, e si indica col simbolo $\gamma_1 \cup \gamma_2$, la curva $\mathrm{r} : [a, c] \to \mathbb{R}^m$ definita da:

$$\mathrm{r}(t) = \begin{cases} \mathrm{r}_1(t) & \text{per } t \in [a, b] \\ \mathrm{r}_2(t) & \text{per } t \in [b, c]. \end{cases}$$

Notiamo che una curva regolare a tratti è unione di un numero finito di curve regolari, nel senso di questa definizione. Viceversa, l'unione di due curve regolari è in generale una curva regolare a tratti; è regolare se risulta anche $\mathrm{r}_1'(b) = \mathrm{r}_2'(b)$.

Il prossimo teorema riassume le principali regole del *calcolo differenziale vettoriale*. Si noti che queste regole si applicano a funzioni derivabili qualsiasi, a valori vettoriali, a prescindere dal fatto che rappresentino curve regolari (in altre parole, la condizione di non annullamento del vettore derivato qui non ha alcuna importanza):

Teorema 2.1 *Sia $I \subseteq \mathbb{R}$. Se $\mathrm{u}, \mathrm{v} : I \to \mathbb{R}^m$ sono derivabili, allora:*

1. $(\mathrm{u} + \mathrm{v})' = \mathrm{u}' + \mathrm{v}'$;
2. *se $c \in \mathbb{R}$ è una costante,* $(c\mathrm{u})' = c\mathrm{u}'$;
3. *se $f : I \to \mathbb{R}$ è una funzione derivabile,* $(f\mathrm{u})' = f'\mathrm{u} + f\mathrm{u}'$;
4. *se $\varphi : \mathbb{R} \to \mathbb{R}$ è una funzione derivabile,* $[\mathrm{u}(\varphi(t))]' = \mathrm{u}'(\varphi(t))\varphi'(t)$ *(purché*[5] $\operatorname{Im}\varphi \subseteq I$*);*
5. $(\mathrm{u} \cdot \mathrm{v})' = \mathrm{u}' \cdot \mathrm{v} + \mathrm{u} \cdot \mathrm{v}'$, *dove $\cdot$ indica il* prodotto scalare *delle due funzioni vettoriali;*
6. *Se $m = 3$,* $(\mathrm{u} \times \mathrm{v})' = \mathrm{u}' \times \mathrm{v} + \mathrm{u} \times \mathrm{v}'$, *dove $\times$ indica il* prodotto vettoriale *in $\mathbb{R}^3$.*

Tutti gli enunciati vanno intesi nel senso seguente: se u, v sono derivabili, allora la derivata scritta a primo membro di ciascuna identità esiste *ed è uguale al secondo membro.*

[5] Per il significato di Im, v. volume 1, capitolo 2, paragrafo 1.

 © 978-88-08-06485-1

DIMOSTRAZIONE. La dimostrazione di queste proprietà è un semplice esercizio: è sufficiente utilizzare la definizione (3.1) di derivata di una funzione $r : \mathbb{R} \to \mathbb{R}^m$, la proprietà (3.2), le regole del calcolo differenziale in una variabile, e, per le 5 e 6, le definizioni di prodotto scalare e vettoriale. Dimostriamo, a titolo d'esempio, la 5:

$$(\mathrm{u}\cdot\mathrm{v})' = \left(\sum_{j=1}^{m} u_j v_j\right)' = \sum_{j=1}^{m} (u_j v_j)' =$$
$$= \sum_{j=1}^{m} \left(u_j' v_j + u_j v_j'\right) = \sum_{j=1}^{m} u_j' v_j + \sum_{j=1}^{m} u_j v_j' = \mathrm{u}'\cdot\mathrm{v} + \mathrm{u}\cdot\mathrm{v}'.$$

3.4 Integrale di una funzione a valori vettoriali

Si può definire in maniera naturale anche l'*integrale definito di una funzione a valori vettoriali*:

DEFINIZIONE 2.7 Se $\mathrm{r} : [a,b] \to \mathbb{R}^m$, poniamo

$$\int_a^b \mathrm{r}(t)\,dt = \left(\int_a^b r_1(t)\,dt, \int_a^b r_2(t)\,dt, \ldots, \int_a^b r_m(t)\,dt\right). \tag{3.3}$$

Più precisamente, diremo che r è integrabile in $[a,b]$ se lo è ogni sua componente r_i $(i = 1,2,\ldots,m)$ e in tal caso definiremo l'integrale di r mediante la (3.3).

Per esempio, se $\mathrm{r} : [a,b] \to \mathbb{R}^m$ è un arco di curva *continua*, certamente è integrabile.

Ragionando componente per componente è immediato dimostrare il

TEOREMA 2.2 (TEOREMA FONDAMENTALE DEL CALCOLO INTEGRALE PER FUNZIONI A VALORI VETTORIALI) *Se* $\mathrm{r} : [a,b] \to \mathbb{R}^m$ *è di classe* $C^1([a,b])$, *allora*

$$\int_a^b \mathrm{r}'(t)\,dt = \mathrm{r}(b) - \mathrm{r}(a).$$

In particolare, se $\mathrm{r} : [a,b] \to \mathbb{R}^m$ *è una curva regolare chiusa,* $\int_a^b \mathrm{r}'(t)\,dt = 0$.

È utile talvolta la seguente disuguaglianza dei moduli, analoga a quella che vale per funzioni a valori reali, ma meno immediata da dimostrare:

LEMMA 2.3 *Se* $\mathrm{r} : [a,b] \to \mathbb{R}^m$ *è integrabile, allora*

$$\left|\int_a^b \mathrm{r}(t)\,dt\right| \le \int_a^b |\mathrm{r}(t)|\,dt.$$

Si noti che l'integrale che compare a primo membro è l'integrale di una funzione a valori vettoriali, mentre quello a secondo membro coinvolge la funzione a valori reali $|\mathrm{r}(t)|$; i moduli che compaiono nella formula sono entrambi moduli di vettori (e non valori assoluti di numeri reali).

© 978-88-08-06485-1

DIMOSTRAZIONE. Entrambi gli integrali esistono, sotto le nostre ipotesi; quello a secondo membro si può calcolare come limite di somme di Cauchy-Riemann, nel modo seguente:

$$\int_a^b |\mathrm{r}(t)|\,dt = \lim_{n\to\infty} s_n = \lim_{n\to\infty} \sum_{j=1}^{n} \left|\mathrm{r}\left(t_j^*\right)\right| (t_j - t_{j-1}) \tag{3.4}$$

con $t_j^* \in [t_{j-1}, t_j]$ e il solito significato di questi simboli[6]. Consideriamo allora le successioni:

$$s_n^{(i)} = \sum_{j=1}^{n} r_i\left(t_j^*\right)(t_j - t_{j-1}), \text{ per } i = 1,2,\ldots,m$$

dove $(r_1, r_2, \ldots, r_m)$ sono le componenti di r. Si ha:

$$s_n^{(i)} \to \int_a^b r_i(t)\,dt \text{ per } n \to \infty, i = 1,2,\ldots,m.$$

Possiamo anche scrivere:

$$\left(s_n^{(1)}, s_n^{(2)}, \ldots, s_n^{(m)}\right) = \sum_{j=1}^{n} \mathrm{r}\left(t_j^*\right)(t_j - t_{j-1})$$

e quindi, per la disuguaglianza triangolare in $\mathbb{R}^m$,

$$\left|\left(s_n^{(1)}, s_n^{(2)}, \ldots, s_n^{(m)}\right)\right| \le \sum_{j=1}^{n} \left|\mathrm{r}\left(t_j^*\right)\right| (t_j - t_{j-1}) = s_n.$$

Passando al limite per $n \to \infty$ nella precedente disuguglianza si ha allora, per il teorema di permanenza del segno e la (3.4),

$$\left|\left(\int_a^b r_1(t)\,dt, \int_a^b r_2(t)\,dt, \ldots, \int_a^b r_m(t)\,dt\right)\right| \le \int_a^b |\mathrm{r}(t)|\,dt,$$

che è la tesi.

3.5 Alcune classi di curve piane

Curve piane, grafico di funzioni

Una classe particolare di curve piane è quella delle curve che si ottengono come grafici di funzioni di una variabile, ossia

$$y = f(x) \text{ per } x \in [a,b], \tag{3.5}$$

che si può pensare scritta in forma parametrica ponendo:

$$\begin{cases} x = t \\ y = f(t) \end{cases} \text{per } t \in [a,b].$$

Una curva espressa come grafico di funzione (3.5) ha le seguenti proprietà:

[6] v. volume 1, capitolo 6, paragrafo 2.

 © 978-88-08-06485-1

- è continua se e solo se f è continua in $[a,b]$;
- è regolare se e solo se f è derivabile con continuità in $[a,b]$ (notiamo che la condizione di non annullamento del vettore velocità è automatica perché $x'(t) \equiv 1$);
- è regolare a tratti se e solo se f è continua in $[a,b]$ e a tratti derivabile con continuità in $[a,b]$;
- non è *mai* chiusa;
- è *sempre* semplice.

Invitiamo il lettore a verificare queste semplici affermazioni.

Curve piane in forma polare

Una forma particolare che possono avere le equazioni parametriche di una *curva piana* è quella *polare*. L'equazione (singola):

$$\rho = f(\theta) \text{ per } \theta \in [\theta_1, \theta_2] \tag{3.6}$$

è una stenografia che sta a indicare (ponendo $x = \rho\cos\theta$, $y = \rho\sin\theta$)

$$\begin{cases} x = f(\theta)\cos\theta \\ y = f(\theta)\sin\theta \end{cases} \quad \theta \in [\theta_1, \theta_2].$$

Osserviamo che:

$$\mathbf{r}'(\theta) = (f'(\theta)\cos\theta - f(\theta)\sin\theta, f'(\theta)\sin\theta + f(\theta)\cos\theta);$$
$$|\mathbf{r}'(\theta)| = \sqrt{f'(\theta)^2 + f(\theta)^2}.$$

Una curva piana espressa in forma polare (3.6) ha le seguenti proprietà:

- è continua se e solo se f è continua in $[\theta_1, \theta_2]$;
- è regolare se e solo se f è derivabile con continuità in $[\theta_1, \theta_2]$ e inoltre f e f' non si annullano contemporaneamente;
- è chiusa se e solo se $f(\theta_1) = f(\theta_2)$ e $\theta_2 - \theta_1 = 2n\pi$ per qualche intero n.

Esempio

3.10 La *spirale di Archimede* (fig. 2.9) ha equazione polare:

$$\rho = A\theta, \ \theta \in [0, +\infty)$$

(A costante positiva fissata). Questa equazione equivale alle equazioni parametriche:

$$\begin{cases} x = A\theta\cos\theta \\ y = A\theta\sin\theta \end{cases} \quad \theta \in [0, +\infty)$$

Il vettore tangente è $\mathbf{r}'(\theta) = A(\cos\theta - \theta\sin\theta, \sin\theta + \theta\cos\theta)$ e non si annulla mai, come si vede calcolando il modulo del vettore:

$$v(\theta) = |\mathbf{r}'(\theta)| = A\sqrt{1+\theta^2} > 0 \quad \forall\theta$$

Perciò la curva è regolare.

© 978-88-08-06485-1

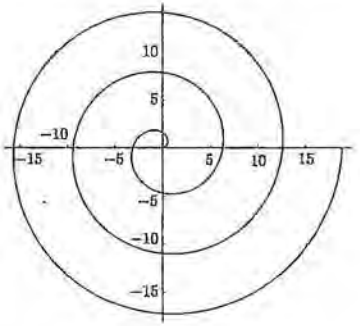


Figura 2.9. Spirale di Archimede.

Coniche in forma polare

Anche le coniche (parabole, ellissi, iperboli), di cui già abbiamo descritto le equazioni parametriche (esempi 3.7, 3.9), ammettono un'equazione in forma polare. Per arrivarci, ricordiamo uno dei modi in cui è possibile definire queste curve come opportuni *luoghi geometrici*. Nel piano x, y, sia d la retta verticale di equazione $x = -p$, e sia $\varepsilon > 0$ un numero fissato. Il luogo dei punti $P = (x, y)$ che soddisfano la condizione

$$\frac{\text{distanza di } P \text{ dall'origine}}{\text{distanza di } P \text{ dalla retta } d} = \varepsilon$$

si dice *conica di eccentricità* ε, con *direttrice* d e *fuoco l'origine*. Poiché in coordinate polari è (fig. 2.10)

$$(\text{distanza di } P \text{ dall'origine}) = \rho$$

e

$$(\text{distanza di } P \text{ dalla retta } d) = p + \rho\cos\theta,$$

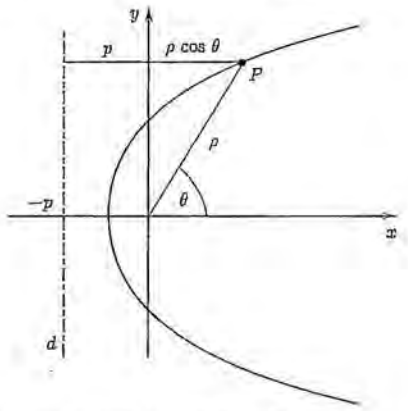


Figura 2.10. Definizione di conica mediante fuoco e direttrice.

 © 978-88-08-06485-1

risulta

$$\frac{p}{p+\rho\cos\theta}=\varepsilon$$

da cui si ricava l'*equazione polare della conica*:

$$\rho=\frac{\varepsilon p}{1-\varepsilon\cos\theta}. \tag{3.7}$$

Lo studente è invitato a verificare che la (3.7) è l'equazione di:

$$\begin{cases}\text{un'ellisse} & \text{se } \varepsilon<1\\ \text{una parabola} & \text{se } \varepsilon=1\\ \text{un'iperbole} & \text{se } \varepsilon>1\end{cases} \tag{3.8}$$

Per esempio, nel caso $\varepsilon>1$, gli asintoti dell'iperbole sono le rette passanti per l'origine con pendenza θ tale da annullare il denominatore $(1-\varepsilon\cos\theta)$.

Per $\varepsilon\to 0$ e $p=R/\varepsilon$ si trova l'equazione della circonferenza

$$\rho=R.$$

Invece, per $\varepsilon\to+\infty$, si trova

$$\rho=-\frac{p}{\cos\theta}$$

che è l'equazione di una retta. (Lo studente verifichi il perché).

Notiamo che nella (3.7) il segno $-$ davanti a $\cos\theta$ non è fondamentale: il cambiamento di parametro $\theta=t+\pi$ trasforma l'equazione (3.7) in

$$\rho=\frac{\varepsilon p}{1+\varepsilon\cos t}; \tag{3.9}$$

se il parametro θ varia per esempio nel periodo $[0,2\pi]$, la curva riparametrizzata sarà definita per $t\in[-\pi,\pi]$ e avrà il medesimo sostegno. Dunque anche la (3.9) descrive una conica.

Nello studio del moto di un pianeta intorno al sole, si arriva a dimostrare, risolvendo le opportune equazioni differenziali, che il corpo si muove lungo una curva di equazione polare di tipo (3.9), da cui si riconosce che le orbite sono curve coniche (nel caso dei pianeti, ellissi). Vedremo nei Complementi in fondo al capitolo come si può effettuare questa deduzione, utilizzando le tecniche di calcolo differenziale vettoriale apprese in questo capitolo.

4 LUNGHEZZA DI UN ARCO DI CURVA

4.1 Curve rettificabili e lunghezza

Un problema molto naturale di calcolo infinitesimale è quello di definire che cosa si intenda per *lunghezza di un arco di curva* e calcolarla. Abbiamo già incontrato un caso particolare di questo problema, nello studio delle funzioni reali di variabile reale, quando abbiamo definito la lunghezza del grafico di una funzione $f:[a,b]\to\mathbb{R}$ (v. volume 1, capitolo 6, paragrafo 6). Ora affronteremo il problema analogo per curve in $\mathbb{R}^m$. Il contesto più generale suggerisce di usare un approccio un po' diverso.

© 978-88-08-06485-1

Sia $\mathrm{r}:[a,b]\to\mathbb{R}^m$ la parametrizzazione di un arco di curva γ continuo e consideriamo una *partizione*

$$\mathcal{P}=\{t_0=a,t_1,\dots,t_{n-1},t_n=b\}$$

dell'intervallo $[a,b]$. Questo significa che $t_0<t_1<\ldots<t_n$, ma i punti non sono necessariamente equispaziati. A $\mathcal{P}$ risulta associata la poligonale "inscritta" in γ costituita dagli n segmenti di estremi $\mathrm{r}(t_{j-1})$, $\mathrm{r}(t_j)$, $j=1,\dots,n$. Indichiamo con $l(\mathcal{P})$ la lunghezza di questa poligonale. Evidentemente si ha

$$l(\mathcal{P})=\sum_{j=1}^{n}|\mathrm{r}(t_j)-\mathrm{r}(t_{j-1})|$$

e l'idea è che $l(\mathcal{P})$ approssimi *per difetto* la lunghezza di γ. Infatti, anche se non abbiamo ancora dato una definizione rigorosa di lunghezza, vogliamo che questa definizione rispetti la nostra idea intuitiva che *il segmento è la più breve linea che congiunge due punti*. Alla definizione di lunghezza di γ si perviene dunque facendo variare in tutti i modi possibili la partizione $\mathcal{P}$ e considerando il "più grande" dei numeri $l(\mathcal{P})$ così ottenuti. Naturalmente non è affatto detto che questo numero "più grande" esista. Precisamente, abbiamo:

Definizione 2.8 Si dice che γ è *rettificabile* se

$$\sup_{\mathcal{P}} l(\mathcal{P})=l(\gamma)<+\infty$$

dove l'estremo superiore è calcolato al variare di tutte le possibili partizioni $\mathcal{P}$ di $[a,b]$. In tal caso $l(\gamma)$ assegna, per definizione, la *lunghezza di* γ.

Vorremmo ora poter calcolare effettivamente tale lunghezza, almeno per le curve regolari, con gli strumenti del calcolo integrale e differenziale, come già visto nel caso dei grafici di funzioni. Per arrivare a capire quale formula ci dobbiamo aspettare, ragioniamo "cinematicamente". Sappiamo che r' rappresenta il vettore velocità lungo la curva e che quindi $v(t)=|\mathrm{r}'(t)|$ è la velocità scalare. D'altra parte, la lunghezza di una curva rappresenta lo spazio percorso nell'intervallo di tempo $[a,b]$. Ricordando la relazione

$$\text{spazio} = \text{velocità} \times \text{tempo}$$

siamo condotti alla formula:

$$l=\int_a^b v(t)\,dt.$$

Questa è proprio la formula che assegna la lunghezza di una curva regolare, come afferma il seguente teorema.

Teorema 2.4 *Sia* $\mathrm{r}:[a,b]\to\mathbb{R}^m$ *la parametrizzazione di un arco di curva* γ *regolare. Allora* γ *è rettificabile e*

$$l(\gamma)=\int_a^b|\mathrm{r}'(t)|\,dt. \tag{4.1}$$

La dimostrazione di questo teorema utilizza il teorema di Cantor-Heine sull'uniforme continuità[7] e sarà fornita nei Complementi, alla fine di questo capitolo.

[7] v. volume 1, capitolo 6, Complementi, teorema 6.14.

 © 978-88-08-06485-1

Lunghezza di un grafico

Osserviamo che nel caso particolare di una curva piana regolare che sia grafico di una funzione, ossia

$$\gamma : \begin{cases} x = t \\ y = f(t) \end{cases} \quad \text{per } t \in [a, b]$$

la formula (4.1) dà:

$$l(\gamma) = \int_a^b \sqrt{1 + f'(t)^2}\, dt,$$

come già noto dal calcolo differenziale in una variabile[8].

Il prossimo risultato è intuitivo e utile. Lo enunciamo senza dimostrazione:

PROPOSIZIONE 2.2 *Se una curva γ è unione di due curve rettificabili γ_1 e γ_2, allora γ è rettificabile e $l(\gamma) = l(\gamma_1) + l(\gamma_2)$.*

La stessa proprietà si estende all'unione di un numero finito qualsiasi di curve rettificabili. In particolare:

se una curva γ è regolare a tratti, allora è rettificabile, e la sua lunghezza si calcola ancora mediante la formula (4.1) (dove ora la funzione integranda sarà continua a tratti).

Vediamo ora qualche esempio di calcolo di lunghezze per curve regolari nel piano o nello spazio. Nello spazio tridimensionale la (4.1) prende la seguente forma:

$$l(\gamma) = \int_a^b \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2}\, dt$$

(per una curva piana manca il termine $z'(t)$).

Esempi

Per un arco di circonferenza

$$\begin{cases} x = R\cos\theta \\ y = R\sin\theta \end{cases} \qquad \theta \in [\theta_1, \theta_2]$$

è $\mathrm{r}'(\theta) = (-R\sin\theta, R\cos\theta)$, $|\mathrm{r}'(\theta)| = R$ e

$$l = \int_{\theta_1}^{\theta_2} R\, d\theta = R(\theta_2 - \theta_1).$$

Per l'arco di elica cilindrica (vedi Esempio 3.4)

$$\begin{cases} x = R\cos t \\ y = R\sin t \\ z = pt \end{cases} \qquad t \in [0, 2\pi]$$

si ha:

$$\mathrm{r}'(t) = (-R\sin t, R\cos t, p);$$

$$|\mathrm{r}'(t)| = \sqrt{x'(t)^2 + y'(t)^2 + z'(t)^2} = \sqrt{R^2 + p^2}$$

$$l = \int_0^{2\pi} \sqrt{R^2 + p^2}\, dt = 2\pi\sqrt{R^2 + p^2}$$

Per una curva piana in forma polare

$$\rho = f(\theta),$$

ovvero $\mathrm{r}(\theta) = (f(\theta)\cos\theta, f(\theta)\sin\theta)$, si ha

$$\mathrm{r}'(\theta) = (f'(\theta)\cos\theta - f(\theta)\sin\theta, f'(\theta)\sin\theta + f(\theta)\cos\theta)$$

dunque

$$|\mathrm{r}'(\theta)| = \sqrt{f(\theta)^2 + f'(\theta)^2}.$$

Per esempio, la lunghezza dell'arco di spirale di Archimede (vedi Esempio 3.15) con $A = 1$ e $\rho = \theta$ per $\theta \in [0, 2\pi]$ è

$$l = \int_0^{2\pi} \sqrt{1 + \theta^2}\, d\theta = [\theta = \mathrm{Sh}t;\ d\theta = \mathrm{Ch}t\, dt] = \int_0^{\mathrm{SettSh}2\pi} (\mathrm{Ch}t)^2\, dt =$$
$$= \frac{1}{2}[\mathrm{Sh}t\,\mathrm{Ch}t + t]_0^{\mathrm{SettSh}2\pi} = \frac{1}{2}(2\pi\sqrt{1 + 4\pi^2} + \log(2\pi + \sqrt{1 + 4\pi^2}))$$

La curva piana che è grafico della funzione

$$f(t) = \begin{cases} t\sin\frac{1}{t} & \text{per } t \in \left(0, \frac{2}{\pi}\right] \\ 0 & \text{per } t = 0 \end{cases}$$

non è rettificabile. Lo dimostriamo provando che l'estremo superiore delle lunghezze delle poligonali inscritte nella curva è infinito.

Consideriamo la seguente partizione di $\left[0, \frac{2}{\pi}\right]$:

$$\mathcal{P}_n = \{t_0, t_1, .., t_n\}, \text{ con}$$
$$t_k = \frac{1}{\frac{\pi}{2} + k\pi} \text{ per } k = 0, 1, 2, \ldots, n-1;\ t_n = 0.$$

La corrispondente poligonale ha lunghezza:

$$l(\mathcal{P}_n) = \sum_{k=1}^{n} \sqrt{1 + (f(t_k) - f(t_{k-1}))^2} \geq$$
$$\geq \sum_{k=1}^{n-1} |f(t_k) - f(t_{k-1})| = \sum_{k=1}^{n-1} \left[\frac{1}{\frac{\pi}{2} + k\pi} + \frac{1}{\frac{\pi}{2} + (k+1)\pi}\right] \tag{4.2}$$

D'altro canto, la stima asintotica

$$\left[\frac{1}{\frac{\pi}{2} + k\pi} + \frac{1}{\frac{\pi}{2} + (k+1)\pi}\right] \sim \frac{2}{\pi k} \text{ per } k \to \infty$$

[8] v. volume 1, capitolo 6, paragrafo 6.

implica che

$$\sum_{k=1}^{n-1}\left[\frac{1}{\frac{\pi}{2}+k\pi}+\frac{1}{\frac{\pi}{2}+(k+1)\pi}\right]\to+\infty,$$

perché la serie armonica $\sum_{k=1}^{\infty}\frac{2}{\pi k}$ è divergente. Dalla (4.2) deduciamo allora che

$$\sup_n l(\mathcal{P}_n)=\infty.$$

4.2 Cambiamenti di parametrizzazione, curve equivalenti

La lunghezza dell'arco di curva percorsa da un punto mobile dipende solo dal sostegno e non, naturalmente, dalla velocità con cui il punto si muove, né dal verso di percorrenza. Come si traduce matematicamente questa affermazione?

Sia $\mathbf{r}=\mathbf{r}(t)$, $t\in[a,b]$, un arco di curva regolare. Supponiamo di cambiare parametrizzazione, ponendo $t=\varphi(u)$, con $\varphi:[c,d]\to[a,b]$, derivabile e invertibile (in particolare, quindi, φ sarà monotona, crescente o decrescente). La nuova curva $\mathbf{r}=r(\varphi(u))$, $u\in[c,d]$, avrà lo stesso sostegno, anche se sarà percorsa con una velocità diversa. Inoltre, a seconda che φ sia crescente o decrescente, la nuova curva sarà percorsa nello stesso verso, o in verso opposto, rispetto alla curva originaria.

Se φ è crescente, le due parametrizzazioni si diranno *equivalenti*. Altrimenti, la seconda si dirà *cambio di orientazione* rispetto alla prima. In entrambi i casi, passando da $\mathbf{r}=\mathbf{r}(t)$ a $\mathbf{r}=r(\varphi(u))$ diremo che abbiamo *riparametrizzato* la curva.

ESEMPIO

Le equazioni parametriche

$$\begin{cases}x=R\cos t\\ y=R\sin t\end{cases}\qquad t\in[0,2\pi]$$

e

$$\begin{cases}x=R\cos(2u)\\ y=R\sin(2u)\end{cases}\qquad u\in[0,\pi]$$

sono due parametrizzazioni equivalenti di una circonferenza. Invece

$$\begin{cases}x=R\cos(-u)\\ y=R\sin(-u)\end{cases}\qquad u\in[0,2\pi]$$

rappresenta un cambio di orientazione.

Possiamo ora enunciare un risultato preciso riguardante la lunghezza di una curva:

PROPOSIZIONE 2.3 *La lunghezza di un arco di curva regolare è invariante per parametrizzazioni equivalenti ed anche per cambiamento di orientazione.*

Questa proposizione è un caso particolare di un risultato più generale che dimostreremo nella prossima sezione (si veda la Proposizione 2.4 nel paragrafo 5), perciò ne rimandiamo la dimostrazione.

4.3 Parametro arco o ascissa curvilinea

La lunghezza dell'arco di curva $\mathbf{r}(\tau)$ per τ da t_0 a t è una funzione di t:

$$s(t)=\int_{t_0}^{t}|\mathbf{r}'(\tau)|d\tau. \tag{4.3}$$

Se si è in grado di calcolare esplicitamente tale funzione e poi di invertirla, esprimendo t come funzione di s, è possibile riparametrizzare la curva in funzione del parametro s, detto *parametro arco o ascissa curvilinea*[9].

ESEMPIO

Per la circonferenza

$$\begin{cases}x=R\cos\theta\\ y=R\sin\theta\end{cases}\qquad\theta\in[0,2\pi]$$

si ha:

$$s(\theta)=\int_0^{\theta}Rd\theta=R\theta,$$

da cui $\theta=s/R$. Si può riparametrizzare la curva come:

$$\begin{cases}x=R\cos\left(\frac{s}{R}\right)\\ y=R\sin\left(\frac{s}{R}\right)\end{cases}\qquad s\in[0,2\pi R]$$

dove s è il parametro arco.

Per l'elica cilindrica dell'esempio 3.4 è:

$$s(t)=\int_0^t\sqrt{R^2+p^2}d\tau=t\sqrt{R^2+p^2}$$

da cui $t=\frac{s}{\sqrt{R^2+p^2}}$. Si può allora riparametrizzare la curva come:

$$\begin{cases}x=R\cos\dfrac{s}{\sqrt{R^2+p^2}}\\ y=R\sin\dfrac{s}{\sqrt{R^2+p^2}}\\ z=p\dfrac{s}{\sqrt{R^2+p^2}}\end{cases}\qquad s\in[0,2\pi\sqrt{R^2+p^2}]$$

dove s è il parametro arco.

Ciò che ha di speciale questa parametrizzazione è che il parametro s è esattamente lo spazio percorso dal punto mobile quando il parametro passa da 0 a s.

Se $\mathbf{r}=\mathbf{r}(s)$ è una curva parametrizzata mediante il parametro arco, il *vettore derivato* $\mathbf{r}'(s)$ coincide col *versore tangente* $\mathbf{T}$. Infatti, $v(s)=|\mathbf{r}'(s)|=1$, come si vede

[9] Per convenzione comunemente accettata, la lettera s si usa per indicare il parametro di una curva se e solo se si tratta proprio del parametro arco.

© 978-88-08-06485-1

dalle relazioni:

$$\frac{d\mathbf{r}}{ds} = \frac{d\mathbf{r}}{dt}\cdot\frac{dt}{ds} = \frac{1}{v(t)}\frac{d\mathbf{r}}{dt}$$
$$\left|\frac{d\mathbf{r}}{ds}\right| = \frac{1}{v(t)}\left|\frac{d\mathbf{r}}{dt}\right| = \frac{v(t)}{v(t)} = 1$$

Il vettore $\mathbf{r}'(s)$ contiene informazioni puramente geometriche: infatti, nonostante sia una "velocità", si tratta di un vettore *adimensionale*.

Se $\mathbf{r}(t)$ è una curva parametrizzata rispetto a un parametro t qualunque (non necessariamente il parametro arco), dalla (4.3) leggiamo la relazione:

$$\frac{ds}{dt} = |\mathbf{r}'(t)| = v(t),$$

che si riscrive anche, simbolicamente, nella forma

$$ds = |\mathbf{r}'(t)|dt = v(t)\,dt.$$

Il simbolo ds si chiama *lunghezza d'arco elementare*. Il suo significato intuitivo è quello di spazio percorso dal punto mobile nell'intervallo di tempo "infinitesimo" dt: nell'intervallo di tempo $[t, t+dt]$ il punto si sposta di un incremento $d\mathbf{r}$ dato da $d\mathbf{r} = \mathbf{r}'(t)\,dt$, a meno di infinitesimi di ordine superiore.

5 INTEGRALI DI LINEA (DI PRIMA SPECIE)

Supponiamo di voler calcolare la massa totale di un filo pesante, non omogeneo, di cui sia nota la densità lineare. Se il filo è rappresentato come curva parametrizzata rispetto al parametro arco, cioè $\mathbf{r} = \mathbf{r}(s)$, $s \in [0, L]$, e $\rho = \rho(s)$ è la densità lineare, se ds è un piccolo elemento di filo, la sua massa sarà $dm = \rho(s)\,ds$. Dunque la massa totale sarà uguale all'integrale:

$$m = \int_0^L \rho(s)\,ds.$$

In pratica è difficile conoscere la parametrizzazione della curva mediante il parametro arco. Se $\mathbf{r} = \mathbf{r}(t)$, $t \in [a, b]$ è una parametrizzazione qualsiasi, calcolando l'integrale precedente mediante cambio di variabili $s = s(t)$, $ds = |\mathbf{r}'(t)|dt$ si ha:

$$m = \int_a^b \rho(s(t))|\mathbf{r}'(t)|dt = \int_a^b \widetilde{\rho}(t)|\mathbf{r}'(t)|dt$$

dove si è indicata con $\widetilde{\rho}$ la funzione densità espressa mediante il parametro t (solitamente è proprio questa la funzione nota, e non $\rho(s)$).

In altri termini, l'elemento di massa del pezzetto di filo corrispondente a $[t, t+dt]$ è dato da:

$$dm = \widetilde{\rho}(t)|\mathbf{r}'(t)|dt.$$

Se poi la densità ρ fosse espressa in funzione delle variabili cartesiane (x, y, z), ossia fosse una funzione $f : \mathbb{R}^3 \to \mathbb{R}$, per calcolare l'integrale di tale densità lungo il filo dovremmo scrivere:

$$m = \int_a^b f(\mathbf{r}(t))|\mathbf{r}'(t)|dt$$

© 978-88-08-06485-1

Abbiamo così motivato la seguente:

Definizione 2.9 Sia $\mathbf{r} : [a, b] \to \mathbb{R}^m$ un arco di curva regolare di sostegno γ e sia f una funzione a valori reali, definita in un sottoinsieme A di $\mathbb{R}^m$ contenente γ, cioè $f : A \subset \mathbb{R}^m \to \mathbb{R}$ con $A \supset \gamma$. Si dice *integrale di linea (di prima specie) di f lungo γ* l'integrale

$$\int_\gamma f\,ds \equiv \int_a^b f(\mathbf{r}(t))|\mathbf{r}'(t)|dt$$

Significato geometrico dell'integrale di linea rispetto al parametro arco
Se $m = 2$, f è positiva e continua e γ è parametrizzata col parametro arco, consideriamo la superficie S in $\mathbb{R}^3$ formata dai segmenti verticali che congiungono i punti di γ ai punti del grafico di f. L'ipotesi che γ sia parametrizzata dal parametro arco significa, geometricamente, che se sviluppiamo in un piano x, y la superficie S (facendo coincidere γ con l'asse x) il parametro arco rappresenta l'ascissa x. Allora l'integrale $\int_\gamma f\,ds$ rappresenta l'area di questa superficie sviluppata e quindi l'area della superficie originale (fig. 2.11).

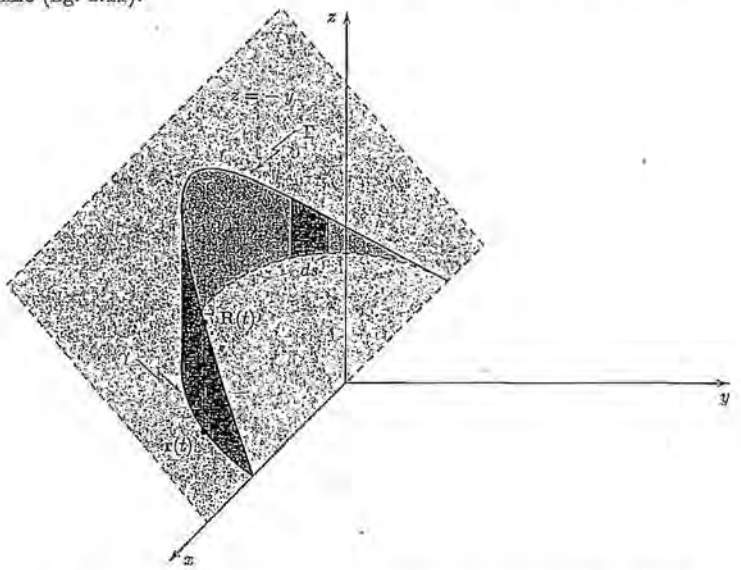


Figura 2.11. L'interpretazione geometrica di $\int_\gamma f\,ds$ con $f(x, y) = -y$ e γ semicirconferenza.

Per l'integrale di linea di prima specie vale un risultato di invarianza per cambi di parametrizzazioni, analogo a quello che abbiamo enunciato per la lunghezza dell'arco di curva:

Proposizione 2.4 *L'integrale di f di prima specie lungo γ è invariante per parametrizzazioni equivalenti ed anche per cambiamento di orientazione su γ.*

In particolare, nel caso $f \equiv 1$, da questa proposizione segue l'*invarianza della lunghezza* di una curva, rispetto a riparametrizzazioni (Proposizione 2.3).

© 978-88-08-06485-1

DIMOSTRAZIONE. La dimostrazione consiste in una semplice applicazione della formula di cambio di variabile nell'integrale e del teorema di derivazione delle funzioni composte. Consideriamo infatti l'integrale di linea

$$\int_\gamma f\,ds \equiv \int_a^b f(\mathrm{r}(t))|\mathrm{r}'(t)|dt \tag{5.1}$$

e supponiamo di cambiare parametrizzazione della curva, ponendo $t=\varphi(u)$, con $\varphi:[c,d]\to[a,b]$, derivabile e invertibile. Distinguiamo inoltre i casi in cui φ è crescente (curve equivalenti) o decrescente (cambio di orientazione).

Nel primo caso, il cambio di variabile nell'integrale (5.1) dà:

$$\int_a^b f(\mathrm{r}(t))|\mathrm{r}'(t)|dt = \int_c^d f(\mathrm{r}(\varphi(u)))|\mathrm{r}'(\varphi(u))|\varphi'(u)\,du.$$

Osserviamo che:

$$\mathrm{r}'(\varphi(u))\varphi'(u) = \frac{d}{du}\left[\mathrm{r}(\varphi(u))\right].$$

Poiché φ è crescente, $\varphi'(u)\geq 0$ perciò:

$$|\mathrm{r}'(\varphi(u))|\varphi'(u) = |\mathrm{r}'(\varphi(u))\varphi'(u)| = \left|\frac{d}{du}\left[\mathrm{r}(\varphi(u))\right]\right|.$$

Indicando con $\overline{\mathrm{r}}(u)=\mathrm{r}(\varphi(u))$ la curva parametrizzata mediante u si ha:

$$\int_a^b f(\mathrm{r}(t))|\mathrm{r}'(t)|dt = \int_c^d f(\mathrm{r}(\varphi(u)))|\mathrm{r}'(\varphi(u))|\varphi'(u)\,du = \int_c^d f(\overline{\mathrm{r}}(u)))|\overline{\mathrm{r}}'(u)|du.$$

L'identità trovata esprime proprio il fatto che la definizione di integrale di linea di prima specie, applicata alla parametrizzazione $\mathrm{r}(t)$ o $\overline{\mathrm{r}}(u)$ dà lo stesso risultato, come si voleva dimostrare.

Nel secondo caso, essendo φ decrescente si ha $\varphi'(u)\leq 0$ perciò:

$$|\mathrm{r}'(\varphi(u))|\varphi'(u) = -|\mathrm{r}'(\varphi(u))\varphi'(u)| = -\left|\frac{d}{du}\left[\mathrm{r}(\varphi(u))\right]\right|.$$

Inoltre, in questo caso sarà $\varphi(c)=b$ e $\varphi(d)=a$, perciò il cambio di variabile $t=\varphi(u)$ nell'integrale (5.1) dà:

$$\begin{aligned}\int_a^b f(\mathrm{r}(t))|\mathrm{r}'(t)|dt &= \int_d^c f(\mathrm{r}(\varphi(u)))|\mathrm{r}'(\varphi(u))|\varphi'(u)\,du = \\ &= -\int_d^c f(\mathrm{r}(\varphi(u)))\left|\frac{d}{du}\left[\mathrm{r}(\varphi(u))\right]\right|du \\ &= \int_c^d f(\overline{\mathrm{r}}(u)))|\overline{\mathrm{r}}'(u)|du,\end{aligned}$$

che è la stessa identità trovata nel primo caso. ◆

© 978-88-08-06485-1

Applicazioni fisiche e geometriche dell'integrale di linea di prima specie

Nello spazio tridimensionale, sia γ una linea materiale non omogenea di *densità lineare* ρ. Sia $\mathrm{r}:[a,b]\to\mathbb{R}^3$ una parametrizzazione regolare di γ. Abbiamo visto che la *massa totale* è data da

$$m=\int_\gamma \rho\,ds = \int_a^b \rho(\mathrm{r}(t))|\mathrm{r}'(t)|dt$$

Se il corpo è *omogeneo* ($\rho=$costante), si ritrova

$$m=\rho\cdot l(\gamma),$$

dove $l(\gamma)$ indica la lunghezza di γ.

Il *baricentro* di γ è il punto $B\equiv(\overline{x},\overline{y},\overline{z})$, con:

$$\begin{cases}\overline{x}=\dfrac{1}{m}\displaystyle\int_\gamma x\,\rho\,ds = \frac{1}{m}\int_a^b x(t)\rho(\mathrm{r}(t))|\mathrm{r}'(t)|dt \\ \overline{y}=\dfrac{1}{m}\displaystyle\int_\gamma y\,\rho\,ds = \frac{1}{m}\int_a^b y(t)\rho(\mathrm{r}(t))|\mathrm{r}'(t)|dt \\ \overline{z}=\dfrac{1}{m}\displaystyle\int_\gamma z\,\rho\,ds = \frac{1}{m}\int_a^b z(t)\rho(\mathrm{r}(t))|\mathrm{r}'(t)|dt\end{cases}$$

dove

$$\mathrm{r}(t)=(x(t),y(t),z(t)).$$

Se il corpo è omogeneo ($\rho=$costante), il baricentro si dice *centroide* e ha coordinate:

$$\overline{x}=\frac{\rho}{m}\int_\gamma x\,ds = \frac{1}{l(\gamma)}\int_a^b x(t)\,|\mathrm{r}'(t)|dt = \frac{\displaystyle\int_a^b x(t)\,|\mathrm{r}'(t)|dt}{\displaystyle\int_a^b |\mathrm{r}'(t)|dt}$$

con analoghe espressioni per $\overline{y},\overline{z}$.

Calcoliamo ora il *momento d'inerzia di γ rispetto a un asse fissato*. Se $\delta(x,y,z)$ indica la distanza di (x,y,z) da quest'asse, si ha:

$$I=\int_\gamma \delta^2\rho\,ds = \int_a^b \delta^2(\mathrm{r}(t))\rho(\mathrm{r}(t))|\mathrm{r}'(t)|dt.$$

Se il corpo è omogeneo:

$$I=\frac{m}{l(\gamma)}\int_\gamma \delta^2\,ds = m\,\frac{\displaystyle\int_a^b \delta^2(\mathrm{r}(t))|\mathrm{r}'(t)|dt}{\displaystyle\int_a^b |\mathrm{r}'(t)|dt}.$$

Osserviamo che il momento d'inerzia ha le dimensioni di *massa* $\times$ *[lunghezza]*^{2}.

Esempi

Determiniamo il centroide di una semicirconferenza omogenea.
Posto $x = R\cos\theta$, $y = R\sin\theta$, $\theta \in [0,\pi]$, ci aspettiamo per simmetria che sia $\overline{x} = 0$, e $\overline{y} \in (0,R)$. In questo caso $ds = R\,d\theta$ e

$$\overline{x} = \frac{1}{\pi R}\int_0^\pi R\cos\theta\, R\,d\theta = 0; \quad \overline{y} = \frac{1}{\pi R}\int_0^\pi R\sin\theta\, R\,d\theta = \frac{2}{\pi}R.$$

Calcoliamo il momento d'inerzia di una circonferenza omogenea che ruota lungo un asse contenente un suo diametro.

Nel piano x, y, sia l'asse y l'asse di rotazione. Detta m la massa totale si ha $\delta(x,y) = x$ e perciò

$$I = \frac{m}{2\pi R}\int_0^{2\pi} R^2\cos^2\theta\, R\,d\theta = \frac{1}{2}mR^2.$$

6 ELEMENTI DI GEOMETRIA DIFFERENZIALE DELLE CURVE

6.1 Curvatura e normale principale per una curva in $\mathbb{R}^m$

In questo paragrafo considereremo sempre una curva regolare $\mathbf{r} : [a,b] \to \mathbb{R}^m$, con m qualsiasi, di classe $C^2([a,b])$, ossia tale che $\mathbf{r}$ sia derivabile due volte in $[a,b]$, con derivata seconda continua. Ci interessa introdurre alcuni concetti e stabilire alcune relazioni che hanno un significato sia geometrico che cinematico. Cominciamo a sintetizzare alcuni fatti e notazioni già utilizzati in precedenza.

Ricordiamo che il *versore tangente* è definito come

$$\mathbf{T} = \operatorname{vers}(\mathbf{r}'),$$

perciò

$$\mathbf{T}(t) = \frac{\mathbf{r}'(t)}{|\mathbf{r}'(t)|}, \text{ mentre } \mathbf{T}(s) = \mathbf{r}'(s)$$

in quanto, come già visto nel paragrafo 4.3, $|\mathbf{r}'(s)| = 1$, se s è il parametro arco.

Con terminologia cinematica, chiamiamo:

vettore velocità istantanea	il vettore	$\mathbf{r}'(t)$;
velocità istantanea scalare	lo scalare	$v(t) = \lvert\mathbf{r}'(t)\rvert$ (perciò $v(s) = 1$);
vettore accelerazione istantanea	il vettore	$\mathbf{a}(t) = \mathbf{r}''(t)$;
accelerazione istantanea scalare	lo scalare	$a(t) = \lvert\mathbf{r}''(t)\rvert$.

In base alle definizioni appena date, si può anche scrivere:

$$\mathbf{r}'(t) = v(t)\,\mathbf{T}(t).$$

Il seguente fatto elementare sarà utilizzato molto spesso in seguito:

PROPOSIZIONE 2.5 *Sia $I \subseteq \mathbb{R}$ un intervallo e $\mathbf{u} : I \to \mathbb{R}^m$ una funzione vettoriale di modulo costante, cioè $|\mathbf{u}(t)| = c$ per ogni t. Allora $\mathbf{u}\cdot\mathbf{u}' = 0$, ossia $\mathbf{u}$ è ortogonale a $\mathbf{u}'$ in ogni istante.*

DIMOSTRAZIONE. Dall'ipotesi $|\mathbf{u}(t)| = c$ abbiamo:

$$\mathbf{u}(t)\cdot\mathbf{u}(t) = c^2,$$

da cui derivando si ha:

$$\mathbf{u}'\cdot\mathbf{u} + \mathbf{u}\cdot\mathbf{u}' = 2\mathbf{u}'\cdot\mathbf{u} = 0$$

(abbiamo applicato il Teorema 2.1, punto 5), che è la tesi. ◆

Normale principale e curvatura, rispetto al parametro arco

Supponiamo ora che la nostra curva sia parametrizzata rispetto al parametro arco, $\mathbf{r}(s)$. Vogliamo definire una grandezza che misuri quantitativamente il modo in cui la curva cambia direzione. È naturale, per far questo, calcolare la derivata del versore tangente, $\mathbf{T}'(s)$.

Poiché $|\mathbf{r}'(s)| = 1$, la Proposizione 2.5 implica che:

$$\mathbf{r}''(s) \text{ è ortogonale a } \mathbf{r}'(s).$$

Dal punto di vista cinematico: *i vettori "velocità" $\mathbf{v}(s)$ e "accelerazione" $\mathbf{a}(s)$ sono ortogonali in ogni istante.*

Dal punto di vista geometrico: *la derivata $\mathbf{T}'(s)$ del versore tangente è ortogonale al versore tangente $\mathbf{T}(s)$.*

Questo giustifica la seguente:

DEFINIZIONE 2.10 Chiamiamo versore *normale principale* della curva $\mathbf{r}(s)$ il versore

$$\mathbf{N}(s) = \operatorname{vers}(\mathbf{T}'(s)) = \frac{\mathbf{T}'(s)}{|\mathbf{T}'(s)|},$$

definito nei punti in cui $\mathbf{T}'(s) \neq \mathbf{0}$.

Chiamiamo *curvatura* della curva la funzione scalare

$$k(s) = |\mathbf{T}'(s)|.$$

Possiamo quindi scrivere:

$$\mathbf{T}'(s) = k(s)\,\mathbf{N}(s). \tag{6.1}$$

Osserviamo che, se ci troviamo in uno spazio con più di due dimensioni ($m \geq 3$), in ogni punto di una curva esistono *infinite* direzioni ortogonali alla curva stessa (cioè alla sua tangente). La *normale principale* è una di queste, che ha la proprietà caratteristica di individuare la direzione in cui avviene la variazione principale del versore tangente; ci dice, in altre parole, la direzione verso cui sta curvando la nostra curva. La *curvatura* invece è uno scalare che misura l'intensità con cui ciò avviene.

Sottolineiamo che è essenziale, in questo procedimento, utilizzare il parametro arco: questo garantisce che la velocità $\mathbf{r}'(s)$ abbia sempre modulo unitario, e quindi sia ortogonale a $\mathbf{r}''(s)$; per una curva parametrizzata rispetto ad un parametro t generico, questo non sarebbe vero, come vedremo meglio in seguito, quando descriveremo la scomposizione dell'accelerazione nelle componenti normale e tangenziale.

Accanto alla curvatura si introduce anche il suo reciproco, che prende il nome di *raggio di curvatura*:

$$\rho(s) = \frac{1}{k(s)},$$

definita nei punti in cui $k(s) \neq 0$; se $k(s) = 0$ si dice anche che il raggio di curvatura è infinito.

Il cerchio di raggio $\rho(s)$ e centro nel punto $\mathrm{r}(s) + \rho(s)\,\mathrm{N}(s)$ si dice *cerchio osculatore* della curva. Il suo significato è illustrato dalla figura 2.12.

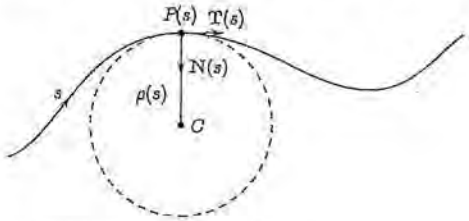


Figura 2.12. Vettori tangente e normale. Cerchio osculatore.

Normale principale e curvatura, rispetto a un parametro qualsiasi. Decomposizione dell'accelerazione

Vediamo ora come si calcolano il versore normale principale e la curvatura per una curva $\mathrm{r}(t)$ descritta da un parametro qualsiasi.

Cominciamo a osservare che dalla definizione (4.3) di parametro arco si ha:

$$\frac{ds}{dt} = |\mathrm{r}'(t)| = v(t),$$

quindi

$$\mathrm{T}'(t) = \mathrm{T}'(s)\frac{ds}{dt} = \mathrm{T}'(s)\,v(t). \tag{6.2}$$

La (6.2) ci dice che i vettori $\mathrm{T}'(t), \mathrm{T}'(s)$ sono paralleli e concordi, perciò hanno lo stesso versore. Dunque possiamo scrivere:

$$\mathrm{N}(t) = \mathrm{vers}\,(\mathrm{T}'(t)) = \frac{\mathrm{T}'(t)}{|\mathrm{T}'(t)|}, \tag{6.3}$$

che definisce il *versore normale principale* in funzione del parametro t qualunque. Inoltre, ancora dalla (6.2) leggiamo:

$$|\mathrm{T}'(t)| = |\mathrm{T}'(s)|\,v(t) = k(s)\,v(t),$$

da cui ricaviamo l'espressione della *curvatura* in funzione di t:

$$k(t) = \frac{|\mathrm{T}'(t)|}{v(t)}. \tag{6.4}$$

Le relazioni (6.3)-(6.4) danno anche:

$$\mathrm{T}'(t) = k(t)\,v(t)\,\mathrm{N}(t), \tag{6.5}$$

che è la forma più generale della relazione (6.1).

Il *raggio di curvatura* si esprime ancora mediante la formula:

$$\rho(t) = \frac{1}{k(t)}.$$

Se la curva viene interpretata in senso cinematico come legge oraria di un moto, t è il tempo e i parametri che compaiono nelle equazioni della curva hanno dimensioni coerenti con questa interpretazione, il raggio di curvatura ha effettivamente le dimensioni di una lunghezza, e la curvatura, di conseguenza, le dimensioni del reciproco di una lunghezza.

ESEMPIO

Calcoliamo i versori T, N e la curvatura k per l'elica cilindrica:

$$\begin{cases} x = R\cos\omega t \\ y = R\sin\omega t \\ z = pt \end{cases} \qquad t \in [0, T].$$

Dimensionalmente, t è il tempo, ω è una pulsazione (ha le dimensioni di una frequenza), R ha le dimensioni di una lunghezza e p quelle di una velocità. Si ha:

$$\mathrm{r}(t) = (R\cos\omega t, R\sin\omega t, pt);$$

$$\mathrm{r}'(t) = (-R\omega\sin\omega t, R\omega\cos\omega t, p);$$

$$v(t) = |\mathrm{r}'| = \sqrt{R^2\omega^2 + p^2};$$

$$\mathrm{T}(t) = \frac{\mathrm{r}'(t)}{v(t)} = \left(-\frac{R\omega}{\sqrt{R^2\omega^2+p^2}}\sin\omega t, \frac{R\omega}{\sqrt{R^2\omega^2+p^2}}\cos\omega t, \frac{p}{\sqrt{R^2\omega^2+p^2}}\right);$$

$$\mathrm{T}'(t) = \left(-\frac{R\omega^2}{\sqrt{R^2\omega^2+p^2}}\cos\omega t, -\frac{R\omega^2}{\sqrt{R^2\omega^2+p^2}}\sin\omega t, 0\right);$$

$$|\mathrm{T}'(t)| = \frac{R\omega^2}{\sqrt{R^2\omega^2+p^2}},$$

e quindi, in base alle (6.3), (6.4),

$$\mathrm{N}(t) = \frac{\mathrm{T}'(t)}{|\mathrm{T}'(t)|} = (-\cos\omega t, -\sin\omega t, 0);$$

$$k(t) = \frac{|\mathrm{T}'(t)|}{v(t)} = \frac{R\omega^2}{R^2\omega^2+p^2}$$

L'elica cilindrica ha quindi curvatura costante (non nulla). Si osservi che questa costante ha effettivamente le dimensioni del reciproco di una lunghezza.

Si può dimostrare che l'elica cilindrica è *l'unica curva nello spazio* $\mathbb{R}^3$ ad avere questa proprietà; di conseguenza la circonferenza, che si può vedere come particolare elica cilindrica con $p = 0$, è *l'unica curva piana* ad avere curvatura costante non nulla.

© 978-88-08-06485-1

La formula (6.5) ha anche un'interessante applicazione cinematica:

PROPOSIZIONE 2.6 (SCOMPOSIZIONE DELL'ACCELERAZIONE) *Vale la seguente formula:*

$$\mathrm{a}(t) = v'(t)\,\mathrm{T}(t) + v^2(t)\,k(t)\,\mathrm{N}(t) \tag{6.6}$$

dove $\mathrm{a}(t) = \mathrm{r}''(t)$, *il primo addendo a secondo membro si chiama* accelerazione tangenziale *e il secondo* accelerazione normale[10].

DIMOSTRAZIONE. Derivando l'identità

$$\mathrm{r}'(t) = v(t)\,\mathrm{T}(t)$$

abbiamo:

$$\mathrm{a}(t) = \mathrm{r}''(t) = v'(t)\,\mathrm{T}(t) + v(t)\,\mathrm{T}'(t) = v'(t)\,\mathrm{T}(t) + v^2(t)\,k(t)\,\mathrm{N}(t),$$

dove nell'ultima uguaglianza si è usata la (6.5). ♦

Introducendo il raggio di curvatura $\rho(t) = 1/k(t)$, il termine dell'accelerazione normale assume la forma, più consueta in cinematica:

$$\frac{v^2(t)}{\rho(t)}\mathrm{N}(t).$$

6.2 Calcolo della curvatura per curve nello spazio $\mathbb{R}^3$ o nel piano

Ricordiamo che tutto quanto abbiamo detto fin qui vale per curve in dimensione m qualunque, in particolare quindi per le curve nello spazio $\mathbb{R}^3$ o nel piano. In questo paragrafo ci concentriamo appunto sul caso delle curve in $\mathbb{R}^3$, o nel piano, per le quali è possibile ottenere formule di calcolo più esplicite, utilizzando le proprietà del prodotto vettoriale.

Cominciamo a stabilire il seguente fatto:

PROPOSIZIONE 2.7 *Valgono le seguenti formule di calcolo per la curvatura:*

$$k(t) = \frac{|\mathrm{r}'(t) \times \mathrm{r}''(t)|}{v^3(t)} \quad \textit{rispetto a un parametro } t \textit{ qualunque,} \tag{6.7}$$

$$k(s) = |\mathrm{r}'(s) \times \mathrm{r}''(s)| \quad \textit{rispetto al parametro arco.}$$

DIMOSTRAZIONE. Dimostriamo la prima (la seconda segue perché rispetto al parametro arco è $v(s) = 1$). Calcoliamo $\mathrm{r}'(t) \times \mathrm{r}''(t)$ facendo uso della (6.6):

$$\begin{aligned}\mathrm{r}' \times \mathrm{r}'' &= \mathrm{r}' \times [v'\mathrm{T} + v^2 k\mathrm{N}] = \\ &= v'\mathrm{r}' \times \mathrm{T} + v^2 k\mathrm{r}' \times \mathrm{N}.\end{aligned}$$

[10] o talvolta, accelerazione centripeta.

© 978-88-08-06485-1

Ora, $\mathrm{r}' \times \mathrm{T} = 0$ perché $\mathrm{T} = \mathrm{vers}(\mathrm{r}')$ è parallelo a r'. Inoltre:

$$\mathrm{r}' \times \mathrm{N} = v\mathrm{T} \times \mathrm{N}.$$

Pertanto:

$$|\mathrm{r}' \times \mathrm{r}''| = v^3 k\,|\mathrm{T} \times \mathrm{N}|.$$

D'altro canto $|\mathrm{T} \times \mathrm{N}| = 1$ perché T, N sono due versori ortogonali, quindi $k = |\mathrm{r}' \times \mathrm{r}''|/v^3$, che è quanto volevamo provare. ♦

Esempio

Es 2.2 Calcoliamo nuovamente la curvatura dell'elica cilindrica dell'Esempio 6.1, facendo uso ora della (6.7). Si ha:

$$\begin{aligned}\mathrm{r}(t) &= (a\cos\omega t, a\sin\omega t, bt);\\ \mathrm{r}'(t) &= (-a\omega\sin\omega t, a\omega\cos\omega t, b);\\ \mathrm{r}''(t) &= (-a\omega^2\cos\omega t, -a\omega^2\sin\omega t, 0);\end{aligned}$$

quindi:

$$\begin{aligned}\mathrm{r}' \times \mathrm{r}'' &= \begin{vmatrix} \mathrm{i} & \mathrm{j} & \mathrm{k} \\ -a\omega\sin\omega t & a\omega\cos\omega t & b \\ -a\omega^2\cos\omega t & -a\omega^2\sin\omega t & 0 \end{vmatrix} = \\ &= (ab\omega^2\sin\omega t, -ab\omega^2\cos\omega t, a^2\omega^3);\\ |\mathrm{r}' \times \mathrm{r}''| &= \sqrt{a^2b^2\omega^4 + a^4\omega^6} = a\omega^2\sqrt{a^2\omega^2 + b^2}.\end{aligned}$$

D'altro canto:

$$v(t) = |\mathrm{r}'| = \sqrt{a^2\omega^2 + b^2}$$

perciò:

$$k(t) = \frac{|\mathrm{r}'(t) \times \mathrm{r}''(t)|}{v(t)^3} = \frac{a\omega^2\sqrt{a^2\omega^2 + b^2}}{(a^2\omega^2 + b^2)^{3/2}} = \frac{a\omega^2}{a^2\omega^2 + b^2},$$

Il caso delle curve piane

Nel caso delle curve piane, le formule ottenute nella Proposizione 2.7 possono essere ulteriormente particolarizzate.

Se

$$\mathrm{r}(t) = (x(t), y(t)), \text{ per } t \in [a, b]$$

è un arco di curva regolare piana, di classe $C^2[a, b]$, possiamo vederla in particolare come curva nello spazio $\mathbb{R}^3$ di equazioni:

$$\mathrm{r}(t) = (x(t), y(t), 0), \text{ per } t \in [a, b].$$

Perciò

$$\mathrm{r}' \times \mathrm{r}'' = \begin{vmatrix} \mathrm{i} & \mathrm{j} & \mathrm{k} \\ x'(t) & y'(t) & 0 \\ x''(t) & y''(t) & 0 \end{vmatrix} = \mathrm{k}\,(x'(t)\,y''(t) - x''(t)\,y'(t))$$

e quindi valgono le seguenti formule di calcolo per la curvatura di una curva piana:

$$k(t) = \frac{|\mathbf{r}'(t) \times \mathbf{r}''(t)|}{v^3(t)} = \frac{|x'(t)\,y''(t) - x''(t)\,y'(t)|}{\left(x'(t)^2 + y'(t)^2\right)^{3/2}}. \tag{6.8}$$

In particolare, rispetto al parametro arco:

$$k(s) = |\mathbf{r}'(s) \times \mathbf{r}''(s)| = |x'(s)\,y''(s) - x''(s)\,y'(s)|.$$

Nel caso particolare di un arco di curva che è *grafico di una funzione*

$$y = f(x) \text{ per } x \in [a, b]$$

si ottiene, sostituendo $x(t) = t, y(t) = f(t)$ nelle (6.8) e quindi chiamando nuovamente x il parametro:

$$k(x) = \frac{|f''(x)|}{\left(1 + f'(x)^2\right)^{3/2}}. \tag{6.9}$$

Riotteniamo così per altra via la formula già ottenuta nel volume 1 (capitolo 4, paragrafo 5.1), quando abbiamo discusso il significato geometrico della derivata seconda di una funzione di una variabile[11].

Esempio

Calcoliamo la curvatura dell'ellisse

$$\begin{cases} x = a\cos\omega t \\ y = b\sin\omega t \end{cases} \quad t \in \left[0, \frac{2\pi}{\omega}\right].$$

Si ha:

$$\begin{aligned} \mathbf{r}'(t) &= (-a\omega\sin\omega t, b\omega\cos\omega t) \\ \mathbf{r}''(t) &= (-a\omega^2\cos\omega t, -b\omega^2\sin\omega t) \\ k(t) &= \frac{|x'(t)\,y''(t) - x''(t)\,y'(t)|}{\left(x'(t)^2 + y'(t)^2\right)^{3/2}} = \frac{ab}{\left(a^2\sin^2\omega t + b^2\cos^2\omega t\right)^{3/2}}. \end{aligned}$$

Nel caso particolare della *circonferenza*, $a = b = R$, si ottiene:

$$k(t) = \frac{R^2}{(R^2)^{3/2}} = \frac{1}{R},$$

ossia la curvatura (in questo caso costante, come già osservato nell'Esempio 6.1) è il reciproco del raggio.

[11] Ricordiamo la conclusione a cui eravamo giunti in quel contesto: se, tra tutte le circoferenze tangenti al grafico di $f(x)$, cerchiamo quella che ha col grafico un contatto del second'ordine, troviamo che il raggio di questa circonferenza dev'essere proprio $R(x) = 1/k(x)$ con k come nella (6.9).

Tornando al caso di un'ellisse generica, chiediamoci ora quando la curvatura è massima o minima. Poiché

$$k(t) = \frac{ab}{\left(a^2\sin^2\omega t + b^2\cos^2\omega t\right)^{3/2}},$$

$k(t)$ sarà massima (rispettivamente, minima) quando

$$a^2\sin^2\omega t + b^2\cos^2\omega t \text{ è minima (massima).}$$

D'altro canto, essendo $\sin^2\omega t + \cos^2\omega t = 1$, la quantità $\left(a^2\sin^2\omega t + b^2\cos^2\omega t\right)$ varia tra b^2 e a^2, perciò supponendo, per fissare le idee, $a > b$, essa è minima quando è uguale a b^2, ossia per $\sin\omega t = 0$, ed è massima quando vale a^2, ossia per $\cos\omega t = 0$. I punti di massima e minima curvatura sono quindi i 4 punti di intersezione dell'ellisse con gli assi, ossia i 4 vertici dell'ellisse; in particolare, sono di curvatura minima i punti di intersezione con l'asse y (corrispondente al semiasse minore dell'ellisse), e di curvatura massima i punti di intersezione con l'asse x (corrispondente al semiasse maggiore dell'ellisse).

In generale:

DEFINIZIONE 2.11. Data una curva piana di classe $C^2([a, b])$, chiusa, si dicono *vertici della curva* i punti in cui $k'(t) = 0$. ■

In particolare, quindi, sono vertici della curva i punti in cui la curvatura è massima e minima. Il significato intuitivo della nozione di vertice è chiarito appunto dall'ultimo esempio fatto.

Si può dimostrare che una curva chiusa e convessa ha sempre almeno 4 vertici. L'ellisse, per esempio, ne ha esattamente 4.

Esempio

Calcoliamo la curvatura della *catenaria*

$$y = \frac{1}{a}\mathrm{Ch}(ax) \text{ per } x \in [-1, 1].$$

Poiché

$$f'(x) = \mathrm{Sh}(ax);\ f''(x) = a\mathrm{Ch}(ax)$$

la formula (6.9) ci dà:

$$k(x) = \frac{a\mathrm{Ch}(ax)}{(1 + \mathrm{Sh}^2(ax))^{3/2}} = \frac{a\mathrm{Ch}(ax)}{\mathrm{Ch}^3(ax)} = \frac{a}{\mathrm{Ch}^2(ax)}.$$

In particolare, per $x = 0$ (vertice della catenaria) osserviamo che la curvatura è massima e pari ad a.

6.3 Torsione e terna intrinseca per curve nello spazio $\mathbb{R}^3$

In questo paragrafo consideriamo una curva regolare $\mathbf{r} : I \to \mathbb{R}^3$, di classe $C^3(I)$ (ossia derivabile tre volte e con $\mathbf{r}'''(t)$ funzione continua). L'obiettivo è definire un sistema di riferimento cartesiano che sia adatto a descrivere la curva nell'intorno di un suo punto.

DEFINIZIONE 2.12 Data una curva $\mathrm{r}: I \to \mathbb{R}^3$ regolare e di classe C^3, definiamo il *versore binormale* della curva come:

$$\mathrm{B} = \mathrm{T} \times \mathrm{N}.$$

I tre versori $\mathrm{T}, \mathrm{N}, \mathrm{B}$ costituiscono un sistema di riferimento ortonormale, detto *terna intrinseca* per la curva. ▩

La terna intrinseca è un sistema di riferimento naturale per la curva, nell'intorno del punto in cui i tre versori sono calcolati. Conosciamo già il significato geometrico del versore tangente e normale: il primo dà la miglior approssimazione locale della curva, il secondo punta nella direzione in cui la curva stessa si sta incurvando; il piano generato da T ed N dovrebbe quindi essere il piano che più si avvicina a contenere (localmente) la curva, anche quando questa non è una curva piana. Si può dimostrare in modo quantitativo che le cose vanno effettivamente in questo modo. Questo piano viene detto *piano osculatore*, e ha per versore normale proprio il terzo elemento della terna intrinseca, B. Se la curva è piana, essa giace identicamente nel suo piano osculatore e in questo caso B è costante.

Le derivate dei tre vettori $\mathrm{T}, \mathrm{N}, \mathrm{B}$ ci danno un'informazione precisa di come la curva sta cambiando direzione e forma nel punto considerato. D'altro canto, essendo $\mathrm{T}, \mathrm{N}, \mathrm{B}$ un sistema di riferimento, possiamo esprimere le derivate $\mathrm{T}', \mathrm{N}', \mathrm{B}'$ come combinazioni lineari di $\mathrm{T}, \mathrm{N}, \mathrm{B}$. Le relazioni trovate sono particolarmente semplici ed espressive se la curva è parametrizzata mediante il parametro arco. Questo è appunto il contenuto del prossimo importante teorema:

TEOREMA 2.5 (FORMULE DI FRENET-SERRET) *Sia* $\mathrm{r}: I \to \mathbb{R}^3$ *una curva regolare e di classe* $C^3(I)$*, parametrizzata rispetto al parametro arco. Valgono allora le seguenti formule:*

$$\begin{cases} \mathrm{T}'(s) = k(s)\,\mathrm{N}(s) \\ \mathrm{N}'(s) = -k(s)\,\mathrm{T}(s) - \tau(s)\,\mathrm{B}(s) \\ \mathrm{B}'(s) = \tau(s)\,\mathrm{N}(s) \end{cases}$$

per un'opportuna funzione scalare $\tau(s)$*, detta* torsione della curva.

DIMOSTRAZIONE. La prima identità, $\mathrm{T}'(s) = k(s)\,\mathrm{N}(s)$, è la (6.1), già nota.

Per stabilire la seconda, osserviamo che N' è ortogonale ad N, perché $|\mathrm{N}(s)| = 1$ (Proposizione 2.5). Perciò N' si può scrivere come combinazione linerare di T e B, con opportuni coefficienti:

$$\mathrm{N}'(s) = a(s)\,\mathrm{T}(s) + b(s)\,\mathrm{B}(s).$$

Notiamo che per definizione $b = \mathrm{N}' \cdot \mathrm{B}$ e $a = \mathrm{N}' \cdot \mathrm{T}$. Derivando l'identità $\mathrm{N} \cdot \mathrm{T} = 0$ abbiamo:

$$\mathrm{N}' \cdot \mathrm{T} + \mathrm{N} \cdot \mathrm{T}' = 0$$

da cui

$$a = \mathrm{N}' \cdot \mathrm{T} = -\mathrm{N} \cdot \mathrm{T}' = -k(s)$$

dove si è sfruttato il fatto che $\mathrm{N} \cdot \mathrm{T}' = k(s)$ essendo $\mathrm{T}' = k\mathrm{N}$. Questo stabilisce la seconda identità, pur di definire $\tau = -b$, ossia

$$\tau(s) = -\mathrm{N}' \cdot \mathrm{B} = \mathrm{N} \cdot \mathrm{B}'. \tag{6.10}$$

Per la terza, procediamo in modo simile: B' è ortogonale a B, perché $|\mathrm{B}(s)| = 1$ e perciò B' si può scrivere come combinazione lineare di T e N, con opportuni coefficienti:

$$\mathrm{B}'(s) = c(s)\,\mathrm{T}(s) + d(s)\,\mathrm{N}(s).$$

Ragionando ancora come in precedenza, vediamo che:

$$c = \mathrm{B}' \cdot \mathrm{T} = -\mathrm{B} \cdot \mathrm{T}' = 0, \text{ dalla prima identità,}$$
$$d = \mathrm{B}' \cdot \mathrm{N} = -\mathrm{B} \cdot \mathrm{N}' = -b = \tau, \text{ dalla seconda identità.}$$

Questo completa la dimostrazione. ∎

Le formule di Frenet-Serret ci dicono che due sole funzioni scalari, la curvatura e la torsione della curva, determinano implicitamente la curva stessa (a meno di rototraslazioni). Infatti, supponendo note k e τ, le formule di Frenet-Serret si possono vedere come un sistema lineare di 9 equazioni differenziali ordinarie in 9 funzioni incognite, che si può dimostrare essere univocamente risolubile, assegnata una posizione iniziale della terna intrinseca[12]. Questo significa che tutta l'informazione rilevante sulla forma della curva è contenuta nelle funzioni curvatura e torsione.

In particolare, la terza identità del teorema ci dice che se la torsione è identicamente nulla, il vettore binormale è costante. In questo caso la curva è piana, e valgono le formule:

$$\begin{cases} \mathrm{T}'(s) = k(s)\,\mathrm{N}(s) \\ \mathrm{N}'(s) = -k(s)\,\mathrm{T}(s) \end{cases}$$

ESEMPIO

Calcoliamo la terna intrinseca e la torsione per l'elica cilindrica, che abbiamo riparametrizzato rispetto al parametro arco nell'Esempio 4.7:

$$\begin{cases} x = R\cos\dfrac{s}{\sqrt{R^2+p^2}} \\ y = R\sin\dfrac{s}{\sqrt{R^2+p^2}} \\ z = p\dfrac{s}{\sqrt{R^2+p^2}} \end{cases} \qquad s \in [0, 2\pi\sqrt{R^2+p^2}]$$

Si ha (lasciando alcuni passaggi al lettore), indicando per brevità $\omega = 1/\sqrt{R^2+p^2}$:

$$\mathrm{T}(s) = \mathrm{r}'(s) = (-\omega R\sin\omega s, \omega R\cos\omega s, \omega p);$$
$$\mathrm{N}(s) = \mathrm{vers}\,(\mathrm{T}'(s)) = (-\cos\omega s, -\sin\omega s, 0);$$
$$\mathrm{B}(s) = \mathrm{T}(s) \times \mathrm{N}(s) = (\omega p\sin\omega s, -\omega p\cos\omega s, \omega R);$$
$$\tau(s) = -\mathrm{N}' \cdot \mathrm{B} = -\omega^2 p.$$

Si osserva che l'elica cilindrica ha torsione costante (oltre ad avere, come già visto, curvatura costante).

Per il calcolo della torsione a partire dalle equazioni parametriche della curva può essere comoda la formula espressa dal prossimo risultato, che non dimostriamo:

[12] Questo fatto sarà discusso in dettaglio nel capitolo 8, paragrafo 2.

© 978-88-08-06485-1

Proposizione 2.8 *Sia $\mathrm{r}: I \to \mathbb{R}^3$ una curva regolare di classe $C^3(I)$. Se $\mathrm{r}' \times \mathrm{r}'' \neq 0$, vale la formula:*

$$\tau(t) = -\frac{\mathrm{r}'(t) \times \mathrm{r}''(t) \cdot \mathrm{r}'''(t)}{|\mathrm{r}'(t) \times \mathrm{r}''(t)|^2}.$$

Esempio

Calcoliamo la curvatura e la torsione dell'*elica conica*:

$$\begin{cases} x = t\cos t \\ y = t\sin t \\ z = t \end{cases} \qquad t \in [0, 6\pi].$$

Si ha:

$$\begin{aligned} \mathrm{r}'(t) &= (\cos t - t\sin t, \sin t + t\cos t, 1); \\ \mathrm{r}''(t) &= (-2\sin t - t\cos t, 2\cos t - t\sin t, 0); \\ \mathrm{r}'''(t) &= (-3\cos t + t\sin t, -3\sin t - t\cos t, 0); \end{aligned}$$

$$\begin{aligned} |\mathrm{r}'(t)| &= \sqrt{2+t^2}; \\ \mathrm{r}'(t) \times \mathrm{r}''(t) &= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \cos t - t\sin t & \sin t + t\cos t & 1 \\ -2\sin t - t\cos t & 2\cos t - t\sin t & 0 \end{vmatrix} = \\ &= (-2\cos t + t\sin t, -2\sin t - t\cos t, 2 + t^2); \\ |\mathrm{r}'(t) \times \mathrm{r}''(t)| &= \sqrt{8 + 5t^2 + t^4}; \\ \mathrm{r}'(t) \times \mathrm{r}''(t) \cdot \mathrm{r}'''(t) &= (-2\cos t + t\sin t, -2\sin t - t\cos t, 2 + t^2) \cdot \\ &\quad \cdot (-3\cos t + t\sin t, -3\sin t - t\cos t, 0) \\ &= 6 + t^2. \end{aligned}$$

Perciò:

$$\begin{aligned} k(t) &= \frac{|\mathrm{r}' \times \mathrm{r}''|}{|\mathrm{r}'|^3} = \frac{\sqrt{8 + 5t^2 + t^4}}{(2+t^2)^{3/2}}; \\ \tau(t) &= -\frac{(\mathrm{r}' \times \mathrm{r}'') \cdot \mathrm{r}'''}{|\mathrm{r}' \times \mathrm{r}''|^2} = -\frac{6 + t^2}{8 + 5t^2 + t^4}. \end{aligned}$$

Esercizi

1. Verificare che il *Folium* di Cartesio

$$\begin{cases} x = t(t-1) \\ y = t(t-1)(2t-1) \end{cases} \qquad t \in (-\infty, +\infty)$$

è una curva regolare.

2. Verificare che la *spirale logaritmica*, di equazione polare $\rho = e^{b\theta}$, $\theta \in (-\infty, +\infty)$, è un arco di curva regolare. Calcolare ds. Esprimere poi θ in funzione del parametro arco s.

© 978-88-08-06485-1

3. Una linea materiale non omogenea è disposta a forma di *spirale di Archimede*, di equazione polare $\rho = A\theta$, $\theta \in [0, 4\pi]$. La sua densità lineare è $d = B\theta$. Calcolare la massa totale della linea materiale.

4. Calcolare la lunghezza totale dell'*astroide*:

$$\begin{cases} x = (\cos\theta)^3 \\ y = (\sin\theta)^3 \end{cases} \qquad \theta \in [0, 2\pi].$$

Si noti che, trattandosi di una curva regolare a tratti, per calcolarne la lunghezza abbiamo bisogno di applicare il Teorema 2.4 e la Proposizione 2.2.

5. Si consideri l'*elica conica*:

$$\begin{cases} x = t\cos t \\ y = t\sin t \\ z = t \end{cases} \qquad t \in [0, 6\pi].$$

Si verifichi che la curva è regolare e se ne calcoli la lunghezza.

6. Calcolare la lunghezza della *cicloide*:

$$\begin{cases} x = R(t - \sin t) \\ y = R(1 - \cos t) \end{cases} \qquad t \in [0, 2\pi].$$

7. Si consideri la *catenaria*:

$$y = a\mathrm{Ch}(bx), \qquad x \in [0, T].$$

Si scriva l'integrale che ne assegna la lunghezza. Lo si calcoli poi nel caso particolare $a = 1/b$.

8. Determinare il centroide dell'arco di *ellisse* omogeneo:

$$\begin{cases} x = 2\cos t \\ y = \sin t \end{cases} \qquad t \in [0, \pi].$$

Indicare con L la lunghezza dell'arco, senza calcolarla.

9. Determinare il momento d'inerzia dell'arco di *parabola* $y = x^2, x \in [-1, 1]$, omogenea di densità ρ, che ruota attorno all'asse y.

10. Sia γ l'arco di *spirale di Archimede* di equazione polare $\rho = A\theta, \theta \in [0, 4\pi]$. Calcolare l'integrale di linea:

$$\int_\gamma \theta^3 ds.$$

11. Determinare la lunghezza dell'arco di curva $y = \log x$, $x \in [1, 2]$.

12. Calcolare l'integrale di linea $\int_\gamma x ds$ dove γ è l'arco di iperbole $y = \frac{1}{x}$, $x \in [\frac{1}{2}, 2]$.

13. Dimostrare tutti i punti del Teorema 2.1.

14. A partire dall'equazione polare della conica (3.7), scrivere le equazioni parametriche corrispondenti; quindi, eliminando il parametro θ, ritrovare l'equazione cartesiana

$$x^2 + y^2 = \varepsilon^2 (p + x)^2.$$

 © 978-88-08-06485-1

Al variare di ε, classificare quindi il tipo di conica, dimostrando l'affermazione (3.8) fatta.

15 Calcolare la curvatura dell'*astroide*:

$$\begin{cases} x = (\cos\theta)^3 \\ y = (\sin\theta)^3 \end{cases} \qquad \theta \in [0, 2\pi].$$

16 Calcolare la curvatura della *cicloide*:

$$\begin{cases} x = R(t - \sin t) \\ y = R(1 - \cos t) \end{cases} \qquad t \in [0, 2\pi].$$

17 Si calcoli la torsione dell'*elica cilindrica*

$$\begin{cases} x = R\cos\omega t \\ y = R\sin\omega t \\ z = pt \end{cases} \qquad t \in [0, T]$$

utilizzando la Proposizione 2.8, anziché il metodo visto nell'Esempio 6.5.

18 Per una curva piana espressa in forma polare

$$\rho = f(\theta), \ \theta \in [\theta_1, \theta_2],$$

si dimostri che la curvatura è espressa da:

$$k(\theta) = \frac{\left|f(\theta)^2 + 2f'(\theta)^2 - f(\theta) f''(\theta)\right|}{\left(f(\theta)^2 + f'(\theta)^2\right)^{3/2}}.$$

Suggerimento: è sufficiente applicare la formula (6.8), ricordando che l'equazione polare equivale alle equazioni parametriche $\mathrm{r}(\theta) = (f(\theta)\cos\theta, f(\theta)\sin\theta)$.

19 Applicare la formula ottenuta nell'esercizio precedente per calcolare la curvatura di una *spirale di Archimede*:

$$f(\theta) = A\theta \text{ con } \theta \in [0, \infty) \text{ e } A > 0 \text{ costante fissata.}$$

Lo stesso per la *spirale logaritmica*:

$$f(\theta) = Ae^{B\theta} \text{ con } \theta \in (-\infty, \infty) \text{ e } A, B > 0 \text{ costanti fissate.}$$

7 COMPLEMENTI

7.1 Lunghezza di una curva regolare

Forniremo ora una dimostrazione del Teorema 2.4 (paragrafo 4.1).

DIMOSTRAZIONE. Sia $\mathcal{P} = \{t_0 = a, t_1, \ldots, t_{n-1}, t_n = b\}$ una partizione dell'intervallo $[a, b]$. Essendo r' continua in $[a, b]$, dal Teorema 2.2 e il Lemma 2.3, si ha, per ogni $j = 1, \ldots, n$:

$$\left|\mathrm{r}(t_j) - \mathrm{r}(t_{j-1})\right| = \left|\int_{t_{j-1}}^{t_j} \mathrm{r}'(t)\, dt\right| \leq \int_{t_{j-1}}^{t_j} \left|\mathrm{r}'(t)\right| dt.$$

Pertanto, indicando con $l(\mathcal{P})$ la lunghezza della corrispondente poligonale, abbiamo:

$$l(\mathcal{P}) = \sum_{j=1}^{n} \left|\mathrm{r}(t_j) - \mathrm{r}(t_{j-1})\right| \leq \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} \left|\mathrm{r}'(t)\right| dt = \int_a^b \left|\mathrm{r}'(t)\right| dt.$$

Essendo $\mathcal{P}$ arbitrariamente scelta, deduciamo che:

$$l(\gamma) = \sup_{\mathcal{P}} l(\mathcal{P}) \leq \int_a^b \left|\mathrm{r}'(t)\right| dt$$

e quindi la curva γ risulta rettificabile.

Per terminare la dimostrazione occorre provare che

$$\int_a^b \left|\mathrm{r}'(t)\right| dt \leq l(\gamma). \tag{7.1}$$

Fissato arbitrariamente $\varepsilon > 0$, mostreremo che

$$\int_a^b \left|\mathrm{r}'(t)\right| dt \leq l(\gamma) + 2\varepsilon(b - a). \tag{7.2}$$

Passando al limite per $\varepsilon \to 0$, si ottiene la (7.1). Per mostrare la (7.2), osserviamo che r' è uniformemente continua, per il teorema di Cantor-Heine[13]. Si può quindi scegliere $\delta = \delta(\varepsilon)$ tale che, se $t, s \in [a, b]$ e $|t - s| \leq \delta$, allora $\left|\mathrm{r}'(t) - \mathrm{r}'(s)\right| \leq \varepsilon$.

Sia ora $\mathcal{P} = \{t_0 = a, t_1, \ldots, t_{n-1}, t_n = b\}$ una partizione di $[a, b]$ tale che

$$\max_{j=1,\ldots,n} |t_j - t_{j-1}| \leq \delta.$$

Allora, se $t \in [t_{j-1}, t_j]$ avremo che $\left|\mathrm{r}'(t) - \mathrm{r}'(t_{j-1})\right| < \varepsilon$, che implica a sua volta

$$\left|\mathrm{r}'(t)\right| \leq \left|\mathrm{r}'(t_{j-1})\right| + \varepsilon.$$

Quindi, possiamo scrivere la seguente catena di disuguaglianze:

$$\int_{t_{j-1}}^{t_j} \left|\mathrm{r}'(t)\right| dt \leq \int_{t_{j-1}}^{t_j} \left(\left|\mathrm{r}'(t_{j-1})\right| + \varepsilon\right) dt = \left|\mathrm{r}'(t_{j-1})\right|(t_j - t_{j-1}) + \varepsilon(t_j - t_{j-1}) =$$

$$= \left|\int_{t_{j-1}}^{t_j} \mathrm{r}'(t_{j-1})\, dt\right| + \varepsilon(t_j - t_{j-1}) =$$

(aggiungendo e togliendo $\mathrm{r}'(t)$)

$$= \left|\int_{t_{j-1}}^{t_j} \left(\mathrm{r}'(t) - \mathrm{r}'(t) + \mathrm{r}'(t_{j-1})\right) dt\right| + \varepsilon(t_j - t_{j-1}) \leq$$

(per la disuguaglianza triangolare)

$$\leq \left|\int_{t_{j-1}}^{t_j} \mathrm{r}'(t)\, dt\right| + \left|\int_{t_{j-1}}^{t_j} \left(\mathrm{r}'(t_{j-1}) - \mathrm{r}'(t)\right) dt\right| + \varepsilon(t_j - t_{j-1}) \leq$$

[13] Più precisamente, applichiamo il teorema di Cantor-Heine (v. volume 1, capitolo 6, Complementi, teorema 6.14) a ciascuna componente $r_i'(t)$ della funzione a valori vettoriali $\mathrm{r}'(t)$ e ne deduciamo l'uniforme continuità anche di quest'ultima.

 © 978-88-08-06485-1

(per il Lemma 2.3)

$$\leq |\mathrm{r}(t_j) - \mathrm{r}(t_{j-1})| + \int_{t_{j-1}}^{t_j} |\mathrm{r}'(t_{j-1}) - \mathrm{r}'(t)|\, dt + \varepsilon(t_j - t_{j-1}) \leq$$

(essendo $|\mathrm{r}'(t_{j-1}) - \mathrm{r}'(t)| < \varepsilon$)

$$\leq |\mathrm{r}(t_j) - \mathrm{r}(t_{j-1})| + 2\varepsilon(t_j - t_{j-1}).$$

Abbiamo dunque:

$$\int_a^b |\mathrm{r}'(t)|\, dt = \sum_{j=1}^{n} \int_{t_{j-1}}^{t_j} |\mathrm{r}'(t)|\, dt \leq \sum_{j=1}^{n} \{|\mathrm{r}(t_j) - \mathrm{r}(t_{j-1})| + 2\varepsilon(t_j - t_{j-1})\} =$$
$$= l(\mathcal{P}) + 2\varepsilon(b-a) \leq l(\gamma) + 2\varepsilon(b-a)$$

che è la (7.2). ◆

7.2 Alcune applicazioni fisiche notevoli

1 **Moto centrale.** Supponiamo che una particella si muova nello spazio soggetta ad un'accelerazione radiale, ovvero tale che se $\mathrm{r} = \mathrm{r}(t)$ è l'equazione della traiettoria, si abbia r parallelo a r''. Per esempio: nel moto di un pianeta attorno al sole, se scegliamo un sistema di riferimento col sole nell'origine, la forza d'attrazione del sole sul pianeta è parallela ad r (posizione del pianeta); d'altro canto tale forza, per la legge fondamentale della dinamica, è proporzionale all'accelerazione r''; dunque r è parallelo a r''.

Sotto queste ipotesi:

a) si dimostri che $\mathrm{r} \times \mathrm{r}'$ è un vettore h costante;

b) se ne deduca che la traiettoria è piana e giace in un piano perpendicolare a h.

2 **Deduzione della prima legge di Keplero.** Questo esempio approfondisce il precedente. La legge di gravitazione universale stabilisce che per un pianeta in moto attorno al sole vale la legge:

$$\mathrm{r}''(t) = -\frac{GM}{r^3(t)}\mathrm{r}(t)$$

dove: $\mathrm{r}(t)$ è la posizione del pianeta, in un sistema di riferimento che ha il sole nell'origine, $r = |\mathrm{r}|$, M è la massa del sole, G la costante di gravitazione universale. Si vuole dedurre che il pianeta si muoverà lungo un'ellisse che ha uno dei due fuochi nell'origine (quindi nel sole). Si segua la seguente traccia:

1. In base all'esercizio precedente (o ripercorrendone il ragionamento), si vede che $\mathrm{r} \times \mathrm{r}' = \mathrm{h}$ =vettore costante. Si scelga il sistema di riferimento in modo che h sia parallelo e concorde all'asse z.

2. Si ponga ora $\mathrm{r} = r\mathrm{u}$, ossia si definisca il versore variabile $\mathrm{u}(t)$ =vers$(\mathrm{r}(t))$. Si dimostri che:

$$\mathrm{r}'' \times \mathrm{h} = -GM\mathrm{u} \times (\mathrm{u} \times \mathrm{u}').$$

3. Sfruttando il fatto che u è un versore, e quindi è perpendicolare a u', si provi che:

$$\mathrm{u} \times (\mathrm{u} \times \mathrm{u}') = -\mathrm{u}'.$$

Quindi:

$$\mathrm{r}'' \times \mathrm{h} = GM\mathrm{u}'.$$

© 978-88-08-06485-1

4. Poiché h è costante, $\mathrm{r}'' \times \mathrm{h} = (\mathrm{r}' \times \mathrm{h})'$. Allora il punto 3 implica:

$$\mathrm{r}' \times \mathrm{h} = GM\mathrm{u} + \mathrm{c}, \tag{7.3}$$

per un certo vettore costante c. Scegliamo, nel piano x, y in cui avviene il moto, gli assi x, y in modo che c sia parallelo e concorde con l'asse x. Sia θ l'angolo formato da r con c, cioè con l'asse x.

5. Facendo prodotto scalare di ambo i membri della (7.3) con r, dedurre che:

$$\mathrm{r} \cdot \mathrm{r}' \times \mathrm{h} = r(GM + c\cos\theta)$$

con $c = |\mathrm{c}|$.

6. Usando le proprietà del prodotto misto, provare che

$$\mathrm{r} \cdot \mathrm{r}' \times \mathrm{h} = |\mathrm{h}|^2 \equiv h^2 = \text{costante}.$$

Di conseguenza

$$r(\theta) = \frac{h^2}{GM + c\cos\theta}$$

è l'equazione in forma polare della legge oraria del pianeta.

7. Riconoscere dal punto precedente che la traiettoria è una conica (confrontare con l'equazione (3.9) nel paragrafo 3), e quindi, dovendo essere chiusa, è un'ellisse, con un fuoco nell'origine.

OSSERVAZIONE Questo procedimento serve a ottenere la curva percorsa dal pianeta, ma non la sua legge oraria: non abbiamo calcolato r in funzione di t, ma $|\mathrm{r}|$ in funzione di θ.

Notiamo anche che il vettore $\mathrm{r} \times \mathrm{r}'$, che si è provato essere costante, è (a meno di una costante moltiplicativa che rappresenta la massa del pianeta) il vettore *momento angolare* del pianeta rispetto al sole, che infatti si mantiene costante. La relazione $\mathrm{r} \times \mathrm{r}' =$ costante, dunque, non è solo un passaggio nel calcolo analitico, ma un punto fisicamente significativo dell'argomentazione ed è in relazione con la *seconda legge di Keplero* (costanza della velocità areolare), nei cui dettagli non entriamo.

3 Calcolo differenziale per funzioni reali di più variabili

Cominciamo a sviluppare il calcolo infinitesimale per le funzioni reali di più variabili, ossia $f : \mathbb{R}^n \to \mathbb{R}$. Per prendere confidenza con questi nuovi oggetti, inizieremo da qualche considerazione elementare sulle possibili rappresentazioni grafiche delle funzioni di due o tre variabili.

1 GRAFICI E INSIEMI DI LIVELLO

Ricordiamo che il *grafico* di una funzione reale di variabile reale, $y = f(x)$, è l'insieme dei punti del piano $\mathbb{R}^2$ di coordinate $(x, f(x))$. Analogamente, il grafico di una *funzione reale di più variabili reali*

$$f : A \subseteq \mathbb{R}^n \to \mathbb{R}$$
$$z = f(\mathrm{x})$$

(dove ora x indica un elemento di $\mathbb{R}^n$ e A il dominio di f) è l'insieme dei punti di $\mathbb{R}^{n+1}$ di coordinate $(\mathrm{x}, f(\mathrm{x}))$. Per $n = 2$ questo grafico "vive" nello spazio tridimensionale e può essere effettivamente visualizzato. Qualche esempio di grafico di funzioni di due variabili è mostrato nelle figure 3.2b-3.5b. Per $n \geq 3$ il grafico vive in uno spazio di dimensione ≥ 4 e purtroppo non si può visualizzare direttamente.

C'è un altro modo di rappresentare graficamente una funzione $z = f(x, y)$ ed è quello di tracciare le sue *linee di livello*. Si pensi alla superficie grafico di f come la superficie terrestre in una regione montuosa. Le linee di livello sono allora quelle che si tracciano nelle carte topografiche. Un grafico a curve di livello è un disegno nel piano in cui si tracciano le linee lungo le quali f ha valore costante, per un insieme sufficientemente fitto di valori di f (nel caso delle carte topografiche, il valore di f è la quota sul livello del mare e le linee di livello rappresentano, per esempio, le quote $f = 100$ m, $f = 150$ m, $f = 200$ m ecc.). È chiaro che un grafico a curve di livello contiene le istruzioni per costruire il grafico di f (sia pur in modo approssimato). Un altro esempio di utilizzo delle linee di livello si trova nelle previsioni del tempo: il grafico delle linee *isobare*, ossia delle linee di livello della funzione di due variabili "pressione atmosferica nel punto di coordinate (x, y), al livello del mare" (fig. 3.1).

 © 978-88-08-06485-1

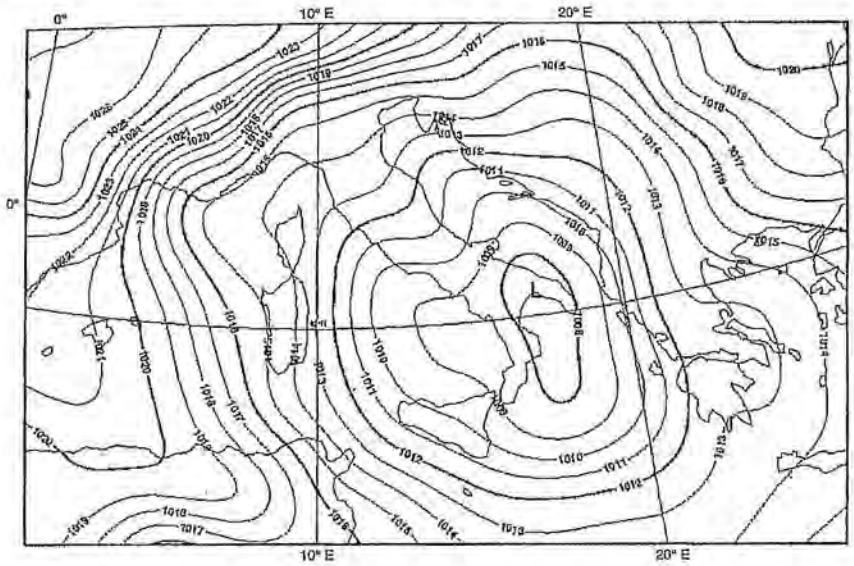


Figura 3.1.

Matematicamente, quindi, le linee di livello sono definite da un'equazione del tipo:

$$f(x,y) = k, \quad \text{con } k \text{ costante.}$$

Se queste linee sono tracciate per un insieme di valori k equispaziati (per esempio $k = 0.1, k = 0.2, k = 0.3, \ldots$), esse saranno più fitte là dove il grafico della funzione è più ripido e più distanziate là dove il grafico della funzione è più pianeggiante.

Esempi

1 La funzione

$$f(x,y) = x^2 + y^2$$

è definita in tutto il piano, è sempre ≥ 0, ha linee di livello $x^2 + y^2 = c$. Per esempio, per $c = 1, 2, 3, \ldots$, le linee di livello sono le circonferenze centrate nell'origine e raggio $1, \sqrt{2}, \sqrt{3}, \ldots$ (fig. 3.2a). Da queste linee di livello leggiamo che: la funzione f ha simmetria radiale (ossia: la funzione ha lo stesso valore nei punti che hanno la stessa distanza dall'origine) e, allontanandosi dall'origine, cresce sempre più velocemente (le linee di livello diventano più dense). Il grafico della funzione sarà quindi del tipo in figura 3.2b (paraboloide).

2 La funzione

$$f(x,y) = \sqrt{x^2 + y^2}$$

è definita in tutto il piano, è sempre ≥ 0, ha linee di livello $\sqrt{x^2 + y^2} = c$. Per esempio, per $c = 1, 2, 3, \ldots$, le linee di livello sono le circonferenze centrate nell'origine e raggio $1, 2, 3, \ldots$(fig. 3.3a). Anche questa funzione ha dunque simmetria radiale; a differenza della precedente, tuttavia, allontanandosi dall'origine cresce a ritmo costante (le linee di livello sono equispaziate), cioè linearmente. Se facciamo una sezione verticale del grafico di $f(x,y)$,

© 978-88-08-06485-1

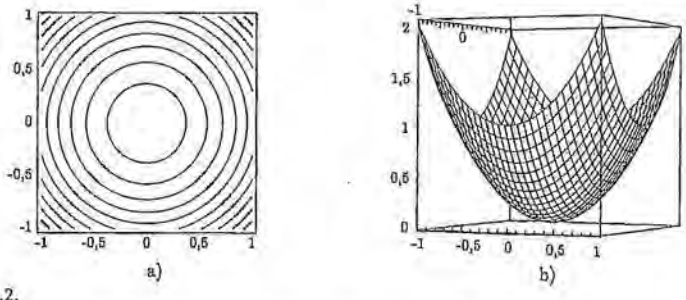


Figura 3.2.

intersecandolo col piano $y = 0$, otteniamo la curva $z = \sqrt{x^2} = |x|$, che ha un punto angoloso nell'origine. La funzione $z = f(x,y)$ è il cono in fig. 3.3b.

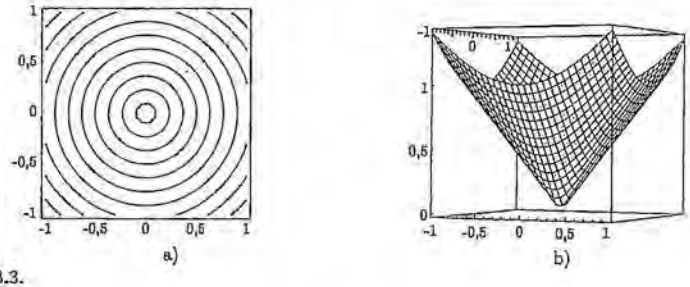


Figura 3.3.

3 La funzione

$$f(x,y) = x^2 - y^2$$

è definita in tutto il piano e assume valori sia positivi che negativi. Le sue linee di livello sono le curve $x^2 - y^2 = c$. Per $c = 0$, esse rappresentano le 2 rette $y = \pm x$; per gli altri valori di c, sono iperboli equilatere aventi le rette $y = \pm x$ come asintoti, precisamente: se $c > 0$, sono iperboli coi vertici sull'asse x, se $c < 0$ sono iperboli coi vertici sull'asse y (fig. 3.4a). Se ne deduce che il grafico della funzione, che si chiama "paraboloide iperbolico", ha la tipica forma a sella illustrata in fig. 3.4b.

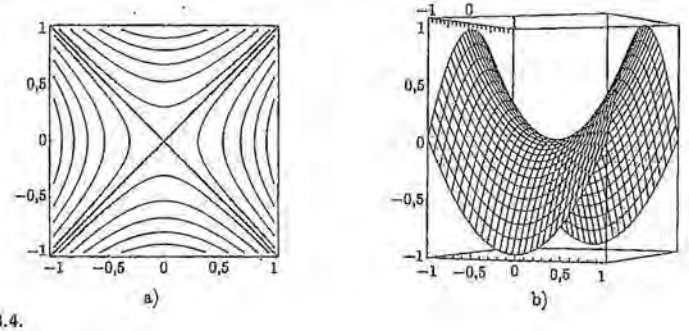


Figura 3.4.

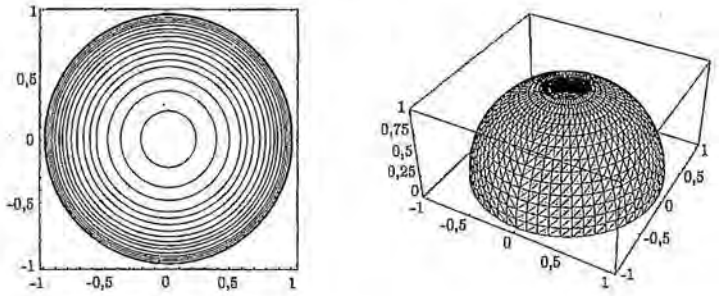
ESEMPIO La funzione

$$f(x,y) = \sqrt{1-x^2-y^2}$$

è definita solo nel cerchio $x^2+y^2 \leq 1$ ed è ≥ 0. Le sue curve di livello sono le curve:

$$\sqrt{1-x^2-y^2} = c, \text{ ossia } x^2+y^2 = 1-c^2.$$

Si tratta quindi di circonferenze di centro l'origine e raggio $\sqrt{1-c^2}$, purché $0 \leq c \leq 1$ (fig. 3.5a). Anche questa è dunque una funzione radiale, che si annulla sul bordo del cerchio su cui è definita e, come segnalato dalle linee di livello, è più ripida al bordo del cerchio. Il grafico è la superficie di una semisfera, come si capisce anche riscrivendo l'equazione $z = \sqrt{1-x^2-y^2}$ nella forma $x^2+y^2+z^2 = 1, z \geq 0$ (fig. 3.5b).



a) b)

Figura 3.5.

Consideriamo ora una funzione di più di due variabili: $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ con $n > 2$. In questo caso, come già osservato, il "grafico" di f non si può visualizzare.

Una (parziale) visualizzazione di una funzione di tre variabili si ottiene mediante le *superfici di livello*, concetto analogo a quello di linee di livello, in una dimensione in più.

Esempio

FISICA Sia $u : \mathbb{R}^3 \to \mathbb{R}$ il potenziale elettrostatico generato nello spazio da un certo sistema di cariche elettriche. In condizioni di equilibrio elettrostatico, il potenziale non dipende dal tempo, perciò è una funzione delle sole tre variabili spaziali. Il luogo dei punti dello spazio in cui $u(x,y,z) =$ costante è una superficie, che si dice *superficie equipotenziale*. Per esempio, nel caso del potenziale generato da una sola carica puntiforme posta nell'origine si ha:

$$u(x,y,z) = \frac{k}{\sqrt{x^2+y^2+z^2}}$$

dove k è una costante che dipende dalle unità di misura scelte; le superfici equipotenziali hanno equazione:

$$\frac{k}{\sqrt{x^2+y^2+z^2}} = \text{costante}$$

cioè

$$x^2+y^2+z^2 = \text{costante (positiva)}$$

e sono quindi superfici sferiche centrate nell'origine.

Generalizzando l'ultimo esempio, diciamo che le superfici di livello di una funzione $f(x,y,z)$ sono definite da un'equazione del tipo $f(x,y,z) = c$, con c costante.

Analogamente, per una funzione $f : \mathbb{R}^n \to \mathbb{R}$ (con n qualunque), si chiamano *insiemi di livello* (o *ipersuperfici di livello*) gli insiemi del tipo

$$\{\mathbf{x} \in \mathbb{R}^n : f(\mathbf{x}) = c\},$$

al variare della costante c. Nel paragrafo 8 considereremo *funzioni definite implicitamente da un'equazione* e saremo più rigorosi su questo argomento.

Esempi

ESEMPIO Le superfici di livello di $f(x,y,z) = x^2+y^2+z$ sono definite dall'equazione $x^2+y^2+z = c$, che si può riscrivere come $z = c-x^2-y^2$. Queste superfici sono dunque paraboloidi ellittici.

ESEMPIO Le superfici di livello di $f(x,y,z) = (x+y)^2-z^2$ sono definite da $(x+y)^2-z^2 = c$. Si noti che per $c = 0$, in particolare, la "superficie" è in realtà l'unione dei due piani $z = \pm(x+y)$.

2 LIMITI E CONTINUITÀ PER FUNZIONI DI PIÙ VARIABILI

2.1 Definizioni e proprietà di limiti e funzioni continue

Anche per le funzioni reali di più variabili si tratta ora di definire e studiare la nozione di limite, su cui si baseranno tutti i successivi concetti del calcolo infinitesimale. Cominciamo col dare la definizione di limite di una successione in $\mathbb{R}^n$, analoga a quella di limite di una funzione $\mathbf{r} : \mathbb{R} \to \mathbb{R}^n$ data nel capitolo 2:

Definizione 3.1 (limite di successione) Data una successione $\{\mathbf{x}_k\}_{k=1}^{\infty}$ di punti di $\mathbb{R}^n$ e un punto $\mathbf{x}_0 \in \mathbb{R}^n$, si dice che

$$\mathbf{x}_k \to \mathbf{x}_0 \text{ per } k \to \infty \text{ se } |\mathbf{x}_k - \mathbf{x}_0| \to 0 \text{ per } k \to \infty.$$

Per introdurre la definizione di limite di una funzione abbiamo bisogno di definire gli "intorni". In particolare, gli intorni sferici di un punto $\mathbf{x}_0$ sono le (iper)sfere centrate in $\mathbf{x}_0$. Precisamente, dato un punto $\mathbf{x}_0 \in \mathbb{R}^n$, si dice *intorno sferico di* $\mathbf{x}_0$ un insieme del tipo:

$$U_r(\mathbf{x}_0) = \{\mathbf{x} \in \mathbb{R}^n : |\mathbf{x}-\mathbf{x}_0| < r\}$$

per qualche $r > 0$, che si dice *raggio dell'intorno*; $\mathbf{x}_0$ si dice anche *centro dell'intorno*.

Definizione 3.2 (definizione successionale di limite di funzione) Sia $f : \mathbb{R}^n \to \mathbb{R}$ definita almeno in un intorno sferico di $\mathbf{x}_0 \in \mathbb{R}^n$ (escluso al più $\mathbf{x}_0$ stesso) e sia[1] $L \in \mathbb{R}^*$. Diremo allora che

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = L$$

[1] Ricordiamo che $L \in \mathbb{R}^*$ significa che $L \in \mathbb{R}$ oppure $L = \pm\infty$.

 © 978-88-08-06485-1

se, per ogni successione $\{x_k\}_{k=1}^{\infty}$ di punti di $\mathbb{R}^n$ tale che $x_k \to x_0$ per $k \to \infty$ (con $x_k \neq x_0$ $\forall k$), si ha che

$$\lim_{k\to\infty} f(x_k) = L.$$

Come nel caso di funzioni di una variabile, si può dare una definizione più diretta:

DEFINIZIONE 3.3 ($\varepsilon - \delta$ DEFINIZIONE DI LIMITE DI FUNZIONE) Sia $f : \mathbb{R}^n \to \mathbb{R}$ definita almeno in un intorno sferico di $x_0 \in \mathbb{R}^n$ (escluso al più x_0 stesso) e sia $L \in \mathbb{R}$. Si dice che

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

se, per ogni $\varepsilon > 0$, esiste un $\delta > 0$ tale che $0 < |x - x_0| < \delta$ implica $|f(x) - L| < \varepsilon$; si dice che

$$\lim_{x\to x_0} f(x) = +\infty \ (o\ -\infty)$$

se, per ogni $k > 0$, esiste un $\delta > 0$ tale che $0 < |x - x_0| < \delta$ implica $f(x) > k$ (rispettivamente, $f(x) < -k$).

Geometricamente, la definizione (nel caso L finito) significa che il valore di $f(x)$ approssima tanto quanto vogliamo il valore L, purché la distanza del punto x dal punto x_0 sia piccola quanto occorre.

Si rifletta subito su un'importante differenza che esiste, a questo riguardo, tra funzioni $f : \mathbb{R} \to \mathbb{R}^m$ e funzioni $f : \mathbb{R}^n \to \mathbb{R}$. Abbiamo visto che assegnare una funzione $f : \mathbb{R} \to \mathbb{R}^m$ è equivalente ad assegnare m funzioni $f_i : \mathbb{R} \to \mathbb{R}$; questo fatto sta alla base della possibilità di calcolare i limiti delle funzioni a valori vettoriali "componente per componente". Invece, una generica funzione $f : \mathbb{R}^n \to \mathbb{R}$ non è "scomponibile" in alcun modo in n funzioni reali di variabile reale; di conseguenza, anche il calcolo dei limiti per funzioni di più variabili non si riduce in modo naturale al calcolo unidimensionale. Nel prossimo paragrafo 2.2 vedremo più da vicino il tipo di difficoltà e di situazioni nuove che si possono creare.

Dal punto di vista formale, comunque, la definizione di limite per funzioni di più variabili è simile a quella data per funzioni reali di variabile reale. In particolare, molte definizioni e proprietà riguardanti i limiti di funzioni di più variabili si possono enunciare e dimostrare in modo perfettamente analogo al caso unidimensionale. Per esempio, per funzioni $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ continuano a valere inalterati nell'enunciato:

- *il teorema di unicità del limite*;
- *il teorema sul limite della somma, del prodotto per una costante, del prodotto e del quoziente di due funzioni*;
- *il teorema del confronto.*
- *La definizione di funzione continua*, in un punto o in un insieme, è analoga al caso unidimensionale:

  si dice che $f : \mathbb{R}^n \to \mathbb{R}$ è continua in x_0 se $\lim_{x\to x_0} f(x) = f(x_0)$.

© 978-88-08-06485-1

Come conseguenza dei teoremi sui limiti, valgono i teoremi sulla *continuità della somma, del prodotto e del quoziente di funzioni continue* (quando hanno senso e il denominatore non si annulla) e della *composizione di funzioni continue* (quando ha senso).

Se f è una funzione reale di variabile reale, continua, risulta continua anche se la consideriamo come funzione di più variabili. (Il lettore è invitato a dimostrare questo semplice asserto in base alle definizioni precedenti).

Quest'ultima affermazione significa, per esempio, che

$$f(x, y) = \sin x \qquad g(x, y) = \sin y$$

considerate come funzioni "di due variabili" sono continue, perché $t \mapsto \sin t$ è una funzione continua di una variabile.

Le proprietà ora enunciate si possono combinare tra loro, permettendo di mostrare la continuità di un gran numero di funzioni, senza dover ricorrere alla definizione di continuità, cioè al calcolo di limiti.

Esempio

La funzione

$$f(x, y) = \frac{\sin y}{1 + x^2}$$

è continua in tutto $\mathbb{R}^2$. Infatti: numeratore e denominatore sono funzioni continue di una variabile, il denominatore non si annulla mai, per cui il quoziente è continuo.

Nel seguito, naturalmente, non espliciteremo ogni volta questi ragionamenti. Quando, applicando i teoremi sopra enunciati, possiamo essere certi che una funzione è continua, per calcolare $\lim_{x\to x_0} f(x)$ sarà sufficiente valutare $f(x_0)$. Come nel caso unidimensionale, sarà solo nel caso di "forme di indeterminazione" che occorrerà ragionare sul limite in base alla definizione.

Enunciamo esplicitamente il teorema di permanenza del segno per funzioni $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, nelle sue varie forme, perché in seguito ci sarà utile:

TEOREMA 3.1 (DI PERMANENZA DEL SEGNO) *Sia $f : \mathbb{R}^n \to \mathbb{R}$ definita almeno in un intorno sferico di $x_0 \in \mathbb{R}^n$ (salvo al più x_0 stesso). Supponiamo che esista*

$$\lim_{x\to x_0} f(x) = L \in \mathbb{R}^*.$$

1. *Se $L > 0$, allora $f(x)$ si mantiene positiva almeno in un intorno di x_0, cioè esiste $\delta > 0$ tale che $f(x) > 0$ purché $0 < |x - x_0| < \delta$.*
2. *Se $f(x) \geq 0$ in un intorno di x_0 (salvo al più x_0 stesso) allora $L \geq 0$. Notiamo che non si può affermare che $L > 0$, anche se $f(x) > 0$!*

   Dal punto 1 segue anche che:
3. *Se f è continua in x_0 e $f(x_0) > 0$, allora $f(x)$ si mantiene positiva almeno in un intorno di x_0, cioè esiste $\delta > 0$ tale che $f(x) > 0$ purché $|x - x_0| < \delta$.*

Analoghi enunciati valgono ovviamente sostituendo il segno $>$ con $<$ e il segno $\geq$ con $\leq$.

Terminiamo il paragrafo con la definizione di limite di una funzione all'infinito. Mentre, per una funzione f di una variabile, abbiamo dato nel volume 1 la definizione di

$$\lim_{x\to\pm\infty} f(x),$$

nel caso di funzioni di n variabili non è così ovvio cosa significhi che la variabile indipendente x tende a infinito: anche nel caso di funzioni di due sole variabili (x, y), è chiaro che un punto del piano può "andare all'infinito" lungo infinite direzioni diverse e la funzione $f(x, y)$ potrebbe avere comportamenti diversi, a priori, in ciascuna di queste direzioni. La definizione che daremo, invece, richiede che la funzione abbia un comportamento uniforme, al tendere di $|\mathrm{x}|$ a $+\infty$, indipendentemente dalla direzione. Precisamente:

DEFINIZIONE 3.4 Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione definita in tutto lo spazio, o almeno per $|\mathrm{x}|$ abbastanza grande. Si dice che

$$\lim_{|\mathrm{x}|\to\infty} f(\mathrm{x}) = L \in \mathbb{R} \text{ se}$$

$$\forall \varepsilon > 0\ \exists R > 0 \text{ tale che } \forall \mathrm{x} \in \mathbb{R}^n, \text{ se } |\mathrm{x}| > R \text{ allora } |f(\mathrm{x}) - L| < \varepsilon.$$

Si dice che

$$\lim_{|\mathrm{x}|\to\infty} f(\mathrm{x}) = +\infty \qquad (\text{o } -\infty) \text{ se}$$

$$\forall K > 0\ \exists R > 0 \text{ tale che } \forall \mathrm{x} \in \mathbb{R}^n, \text{ se } |\mathrm{x}| > R \text{ allora } f(\mathrm{x}) > K$$

(rispettivamente $f(\mathrm{x}) < -K$). ▩

2.2 Calcolo dei limiti in più variabili: analisi delle forme di indeterminazione

Presenteremo ora, mediante esempi, alcune idee fondamentali per l'analisi delle forme di indeterminazione nel calcolo dei limiti delle funzioni di più variabili. Premettiamo un'osservazione sulla definizione di limite. Supponiamo che f abbia limite L finito per $\mathrm{x} \to \mathrm{x}_0$. L'essenza della definizione sta nel fatto che $f(\mathrm{x})$ si avvicina indefinitamente a L quando la *distanza* tra x e x_0 tende a zero, indipendentemente dalla *direzione* con cui x si avvicina a x_0. I seguenti esempi illustrano due tecniche tipiche che si usano per dimostrare, rispettivamente, l'esistenza o la non esistenza di un limite.

Restrizione di una funzione a una curva e non esistenza del limite

Se $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ è una funzione reale di n variabili, $\mathrm{r} : I \subseteq \mathbb{R} \to \mathbb{R}^n$ è un arco di curva in $\mathbb{R}^n$ ed esiste la funzione composta

$$g(t) = f(\mathrm{r}(t)),$$

questa si dice *restrizione di f alla curva* r ed è una funzione reale di variabile reale. Il termine restrizione, che denota questo tipo di composizione, deriva dalla seguente idea geometrica: invece di far variare x in ogni modo nel dominio n-dimensionale in

cui è definita f, ci restringiamo ai punti di $\mathbb{R}^n$ che stanno sull'arco di curva $\mathrm{r}(t)$. È chiaro che, se l'arco di curva è continuo e f è continua, anche la sua restrizione g all'arco di curva sarà continuo (continuità della funzione composta). Essendo funzioni di una variabile, le restrizioni sono facili da studiare; l'idea è di ottenere informazioni sul comportamento di f esaminando le sue restrizioni a curve differenti.

In particolare, per mostrare che il limite per $\mathrm{x} \to \mathrm{x}_0$ di una certa funzione $f(\mathrm{x})$ *non esiste* è sufficiente determinare due curve che passano da x_0, lungo le quali la funzione tende a due limiti diversi. La stessa conclusione vale se la restrizione di $f(\mathrm{x})$ a *una* particolare curva non ammette limite.

Esempio

Calcolare

$$\lim_{(x,y)\to(0,0)} \frac{xy}{x^2+y^2}.$$

Il limite non esiste. Infatti, la restrizione di $f(x, y)$ alla retta $y = x$ è

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

perciò tende a $\frac{1}{2}$. Invece, la restrizione di $f(x, y)$ alla retta $y = -x$ è

$$f(x, -x) = \frac{-x^2}{2x^2} = -\frac{1}{2},$$

perciò tende a $-\frac{1}{2}$. Essendo i due limiti diversi, il limite di f non esiste.

Questo è il metodo comunemente seguito per dimostrare che una funzione di più variabili non ammette limite. Si badi che, per questa via, non si può invece dimostrare l'*esistenza* del limite. Per esempio, l'esistenza non è garantita dal fatto che lungo qualsiasi retta uscente dall'origine f abbia lo stesso limite, come mostreranno altri esempi.

Uso di maggiorazioni con funzioni radiali per provare l'esistenza del limite

Spesso col passaggio a coordinate polari (nel caso bidimensionale) si riesce a mettere in evidenza la dipendenza di $f(x, y)$ dalla distanza tra (x, y) e $(0, 0)$ attraverso $\rho = \sqrt{x^2+y^2}$.

Esempio

Calcolare

$$\lim_{(x,y)\to(0,0)} \frac{2x^2y}{x^2+y^2}.$$

Affermiamo che questo limite esiste e vale 0. Per dimostrarlo, riscriviamo la funzione in coordinate polari:

$$\frac{2x^2y}{x^2+y^2} = \frac{2\rho^3\cos^2\theta\sin\theta}{\rho^2} = 2\rho\cos^2\theta\sin\theta.$$

Poiché $|\cos\theta|, |\sin\theta| \leq 1$, si può scrivere la maggiorazione:

$$\left|\frac{2x^2y}{x^2+y^2}\right| = 2\rho\left|\cos^2\theta\sin\theta\right| \leq 2\rho.$$

 © 978-88-08-06485-1

Poiché $|f(x,y)|$ è compreso tra 0 e 2ρ, la funzione è arbitrariamente vicina a zero quando ρ, cioè la distanza tra (x,y) e $(0,0)$, è sufficientemente piccolo. Ne segue che il limite di f è 0, proprio per definizione di limite.

Questo esempio contiene l'idea di un *criterio valido in generale per provare l'esistenza di un limite*. Enunciamolo prima nel caso bidimensionale e poi in generale:

Per dimostrare che $f(x,y) \to L$ per $(x,y) \to (0,0)$, è sufficiente riuscire a scrivere una maggiorazione del tipo

$$|f(\rho,\theta) - L| \le g(\rho) \text{ dove } g(\rho) \to 0 \text{ per } \rho \to 0.$$

L'essenziale è che la funzione g *non dipenda da* θ. Più in generale, se il punto (x,y) tende a (x_0,y_0), si applica lo stesso criterio con $\rho = \sqrt{(x-x_0)^2+(y-y_0)^2}$, cioè si pone:

$$\begin{cases} x = x_0 + \rho\cos\theta \\ y = y_0 + \rho\sin\theta. \end{cases}$$

Naturalmente, *non riuscire* a dimostrare una maggiorazione del genere non dimostra che il limite non esiste!

Nel caso di una funzione di n variabili, il criterio assume questa forma:

TEOREMA 3.2 *Sia $f : \mathbb{R}^n \to \mathbb{R}$ definita almeno in un intorno di x_0 (salvo al più x_0 stesso) e sia $L \in \mathbb{R}$. Se $g : (0,+\infty) \to \mathbb{R}$ è una funzione tale che $g(\rho) \to 0$ per $\rho \to 0$ e*

$$|f(\mathrm{x}) - L| \le g(|\mathrm{x} - \mathrm{x}_0|)$$

per ogni x in un opportuno intorno sferico di x_0 (salvo al più x_0 stesso), allora

$$\lim_{\mathrm{x}\to\mathrm{x}_0} f(\mathrm{x}) = L.$$

L'essenziale, nell'enunciato precedente, è aver maggiorato la differenza $|f(\mathrm{x}) - L|$ con una quantità che dipende da x solo mediante la sua distanza da x_0. La dimostrazione del teorema è un'immediata conseguenza della definizione di limite.

Esempio

234 Sia $f : \mathbb{R}^n \to \mathbb{R}$,

$$f(\mathrm{x}) = \frac{x_1 x_2}{|\mathrm{x}|} e^{|\mathrm{x}|}, \text{ definita per } \mathrm{x} = (x_1, x_2, \ldots, x_n) \neq 0.$$

Calcolare

$$\lim_{\mathrm{x}\to 0} f(\mathrm{x}).$$

Scriviamo, per ogni $\mathrm{x} \neq 0$:

$$|f(\mathrm{x})| = \frac{|x_1 x_2|}{|\mathrm{x}|} e^{|\mathrm{x}|} \le |\mathrm{x}|\, e^{|\mathrm{x}|} \equiv g(|\mathrm{x}|).$$

Poiché $g(\rho) = \rho e^{\rho} \to 0$ per $\rho \to 0$, il limite cercato è 0.

235 Calcolare

$$\lim_{(x,y)\to(0,0)} \frac{x^2 y}{x^4 + y^2}.$$

© 978-88-08-06485-1

Dimostriamo che il limite non esiste. Infatti, la restrizione di f alle due curve $y = \pm x^2$ è:

$$f(x, x^2) = \frac{\pm x^2 x^2}{x^4 + x^4} = \pm\frac{1}{2},$$

pertanto lungo le due curve la funzione ha due limiti diversi e il limite in due variabili non esiste.

Notiamo che, se scrivessimo la funzione in coordinate polari,

$$f(\rho,\theta) = \frac{\rho^3 \cos^2\theta \sin\theta}{\rho^4 \cos^4\theta + \rho^2 \sin^2\theta} = \frac{\rho \cos^2\theta \sin\theta}{\rho^2 \cos^4\theta + \sin^2\theta},$$

troveremmo che, per ogni θ fissato, $f(\rho,\theta) \to 0$ per $\rho \to 0$. Questo significa che lungo ogni retta uscente dall'origine la funzione ha lo stesso limite (zero). Tuttavia il limite in due variabili non esiste.

Nei Complementi, paragrafo 9.2, studiando le proprietà delle *funzioni omogenee*, vedremo qualche altro criterio utile a decidere se una funzione è continua o meno in un punto in cui il calcolo del limite presenta una forma di indeterminazione.

Infine, illustriamo mediante esempi come si adattino al caso dei limiti per $|\mathrm{x}| \to \infty$ i criteri discussi in questo paragrafo per dimostrare l'esistenza o non esistenza del limite.

Esempio

236 Dimostriamo che:

(a) $\lim_{|(x,y)|\to\infty} xye^{-(x^2+y^2)} = 0$;

(b) $\lim_{|(x,y)|\to\infty} xye^{x^2+y^2}$ non esiste.

(a) È sufficiente passare in coordinate polari e maggiorare come segue:

$$\left|xye^{-(x^2+y^2)}\right| = \left|\rho^2 \cos\theta \sin\theta e^{-\rho^2}\right| \le \rho^2 e^{-\rho^2};$$

poiché l'ultima funzione scritta dipende solo da ρ e tende a zero per $\rho \to +\infty$, segue la tesi.
(b) Osserviamo che lungo la retta $y = x$ si ha:

$$f(x,x) = x^2 e^{2x^2} \to +\infty \text{ per } x \to \pm\infty,$$

mentre lungo la retta $y = 0$ è $f(x,0) = 0$. Dunque il limite di f per $\mathrm{x} \to \infty$ non esiste.

Esercizi

237 Delle seguenti funzioni reali di due variabili, si provi a capire com'è fatto il grafico, studiando le linee di livello ed eventualmente alcune sezioni con piani opportuni. Si controlli poi quanto previsto disegnando i grafici col computer:

$$f(x,y) = 1 - (x^2+y^2); \quad f(x,y) = xy; \quad f(x,y) = e^{-x^2-y^2};$$
$$f(x,y) = e^x \sin y; \quad f(x,y) = \sin x \sin y; \quad f(x,y) = \frac{1}{x+y}.$$

 © 978-88-08-06485-1

Per ciascuna delle seguenti funzioni, dire su quale sottoinsieme del piano si può affermare che la funzione è continua, senza necessità di calcolare limiti:

$$e^{x^2/y}; \qquad \operatorname{arctg}\left(\tfrac{1}{x^2+y^2}\right); \qquad \log\left(1+\sqrt{x^2+y^2}\right);$$

$$\frac{x^2+3yx+2}{x^2+2xy+y^2}; \quad \log([\sin(x^2+y^2)]^2); \quad \operatorname{arctg}(e^{y/x}).$$

Si provi poi a calcolare i limiti alla frontiera dell'insieme di definizione, distinguendo i casi in cui si ha una effettiva forma di indeterminazione da quelli in cui il valore del limite si può calcolare in base ai teoremi noti.

Dimostrare che

$$\lim_{(x,y)\to(0,0)} \frac{x^2+y^2}{x}$$

non esiste. *Suggerimento*: utilizzare la tecnica delle restrizioni.

Dimostrare che

$$\lim_{(x,y)\to(0,0)} \frac{\sin^2(xy)}{x^2+y^2} = 0.$$

Suggerimento: utilizzare le coordinate polari, come nell'esempio 2.3.

Dimostrare che

$$\lim_{(x,y)\to(0,0)} \frac{x^3+y^5}{x^2+y^4} = 0.$$

Dimostrare che

$$\lim_{(x,y)\to(2,1)} \frac{(y-1)^2\sin(\pi x)}{(x-2)^2+(y-1)^2} = 0.$$

Dimostrare che

$$\lim_{(x,t)\to(0,0)} \frac{e^{-x^2/t}}{\sqrt{|t|}} \text{ non esiste,}$$

mentre, per ogni $x_0 \neq 0$,

$$\lim_{(x,t)\to(x_0,0)} \frac{e^{-x^2/t}}{\sqrt{|t|}} = 0.$$

Dimostrare che, se $p_n(x,y)$ è un polinomio di grado $n \geq 1$ in x, y, si ha:

$$\lim_{|(x,y)|\to\infty} |p_n(x,y)| = +\infty.$$

Mostrare con esempi che, invece, può non esistere

$$\lim_{|(x,y)|\to\infty} p_n(x,y).$$

3 TOPOLOGIA IN $\mathbb{R}^n$ E PROPRIETÀ DELLE FUNZIONI CONTINUE

Nello studio delle funzioni reali di variabile reale (v. volume 1) abbiamo solitamente considerato funzioni definite su un intervallo, o sull'unione di più intervalli: considerare funzioni definite su insiemi più complicati sarebbe scarsamente motivato, almeno dalle applicazioni che interessano in questo corso. Non appena si considerano funzioni di più variabili, invece, la situazione cambia drasticamente, come mostra il prossimo

Esempio

Determiniamo l'insieme di definizione della funzione di due variabili:

$$f(x,y) = \frac{\sqrt{x^2-y}}{\log(x+y)}.$$

La funzione è definita nell'insieme degli (x,y) per cui si ha:

$x^2 - y \geq 0$ (radicando non negativo)
$x+y > 0$ (argomento del logaritmo positivo)
$x+y \neq 1$ (denominatore diverso da zero).

Perciò l'insieme di definizione è

$$E = \{(x,y) \in \mathbb{R}^2 : x^2 - y \geq 0, x+y > 0, x+y \neq 1\}.$$

Rimane a questo punto il problema di disegnare questo insieme (fig. 3.6) e di capire quali siano le sue proprietà rilevanti per lo studio di f.

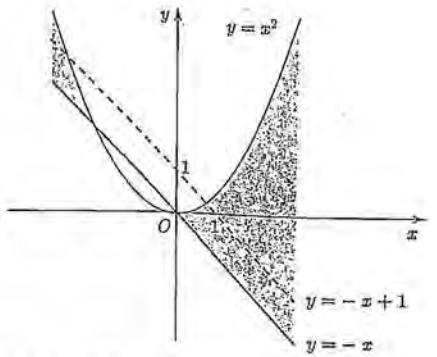


Figura 3.6: L'insieme E è la regione ombreggiata.

Come si vede già da questo semplice esempio, il tipo di sottoinsiemi di $\mathbb{R}^2$ che è naturale considerare nello studio delle funzioni di due variabili è molto vario; in particolare, notiamo che l'insieme E dell'esempio non è descrivibile come unione di un numero finito di insiemi di forma semplice (es. rettangoli), mentre nel caso di funzioni $f : \mathbb{R} \to \mathbb{R}$ spesso l'insieme di definizione era semplicemente l'unione di un numero finito di intervalli.

La necessità di considerare sottoinsiemi di $\mathbb{R}^n$ piuttosto generali porta con sé a sua volta la necessità di introdurre certe proprietà degli insiemi, il cui studio prende il nome di *topologia*. Come vedremo, le proprietà di una funzione continua (o più regolare) definita su un insieme E dipendono anche da alcune *proprietà topologiche* dell'insieme. Introduciamo quindi anzitutto alcuni concetti e proprietà che ci saranno utili.

3.1 Concetti fondamentali

La seguente definizione classifica tre tipi di punti in relazione a un dato sottoinsieme di $\mathbb{R}^n$.

DEFINIZIONE 3.5 Sia E un sottoinsieme di $\mathbb{R}^n$. Un punto $x_0 \in \mathbb{R}^n$ si dice:
interno a E, se esiste un intorno sferico[2] di x_0 contenuto in E;
esterno a E, se esiste un intorno sferico di x_0 contenuto in E^c (complementare di E);
di frontiera per E, se ogni intorno sferico di x_0 contiene almeno un punto di E e un punto di E^c.

ESEMPIO

Sia

$$E = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1 \text{ e } y > x^2\}$$

(fig. 3.7). Notiamo che:
il punto $\left(0, \frac{1}{2}\right)$ è interno a E;
il punto $(1,0)$ è esterno a E;
il punto $(0,0)$ è di frontiera per E (e non appartiene a E);
il punto $(0,1)$ è di frontiera per E (e appartiene a E).

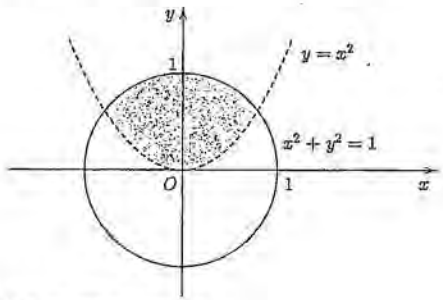


Figura 3.7. L'insieme E è la regione ombreggiata.

OSSERVAZIONE Il lettore è invitato a rendersi conto che, dalle definizioni, seguono subito i seguenti fatti:

[2] La definizione di *intorno sferico* è stata data nel paragrafo 2.1.

1. dato un insieme $E \subseteq \mathbb{R}^n$ e un punto $x_0 \in \mathbb{R}^n$, si verifica necessariamente una e una sola delle tre condizioni: x_0 è interno a E; x_0 è esterno a E; x_0 è di frontiera per E;
2. se x_0 è interno a E, necessariamente $x_0 \in E$;
3. se x_0 è esterno a E, necessariamente $x_0 \notin E$.

Dall'esempio 3.2 si osserva invece che:

4. un punto di frontiera per E può appartenere o non appartenere a E.

DEFINIZIONE 3.6 Un insieme $E \subseteq \mathbb{R}^n$ si dice:
aperto, se ogni suo punto è interno a E;
chiuso, se il suo complementare è aperto.

ESEMPIO

Consideriamo i seguenti sottoinsiemi di $\mathbb{R}^2$.

$E_1 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 < 1\}$ è aperto;
$E_2 = \{(x,y) \in \mathbb{R}^2 : y \neq 0\}$ è aperto;
$E_3 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 4\}$ è chiuso;
$E_4 = \{(x,y) \in \mathbb{R}^2 : 2x + 3y - 1 = 0\}$ è chiuso;
$E_5 = \{(x,y) \in \mathbb{R}^2 : x^2 + y^2 \leq 1 \text{ e } y < x^2\}$ non è né aperto, né chiuso;
$\mathbb{R}^2$ è aperto (quindi $\varnothing$ è chiuso);
$\varnothing$ è aperto (quindi $\mathbb{R}^2$ è chiuso).

Il lettore è invitato a rendersi conto delle precedenti affermazioni, disegnando gli insiemi in questione e applicando le definizioni. L'ultima affermazione coinvolge le proprietà un po' paradossali dell'insieme vuoto: poiché $\varnothing$ non ha punti, è lecito affermare che *ogni suo punto* è interno (infatti, quest'affermazione non può essere contraddetta).

OSSERVAZIONE Come si vede dagli esempi qui sopra, un insieme qualunque può essere né aperto né chiuso. Inoltre esistono almeno due insiemi (e in realtà solo questi due) che sono simultaneamente aperti e chiusi: il vuoto e tutto lo spazio $\mathbb{R}^n$. Notiamo anche che, poiché per definizione il complementare di un aperto è chiuso, è anche vero che *il complementare di un chiuso è aperto.*

ESEMPIO

Intervalli aperti e chiusi. Tutte le definizioni topologiche viste per sottoinsiemi di $\mathbb{R}^n$ si applicano naturalmente anche al caso $n = 1$. Il lettore è invitato a controllare che il linguaggio introdotto a suo tempo nello studio delle funzioni reali di variabile reale[3] è coerente con quello introdotto in questo capitolo:

[3] v. volume 1, capitolo 1, paragrafo 4.4.

gli intervalli (a,b), $(a,+\infty)$, $(-\infty,b)$ (con $a,b\in\mathbb{R}$) sono sottoinsiemi aperti di $\mathbb{R}$;
gli intervalli $[a,b]$, $[a,+\infty)$, $(-\infty,b]$ (con $a,b\in\mathbb{R}$) sono sottoinsiemi chiusi di $\mathbb{R}$.

Tutte le definizioni topologiche date fin qui si basano sul concetto di *intorno sferico*. Possiamo ora dare una definizione più generale di intorno, che risulterà utile nello sviluppo del calcolo differenziale.

DEFINIZIONE 3.7 (INTORNO) Dato un punto $x_0\in\mathbb{R}^n$, si dice *intorno* di x_0 un qualsiasi insieme aperto contenente x_0.

Naturalmente un intorno sferico è un particolare intorno. Inoltre, ogni intorno di x_0 contiene un intorno sferico di x_0. Utilizzando queste due osservazioni, si dimostra immediatamente la

PROPOSIZIONE 3.1 *Sia E un sottoinsieme di $\mathbb{R}^n$. Un punto $x_0\in\mathbb{R}^n$:*

- *è interno a E se e solo se esiste un intorno di x_0 contenuto in E;*
- *è esterno a E se e solo se esiste un intorno di x_0 contenuto in E^c;*
- *è di frontiera per E se ogni intorno di x_0 contiene almeno un punto di E e un punto di E^c.*

Si confronti questa Proposizione con la definizione che abbiamo dato in precedenza di punto interno, esterno e di frontiera.

A partire da insiemi aperti o chiusi, se ne possono costruire altri aperti o chiusi, mediante le operazioni insiemistiche:

TEOREMA 3.3 (OPERAZIONI INSIEMISTICHE SU INSIEMI APERTI O CHIUSI) *L'unione di una famiglia qualsiasi (anche infinita) di insiemi aperti e l'intersezione di un numero finito di insiemi aperti sono insiemi aperti.*
L'intersezione di una famiglia qualsiasi di insiemi chiusi e l'unione di un numero finito di insiemi chiusi sono insiemi chiusi.

DIMOSTRAZIONE. Proviamo il primo asserto. Sia $\{A_\alpha\}_{\alpha\in I}$ una famiglia qualsiasi[4] di sottoinsiemi aperti di $\mathbb{R}^n$ e consideriamo

$$E=\bigcup_{\alpha\in I}A_\alpha.$$

Sia $x_0\in E$. Per definizione, questo significa che $x_0\in A_{\alpha_0}$ per qualche α_0; poiché A_{α_0} è aperto, esiste un intorno sferico $U_r(x_0)\subseteq A_{\alpha_0}\subseteq E$, perciò $U_r(x_0)\subseteq E$. Questo dimostra che x_0 è interno a E. Poiché x_0 era un generico punto di E, questo prova che E è aperto.

[4] Quindi I denota un qualunque insieme di indici. Per esempio, se fosse $I=\mathbb{N}$, avremmo una successione di aperti e

$$\bigcup_{\alpha\in I}A_\alpha=\bigcup_{k=1}^{\infty}A_k,$$

ma I potrebbe anche essere finito, oppure più numeroso di $\mathbb{N}$.

Siano ora

$$F=\bigcap_{i=1}^{k}A_i$$

con $A_1,A_2,\dots,A_k$ insiemi aperti e sia $x_0\in F$. Per definizione, questo significa che $x_0\in A_i$ per ogni i; poiché A_i è aperto, esiste un intorno sferico $U_{r_i}(x_0)\subseteq A_i$; poniamo

$$r=\min\{r_i : i=1,2,\dots,k\}.$$

Si avrà allora $U_r(x_0)\subseteq U_{r_i}(x_0)\subseteq A_i$ per ogni i e quindi

$$U_r(x_0)\subseteq\bigcap_{i=1}^{k}A_i.$$

Questo mostra che x_0 è interno a F, perciò F è aperto.

Il secondo asserto (riguardo all'unione e intersezione di chiusi) si dimostra ora *per dualità*: siano $C_1,C_2,\dots,C_k$ chiusi; allora (legge di De Morgan di unione e intersezione):

$$\left(\bigcup_{i=1}^{k}C_i\right)^c=\bigcap_{i=1}^{k}C_i^c;$$

poiché C_i è chiuso, C_i^c è aperto, perciò $\bigcap_{i=1}^{k}C_i^c$ è aperto, per quanto sopra dimostrato; ma allora

$$\left(\bigcup_{i=1}^{k}C_i\right)^c \text{ è aperto, ossia}$$

$$\bigcup_{i=1}^{k}C_i \text{ è chiuso,}$$

che è quanto volevamo dimostrare. Analogamente, si dimostra che l'intersezione di una famiglia qualsiasi di chiusi è chiusa (sfruttando il fatto che l'unione di una famiglia qualsiasi di aperti è un aperto).

OSSERVAZIONE Il lettore è invitato a mostrare con esempi opportuni che:
– l'unione di una famiglia infinita di chiusi può non essere un chiuso;
– l'intersezione di una famiglia infinita di aperti può non essere un aperto.
(È sufficiente ragionare in $\mathbb{R}$, con opportune successioni di intervalli aperti o chiusi).

Abbiamo definito gli insiemi chiusi come i complementari degli aperti. È molto utile la seguente caratterizzazione alternativa degli insiemi chiusi:

TEOREMA 3.4 (INSIEMI CHIUSI E LIMITI DI SUCCESSIONI) *Sia $C\subseteq\mathbb{R}^n$. Allora C è chiuso se e solo se ha la seguente proprietà:*

per ogni successione $\{x_k\}_{k=1}^{\infty}\subseteq C$ tale che x_k converga a un certo limite $x\in\mathbb{R}^n$, si ha che $x\in C$.

In altre parole: un insieme è chiuso se e solo se contiene i limiti delle sue successioni convergenti.

DIMOSTRAZIONE. Sia prima C chiuso e proviamo che contiene i limiti delle sue successioni convergenti. Sia dunque

$$\{x_k\}_{k=1}^{\infty} \subseteq C;\ x_k \to x,$$

e, per assurdo, supponiamo che $x \notin C$. Allora $x \in C^c$, che è un insieme aperto. Sia $U_r(x)$ un intorno sferico di x contenuto in C^c. Per definizione di limite, $x_k \in U_r(x)$ definitivamente, in particolare $x_k \in C^c$ definitivamente. Ma questo è assurdo, perché $x_k \in C$ per ogni k.

Viceversa, supponiamo che C contenga i limiti delle sue successioni convergenti e proviamo che è chiuso, mostrando che C^c è aperto. Sia dunque $x \in C^c$ e mostriamo che x è interno a C^c. Per assurdo, non lo sia; questo significa che per ogni intorno sferico $U_r(x)$, questo intorno non è contenuto in C^c. Consideriamo la successione di intorni $U_{1/k}(x)$ per $k = 1, 2, 3, \ldots$; per ogni k esisterà $x_k \in U_{1/k}(x)$ tale che $x_k \notin C^c$, ossia $x_k \in C$. D'altro canto, la costruzione degli intorni mostra che $x_k \to x$ (poiché $|x_k - x| < 1/k$). Poiché, per ipotesi, C contiene i limiti delle sue successioni convergenti, $x \in C$ e otteniamo una contraddizione. Il teorema è dimostrato. ◆

I sottoinsiemi di $\mathbb{R}^n$ che ci capiterà di considerare nello studio delle funzioni di più variabili sono normalmente assegnati mediante un certo numero di *equazioni o disequazioni* (come nei casi considerati nell'Esempio 3.3). In questi casi c'è un criterio molto utile per decidere facilmente se l'insieme è aperto, o se è chiuso, senza applicare esplicitamente la definizione. Tale criterio è fornito dal seguente:

TEOREMA 3.5 (INSIEMI APERTI E CHIUSI DEFINITI DA FUNZIONI CONTINUE)
Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione definita e continua in tutto $\mathbb{R}^n$. Allora:

a) gli insiemi

$$\{x \in \mathbb{R}^n : f(x) > 0\};$$
$$\{x \in \mathbb{R}^n : f(x) < 0\};$$
$$\{x \in \mathbb{R}^n : f(x) \neq 0\}$$

sono aperti;

b) gli insiemi

$$\{x \in \mathbb{R}^n : f(x) \leq 0\};$$
$$\{x \in \mathbb{R}^n : f(x) \geq 0\};$$
$$\{x \in \mathbb{R}^n : f(x) = 0\}$$

sono chiusi.

DIMOSTRAZIONE. Sia $x_0 \in E_1 = \{x \in \mathbb{R}^n : f(x) > 0\}$ e proviamo che x_0 è interno (da cui, per la genericità di x_0, seguirà che E è aperto). Poiché f è continua, per il teorema di permanenza del segno (teorema 3.1, paragrafo 2.1), esiste un intorno sferico $U_r(x_0)$ in cui è $f(x) > 0$. Pertanto $U_r(x_0) \subseteq E$, perciò x_0 è interno.

Analogamente si prova che

$$E_2 = \{x \in \mathbb{R}^n : f(x) < 0\}$$

è aperto. Infine,

$$E_3 = \{x \in \mathbb{R}^n : f(x) \neq 0\} = E_1 \cup E_2,$$

perciò E_3 è aperto perché unione di aperti.

I tre insiemi elencati nel punto *b)* sono chiusi perché i loro complementari sono, rispettivamente, i tre insiemi del punto *a)*, che abbiamo provato essere aperti. ◆

OSSERVAZIONE Per la validità del teorema precedente, è essenziale che la funzione f che definisce l'insieme sia *continua*. Senza la richiesta di continuità di f, infatti, *qualsiasi* insieme si può rappresentare, per esempio, nella forma $\{x \in \mathbb{R}^n : f(x) > 0\}$. Un esempio di questo fatto sarà fornito nell'Esercizio 18.

Se un insieme E è definito da più equazioni o disequazioni (in numero finito), che devono essere verificate *simultaneamente*, questo significa che E è assegnato come *intersezione* di più insiemi, ciascuno definito da una sola equazione o disequazione; se questi insiemi sono tutti aperti o tutti chiusi, E sarà aperto o chiuso, rispettivamente, in base al teorema 3.3.

Esempio

In $\mathbb{R}^2$, l'insieme di definizione di

$$g(x, y) = \frac{1}{\log(x + y)}$$

è

$$F = \{(x, y) : 0 < x + y \neq 1\},$$

perciò è aperto, perché intersezione di due aperti (stiamo sfruttando il fatto che la funzione $x + y$ è continua in tutto $\mathbb{R}^2$).

L'insieme di definizione di

$$h(x, y) = \sqrt{x^2 - y} + \sqrt{1 - x^2 - y^2}$$

è

$$G = \{(x, y) : x^2 - y \geq 0 \text{ e } 1 - x^2 - y^2 \geq 0\},$$

perciò è chiuso (intersezione di due chiusi).

Una curva piana definita da un'equazione del tipo $f(x, y) = 0$ è un insieme chiuso (se f è continua!). Per esempio, la circonferenza di equazione $x^2 + y^2 = R^2$ o la parabola di equazione $y = x^2$ sono insiemi chiusi.

DEFINIZIONE 3.8 Sia E un sottoinsieme di $\mathbb{R}^n$. Si dice:
interno di E, e si indica con E°, l'insieme dei punti interni di E;
frontiera o *bordo* di E, e si indica con ∂E, l'insieme dei punti di frontiera di E;
chiusura di E, e si indica con $\overline{E}$, l'insieme $E \cup \partial E$.
In particolare, si ha sempre:

$$E^\circ \subseteq E \subseteq \overline{E}.$$

Esempio

Sia E il "cerchio bucato":

$$E = \{(x, y) \in \mathbb{R}^2 : 0 < x^2 + y^2 \leq 1\}.$$

L'insieme E non è né aperto né chiuso. Si ha:

$$\partial E = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 = 1\} \cup \{(0,0)\}$$
$$\overline{E} = \{(x,y) \in \mathbb{R}^2 : x^2+y^2 \leq 1\}$$
$$E^\circ = \{(x,y) \in \mathbb{R}^2 : 0 < x^2+y^2 < 1\}$$

Si noti che in questo caso le inclusioni $E^\circ \subset E \subset \overline{E}$ sono entrambe strette.

Sia D la circonferenza $\{(x,y) \in \mathbb{R}^2 : x^2+y^2 = 1\}$. Ogni punto di D è di frontiera, perciò l'interno di D è vuoto: $D^\circ = \emptyset$.

OSSERVAZIONE Ricordando che un punto di E o è interno o è di frontiera, si vede subito che:

$$E^\circ \cup \partial E = \overline{E}.$$

Inoltre, dalle definizioni precedenti si vede che:

- un insieme è aperto se e solo se coincide con il suo interno;
- un insieme è chiuso se e solo se coincide con la sua chiusura;
- un insieme aperto non contiene nessuno dei suoi punti di frontiera;
- un insieme chiuso contiene tutti i suoi punti di frontiera (fig. 3.8).

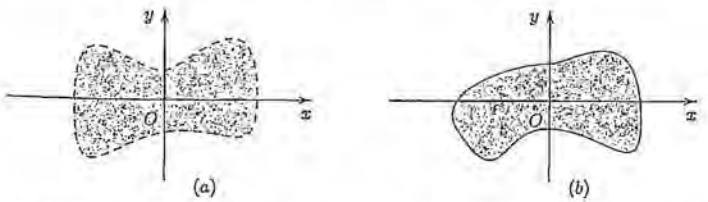


Figura 3.8. (a) Un insieme aperto; il bordo *non è contenuto* nell'insieme; (b) Un insieme chiuso; il bordo è *contenuto* nell'insieme.

DEFINIZIONE 3.9 Sia E un sottoinsieme di $\mathbb{R}^n$. E si dice *limitato* se esiste una costante $k > 0$ tale che

$$|x| \leq k \text{ per ogni } x \in E$$

(in altre parole, se è contenuto in una sfera di raggio abbastanza grande); *illimitato* altrimenti.

Con riferimento agli insiemi dell'Esempio 3.5, G è limitato, F non è limitato. Per convincersene, disegnare gli insiemi.

Come vedremo nel prossimo paragrafo, gli insiemi *chiusi* e *limitati* giocheranno un ruolo importante nello studio delle proprietà delle funzioni continue.

DEFINIZIONE 3.10 Un insieme E si dice *connesso* (per archi) se, per ogni coppia di punti $x, y \in E$, esiste un arco di curva continuo contenuto in E, che ha per estremi x e y.

Sostanzialmente, questo significa che E è composto "da un pezzo solo".

Esempio

Con riferimento agli insiemi dell'Esempio 3.5, G è connesso, F non è connesso, poiché la retta $x+y=1$ lo spezza in due parti "che non comunicano tra loro".

Il sostegno di un arco di curva continuo è un insieme connesso (esso stesso fornisce l'arco che congiunge due suoi punti qualsiasi!).

Si può dimostrare che in $\mathbb{R}$ gli insiemi connessi sono (tutti e soli) gli intervalli.

3.2 Proprietà topologiche delle funzioni continue

Vediamo ora alcuni dei motivi per cui le nozioni appena introdotte sono utili. Anzitutto, il teorema di Weierstrass sui massimi e minimi di una funzione continua, che abbiamo incontrato nello studio del calcolo infinitesimale in una variabile, ha una notevole generalizzazione alle funzioni di più variabili:

TEOREMA 3.6 (DI WEIERSTRASS) *Sia $E \subset \mathbb{R}^n$ un insieme chiuso e limitato e $f : E \to \mathbb{R}$ sia continua. Allora f ammette massimo e minimo in E, ossia esistono $x_m, x_M \in E$ tali che*

$$f(x_m) \leq f(x) \leq f(x_M) \text{ per ogni } x \in E.$$

Esempio

La funzione $h(x,y)$ dell'Esempio 3.5 ha senz'altro massimo e minimo nel suo insieme di definizione G, perché G è chiuso e limitato e h è continua in G.

Lo stesso non si può invece affermare, almeno a priori, per la funzione g dell'Esempio 3.5, che è definita su un insieme né chiuso né limitato.

DIMOSTRAZIONE. La dimostrazione ricalca quella vista nel caso unidimensionale[6]. La presenteremo nel caso bidimensionale, sufficiente a illustrare l'idea generale. Poiché $E \subset \mathbb{R}^2$ è limitato, è senz'altro contenuto in un rettangolo

$$R = [a,b] \times [c,d].$$

Proveremo che f ammette massimo in E; sia

$$\Lambda = \sup_E f = \sup\{f(x) : x \in E\}$$

(a priori Λ è finito o infinito; proveremo che in realtà è finito). Procediamo per *dicotomia*: dividiamo ognuno dei due intervalli $[a,b]$, $[c,d]$ in parti uguali, ottenendo così una suddivisione di R in 4 rettangoli uguali. Per almeno uno di questi, che chiamiamo R_1, sarà vero che

$$\Lambda = \sup_{R_1 \cap E} f.$$

Ora suddividiamo R_1 in 4 rettangolini uguali; per uno di questi, che chiamiamo R_2, sarà vero che

$$\Lambda = \sup_{R_2 \cap E} f.$$

Così facendo costruiamo una successione di rettangoli $R_k = [a_k, b_k] \times [c_k, d_k]$ ciascuno contenuto nei precedenti, con le proprietà:

[6] v. volume 1, capitolo 3, Teorema 3.27.

 © 978-88-08-06485-1

1. le successioni a_k e c_k sono monotone crescenti e limitate; b_k e d_k sono monotone decrescenti e limitate;
2. $b_k - a_k = \frac{b-a}{2^k} \to 0$; $d_k - c_k = \frac{d-c}{2^k} \to 0$;
3. $\Lambda = \sup_{R_k \cap E} f$.

Per il teorema di esistenza del limite per una successione monotona e limitata[6], dai punti 1) e 2) segue che le successioni a_k e b_k convergono a uno stesso limite $x_0 \in [a,b]$, mentre c_k e d_k convergono a uno stesso limite $y_0 \in [c,d]$.

Distinguiamo ora due casi: $\Lambda < +\infty, \Lambda = +\infty$ (proveremo che questo secondo caso in realtà non può verificarsi).

Se $\Lambda < +\infty$, per ogni k esiste un punto[7] $(x_k, y_k) \in R_k \cap E$ tale che

$$\Lambda - \frac{1}{k} < f(x_k, y_k) \leq \Lambda. \tag{3.1}$$

Poiché $x_k \in [a_k, b_k]$, $a_k \to x_0$ e $b_k \to x_0$, per il teorema del confronto anche $x_k \to x_0$; analogamente si mostra che $y_k \to y_0$. Dunque $(x_k, y_k) \to (x_0, y_0)$; inoltre, poiché i punti (x_k, y_k) appartengono a E ed E è chiuso, anche $(x_0, y_0) \in E$ (per il teorema 3.4).

Ancora per il teorema del confronto, la (3.1) dà

$$\lim_{k\to\infty} f(x_k, y_k) = \Lambda.$$

D'altro canto, poiché f è continua e $(x_k, y_k) \to (x_0, y_0)$, si ha che

$$\lim_{k\to\infty} f(x_k, y_k) = f(x_0, y_0),$$

perciò $f(x_0, y_0) = \Lambda$. Quindi Λ è il *massimo* di f in E ed è assunto nel punto (x_0, y_0). In questo caso perciò il teorema è dimostrato.

Supponiamo ora che sia $\Lambda = +\infty$. Allora, per ogni k, esiste un punto $(x_k, y_k) \in R_k \cap E$ tale che

$$f(x_k, y_k) \geq k. \tag{3.2}$$

Ragionando come sopra, si prova che $(x_k, y_k) \to (x_0, y_0)$ per un certo $(x_0, y_0) \in E$. Poiché f è continua in (x_0, y_0),

$$\lim_{k\to\infty} f(x_k, y_k) = f(x_0, y_0),$$

ma per la (3.2) $\lim_{k\to\infty} f(x_k, y_k) = +\infty$, assurdo perché in (x_0, y_0) la funzione deve avere un valore finito. ◊

Anche il teorema degli zeri per le funzioni continue ha un'utile estensione a più variabili.

TEOREMA 3.7 (DEGLI ZERI) *Sia E un insieme connesso di $\mathbb{R}^n$ e $f : E \to \mathbb{R}$ sia continua. Se x, y sono due punti di E tali che $f(\mathrm{x}) > 0$ e $f(\mathrm{y}) < 0$, allora esiste un terzo punto $\mathrm{z} \in E$ in cui f si annulla. In particolare, lungo ogni arco di curva continua (contenuto in E) che congiunge x e y, c'è almeno un punto in cui f si annulla.*

[6] v. volume 1, capitolo 3, paragrafo 1.2.

[7] Infatti, poiché $\Lambda - \frac{1}{k}$ è minore di Λ, che è il minimo dei maggioranti dei valori di $f(x,y)$ in $R_k \cap E$, $\Lambda - \frac{1}{k}$ non è un maggiorante, quindi esiste $(x_k, y_k) \in R_k \cap E$ con la proprietà precedente.

In questa versione multidimensionale, quindi, il concetto di *insieme connesso* si rivela essere una generalizzazione opportuna del concetto unidimensionale di *intervallo*.

DIMOSTRAZIONE. Sia $\mathrm{r} : [a,b] \to \mathbb{R}^n$ un arco di curva continua, contenuto in E, che congiunge x e y, ossia $\mathrm{r}(a) = \mathrm{x}, \mathrm{r}(b) = \mathrm{y}$. La funzione composta

$$g(t) = f(\mathrm{r}(t)), \ t \in [a,b],$$

è una funzione (reale di variabile reale) continua in $[a,b]$; inoltre

$$g(a) = f(\mathrm{r}(a)) = f(\mathrm{x}) > 0;$$
$$g(b) = f(\mathrm{r}(b)) = f(\mathrm{y}) < 0.$$

Per il teorema degli zeri unidimensionale[8], esiste $t_0 \in [a,b]$ tale che $g(t_0) = 0$, ossia tale che $f(\mathrm{r}(t_0)) = 0$. Il teorema è dimostrato, con $\mathrm{z} = r(t_0) \in E$. ◊

Il teorema precedente ha un'applicazione pratica nello *studio del segno di una funzione di più variabili*. Se f è una funzione continua, definita su un dominio qualunque, l'insieme degli zeri di f spezza il dominio in un certo numero di insiemi connessi (in cui f è diversa da zero), *su ciascuno dei quali il segno di f è costante* (se cambiasse, all'interno del connesso ci dovrebbe essere un punto in cui la funzione si annulla). È sufficiente allora valutare il segno di f in un punto solo di ogni connesso per conoscere il segno in tutta quella componente.

Esempio

Si studi il segno di

$$f(x,y) = x^2 - y^2 - 1$$

La funzione è definita in tutto il piano; l'insieme in cui si annulla è l'iperbole equilatera $x^2 - y^2 = 1$. Questa curva spezza il piano in tre componenti connesse, R_1, R_2, R_3 (fig. 3.9), su ognuna delle quali il segno di f è costante: basta calcolare f in un punto per conoscerne il segno in tutta la componente. Per esempio: $f(2,0) > 0$, perciò $f > 0$ in tutto R_1; $f(0,0) < 0$, perciò $f < 0$ in R_2; $f(-2,0) > 0$, perciò $f > 0$ in R_3.

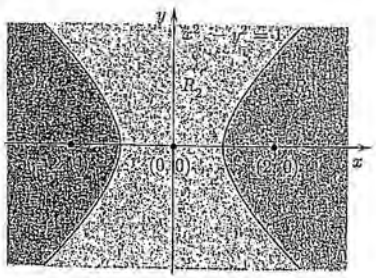


Figura 3.9.

[8] v. volume 1, capitolo 3, teorema 3.26.

Esercizi

(14) Dimostrare che, se A è un insieme aperto e C un insieme chiuso, allora $A \setminus C$ è aperto e $C \setminus A$ è chiuso.

(15) Dimostrare che se A, B sono insiemi connessi e $A \cap B \neq \varnothing$, allora $A \cup B$ è connesso.
Mostrare poi con un esempio che se A, B sono insiemi connessi, $A \cap B$ può non essere connesso.

(16) Dare un esempio di insieme connesso la cui frontiera non è connessa.

(17) Dimostrare che se $A \subset B \subset \mathbb{R}^n$, allora:

$$A^\circ \subset B^\circ;$$
$$\overline{A} \subset \overline{B}.$$

(18) Dimostrare che per ogni sottoinsieme $E \subset \mathbb{R}^n$ valgono le seguenti identità:

$$(E^\circ)^\circ = E^\circ;$$
$$\overline{(E)} = \overline{E}.$$

(19) Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione continua su tutto lo spazio e sia

$$E = \{x \in \mathbb{R}^n : f(x) > 0\}.$$

Stabilire se è vera o falsa ciascuna delle seguenti affermazioni, fornendo una dimostrazione (se vera) o un contresempio (se falsa):

$$\partial E \subseteq \{x \in \mathbb{R}^n : f(x) = 0\};$$
$$\partial E = \{x \in \mathbb{R}^n : f(x) = 0\};$$
$$\overline{E} = \{x \in \mathbb{R}^n : f(x) \geq 0\};$$
$$\overline{E} \subseteq \{x \in \mathbb{R}^n : f(x) \geq 0\}.$$

(20) In $\mathbb{R}$, dare un esempio di insieme E tale che $\partial E = \mathbb{R}$.

(21) Per ciascuna delle seguenti funzioni:
a) Determinare (e se possibile disegnare) l'insieme E di definizione.
b) Rispondere alle domande: E è aperto? È chiuso? È limitato? È connesso?
c) Determinare (e se possibile disegnare) l'insieme in cui f si annulla e studiare il segno di f.

$$(i)\quad f(x,y) = \sqrt{x - \sqrt{1 - x^2 - y^2}};$$
$$(ii)\quad f(x,y) = \frac{x \log(1 + x^2 - y)}{y};$$
$$(iii)\quad f(x,y) = \log \sin(x^2 + y^2);$$
$$(iv)\quad f(x,y) = \frac{\sqrt{x + y + 1}}{2x - y + 2}.$$

(22) Dire se i seguenti insiemi, in $\mathbb{R}^3$, sono o non sono aperti, chiusi, limitati, connessi:

(i) $\{(x,y,z) \in \mathbb{R}^3 : xyz = 0\}$;
(ii) $\{(x,y,z) \in \mathbb{R}^3 : x^2 + 2y^2 + z^2 < 1\}$;
(iii) $\{(x,y,z) \in \mathbb{R}^3 : x^2 + y^2 = 1\}$.

(23) Sia E l'insieme del piano definito da:

$$E = \{(x,y) \in \mathbb{R}^2 : f(x,y) > 0\},$$

dove

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

Determinare l'insieme E e disegnarlo; rispondere quindi alle domande:
– l'insieme E è aperto?
– È chiuso?
– La funzione f è continua?
Riflettere, alla luce di quest'esempio, sul ruolo della continuità di f nel teorema 3.5.

(24) Dimostrare le seguenti versioni del teorema di Weierstrass su un dominio illimitato:
(a) "Sia $f : \mathbb{R}^n \to \mathbb{R}$ continua. Se $f(x) \geq 0\ \forall x \in \mathbb{R}^n$ e $\lim_{|x| \to +\infty} f(x) = 0$, allora f ammette massimo".
Mostrare con un esempio che f può non ammettere minimo; mostrare con un altro esempio che, senza l'ipotesi $f(x) \geq 0\ \forall x \in \mathbb{R}^n$, la conclusione è falsa in generale.
(b) "Sia $f : \mathbb{R}^n \to \mathbb{R}$ continua. Se $\lim_{|x| \to +\infty} f(x) = +\infty$ (o $-\infty$), allora f ammette minimo (rispettivamente, massimo)".

4 DERIVATE PARZIALI, PIANO TANGENTE, DIFFERENZIALE

4.1 Derivate parziali

Ci interessa ora definire il concetto di *derivata* per una funzione $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$.

Il caso bidimensionale

Il caso $n = 2$ permette già di illustrare quali siano i problemi legati alla definizione di derivata per funzioni di più variabili e come li si superi. La "solita" definizione di derivata non è applicabile in modo ovvio perché, se (x_0, y_0) è il punto in cui vogliamo derivare $f(x,y)$, che cosa si intende per incremento della coppia di variabili (x,y)? Non può essere un vettore, perché non avrebbe senso dividere poi l'incremento di f per un vettore.

La soluzione sta nell'incrementare solo una variabile alla volta: si può tenere costante il valore di $y = y_0$ e derivare la *funzione di una variabile* $x \mapsto f(x, y_0)$ nel punto x_0; oppure tenere costante il valore di $x = x_0$ e derivare la *funzione di una variabile* $y \mapsto f(x_0, y)$ nel punto y_0. Il risultato è espresso, rispettivamente, dalle seguenti formule, ove si intende che i limiti presenti siano *finiti*:

$$\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h, y_0) - f(x_0, y_0)}{h} \tag{4.1}$$

oppure

$$\frac{\partial f}{\partial y}(x_0, y_0) = \lim_{k \to 0} \frac{f(x_0, y_0 + k) - f(x_0, y_0)}{k} \tag{4.2}$$

La (4.1) e la (4.2) si dicono, rispettivamente, *derivata parziale di f rispetto a x*, o a y, *calcolata in* (x_0, y_0) e si indicano rispettivamente con uno dei seguenti simboli:

$$\frac{\partial f}{\partial x}(x_0, y_0), \qquad \partial_x f(x_0, y_0), \qquad D_x f(x_0, y_0), \qquad D_1 f(x_0, y_0), \qquad f_x(x_0, y_0)$$

e

$$\frac{\partial f}{\partial y}(x_0, y_0), \qquad \partial_y f(x_0, y_0), \qquad D_y f(x_0, y_0), \qquad D_2 f(x_0, y_0), \qquad f_y(x_0, y_0).$$

Il vettore che ha per componenti le derivate parziali di f in (x_0, y_0) si dice *gradiente* di f, calcolato in (x_0, y_0) e si indica con uno dei seguenti simboli:

$$\nabla f(x_0, y_0), \qquad Df(x_0, y_0), \qquad \mathrm{grad} f(x_0, y_0).$$

Esempi

4.10 Sia $f(x, y) = x^2 y^3$.

$$\frac{\partial f}{\partial x}(1, 2) = \frac{d}{dx}[f(x, 2)]_{|x=1} = \frac{d}{dx}[8x^2]_{|x=1} = (16x)_{|x=1} = 16.$$

Abbiamo applicato la definizione, scrivendo prima la funzione di una variabile che si ottiene restringendo f alla retta $y = 2$, quindi calcolando la derivata di questa rispetto a x e valutando infine la derivata in $x = 1$. Ma si può procedere più speditamente ragionando così: la derivata $\frac{\partial f}{\partial x}$ in un generico punto (x, y) è la derivata di $x^2 y^3$ rispetto a x, quando si pensi y costante, perciò, per le regole di calcolo differenziale in una variabile,

$$\frac{\partial f}{\partial x}(x, y) = \frac{\partial}{\partial x}(x^2 y^3) = 2xy^3.$$

Questa è la derivata desiderata, in ogni punto (x, y); in particolare

$$\frac{\partial f}{\partial x}(1, 2) = (2xy^3)_{|(x,y)=(1,2)} = 16.$$

4.11 Talvolta è *necessario* applicare il primo metodo (ossia la definizione). Si voglia calcolare

$$\frac{\partial}{\partial x}(y\sqrt{x})_{|(x,y)=(0,0)}.$$

Se si calcola, formalmente,

$$\frac{\partial}{\partial x}(y\sqrt{x}) = \frac{y}{2\sqrt{x}}$$

e poi si tenta di valutare l'espressione trovata in $(0, 0)$, si trova $\frac{0}{0}$, che non ha senso. Invece, dalla definizione si ha che:

$$\frac{\partial}{\partial x}(y\sqrt{x})_{|(x,y)=(0,0)} = \frac{d}{dx}(0\sqrt{x})_{|x=0} = \frac{d}{dx}(0) = 0.$$

La derivata parziale rispetto a x nell'origine dunque esiste e vale zero.

Il caso n-dimensionale

Possiamo ora considerare il caso generale di una funzione di n variabili. La generalizzazione delle idee precedenti è, a questo punto, più che altro una questione di notazioni.

Se $f : \mathbb{R}^n \to \mathbb{R}$, per calcolare la derivata parziale di f rispetto alla i-esima variabile x_i, in un punto x_0 fissato, terremo fisse tutte le coordinate di x eccetto la i-esima e calcoleremo il limite del rapporto incrementale di f in questa direzione. In simboli:

$$\frac{\partial f}{\partial x_i}(\mathrm{x}_0) = \lim_{h\to 0} \frac{f(\mathrm{x}_0 + h\mathrm{e}_i) - f(\mathrm{x}_0)}{h} \quad (\text{per } i = 1, 2, \ldots, n)$$

(si noti che h è ancora una grandezza scalare). Si rifletta sul significato della precedente notazione vettoriale compatta: se

$$\mathrm{x}_0 = (x_1^0, x_2^0, \ldots, x_n^0),$$

ricordando che

$$\mathrm{e}_i = \Big(0, 0, \ldots 0, \underset{i}{1}, 0, \ldots, 0\Big),$$

si ha, più esplicitamente,

$$\frac{\partial f}{\partial x_i}(x_1^0, x_2^0, \ldots, x_n^0) = \lim_{h\to 0} \frac{f(x_1^0, x_2^0, \ldots, x_i^0 + h, \ldots, x_n^0) - f(x_1^0, x_2^0, \ldots, x_i^0, \ldots, x_n^0)}{h}.$$

Osservazione Per poter calcolare le derivate parziali di f in un punto x_0, è necessario che anche il punto incrementato appartenga al dominio di f, almeno per un incremento abbastanza piccolo. Questo è certamente vero se il dominio di f è un insieme *aperto*. D'ora in poi, quindi, considereremo $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ con A aperto.

Definizione 3.11 Una funzione $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ si dice *derivabile* in un punto del suo dominio se in quel punto esistono tutte le sue derivate parziali; si dice derivabile in A se è derivabile in ogni punto di A.

Se f è derivabile in un punto, chiameremo ancora *gradiente* il vettore delle sue derivate parziali:

$$\nabla f(\mathrm{x}) = \left(\frac{\partial f}{\partial x_i}(\mathrm{x}), \frac{\partial f}{\partial x_2}(\mathrm{x}), \ldots, \frac{\partial f}{\partial x_n}(\mathrm{x})\right).$$

Esempi

4.12 La funzione $f : \mathbb{R}^n \to \mathbb{R}$

$$f(\mathrm{x}) = e^{-|\mathrm{x}|^2}$$

è derivabile in tutto $\mathbb{R}^n$ e, infatti, per ogni $i = 1, 2, \ldots, n$ e ogni $\mathrm{x} \in \mathbb{R}^n$ si ha:

$$\frac{\partial f}{\partial x_i}(\mathrm{x}) = \frac{\partial}{\partial x_i}\left(e^{-(x_1^2 + x_2^2 + \ldots + x_n^2)}\right) = -2x_i\left(e^{-(x_1^2 + x_2^2 + \ldots + x_n^2)}\right) = -2x_i e^{-|\mathrm{x}|^2}$$

ESEMPIO La funzione $f : \mathbb{R}^n \to \mathbb{R}$

$$f(\mathbf{x}) = e^{|\mathbf{x}|}$$

è derivabile in tutti i punti di $\mathbb{R}^n$, tranne l'origine. Infatti, per ogni $i = 1, 2, \ldots, n$ e ogni $\mathbf{x} \in \mathbb{R}^n$, si ha:

$$\frac{\partial f}{\partial x_i}(\mathbf{x}) = \frac{\partial}{\partial x_i}\left(e^{\sqrt{x_1^2+x_2^2+\ldots+x_n^2}}\right) = \frac{x_i}{\sqrt{x_1^2+x_2^2+\ldots+x_n^2}}\left(e^{\sqrt{x_1^2+x_2^2+\ldots+x_n^2}}\right) = \frac{x_i e^{|\mathbf{x}|}}{|\mathbf{x}|},$$

espressione che perde significato per $\mathbf{x} = \mathbf{0}$. In questo punto si ha:

$$\frac{\partial f}{\partial x_i}(0) = \frac{\partial}{\partial x_i}\left(e^{|x_i|}\right)_{/x_i=0} = \lim_{h\to 0}\frac{e^{|h|}-1}{h}, \text{ che non esiste.}$$

4.2 Piano tangente

Per funzioni di una variabile, definire la derivata di una funzione equivale sostanzialmente a definire la retta tangente al suo grafico. Per una funzione di *due* variabili il problema analogo è definire il *piano tangente al grafico della funzione*. Ci proponiamo ora di determinare tale piano. Questo problema geometrico sarà la strada naturale per arrivare alla nozione generale di *funzione differenziabile*, concetto centrale in questo paragrafo.

Lasciamoci guidare dall'intuizione geometrica. Se sezioniamo il grafico di $z = f(x, y)$ con il piano verticale $y = y_0$, troveremo una curva in questo stesso piano, descritta dall'equazione $z = f(x, y_0)$. La retta r_1 tangente a tale curva in x_0 starà anche sul piano tangente che cerchiamo; ripetendo la stessa costruzione lungo il piano verticale $x = x_0$, troviamo una seconda retta r_2 passante per il punto $(x_0, y_0, f(x_0, y_0))$ e appartenente al piano tangente: poiché in $\mathbb{R}^3$ due rette intersecantesi individuano un unico piano, questo dovrebbe essere il piano tangente (fig. 3.10).

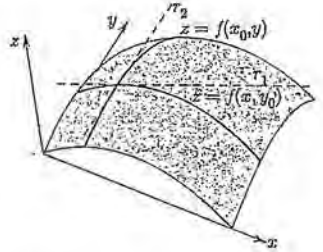


Figura 3.10.

L'equazione della retta r_1 nel piano $y = y_0$ è:

$$z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0).$$

La stessa retta, considerata nello spazio (x, y, z) anziché nel piano $y = y_0$, è individuata dal *sistema*:

$$\text{(4.3)} \qquad \begin{cases} z = f(x_0, y_0) + \dfrac{\partial f}{\partial x}(x_0, y_0)(x - x_0) \\ y = y_0. \end{cases}$$

Analogamente, la retta r_2 è individuata dal sistema:

$$\text{(4.4)} \qquad \begin{cases} z = f(x_0, y_0) + \dfrac{\partial f}{\partial y}(x_0, y_0)(y - y_0) \\ x = x_0. \end{cases}$$

Il piano che contiene entrambe le rette è:

$$\text{(4.5)} \qquad z = f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0, y_0)(y - y_0),$$

come si verifica immediatamente (ogni punto (x, y, z) che soddisfa (4.3) o (4.4) soddisfa anche (4.5)). Possiamo affermare che la (4.5) è l'*equazione del piano tangente al grafico di* $z = f(x, y)$ *nel punto* $(x_0, y_0, f(x_0, y_0))$? Prima di rispondere, vediamo un paio di esempi.

ESEMPI

ESEMPIO Determiniamo il piano tangente al grafico di $z = x^2 + y^2$ nel punto $(1, 1)$.

$$\frac{\partial f}{\partial x} = 2x; \ \frac{\partial f}{\partial y} = 2y$$

$$\frac{\partial f}{\partial x}(1,1) = 2; \ \frac{\partial f}{\partial y}(1,1) = 2; \ f(1,1) = 2;$$

l'equazione è:

$$z = 2 + 2(x-1) + 2(y-1)$$

cioè

$$z = 2x + 2y - 2.$$

Dal grafico in figura 3.11, sembra ragionevole considerare questo piano come "tangente".

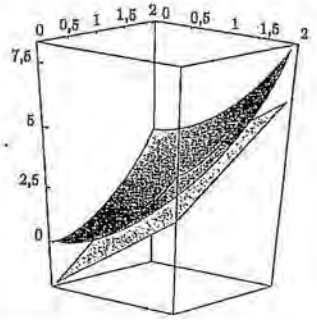


Figura 3.11. Grafico di $f(x, y) = x^2 + y^2$ e del piano $z = 2x + 2y - 2$.

ESEMPIO Sia

$$f(x, y) = \begin{cases} \dfrac{xy}{x^2+y^2} & \text{se } (x, y) \neq (0, 0) \\ 0 & \text{se } (x, y) = (0, 0). \end{cases}$$

 © 978-88-08-06485-1

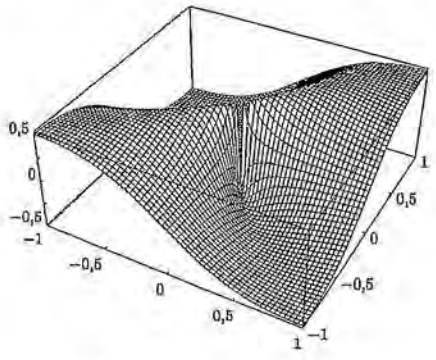


Figura 3.12. Grafico di $f(x,y) = \frac{xy}{x^2+y^2}x$.

La funzione è identicamente nulla sugli assi coordinati, perciò, calcolando le derivate parziali di f nell'origine mediante la definizione, vediamo subito che esse esistono e sono nulle. Poiché anche f è nulla in $(0,0)$, la (4.5) coincide col piano $z = 0$. Uno sguardo al grafico di $f(x,y)$ in figura 3.12 mostra però che questo piano non ha il diritto di chiamarsi "tangente": la funzione non è neppure continua nell'origine!

Occorre ora una riflessione. L'equazione (4.5) si può scrivere ogni volta che $f(x,y)$ è derivabile; il piano che essa individua, per costruzione, contiene le rette tangenti alle sezioni che il grafico di $f(x,y)$ forma con i piani $x = x_0$, $y = y_0$. Se questo piano è il piano tangente, deve contenere anche tutte le rette tangenti alle altre curve che si ottengono sezionando il grafico con un piano verticale qualsiasi. Purtroppo, però, non c'è garanzia che questo accada nella sola ipotesi di derivabilità di $f(x,y)$: in altre parole, il procedimento che abbiamo seguito individua il piano tangente *nell'ipotesi che questo ci sia*; potrebbe però non esserci, come accade nell'esempio 4.6.

4.3 Differenziabilità e approssimazione lineare

Il caso bidimensionale

L'ultimo esempio mostra un problema tipico del calcolo differenziale in più variabili, che non ha analogo nel caso unidimensionale: mentre in una variabile la derivabilità è equivalente all'esistenza della retta tangente e implica la continuità, in due (o più) variabili *la sola derivabilità* (nel senso della definizione precedente) *non implica né la continuità né l'esistenza del piano tangente.*

Qual è allora la proprietà di f che garantisce che il piano (4.5) sia "realmente" tangente? Per rispondere, basta osservare qual è la proprietà fondamentale della retta tangente nel caso unidimensionale. Questa è espressa dalla relazione:

$$f(x+h) - f(x) = f'(x)h + o(h) \text{ per } h \to 0.$$

L'incremento della funzione coincide al primo ordine con il *differenziale*, cioè con l'incremento calcolato lungo la retta tangente; i due incrementi differiscono per infinitesimi di ordine superiore rispetto all'incremento h della variabile indipendente[9].

Il concetto analogo è quello di *differenziabilità in più variabili*: l'incremento di f è uguale all'incremento calcolato lungo il piano tangente, più un infinitesimo di ordine superiore rispetto alla lunghezza dell'incremento (h,k) delle variabili indipendenti. In formule:

(4.6)

$$f(x_0+h, y_0+k) - f(x_0,y_0) = \\ = \frac{\partial f}{\partial x}(x_0,y_0)(x-x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y-y_0) + o\left(\sqrt{h^2+k^2}\right) \quad \text{per } (h,k) \to (0,0).$$

Nella formula precedente, il primo membro rappresenta l'incremento della funzione; i primi due addendi del secondo membro rappresentano l'incremento calcolato lungo il piano tangente; l'ultimo addendo costituisce l'errore che si commette valutando l'incremento di f come incremento lungo il piano tangente. Il simbolo $o\left(\sqrt{h^2+k^2}\right)$ indica, per definizione, una funzione di (h,k) tale che

$$\lim_{(h,k)\to(0,0)} \frac{o\left(\sqrt{h^2+k^2}\right)}{\sqrt{h^2+k^2}} = 0.$$

Nel caso particolare $(x_0,y_0) = (0,0)$, la (4.6) assume la forma:

$$f(x,y) - f(0,0) = \frac{\partial f}{\partial x}(0,0)x + \frac{\partial f}{\partial y}(0,0)y + o\left(\sqrt{x^2+y^2}\right) \tag{4.7}$$

per $(x,y) \to (0,0)$.

Se la (4.6) è verificata, diremo che f è *differenziabile* in (x_0,y_0). La condizione (4.6) garantisce che il piano (4.5) sia tangente al grafico di f (anzi, è una definizione precisa di cosa significhi piano tangente). Il piano tangente è dunque il piano (4.5), ma *solo se* la (4.6) è verificata; altrimenti non esiste.

Il caso n-dimensionale

Diamo ora la definizione generale di differenziabilità per una funzione di n variabili:

DEFINIZIONE 3.12 Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ con A aperto e sia $\mathbf{x}_0 \in A$. Diremo che f è *differenziabile* in $\mathbf{x}_0$ se esiste un vettore $\mathbf{a} \in \mathbb{R}^n$ tale che:

$$f(\mathbf{x}_0+\mathbf{h}) - f(\mathbf{x}_0) = \mathbf{a}\cdot\mathbf{h} + o(|\mathbf{h}|) \text{ per } \mathbf{h} \to \mathbf{0}. \tag{4.8}$$

Nella formula precedente, il vettore $\mathbf{h} \in \mathbb{R}^n$ è l'incremento della variabile indipendente e $\mathbf{a}\cdot\mathbf{h}$ il prodotto scalare dei due vettori; la funzione f è definita nel punto $\mathbf{x}_0+\mathbf{h}$, almeno quando $|\mathbf{h}|$ è abbastanza piccolo, perché per ipotesi il punto $\mathbf{x}_0$ è interno al dominio di definizione di f. Più esplicitamente, la (4.8) significa che:

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{f(\mathbf{x}_0+\mathbf{h}) - f(\mathbf{x}_0) - \mathbf{a}\cdot\mathbf{h}}{|\mathbf{h}|} = 0. \tag{4.9}$$

[9] Si veda il volume 1, capitolo 4, paragrafo 7.1.

Dimostriamo subito una semplice proposizione che rende più comprensibile la definizione precedente:

PROPOSIZIONE 3.2 *Se f è differenziabile in x_0, allora f è anche derivabile in x_0 e il vettore* a *che compare nella* (4.8) *è il gradiente di f calcolato in x_0:*

$$a = \nabla f(x_0).$$

Inoltre, f è continua in x_0.

DIMOSTRAZIONE. Nella (4.9), scegliamo come incremento un vettore $h = he_i$ con $h \in \mathbb{R}$ (abbastanza piccolo) e $i = 1, 2, \ldots, n$ (in altre parole, incrementiamo la variabile solo nella direzione parallela a un asse coordinato). Otteniamo, ponendo $a = (a_1, a_2, \ldots, a_n)$:

$$\lim_{h\to 0} \frac{f(x_0 + he_i) - f(x_0) - a_i h}{h} = 0,$$

da cui leggiamo che esiste

$$\lim_{h\to 0} \frac{f(x_0 + he_i) - f(x_0)}{h} = a_i.$$

Ma questo limite per definizione è $\frac{\partial f}{\partial x_i}(x_0)$, perciò $a_i = \frac{\partial f}{\partial x_i}(x_0)$ per ogni $i = 1, 2, \ldots, n$ e

$$a = \nabla f(x_0).$$

In particolare, f è derivabile in x_0. Per mostrare che è continua, basta osservare che, direttamente dalla (4.8), se $h \to 0$ il secondo membro della (4.8) tende a zero; di conseguenza tende a zero anche il primo membro, ossia $f(x_0 + h) \to f(x_0)$, che esprime la continuità di f in x_0. ◆

In base alla proposizione precedente, la definizione di differenziabilità di f in x_0 assume la forma:

$$f(x_0 + h) - f(x_0) = \nabla f(x_0) \cdot h + o(|h|) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)\, h_i + o(|h|) \tag{4.10}$$

per $h \to 0$.

DEFINIZIONE 3.13 Se f è differenziabile in x_0, si dice *differenziale di f calcolato in x_0* la funzione lineare $df(x_0) : \mathbb{R}^n \to \mathbb{R}$ definita da

$$df(x_0) : h \longmapsto \nabla f(x_0) \cdot h.$$

Nel caso $n = 2$, il numero $\nabla f(x_0) \cdot h$ rappresenta l'incremento della funzione nel passare da x_0 a $x_0 + h$, calcolato lungo il piano tangente al grafico di f in x_0.

L'approssimazione dell'incremento di f mediante il suo differenziale (applicato ad h) prende il nome di *linearizzazione di f* e, per le funzioni di più variabili come già per quelle di una variabile sola, costituisce la più semplice e tipica applicazione del calcolo differenziale. Spesso, sottointendendo l'incremento h, con abuso di linguaggio si scrive semplicemente $df(x_0)$ intendendo $df(x_0)(h)$, cioè $\nabla f(x_0) \cdot h$. La (4.10) prende allora la forma sintetica

$$\Delta f(x_0) = df(x_0) + o(|h|),$$

che esprime appunto la linearizzazione di f.

Possiamo anche dare la seguente:

DEFINIZIONE 3.14 Se f è differenziabile in $x_0 = (x_1^0, x_2^0, \ldots, x_n^0)$, si dice *iperpiano tangente* al grafico di f, in x_0, l'iperpiano

$$z = f(x_0) + \nabla f(x_0) \cdot (x - x_0)$$

ossia

$$z = f(x_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)\,(x_i - x_i^0).$$

Questa è effettivamente l'equazione di un iperpiano nello spazio $\mathbb{R}^{n+1}$ delle variabili $(x_1, x_2, \ldots, x_n, z)$. Naturalmente, nel caso $n = 2$ ritroviamo la definizione di piano tangente data in precedenza.

Riassumendo: in questo paragrafo abbiamo dato la definizione di differenziabilità di una funzione in un punto; abbiamo visto che questa implica la derivabilità, la continuità e, per definizione stessa, l'esistenza dell'iperpiano tangente. Abbiamo visto invece, nel paragrafo precedente, che la sola derivabilità (cioè esistenza delle derivate parziali) non è sufficiente a garantire neppure la continuità della funzione.

Dunque la differenziabilità è una condizione più forte sia della continuità che della derivabilità.

Purtroppo, tuttavia, la differenziabilità non è facile da verificare direttamente. Per esempio, nel caso $n = 2$ dimostrare la differenziabilità in (x_0, y_0), ossia la (4.6), significa provare che:

$$\lim_{(h,k)\to(0,0)} \frac{f(x_0 + h, y_0 + k) - \{f(x_0, y_0) + \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k\}}{\sqrt{h^2 + k^2}} = 0$$

e questo è un limite, in due variabili, che dà certamente una forma di indeterminazione e va quindi trattato con i metodi illustrati nel paragrafo 2.2. È desiderabile perciò conoscere qualche condizione *sufficiente* a garantire che la (4.6) valga e di più facile verifica. Un criterio molto utile a questo riguardo è indicato nel prossimo teorema.

TEOREMA 3.8 (CONDIZIONE SUFFICIENTE DI DIFFERENZIABILITÀ)
Siano $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, con A aperto e $x_0 \in A$. Supponiamo che le derivate parziali di f esistano in un intorno di x_0 e siano continue in x_0. Allora f è differenziabile in x_0.

In particolare, se le derivate parziali di f esistono e sono continue in tutto A, allora f è differenziabile in tutti i punti di A.

Una funzione le cui derivate parziali esistono e sono continue in tutto un aperto A si dice *di classe* $C^1(A)$ e si scrive $f \in C^1(A)$. Il secondo enunciato del teorema afferma quindi che

$$f \in C^1(A) \Rightarrow f \text{ differenziabile in } A$$

(L'implicazione inversa non vale, come vedremo negli esempi).

Il teorema precedente detta il metodo comunemente seguito per stabilire la differenziabilità di una funzione, senza calcolare limiti.

Esempio

Riprendiamo la funzione dell'Esempio 4.5: $z = x^2 + y^2$. Abbiamo calcolato:

$$\frac{\partial f}{\partial x} = 2x; \quad \frac{\partial f}{\partial y} = 2y$$

queste funzioni sono evidentemente continue in tutto il piano, quindi $f \in C^1(\mathbb{R}^2)$, in particolare è differenziabile in tutto il piano. Ciò significa, per esempio, che il piano che abbiamo calcolato nell'esempio 4.5 è effettivamente il piano tangente alla funzione nel punto assegnato.

In pratica, per decidere che una funzione è C^1 in un aperto, spesso non occorre neppure *calcolare le derivate parziali*, ma è sufficiente *osservare la forma della funzione*, ricordando le regole di derivazione e i teoremi sulla continuità delle funzioni elementari. Per esempio, è ovvio che la funzione

$$f(x,y) = \frac{e^{3x^2y} - 5\cos(xy)}{x^2 + y^2}$$

al di fuori dell'origine è continua *e ha derivate parziali continue*, perché, in base alle regole di derivazione, anche le derivate parziali saranno date da un quoziente di funzioni continue, il cui denominatore si annulla solo nell'origine. Pertanto $f \in C^1(\mathbb{R}^n \setminus \{0\})$ e in tale insieme f è anche differenziabile.

DIMOSTRAZIONE DEL TEOREMA 3.8. Naturalmente il secondo enunciato del teorema segue dal primo, applicandolo a ogni punto di A.

Proviamo il primo enunciato, nel caso $n = 2$. Il caso generale procede sulla stessa linea. La dimostrazione si basa su un ragionamento sostanzialmente unidimensionale: per valutare

$$f(x_0 + h, y_0 + k) - f(x_0, y_0)$$

muoviamoci dal punto (x_0, y_0) al punto $(x_0 + h, y_0 + k)$ incrementando le variabili una alla volta, ossia prima sul segmento orizzontale che unisce (x_0, y_0) a $(x_0 + h, y_0)$, poi sul segmento verticale che unisce $(x_0 + h, y_0)$ a $(x_0 + h, y_0 + k)$.

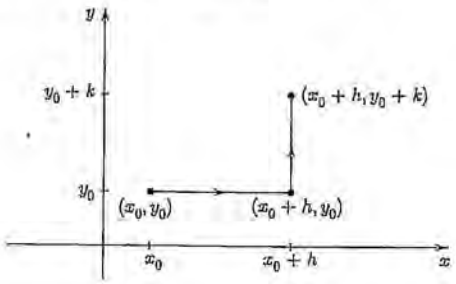


Figura 3.13.

La funzione f ristretta a ciascuno dei due segmenti è una funzione di una sola variabile.

Applicando il teorema di Lagrange[10] alla funzione $x \mapsto f(x, y_0)$ della sola variabile x, possiamo scrivere

$$f(x_0 + h, y_0) - f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0 + \delta_1 h, y_0)h \tag{4.11}$$

per un opportuno $\delta_1 \in (0,1)$ (δ_1 dipendente da h). Applicando ancora il teorema di Lagrange alla funzione $y \mapsto f(x_0 + h, y)$ della variabile y, abbiamo invece

$$f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) = \frac{\partial f}{\partial y}(x_0 + h, y_0 + \delta_2 k)k \tag{4.12}$$

per un opportuno $\delta_2 \in (0,1)$ (δ_2 dipendente da h e k). Sommando membro a membro (4.11) e (4.12):

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0 + \delta_1 h, y_0)h + \frac{\partial f}{\partial y}(x_0 + h, y_0 + \delta_2 k)k.$$

Poiché le derivate parziali sono continue in (x_0, y_0), possiamo scrivere:

$$\frac{\partial f}{\partial x}(x_0 + \delta_1 h, y_0) = \frac{\partial f}{\partial x}(x_0, y_0) + \varepsilon_1(h)$$
$$\frac{\partial f}{\partial y}(x_0 + h, y_0 + \delta_2 k)k = \frac{\partial f}{\partial y}(x_0, y_0) + \varepsilon_2(h,k)$$

con $\varepsilon_1(h), \varepsilon_2(h,k) \to 0$ per $(h,k) \to (0,0)$; otteniamo così:

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = \frac{\partial f}{\partial x}(x_0, y_0)h + \frac{\partial f}{\partial y}(x_0, y_0)k + \varepsilon_1(h)h + \varepsilon_2(h,k)k.$$

Basta ora osservare che la quantità $\varepsilon_1(h)h + \varepsilon_2(h,k)k$ è anche $o(\sqrt{h^2 + k^2})$ per $(h,k) \to (0,0)$. Infatti:

$$\frac{|\varepsilon_1(h)h + \varepsilon_2(h,k)k|}{\sqrt{h^2 + k^2}} \le \left|\frac{\varepsilon_1(h)h}{h}\right| + \left|\frac{\varepsilon_2(h,k)k}{k}\right| = |\varepsilon_1| + |\varepsilon_2| \to 0.$$

Quindi la (4.6) è verificata. ∎

Ricordiamo che il teorema precedente, pur molto utile, dà una *condizione sufficiente ma non necessaria per la differenziabilità*.

Esempio

La funzione

$$f(x,y) = xy^{1/3}$$

è definita in tutto $\mathbb{R}^2$. Chiediamoci in quali punti è differenziabile. Calcoliamo anzitutto:

$$\frac{\partial f}{\partial x}(x,y) = y^{1/3} \text{ definita e continua in tutto } \mathbb{R}^2;$$
$$\frac{\partial f}{\partial y}(x,y) = \frac{x}{3y^{2/3}} \text{ purché } y \neq 0.$$

Nei punti $(x_0, 0)$ la derivata $\frac{\partial f}{\partial y}$ va calcolata in base alla definizione. Notiamo che se $x_0 \neq 0$,

$$\frac{\partial f}{\partial y}(x_0, 0) = \frac{\partial}{\partial y}\left(x_0 y^{1/3}\right)_{/y=0} = x_0 \frac{\partial}{\partial y}\left(y^{1/3}\right)_{/y=0} \text{ non esiste;}$$

[10] v. volume 1, capitolo 4, teorema 4.6.

mentre se $x_0 = 0$,

$$\frac{\partial f}{\partial y}(0,0) = \frac{\partial}{\partial y}\left(0 \cdot y^{1/3}\right)_{/y=0} = 0$$

in quanto $f(0,y)$ è identicamente nulla.

Ricapitolando abbiamo:

nell'aperto $A = \{(x,y) \in \mathbb{R}^2 : y \neq 0\}$ entrambe le derivate parziali esistono e sono continue; dunque $f \in C^1(A)$ e, per il teorema precedente, f è differenziabile in A;

nei punti $(x_0, 0)$ con $x_0 \neq 0$ la funzione f non è derivabile (perché non esiste $\frac{\partial f}{\partial y}$), quindi a maggior ragione non è differenziabile;

rimane da studiare l'origine, in cui f è derivabile, ma il teorema precedente non è applicabile (perché l'origine non appartiene a un insieme aperto in cui le derivate parziali sono continue).

Nell'origine, quindi, è necessario applicare direttamente la definizione di differenziabilità. In questo caso non è difficile, in quanto

$$f(0,0) = \frac{\partial f}{\partial x}(0,0) = \frac{\partial f}{\partial y}(0,0) = 0,$$

per cui occorre provare che

$$f(x,y) = o\left(\sqrt{x^2+y^2}\right),$$

ovvero che

$$\lim_{(x,y)\to(0,0)} \frac{xy^{1/3}}{\sqrt{x^2+y^2}} = 0.$$

Ma questo, passando in coordinate polari e utilizzando i metodi visti nel paragrafo 2.2., è immediato. Concludiamo che f è differenziabile anche in $(0,0)$. Nell'origine però non è stato possibile applicare la condizione sufficiente espressa dal teorema.

4.4 Derivate direzionali

Abbiamo visto che, per una funzione $f : \mathbb{R}^n \to \mathbb{R}$, la derivata parziale $\frac{\partial f}{\partial x_i}$ misura la velocità di crescita di f nella direzione dell'asse x_i. Un problema naturale è quello di misurare la velocità di crescita di f anche in direzioni diverse da quelle degli assi.

Supponiamo quindi di considerare un punto x_0 e di muoverci da lì lungo la retta individuata da un versore v. Tale retta ha equazione parametrica

$$\mathrm{x}(t) = \mathrm{x}_0 + t\mathrm{v}$$

con t parametro reale. Si noti che $t = 0$ corrisponde al punto x_0. In seguito a tale spostamento, la funzione f subirà l'incremento

$$f(\mathrm{x}_0 + t\mathrm{v}) - f(\mathrm{x}_0).$$

Siamo ora in una situazione unidimensionale. Basterà dividere per t e far tendere t a zero per ottenere il tasso di incremento di f lungo v in x_0. Precisamente:

Definizione 3.15 Siano $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, con A aperto, $\mathrm{x}_0 \in A$ e v un versore. Si dice *derivata direzionale di f rispetto al versore* v, *nel punto* x_0, il limite

$$D_{\mathrm{v}} f(\mathrm{x}_0) = \lim_{t\to 0} \frac{f(\mathrm{x}_0 + t\mathrm{v}) - f(\mathrm{x}_0)}{t} \tag{4.13}$$

purché esista finito.

Detto altrimenti: si considera la restrizione di f alla retta uscente da x_0 in direzione v, ossia

$$g(t) = f(\mathrm{x}_0 + t\mathrm{v}).$$

Il limite (4.13) coincide con

$$D_{\mathrm{v}} f(\mathrm{x}_0) = g'(0).$$

Confrontando questa definizione con quella di derivata parziale data nel paragrafo 4.1, si vede subito che *le derivate parziali sono particolari derivate direzionali*, corrispondenti ai versori canonici e_i:

$$D_{\mathrm{e}_i} f(\mathrm{x}_0) = \frac{\partial f}{\partial x_i}(\mathrm{x}_0).$$

Nel caso particolare $n = 2$, un versore v assume la forma

$$\mathrm{v} = (\cos\theta, \sin\theta),$$

dove θ è l'angolo formato da v con l'asse x. Perciò si ha:

$$D_{\mathrm{v}} f(x_0, y_0) = \lim_{t\to 0} \frac{f(x_0 + t\cos\theta, y_0 + t\sin\theta) - f(x_0, y_0)}{t},$$

o anche, ponendo $g(t) = f(x_0 + t\cos\theta, y_0 + t\sin\theta)$,

$$D_{\mathrm{v}} f(x_0, y_0) = g'(0).$$

In questo caso si ha:

$$\frac{\partial f}{\partial x}(x_0, y_0) = D_{\mathrm{v}} f(x_0, y_0) \text{ con } \mathrm{v} = \mathrm{i}, \text{ cioè } \theta = 0;$$
$$\frac{\partial f}{\partial y}(x_0, y_0) = D_{\mathrm{v}} f(x_0, y_0) \text{ con } \mathrm{v} = \mathrm{j}, \text{ cioè } \theta = \frac{\pi}{2}.$$

ESEMPI

ESEMPIO Per la funzione $z = xe^{xy}$, calcoliamo $D_{\mathrm{v}} f(0,0)$ per $\mathrm{v} = (\cos\theta, \sin\theta)$ generico. Si ha:

$$g(t) = t\cos\theta e^{t^2\cos\theta\sin\theta}$$
$$g'(t) = \cos\theta e^{t^2\cos\theta\sin\theta}(1 + 2t^2\cos\theta\sin\theta)$$

e quindi

$$D_{\mathrm{v}} f(0,0) = g'(0) = \cos\theta.$$

Per esempio, se $\theta = \pi/4$ si ha $\mathrm{v} = (\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}})$ e

$$D_{\mathrm{v}} f(0,0) = \frac{1}{\sqrt{2}}.$$

Si noti che questa funzione è C^1, quindi differenziabile nell'origine.

Per la funzione $f(x,y) = \sqrt[3]{x^2 y}$, calcoliamo $D_{\mathrm{v}} f(0,0)$ con $\mathrm{v} = (\cos\theta, \sin\theta)$:

$$g(t) = t(\cos\theta)^{2/3}(\sin\theta)^{1/3}$$
$$g'(t) = (\cos\theta)^{2/3}(\sin\theta)^{1/3},$$

per cui

$$D_{\mathrm{v}} f(0,0) = g'(0) = (\cos\theta)^{2/3}(\sin\theta)^{1/3}.$$

Perciò la funzione ha tutte le derivate direzionali in $(0,0)$. In particolare, si noti che $f_x(0,0) = f_y(0,0) = 0$. Tuttavia, questa funzione non è differenziabile nell'origine. (Questo fatto seguirà facilmente dal prossimo teorema, oppure può essere verificato direttamente dalla definizione.)

L'ultimo esempio mostra che l'esistenza di tutte le derivate direzionali, di per sé, dà un'informazione piuttosto povera sulla regolarità di una funzione: la funzione infatti potrebbe non essere differenziabile pur possedendo tutte le derivate direzionali.

Un fatto notevole è che, se la funzione è differenziabile, tutte le informazioni date dalle derivate direzionali sono implicitamente contenute nelle derivate parziali; infatti vale l'importante

TEOREMA 3.9 (FORMULA DEL GRADIENTE) *Sia $f : A \to \mathbb{R}$ con A aperto di $\mathbb{R}^n$, f differenziabile in $\mathrm{x}_0 \in A$. Allora per ogni versore v esiste la derivata direzionale $D_{\mathrm{v}} f(\mathrm{x}_0)$ e vale l'identità:*

$$D_{\mathrm{v}} f(\mathrm{x}_0) = \nabla f(\mathrm{x}_0) \cdot \mathrm{v} = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathrm{x}_0)\, v_i. \tag{4.14}$$

La derivata direzionale è in questo caso il prodotto scalare del gradiente con il versore nella cui direzione si deriva; *tutte le derivate direzionali risultano così combinazione lineare delle derivate parziali*. La (4.14) prende il nome di *formula del gradiente*.

Notiamo che nel caso $n = 2$ la formula del gradiente assume la forma:

$$D_{\mathrm{v}} f(x_0, y_0) = \nabla f(x_0, y_0) \cdot \mathrm{v} = \frac{\partial f}{\partial x}(x_0, y_0)\cos\theta + \frac{\partial f}{\partial y}(x_0, y_0)\sin\theta,$$

con $\mathrm{v} = (\cos\theta, \sin\theta)$.

DIMOSTRAZIONE. Basta ricordare la definizione di differenziabilità (4.10):

$$f(\mathrm{x}_0 + \mathrm{h}) - f(\mathrm{x}_0) = \nabla f(\mathrm{x}_0)\cdot \mathrm{h} + o(|\mathrm{h}|) \quad \text{per } \mathrm{h} \to 0$$

e applicarla all'incremento $\mathrm{h} = t\mathrm{v}$; si ha:

$$f(\mathrm{x}_0 + t\mathrm{v}) - f(\mathrm{x}_0) = \nabla f(\mathrm{x}_0)\cdot t\mathrm{v} + o(t)$$

(ricordando che $|\mathrm{v}| = 1$). Dividendo per t,

$$\frac{f(\mathrm{x}_0 + t\mathrm{v}) - f(\mathrm{x}_0)}{t} = \nabla f(\mathrm{x}_0)\cdot \mathrm{v} + \frac{o(t)}{t}$$

e facendo tendere t a zero, otteniamo che

$$D_{\mathrm{v}} f(\mathrm{x}_0) = \lim_{t\to 0} \frac{f(\mathrm{x}_0 + t\mathrm{v}) - f(\mathrm{x}_0)}{t} = \nabla f(\mathrm{x}_0)\cdot \mathrm{v},$$

poiché $\frac{o(t)}{t} \to 0$. ♦

COROLLARIO 3.10 (DIREZIONI DI MASSIMA E MINIMA CRESCITA) *Sia $f : A \to \mathbb{R}$ con A aperto di $\mathbb{R}^n$, f differenziabile in $\mathrm{x}_0 \in A$. Allora il vettore $\nabla f(\mathrm{x}_0)$ indica la direzione (e il verso) di massimo accrescimento di f, ossia la direzione corrispondente alla massima derivata direzionale; $-\nabla f(\mathrm{x}_0)$ indica la direzione corrispondente alla minima derivata direzionale (che in generale è negativa); infine, nella direzione ortogonale al gradiente le derivate direzionali sono nulle.*

DIMOSTRAZIONE. Basta applicare la formula del gradiente e chiedersi per quale versore v il prodotto scalare $\nabla f(\mathrm{x}_0)\cdot \mathrm{v}$ è, rispettivamente, massimo, minimo, nullo. ♦

Esempi

Sia $f(x,y) = e^x + \sin 2y$. Calcoliamo le derivate direzionali nell'origine.
Poiché $f \in C^1(\mathbb{R}^2)$, è sufficiente calcolare il gradiente di f,

$$\nabla f(x,y) = (e^x, 2\cos 2y);$$
$$\nabla f(0,0) = (1,2),$$

e quindi applicare la formula del gradiente: per ogni versore $\mathrm{v} = (\cos\theta, \sin\theta)$, si ha:

$$D_{\mathrm{v}} f(0,0) = \nabla f(0,0)\cdot \mathrm{v} = (1,2)\cdot(\cos\theta, \sin\theta) = \cos\theta + 2\sin\theta.$$

In particolare, la derivata direzionale massima si ha per $\mathrm{v} = \mathrm{vers}(\nabla f(0,0)) = \left(\frac{1}{\sqrt{5}}, \frac{2}{\sqrt{5}}\right)$, quella minima per $\mathrm{v} = \left(-\frac{1}{\sqrt{5}}, -\frac{2}{\sqrt{5}}\right)$, mentre nelle direzioni di $\mathrm{v} = \pm\left(\frac{2}{\sqrt{5}}, -\frac{1}{\sqrt{5}}\right)$ la derivata direzionale è nulla.

La funzione $f(x,y) = \sqrt[3]{x^2 y}$, considerata nell'esempio 4.10, *non soddisfa* la formula del gradiente nell'origine: infatti le sue derivate parziali sono nulle, mentre le derivate direzionali non sono tutte nulle (e quindi non possono essere combinazioni lineari delle derivate parziali). Ne segue che la funzione non è differenziabile nell'origine.

Sia

$$f(x,y) = \begin{cases} \frac{x^2 y}{x^4 + y^2} & \text{se } (x,y) \neq (0,0) \\ 0 & \text{se } (x,y) = (0,0) \end{cases}$$

e calcoliamo le derivate direzionali in $(0,0)$. Posto $\mathrm{v} = (\cos\theta, \sin\theta)$:

$$g(t) = f(t\cos\theta, t\sin\theta) = (\text{per } t \neq 0)\ \frac{t^3 \cos^2\theta \sin\theta}{t^4\cos^4\theta + t^2\sin^2\theta} =$$
$$= \frac{t\cos^2\theta\sin\theta}{t^2\cos^4\theta + \sin^2\theta};$$

osserviamo che l'espressione trovata rappresenta $g(t)$ anche per $t = 0$, purché sia $\sin\theta \neq 0$ (tratteremo a parte il caso $\sin\theta = 0$). Si ha:

$$g'(t) = \frac{\cos^2\theta\sin\theta\left(t^2\cos^4\theta + \sin^2\theta\right) - 2t\cos^4\theta(t\cos^2\theta\sin\theta)}{\left(t^2\cos^4\theta + \sin^2\theta\right)^2};$$

$$D_{\mathrm{v}} f(0,0) = g'(0) = \frac{\cos^2\theta\sin^3\theta}{\sin^4\theta} = \frac{\cos^2\theta}{\sin\theta} \ (\text{purché } \sin\theta \neq 0).$$

 © 978-88-08-06485-1

Se $\sin\theta = 0$, stiamo considerando l'asse x, su cui f è identicamente nulla e le relative derivate direzionali ($\pm f_x(0,0)$) sono nulle. In conclusione, esistono tutte le derivate direzionali di f nell'origine.

Osserviamo che, per contro, questa funzione non è neppure continua in $(0,0)$. Infatti, $f(0,0) = 0$, ma se calcoliamo i valori di f lungo la curva $y = x^2$ (che passa per l'origine) troviamo:

$$f(x, x^2) = \frac{x^2 x^2}{x^4 + x^4} = \frac{1}{2} \neq f(0,0) = 0$$

e pertanto f è discontinua in $(0,0)$ (si veda il paragrafo 2.2). Notiamo quindi che la sola esistenza di tutte le derivate direzionali in un punto non comporta neppure la continuità della funzione.

Riepiloghiamo:

$f \in C^1(A) \Longrightarrow$ f differenziabile in A (cioè f ha iperpiano tangente) $\Longrightarrow$ { f continua in A; f derivabile in A; f ha derivate direzionali e vale la formula del gradiente.

Invece (v. esempi 4.12 e 4.13):

f continua, derivabile, dotata di tutte le derivate direzionali $\not\Rightarrow$ f differenziabile

f derivabile, dotata di tutte le derivate direzionali $\not\Rightarrow$ f continua.

4.5 Calcolo delle derivate

TEOREMA 3.11 (FORMULE DI CALCOLO PER LE DERIVATE) *Per ogni coppia di funzioni derivabili $f, g : \mathbb{R}^n \to \mathbb{R}$, per ogni coppia di costanti $\alpha, \beta \in \mathbb{R}$, valgono le seguenti proprietà del gradiente:*

$$\nabla(\alpha f + \beta g) = \alpha \nabla f + \beta \nabla g;$$
$$\nabla(fg) = g\nabla f + f\nabla g;$$
$$\nabla\left(\frac{f}{g}\right) = \frac{g\nabla f - f\nabla g}{g^2}.$$

Esplicitamente (cioè componente per componente), le precedenti formule significano:

$$\frac{\partial}{\partial x_i}(\alpha f + \beta g) = \alpha \frac{\partial f}{\partial x_i} + \beta \frac{\partial g}{\partial x_i};$$
$$\frac{\partial}{\partial x_i}(fg) = g\frac{\partial f}{\partial x_i} + f\frac{\partial g}{\partial x_i};$$
$$\frac{\partial}{\partial x_i}\left(\frac{f}{g}\right) = \frac{g\frac{\partial f}{\partial x_i} - f\frac{\partial g}{\partial x_i}}{g^2}.$$

© 978-88-08-06485-1

Analoghe proprietà valgono per i differenziali, per esempio

$$d(\alpha f + \beta g) = \alpha df + \beta dg, \quad ecc.$$

La dimostrazione è un semplice esercizio che lasciamo allo studente.

Molto importante è anche la regola di derivazione delle funzioni composte, che in più variabili assume varie forme, a seconda del tipo di funzioni che si compongono. Vediamo ora due situazioni tipiche, mentre il caso generale sarà trattato in seguito (capitolo 4, paragrafo 2).

TEOREMA 3.12 (DERIVAZIONE DELLE FUNZIONI COMPOSTE)

1. *Siano $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, $g : I \subseteq \mathbb{R} \to \mathbb{R}$ e supponiamo che la funzione composta $h(\mathrm{x}) = g(f(\mathrm{x}))$ sia definita almeno in un intorno U di $\mathrm{x}_0 \in A$. Se f è differenziabile in x_0 e g è derivabile in $f(\mathrm{x}_0)$, allora la funzione composta*

$$h = g \circ f : U \subseteq \mathbb{R}^n \to \mathbb{R}$$

è differenziabile in x_0 e

$$\nabla h(\mathrm{x}_0) = g'(f(\mathrm{x}_0))\nabla f(\mathrm{x}_0). \tag{4.15}$$

2. *Siano $\mathrm{r} : I \subseteq \mathbb{R} \to \mathbb{R}^n$ e $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ e supponiamo che la funzione composta $g(t) = f(\mathrm{r}(t))$ sia definita almeno in un intorno J di $t_0 \in I$. Se r è derivabile in t_0 e f è differenziabile in $\mathrm{r}(t_0)$, allora la funzione composta*

$$g = f \circ \mathrm{r} : J \subseteq \mathbb{R} \to \mathbb{R}$$

è derivabile in t_0 e

$$g'(t_0) = \nabla f(\mathrm{r}(t_0)) \cdot \mathrm{r}'(t_0) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(\mathrm{r}(t_0))\, r_i'(t_0). \tag{4.16}$$

DIMOSTRAZIONE. 1. Osserviamo che, se f è differenziabile in x_0, in particolare f è derivabile in x_0; poiché g è derivabile in $f(\mathrm{x}_0)$, dal teorema sulla derivata della funzione composta per funzioni reali di variabile reale[11], si ha subito che esistono le derivate parziali

$$\frac{\partial h}{\partial x_i}(\mathrm{x}_0) = g'(f(\mathrm{x}_0))\frac{\partial f}{\partial x_i}(\mathrm{x}_0)$$

(infatti ogni derivata parziale si può vedere come derivata di una funzione di una variabile), da cui segue la formula (4.15). Questo tuttavia prova la *derivabilità* di h, ma non ancora la sua *differenziabilità* in x_0. Per dimostrare questa, occorre ripercorrere anche la dimostrazione del teorema unidimensionale che abbiamo citato.

Poiché g è differenziabile in $y_0 = f(\mathrm{x}_0)$, possiamo scrivere:

$$g(y_0 + k) = g(y_0) + kg'(y_0) + \omega(k),$$

[11] v. volume 1, capitolo 4, Teorema 4.3.

con $\frac{\omega(k)}{k} \to 0$ per $k \to 0$. Se poniamo $\omega(0) = 0$, la precedente identità risulta verificata anche per $k = 0$ e possiamo allora scrivere $\omega(k) = k\varepsilon(k)$ con $\varepsilon(k) \to 0$ per $k \to 0$, $\varepsilon(0) = 0$. Applicando questa identità a

$$k = f(x_0 + h) - f(x_0),$$

otteniamo:

$$g(f(x_0+h)) - g(f(x_0)) = [g'(f(x_0)) + \varepsilon(k)]\,[f(x_0+h) - f(x_0)] =$$

per la differenziabilità di f in x_0

$$= [g'(f(x_0)) + \varepsilon(k)]\,[\nabla f(x_0)\cdot h + o(h)]$$
$$= g'(f(x_0))\,\nabla f(x_0)\cdot h + \{g'(f(x_0))\,o(h) + \nabla f(x_0)\cdot h\varepsilon(k) + \varepsilon(k)\,o(h)\}.$$

Si tratta ora di mostrare che la quantità entro $\{\ldots\}$ è $o(h)$ per $h \to 0$. Il primo addendo lo è per definizione; quanto agli altri due, osserviamo che per $h \to 0$ anche $k \to 0$ (poiché f è continua in x_0), perciò $\varepsilon(k) \to 0$; dunque,

$$\frac{|\nabla f(x_0)\cdot h\varepsilon(k)|}{|h|} \le \frac{|\nabla f(x_0)|\,|h|\,|\varepsilon(k)|}{|h|} = |\nabla f(x_0)|\,|\varepsilon(k)| \to 0;$$
$$\frac{|\varepsilon(k)\,o(h)|}{|h|} = |\varepsilon(k)|\,\frac{|o(h)|}{|h|} \to 0,$$

e il primo punto del teorema è dimostrato.

2. Il ragionamento è analogo, anche se questa volta la formula (4.16) non è ovvia, ma discenderà proprio dalla dimostrazione. Poiché f è differenziabile in $x_0 = r(t_0)$, si ha:

$$f(x_0 + k) = f(x_0) + \nabla f(x_0)\cdot k + \omega(k),$$

con $\frac{\omega(k)}{|k|} \to 0$ per $k \to 0$. Ragionando come sopra, possiamo scrivere $\omega(k) = |k|\,\varepsilon(k)$ con $\varepsilon(k) \to 0$ per $k \to 0$, $\varepsilon(0) = 0$. Applicando questa identità a

$$k = r(t_0 + h) - r(t_0),$$

otteniamo:

$$f(r(t_0+h)) - f(r(t_0)) = \nabla f(r(t_0))\cdot[r(t_0+h) - r(t_0)] + |r(t_0+h) - r(t_0)|\,\varepsilon(k) =$$

per la derivabilità di r in t_0

$$= \nabla f(r(t_0))\cdot[r'(t_0)h + o(h)] + |r'(t_0)h + o(h)|\,\varepsilon(k)$$
$$= \nabla f(r(t_0))\cdot r'(t_0)h + \{\nabla f(r(t_0))\cdot o(h) + |r'(t_0)|\,h\varepsilon(k) + o(h)\,\varepsilon(k)\}.$$

Ragionando come sopra, si mostra ancora che la quantità $\{\ldots\}$ è $o(h)$ per $h \to 0$, il che mostra che $f(r(t))$ è differenziabile in t_0 e vale la (4.16). ϕ

Vediamo ora qualche applicazione significativa del teorema precedente.

Gradiente di una funzione radiale

Si chiama *funzione radiale* una funzione $h : \mathbb{R}^n \to \mathbb{R}$ che dipende solo dalla distanza di x dall'origine, ossia del tipo

$$h(x) = g(|x|),$$

dove $g : [0, \infty) \to \mathbb{R}$. Ponendo

$$\rho = |x| = \sqrt{\sum_{j=1}^{n} x_j^2},$$

si ha:

$$\frac{\partial}{\partial x_i}|x| = \frac{x_i}{\sqrt{\sum_{j=1}^n x_j^2}} \quad (\text{per } x \neq 0).$$

Sinteticamente,

$$\nabla\rho = \left(\frac{x_1}{\rho}, \ldots, \frac{x_n}{\rho}\right).$$

Notiamo che $|\nabla\rho| = 1$. Allora, per il punto 1 del teorema precedente,

$$\nabla h(x) = g'(|x|)\left(\frac{x_1}{|x|}, \ldots, \frac{x_n}{|x|}\right) \quad (\text{per } x \neq 0);$$
$$|\nabla h(x)| = |g'(|x|)|.$$

Esempio

Sia

$$h(x, y) = \log(x^2 + y^2) = 2\log\rho.$$

Si può affermare, senza troppi calcoli, che:

$$\nabla h(x, y) = \left(\frac{x}{\rho}, \frac{y}{\rho}\right)\cdot\frac{d}{d\rho}(2\log\rho) = \left(\frac{2x}{\rho^2}, \frac{2y}{\rho^2}\right),$$
$$|\nabla h(x, y)| = \frac{2}{\rho}.$$

Ortogonalità del gradiente con le curve di livello

Sia $f : \mathbb{R}^2 \to \mathbb{R}$ differenziabile e $f(x, y) = c$ l'equazione di una sua linea di livello. Supponiamo che questa linea ammetta una rappresentazione parametrica regolare $r = r(t)$. Allora, posto

$$g(t) = f(r(t)),$$

è, per definizione, $g(t) = c$ e dunque $g'(t) = 0$. D'altro canto

$$g'(t) = \nabla f(r(t))\cdot r'(t),$$

per cui $\nabla f(r(t))\cdot r'(t) = 0$, cioè: *il gradiente è ortogonale in ogni punto alle linee di livello della funzione*. (Infatti, $r'(t)$ è tangente alla linea di livello).

Una proprietà analoga per funzioni di n variabili è un po' più delicata da enunciarsi. Sia $f : \mathbb{R}^n \to \mathbb{R}$ differenziabile in x_0 e $f(x) = c$ l'equazione di una sua superficie di livello. Supponiamo che questa superficie ammetta piano tangente in x_0 (accontentandoci per il momento del significato intuitivo di questa affermazione[12]). Provare che

[12] Una trattazione più rigorosa delle superfici in $\mathbb{R}^n$ sarà fornita nel capitolo 4.

 © 978-88-08-06485-1

$\nabla f(x_0)$ è ortogonale alla superficie di livello in x_0 significa provare che è ortogonale a ciascuna curva regolare tracciata sulla superficie e passante per x_0. Sia $\mathbf{r} = \mathbf{r}(t)$ una di queste curve; ciò significa che $\mathbf{r}(t_0) = x_0$ e che

$$g(t) = f(\mathbf{r}(t)) = c \text{ per ogni } t.$$

Dunque $g'(t) = 0$ e, poiché

$$g'(t_0) = \nabla f(\mathbf{r}(t_0)) \cdot \mathbf{r}'(t_0),$$

ricaviamo appunto l'ortogonalità del *gradiente* $\nabla f(x_0)$ con il vettore tangente alla curva $\mathbf{r}'(t_0)$; poiché questo vale per ogni curva con le proprietà dette, ne deduciamo l'asserto.

Equazione del trasporto

Consideriamo l'*equazione differenziale a derivate parziali*[13]:

$$c\frac{\partial u}{\partial x} + \frac{\partial u}{\partial t} = 0 \tag{4.17}$$

con c costante. La (4.17), detta *equazione del trasporto*, può avere vari significati fisici. Per esempio, supponiamo che una sostanza inquinante si muova a velocità costante c con la corrente di un fiume. Sia u la concentrazione di tale sostanza nell'acqua e supponiamo che essa sia costante all'interno di ogni sezione del fiume; dunque u sarà funzione del tempo t e di un'unica variabile spaziale, x, che rappresenta l'ascissa lungo il fiume, schematizzato come una retta. Sotto tali ipotesi, si può dimostrare che u soddisfa proprio la (4.17).

Questa equazione si può risolvere, osservando che il primo membro è la derivata direzionale di $u(x,t)$ nella direzione del vettore $(c,1)$. Perciò, se tale derivata è zero, significa che u è costante lungo ogni retta del tipo $x = ct + q$, per cui u dipende solo dalla quantità $(x - ct)$:

$$u(x,t) = v(x - ct)$$

con $v : \mathbb{R} \to \mathbb{R}$ generica funzione derivabile. (Come controprova, si calcoli $cu_x - u_t$ in funzione di v, utilizzando il teorema di derivazione delle funzioni composte, e si verifichi che questa quantità si annulla qualunque sia v).

La funzione v ha il significato di profilo iniziale, in quanto

$$u(x,0) = v(x).$$

Perciò, nota la concentrazione $v(x)$ all'istante $t = 0$, la concentrazione in ogni istante è data da $u(x,t) = v(x - ct)$.

Come si vede dalla figura 3.14(b), il grafico di $u(x,t)$ come funzione di x, ad istanti successivi, è semplicemente traslato rispetto al grafico di $v(x)$ e prende il nome di *onda progressiva*. (La concentrazione di inquinante è semplicemente "trasportata" dalla corrente, da cui il nome dell'equazione).

[13] Qualcosa di più su queste equazioni differenziali sarà detto nel paragrafo 5.1.

© 978-88-08-06485-1

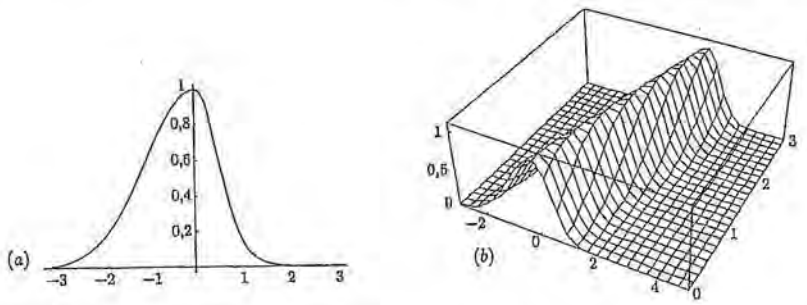


Figura 3.14. (a) Grafico della concentrazione iniziale $v(x)$; (b) Grafico della funzione $u(x,t) = v(x - ct)$.

Esempio

La temperatura in una certa regione del piano è assegnata da:

$$u(x,y) = e^{-x+2y}.$$

Un punto si muove nel piano secondo la legge oraria:

$$\begin{cases} x = t(1-t) \\ y = t^3 \end{cases} \qquad t \in (0, +\infty).$$

Calcolare la velocità con cui un osservatore posto nel punto mobile sente variare la temperatura.

Possiamo ragionare in due modi. Il primo è applicare la formula di derivazione della funzione composta, e, posto

$$g(t) = u(x(t), y(t)),$$

calcolare:

$$u_x = -e^{-x+2y};\ u_y = 2e^{-x+2y};$$
$$x'(t) = 1 - 2t;\ y'(t) = 3t^2$$

e quindi

$$g'(t) = u_x(x(t),y(t))x'(t) + u_y(x(t),y(t))y'(t) =$$
$$= -e^{-t(1-t)+2t^3}(1-2t) + 2e^{-t(1-t)+2t^3}(3t^2) = e^{-t(1-t)+2t^3}(-1+2t+6t^2).$$

Il secondo metodo è scrivere per prima cosa esplicitamente $g(t)$ come composizione di funzioni:

$$g(t) = e^{-t(1-t)+2t^3}$$

e quindi calcolare direttamente

$$g'(t) = e^{-t(1-t)+2t^3}(-1+2t+6t^2).$$

Ovviamente, il risultato non cambia.

Come mostra il confronto tra l'ultimo esempio e i precedenti, le formule di derivazione della funzione composta in più variabili sono utili soprattutto quando una almeno delle

 © 978-88-08-06485-1

due funzioni non si possa scrivere esplicitamente (ossia quando vogliamo ricavare una formula avente un significato generale, valido per ogni funzione di un certo tipo).

Un'altra applicazione di tipo teorico del teorema precedente consiste nel provare un analogo del teorema di Lagrange per funzioni di più variabili.

Ricordiamo anzitutto che il segmento di estremi $x_0, x_1 \in \mathbb{R}^n$ (percorso da x_0 a x_1) ha equazioni parametriche:

$$r(t) = tx_1 + (1-t)x_0 \text{ per } t \in [0,1];$$

l'insieme dei punti di $\mathbb{R}^n$ che giacciono sul segmento (ossia il sostegno di $r(t)$) verrà denotato col simbolo $[x_0, x_1]$. Possiamo allora enunciare il:

TEOREMA 3.13 (DEL VALOR MEDIO) *Sia $A \subset \mathbb{R}^n$ un aperto e $f : A \to \mathbb{R}$ una funzione differenziabile in A. Allora, per ogni coppia di punti $x_0, x_1 \in A$, esiste un punto $x^* \in [x_0, x_1]$ tale che:*

$$f(x_1) - f(x_0) = \nabla f(x^*) \cdot (x_1 - x_0).$$

In particolare:

$$|f(x_1) - f(x_0)| \leq |\nabla f(x^*)| \cdot |x_1 - x_0|.$$

DIMOSTRAZIONE. Sia

$$r(t) = tx_1 + (1-t)x_0 \text{ per } t \in [0,1]$$

il segmento di estremi x_0, x_1 e consideriamo la funzione

$$g(t) = f(r(t)), \text{ per } t \in [0,1]$$

continua e derivabile in $[0,1]$ (per il punto 2 del teorema 3.12); per il teorema di Lagrange unidimensionale, esiste $t_0 \in (0,1)$ tale che

$$g(1) - g(0) = g'(t_0).$$

Ma:

$$g(1) - g(0) = f(r(1)) - f(r(0)) = f(x_1) - f(x_0);$$
$$g'(t_0) = \nabla f(r(t_0)) \cdot r'(t_0) = \nabla f(r(t_0)) \cdot (x_1 - x_0),$$

da cui la tesi, ponendo $x^* = r(t_0)$. ∎

Esercizi

Derivate parziali, piano tangente, linee di livello

1 Si calcolino le derivate parziali della funzione

$$f(x,y) = xe^{x^2}\sqrt{y}$$

in tutti i punti in cui esistono. Si dica qual è il più grande insieme aperto del piano in cui le derivate parziali sono continue (e quindi certamente f è differenziabile).

2 Si scriva l'equazione del piano tangente alla superficie grafico di

$$f(x,y) = e^x \sin y$$

nel generico punto (x_0, y_0); si particolarizzi poi la formula nel punto $(1, \pi)$.

3 Si calcoli la derivata direzionale della funzione dell'esercizio precedente nel punto $(1,\pi)$ rispetto al versore $(\frac{3}{5}, \frac{4}{5})$, prima usando la definizione di derivata direzionale e poi usando la formula del gradiente.

4 Scrivere lo sviluppo (4.7) per le seguenti funzioni:
a) $f(x,y) = e^{x+y}$
b) $f(x,y) = \sin 2x + y\cos x$
c) $f(x,y) = \sqrt{1+x+x^2+y^2}$
d) $f(x,y) = \sqrt[3]{1+x^2+y^2}$

5 Si disegnino alcune linee di livello della funzione

$$f(x,y) = x^2 + \frac{y^2}{4}.$$

Si traccino poi alcuni vettori spiccati da punti di tali linee, aventi la direzione e il verso del gradiente di f (non occorrono calcoli!).

6 Si calcolino le derivate parziali della funzione

$$f(x,y) = \begin{cases} \dfrac{xy^2}{x^2+y^2} & \text{se } (x,y) \neq (0,0) \\ 0 & \text{se } (x,y) = (0,0) \end{cases}$$

in ogni punto del piano (per l'origine, applicare la definizione).

Si calcolino poi, in base alla definizione, le derivate direzionali di f, nell'origine, rispetto al generico versore $(\cos\theta, \sin\theta)$.

La formula del gradiente è verificata? Spiegare il risultato.

7 Si scriva l'equazione dell'*iperpiano tangente all'ipersuperficie* grafico della funzione di tre variabili $w = \sqrt{x^2+y^2+z^2}$ nel punto $(1,1,2)$.

8 In elettrostatica, il campo elettrico $\mathbf{E}$ è, in ogni punto dello spazio fisico, il gradiente di una funzione $U(x,y,z)$, detta potenziale elettrostatico. Il potenziale elettrostatico generato da una carica puntiforme è

$$U(x,y,z) = \frac{k}{r},$$

con $r = \sqrt{x^2+y^2+z^2}$ e k costante. Calcolare il campo elettrico generato dalla carica puntiforme; calcolarne poi il modulo.

9 La forza di attrazione gravitazionale $\mathbf{F}$ generata da un corpo è, in ogni punto dello spazio fisico, il gradiente di una funzione $U(x,y,z)$, detta potenziale gravitazionale. Il potenziale gravitazionale generato da una sfera omogenea di raggio R è:

$$U(x,y,z) = \begin{cases} \frac{4}{3}\pi R^3 \frac{\rho}{r} & \text{se } r \geq R \\ 2\pi\rho(R^2 - \frac{r^2}{3}) & \text{se } r < R \end{cases}$$

dove ρ è una costante (la densità della sfera) e $r = \sqrt{x^2+y^2+z^2}$.

a) Verificare che il potenziale è continuo in tutto lo spazio.

b) Calcolare la forza $\mathbf{F}$ in tutto lo spazio.
c) Stabilire se il campo vettoriale $\mathbf{F}$ è continuo[14] in tutto lo spazio oppure no.

29 (Si veda l'esercizio precedente per il significato dei termini). Il potenziale gravitazionale generato da uno strato sferico (o corona sferica) di raggio interno R_1 e raggio esterno R_2 è:

$$U(x,y,z) = \begin{cases} \frac{4}{3}\pi\rho(R_2^3 - R_1^3)\frac{1}{r} & \text{se } r > R_2 \\ 2\pi\rho(R_2^2 - \frac{r^2}{3}) - \frac{4}{3}\pi\rho\frac{R_1^3}{r} & \text{se } R_1 < r \le R_2 \\ 2\pi\rho(R_2^2 - R_1^2) & \text{se } r \le R_1. \end{cases}$$

a) Verificare che il potenziale è continuo in tutto lo spazio.
b) Calcolare la forza $\mathbf{F}$ in tutto lo spazio.
c) Stabilire se il campo vettoriale $\mathbf{F}$ è continuo in tutto lo spazio oppure no.

30 Se $u(x,y,z)$ è la temperatura nel punto (x,y,z), in regime stazionario (ossia: in ogni punto la temperatura non cambia nel tempo), si definisce densità di corrente termica il vettore

$$\mathbf{q} = -\chi\nabla u$$

dove $\chi > 0$ è la conducibilità termica del mezzo (può essere una costante, oppure anch'essa una funzione di (x,y,z), a valori reali). Supponiamo che la temperatura sia

$$u(x,y,z) = e^{-2x^2-3y^2-z^2}.$$

Si calcoli la densità di corrente termica e il modulo di tale vettore, nell'ipotesi $\chi = 1$.

31 Supponiamo che una collina sia rappresentata, in prima approssimazione, dal grafico della funzione

$$z = f(x,y) = 1 - 2x^2 - y^4.$$

Un sentiero che scende dalla cima della collina è disegnato su una mappa di quella regione come una curva di equazione polare $\rho = 2\theta$, $\theta \in [0, 3\pi]$. Scrivere, in funzione di θ, la velocità con cui il sentiero fa perdere quota, cioè $\frac{dz}{d\theta}$. Ottenere il risultato utilizzando entrambi i metodi illustrati nell'esempio 4.15. (*Suggerimento*: scrivere per prima cosa le equazioni parametriche della curva nella forma $x = x(\theta), y = y(\theta)$).

32 Generalizzando il risultato visto nel paragrafo 4.5 sull'equazione del trasporto unidimensionale, mostrare che se $f : \mathbb{R}^n \to \mathbb{R}$ è una qualunque funzione differenziabile, la funzione $u : \mathbb{R}^{n+1} \to \mathbb{R}$ definita da $u(\mathrm{x}, t) = f(\mathrm{x} - \mathrm{c}t)$, con c vettore costante in $\mathbb{R}^n$, soddisfa l'equazione del trasporto:

$$\mathrm{c} \cdot \nabla_x u + \frac{\partial u}{\partial t} = 0.$$

Differenziabilità

Alcuni dei prossimi esercizi, oltre a potersi svolgere coi metodi illustrati in questo capitolo, possono essere affrontati in modo più semplice utilizzando i risultati sulle funzioni positivamente omogenee che sono contenuti nei Complementi in fondo al capitolo (paragrafo 9.2).

33 Si verifichino i seguenti fatti per la funzione $f(x,y) = x\sqrt[3]{y}$.

[14] Cioè se ogni sua componente è continua. Nel capitolo 4 ci occuperemo sistematicamente di estendere alle funzioni a valori vettoriali i concetti studiati in questo capitolo su limiti, continuità e calcolo differenziale.

a) La funzione è derivabile anche nell'origine.
b) Le derivate parziali non sono continue nell'origine (perciò non si può applicare la condizione sufficiente di differenziabilità).
c) f è differenziabile nell'origine (occorre applicare la definizione e calcolare il limite opportuno, coi metodi visti nel paragrafo 2.2, oppure applicare i risultati sulle funzioni positivamente omogenee che si trovano nei Complementi).

34 Si verifichi, direttamente dalla definizione, che la seguente funzione è differenziabile nell'origine:

$$f(x,y) = \begin{cases} \frac{xy^3}{x^2+y^2} & \text{per } (x,y) \neq (0,0) \\ 0 & \text{per } (x,y) = (0,0). \end{cases}$$

Si ottenga poi lo stesso risultato dimostrando che le funzioni f_x, f_y sono continue nell'origine. (Notare che la continuità di queste funzioni non è ovvia, ma va provata calcolandone il limite).

Infine, si riottenga ancora lo stesso risultato applicando i risultati sulle funzioni positivamente omogenee (v. Complementi).

35 Si consideri la seguente funzione:

$$f(x,y) = \begin{cases} \frac{x^3y}{x^2+y^4} & \text{per } (x,y) \neq (0,0) \\ 0 & \text{per } (x,y) = (0,0). \end{cases}$$

Dire in quale insieme del piano f è continua, derivabile, differenziabile.

36 Si consideri la seguente funzione:

$$f(x,y) = \begin{cases} \frac{x^{5/3}y^{4/3}}{x^2+y^2} & \text{per } (x,y) \neq (0,0) \\ 0 & \text{per } (x,y) = (0,0). \end{cases}$$

Calcolare f_x, f_y in tutti i punti di $\mathbb{R}^2$; stabilire se le derivate parziali sono continue in $(0,0)$; stabilire se f è differenziabile in $(0,0)$ direttamente dalla definizione, oppure utilizzando i risultati sulle funzioni positivamente omogenee (v. Complementi).

5 DERIVATE DI ORDINE SUPERIORE E APPROSSIMAZIONI SUCCESSIVE

5.1 Derivate di ordine superiore

Supponiamo che una funzione reale di due variabili, $f(x,y)$, possieda per esempio la derivata parziale f_x in tutto un insieme aperto A. Possiamo allora chiederci se, a sua volta, la funzione f_x sia derivabile nei punti di tale insieme. In caso affermativo calcoliamo le derivate parziali di f_x, che si indicano coi simboli

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) \qquad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}\right)$$

Analogamente, le derivate parziali di f_y (se esistono) sono:

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}\right) \qquad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)$$

 © 978-88-08-06485-1

Queste quattro funzioni si dicono *derivate parziali seconde di* f. Si usano anche i simboli:

$$f_{xy} \qquad D^2_{xy}f \qquad \partial^2_{xy}f \qquad D^2_{12} \qquad f_{xx}\dots$$

Più in generale, data una funzione $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ con A aperto, per cui una derivata parziale f_{x_i} esiste in tutto A, possiamo chiederci se questa funzione sia derivabile rispetto a una certa variabile x_j in A. In caso affermativo, si indicherà tale derivata col simbolo

$$\frac{\partial^2 f}{\partial x_j \partial x_i} = \frac{\partial}{\partial x_j}\left(\frac{\partial f}{\partial x_i}\right).$$

Esempio

Sia $f(x,y) = x^2 \sin y$. Calcoliamo:

$$\frac{\partial f}{\partial x} = 2x\sin y; \qquad \frac{\partial f}{\partial y} = x^2\cos y;$$

$$\frac{\partial^2 f}{\partial x^2} = \frac{\partial}{\partial x}(2x\sin y) = 2\sin y; \qquad \frac{\partial^2 f}{\partial y \partial x} = \frac{\partial}{\partial y}(2x\sin y) = 2x\cos y;$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial}{\partial x}(x^2\cos y) = 2x\cos y; \qquad \frac{\partial^2 f}{\partial y^2} = \frac{\partial}{\partial y}(x^2\cos y) = -x^2\sin y.$$

Notiamo che le due derivate *miste*, f_{xy} e f_{yx}, sono uguali, nonostante la funzione di partenza non mostri particolari simmetrie nelle due variabili. Questo fatto ha una validità generale:

Teorema 3.14 (di Schwarz) *Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ con A aperto. Supponiamo che (per certi indici $i, j \in \{1, 2, \dots, n\}$) le derivate seconde miste $f_{x_i x_j}, f_{x_j x_i}$ esistano in un intorno di un punto x_0 e siano entrambe continue in x_0; allora esse coincidono in x_0.*

In particolare, se le derivate seconde miste $f_{x_i x_j}, f_{x_j x_i}$ esistono e sono continue in A, allora esse coincidono in tutto A.

Una funzione che ha tutte le derivate parziali seconde continue in un aperto A si dice *di classe* $C^2(A)$. Sotto tale ipotesi, le derivate parziali prime hanno derivate parziali continue, quindi sono differenziabili, per il Teorema 3.8; in particolare, anche le derivate parziali prime sono continue e allora f è differenziabile, anzi di classe $C^1(A)$. Riassumendo:

se $f \in C^2(A)$, allora $f \in C^1(A)$ (in particolare, f è differenziabile); le derivate parziali prime di f sono differenziabili; le derivate parziali seconde di f sono continue; in particolare, le derivate seconde miste di f sono uguali.

Dimostrazione del teorema di Schwarz. Dimostriamo il teorema nel caso di una funzione di due variabili, $f(x,y)$. Dalla dimostrazione stessa risulterà evidente che questo ragionamento ha un valore generale (n qualsiasi).

© 978-88-08-06485-1

Sia $x_0 = (x_0, y_0)$. Almeno in un intorno di x_0 la funzione f è definita, esistono le derivate seconde miste di f e, a maggior ragione, le derivate parziali prime; scegliamo un quadrato $[x_0 - \delta, x_0 + \delta] \times [y_0 - \delta, y_0 + \delta]$ contenuto in questo intorno e consideriamo la seguente quantità:

$$Q = f(x_0 + h, y_0 + k) - f(x_0 + h, y_0) - f(x_0, y_0 + k) + f(x_0, y_0),$$

ben definita purché $|h| < \delta, |k| < \delta$.

Per un h fissato, consideriamo la funzione di una variabile:

$$u(t) = f(x_0 + h, t) - f(x_0, t).$$

Notiamo che $Q = u(y_0 + k) - u(y_0)$. Inoltre, nelle nostre ipotesi la funzione u risulta derivabile in tutto $[y_0, y_0 + k]$. Applicando allora il Teorema di Lagrange a u su questo intervallo, otteniamo:

$$Q = u(y_0 + k) - u(y_0) = ku'(y_0 + \delta_1 k)$$

(per un opportuno $\delta_1 \in (0,1)$, dipendente da h e k)

$$= k\left[\frac{\partial f}{\partial y}(x_0 + h, y_0 + \delta_1 k) - \frac{\partial f}{\partial y}(x_0, y_0 + \delta_1 k)\right] =$$

(applicando ancora il teorema di Lagrange alla funzione $v(s) = \frac{\partial f}{\partial y}(s, y_0 + \delta_1 k)$)

$$= kh\left[\frac{\partial}{\partial x}\frac{\partial f}{\partial y}(x_0 + \delta_2 h, y_0 + \delta_1 k)\right]$$

(per un opportuno $\delta_2 \in (0,1)$, dipendente da h e k). Ora ripetiamo un discorso simile, ma scrivendo questa volta:

$$Q = w(x_0 + h) - w(x_0)$$

con

$$w(t) = f(t, y_0 + k) - f(t, y_0).$$

Con passaggi analoghi si ha:

$$Q = hw'(x_0 + \delta_3 h) = h\left[\frac{\partial f}{\partial x}(x_0 + \delta_3 h, y_0 + k) - \frac{\partial f}{\partial x}(x_0 + \delta_3 h, y_0)\right]$$
$$= hk\left[\frac{\partial}{\partial y}\frac{\partial f}{\partial x}(x_0 + \delta_3 h, y_0 + \delta_4 k)\right]$$

per opportuni $\delta_3, \delta_4 \in (0,1)$, dipendenti da h e k. Uguagliando le due diverse espressioni trovate per Q e semplificando per hk, abbiamo:

$$\frac{\partial^2 f}{\partial x \partial y}(x_0 + \delta_2 h, y_0 + \delta_1 k) = \frac{\partial^2 f}{\partial y \partial x}(x_0 + \delta_3 h, y_0 + \delta_4 k).$$

Facciamo ora tendere h, k a zero. Si ha

$$(x_0 + \delta_2 h, y_0 + \delta_1 k) \to (x_0, y_0);$$
$$(x_0 + \delta_3 h, y_0 + \delta_4 k) \to (x_0, y_0)$$

(perché $\delta_i \in (0,1)$, per $i = 1, \dots, 4$) e, per la continuità delle derivate seconde miste, otteniamo

$$\frac{\partial^2 f}{\partial x \partial y}(x_0, y_0) = \frac{\partial^2 f}{\partial y \partial x}(x_0, y_0),$$

che è la tesi.

 © 978-88-08-06485-1

Osserviamo infine che, poiché l'enunciato del teorema coinvolge le derivate parziali di f rispetto a sole due variabili, anche nel caso di funzioni di n variabili questa stessa dimostrazione è valida: le $n-2$ variabili rispetto alle quali non si calcolano le derivate giocano il ruolo (ininfluente) di parametri fissati.

Notiamo che, per quanto l'ipotesi di questo teorema sia solitamente verificata, esistono casi in cui una funzione possiede derivate seconde miste diverse tra loro:

Esempio

Sia

$$f(x,y)=\begin{cases}\frac{xy(x^2-y^2)}{x^2+y^2} & \text{se } (x,y)\neq(0,0)\\ 0 & \text{se } (x,y)=(0,0).\end{cases}$$

Si verifica che $f\in C^1(\mathbb{R}^2)$, mentre le derivate parziali seconde sono discontinue nell'origine[15].

Il teorema di Schwarz non è quindi applicabile e non possiamo affermare a priori che sia $f_{xy}(0,0)=f_{yx}(0,0)$. Calcoliamo dunque esplicitamente queste derivate. Un calcolo un po' noioso dà:

$$\frac{\partial f}{\partial x}(x,y)=\frac{y\left(x^4+4x^2y^2-y^4\right)}{\left(x^2+y^2\right)^2};\qquad \frac{\partial f}{\partial y}(x,y)=\frac{x\left(x^4-4x^2y^2-y^4\right)}{\left(x^2+y^2\right)^2};$$

$$\frac{\partial f}{\partial x}(0,y)=-y;\qquad \frac{\partial^2 f}{\partial y\partial x}(0,0)=\frac{\partial}{\partial y}\left(\frac{\partial f}{\partial x}(0,y)\right)_{/y=0}=-1;$$

$$\frac{\partial f}{\partial y}(x,0)=x;\qquad \frac{\partial^2 f}{\partial x\partial y}(0,0)=\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial y}(x,0)\right)_{/x=0}=1$$

pertanto $f_{xy}(0,0)\neq f_{yx}(0,0)$.

Le definizioni di derivate parziali successive e il teorema di Schwarz si generalizzano a derivate di ordine k. Per esempio, per una funzione di tre variabili $f(x,y,z)$, di classe $C^4(A)$, si può affermare l'identità:

$$\frac{\partial^4 f}{\partial^2 x\partial y\partial z}=\frac{\partial^4 f}{\partial z\partial x\partial y\partial x}.$$

Più in generale:

Definizione 3.16 Diremo che una funzione $f:A\subseteq\mathbb{R}^n\to\mathbb{R}$, con A aperto, è di classe $C^k(A)$ se tutte le derivate parziali fino all'ordine k esistono e sono continue in A:

$$\frac{\partial^r f}{\partial x_{i_1}\partial x_{i_2}\dots\partial x_{i_r}}\text{ continue in } A,$$

per ogni $i_1,i_2,\dots,i_r\in\{1,2,\dots,n\}$, $1\le r\le k$.

[15] Tutto ciò può essere verificato in base alle definizioni e ai metodi visti in questo capitolo, oppure, più semplicemente, applicando i risultati sulle funzioni positivamente omogenee (Complementi, paragrafo 9.2): f è omogenea di grado 2 e regolare fuori dall'origine; quindi le sue derivate parziali prime sono funzioni omogenee di grado 1, perciò sono continue (teorema 3.32); dunque, per il teorema 3.8 si ha che $f\in C^1(\mathbb{R}^2)$; le derivate parziali seconde, invece, sono funzioni omogenee di grado zero (e non costanti), perciò discontinue nell'origine.

© 978-88-08-06485-1

Solitamente, quando si considerano derivate di ordine k di una funzione f di più variabili, ci si mette proprio in questa ipotesi; per il teorema di Schwarz, questo significa che nelle derivate parziali miste di ogni ordine $\le k$ si può scambiare l'ordine di derivazione. Questo consente anche di introdurre una notazione sintetica, per le derivate successive di una funzione di n variabili, detta *notazione a multiindice*:

Definizione 3.17 Chiamiamo *multiindice* (a n componenti) ogni n-upla di interi non negativi

$$\alpha=(\alpha_1,\alpha_2,\dots,\alpha_n).$$

L'intero

$$|\alpha|=\alpha_1+\alpha_2+\dots+\alpha_n$$

si chiama *altezza del multiindice* α. Data una funzione $f\in C^k(A)$ si pone allora, se $|\alpha|\le k$,

$$D^\alpha f(x)=\frac{\partial^\alpha f}{\partial x^\alpha}(x)=\frac{\partial^{|\alpha|}f(x)}{\partial_{x_1}^{\alpha_1}\partial_{x_2}^{\alpha_2}\dots\partial_{x_n}^{\alpha_n}}\text{ per ogni } x\in A.$$

Il numero $|\alpha|$ è l'*ordine della derivata*.

Esempio

Se $f\in C^4(\mathbb{R}^5)$, la derivata parziale

$$\frac{\partial^4 f(x)}{\partial_{x_1}\partial_{x_3}^2\partial_{x_5}}$$

si può indicare con $D^\alpha f(x)$ con $\alpha=(1,0,2,0,1)$ e $|\alpha|=4$. Infatti,

$$D^{(1,0,2,0,1)}f(x)=\frac{\partial^{1+0+2+0+1}f(x)}{\partial_{x_1}^1\partial_{x_2}^0\partial_{x_3}^2\partial_{x_4}^0\partial_{x_5}^1}=\frac{\partial^4 f(x)}{\partial_{x_1}\partial_{x_3}^2\partial_{x_5}}$$

(per esempio, $\partial_{x_2}^0$ significa che non calcoliamo nessuna derivata rispetto a x_2).

Si noti che la notazione a multiindice *presuppone* che l'ordine in cui eseguiamo le derivate sia indifferente (teorema di Schwarz), il che consente di raggruppare le derivate rispetto alla medesima variabile. In altre parole, per precisare di quale derivata parziale si tratta, è sufficiente dire quante volte deriviamo rispetto a ciascuna delle variabili $x_1,x_2,\dots,x_n$ (non avendo invece importanza l'ordine in cui eseguiamo tali derivate).

Equazioni a derivate parziali

Le derivate parziali successive (e in particolare le seconde) intervengono nella formulazione di numerosissime leggi fisiche. Si dice *equazione a derivate parziali* un'equazione differenziale in cui la funzione incognita sia una funzione di più variabili, che compare nell'equazione attraverso alcune sue derivate, fatte (almeno) rispetto a due variabili diverse. Abbiamo già incontrato un'equazione di questo tipo (equazione del trasporto) nel paragrafo 4.5; ora ne vediamo altre di particolare interesse applicativo.

Esempi

5.4 L'*equazione di Laplace* (in tre variabili):

$$\Delta u(x,y,z) = 0$$

ossia

$$\frac{\partial^2 u}{\partial x^2} + \frac{\partial^2 u}{\partial y^2} + \frac{\partial^2 u}{\partial z^2} = 0$$

è soddisfatta dal potenziale elettrostatico (o gravitazionale) u nei punti dello spazio privi di carica elettrica (o di materia, rispettivamente). In generale, si dice *operatore di Laplace* (o *laplaciano*) l'operatore differenziale (in n variabili)

$$\Delta = \sum_{i=1}^{n} \frac{\partial^2}{\partial x_i^2},$$

e le funzioni che soddisfano l'equazione $\Delta u = 0$ si dicono *funzioni armoniche*.

5.5 L'*equazione del calore* nelle tre variabili spaziali (x,y,z) e nella variabile temporale t ha la forma

$$u_t = \kappa \Delta u$$

ed è soddisfatta dalla temperatura u di un corpo omogeneo e isotropo rispetto alla conduzione, sotto l'ipotesi che all'istante t e nel punto (x,y,z) non vi siano sorgenti o pozzi di calore distribuiti; κ è il coefficiente di diffusione e dipende dalle caratteristiche termiche del materiale.

5.6 L'*equazione della corda vibrante* è:

$$u_{tt} = c^2 u_{xx}.$$

Questa equazione descrive le piccole vibrazioni verticali di una corda perfettamente elastica e flessibile (ad es. una corda di violino) che, in posizione di riposo, coincide con l'asse x. La funzione $u(x,t)$ rappresenta l'ordinata al tempo t del punto di ascissa x. La costante c è la velocità di propagazione dell'onda.

5.2 Differenziale secondo, matrice hessiana, formula di Taylor al secondo ordine

Come già nel caso delle funzioni di una variabile, uno degli strumenti più utili del calcolo differenziale che coinvolge l'uso delle derivate di ordine superiore è la formula di Taylor. In questo paragrafo ci concentreremo sulla formula di Taylor del second'ordine. Questa peraltro è già sufficiente per due applicazioni significative: lo studio dei punti di massimo e minimo di una funzione (che tratteremo nel prossimo paragrafo 6) e quello delle funzioni convesse, a cui dedicheremo il paragrafo 7. Nei Complementi in fondo al capitolo daremo invece un cenno sulla formula di Taylor di ordine qualunque (paragrafo 9.3).

DEFINIZIONE 3.18 Se $f \in C^2(A)$ e $x_0 \in A$, si dice *differenziale secondo di f in x_0* la funzione[16]

$$d^2 f(x_0) : \mathbb{R}^n \to \mathbb{R}$$

[16] Come vedremo nel paragrafo 6.3, una funzione di questo tipo si chiama *forma quadratica*.

definita da

$$d^2 f(x_0) : h \longmapsto \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) h_i h_j. \tag{5.1}$$

Con abuso di scrittura, si indica spesso col medesimo simbolo $d^2 f(x_0)$ anche il valore del differenziale calcolato in corrispondenza dell'incremento h, ossia si scrive (analogamente a quanto si fa con il differenziale primo):

$$d^2 f(x_0) = \sum_{i=1}^{n}\sum_{j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0) h_i h_j.$$

I coefficienti $\frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)$ che compaiono nel differenziale secondo possono essere ordinati in una matrice $n \times n$ detta *matrice hessiana di f in x_0*:

$$H_f(x_0) = \begin{pmatrix} f_{x_1x_1}(x_0) & f_{x_1x_2}(x_0) & \dots & f_{x_1x_n}(x_0) \\ f_{x_2x_1}(x_0) & f_{x_2x_2}(x_0) & \dots & f_{x_2x_n}(x_0) \\ \dots & \dots & \dots & \dots \\ f_{x_nx_1}(x_0) & f_{x_nx_2}(x_0) & \dots & f_{x_nx_n}(x_0) \end{pmatrix}.$$

In particolare, se f è una funzione di due variabili:

$$H_f(x_0,y_0) = \begin{pmatrix} f_{xx}(x_0,y_0) & f_{xy}(x_0,y_0) \\ f_{yx}(x_0,y_0) & f_{yy}(x_0,y_0) \end{pmatrix}$$

e il differenziale secondo applicato all'incremento (h,k) assume la forma:

$$d^2 f(x_0,y_0) = f_{xx}(x_0,y_0)h^2 + 2f_{xy}(x_0,y_0)hk + f_{yy}(x_0,y_0)k^2.$$

Per il teorema di Schwarz, se f è di classe $C^2(A)$ la matrice hessiana è *simmetrica* in ogni punto di A.

Se la differenziabilità di una funzione di più variabili permette di approssimarne il grafico con quello del suo piano tangente, il fatto che una funzione sia di classe $C^2(A)$ (o più regolare ancora) permette di migliorare l'approssimazione. Cominciamo a stabilire una formula di questo tipo che coinvolge le derivate fino al second'ordine:

TEOREMA 3.15 (FORMULA DI TAYLOR, RESTO SECONDO LAGRANGE)
Sia $f \in C^2(A)$; per ogni $x_0 \in A$ e $h \in \mathbb{R}^n$ tale che $x_0 + h \in A$, esiste un numero reale $\delta \in (0,1)$, dipendente da x_0 e h, tale che:

$$f(x_0+h) = f(x_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0) h_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0 + \delta h) h_i h_j. \tag{5.2}$$

DIMOSTRAZIONE. Consideriamo la funzione di una variabile

$$g(t) = f(x_0 + th),$$

 © 978-88-08-06485-1

che sotto le nostre ipotesi risulta di classe $C^2[0,1]$. Scriviamo la formula di Taylor (unidimensionale) del second'ordine per g, con resto secondo Lagrange:

$$g(1) = g(0) + g'(0) + \frac{1}{2}g''(\delta) \tag{5.3}$$

per un opportuno $\delta \in (0,1)$ che dipende da g, perciò da f, x_0 e h. Calcoliamo ora:

$$g(0) = f(x_0); \quad g(1) = f(x_0 + h);$$
$$g'(t) = \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0 + th)\, h_i;$$
$$g''(t) = \sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0 + th)\, h_i h_j.$$

Sostituendo $g(1)$, $g(0)$, $g'(0)$, $g''(\delta)$ nello sviluppo (5.3) si ha la (5.2). ♦

TEOREMA 3.16 (FORMULA DI TAYLOR, RESTO SECONDO PEANO)
Sia $f \in C^2(A)$. Per ogni $x_0 \in A$ vale la formula:

$$f(x_0 + h) = f(x_0) + \sum_{i=1}^{n} \frac{\partial f}{\partial x_i}(x_0)\, h_i + \frac{1}{2}\sum_{i,j=1}^{n} \frac{\partial^2 f}{\partial x_i \partial x_j}(x_0)\, h_i h_j + o\left(|h|^2\right) \tag{5.4}$$

per $h \to 0$.

Il simbolo $o\left(|h|^2\right)$ indica, analogamente al caso unidimensionale, una quantità tale che

$$\lim_{h\to 0} \frac{o\left(|h|^2\right)}{|h|^2} = 0,$$

dove il limite va ora inteso in $\mathbb{R}^n$, con le avvertenze che questo comporta trattandosi ovviamente di una forma di indeterminazione (si veda la discussione del paragrafo 2.2).

Esaminiamo i termini a secondo membro della formula di Taylor (5.4): dopo $f(x_0)$, la prima sommatoria è il differenziale di f applicato all'incremento h, mentre la seconda sommatoria è una *forma quadratica* nelle componenti h_i dell'incremento, che rappresenta il differenziale secondo di f in x_0 applicato ad h.

DIMOSTRAZIONE. La nostra tesi consiste nel dimostrare che per $h \to 0$ è

$$\frac{f(x_0 + h) - \left[f(x_0) + \nabla f(x_0)\cdot h + \frac{1}{2}\sum_{i,j=1}^{n} f_{x_i x_j}(x_0)\, h_i h_j\right]}{|h|^2} \to 0. \tag{5.5}$$

Calcoliamo la differenza che compare a numeratore della (5.5) usando la (5.2):

$$\left|\frac{f(x_0 + h) - [\ldots]}{|h|^2}\right| = \left|\frac{\frac{1}{2}\sum_{i,j=1}^{n}\left[f_{x_i x_j}(x_0 + \delta h) - f_{x_i x_j}(x_0)\right] h_i h_j}{|h|^2}\right| \leq$$

essendo $\frac{|h_i h_j|}{|h|^2} \leq 1$,

$$\leq \frac{1}{2}\sum_{i,j=1}^{n} \left|f_{x_i x_j}(x_0 + \delta h) - f_{x_i x_j}(x_0)\right|$$

dove δ dipende da x_0 e h, ma in ogni modo $\delta \in (0,1)$. Pertanto $|\delta h| \leq |h|$ e, per $h \to 0$, si ha $x_0 + \delta h \to x_0$.

D'altro canto per ipotesi $f \in C^2(A)$, perciò le funzioni $f_{x_i x_j}$ sono continue in x_0 e, per $h \to 0$,

$$\left|f_{x_i x_j}(x_0 + \delta h) - f_{x_i x_j}(x_0)\right| \to 0,$$

da cui la tesi. ♦

ESERCIZI

1. Verificare che le seguenti funzioni soddisfano l'equazione di Laplace in due variabili:

$$f_{xx} + f_{yy} = 0.$$

a) $f(x,y) = e^x \sin y$
b) $f(x,y) = \cos x \operatorname{Sh} y$
c) $f(x,y) = x^3 - 3xy^2$
d) $f(x,y) = x^4 - 6x^2y^2 + y^4$

2. Scrivere la matrice hessiana di ciascuna delle funzioni dell'esercizio precedente.

3. Scrivere il differenziale secondo nell'origine per la funzione a) dell'esercizio 37.

4. Scrivere la matrice hessiana della funzione:

$$f(x,y) = \sqrt{1 + x^2 + y^2}$$

precisando l'insieme in cui è definita.

5. Verificare che la funzione

$$u(x,t) = \frac{1}{\sqrt{t}} e^{-\frac{x^2}{4t}}$$

soddisfa l'equazione del calore

$$u_t - u_{xx} = 0$$

per ogni $t > 0, x \in \mathbb{R}$.

6. *Soluzioni radiali dell'equazione di Laplace*
a) Posto $\rho(x) = |x|$ per $x \in \mathbb{R}^n$, calcolare le seguenti espressioni contenenti derivate parziali:

$$\rho_{x_i}; \qquad \sum_{i=1}^{n}(\rho_{x_i})^2; \qquad \rho_{x_i x_i}; \qquad \sum_{i=1}^{n} \rho_{x_i x_i}.$$

b) Sia $g(t)$ una funzione di una variabile e $h(x) = g(|x|)$ una funzione radiale di n variabili. Sfruttando il teorema di derivazione delle funzioni composte e i calcoli del punto a), calcolare $h_{x_i x_i}$ e provare che

$$\Delta h(x) = g''(|x|) + \frac{n-1}{|x|} g'(|x|).$$

 © 978-88-08-06485-1

c) Risolvendo l'equazione differenziale ordinaria

$$g''(\rho) + \frac{n-1}{\rho} g'(\rho) = 0,$$

determinare le soluzioni radiali dell'equazione $\Delta u = 0$ in $\mathbb{R}^n \setminus \{0\}$.

(*Suggerimento*: porre $v(\rho) = g'(\rho)$, ricavare prima v e quindi g, per integrazione; distinguere i casi $n = 2, n > 2$). *Per $n = 2$ e $n = 3$ le soluzioni trovate hanno un interesse fisico: sono il potenziale logaritmico $c_1 \log(x^2 + y^2) + c_2$ e il potenziale newtoniano $c_1 + c_2/\sqrt{x^2 + y^2 + z^2}$.*

6 OTTIMIZZAZIONE. ESTREMI LIBERI

6.1 Generalità sui problemi di ottimizzazione

Il termine *ottimizzazione* si riferisce a un'ampia categoria di problemi che rivestono enorme importanza nelle scienze applicate. In meccanica, minimizzare l'energia potenziale di un sistema non dissipativo conduce a individuarne le posizioni di equilibrio stabile. In economia si sviluppano processi decisionali finalizzati a ottenere il massimo (di utilità o di profitto) da una data quantità di beni o risorse o il minimo dei loro costi di produzione. In numerosi problemi tecnologici intervengono problemi di controllo allo scopo di minimizzare effetti indesiderati come per esempio attriti e/o turbolenze (problemi di ottimizzazione di forma in scafi o profili d'ala) o emissioni di gas tossici (marmitta catalitica).

In generale si tratta di massimizzare o minimizzare una quantità, in gergo *un obiettivo*, sotto certe condizioni la cui modellizzazione matematica dipende dalla natura del problema e di conseguenza presenta vari gradi di complessità.

Nei modelli più semplici l'obiettivo è costituito da una funzione reale f di n variabili reali:

$$f : A \to \mathbb{R}, \text{ con } A \subseteq \mathbb{R}^n.$$

Abbiamo già trattato il caso $n = 1$ nel primo volume. Qui e nel capitolo 4 ci occuperemo del caso $n > 1$. Occorre subito convincersi che la presenza di due o più variabili rende un problema di ottimizzazione molto più complesso rispetto al caso di una variabile. Innanzitutto, conviene distinguere fra i seguenti casi, in cui si vede che le proprietà topologiche di A assumono un ruolo rilevante.

a) Ricerca di estremi assunti in punti *interni al dominio* A di f. Si parla in questo caso di *estremi liberi*. In particolare, si è in questo caso se A è un insieme *aperto*.

b) Ricerca di estremi di f assunti su un sottoinsieme *non aperto* di A o sulla frontiera di A. Si parla allora di *estremi vincolati*. Per esempio, la massimizzazione di una funzione su un cerchio chiuso del piano o sul suo bordo.

Prima di passare a definizioni precise, illustriamo con alcuni esempi la distinzione tra i problemi di tipo a) e b).

Un problema di allocazione ottima

Siano assegnati nel piano xy, tre punti $\mathrm{p}_1 = (x_1, y_1)$, $\mathrm{p}_2 = (x_2, y_2)$, $\mathrm{p}_3 = (x_3, y_3)$, che pensiamo disposti ai vertici di un triangolo acutangolo come in figura 3.15.

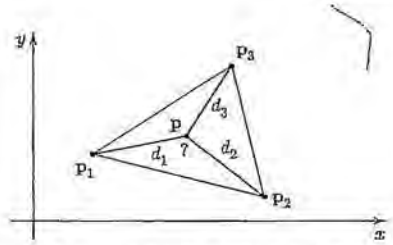


Figura 3.15. Dove deve essere collocato p in modo da minimizzare $d_1 + d_2 + d_3$?

Si vogliono collegare questi punti con un quarto punto $\mathrm{p} = (x, y)$ mediante un cavo, in modo da utilizzare meno materiale possibile. Dove deve essere collocato il punto p?

Posto $d_i(x, y) = |\mathrm{p} - \mathrm{p}_i| = \sqrt{(x_i - x)^2 + (y_i - y)^2}$, per $i = 1, 2, 3$, si tratta di minimizzare la funzione $f : \mathbb{R}^2 \to \mathbb{R}$ definita da

$$f(x, y) = \sum_{i=1}^{n} d_i(x, y)$$

corrispondente alla somma delle distanze di p dai tre punti assegnati.

In questo caso $A = \mathbb{R}^2$ e quindi si tratta della ricerca di estremi liberi.

Un problema isoperimetrico

Fra tutti i parallelepipedi aventi superficie totale assegnata $2S$, determinare quello di volume massimo.

Se indichiamo con x, y, z le lunghezze (positive) degli spigoli del parallelepipedo, il problema equivale a massimizzare la funzione

$$V(x, y, z) = xyz$$

sull'ottante positivo di $\mathbb{R}^3$ definito da

$$\Omega = \{(x, y, z) \in \mathbb{R}^3 : x > 0, y > 0, z > 0\}$$

sotto la condizione "isoperimetrica"

$$g(x, y, z) \equiv xy + yz + xz = S. \tag{6.1}$$

Si tratta di un problema di massimizzazione vincolata: massimizzare la funzione $V(x, y, z)$ in Ω, sotto il *vincolo di uguaglianza* (6.1).

Problema della dieta

Supponiamo di avere a disposizione n tipi di cibo $A_1, \ldots, A_n$, ciascuno dei quali contenga m sostanze nutritive $N_1, \ldots, N_m$. Siano, per $i = 1, \ldots, n$, $j = 1, \ldots, m$:

a_{ij} = contenuto calorico della sostanza nutritiva N_j per unità di peso dell'alimento A_i;

c_i = costo dell'alimento A_i per unità di peso;

b_j = fabbisogno stimato giornaliero nell'organismo della sostanza N_j (espresso in peso).

Il *problema della dieta* consiste nello scegliere la quantità x_i di cibo A_i, $i = 1, \dots, n$ in modo da minimizzare il costo totale, sotto la condizione che la dieta sia sufficiente al fabbisogno di ogni sostanza.

Si tratta dunque di minimizzare la funzione

$$C(x_1, \dots, x_n) = \sum_{i=1}^{n} c_i x_i$$

sotto le condizioni

$$\begin{cases} x_i \geq 0 & \text{per } i = 1, \dots, n \\ \sum_{i=1}^{n} x_i a_{ij} \geq b_j & \text{per } j = 1, \dots, m. \end{cases}$$

Si tratta di un tipico problema di *programmazione*, ovvero di ottimizzazione in presenza di vincoli di *disuguaglianza*. In questo caso, data la linearità della funzione obiettivo e dei vincoli, il problema è più precisamente di *programmazione lineare*.

Formuliamo ora con precisione i concetti di estremo illustrati negli esempi precedenti. In particolare è necessario distinguere bene *estremi globali e locali* ed evitare confusioni tra massimi e/o minimi e *punti di* massimo e/o minimo.

DEFINIZIONE 3.19 Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$ e $x_0 \in A$. Diciamo che:
(*a*) x_0 è *punto di massimo (risp. minimo) assoluto* per f in A e che $f(x_0)$ è *il massimo (risp. minimo) assoluto o globale* di f in A se

$$\text{per ogni } x \in A \text{ si ha } f(x) \leq f(x_0) \quad (\text{risp. } f(x) \geq f(x_0)). \tag{6.2}$$

(*b*) Diciamo che x_0 è *punto di massimo (minimo) relativo o locale* per f e che $f(x_0)$ è *un massimo (risp. minimo) relativo o locale* di f se *esiste* un intorno U di x_0 tale che

$$\text{per ogni } x \in U \text{ si ha } f(x) \leq f(x_0) \quad (\text{risp. } f(x) \geq f(x_0)). \tag{6.3}$$

Se nelle (6.2) e (6.3) per $x \neq x_0$ le disuguaglianze sono strette ($<$ o $>$), i punti di estremo e gli estremi si dicono *forti*, altrimenti si dicono *deboli*. ▩

Le domande spontanee sono: ci sarà davvero un punto x_0 che minimizza f? Se c'è, ce ne potrebbe essere più d'uno? Come si fa a determinarli?

Dobbiamo dunque affrontare le seguenti questioni:

1. *Esistenza.* Ovviamente non si può garantire che esistano sempre punti in cui una funzione assume massimo o minimo, sia locali che (a maggior ragione) globali. In molti casi, il Teorema di Weierstrass[17] si rivela utile per provare l'esistenza di estremi globali: *se A è chiuso e limitato ed f è continua in A, allora f possiede sia massimo sia minimo in A.* Nell'esempio del problema di allocazione ottima non ha neppure senso cercare massimi globali, in quanto la funzione f diventa grande quanto si vuole se x_0 si allontana dai vertici del triangolo. D'altra parte, proprio questo fatto ci dice che, se c'è, il punto di minimo starà all'interno di un cerchio chiuso C di raggio abbastanza grande. Ma allora ci si può ridurre a cercare il minimo su C, che è chiuso

[17] v. paragrafo 3.2.

e limitato. Essendo f continua, il teorema di Weierstrass assicura l'esistenza di un punto x_0 nel quale f realizza il suo minimo globale[18]; ovvero, il nostro problema *ha soluzione!*

Confortati da questa buona notizia, si può procedere con la ricerca di x_0.

2. *Unicità.* È chiaro che se $m = f(x_0)$ è minimo globale di f in A, questo è unico. Non così per il punto in cui f assume il valore m. Già nel caso unidimensionale, si vede che $\sin x$ assume il suo minimo globale ($= -1$) in infiniti punti. Analogo discorso vale per i massimi.

Naturalmente, poter dimostrare che esiste un unico punto di minimo o di massimo può rivelarsi molto utile, per esempio nel calcolo approssimato con metodi numerici. C'è in effetti un caso in cui il punto di minimo (massimo), se esiste, è unico: si tratta, come vedremo nel paragrafo 7.2, del caso delle funzioni *strettamente convesse (o concave).*

Nel problema di allocazione ottima la funzione f è strettamente convessa (come vedremo in seguito) e perciò il punto di minimo *esiste ed è unico!*

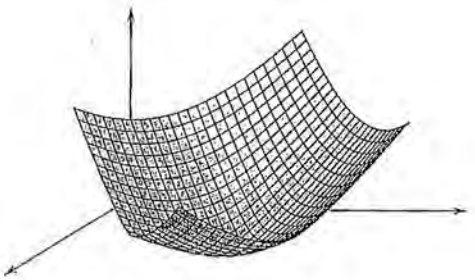


Figura 3.16. La funzione obiettivo nel problema di allocazione ottima.

3. *Ricerca e caratterizzazione dei punti di estremo.* Come già nel caso unidimensionale, la ricerca dei punti di massimo/minimo è basata su condizioni analitiche necessarie e/o sufficienti che tali punti devono soddisfare.

Nella ricerca di estremi liberi, lo strumento principale d'indagine è il calcolo differenziale sviluppato nelle sezioni precedenti. È così possibile estendere al caso multidimensionale concetti come quello di punto stazionario e classificarne la natura.

La ricerca di estremi vincolati richiede invece l'uso del calcolo differenziale per funzioni a valori vettoriali e quindi sarà trattata più avanti, nel capitolo 4. In particolare, svilupperemo il *metodo dei moltiplicatori di Lagrange* per il caso di vincoli di uguaglianza e daremo un cenno a quello dei *moltiplicatori di Kuhn-Tucker* per il caso dei vincoli di disuguaglianza.

4. *Algoritmi di calcolo.* Tranne che in casi particolarmente semplici, l'uso dei metodi indicati al punto 3 porta a sistemi di equazioni non lineari che difficilmente si possono risolvere senza l'uso di un computer. Diventa quindi essenziale poter disporre di algoritmi che permettano il calcolo approssimato della soluzione, con il gra-

[18] Si veda in proposito l'Esercizio 19: abbiamo applicato un argomento di quel tipo.

 © 978-88-08-06485-1

do di approssimazione desiderato. Qui entriamo propriamente nel campo dell'Analisi Numerica, per cui rimandiamo ai testi specializzati sull'argomento.

6.2 Estremi liberi. Condizioni necessarie del prim'ordine

Iniziamo a discutere il problema della ricerca dei massimi e minimi *liberi*, locali e/o globali per funzioni sufficientemente regolari. I punti x che si candidano a essere di massimo o minimo sono interni al dominio A di f e, quindi, ci si può muovere attorno a essi in tutte le direzioni possibili. Per questo motivo, possiamo supporre direttamente che A sia *aperto*.

Sia dunque $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, con A aperto. Procedendo come nel caso unidimensionale, cerchiamo di selezionare i candidati punti di estremo locale, individuando una condizione necessaria che essi devono soddisfare. Vale il seguente teorema.

TEOREMA 3.17 (DI FERMAT) *Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}$, con A aperto e $x_0 \in A$ un punto di massimo o minimo locale per f. Se f è derivabile in x_0, allora $\nabla f(x_0) = 0$.*

DIMOSTRAZIONE. Supponiamo, per fissare le idee, che x_0 sia punto di minimo locale. Occorre dimostrare che ogni derivata parziale di f si annulla in x_0. Infatti, sia $e_1, \dots, e_n$ la base canonica in $\mathbb{R}^n$. Muoviamoci da x_0 lungo l'asse x_j considerando la funzione della variabile reale t data da:

$$g(t) = f(x_0 + te_j).$$

La funzione g è ben definita in un intorno di $t = 0$ e, poiché x_0 è un punto di minimo locale per f, deduciamo che $t = 0$ è punto di minimo locale per g. D'altra parte, g è derivabile in $t = 0$ e infatti, per definizione di derivata parziale, si ha

$$\frac{\partial f}{\partial x_j}(x_0) = g'(0).$$

Per il teorema di Fermat in una variabile[19], deve essere allora $g'(0) = 0$. In conclusione:

$$\frac{\partial f}{\partial x_j}(x_0) = 0, \quad \text{per ogni } j = 1, \dots, n. \qquad \blacklozenge$$

I punti in cui il gradiente di una funzione f si annulla si dicono *punti critici* o *stazionari* di f. Il teorema precedente afferma quindi che, per cercare i punti di massimo o minimo locale nei quali una funzione è derivabile, occorre anzitutto determinare tutti i suoi punti critici, ossia i punti $x = (x_1, x_2, \dots, x_n)$ che risolvono il sistema di n equazioni in n incognite:

$$\begin{cases} f_{x_1}(x_1, x_2, \dots, x_n) = 0 \\ f_{x_2}(x_1, x_2, \dots, x_n) = 0 \\ \dots \\ f_{x_n}(x_1, x_2, \dots, x_n) = 0. \end{cases} \tag{6.4}$$

Naturalmente, come già nel caso di una variabile, non è detto che ogni punto critico sia effettivamente un punto di massimo o minimo. Il secondo passo sarà quindi, una

[19] v. volume 1, capitolo 4, teorema 4.5.

© 978-88-08-06485-1

volta determinati i punti stazionari, studiarne la natura, ossia stabilire per ciascuno di essi se è punto di minimo, di massimo o nessuna delle due cose. In quest'ultimo caso, un punto stazionario si dice *punto di sella* o di *colle*. Va osservato che, per esempio nei testi di programmazione, il termine punto di *sella* è riservato a punti critici nei quali f è minima lungo alcune direzioni e massima lungo altre. La figura 3.17d mostra un tipico punto di sella.

OSSERVAZIONE Vi possono essere anche punti di estremo in cui f non è derivabile. Per esempio la funzione

$$f(x, y) = \sqrt{x^2 + y^2},$$

che ha come grafico un cono con vertice nell'origine, ha ivi un minimo globale. Ma le derivate parziali di f non esistono nell'origine. Questi punti vanno esaminati a parte.

A volte uno studio diretto può bastare per studiare la natura dei punti stazionari, come per l'origine nell'esempio seguente.

Esempio

Cerchiamo i punti critici di

$$f(x, y) = 3x^2 + y^2 - x^3 y.$$

Occorre risolvere il sistema:

$$\begin{cases} \dfrac{\partial f}{\partial x} = 6x - 3x^2 y = 0 \\ \dfrac{\partial f}{\partial y} = 2y - x^3 = 0. \end{cases}$$

La prima equazione dà $x = 0$ oppure $y = 2/x$. Sostituendo queste nella seconda si trovano le soluzioni:

$$(0, 0) \qquad (\sqrt{2}, \sqrt{2}) \qquad (-\sqrt{2}, -\sqrt{2}).$$

Questi sono tutti e soli i punti critici di f. Poiché f è differenziabile in tutti i punti del piano, per il teorema di Fermat non possono esservi altri punti di massimo e minimo.

Si tratta ora di decidere, per ciascuno di questi tre punti, se sia effettivamente un punto di estremo oppure no. Per l'origine, è facile convincersi da un esame diretto che si tratta di un punto di minimo relativo; infatti $f(0, 0) = 0$ e in un intorno dell'origine f è non negativa, poiché il termine $-x^3 y$ (di quarto grado) è trascurabile rispetto a $3x^2 + y^2$, che è positivo. (Quest'affermazione sarà giustificata più rigorosamente nel paragrafo 6.5).

Vedremo nei prossimi paragrafi un *criterio generale per lo studio dei punti critici*, che ci consentirà di studiare gli altri due punti critici di f.

6.3 Forme quadratiche. Classificazione

In assenza di informazioni aggiuntive, un modo per determinare la natura di un punto critico è quello di utilizzare la formula di Taylor per analizzare il segno dell'incremento $\Delta f(x_0) = f(x_0 + h) - f(x_0)$. Se infatti si riesce a stabilire che $\Delta f(x_0)$ si mantiene di segno positivo (negativo) per ogni h di modulo abbastanza piccolo, possiamo dedurre che x_0 è punto di minimo (massimo) locale. Se invece, al variare di h, $\Delta f(x_0)$ cambia segno, siamo in presenza di un colle.

 © 978-88-08-06485-1

Rivediamo brevemente il caso unidimensionale. Sia f due volte derivabile e $f'(x_0) = 0$. Arrestiamo la formula di Taylor al second'ordine. Si trova:

$$\Delta f(x_0) = f(x_0 + h) - f(x_0) = \frac{1}{2} f''(x_0) h^2 + o(h^2) \quad \text{per } h \to 0.$$

Se ora $f''(x_0) > 0$ o < 0, il segno di $\Delta f(x_0)$ sarà > 0 o < 0, definitivamente per $h \to 0$ e possiamo concludere che x_0 è punto di minimo o massimo, rispettivamente. Se invece $f''(x_0) = 0$, occorre un'analisi più profonda.

Nel caso multidimensionale, sia $f \in C^2(A)$ e $\nabla f(\mathrm{x}_0) = 0$. Procedendo allo stesso modo, si trova (applicando ora il teorema 3.16, paragrafo 5.2):

$$\Delta f(\mathrm{x}_0) = f(\mathrm{x}_0 + \mathrm{h}) - f(\mathrm{x}_0) = \frac{1}{2} \sum_{i,j=1}^{n} f_{x_i x_j}(\mathrm{x}_0) h_i h_j + o(|\mathrm{h}|^2). \tag{6.5}$$

La (6.5) indica che la determinazione del segno dell'incremento $\Delta f(\mathrm{x}_0)$ conduce ancora all'analisi del differenziale secondo di f in x_0, ossia all'analisi del polinomio omogeneo di secondo grado nelle componenti di h (che si chiama *forma quadratica*) dato da

$$\sum_{i,j=1}^{n} f_{x_i x_j}(\mathrm{x}_0) h_i h_j.$$

L'analisi della natura dei punti stazionari conduce quindi a un problema che, come vedremo, è di tipo puramente algebrico: lo *studio del segno di una forma quadratica*. Apriamo dunque una parentesi su questo problema.

Consideriamo una *forma quadratica* (in breve, f.q.) su $\mathbb{R}^n$, definita da:

$$q(\mathrm{h}) = q(h_1, h_2, \ldots, h_n) = \sum_{i,j=1}^{n} a_{ij} q_i q_j \tag{6.6}$$

dove gli a_{ij} sono numeri reali, detti coefficienti della f.q.. Se tutti gli a_{ij} sono nulli, la (6.6) definisce la f.q. *nulla*.

Si può sempre supporre che $a_{ij} = a_{ji}$. Se così non fosse, basterebbe sostituire ai due coefficienti a_{ij} e a_{ji} la loro semisomma $(a_{ij} + a_{ji})/2$, lasciando così inalterata la (6.6). Ogni f.q. risulta così associata biunivocamente a una matrice simmetrica $\mathrm{M} = \{a_{ij}\}_{i,j=1,\ldots,n}$. Nel caso del differenziale secondo di una funzione f in un punto x, la matrice M coincide con la matrice hessiana di f in x.

Esempio

6.2 Le forme quadratiche seguenti corrispondono alle matrici indicate:

$$q(h_1, h_2) = h_1^2 + 2h_2^2 - 6h_1h_2 \qquad \mathrm{M} = \begin{pmatrix} 1 & -3 \\ -3 & 2 \end{pmatrix};$$

$$q(h_1, h_2, h_3) = h_1^2 + 2h_2^2 + h_3^2 - 6h_1h_2 + 3h_1h_3 \qquad \mathrm{M} = \begin{pmatrix} 1 & -3 & 3/2 \\ -3 & 2 & 0 \\ 3/2 & 0 & 1 \end{pmatrix};$$

$$q(h_1, h_2) = f_{xx}(x,y)h_1^2 + 2f_{xy}(x,y)h_1h_2 + f_{yy}(x,y)h_2^2 \qquad \mathrm{M} = \mathrm{H}_f(x,y).$$

© 978-88-08-06485-1

A volte è più comodo o significativo scrivere una f.q. come prodotto scalare

$$q(\mathrm{h}) = \langle \mathrm{Mh}, \mathrm{h} \rangle$$

oppure matriciale:

$$q(\mathrm{h}) = \mathrm{h}^{\top} \mathrm{Mh}$$

intendendo che h sia un vettore colonna e $\mathrm{h}^{\top}$ il suo trasposto (vettore riga).

Siamo interessati al *segno* che una f.q. assume al variare di h. Una prima osservazione è che, essendo q un polinomio omogeneo di secondo grado, si ha $q(t\mathrm{h}) = t^2 q(\mathrm{h})$ per ogni $t \in \mathbb{R}$, perciò:

$q(\mathrm{h})$ assume segno costante su ogni retta passante per l'origine (origine esclusa; infatti si ha sempre $q(0) = 0$).

I seguenti semplici esempi in $\mathbb{R}^2$ illustrano tutte le possibilità di comportamento:

a) $h_1^2 + h_2^2$ è positiva per ogni $(h_1, h_2) \neq (0,0)$;

b) $-h_1^2 + h_2^2$ è positiva per $(h_1, h_2) = (0,1)$ e negativa per $(h_1, h_2) = (1,0)$;

c) $-h_1^2 - h_2^2$ è negativa per ogni $(h_1, h_2) \neq (0,0)$;

d) h_1^2 è sempre positiva tranne che nei vettori del tipo $(0, h_2)$;

e) $-h_1^2$ è sempre negativa tranne che nei vettori del tipo $(0, h_2)$.

Tali esempi suggeriscono la seguente classificazione che si applica sia alla f.q. che alla matrice simmetrica ad essa associata.

Definizione 3.20 Una forma quadratica $q(\mathrm{h})$, $\mathrm{h} \in \mathbb{R}^n$ (o la matrice simmetrica corrispondente) si dice:

i) definita positiva (negativa) se, per ogni $\mathrm{h} \neq 0$, $q(\mathrm{h}) > 0\ (< 0)$;

ii) semidefinita positiva (negativa) se, per ogni $\mathrm{h} \neq 0$, $q(\mathrm{h}) \geq 0\ (\leq 0)$ ed esiste $\mathrm{h} \neq 0$ tale che $q(\mathrm{h}) = 0$;

iii) indefinita, se esistono $\mathrm{h}_1, \mathrm{h}_2$ tali che $q(\mathrm{h}_1) > 0$ e $q(\mathrm{h}_2) < 0$. ▩

Si noti che le condizioni *i)*, *ii)*, *iii)* sono mutuamente esclusive. Le figure 3.17 illustrano i grafici di alcune tipiche forme quadratiche, appartenenti nell'ordine alle tre categorie *i)*, *ii)*, *iii)*. Si noti la linea di punti di minimo nel caso *c)*.

Osservazione È sempre bene specificare lo spazio $\mathbb{R}^n$ in cui si opera quando si vuole classificare una forma quadratica. Per esempio, la f.q. in *a)* è definita positiva in $\mathbb{R}^2$, ma è semidefinita positiva in $\mathbb{R}^3$. Per le altre, considerate come f.q. su $\mathbb{R}^2$, si ha: *b)* indefinita; *c)* definita negativa; *d)* semidefinita positiva; *e)* semidefinita negativa.

Occorrono ora dei criteri per classificare una f.q. senza dover ricorrere alla definizione. Lasciamoci guidare dal caso bidimensionale. Sia dunque:

$$q(h_1, h_2) = ah_1^2 + 2bh_1h_2 + ch_2^2 \tag{6.7}$$

con a, b, c *non tutti nulli*. La matrice associata è

$$\mathrm{M} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

 © 978-88-08-06485-1

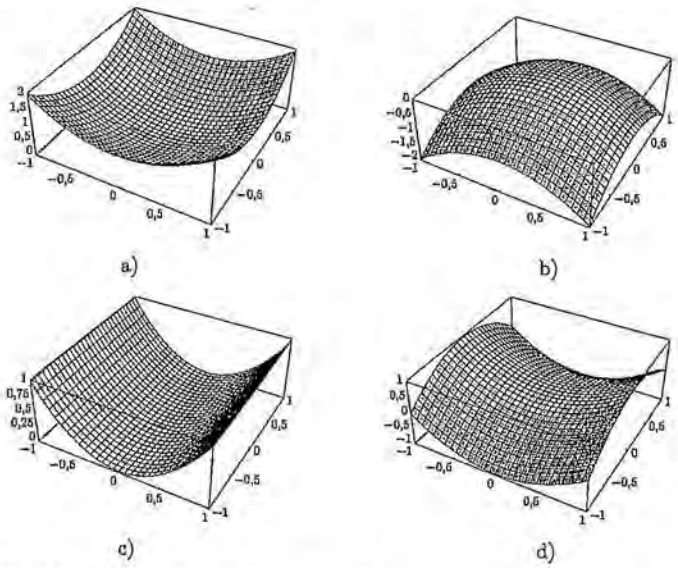


Figura 3.17. a) Forma quadratica definita positiva. b) Forma quadratica definita negativa. c) Forma quadratica semidefinita positiva. d) Forma quadratica indefinita.

Se $a = c = 0$, q è certamente indefinita: basta osservare che $q(1,1) = 2b$ e $q(-1,1) = -2b$. Se $a \neq 0$ (nel caso $a = 0, c \neq 0$ si procede analogamente), q si può riscrivere al modo seguente:

$$q(h_1, h_2) = a\left(h_1 + \frac{b}{a}h_2\right)^2 + \frac{ac - b^2}{a}h_2^2$$

I coefficienti dei quadrati sono a e $\frac{\det \mathrm{M}}{a}$; ragionando sul segno di questi due coefficienti, si deduce pertanto il seguente risultato:

TEOREMA 3.18 (SEGNO DELLE F.Q. IN DUE VARIABILI) *Se $a \neq 0$, la forma quadratica (6.7) è:*

i) definita positiva (negativa) se e solo se $\det \mathrm{M} > 0$ *e* $a > 0$ $(a < 0)$;

ii) indefinita se e solo se $\det \mathrm{M} < 0$;

iii) semidefinita positiva (negativa) se e solo se $\det \mathrm{M} = 0$ *e* $a > 0$ $(a < 0)$.

Se $a = 0$ e $c \neq 0$, nelle affermazioni precedenti occorre sostituire a con c.

La forma quadratica nulla si può considerare semidefinita sia positiva che negativa. Il punto importante nel teorema precedente è che il test fa intervenire non solo la matrice M, ma anche una matrice 1×1, per esempio la matrice $\mathrm{M}_1 = (a)$. La naturale generalizzazione del teorema 3.18 alle forme quadratiche di n variabili fa

© 978-88-08-06485-1

allora intervenire tutte le n sottomatrici M_k composte mediante le prime k righe e k colonne di M, chiamate talvolta *sottomatrici principali di nord-ovest*:

$$\mathrm{M}_1 = (a_{11}), \quad \mathrm{M}_2 = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \quad \mathrm{M}_3 = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}, \ldots, \quad \mathrm{M}_n = \mathrm{M}$$

I determinanti di queste matrici si chiamano *minori principali di nord-ovest*. Vale il seguente teorema, che ci limitiamo a enunciare:

TEOREMA 3.19 (SEGNO DELLE F.Q. IN n VARIABILI) *Sia* $q(\mathrm{h}) = \mathrm{h}^\top \mathrm{M} \mathrm{h}$, $\mathrm{h} \in \mathbb{R}^n$. *Allora:*

i) q è definita positiva se e solo se $\det \mathrm{M}_k > 0$ *per ogni* $k = 1, 2, \ldots, n$;

ii) q è definita negativa se e solo se $(-1)^k \det \mathrm{M}_k > 0$ *per ogni* $k = 1, 2, \ldots, n$.

(*ovvero se* $a_{11} < 0$, $\det \mathrm{M}_2 > 0$, $\det \mathrm{M}_3 < 0$, ...)

Le *sottomatrici principali di nord-ovest* non sono le uniche che si possono usare. Il teorema 3.19 vale anche se si usano le *sottomatrici principali di sud-est*, costruite in modo analogo a partire dall'elemento a_{nn}. Esiste anche una caratterizzazione delle forme quadratiche semidefinite ma, poiché è di difficile uso, ne tralasciamo la presentazione.

ESEMPI

Sia

$$q(h_1, h_2) = -2h_1^2 + 2h_1h_2 - 3h_2^2.$$

La matrice corrispondente è

$$\mathrm{M} = \begin{pmatrix} -2 & 1 \\ 1 & -3 \end{pmatrix}.$$

Poiché $a_{11} = -2 < 0$ e $\det \mathrm{M} = 5 > 0$, la f.q. è definita negativa.

Sia

$$q(h_1, h_2, h_3) = 5h_1^2 - 8h_1h_3 + 3h_2^2 + 4h_3^2.$$

La matrice corrispondente è

$$\mathrm{M} = \begin{pmatrix} 5 & 0 & -4 \\ 0 & 3 & 0 \\ -4 & 0 & 4 \end{pmatrix}.$$

Essendo:

$$a_{11} = 5 > 0, \quad \det \mathrm{A}_2 = \det \begin{pmatrix} 5 & 0 \\ 0 & 3 \end{pmatrix} = 15 > 0, \quad \det \mathrm{A} = 12 > 0,$$

la f.q. è definita positiva.

 © 978-88-08-06485-1

6.4 Forme quadratiche. Test degli autovalori

Un importante test per determinare il segno di una forma quadratica in $\mathbb{R}^n$ è basato sul segno degli autovalori della matrice M. Ricordiamo che un numero complesso λ e un vettore *non nullo* $\mathrm{v} \in \mathbb{C}^n$ si dicono, rispettivamente, *autovalore ed autovettore* (corrispondente a λ) di una matrice quadrata M di ordine n, se soddisfano la relazione:

$$\mathrm{Mv} = \lambda \mathrm{v}$$

o equivalentemente

$$(\mathrm{M} - \lambda \mathrm{I}_n)\mathrm{v} = 0 \tag{6.8}$$

dove I_n indica la matrice identità in $\mathbb{R}^n$.

La (6.8) ha soluzioni v non nulle se e solo se la matrice dei coefficienti è singolare, ovvero se e solo se λ è soluzione dell'equazione, detta *equazione caratteristica:*

$$\det(\mathrm{M} - \lambda \mathrm{I}_n) = 0. \tag{6.9}$$

Il primo membro della (6.9) è un polinomio di grado n in λ e pertanto, in base al teorema fondamentale dell'algebra, *esistono esattamente n autovalori di* M *(in* $\mathbb{C}$*) ciascuno contato secondo la propria molteplicità.*

Le matrici M *simmetriche* posseggono importanti proprietà, che elenchiamo di seguito:

(P_1) *Gli autovalori di* M *sono numeri reali e possiedono autovettori reali.*

(P_2) *Esistono n autovettori linearmente indipendenti* $\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_n$ *che costituiscono una base ortonormale in* $\mathbb{R}^n$, *ovvero*[20] $\mathrm{w}_i \cdot \mathrm{w}_j = \delta_{ij}$.

(P_3) La matrice $\mathrm{S} = (\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_n)$, *le cui colonne sono gli autovettori di cui al punto* (P_2), *è ortogonale (cioè* $\mathrm{S}^\top = \mathrm{S}^{-1}$*) e diagonalizza* M; *precisamente:*

$$\mathrm{S}^\top \mathrm{M} \mathrm{S} = \Lambda \equiv \begin{pmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{pmatrix} \tag{6.10}$$

dove λ_j *è l'autovalore corrispondente a* w_j.

Sia ora $q(\mathrm{h}) = \mathrm{h}^\top \mathrm{M} \mathrm{h}$. L'idea è di sfruttare le proprietà (P_1), (P_2) e (P_3) per scrivere q come somma di termini al quadrato moltiplicati per gli autovalori. Risalire da qui al segno di q diventa allora molto semplice.

Dalla (6.10) si ricava

$$\mathrm{M} = \mathrm{S} \Lambda \mathrm{S}^\top,$$

per cui si può scrivere:

$$q(\mathrm{h}) = \mathrm{h}^\top \mathrm{M} \mathrm{h} = \mathrm{h}^\top \left(\mathrm{S} \Lambda \mathrm{S}^\top\right) \mathrm{h} = \left(\mathrm{S}^\top \mathrm{h}\right)^\top \Lambda (\mathrm{S}^\top \mathrm{h}). \tag{6.11}$$

[20] $\delta_{ij} = 1$ se $i \neq j$ e $\delta_{ii} = 0$ (simbolo di Kronecker).

© 978-88-08-06485-1

Operiamo ora la trasformazione $\mathrm{k} = \mathrm{S}^\top \mathrm{h}$, che è biunivoca in $\mathbb{R}^n$. Sostituendo nella (6.11), si trova:

$$q(\mathrm{h}) = q(\mathrm{Sk}) = \mathrm{k}^\top \Lambda \mathrm{k} = \sum_{i=1}^{n} \lambda_i k_i^2. \tag{6.12}$$

Il procedimento che porta a scrivere q nella forma (6.12) si chiama *riduzione a forma canonica di* q. In altri termini, rispetto alla base ortonormale $\mathrm{w}_1, \mathrm{w}_2, \ldots, \mathrm{w}_n$, la forma quadratica assume la forma (6.12).

Questo ci consente di provare il seguente risultato:

TEOREMA 3.20 *La forma quadratica* $q(\mathrm{h}) = \mathrm{h}^T \mathrm{M} \mathrm{h}$ *è:*

- *definita positiva (negativa) se e solo se tutti gli autovalori di* M *sono positivi (negativi);*
- *semidefinita positiva (negativa) se e solo se tutti gli autovalori di* M *sono* ≥ 0 (≤ 0) *e almeno uno di essi è nullo;*
- *indefinita se* M *ha almeno un autovalore positivo e uno negativo.*

La classificazione delle forme quadratiche si può basare quindi sullo studio del segno degli autovalori.

DIMOSTRAZIONE. Cominciamo a considerare la f.q.

$$\widetilde{q}(\mathrm{k}) = q(\mathrm{Sk}).$$

Dall'uguaglianza (6.12)

$$\widetilde{q}(\mathrm{k}) = \sum_{i=1}^{n} \lambda_i k_i^2,$$

leggiamo subito che per $\widetilde{q}$ la tesi del teorema è vera. D'altro canto il fatto che la trasformazione $\mathrm{k} \mapsto \mathrm{Sk}$ sia biunivoca e in particolare che

$$\mathrm{Sk} = 0 \Leftrightarrow \mathrm{k} = 0$$

mostra che le f.q. q e $\widetilde{q}$ hanno lo stesso segno (cioè q è definita positiva se e solo se $\widetilde{q}$ è definita positiva e così via per gli altri casi). Questo conclude la dimostrazione. ◆

OSSERVAZIONE Nel caso bidimensionale si può riottenere per questa via il risultato del teorema 3.18: consideriamo ancora la forma quadratica (6.7) con matrice associata M.

$$\det(\mathrm{M} - \lambda \mathrm{I}_2) = \begin{vmatrix} a - \lambda & b \\ b & c - \lambda \end{vmatrix} = \lambda^2 - \lambda(a + c) + (ac - b^2) =$$
$$= \lambda^2 - \lambda \,\mathrm{Tr}\mathrm{M} + \det \mathrm{M} = 0.$$

Detti λ_1, λ_2 gli autovalori di M, si ha dunque: $\lambda_1 \cdot \lambda_2 = \det \mathrm{M}$ e $\lambda_1 + \lambda_2 = \mathrm{Tr}\mathrm{M}$. Interpretando allora quanto afferma il teorema 3.20 in termini di segno di $\det \mathrm{M}$ e $\mathrm{Tr}\mathrm{M}$, si riottiene il teorema 3.18.

Concludiamo evidenziando la seguente proprietà delle f.q. *definite*, che ci sarà utile nel prossimo paragrafo.

 © 978-88-08-06485-1

TEOREMA 3.21 *Sia $q(\mathrm{h}) = \mathrm{h}^T \mathrm{M} \mathrm{h}$ una f.q. in $\mathbb{R}^n$. Se q è definita positiva, allora*

$$q(\mathrm{h}) \geq \lambda_{min} |\mathrm{h}|^2 \quad \textit{per ogni } \mathrm{h} \in \mathbb{R}^n$$

dove λ_{min} è il minimo autovalore di M. Analogamente, se q è definita negativa, allora

$$q(\mathrm{h}) \leq \lambda_{max} |\mathrm{h}|^2 \quad \textit{per ogni } \mathrm{h} \in \mathbb{R}^n$$

dove λ_{max} (che è negativo) è il massimo autovalore di M.

DIMOSTRAZIONE. Dalla (6.12), abbiamo:

$$q(\mathrm{h}) = q(\mathrm{Sk}) = \mathrm{k}^T \Lambda \mathrm{k} = \sum_{i=1}^{n} \lambda_i k_i^2$$

e perciò, se q è definita positiva, si ha

$$q(\mathrm{h}) \geq \lambda_{\min} \sum_{i=1}^{n} k_i^2 = \lambda_{\min} |\mathrm{k}|^2 ,$$

mentre, se q è definita negativa, si ha

$$q(\mathrm{h}) = \sum_{i=1}^{n} \lambda_i k_i^2 \leq \lambda_{\max} \sum_{i=1}^{n} k_i^2 = \lambda_{\max} |\mathrm{k}|^2 .$$

Per concludere, basta osservare che $|\mathrm{k}|^2 = |\mathrm{h}|^2$. Infatti, per l'ortogonalità di S abbiamo $\mathrm{S}\mathrm{S}^T = \mathrm{I}_n$, mentre la definizione di matrice trasposta equivale a:

$$\left\langle \mathrm{S}^T \mathrm{x}, \mathrm{y} \right\rangle = \langle \mathrm{x}, \mathrm{S}\mathrm{y} \rangle \quad \text{per ogni } \mathrm{x}, \mathrm{y} \in \mathbb{R}^n .$$

Possiamo allora scrivere:

$$|\mathrm{k}|^2 = \langle \mathrm{k}, \mathrm{k} \rangle = \left\langle \mathrm{S}^T \mathrm{S}\mathrm{k}, \mathrm{k} \right\rangle = \langle \mathrm{S}\mathrm{k}, \mathrm{S}\mathrm{k} \rangle = |\mathrm{h}|^2 .$$

(Del resto, la trasformazione $\mathrm{k} \mapsto \mathrm{S}\mathrm{k}$ rappresenta geometricamente una rotazione nello spazio $\mathbb{R}^n$, perciò le lunghezze dei vettori si conservano). ◆

OSSERVAZIONE Interpretiamo i risultati raggiunti in termini di ottimizzazione di una forma quadratica. Se $q = q(\mathrm{h})$ è una f.q. in $\mathbb{R}^n$, si ha sempre $q(0) = 0$ e $\nabla q(0) = 0$.

Quindi l'origine è sempre punto critico per q e la sua natura dipende dal segno di q (v. figure 3.17)

- Se q è definita positiva (negativa), $\mathrm{h} = 0$ è punto di *minimo (massimo) globale forte.*
- Se q è semidefinita positiva (negativa), $\mathrm{h} = 0$ è punto di *minimo (massimo) globale debole.*
- Se q è indefinita, $\mathrm{h} = 0$ è punto di *sella.*

© 978-88-08-06485-1

6.5 Studio della natura dei punti critici

Ci occupiamo ora di fornire strumenti utili a determinare la natura di un punto stazionario x_0 per una funzione $f \in C^2(A)$. Come già visto, essendo $\nabla f(\mathrm{x}_0) = 0$, la formula di Taylor al secondo ordine ha la forma seguente:

$$\Delta f(\mathrm{x}_0) \equiv f(\mathrm{x}_0 + \mathrm{h}) - f(\mathrm{x}_0) = \frac{1}{2} \sum_{i,j=1}^{n} f_{x_i x_j}(\mathrm{x}_0) h_i h_j + o(|\mathrm{h}|^2) \tag{6.13}$$

per $\mathrm{h} \to \mathbf{0}$. Per estrarre informazioni sulla natura del punto critico x_0 occorre studiare e classificare la forma quadratica

$$q(\mathrm{h}) = \sum_{i,j=1}^{n} f_{x_i x_j}(\mathrm{x}_0) h_i h_j = \mathrm{h}^T \mathrm{H}_f(\mathrm{x}_0)\, \mathrm{h} \tag{6.14}$$

dove $\mathrm{H}_f(\mathrm{x}_0)$ è la matrice hessiana di f in x_0. Nel caso in cui q sia definita oppure indefinita, la conclusione è indicata nel seguente teorema.

TEOREMA 3.22 *Siano $f \in C^2(A)$ e $\mathrm{x}_0 \in A$ un punto critico per f. Se la forma quadratica* (6.14) *è:*

(i) definita positiva (negativa), allora x_0 è punto di minimo (massimo) locale forte.

(ii) indefinita, allora x_0 è punto di sella.

DIMOSTRAZIONE. (*i*) Sia $q(\mathrm{h}) = \mathrm{h}^T \mathrm{H}_f(\mathrm{x}_0)\, \mathrm{h}$ definita positiva. Dal teorema 3.21 abbiamo:

$$q(\mathrm{h}) \geq \lambda_{\min} |\mathrm{h}|^2 \quad \text{per ogni } \mathrm{h} \in \mathbb{R}^n$$

dove $\lambda_{\min}$ è il minimo autovalore di $\mathrm{H}_f(\mathrm{x}_0)$. Dalla (6.13) otteniamo allora, per $\mathrm{h} \to 0$:

$$\begin{aligned} f(\mathrm{x}_0 + \mathrm{h}) - f(\mathrm{x}_0) &= \frac{1}{2} q(\mathrm{h}) + o(|\mathrm{h}|^2) \geq \frac{1}{2} \lambda_{\min} |\mathrm{h}|^2 + o(|\mathrm{h}|^2) \\ &= \frac{1}{2} \lambda_{\min} |\mathrm{h}|^2 (1 + o(1)) . \end{aligned}$$

Poiché $\lambda_{\min} > 0$ e $1 + o(1) > 0$ definitivamente per $\mathrm{h} \to 0$, si deduce che $f(\mathrm{x}_0 + \mathrm{h}) - f(\mathrm{x}_0) > 0$ per $|\mathrm{h}| \neq 0$, abbastanza piccolo e quindi x_0 è punto di minimo locale forte.

Analogamente, se $q(\mathrm{h}) = \mathrm{h}^T \mathrm{H}_f(\mathrm{x}_0)\, \mathrm{h}$ definita negativa, dal teorema 3.21 abbiamo:

$$q(\mathrm{h}) \leq \lambda_{\max} |\mathrm{h}|^2 \quad \text{per ogni } \mathrm{h} \in \mathbb{R}^n$$

dove $\lambda_{\max}$ (che è negativo) è il massimo autovalore di $\mathrm{H}_f(\mathrm{x}_0)$. Dalla (6.13) otteniamo allora, per $\mathrm{h} \to 0$:

$$\begin{aligned} f(\mathrm{x}_0 + \mathrm{h}) - f(\mathrm{x}_0) &= \frac{1}{2} q(\mathrm{h}) + o(|\mathrm{h}|^2) \leq \frac{1}{2} \lambda_{\max} |\mathrm{h}|^2 + o(|\mathrm{h}|^2) \\ &= \frac{1}{2} \lambda_{\max} |\mathrm{h}|^2 (1 + o(1)) . \end{aligned}$$

Poiché $\lambda_{\max} < 0$ e $1 + o(1) > 0$ definitivamente per $\mathrm{h} \to 0$, si deduce che $f(\mathrm{x}_0 + \mathrm{h}) - f(\mathrm{x}_0) < 0$ per $|\mathrm{h}| \neq 0$, abbastanza piccolo e quindi x_0 è punto di massimo locale forte.

(*ii*) Sia ora $q(\mathrm{h}) = \mathrm{h}^T \mathrm{H}_f(\mathrm{x}_0)\, \mathrm{h}$ indefinita. Esistono allora due vettori v e w tali che $q(\mathrm{v}) > 0$ e $q(\mathrm{w}) < 0$. Valutiamo $\Delta f(\mathrm{x}_0)$ lungo la retta $\mathrm{x}_0 + t\mathrm{v}$, $t \in \mathbb{R}$. Si ha, per $t \to 0$:

$$f(\mathrm{x}_0 + t\mathrm{v}) - f(\mathrm{x}_0) = \frac{1}{2} q(t\mathrm{v}) + o(|t\mathrm{v}|^2) = \frac{1}{2} t^2 (q(\mathrm{v}) + o(1))$$

e quindi, essendo $g(v)+o(1)>0$ definitivamente per $t\to 0$, si deduce che $f(x_0+tv)-f(x_0)>0$ per $|t|\neq 0$, abbastanza piccolo.

Ragionando analogamente, si mostra che $f(x_0+tw)-f(x_0)<0$ per $|t|\neq 0$, abbastanza piccolo. Poiché $\Delta f(x_0)$ cambia segno in ogni intorno di x_0, concludiamo che x_0 è punto di sella. ◆

Rimane dubbio il caso in cui la (6.13) risulta una f.q. semidefinita. Ragionando come nella dimostrazione del teorema precedente, possiamo comunque dedurre che: *se* $g(h)=h^T H_f(x_0)h$ *è non nulla e semidefinita positiva (negativa), allora* x_0 *è di minimo (massimo) debole oppure di colle.*

La situazione cambia se $g(h)=h^T H_f(x)h$ è semidefinita positiva o negativa *non solo in* x_0 ma per ogni x in un intorno di x_0. In tal caso, infatti, vedremo nel prossimo paragrafo che in quell'intorno f è *convessa* (o *concava*, rispettivamente) e quindi x_0 risulta punto di minimo (o massimo), debole.

Se $H_f(x_0)$ è nulla, ovviamente q non può dare alcuna informazione sulla natura di x_0 e occorrerebbe ricorrere ai differenziali di ordine superiore.

Sintetizziamo brevemente una possibile strategia per la ricerca di estremi liberi di una funzione $f: A\to\mathbb{R}$, con A sottoinsieme aperto di $\mathbb{R}^n$.

Passo 1. Si isolano i punti in cui f non è regolare (ad es. non derivabile una o due volte). Questi punti vanno esaminati a parte con uno studio in generale diretto ed adattato al caso concreto.

Passo 2. Tolti i punti anomali trovati nel passo 1, si determinano gli eventuali punti critici di f risolvendo il sistema (6.4).

Passo 3. Si esamina la natura di ogni punto stazionario studiando il segno della forma quadratica (6.14). Se questa è definita o indefinita, si conclude usando il teorema 3.22. Se la (6.14) è nulla o è semidefinita, in generale si ricorre a uno studio diretto del segno dell'incremento $\Delta f(x_0)$ in un intorno di x_0.

Nel caso bidimensionale, le considerazioni ai passi 1 e 2 conducono alla seguente regola.

PROPOSIZIONE 3.3 *Siano* $f\in C^2(A)$, A *aperto in* $\mathbb{R}^2$, $(x_0,y_0)\in A$ *un punto critico per* f *e*

$$H_f(x_0,y_0)=\begin{pmatrix} f_{xx}(x_0,y_0) & f_{xy}(x_0,y_0)\\ f_{xy}(x_0,y_0) & f_{yy}(x_0,y_0)\end{pmatrix}$$

la matrice hessiana di f *nel punto critico. Allora:*

a) Se $\det H_f(x_0,y_0)>0$ *e*

$f_{xx}(x_0,y_0)>0$ *allora* (x_0,y_0) *è di minimo locale forte*
$f_{xx}(x_0,y_0)<0$ *allora* (x_0,y_0) *è di massimo locale forte*

(si noti che in questo caso $f_{xx}(x_0,y_0)$ *e* $f_{yy}(x_0,y_0)$ *hanno lo stesso segno).*

b) Se $\det H_f(x_0,y_0)<0$ *allora* (x_0,y_0) *è punto di sella.*

c) Se $\det H_f(x_0,y_0)=0$ *occorre un'analisi ulteriore.*

Presentiamo ora alcuni esempi e osservazioni su come si utilizzano in pratica le idee precedenti.

Esempio

6.5 Possiamo ora completare la ricerca degli estremi per la funzione dell'esempio 6.1:

$$f(x,y)=3x^2+y^2-x^3y$$

Poiché f ha derivate di ogni ordine in $\mathbb{R}^2$, occorre considerare solo i punti critici $(0,0)$, $(\sqrt{2},\sqrt{2})$ e $(-\sqrt{2},-\sqrt{2})$.

La matrice hessiana di f è

$$H_f(x,y)=\begin{pmatrix} 6-6xy & -3x^2\\ -3x^2 & 2\end{pmatrix}.$$

In $(0,0)$ abbiamo

$$H_f(0,0)=\begin{pmatrix} 6 & 0\\ 0 & 2\end{pmatrix}.$$

La matrice ha autovalori 6 e 2 per cui è definita positiva: $(0,0)$ è punto di minimo locale forte. Non può essere di minimo globale poiché, per esempio, $f(x,1)\to-\infty$ se $x\to+\infty$.

Nel punto $(\sqrt{2},\sqrt{2})$ abbiamo

$$H_f(\sqrt{2},\sqrt{2})=\begin{pmatrix} -6 & -6\\ -6 & 2\end{pmatrix}.$$

La matrice ha determinante negativo e perciò è indefinita: $(\sqrt{2},\sqrt{2})$ è un punto di sella.

Infine, nel punto $(-\sqrt{2},-\sqrt{2})$ abbiamo ancora

$$H_f(-\sqrt{2},-\sqrt{2})=\begin{pmatrix} -6 & -6\\ -6 & 2\end{pmatrix},$$

per cui anche $(-\sqrt{2},-\sqrt{2})$ è un punto di sella.

La tecnica delle restrizioni per provare che un punto è di sella

Esempio

6.6 Cerchiamo massimi e minimi locali per la funzione

$$f(x,y)=x^4-6x^2y^2+y^4.$$

Poiché f ha derivate di ogni ordine in $\mathbb{R}^2$, passiamo alla ricerca dei punti critici, soluzioni del sistema:

$$\begin{cases} f_x(x,y)=4x^3-12xy^2=4x(x^2-3y^2)=0\\ f_y(x,y)=-12x^2y+4y^3=4y(y^2-3x^2)=0.\end{cases}$$

Si trova l'unico punto critico $(0,0)$. Poiché la funzione è un polinomio omogeneo di grado 4, la sua matrice hessiana nell'origine ha tutti gli elementi nulli (perché?), quindi siamo nel caso dubbio. Come si può studiare la natura di questo punto? Osserviamo che lungo la retta $y=0$ si ha

$$f(x,0)=x^4$$

 © 978-88-08-06485-1

che ha un punto di minimo in $x = 0$: lungo l'asse x, perciò, la funzione ha un punto di minimo nell'origine. Lo stesso accade lungo l'asse y: $f(0, y) = y^4$. Invece, lungo la retta $y = x$ si ha

$$f(x, x) = -4x^4$$

che ha un punto di massimo in $x = 0$. Pertanto $(0, 0)$ non può essere né di massimo né di minimo: è un punto di sella.

L'esempio precedente contiene un'idea generale: se in un caso dubbio si riescono a trovare due rette (o due curve) diverse, lungo le quali la restrizione di f ha in un caso un massimo e nell'altro un minimo nel punto critico, allora quest'ultimo è un punto di sella. Alla stessa conclusione si arriva se, lungo una particolare curva, la funzione ha un flesso nel punto critico.

Si noti che, invece, il fatto che lungo due rette diverse la funzione abbia lo stesso comportamento (per esempio, punto di massimo) non prova nulla.

Perturbazione di una forma quadratica

Esempio

Siano

$$f(x, y) = x^2 + 2y^2 - x^3 \quad \text{e} \quad g(x, y) = x^2 - y^3.$$

Poiché

$$\nabla f(x, y) = (2x - 3x^2, 4y) \quad \text{e} \quad \nabla g(x, y) = (2x, -3y^2),$$

l'origine è punto critico[21] per entrambe. La matrice hessiana di f è

$$\mathrm{H}_f(x, y) = \begin{pmatrix} 2 - 6x & 0 \\ 0 & 4 \end{pmatrix}$$

per cui

$$\mathrm{H}_f(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 4 \end{pmatrix}.$$

Gli autovalori sono uguali a 2 e 4, per cui $\mathrm{H}_f(0, 0)$ è definita positiva. Da ciò deduciamo che $(0, 0)$ è un punto di *minimo locale forte per* f.

La matrice hessiana di g è

$$\mathrm{H}_g(x, y) = \begin{pmatrix} 2 & 0 \\ 0 & -6y \end{pmatrix}$$

per cui

$$\mathrm{H}_g(0, 0) = \begin{pmatrix} 2 & 0 \\ 0 & 0 \end{pmatrix}.$$

Gli autovalori sono uguali a 2 e 0, per cui $\mathrm{H}_g(0, 0)$ è *semidefinita* positiva. Da ciò deduciamo che $(0, 0)$ è un punto di *minimo debole o di sella* per g. Come si fa a decidere? In questo caso è semplice: muovendoci lungo l'asse $x = 0$, si ottiene $f(0, y) = -y^3$, che ha un flesso in $y = 0$. Deduciamo subito che $(0, 0)$ è punto di sella per g.

[21] Se f è un polinomio privo di termini di ordine uno, l'origine è sempre punto stazionario. Il differenziale secondo $d^2 f(0)$ è dato dal complesso dei termini di secondo grado moltiplicato per due.

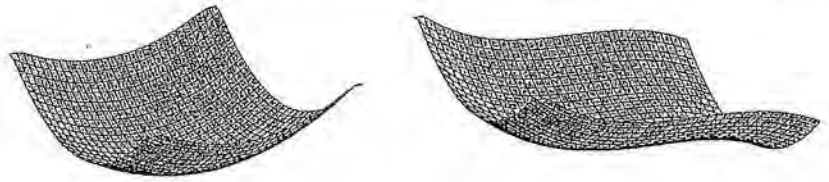
© 978-88-08-06485-1

Esaminiamo i due esempi precedenti da un altro punto di vista, utilizzando semplici esperimenti al computer per illustrare meglio l'idea principale che sta alla base dei ragionamenti precedenti.

La funzione f si scrive come somma:

- di una forma quadratica definita positiva ($q(x, y) = x^2 + 2y^2$) e
- di un termine cubico ($= -x^3$), in teoria molto più piccolo quando si sta vicino all'origine.

Possiamo dunque considerare f, vicino all'origine, come una piccola deformazione della forma quadratica q. Il calcolo fatto indica che, poiché nell'origine quest'ultima ha un *minimo forte*, deformandola poco lì vicino, le cose sostanzialmente non cambiano. Alla peggio, il minimo non sarà più globale ma locale. L'effetto è mostrato in figura 3.18.



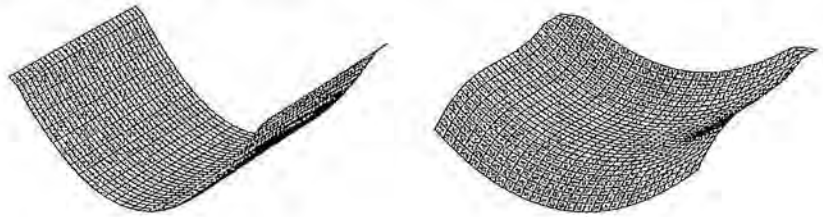
a) b)

Figura 3.18. Perturbazione di una forma quadratica definita positiva vicino all'origine. a) $q(x, y) = x^2 + 2y^2$. b) $f(x, y) = x^2 + 2y^2 - x^3$.

La funzione g si scrive invece come somma:

- di una forma quadratica *semidefinita* positiva ($q_1(x, y) = x^2$) e
- di un termine cubico ($= -y^3$).

La f.q. $q_1(x, y) = x^2$ ha una linea di minimi deboli, tra cui l'origine. Perturbandola con il termine $-y^3$, si distrugge la configurazione di minimo debole. In figura 3.19 è mostrato il grafico di q_1 e della sua deformata $g(x, y) = x^2 - y^3$. L'origine non è più punto d'estremo, neanche locale.



a) b)

Figura 3.19. a) $q_1(x, y) = x^2$. b) $g(x, y) = x^2 - y^3$.

 © 978-88-08-06485-1

Invitiamo il lettore a fare qualche altro esperimento al computer, per esempio con le funzioni:

$$f_1(x,y) = x^2 + 2y^2 - x^2y, \qquad f_2(x,y) = -x^2 - 2y^2 - xy^4, \qquad f_3(x,y) = x^2 - xy^5$$

ottenute perturbando una forma quadratica con termini di grado superiore. I grafici mostreranno che l'origine è ancora punto di estremo (locale) per le prime due. Le cose funzionano bene con le forme quadratiche definite positive o negative. La terza, nonostante si deformi la f.q. semidefinita positiva $q_3(x,y) = x^2$ con un termine di grado elevato, non ha un punto di estremo in $(0,0)$.

Massimi e minimi in più di due variabili

Esempio

Cerchiamo gli estremi liberi della *funzione di tre variabili*:

$$f(x,y,z) = 3x^2 + 2y^2 + z^2 - 2xz + 2x + 2y + 1.$$

I punti critici risolvono il sistema:

$$\begin{cases} f_x(x,y,z) = 6x + 2 - 2z = 0 \\ f_y(x,y,z) = 4y + 2 = 0 \\ f_z(x,y,z) = 2z - 2x = 0 \end{cases} \quad \text{ossia} \quad \begin{cases} z = x \\ y = -\frac{1}{2} \\ 2x + 1 = 0. \end{cases}$$

Si trova l'unico punto critico

$$x_0 = \left(-\frac{1}{2}, -\frac{1}{2}, -\frac{1}{2}\right).$$

Calcoliamo la matrice hessiana di f:

$$\mathrm{H}_f(x,y,z) = \begin{pmatrix} 6 & 0 & -2 \\ 0 & 4 & 0 \\ -2 & 0 & 2 \end{pmatrix}.$$

Dunque l'hessiana è costante. Per studiarne il segno, usiamo il teorema 3.19:

$$\det \mathrm{H}_f = 32 > 0 \qquad \det \begin{pmatrix} 6 & 0 \\ 0 & 4 \end{pmatrix} = 24 > 0 \qquad 6 > 0.$$

Poiché i minori principali di nord-ovest sono positivi, la forma quadratica è definita positiva. Il punto x_0 è di minimo locale forte.

Vediamo ora due esempi un po' meno accademici.

Baricentro di n punti nel piano

Siano dati n punti nel piano $\mathrm{p}_i = (x_i, y_i)$, $i = 1, \ldots, n$. Cerchiamo un punto $\mathrm{p} = (x,y)$ tale che la somma dei quadrati delle distanze dagli n punti assegnati sia minima.

Si tratta dunque di minimizzare la funzione:

$$f(x,y) = \sum_{i=1}^{n} [(x_i - x)^2 + (y_i - y)^2].$$

© 978-88-08-06485-1

Poiché f ha derivate continue di ogni ordine, passiamo alla ricerca dei punti critici, soluzioni del sistema:

$$\begin{cases} f_x(x,y) = \sum_{i=1}^{n} 2(x - x_i) = 2(nx - \sum_{i=1}^{n} x_i) = 0 \\ f_y(x,y) = \sum_{i=1}^{n} 2(y - y_i) = 2(ny - \sum_{i=1}^{n} y_i) = 0. \end{cases}$$

Si trova l'unico punto di coordinate:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i, \quad \bar{y} = \frac{1}{n}\sum_{i=1}^{n} y_i,$$

che coincide dunque col *baricentro* del sistema degli n punti. La matrice hessiana è costante:

$$\begin{pmatrix} 2n & 0 \\ 0 & 2n \end{pmatrix}$$

ed è definita positiva. Dunque, il punto è effettivamente di *minimo*. Abbiamo quindi scoperto che il baricentro di n punti nel piano minimizza la somma dei quadrati delle distanze dagli n punti.

Metodo dei minimi quadrati, retta di regressione

Molti problemi interessanti nell'ambito della statistica, dell'economia, delle scienze sperimentali, si riducono alla ricerca di una funzione

$$y = \beta x + \alpha,$$

detta *retta di regressione*, che sia il più possibile in accordo con una serie d'osservazioni congiunte di due variabili

$$(x_1, y_1), \quad (x_2, y_2), \quad \ldots, \quad (x_n, y_n).$$

L'idea generale soggiacente è la seguente: per ragioni teoriche a priori, oppure perché lo constatiamo empiricamente, siamo convinti che due grandezze siano correlate tra loro: per esempio, all'aumentare di una aumenta anche l'altra (correlazione diretta) oppure all'aumentare di una l'altra diminuisce (correlazione inversa). Vorremmo trovare un'espressione analitica esplicita della legge che lega le due grandezze e l'ipotesi più semplice è che possa trattarsi di una dipendenza di tipo *lineare affine*[22], $y = \beta x + \alpha$.

Il fatto di considerare una retta eventualmente non passante per l'origine (invece di supporre la più semplice relazione $y = kx$) corrisponde a pensare che, almeno approssimativamente, y sia costituito da una parte α indipendente da x e da una parte βx, proporzionale a x. Per scrivere l'equazione (ossia per determinare i coefficienti α e β), consideriamo quindi una serie di coppie di valori osservati (x_i, y_i) e cerchiamo la retta che passi "il più possibile vicino" (in un senso da precisarsi) a tutti questi

[22] Si chiama legge *lineare affine* una legge $y = \beta x + \alpha$, mentre *lineare*, come sappiamo, è una legge $y = \beta x$ (cioè con $\alpha = 0$).

punti, che in generale non saranno perfettamente allineati. Una volta determinata questa retta (mediante il metodo che ora illustreremo), questa potrà essere usata per *prevedere* il valore che y assumerà in corrispondenza di un valore x diverso da quelli già osservati. Prima di proseguire, facciamo qualche esempio.

Esempio

(a) Sia x il reddito nazionale e y l'ammontare dei consumi totali. Ci aspettiamo che i consumi aumentino all'aumentare del reddito nazionale (correlazione diretta), pur senza esserci una proporzionalità esatta tra le due grandezze. Il coefficiente α (consumi che in teoria corrispondono a un reddito nullo) quantifica i "consumi di sopravvivenza", mentre il coefficiente β indica il tasso di crescita dei consumi all'aumentare del reddito. Le coppie (x_i, y_i) osservate potrebbero riguardare gli anni passati $i = 1, 2, \ldots, n$.

(b) Sia x l'ammontare speso in pubblicità da una certa azienda in un certo distretto e intervallo temporale e y il corrispondente volume di vendite realizzato. Di nuovo, ci aspettiamo una correlazione diretta tra le due variabili. Le coppie (x_i, y_i) osservate potrebbero riguardare diversi distretti e periodi passati.

(c) Nello studio dei mercati azionari si cerca spesso di stimare una relazione del genere, che lega y, la differenza tra il rendimento di un'azione e quello di un investimento sicuro (per esempio in buoni del Tesoro), con x, la differenza tra il rendimento medio del complesso delle azioni e ancora quello dei B.O.T. In gergo finanziario, si chiamano proprio *coefficienti beta* e *coefficienti alfa* i parametri che entrano nell'equazione della retta; il coefficiente β è usato per classificare i titoli in *aggressivi* (che con $\beta > 1$ hanno variazioni di rendimento più marcate delle medie di mercato) e in *difensivi* (che con $\beta < 1$ smorzano le variazioni dei rendimenti di mercato).

Proseguendo con altri esempi ancora più schematicamente:

(d) Sia x la statura di un individuo maschio adulto e y il suo peso. Correlazione diretta.

(e) Sia x il numero di pezzi prodotti da un'industria e y il costo di produzione corrispondente. Correlazione diretta.

Tornando al discorso generale, supponiamo dunque di avere n osservazioni congiunte di due variabili (x, y), ossia

$$(x_1, y_1), \quad (x_2, y_2), \quad \ldots, \quad (x_n, y_n),$$

dove *supporremo i valori $x_1, x_2, \ldots, x_n$ non tutti uguali*, per non svuotare d'interesse il problema, e proponiamoci di determinare la retta di regressione. Naturalmente, in nessuno degli esempi fatti si pretende che la relazione lineare affine tra y e x sia esatta, ossia:

$$y_i = \beta x_i + \alpha \quad \text{per ogni } i.$$

L'ipotesi è che essa valga, a meno d'un errore ε_i, ossia

$$y_i = \beta x_i + \alpha + \varepsilon_i \quad \text{per ogni } i$$

e cioè che il valore "vero" y_i sia pari al valore teorico $\beta x_i + \alpha$, "sporcato" però dall'errore ε_i.

Come determinare i parametri α e β in modo da minimizzare il "complesso" degli ε_i? La prima idea che viene in mente è minimizzare la somma $\sum_{i=1}^n \varepsilon_i$, ma ci si rende conto immediatamente che errori consistenti di segno opposto potrebbero compensarsi e rendere piccola la somma. La somma dei quadrati di tali errori (ovviamente

funzione di β e α) non permette compensazioni ed è una buona misura della bontà dell'approssimazione. Tale somma è:

$$E(\beta, \alpha) = \sum_{i=1}^n \varepsilon_i^2 = \sum_{i=1}^n (\beta x_i + \alpha - y_i)^2$$

e aggrega i quadrati delle "distanze verticali" dei punti osservati (x_i, y_i) dalla retta. Con l'espressione distanza verticale intendiamo la differenza tra l'ordinata del punto $(x_i, \beta x_i + \alpha)$, che appartiene alla retta e l'ordinata del punto (x_i, y_i).

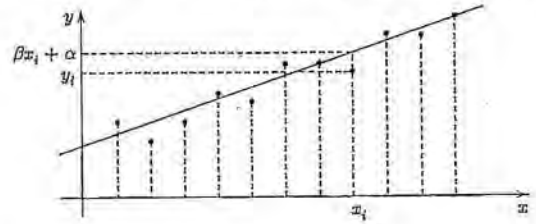


Figura 3.20. Retta che approssima un insieme di punti.

Il problema è allora minimizzare la funzione $E(\beta, \alpha)$, che è definita su $\mathbb{R}^2$ e ha derivate parziali continue in ogni punto. Per determinare i punti stazionari, calcoliamo le derivate parziali di E e uguagliamole a zero. Si ha:

$$\begin{cases} \dfrac{\partial E}{\partial \beta}(\beta, \alpha) \displaystyle\sum_{i=1}^n 2x_i(\beta x_i + \alpha - y_i) = 0 \\ \dfrac{\partial E}{\partial \alpha}(\beta, \alpha) = \displaystyle\sum_{i=1}^n 2(\beta x_i + \alpha - y_i) = 0 \end{cases}$$

ossia

$$\text{(6.15)} \qquad \begin{cases} \beta \displaystyle\sum_{i=1}^n x_i + n\alpha = \sum_{i=1}^n y_i \\ \beta \displaystyle\sum_{i=1}^n x_i^2 + \alpha \sum_{i=1}^n x_i = \sum_{i=1}^n x_i y_i. \end{cases}$$

Il sistema ottenuto è lineare. Poniamo:

$$\overline{x} = \frac{1}{n}\sum_{i=1}^n x_i; \quad \overline{y} = \frac{1}{n}\sum_{i=1}^n y_i; \quad \overline{q} = \frac{1}{n}\sum_{i=1}^n x_i^2; \quad \overline{p} = \frac{1}{n}\sum_{i=1}^n x_i y_i.$$

I numeri $\overline{x}$, $\overline{y}$, $\overline{q}$ e $\overline{p}$ rappresentano le medie aritmetiche delle ascisse, delle ordinate, dei quadrati delle ascisse e dei prodotti ascissa per ordinata, rispettivamente. Dopo averne diviso le equazioni membro a membro per n, il sistema (6.15) si può riscrivere nel modo seguente:

$$\text{(6.16)} \qquad \begin{cases} \overline{q}\beta + \overline{x}\alpha = \overline{p} \\ \overline{x}\beta + \alpha = \overline{y}. \end{cases}$$

Per procedere nella risoluzione del sistema, è opportuno ora introdurre due quantità significative in statistica:

– la *varianza della variabile* x, definita da:

$$\sigma_X^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2;$$

è una misura di quanto le osservazioni x_i si discostino dal valor medio $\overline{x}$; dalla definizione è evidente che si tratta di una quantità ≥ 0, che si annulla se e solo se tutte le x_i sono uguali a $\overline{x}$ (e quindi uguali tra loro, circostanza che abbiamo escluso); il seguente calcolo elementare mostra un altro modo utile di scrivere la varianza:

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \frac{1}{n}\sum_{i=1}^{n}(x_i^2 + \overline{x}^2 - 2x_i\overline{x}) = \left(\frac{1}{n}\sum_{i=1}^{n}x_i^2\right) + \overline{x}^2 - 2\overline{x}\frac{1}{n}\sum_{i=1}^{n}x_i =$$
$$= \left(\frac{1}{n}\sum_{i=1}^{n}x_i^2\right) - \overline{x}^2 = \overline{q} - \overline{x}^2;$$

– la *covarianza delle variabili* x *e* y, definita da:

$$\sigma_{XY} = \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}).$$

Il significato della covarianza è il seguente: una covarianza positiva significa che le variabili x e y sono direttamente correlate (ossia, all'aumentare di x aumenta anche y); una covarianza negativa significa che le variabili x e y sono inversamente correlate (ossia, all'aumentare di x diminuisce y)[23]. Un calcolo elementare analogo a quello fatto per la varianza mostra che:

$$\frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})(y_i - \overline{y}) = \left(\frac{1}{n}\sum_{i=1}^{n}x_iy_i\right) - \overline{x}\overline{y} = \overline{p} - \overline{x}\overline{y}.$$

Torniamo ora al sistema (6.16). Il determinante della matrice dei coefficienti del sistema è:

$$\begin{vmatrix} \overline{q} & \overline{x} \\ \overline{x} & 1 \end{vmatrix} = \overline{q} - \overline{x}^2 = \sigma_X^2 ,$$

pertanto nelle nostre ipotesi è senz'altro positivo. Se ne deduce che il sistema (6.16) ammette l'unica soluzione $(\overline{\beta}, \overline{\alpha})$ data da:

$$\beta = \frac{\overline{p} - \overline{x}\overline{y}}{\overline{q} - \overline{x}^2} = \frac{\sigma_{XY}}{\sigma_X^2}, \qquad \alpha = \frac{\overline{q}\overline{y} - \overline{p}\overline{x}}{\overline{q} - \overline{x}^2} = \overline{y} - \frac{\sigma_{XY}}{\sigma_X^2}\overline{x}. \tag{6.17}$$

[23] Per capire l'affermazione appena fatta si può ragionare così: la covarianza è positiva se, mediamente, le quantità $(x_i - \overline{x})$ e $(y_i - \overline{y})$ hanno lo stesso segno; questo accade se a valori "grandi" di x_i (cioè maggiori del valor medio $\overline{x}$) corrispondono valori grandi di y_i (cioè maggiori del valor medio $\overline{y}$) e a valori piccoli di x_i corrispondono valori piccoli di y_i. Analogamente si ragiona nel caso della covarianza negativa.

Il punto $(\overline{\beta}, \overline{\alpha})$, che è l'unico punto stazionario della funzione E, è punto di minimo assoluto, dato che E è somma di funzioni convesse[24]. In ogni caso, controlliamo esaminando la matrice hessiana di E, data da

$$\mathbf{H}_E(\beta, \alpha) = 2n\begin{pmatrix} \overline{q} & \overline{x} \\ \overline{x} & 1 \end{pmatrix}.$$

Si ha $\overline{q} > 0$ e $\det\mathbf{H}_E(\beta, \alpha) = \overline{q} - \overline{x}^2 > 0$ e quindi $(\overline{\beta}, \overline{\alpha})$ è effettivamente punto di minimo globale.

La retta d'equazione $y = \beta x + \alpha$ è pertanto la *retta di regressione* cercata. In base alle (6.17), la sua equazione si può scrivere nella forma:

$$y - \overline{y} = \frac{\sigma_{XY}}{\sigma_X^2}(x - \overline{x})$$

da cui risulta evidente che questa retta passa per il *baricentro* $(\overline{x}, \overline{y})$ dei punti $(x_1, y_1), \ldots, (x_n, y_n)$ e ha pendenza

$$\beta = \frac{\sigma_{XY}}{\sigma_X^2} = \frac{\text{covarianza tra } x \text{ e } y}{\text{varianza di } x}. \tag{6.18}$$

Esempio

Applichiamo il procedimento precedente alle coppie altezza / peso della seguente tabella:

x = altezza in cm	160	163	168	170	175	178	185	190	193
y = peso in kg	60	61	70	72	73	74	81	83	87

Dai dati precedenti, si ha:

$$\overline{x} = 175{,}79;\ \overline{y} = 73{,}44;\ \overline{q} = \frac{1}{n}\sum_{i=1}^{n}x_i^2 = 31\,019{,}56;\quad \overline{p} = \frac{1}{n}\sum_{i=1}^{n}x_iy_i = 13\,004.$$

Possiamo quindi calcolare:

$$\begin{array}{lll} \sigma_X^2 & = \overline{q} - \overline{x}^2 & = 121{,}73; \\ \sigma_{XY} & = \overline{p} - \overline{x}\overline{y} & = 94{,}1; \\ \beta & = \frac{\sigma_{XY}}{\sigma_X^2} & = 0{,}773; \\ \alpha & = \overline{y} - \beta\overline{x} & = -62{,}44. \end{array}$$

Perciò la retta di regressione è

$$y = 0{,}773x - 62{,}435.$$

Quest'ultima equazione può essere interpretata nel presente caso come una relazione "ottimale" tra peso e altezza, dove il criterio di ottimalità è quello di rendere minimo lo scarto quadratico totale dai dati della tabella. Per esempio, secondo questo legame, un'altezza di cm 180 dovrebbe essere corrispondente a un peso "ideale" di 76,705 kg.

La figura 3.21 mostra i punti e la retta di regressione calcolata.

[24] Quest'ultima affermazione sarà giustificata nel paragrafo 7.2.

 © 978-88-08-06485-1

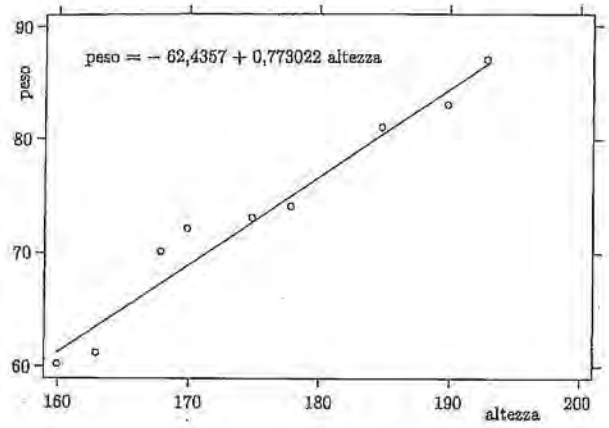


Figura 3.21. Retta che approssima un insieme di punti.

Esercizi

46 Classificare le seguenti forme quadratiche in $\mathbb{R}^2$:

$$q_1(x,y) = 2x^2 + 3xy - 5y^2$$
$$q_2(x,y) = 2x^2 + 4xy + 3y^2$$
$$q_3(x,y) = x^2 + 4xy + 4y^2$$
$$q_4(x,y) = 3xy$$
$$q_5(x,y) = -2x^2 + 2xy - 5y^2$$

47 Classificare, al variare del parametro reale α, la forma quadratica rappresentata dalla matrice:

$$\begin{pmatrix} 1+\alpha & \alpha \\ 1 & 2\alpha \end{pmatrix}.$$

48 Stabilire se le seguenti forme quadratiche su $\mathbb{R}^3$ possono essere classificate in base al teorema 3.19 e in caso affermativo classificarle:

$$q_1(x,y,z) = x^2 + 3y^2 + 2z^2 + 2xy - 2yz - 2xz$$
$$q_2(x,y,z) = -x^2 - 2y^2 - 5z^2 + 2xy + 4yz$$
$$q_3(x,y,z) = 2xz - 2xy - y^2 - 2yz$$
$$q_4(x,y,z) = 5x^2 + y^2 + z^2 + 2xy - 4xz$$

© 978-88-08-06485-1

49 Studiare i punti di massimo e minimo delle seguenti funzioni:

$$f_1(x,y) = xye^{-x^2-y^2}$$
$$f_2(x,y) = x^4 + y^4 - 3(x-y)^2$$
$$f_3(x,y) = x^2 + y^2 + x^3y$$
$$f_4(x,y) = \log(1+x^2+y^2) - 3xy$$
$$f_5(x,y) = e^{-2(x^2+y^2)} + x^2 + y^2$$
$$f_6(x,y) = xy^2e^{-x^4-y^2}$$
$$f_7(x,y) = x^3 - x^2y$$
$$f_8(x,y) = (\sin x)^2 + \cos y$$

Suggerimenti:
per f_6: oltre ad alcuni punti critici isolati, che si studiano coi soliti metodi, la funzione presenta un'intera retta di punti critici: decidere la natura di questi punti con considerazioni sul segno della funzione;
per f_7: studiare il segno della funzione;
per f_8: ci si può limitare a studiare i punti critici che cadono in $[0, 2\pi) \times [0, 2\pi)$.

50 Controllare che l'origine è punto critico per la funzione

$$f(x,y) = \log(1+x^2) - x^2 + xy^2 + y^3 + 2$$

e poi determinarne la natura.

51 Determinare gli estremi liberi delle seguenti funzioni:

$$f(x,y,z) = y^2 + z^2 - 2x^2 + 2xy - 2xz - 4x$$
$$f(x,y,z) = xyze^{-x^2-2y^2-3z^2}$$

52 Generalizzare il risultato sul baricentro di n punti nel piano (discusso nel paragrafo 6.5) al caso di n punti nello spazio tridimensionale.

53 Si può dimostrare che le funzioni che soddisfano l'equazione di Laplace non hanno massimi e minimi locali. Verificare questo fatto per le seguenti funzioni (nell'esercizio 37 si è provato che esse soddisfano tale equazione).
a) $f(x,y) = e^x \sin y$;
b) $f(x,y) = \cos x \operatorname{Sh} y$.
Suggerimento: per studiare la natura dei punti critici dubbi, studiare il segno di f);
c) $f(x,y) = x^3 - 3xy^2$
Suggerimento: per studiare la natura dei punti critici dubbi, usare la tecnica delle restrizioni.

54 Si vuole stabilire la relazione tra il costo per la produzione in serie di un certo prodotto e le dimensioni del processo (cioè il numero di pezzi prodotti). Da produzioni passate sono noti i seguenti dati:

n° di pezzi prodotti	100	120	240	98	132	180	170	200
costo (in euro)	528	600	1240	520	700	920	850	1020

a) Determinare la retta di regressione.

b) In base all'equazione della retta di regressione, determinare il costo fisso e il costo unitario di produzione.

c) Per una produzione di 300 pezzi, usare la retta di regressione per prevedere: il costo totale di produzione; il costo medio per pezzo prodotto.

7 FUNZIONI CONVESSE DI n VARIABILI

7.1 Generalità sulle funzioni convesse

Ci occupiamo ora brevemente delle *funzioni convesse*. Si tratta di una classe di funzioni importante per varie questioni, in particolare legate ai problemi di ottimizzazione e per questo motivo le trattiamo a questo punto del discorso.

Ricordiamo che un *insieme* $\Omega \subseteq \mathbb{R}^n$ si dice *convesso* se per ogni coppia di punti $x_1, x_2 \in \Omega$ si ha che $[x_1, x_2] \subseteq \Omega$ (dove il simbolo $[x_1, x_2]$ denota il segmento di estremi x_1, x_2); si dice *strettamente convesso* se per ogni coppia di punti $x_1, x_2 \in \Omega$ il segmento (x_1, x_2) privato degli estremi è strettamente contenuto in Ω.

DEFINIZIONE 3.21 Si dice *epigrafico* di una funzione $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ l'insieme

$$\text{epi} f = \{(x, z) \in \mathbb{R}^{n+1} : z \geq f(x), x \in \Omega\}.$$

Si dice che una funzione $f : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}$ è *convessa* (strettamente convessa) se epi$f$ è un sottoinsieme convesso (rispettivamente, strettamente convesso) di $\mathbb{R}^{n+1}$; si dice che f è *concava* se $-f$ è convessa.

Osserviamo che se l'insieme Ω su cui f è definita non è convesso, epi$f$ non è mai convesso (il lettore è invitato a convincersene con un disegno), quindi d'ora in poi supporremo sempre che Ω sia convesso. Osserviamo che:

PROPOSIZIONE 3.4 *Una funzione $f : \Omega \to \mathbb{R}$ (con $\Omega \subseteq \mathbb{R}^n$ convesso) è convessa se e solo se per ogni $x_1, x_2 \in \Omega, t \in [0,1]$ vale la condizione*

$$f(tx_2 + (1-t)x_1) \leq tf(x_2) + (1-t)f(x_1). \tag{7.1}$$

Se la (7.1) vale col segno $<$ per $t \in (0,1)$, la f è strettamente convessa.

Si invita il lettore a verificare questa proposizione. Si noti che il punto $x_t = tx_2 + (1-t)x_1$ percorre il segmento $[x_1, x_2]$ al variare di $t \in [0,1]$. La (7.1) è l'analogo multidimensionale della proprietà di "convessità per corde" che abbiamo studiato per le funzioni di una variabile[25].

Esempio

La funzione

$f(x) = |x|^2$ è strettamente convessa in $\mathbb{R}^n$;

$f(x) = |x|$ è convessa in $\mathbb{R}^n$ (ma non strettamente).

[25] v. volume 1, capitolo 4, paragrafo 5.2.

Anche se la condizione (7.1), a priori, ha senso per una funzione f qualsiasi (ossia non richiede, per aver senso, che la funzione sia continua, o derivabile), si dimostra che in effetti la (7.1) implica una certa regolarità:

TEOREMA 3.23 (REGOLARITÀ DELLE FUNZIONI CONVESSE) *Sia $f : \Omega \to \mathbb{R}$ (con $\Omega \subseteq \mathbb{R}^n$ aperto convesso). Se f è convessa, allora:*

a) f è continua;

b) f ha derivate parziali destre e sinistre in ogni punto;

c) nei punti in cui è derivabile, f è differenziabile.

Non dimostriamo questo teorema. Vediamo invece qualche risultato che vale per funzioni convesse che a priori sono un po' più regolari:

TEOREMA 3.24 (CONVESSITÀ E PIANO TANGENTE) *Sia $f : \Omega \to \mathbb{R}$ (con $\Omega \subseteq \mathbb{R}^n$ aperto convesso), differenziabile in Ω. Allora f è convessa in Ω se e solo se per ogni coppia di punti $x_0, x \in \Omega$ si ha:*

$$f(x) \geq f(x_0) + \nabla f(x_0) \cdot (x - x_0). \tag{7.2}$$

Il significato geometrico della (7.2) si vede bene in due dimensioni. Infatti, scritta esplicitamente in questo caso, si ha:

$$f(x,y) \geq f(x_0,y_0) + \frac{\partial f}{\partial x}(x_0,y_0)(x - x_0) + \frac{\partial f}{\partial y}(x_0,y_0)(y - y_0)$$

che equivale ad affermare che il piano tangente in (x_0, y_0) sta sotto il grafico di f. Questo teorema risulta quindi l'analogo multidimensionale della proprietà di "convessità per tangenti" enunciata per le funzioni di una variabile[26].

DIMOSTRAZIONE. Sia f convessa e proviamo la (7.2). Riscriviamo la (7.1) nella forma

$$\frac{f(x_0 + t(x - x_0)) - f(x_0)}{t} \leq f(x) - f(x_0).$$

Per $t \to 0$ il primo membro tende a

$$\frac{d}{dt} f(x_0 + t(x - x_0))_{/t=0} = \nabla f(x_0) \cdot (x - x_0),$$

poiché f è differenziabile (abbiamo applicato il teorema di derivazione delle funzioni composte); quindi

$$\nabla f(x_0) \cdot (x - x_0) \leq f(x) - f(x_0)$$

ossia

$$f(x) \geq f(x_0) + \nabla f(x_0) \cdot (x - x_0),$$

che è la (7.2).

Viceversa, supponiamo ora valida la (7.2) e proviamo che f è convessa. Ponendo $x_t = tx_2 + (1-t)x_1$, applichiamo la (7.2) nei punti x_2, x_t:

$$f(x_2) \geq f(x_t) + \nabla f(x_t) \cdot (x_2 - x_t), \tag{7.3}$$

[26] Si veda il volume 1, capitolo 4, paragrafo 5.2, teorema 4.12.

Applicandola invece ai punti x_1, x_t, abbiamo:

$$f(x_1) \geq f(x_t) + \nabla f(x_t) \cdot (x_1 - x_t). \tag{7.4}$$

Ora moltiplichiamo ambo i membri di (7.3) per t, ambo i membri di (7.4) per $1-t$ e sommiamo:

$$\begin{aligned} t f(x_2) + (1-t) f(x_1) &\geq f(x_t) + \nabla f(x_t) \cdot [t(x_2 - x_t) + (1-t)(x_1 - x_t)] \\ &= f(x_t) \end{aligned}$$

in quanto $t(x_2 - x_t) + (1-t)(x_1 - x_t) = 0$, come si verifica. Questo completa la dimostrazione. ♦

Come per le funzioni di una variabile il segno della derivata seconda fornisce un criterio per studiare la convessità di una funzione, così per le funzioni di più variabili un analogo criterio è dato dallo studio del segno della forma quadratica data dal differenziale secondo:

TEOREMA 3.25 (CONVESSITÀ E MATRICE HESSIANA) *Sia $f \in C^2(\Omega)$ con Ω aperto convesso di $\mathbb{R}^n$. Se per ogni $x_0 \in \Omega$ la forma quadratica $d^2 f(x_0)$ è semidefinita positiva, allora f è convessa in Ω.*

DIMOSTRAZIONE. Applichiamo la formula di Taylor al second'ordine con resto secondo Lagrange a due generici punti $x_0, x \in \Omega$:

$$f(x) = f(x_0) + \nabla f(x_0) \cdot (x - x_0) + \frac{1}{2} \sum_{i,j=1}^{n} f_{x_i x_j}(x_0 + \delta h) h_i h_j,$$

con $h = x - x_0$, per un opportuno $\delta \in (0,1)$ dipendente da x_0 e x. Ora, poiché in ogni punto di Ω la forma quadratica $d^2 f$ è semidefinita positiva, si ha:

$$\frac{1}{2} \sum_{i,j=1}^{n} f_{x_i x_j}(x_0 + \delta h) h_i h_j \geq 0 \quad \text{per ogni } h,$$

perciò

$$f(x) \geq f(x_0) + \nabla f(x_0) \cdot (x - x_0),$$

da cui si deduce la convessità di f, in base al teorema 3.24. ♦

7.2 Ottimizzazione di funzioni convesse e concave

Nel paragrafo 6 abbiamo visto che nella ricerca dei punti di massimo e minimo di una funzione (derivabile), il primo passo consiste nella ricerca dei punti stazionari, ossia quelli in cui il gradiente si annulla (condizione del prim'ordine); successivamente, occorre in generale studiare ciascuno dei punti stazionari mediante la matrice hessiana (condizioni del second'ordine). Un importante caso in cui le condizioni del prim'ordine risultano sufficienti alla ricerca dei punti di massimo e minimo è quello delle funzioni *convesse* o *concave*, per le quali i punti stazionari, se esistono, sono di *minimo globale* o *di massimo globale*, rispettivamente. Infatti:

PROPOSIZIONE 3.5 *Sia Ω un aperto convesso, $f : \Omega \to \mathbb{R}$ convessa (concava) e differenziabile in Ω, $x_0 \in \Omega$. Se x_0 è un punto critico per f, allora x_0 è punto di minimo (massimo) globale. Inoltre, se f è strettamente convessa (concava), x_0 è di minimo (massimo) globale forte e quindi in particolare il punto di minimo (massimo) globale è unico.*

DIMOSTRAZIONE. Se f è convessa e differenziabile, si ha:

$$f(x) \geq f(x_0) + \nabla(x_0) \cdot (x - x_0) \qquad \text{per ogni } x \in \Omega. \tag{7.5}$$

Se x_0 è critico, allora $\nabla(x_0) = 0$ e quindi $f(x) \geq f(x_0)$ per ogni $x \in \Omega$, da cui segue che x_0 è punto di minimo globale. Se, infine, f è strettamente convessa, la (7.5) vale in senso stretto per $x \neq x_0$, per cui x_0 è punto di minimo globale forte ed è unico. ♦

Esempio

Il problema di allocazione ottima. La proposizione precedente permette ora di completare l'analisi del problema di allocazione ottima, presentato nel paragrafo 6.1. Ricordiamo che l'obiettivo da minimizzare è

$$f(x,y) = \sum_{i=1}^{n} d_i(x,y)$$

dove $d_i(x,y) = |p - p_i| = \sqrt{(x_i - x)^2 + (y_i - y)^2}$ con $i = 1,2,3$ e $p_i = (x_i, y_i)$ sono tre punti disposti ai vertici di un triangolo acutangolo. Sappiamo già che il minimo esiste (v. paragrafo 6.1). Mostriamo che il punto di minimo è unico provando che f è strettamente convessa nel piano. Infatti, ogni addendo è convesso, non strettamente, poiché, per ogni $t \in (0,1)$ e ogni coppia di punti p, q, si ha:

$$|tp + (1-t)q - p_i| = |t(p - p_i) + (1-t)(q - p_i)| \leq |p - p_i| + (1-t)|q - p_i|.$$

Il segno di uguaglianza vale solo nel caso in cui i vettori $p - p_i$ e $q - p_i$ sono linearmente dipendenti. Ma allora deve essere, per ogni p, q, $p \neq q$,

$$\begin{aligned} f(tp + (1-t)q) &= \sum_{i=1}^{3} |tp + (1-t)q - p_i| < \\ &< t \sum_{i=1}^{3} |p - p_i| + (1-t) \sum_{i=1}^{3} |q - p_i| = t f(p) + (1-t) f(q) \end{aligned}$$

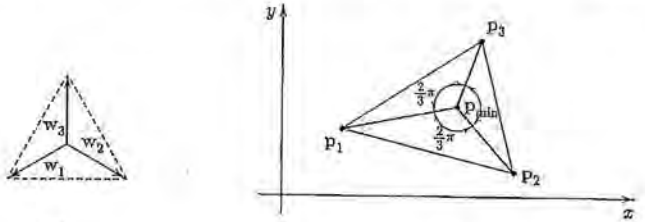
poiché il segno di uguale può valere solo nel caso in cui si ha *simultaneamente* che i vettori $p - p_i$ e $q - p_i$, per ogni $i = 1,2,3$, siano linearmente dipendenti e ciò è impossibile. Dunque f è strettamente convessa e per la proposizione precedente il punto di minimo è *unico*.

Vogliamo ora determinare effettivamente tale punto di minimo. Osserviamo che f è differenziabile in tutti i punti del piano tranne che nei vertici del triangolo. Esaminiamo subito questi tre punti e facciamo vedere che nessuno di essi può essere il punto di minimo. Basta considerare il vertice comune ai due lati più corti. Infatti, sia per esempio p_1 tale vertice e sia q il piede della perpendicolare condotta dal vertice p_2 al lato opposto. Allora si controlla facilmente che:

$$\begin{aligned} f(q) &= |q - p_1| + |q - p_2| + |q - p_3| = |q - p_2| + |p_1 - p_3| \\ &< |p_1 - p_2| + |p_1 - p_3| = f(p_1) \end{aligned}$$

per cui p_1 non può essere il punto di minimo.

 © 978-88-08-06485-1



a) b)

Figura 3.22. a) L'unica disposizione (a meno di rotazioni del piano) di tre versori in $\mathbb{R}^2$, affinché la loro somma sia 0. b) La soluzione del problema di minimo per il triangolo di figura 3.15.

Possiamo dunque limitarci a considerare gli altri punti. Cerchiamo allora i punti critici annullando il gradiente di f. Si trova il sistema seguente:

$$\begin{cases} \dfrac{\partial f}{\partial x} = \dfrac{x-x_1}{d_1(x,y)} + \dfrac{x-x_2}{d_2(x,y)} + \dfrac{x-x_3}{d_3(x,y)} = 0 \\ \dfrac{\partial f}{\partial y} = \dfrac{y-y_1}{d_1(x,y)} + \dfrac{y-y_2}{d_2(x,y)} + \dfrac{y-y_1}{d_3(x,y)} = 0. \end{cases}$$

Per risolverlo, notiamo che, introdotti i tre vettori (anzi, *versori*)

$$\mathrm{w}_1 = \left(\frac{x-x_1}{d_1(x,y)}, \frac{y-y_1}{d_1(x,y)}\right), \quad \mathrm{w}_2 = \left(\frac{x-x_2}{d_2(x,y)}, \frac{y-y_2}{d_2(x,y)}\right), \quad \mathrm{w}_3 = \left(\frac{x-x_3}{d_3(x,y)}, \frac{y-y_3}{d_3(x,y)}\right),$$

il sistema equivale all'equazione vettoriale

$$\mathrm{w}_1 + \mathrm{w}_2 + \mathrm{w}_3 = 0.$$

Ora, tre versori non nulli nel piano possono avere per somma il vettore nullo solo se i punti del piano corrispondenti sono disposti ai vertici di un triangolo equilatero, ovvero se l'angolo che ciascuno di essi forma con gli altri due è di 120 gradi (fig. 3.22). Esiste un unico punto che soddisfa questa condizione ed è interno al triangolo, come mostrato in figura 3.22.
A titolo di confronto, il lettore è invitato a calcolare la matrice hessiana di f, per rendersi conto che in questo caso l'analisi del segno dell'hessiana risulterebbe una strada più laboriosa per lo studio dei punti di massimo e minimo, rispetto alle considerazioni che abbiamo presentato qui.

Esempi

Sia $f(x,y) = g\left(\sqrt{x^2+y^2}\right)$ una funzione di due variabili, radiale, C^2.

a) Calcolare la matrice hessiana di f, verificando che ha la seguente forma:

$$\mathrm{H}_f(x,y) = \frac{g''(\rho)}{\rho^2}\begin{bmatrix} x^2 & xy \\ xy & y^2 \end{bmatrix} + \frac{g'(\rho)}{\rho^3}\begin{bmatrix} y^2 & -xy \\ -xy & x^2 \end{bmatrix}.$$

b) Dedurne che f è convessa nell'insieme in cui g' e g'' sono ≥ 0 ed è concava nell'insieme in cui g' e g'' sono ≤ 0.

© 978-88-08-06485-1

Determinare gli insiemi del piano in cui le seguenti funzioni sono convesse, concave o nessuna delle due (calcolando direttamente la matrice hessiana, oppure utilizzando il risultato dell'esercizio precedente):

i) $f(x,y) = e^{x^2+y^2}$; *ii)* $f(x,y) = e^{-x^2-y^2}$; *iii)* $f(x,y) = \sin x \sin y$.

8 FUNZIONI DEFINITE IMPLICITAMENTE

8.1 Funzione implicita di una variabile

Nel paragrafo 1 abbiamo affermato, giustificando l'affermazione attraverso alcuni esempi, che gli insiemi definiti nel piano da un'equazione $f(x,y) = cost.$ sono solitamente delle curve, dette linee di livello della funzione $z = f(x,y)$. Questa affermazione va presa con una certa cautela, come mostrano i prossimi esempi.

Esempi

L'insieme definito dall'equazione $x^2+y^2 = 1$ è una circonferenza (curva regolare);
l'insieme definito dall'equazione $x^2+y^2 = 0$ è un punto (l'origine);
l'insieme definito dall'equazione $x^2+y^2 = -1$ è vuoto;
l'insieme definito dall'equazione $x^2-y^2 = 0$ è l'unione di due rette: $y = \pm x$;
l'insieme definito dall'equazione $x^3-y^2 = 0$ è una curva non regolare (fig. 3.23).

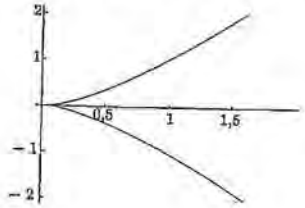


Figura 3.23.

Come si vede, anche con f molto regolare, l'insieme di livello definito da $f(x,y) = cost.$ non è necessariamente una curva, né una curva regolare.

Poniamoci il problema di specificare sotto quali condizioni l'equazione $f(x,y) = 0$ definisce una curva regolare. Più precisamente, ci poniamo un problema più ristretto (e locale):

Data una funzione $f(x,y)$ definita in un aperto del piano e ivi regolare (almeno C^1), precisare le condizioni sotto le quali l'equazione $f(x,y) = 0$ definisce implicitamente una funzione $y = g(x)$.

Precisiamo: una funzione $g : I \to \mathbb{R}$ (con $I \subseteq \mathbb{R}$ intervallo) tale che

$$f(x, g(x)) = 0 \text{ per ogni } x \in I$$

si dice "definita implicitamente dall'equazione $f(x,y)=0$", o più brevemente "funzione implicita".

Affinché una tale funzione g esista, è necessario che l'equazione $f(x,y)=0$ sia soddisfatta almeno in un punto (x_0,y_0). In tal caso sarà $g(x_0)=y_0$ e il problema è capire se una tale g può essere definita in tutto un intorno I di x_0. Inoltre, vorremmo che $g(x)$ fosse abbastanza regolare (almeno derivabile).

Ragioniamo a ritroso: se una tale funzione $g(x)$ esiste ed è derivabile in I, essendo anche $f(x,y)$ differenziabile per ipotesi, possiamo derivare rispetto ad x l'identità

$$f(x,g(x))=0$$

ottenendo, per il teorema di derivazione delle funzioni composte:

$$f_x(x,g(x))+f_y(x,g(x))\cdot g'(x)=0$$

da cui si ricava che

$$g'(x)=-\frac{f_x(x,g(x))}{f_y(x,g(x))}$$

per ogni $x\in I$ in cui il denominatore non si annulla. In particolare,

$$g'(x_0)=-\frac{f_x(x_0,y_0)}{f_y(x_0,y_0)}$$

purché sia $f_y(x_0,y_0)\neq 0$. Sotto ipotesi ragionevoli, l'esistenza e la derivabilità di g possono essere effettivamente dimostrate. Precisamente, abbiamo:

TEOREMA 3.26 (DI DINI, DELLA FUNZIONE IMPLICITA) *Sia A un aperto di $\mathbb{R}^2$ e $f:A\to\mathbb{R}$ una funzione $C^1(A)$. Supponiamo che in un punto $(x_0,y_0)\in A$ sia:*

$$f(x_0,y_0)=0 \text{ e } f_y(x_0,y_0)\neq 0.$$

Allora esiste un intorno I di x_0 in $\mathbb{R}$ e un'unica funzione $g:I\to\mathbb{R}$, tale che $y_0=g(x_0)$ e

$$f(x,g(x))=0 \text{ per ogni } x\in I.$$

Inoltre, $g\in C^1(I)$ e

$$g'(x)=-\frac{f_x(x,g(x))}{f_y(x,g(x))} \text{ per ogni } x\in I. \tag{8.1}$$

Notiamo anche che, se $f(x_0,y_0)=0$ e $f_y(x_0,y_0)=0$, ma in compenso $f_x(x_0,y_0)\neq 0$, si può applicare il teorema scambiando i ruoli di x e y, ossia affermare che esiste un intorno J di y_0 e un'unica funzione $x=h(y)$ definita in J, tale che $x_0=h(y_0)$ e

$$f(h(y),y)=0 \text{ per ogni } y\in J.$$

Inoltre $h\in C^1(J)$ e

$$h'(y)=-\frac{f_y(x,y)}{f_x(x,y)} \text{ per ogni } y\in J. \tag{8.2}$$

In sostanza, i punti in cui il teorema del Dini *non* è applicabile sono quelli in cui il gradiente di f si annulla, ossia i punti critici di f.

DIMOSTRAZIONE. La dimostrazione si articola in vari passi.

1. *Esistenza e unicità della funzione implicita.* Per ipotesi $f_y(x_0,y_0)\neq 0$, supponiamo per fissare le idee che sia $f_y(x_0,y_0)>0$. Poiché per ipotesi f_y è continua in A, per il teorema di permanenza del segno (teorema 3.1 paragrafo 2.1) esiste un rettangolo

$$R=[x_0-a,x_0+a]\times[y_0-b,y_0+b]$$

in cui $f_y(x,y)>0$. Allora la funzione di una variabile

$$y\longmapsto f(x_0,y)$$

è continua e strettamente crescente in $[y_0-b,y_0+b]$. In particolare, poiché per ipotesi $f(x_0,y_0)=0$, questo implica che

$$f(x_0,y_0-b)<0 \text{ e } f(x_0,y_0+b)>0.$$

Ancora per il teorema di permanenza del segno, esiste allora un intervallo

$$I\equiv(x_0-\delta,x_0+\delta)\subset[x_0-a,x_0+a]$$

tale che

$$f(x,y_0-b)<0 \text{ e } f(x,y_0+b)>0 \text{ per ogni } x\in I.$$

Se ora fissiamo un punto qualsiasi $\overline{x}\in I$, la funzione di una variabile

$$y\longmapsto f(\overline{x},y)$$

è continua su $[y_0-b,y_0+b]$, negativa in y_0-b, positiva in y_0+b e strettamente crescente. Per il teorema degli zeri (per funzioni di una variabile[27]) esiste uno e un sol punto $\overline{y}\in(y_0-b,y_0+b)$ tale che $f(\overline{x},\overline{y})=0$. Ponendo $g(\overline{x})=\overline{y}$ abbiamo dimostrato esistenza e unicità della funzione implicita, sull'intervallo I.

2. *Continuità di g.* Mostriamo prima che g è continua in x_0. Fissato $\varepsilon>0$, ripetiamo la dimostrazione precedente con ε al posto di b; troveremo un δ_ε tale che per ogni $\overline{x}\in(x_0-\delta_\varepsilon,x_0+\delta_\varepsilon)$ esiste uno e un sol punto $\overline{y}\in(y_0-\varepsilon,y_0+\varepsilon)$ tale che $f(\overline{x},\overline{y})=0$. Per l'unicità della funzione implicita dev'essere $\overline{y}=g(\overline{x})$, da cui l'implicazione

$$\overline{x}\in(x_0-\delta_\varepsilon,x_0+\delta_\varepsilon)\Longrightarrow g(\overline{x})\in(y_0-\varepsilon,y_0+\varepsilon),$$

che esprime la continuità di g in x_0. D'altro canto, la condizione

$$f_y(x,y)>0 \text{ in } I\times[y_0-b,y_0+b]$$

consente di ripetere lo stesso ragionamento in qualsiasi punto $x\in I$, perciò g è continua in I.

3. *Derivabilità di g.* Applichiamo il teorema del valor medio (teorema 3.13 paragrafo 4.5) a f in due punti qualsiasi

$$(x,y),(x_1,y_1)\in I\times[y_0-b,y_0+b].$$

[27] v. volume 1, capitolo 3, paragrafo 4.1.

 © 978-88-08-06485-1

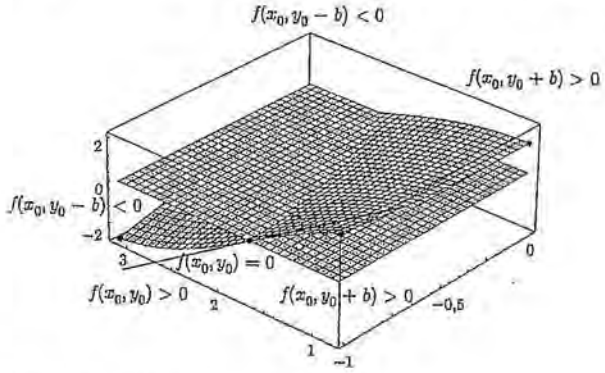


Figura 3.24. Esistenza della funzione implicita.

Esisterà dunque un punto (x^*, y^*), sul segmento di estremi (x, y), (x_1, y_1), tale che

$$f(x_1, y_1) - f(x, y) = f_x(x^*, y^*)(x_1 - x) + f_y(x^*, y^*)(y_1 - y).$$

Scegliendo ora $y_1 = g(x_1)$ e $y = g(x)$, avremo $f(x_1, y_1) = f(x, y) = 0$ e quindi:

$$f_x(x^*, y^*)(x_1 - x) + f_y(x^*, y^*)(g(x_1) - g(x)) = 0$$

da cui

$$\frac{g(x_1) - g(x)}{x_1 - x} = -\frac{f_x(x^*, y^*)}{f_y(x^*, y^*)}.$$

Essendo

$$f_y(x^*, y^*) \geq \min_R f_y(x, y) > 0,$$

l'ultimo denominatore scritto non si annulla mai.

Se ora (x, y) è fissato e $(x_1, y_1) \to (x, y)$, si avrà anche $(x^*, y^*) \to (x, y) = (x, g(x))$ e, per la continuità delle derivate prime,

$$-\frac{f_x(x^*, y^*)}{f_y(x^*, y^*)} \to -\frac{f_x(x, g(x))}{f_y(x, g(x))}$$

da cui si ricava che esiste

$$g'(x) = \lim_{x_1 \to x} \frac{g(x_1) - g(x)}{x_1 - x} = -\frac{f_x(x, g(x))}{f_y(x, g(x))},$$

che è la formula desiderata per la derivata di g.

4. *Continuità di g'.* La (8.1) esprime g' come funzione composta di funzioni continue, per cui essa stessa è continua. ◆

Esempi

Consideriamo l'equazione

$$f(x, y) = e^{xy} + x - y - 1 = 0$$

© 978-88-08-06485-1

soddisfatta per esempio in $(0, 0)$. Poiché

$$f_y(x, y) = xe^{xy} - 1, f_y(0, 0) = -1 \neq 0$$

possiamo affermare che esiste un intorno I di 0 e una funzione $y = y(x)$ definita implicitamente dall'equazione, in I, con $y(0) = 0$. Derivando rispetto a x l'identità

$$f(x, y(x)) = 0$$

si ottiene

$$e^{xy}(y + xy') + 1 - y' = 0 \tag{8.3}$$

da cui

$$y'(x) = -\frac{y(x)e^x + 1}{xe^{xy(x)} - 1}.$$

Per esempio, $y'(0) = 1$. In questo caso la funzione $y = y(x)$ non si sa scrivere esplicitamente: è proprio questo il caso più interessante, che mostra la potenza del teorema. Derivando ancora ambo i membri della (8.3) (pensando $y = y(x)$), si trova:

$$e^{xy}\left((y + xy')^2 + 2y' + xy''\right) - y'' = 0$$

da cui, inserendo $x = 0, y(0) = 0, y'(0) = 1$, si ricava:

$$y''(0) = 2.$$

Iterando il procedimento, possiamo calcolare le derivate successive di $y(x)$ in $x = 0$ e, mediante lo sviluppo di MacLaurin, avere una buona approssimazione di $y(x)$, pur non conoscendone l'espressione esplicita.

Consideriamo l'equazione

$$f(x, y) = x^2 - y^2 - 1 = 0$$

soddisfatta, per esempio, nel punto $(1, 0)$. Si ha:

$$f_x(x, y) = 2x; \quad f_y(x, y) = -2y.$$

Poiché $f_x(x, y) = 2 \neq 0$, si può affermare che esiste un intorno J di $y_0 = 0$ e una funzione $x = h(y)$ definita implicitamente dall'equazione; inoltre

$$h'(y) = -\frac{-2y}{2x} = \frac{y}{x}.$$

In questo caso la funzione h si può scrivere esplicitamente:

$$h(y) = \sqrt{y^2 + 1}$$

(si noti che l'altro ramo della curva, $x = -\sqrt{y^2 + 1}$, non passa dal punto $(1, 0)$ che abbiamo fissato).

L'equazione

$$f(x, y) = x^2 - y^2 = 0 \tag{8.4}$$

è soddisfatta, per esempio, in $(0, 0)$. Poiché $\nabla f(0, 0) = \mathbf{0}$, non è possibile applicare il teorema in questo punto. Infatti, la (8.4) definisce le *due* funzioni implicite $y = \pm x$, e l'*unicità* della funzione implicita viene a cadere.

Analogamente, l'equazione

$$f(x, y) = x^2 + y^2 = 0$$

è soddisfatta in $(0, 0)$ ma, poiché $\nabla f(0, 0) = \mathbf{0}$, non è possibile applicare il teorema in questo punto. In questo caso l'equazione è soddisfatta solo nell'origine ed è l'*esistenza* della funzione implicita che viene a cadere.

 © 978-88-08-06485-1

Linee di livello

Torniamo ora al problema da cui siamo partiti all'inizio di questo paragrafo: descrivere gli insiemi di livello di una funzione $z = f(x, y)$, regolare. Sia

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

Calcoliamo $\nabla f(x, y)$ e cerchiamo i *punti critici* di f, ossia i punti in cui il gradiente si annulla. Supponiamo che questi siano in numero finito, o per lo meno che ogni insieme E_c contenga solo un numero finito di questi. Allora, salvo un numero finito di eccezioni, in ogni punto di E_c risulterà applicabile il teorema di Dini, ossia l'insieme E_c si potrà rappresentare, in un intorno di quei punti, come grafico di una funzione C^1 del tipo $y = g(x)$ oppure del tipo $x = h(y)$. I valori di c per cui E_c contiene almeno un punto critico di f si dicono *valori critici*. Se c non è un valore critico, nell'intorno di ogni suo punto l'insieme E_c è effettivamente una curva regolare. Questa è una giustificazione precisa del motivo per cui gli insiemi E_c sono generalmente delle linee (linee di livello).

8.2 Funzione implicita di n variabili

Tutti i discorsi fatti sulle funzioni definite implicitamente da un'equazione del tipo $f(x, y) = 0$ si possono generalizzare al caso di funzioni di n variabili. Per esempio, consideriamo l'equazione

$$f(x, y, z) = 0, \tag{8.5}$$

con f di classe C^1 e supponiamo che sia soddisfatta in un certo punto (x_0, y_0, z_0). Ci chiediamo se esistono un intorno U di (x_0, y_0) nel piano e una funzione $z = g(x, y)$, definita e regolare in U, per cui risulti

$$f(x, y, g(x, y)) = 0 \text{ per ogni } (x, y) \in U.$$

Diremo che g è definita implicitamente dalla (8.5). Se questo è vero, dovrà essere (derivando rispetto a x l'identità precedente):

$$f_x(x, y, g(x, y)) + f_z(x, y, g(x, y)) \cdot g_x(x, y) = 0$$

da cui

$$g_x(x, y) = -\frac{f_x(x, y, g(x, y))}{f_z(x, y, g(x, y))} \quad \text{e} \quad g_x(x_0, y_0) = -\frac{f_x(x_0, y_0, z_0)}{f_z(x_0, y_0, z_0)} \tag{8.6}$$

Analogamente

$$g_y(x, y) = -\frac{f_y(x, y, g(x, y))}{f_z(x, y, g(x, y))} \quad \text{e} \quad g_y(x_0, y_0) = -\frac{f_y(x_0, y_0, z_0)}{f_z(x_0, y_0, z_0)} \tag{8.7}$$

purché sia $f_z(x_0, y_0, z_0) \neq 0$. Si capisce allora come si generalizzi il teorema di Dini: se $f(x_0, y_0, z_0) = 0$ e $f_z(x_0, y_0, z_0) \neq 0$, esistono un intorno di (x_0, y_0) e un'unica funzione $z = g(x, y)$ definita implicitamente dall'equazione $f(x, y, z) = 0$, per cui valgono le (8.6)-(8.7). Se la derivata $f_z(x_0, y_0, z_0)$ si annulla, ma in compenso un'altra delle due non si annulla, si potrà esplicitare la variabile corrispondente, rispetto alle

© 978-88-08-06485-1

rimanenti due e si otterranno formule analoghe. Il lettore è invitato, per esercizio, a scrivere le formule che assegnano le due derivate parziali della funzione definita implicitamente, in ciascuno dei casi $f_x(x_0, y_0, z_0) \neq 0$, $f_y(x_0, y_0, z_0) \neq 0$.

L'enunciato preciso che vale in generale è il seguente e si può dimostrare sulla falsariga del teorema 3.26:

Teorema 3.27 (di Dini: caso n-dimensionale) *Sia A un aperto di $\mathbb{R}^{n+1}$, $f : A \to \mathbb{R}$, $f \in C^1(A)$ e supponiamo che*

$$f(x_0, y_0) = 0; \quad f_y(x_0, y_0) \neq 0$$

per un certo punto $(x_0, y_0) \in A$ (dove $x_0 \in \mathbb{R}^n$, $y_0 \in \mathbb{R}$). Allora esistono un intorno $U \subset \mathbb{R}^n$ di x_0 e un'unica funzione $g : U \to \mathbb{R}$, $g \in C^1(U)$, tale che

$$f(x, g(x)) = 0 \text{ per ogni } x \in U;$$
$$g_{x_j}(x) = -\frac{f_{x_j}(x, g(x))}{f_y(x, g(x))} \text{ per ogni } x \in U, j = 1, 2, \ldots, n.$$

Esercizi

Verificare che l'equazione

$$xe^y + ye^x = 0$$

definisce implicitamente una funzione $y = g(x)$ in un intorno di $x = 0$. Calcolare $g'(0)$ e $g''(0)$ e scrivere lo sviluppo di MacLaurin al secondo ordine per g.

Si provi che, per tutti i valori di c tranne due, l'insieme di livello E_c della funzione

$$f(x, y) = x^3 + y^3 - 3xy$$

risulta una curva regolare. Si determinino i due valori critici di c.

Verificare che l'equazione

$$e^{x-y} + x^2 - y^2 - e(x+1) - 1 = 0$$

definisce implicitamente $y = g(x)$ in un intorno di $x = 0$, con $f(0) = -1$. Provare che $x = 0$ è punto di minimo per f.

Verificare che l'equazione

$$\operatorname{arctg} z + xy^2 + xz - y^3 - 1 = 0$$

definisce implicitamente $z = g(x, y)$ in un intorno di $(0, -1, 0)$. Scrivere l'equazione del piano tangente al grafico di $z = g(x, y)$ in tale punto.

 © 978-88-08-06485-1

9 COMPLEMENTI

9.1 Topologia e funzioni continue

Connessione

Dimostrare che gli insiemi connessi in $\mathbb{R}$ sono esattamente gli intervalli. *Suggerimento*: utilizzare la definizione generale di intervallo:

"$I \subseteq \mathbb{R}$ si dice intervallo se $\forall x_1, x_2, x_3 \in \mathbb{R}$, se $x_1 < x_2 < x_3$ e $x_1, x_3 \in I$ allora anche $x_2 \in I$".

Applicare la definizione di insieme connesso, chiedendosi cos'è una curva continua in $\mathbb{R}$; applicare il teorema dei valori intermedi unidimensionale.

Dimostrare che:

1. Se $E \subseteq \mathbb{R}^n$ è un aperto connesso, allora non esistono due insiemi $A_1, A_2 \subseteq \mathbb{R}^n$ aperti, disgiunti e non vuoti, tali che $A_1 \cup A_2 = E$.

Suggerimento: per assurdo, esistano tali insiemi; sia $\mathrm{x}_1 \in A_1, \mathrm{x}_2 \in A_2$ e sia $\mathrm{r} : [a, b] \to \mathbb{R}^n$ un arco di curva continuo, contenuto in E, tale che $\mathrm{r}(a) = \mathrm{x}_1, \mathrm{r}(b) = \mathrm{x}_2$. Poniamo $t_0 = \sup\{t \in [a, b] : \mathrm{r}(t) \in A_1\}$. Mostrare che $\mathrm{r}(t_0) \in E$, ma che entrambe le possibilità $\mathrm{r}(t_0) \in A_1, \mathrm{r}(t_0) \in A_2$, portano a una contraddizione.

2. Gli unici sottoinsiemi di $\mathbb{R}^n$ che sono simultaneamente aperti e chiusi sono il vuoto e l'intero spazio $\mathbb{R}^n$. *Suggerimento*: dedurlo dal punto precedente.

Proprietà delle successioni in $\mathbb{R}^n$.

Nel volume 1 (complementi al capitolo 6) abbiamo incontrato il teorema di Bolzano-Weierstrass in $\mathbb{R}$, che afferma:

"Se $\{x_k\}_{k=1}^{\infty} \subset \mathbb{R}$ è una successione limitata, allora essa ammette una sottosuccessione convergente".

Un enunciato perfettamente analogo vale in $\mathbb{R}^n$:

TEOREMA 3.28 (DI BOLZANO-WEIERSTRASS IN $\mathbb{R}^n$) *Sia $\{\mathrm{x}_k\}_{k=1}^{\infty} \subset \mathbb{R}^n$ una successione limitata. Allora essa ammette una sottosuccessione convergente.*

Diamo qui la linea della dimostrazione nel caso $n = 2$; la generalizzazione al caso n qualunque è immediata. Sia dunque $\mathrm{x}_k = (x_k, y_k)$ una successione limitata in $\mathbb{R}^2$. Questo significa che le due successioni $\{x_k\}, \{y_k\} \subset \mathbb{R}$ sono entrambe limitate. Per il teorema di Bolzano-Weierstrass unidimensionale, da $\{x_k\}$ si può estrarre una sottosuccessione $\{x_{k_h}\}_{h=1}^{\infty}$ convergente; consideriamo la sottosuccessione $\{y_{k_h}\}_{h=1}^{\infty}$ corrispondente ai medesimi indici; anche questa, essendo limitata, avrà una sottosuccessione convergente, sia essa $\left\{y_{k_{h_i}}\right\}_{i=1}^{\infty}$. Consideriamo la corrispondente successione $\left\{x_{k_{h_i}}\right\}_{i=1}^{\infty}$: poiché questa successione è estratta da $\{x_{k_h}\}_{h=1}^{\infty}$ che già era convergente, sarà ancora convergente. Abbiamo quindi individuato una sottosuccessione convergente

$$\left\{\left(x_{k_{h_i}}, y_{k_{h_i}}\right)\right\}_{i=1}^{\infty}$$

estratta da $\{(x_k, y_k)\}_{k=1}^{\infty}$ e il teorema è dimostrato. (Si rifletta sul motivo per cui non è stato possibile estrarre "in un colpo solo" la sottosuccessione convergente di $\{(x_k, y_k)\}_{k=1}^{\infty}$, applicando simultaneamente il teorema di Bolzano-Weierstrass alle successioni $\{x_k\}$, $\{y_k\}$. Il punto delicato della dimostrazione è tutto qui).

© 978-88-08-06485-1

Questo teorema serve tra l'altro a provare il prossimo importante risultato.

Premettiamo la

DEFINIZIONE 3.22 (SUCCESSIONI DI CAUCHY IN $\mathbb{R}^n$) Sia $\{\mathrm{x}_k\}_{k=1}^{\infty}$ una successione in $\mathbb{R}^n$. Si dice che la successione soddisfa la condizione di Cauchy (o semplicemente che "è una successione di Cauchy") se

$$\forall \varepsilon > 0\ \exists n_0 : \forall h, k \geq n_0 \text{ si ha } |\mathrm{x}_h - \mathrm{x}_k| < \varepsilon.$$

(Più brevemente: $\forall \varepsilon > 0$ si ha $|\mathrm{x}_h - \mathrm{x}_k| < \varepsilon$ definitivamente).

Vale in proposito il:

TEOREMA 3.29 (COMPLETEZZA DI $\mathbb{R}^n$) *Se $\{\mathrm{x}_k\}_{k=1}^{\infty}$ è una successione di Cauchy in $\mathbb{R}^n$, allora converge.*

Questo risultato è noto come *completezza di* $\mathbb{R}^n$.

Per esercizio, si dimostri il teorema precedente. (Suggerimento: utilizzare il teorema di Bolzano-Weierstrass sopra enunciato e seguire la stessa linea dimostrativa usata per provare questo risultato per le successioni di numeri reali[28]).

Proprietà topologiche delle funzioni continue

Oltre al teorema di Weierstrass e al teorema degli zeri, di cui ci siamo occupati nel paragrafo 3.2, il terzo importante risultato per le funzioni continue di più variabili è, come già nel caso unidimensionale, il teorema di Cantor-Heine sull'uniforme continuità[29]. Questo risultato gioca un ruolo importante, per esempio, nella teoria dell'integrazione in più variabili (v. capitolo 5).

DEFINIZIONE 3.23 (UNIFORME CONTINUITÀ) Si dice che $f : \Omega \to \mathbb{R}$ è uniformemente continua in $\Omega \subseteq \mathbb{R}^n$ se:

$$\forall \varepsilon > 0\ \exists \delta > 0 : \forall \mathrm{x}_1, \mathrm{x}_2 \in \Omega, \text{ se } |\mathrm{x}_1 - \mathrm{x}_2| < \delta \text{ allora } |f(\mathrm{x}_1) - f(\mathrm{x}_2)| < \varepsilon.$$

La continuità uniforme implica la continuità, ma è una condizione più forte, in generale. Vale tuttavia, anche per le funzioni di più variabili, il seguente importante

TEOREMA 3.30 (DI CANTOR-HEINE) *Sia $K \subset \mathbb{R}^n$ un insieme chiuso e limitato e $f : K \to \mathbb{R}$ una funzione continua. Allora f è uniformemente continua in K.*

In altre parole, una funzione continua può violare la condizione di uniforme continuità solo se il suo insieme di definizione non è chiuso oppure non è limitato.

La dimostrazione di questo teorema si può condurre sulla falsariga di quella dell'analogo teorema per le funzioni di una variabile, sfruttando ora il teorema di Bolzano-Weierstrass in $\mathbb{R}^n$ (vedi sopra).

Vale anche il prossimo risultato:

[28] v. volume 1, capitolo 6, complementi.

[29] Per le funzioni di una variabile abbiamo incontrato questo teorema nel volume 1, capitolo 6, complementi.

 © 978-88-08-06485-1

TEOREMA 3.31 *Sia $\Omega \subset \mathbb{R}^n$ un aperto e sia $f : \Omega \to \mathbb{R}$ uniformemente continua in Ω. Allora f è prolungabile con continuità fino alla frontiera di Ω, ossia esiste una funzione $\tilde{f} : \overline{\Omega} \to \mathbb{R}$ continua in $\overline{\Omega}$ e tale che in Ω coincide con f.*

Esempio

Dimostrare che la funzione $f(x, y) = x^y$ è continua ma non uniformemente continua in $(0, 1) \times (0, 1)$.

9.2 Funzioni omogenee

In questo paragrafo ci occupiamo brevemente di una particolare classe di funzioni per le quali è possibile stabilire dei criteri semplici per studiare la continuità e la differenziabilità in un punto, anche in presenza di forme di indeterminazione, che non richiedano l'utilizzo esplicito della definizione di limite in più variabili.

DEFINIZIONE 3.24 Una funzione $f : \mathbb{R}^n \to \mathbb{R}$ (eventualmente definita solo per $x \neq 0$), non identicamente nulla, si dice *positivamente omogenea di grado* $\alpha \in \mathbb{R}$ se

$$f(\lambda x) = \lambda^\alpha f(x) \text{ per ogni } x \in \mathbb{R}^n, x \neq 0, \lambda > 0; \tag{9.1}$$

si dice *omogenea di grado* α se la (9.1) vale anche per $\lambda < 0$; in tal caso però ci sono delle ovvie restrizioni ai valori ammissibili di α; a noi interesseranno solo funzioni omogenee di grado α intero.

Notiamo che, se f è omogenea, è anche positivamente omogenea; il viceversa non è vero, come mostreranno gli esempi.

Esempio

- $f(x, y) = \dfrac{x^4 + 3x^3y}{x^2 + y^2}$ è omogenea di grado 2;
- $f(x, y) = \sqrt[3]{3x - 5y}$ è omogenea di grado $1/3$;
- $f(x, y) = \dfrac{xy}{x^2 + y^2}$ è omogenea di grado 0;
- $f(x, y) = \sqrt{x^2 + y^2}$ è positivamente omogenea di grado 1 (ma non è omogenea);
- $f(x, y) = \dfrac{xy}{(x^2 + y^2)^5}$ è omogenea di grado -8.

Naturalmente la maggior parte delle funzioni, come per esempio

$$xe^y;\ x + y^2;\ \sin(xy)$$

non è omogenea di nessun grado.

Tra le funzioni omogenee, rivestono una certa importanza i *polinomi omogenei*: un polinomio (in qualsiasi numero di variabili) è una funzione omogenea di grado α (intero positivo) se e solo se tutti i suoi termini (monomi) hanno grado complessivo α. Per esempio:

$x^3yz + 4x^2z^3$ è un polinomio omogeneo di grado 5 (in 3 variabili).

© 978-88-08-06485-1

Se una funzione f è positivamente omogenea, i suoi valori sono completamente determinati da quelli che f assume nei punti x di modulo unitario. Infatti, dall'identità

$$f(x) = f\left(|x| \cdot \frac{x}{|x|}\right) = |x|^\alpha f\left(\frac{x}{|x|}\right) \tag{9.2}$$

leggiamo che, per conoscere il valore di $f(x)$, è sufficiente conoscere il valore di $f(x/|x|)$; ma $x/|x|$ è un punto di modulo unitario e questo significa appunto che i valori di f nei punti di modulo unitario determinano tutti gli altri.

In particolare, se una funzione è omogenea (o positivamente omogenea) di grado zero, significa che è costante su ogni retta (rispettivamente, semiretta) uscente dall'origine. Infatti, indicata con

$$r(t) = tv$$

l'equazione parametrica della retta, con v versore fissato e $t \in \mathbb{R}$, sarà:

$$f(r(t)) = f(tv) = t^0 f(v) = f(v) = \text{costante}$$

per ogni $t \in \mathbb{R}$ se la funzione è omogenea (e in tal caso la funzione è costante sull'intera retta), per ogni $t > 0$ se la funzione è positivamente omogenea (e in tal caso la funzione è costante sulla semiretta).

Esempio

Riprendiamo l'esempio $f(x, y) = \sqrt[3]{3x - 5y}$. Introducendo le coodinate polari (ρ, θ), si ha:

$$f(\rho, \theta) = \rho^{1/3} \cdot \sqrt[3]{3\cos\theta - 5\sin\theta}.$$

La funzione $g(\theta) = \sqrt[3]{3\cos\theta - 5\sin\theta}$ coincide con la funzione f se $\rho = 1$ e determina i valori di f in tutto il piano, una volta noto che f è omogenea di grado $1/3$. Se, per esempio, sapessimo che un'altra funzione h coincide con f per $\rho = 1$ ed è omogenea di grado -2, avremmo necessariamente:

$$h(\rho, \theta) = \rho^{-2} \cdot \sqrt[3]{3\cos\theta - 5\sin\theta}.$$

Più in generale, per una generica funzione di due variabili positivamente omogenea di grado α vale la seguente rappresentazione in coordinate polari:

$$f(\rho, \theta) = \rho^\alpha f(1, \theta).$$

In altri termini, *le funzioni omogenee di due variabili sono tutte e sole quelle che, in coordinate polari, si scrivono nella forma*

$$f(\rho, \theta) = \rho^\alpha g(\theta)$$

per qualche $\alpha \in \mathbb{R}$ e qualche funzione

$$g : [0, 2\pi) \to \mathbb{R}.$$

Le funzioni omogenee sono più semplici da studiare, rispetto alle altre, dal punto di vista della continuità e della differenziabilità. Vale infatti il seguente:

 © 978-88-08-06485-1

TEOREMA 3.32 *Sia $f : \mathbb{R}^n \to \mathbb{R}$ una funzione positivamente omogenea di grado α, definita e continua per $x \neq 0$. Allora:*

1. *f è continua anche nell'origine se $\alpha > 0$; in questo caso $f(0) = 0$; f è discontinua nell'origine se $\alpha < 0$; è discontinua anche se $\alpha = 0$, tranne il caso banale in cui f è costante.*

2. *f è differenziabile nell'origine se $\alpha > 1$; non è differenziabile nell'origine se $\alpha < 1$, tranne il caso banale in cui $\alpha = 0$ e f è costante; se $\alpha = 1$, f è differenziabile se e solo se è una funzione lineare, (ossia $f(x) = a \cdot x$ per qualche vettore costante $a \in \mathbb{R}^n$).*

DIMOSTRAZIONE. 1. Proviamo che se $\alpha > 0$ risulta

$$\lim_{x \to 0} f(x) = 0.$$

Riscriviamo l'identità già vista:

$$f(x) = f\left(|x| \cdot \frac{x}{|x|}\right) = |x|^\alpha f\left(\frac{x}{|x|}\right). \tag{9.3}$$

Ora: al variare di $x \in \mathbb{R}^n, x \neq 0$, il punto $\frac{x}{|x|}$ varia nell'insieme chiuso e limitato

$$C = \{y \in \mathbb{R}^n : |y| = 1\};$$

poiché f è continua in C, per il teorema di Weierstrass (v. teorema 3.6 paragrafo 3.2) ammette massimo e minimo; in particolare quindi esiste una costante positiva c tale che

$$\left|f\left(\frac{x}{|x|}\right)\right| \leq c \ \forall x \neq 0.$$

Dalla (9.3) segue allora:

$$|f(x)| \leq c|x|^\alpha,$$

e questo mostra che $f(x) \to 0$ per $x \to 0$ (si veda il teorema 3.2 nel paragrafo 2.2).

Sia ora $\alpha < 0$. Sia x_0 un punto qualsiasi in cui $f(x_0) \neq 0$. Allora

$$f(tx_0) = t^\alpha f(x_0) \to +\infty \ (\text{o } -\infty) \text{ per } t \to 0^+,$$

perciò f non può essere continua.

Se infine $\alpha = 0$, per ogni versore v si ha:

$$f(tv) = f(v), \text{ da cui}$$
$$\lim_{t \to 0^+} f(tv) = f(v).$$

L'unica possibilità affinché esista $\lim_{x \to 0} f(x)$ è allora che $f(v)$ sia costante al variare del versore v; ma questo significa che f è la funzione costante.

2. Chiediamoci anzitutto se esistono le derivate parziali nell'origine. Per calcolare $f_{x_i}(0)$, consideriamo

$$g(t) = f(0 + te_i) = \begin{cases} t^\alpha f(e_i) & \text{se } t > 0 \\ |t|^\alpha f(-e_i) & \text{se } t < 0. \end{cases}$$

© 978-88-08-06485-1

Allora, se $\alpha > 1$, esiste $g'(0) = 0$. Perciò, se $\alpha > 1$, la funzione f è derivabile in 0 con gradiente nullo. Poiché in questo caso è anche $f(0) = 0$, provare la differenziabilità equivale a provare che

$$f(x) = o(x) \text{ per } x \to 0, \text{ ossia che}$$
$$\frac{f(x)}{|x|} \to 0 \text{ per } x \to 0.$$

Ma se f è positivamente omogenea di grado $\alpha > 1$, $\frac{f(x)}{|x|}$ sarà positivamente omogenea di grado $\alpha - 1 > 0$ (verificarlo dalla definizione!), perciò per il punto precedente del teorema $\frac{f(x)}{|x|} \to 0$ per $x \to 0$.

Se $0 < \alpha < 1$, per il calcolo fatto sopra, f non è derivabile nell'origine (e quindi a maggior ragione non è differenziabile), a meno che sia $f(\pm e_i) = 0$ per $i = 1, 2, \ldots, n$. In questo caso la funzione è identicamente nulla sugli assi, perciò ha derivate parziali nulle nell'origine.

In questa situazione, affinché f sia differenziabile, dovrebbe essere, ancora, $\frac{f(x)}{|x|} \to 0$ per $x \to 0$; e poiché $\frac{f(x)}{|x|}$ è positivamente omogenea di grado $\alpha - 1 < 0$, per il primo punto del teorema questo non accade, perciò f non è differenziabile.

Se poi $\alpha \leq 0$, f non è differenziabile perché è discontinua, tranne il caso banale in cui f è costante, che abbiamo escluso.

Sia infine $\alpha = 1$. Supponiamo prima che f sia omogenea (e non solo positivamente omogenea) di grado 1; allora:

$$g(t) = f(0 + te_i) = tf(e_i) \ \forall t \in \mathbb{R} \text{ perciò}$$
$$g'(0) = f(e_i) \text{ ossia esiste } f_{x_i}(0) = f(e_i).$$

Per provare che f è differenziabile in 0 occorre allora provare che:

$$f(x) - \sum_{i=1}^{n} f(e_i) x_i = o(x) \text{ per } x \to 0.$$

Ma la funzione $f(x) - \sum_{i=1}^{n} f(e_i) x_i$ è omogenea di grado 1, essendo differenza tra due funzioni omogenee di grado 1, quindi l'unica possibilità affinchè risulti $o(x)$ è che sia identicamente nulla. Ma allora

$$f(x) = \sum_{i=1}^{n} f(e_i) x_i$$

cioè f è una funzione lineare (ovviamente differenziabile).

Se poi f è positivamente omogenea di grado 1, ma non è omogenea di grado 1, significa che, per qualche versore v, risulta:

$$f(v) \neq f(-v) \text{ e quindi}$$
$$f(tv) \neq f(-tv) \text{ per ogni } t > 0.$$

Questo implica che non esiste la derivata direzionale $D_v f(0)$, pertanto f non è differenziabile nell'origine. ♦

 © 978-88-08-06485-1

Esempio

93 $f(x,y) = \dfrac{x^4+3x^3y}{x^2+y^2}$ è omogenea di grado 2:

f è differenziabile in 0.

$f(x,y) = \sqrt[3]{3x-5y}$ è omogenea di grado 1/3:

f è continua ma non differenziabile in 0.

$f(x,y) = \dfrac{xy}{x^2+y^2}$ è omogenea di grado 0 e non costante:

f è discontinua in 0 (ma limitata).

$f(x,y) = \sqrt{x^2+y^2}$ è positivamente omogenea di grado 1 (ma non omogenea e non lineare):

f è continua ma non differenziabile in 0.

$f(x,y) = \dfrac{xy}{(x^2+y^2)^5}$ è omogenea di grado -8:

f è discontinua in 0 (e illimitata in un intorno di 0).

Come si vede, il teorema precedente è uno strumento potente per decidere la continuità o la differenziabilità di una funzione omogenea, senza dover ogni volta studiare esplicitamente un limite in più variabili.

Inoltre, per una funzione che pur non essendo omogenea sia somma di funzioni omogenee di gradi diversi, si possono ripetere analoghi ragionamenti:

Esempio

94 Sia

$$f(x,y) = \begin{cases} \dfrac{x^2}{\sqrt{x^2+y^2}} + \dfrac{x^4+3x^3y}{x^2+y^2} & \text{per } (x,y) \neq (0,0) \\ 0 & \text{per } (x,y) = (0,0). \end{cases}$$

La funzione è somma di un termine positivamente omogeneo di grado 1 (ma non lineare) e un termine omogeneo di grado 2, entrambi continui fuori dall'origine. In base al teorema visto, il primo addendo è una funzione continua ma non differenziabile in $(0,0)$, il secondo è una funzione differenziabile in $(0,0)$. Ne segue che f è una funzione continua ma non differenziabile in $(0,0)$.

Esercizi

94 Dimostrare che se $f \in C^1(\mathbb{R}^n)$ è una funzione positivamente omogenea di grado α, allora f_{x_i} è positivamente omogenea di grado $\alpha-1$. (Suggerimento: derivare ambo i membri della (9.1) rispetto a x_i).

95 Dimostrare che se $f \in C^1(\mathbb{R}^n)$ è una funzione positivamente omogenea di grado α, allora f soddisfa l'*equazione di Eulero*:

$$\nabla f(\mathrm{x}) \cdot \mathrm{x} = \alpha f(\mathrm{x}).$$

Suggerimento: derivare ambo i membri della (9.1) rispetto a λ.

© 978-88-08-06485-1

96 Provare che una *funzione* $f : \mathbb{R}^2 \to \mathbb{R}$ *omogenea di grado zero* si può scrivere anche nella forma

$$f(x,y) = h\left(\frac{y}{x}\right)$$

con $h(t) : \mathbb{R} \to \mathbb{R}$. Scrivere l'espressione del gradiente di f in funzione di h, nell'ipotesi che h sia derivabile. Nel caso particolare in cui $h(t) = \frac{t}{1+t^2}$, calcolare quindi il gradiente di f:
a) applicando la formula appena trovata;
b) scrivendo in funzione di x, y la funzione f e calcolandone direttamente il gradiente.
(Ovviamente, i risultati devono coincidere).

97 Dimostrare il seguente fatto, che si è utilizzato nella dimostrazione del teorema 3.32: se una funzione omogenea di grado 1 è $o(\mathrm{x})$ per $\mathrm{x} \to 0$, allora è identicamente nulla.

9.3 Differenziali e formula di Taylor di ordine superiore

Per comprendere che forma assume la formula di Taylor di ordine superiore al secondo, per una funzione di n variabili sufficientemente regolare, il modo più semplice è riprendere la dimostrazione del teorema 3.15 e proseguire oltre. Con le stesse notazioni di quella dimostrazione, se scriviamo per esempio lo sviluppo di Taylor al 3° ordine per la funzione di una variabile

$$g(t) = f(\mathrm{x}_0 + t\mathrm{h}) \text{ per } t \in (0,1)$$

(con resto secondo Lagrange), abbiamo:

$$g(1) = g(0) + tg'(0) + \frac{t^2}{2}g''(0) + \frac{t^3}{3!}g'''(\delta)$$

per un opportuno $\delta \in (0,1)$. Ora,

$$g'''(t) = \sum_{i,j,k=1}^{n} \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}(\mathrm{x}_0 + t\mathrm{h})\, h_i h_j h_k.$$

Ragionando in modo perfettamente analogo a quanto abbiamo fatto nel caso di ordine 2, si ottiene il seguente risultato (si invita il lettore a ripercorrere per proprio conto il ragionamento relativo):

TEOREMA 3.33 (FORMULA DI TAYLOR DI ORDINE m, CON RESTO SECONDO PEANO) *Sia $f \in C^m(A)$, $A \subseteq \mathbb{R}^n$ aperto; allora per ogni $\mathrm{x}_0 \in A$ vale la formula:*

(9.4)
$$f(\mathrm{x}_0+\mathrm{h}) = f(\mathrm{x}_0) + df(\mathrm{x}_0) + \frac{1}{2}d^2f(\mathrm{x}_0) + \frac{1}{3!}d^3f(\mathrm{x}_0) + \ldots + \frac{1}{m!}d^mf(\mathrm{x}_0) + o(|\mathrm{h}|^m)$$
per $\mathrm{h} \to 0$.

Nella (9.4), il simbolo $d^kf(\mathrm{x}_0)$ $(k = 1, 2, \ldots, m)$ denota il *differenziale di ordine k di f, calcolato in x_0* e *applicato all'incremento* h, dato per definizione da:

$$d^kf(\mathrm{x}_0) = \sum_{i_1,i_2,\ldots,i_k=1}^{n} \frac{\partial^k f}{\partial x_{i_1}\partial x_{i_2}\ldots\partial x_{i_k}}(\mathrm{x}_0)\, h_{i_1}h_{i_2}\ldots h_{i_k}.$$

 © 978-88-08-06485-1

Ora, l'ultima espressione scritta per il differenziale k-esimo è poco pratica da utilizzare, perché richiederebbe di scrivere distintamente tutte le derivate miste (ad es. $f_{xxy}, f_{yxx}, f_{xyx}$) e a posteriori contare quante sono quelle uguali, per raggrupparle. Conviene quindi stabilire una volta per tutte una formula più sintetica. Ricordiamo, a questo scopo, la definizione di *multiindice* data nel paragrafo 5.1; per

$$\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n) \text{ con } \alpha_i \text{ interi non negativi,}$$
$$|\alpha| = \alpha_1 + \alpha_2 + \dots + \alpha_n,$$

poniamo anche

$$\alpha! = \alpha_1! \alpha_2! \dots \alpha_n!$$

e, per ogni $\mathrm{x} = (x_1, x_2, \dots, x_n) \in \mathbb{R}^n$,

$$\mathrm{x}^\alpha = x_1^{\alpha_1} x_2^{\alpha_2} \dots x_n^{\alpha_n}.$$

Con queste notazioni, enunciamo anzitutto il

LEMMA 3.34 (SVILUPPO MULTINOMIALE) *Per ogni $x_1, x_2, \dots, x_n \in \mathbb{R}$, k intero positivo, si ha:*

$$(x_1 + x_2 + \dots + x_n)^k = k! \sum_{|\alpha|=k} \frac{x^\alpha}{\alpha!}.$$

Il lettore è invitato a dimostrare questo lemma per esercizio. (Suggerimento: procedere per induzione su n; per $n = 2$ si riconosce la formula del binomio di Newton; il passo induttivo richiede qualche calcolo con le sommatorie e l'uso delle definizioni precedenti).

Riscriviamo ora il differenziale k-esimo come segue:

$$d^k f(\mathrm{x}_0) = \sum_{i_1, i_2, \dots, i_k = 1}^{n} \frac{\partial^k f}{\partial x_{i_1} \partial x_{i_2} \dots \partial x_{i_k}} (\mathrm{x}_0)\, h_{i_1} h_{i_2} \dots h_{i_k}$$

$$= \left(\frac{\partial}{\partial x_1} h_1 + \frac{\partial}{\partial x_2} h_2 + \dots + \frac{\partial}{\partial x_n} h_n \right)^k f(\mathrm{x}_0) \tag{9.5}$$

dove la potenza k-esima va intesa come la composizione (eseguita k volte) dell'operatore

$$\frac{\partial}{\partial x_1} h_1 + \frac{\partial}{\partial x_2} h_2 + \dots + \frac{\partial}{\partial x_n} h_n$$

(che associa a f il differenziale $\frac{\partial f}{\partial x_1} h_1 + \frac{\partial f}{\partial x_2} h_2 + \dots + \frac{\partial f}{\partial x_n} h_n$).

Esempio

Per una funzione $f : \mathbb{R}^2 \to \mathbb{R}$ il differenziale terzo è:

$$d^3 f(\mathrm{x}_0) = \left(\frac{\partial}{\partial x_1} h_1 + \frac{\partial}{\partial x_2} h_2 \right)^3 f(\mathrm{x}_0) =$$
$$= \left(\frac{\partial^3}{\partial x_1^3} h_1^3 + 3 \frac{\partial^3}{\partial x_1^2 \partial x_2} h_1^2 h_2 + 3 \frac{\partial^3}{\partial x_1 \partial x_2^2} h_1 h_2^2 + \frac{\partial^3}{\partial x_2^3} h_2^3 \right) f(\mathrm{x}_0)$$
$$= \frac{\partial^3 f}{\partial x_1^3}(\mathrm{x}_0)\, h_1^3 + 3 \frac{\partial^3 f}{\partial x_1^2 \partial x_1}(\mathrm{x}_0)\, h_1^2 h_2 + 3 \frac{\partial^3 f}{\partial x_1 \partial x_1^2}(\mathrm{x}_0)\, h_1 h_2^2 + \frac{\partial^3 f}{\partial x_2^3}(\mathrm{x}_0)\, h_2^3.$$

© 978-88-08-06485-1

In generale, l'espressione (9.5) si può riscrivere, in base allo sviluppo multinomiale:

$$d^k f(\mathrm{x}_0) = k! \sum_{|\alpha|=k} \frac{D^\alpha f(\mathrm{x}_0)}{\alpha!} \mathrm{h}^\alpha$$

e la formula di Taylor si riscrive più sinteticamente così:

$$f(\mathrm{x}_0 + \mathrm{h}) = \sum_{k=0}^{m} \left(\sum_{|\alpha|=k} \frac{D^\alpha f(\mathrm{x}_0)}{\alpha!} \mathrm{h}^\alpha \right) + o(|\mathrm{h}|^m) \text{ per } \mathrm{h} \to \mathbf{0}.$$

4 Calcolo differenziale per funzioni di più variabili a valori vettoriali

1 FUNZIONI DI PIÙ VARIABILI A VALORI VETTORIALI: GENERALITÀ

In questo capitolo consideriamo *funzioni di più variabili a valori vettoriali*, ossia funzioni

$$f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m.$$

Ci sono molti esempi di oggetti fisici, geometrici ecc. rappresentati da questi oggetti matematici. In particolare, alcune classi di esempi di questo tipo che incontreremo sono i seguenti:

- superfici in forma parametrica nello spazio $\mathbb{R}^3$ (v. paragrafo 3) e più in generale varietà k-dimensionali in $\mathbb{R}^n$ (v. paragrafo 4);
- trasformazioni di coordinate in $\mathbb{R}^n$ (v. paragrafo 5);
- campi vettoriali (di cui ci occuperemo nel capitolo 6, nel contesto delle questioni legate all'integrazione su curve e superfici).

Lo scopo di questo primo paragrafo è fare una carrellata di esempi introduttivi per illustrare la varietà di oggetti che si possono descrivere mediante funzioni di più variabili a valori vettoriali.

1.1 Superfici in forma parametrica

Abbiamo visto (capitolo 2) che una *curva piana in forma parametrica* è una funzione che descrive una linea più generale di quella che può essere descritta come grafico di una funzione di una variabile. Per esempio, la circonferenza è rappresentabile come curva in forma parametrica, ma non come grafico di funzione.

Analogamente, il grafico di una funzione $z = f(x, y)$ è una superficie in $\mathbb{R}^3$, ma non ogni superficie "ragionevole" può essere grafico di una funzione di due variabili. Per esempio, una sfera non è il grafico di una funzione di due variabili (perché uno stesso punto (x, y) individuerebbe due punti aventi ordinate z diverse, contro la definizione

di funzione). Per rappresentare una superficie come una sfera, un ellissoide, ecc., si può utilizzare una rappresentazione *implicita*, come

$$x^2 + y^2 + z^2 = 1,$$

oppure una *rappresentazione parametrica*, analogamente a quanto accade per le curve. Le tre coordinate del punto mobile sulla superficie dipenderanno questa volta da *due parametri* (u, v), coerentemente al fatto che un punto vincolato a muoversi su una superficie assegnata ha due gradi di libertà (può scegliere, in ogni punto, tra due direzioni ortogonali di moto).

L'oggetto matematico che otteniamo, detto *superficie in forma parametrica*, è quindi una funzione:

$$\mathrm{r} : A \subseteq \mathbb{R}^2 \to \mathbb{R}^3$$

con $\mathrm{r} = (x, y, z)$ e

$$\begin{cases} x = x(u, v) \\ y = y(u, v) \\ z = z(u, v) \end{cases} \qquad (u, v) \in A$$

dove A è un'opportuna regione del piano $\mathbb{R}^2$.

Esempio

La sfera di raggio R e centro l'origine ha rappresentazione parametrica:

$$\begin{cases} x = R\sin\varphi\cos\theta \\ y = R\sin\varphi\sin\theta \\ z = R\cos\varphi \end{cases} \quad \text{con} \quad \varphi \in [0, \pi],\ \theta \in [0, 2\pi).$$

Si noti che R è fissato, mentre i due parametri (φ, θ) sono corrispondenti, rispettivamente, alla *colatitudine* e alla *longitudine* (fig. 4.1).

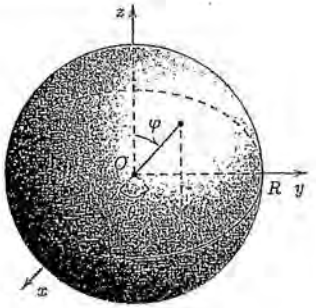


Figura 4.1. La superficie sferica di raggio R e centro l'origine.

Vedremo in seguito come si scrivono le equazioni parametriche di altre superfici notevoli. Al solito, è fondamentale precisare l'insieme in cui variano i parametri. Per esempio, le stesse equazioni parametriche dell'esempio precedente, con $\varphi \in [0, \pi/4]$, $\theta \in [0, 2\pi)$ rappresenterebbero una calotta sferica, mentre con le condizioni $\varphi \in [0, \pi]$, $\theta \in [0, \pi/6]$ rappresenterebbero uno "spicchio" di sfera.

Osservazione Il lettore potrebbe chiedersi perché ci sia bisogno di usare le "complicate" equazioni parametriche per rappresentare un oggetto come una sfera, anziché la semplice equazione $x^2 + y^2 + z^2 = R^2$. Il seguito del discorso mostrerà vari utilizzi delle formule parametriche, ma fin d'ora si può osservare che le equazioni parametriche contengono un insieme di *istruzioni esplicite* con cui per esempio un computer può tracciare la superficie, unendo i punti (x, y, z) che corrispondono a un insieme abbastanza fitto di coppie (u, v). Per contro, non esiste alcun metodo semplice e *generale* per determinare il grafico di una superficie assegnata nella forma implicita $f(x, y, z) = 0$.

Le *superfici in forma parametrica*, che tratteremo nel paragrafo 3, sono solo un caso particolare del concetto più astratto e generale di *varietà k-dimensionale in forma parametrica in* $\mathbb{R}^n$, di cui ci occuperemo nel paragrafo 4.

1.2 Trasformazioni di coordinate

Un esempio importante di funzione $\mathrm{f} : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ (cioè con dominio e codominio della stessa dimensione) è dato dalle *trasformazioni di coordinate* in $\mathbb{R}^n$. Gli esempi tipici in $\mathbb{R}^2$ e $\mathbb{R}^3$, che saranno utili successivamente, sono i seguenti.

Coordinate polari nel piano

Il punto $(x, y) \in \mathbb{R}^2$ può essere individuato anche assegnandone le *coordinate polari* (ρ, θ), che abbiamo già utilizzato in precedenza. Ricordiamo che:

$$\rho = \sqrt{x^2 + y^2}$$

rappresenta la distanza di (x, y) dall'origine e θ è l'angolo formato tra il vettore (x, y) e l'asse x, orientato positivamente. Perciò si ha:

$$\begin{cases} x = \rho\cos\theta \\ y = \rho\sin\theta \end{cases} \quad \text{con } \rho \in [0, +\infty),\ \theta \in [0, 2\pi)$$

(oppure $\theta \in [-\pi, \pi)$, o qualche altro intervallo di ampiezza 2π). Questa trasformazione di coordinate si può vedere come una funzione $\mathrm{f} : \mathbb{R}^2 \to \mathbb{R}^2$, $(x, y) = \mathrm{f}(\rho, \theta)$. Le coordinate polari sono utili per rappresentare insiemi o funzioni aventi qualche simmetria rispetto all'origine.

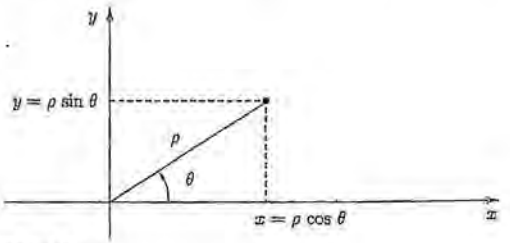


Figura 4.2. Coordinate polari.

Esempi

ESEMPIO L'insieme definito in coordinate polari da:

$$E = \left\{(\rho, \theta) : \rho < 1;\ \frac{\pi}{4} < \theta < \frac{\pi}{2}\right\}$$

è illustrato in figura 4.3.

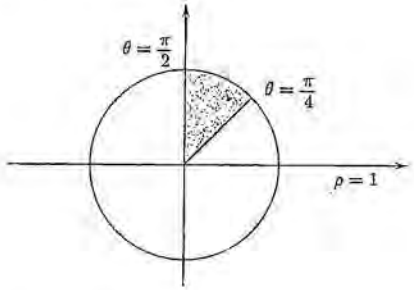


Figura 4.3.

ESEMPIO Si voglia rappresentare in coordinate polari l'insieme ottenuto intersecando un settore circolare di angolo $\frac{\pi}{2}$ e raggio 2 con la corona circolare concentrica di raggio interno ed esterno, rispettivamente, uguale a 1 e 3.

Il settore circolare è definito da $\{\rho \leq 2,\ 0 \leq \theta \leq \frac{\pi}{2}\}$; la corona circolare è definita da $\{1 \leq \rho \leq 3,\ 0 \leq \theta < 2\pi\}$, la loro intersezione è definita quindi da $\{1 \leq \rho \leq 2,\ 0 \leq \theta \leq \frac{\pi}{2}\}$.

Coordinate cilindriche nello spazio

Per descrivere insiemi o funzioni in $\mathbb{R}^3$ che abbiano qualche simmetria rispetto all'asse z sono utili le *coordinate cilindriche* (ρ, θ, t), per le quali si ha:

$$\begin{cases} x = \rho\cos\theta \\ y = \rho\sin\theta \quad \text{con } \rho \in [0, +\infty),\ \theta \in [0, 2\pi),\ t \in \mathbb{R}. \\ z = t \end{cases}$$

Questa trasformazione di coordinate (fig. 4.4) si può vedere come una funzione $\mathrm{f} : \mathbb{R}^3 \to \mathbb{R}^3$.

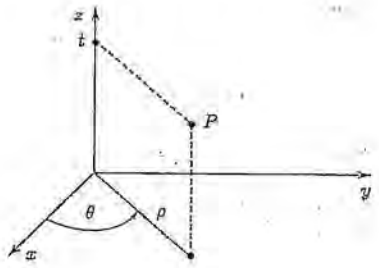


Figura 4.4. Coordinate cilindriche.

Coordinate sferiche nello spazio

Per descrivere insiemi o funzioni in $\mathbb{R}^3$ che abbiano qualche simmetria rispetto all'origine, sono utili le *coordinate sferiche* (ρ, θ, φ), per le quali si ha:

$$\begin{cases} x = \rho\sin\varphi\cos\theta \\ y = \rho\sin\varphi\sin\theta \quad \text{con } \rho > 0,\ \varphi \in [0, \pi],\ \theta \in [0, 2\pi). \\ z = \rho\cos\varphi \end{cases}$$

Anche queste si possono vedere come una funzione $\mathrm{f} : \mathbb{R}^3 \to \mathbb{R}^3$. Si noti che la funzione $\mathrm{g} : \mathbb{R}^2 \to \mathbb{R}^3$ che si ottiene da questa fissando il valore della prima variabile ρ rappresenta la superficie della sfera di raggio ρ, in forma parametrica (fig. 4.5).

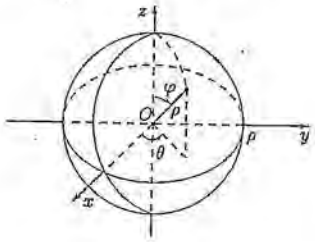


Figura 4.5. Coordinate sferiche.

Tratteremo le trasformazioni di coordinate in $\mathbb{R}^n$ in generale nel paragrafo 5, affrontando in particolare il problema della loro invertibilità.

1.3 Campi vettoriali

Esempio

ESEMPIO Il moto di un fluido (per esempio l'acqua in un canale) può essere descritto assegnando, istante per istante e punto per punto, la velocità della "particella d'acqua" che si trova in quell'istante in quel punto. La velocità, d'altro canto, è un vettore a tre componenti, perciò sarà una funzione:

$$\mathrm{v}(t, x, y, z) : \mathbb{R}^4 \to \mathbb{R}^3.$$

Il moto si dice *stazionario* se v non dipende da t; si dice *piano* se v ha solo le prime due componenti non nulle e il punto da cui dipende è specificato dalle due sole componenti (x, y). Un moto stazionario piano è quindi descritto da una funzione:

$$\mathrm{v}(x, y) : \mathbb{R}^2 \to \mathbb{R}^2.$$

In questo caso (campo bidimensionale di due variabili) è possibile visualizzare in parte questa funzione disegnando un certo numero di vettori $\mathrm{v}(x, y)$, spiccati dai punti corrispondenti. Per esempio il campo di velocità rotatorio

$$\mathrm{v}(x, y) = (y, -x) = -y\mathbf{i} - x\mathbf{j}$$

può essere visualizzato con un disegno del tipo in figura 4.6.

 © 978-88-08-06485-1

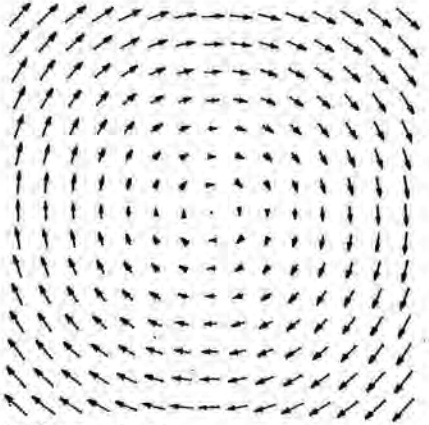


Figura 4.6. Il campo di velocità $\mathbf{v} = -y\mathbf{i} - x\mathbf{j}$

Generalizzando questo esempio, diciamo che un campo vettoriale è una funzione che ad ogni punto dello spazio fisico (o dello spazio-tempo) assegna un vettore, che può avere il significato di *velocità*, come nell'esempio precedente (campo di velocità) o di *forza* agente nel punto assegnato (campo gravitazionale, campo elettrostatico, ...). Gli spazi $\mathbb{R}^n$ dominio e codominio sono quindi "pensati" in due modi differenti: il primo come spazio fisico (insieme di punti), il secondo come insieme di vettori.

ESEMPIO

Un campo elettromagnetico variabile è una funzione che associa una coppia di vettori tridimensionali ad ogni punto (x, y, z) e istante t: il vettore $\mathbf{E}$ che descrive il campo elettrico e il vettore $\mathbf{H}$ che descrive il campo magnetico. Perciò è una funzione $(\mathbf{E}, \mathbf{H}) : \mathbb{R}^4 \to \mathbb{R}^6$.

Esempi come il precedente mostrano che considerare funzioni $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ con n, m qualunque non è un puro capriccio matematico, ma risponde a esigenze concrete.

ESERCIZI

1. Un'ellisse in forma parametrica è espressa da una funzione $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ con n, m uguali rispettivamente a ...

2. Un cilindro in forma parametrica è descritto da una funzione $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ con n, m uguali rispettivamente a ...

3. Se vogliamo descrivere simultaneamente pressione e temperatura di un gas racchiuso in un cilindro, punto per punto e istante per istante, utilizziamo una funzione $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$ con n, m uguali rispettivamente a ...

4. La superficie di un cono a due falde (fig. 4.7a) può essere rappresentata:
   - come superficie in forma parametrica (vero o falso?);
   - come grafico di una funzione di due variabili (vero o falso?).

   Rispondere alle stesse domande per un cono a una falda (fig. 4.7b).

© 978-88-08-06485-1

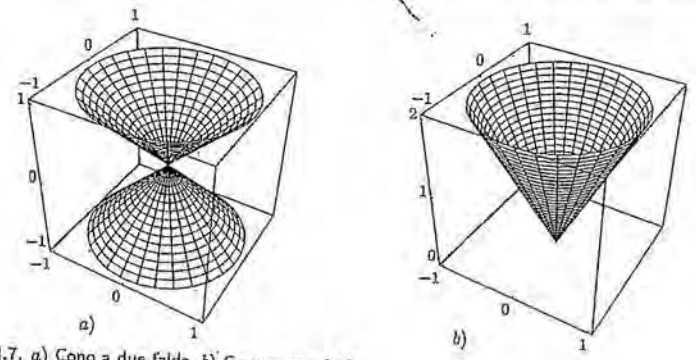


Figura 4.7. *a*) Cono a due falde. *b*) Cono a una falda.

5. Sia $U(x, y, z)$ il potenziale gravitazionale nel punto (x, y, z). Il luogo dei punti in cui $U(x, y, z) = 0$ sarà generalmente:
   a. una linea piana;
   b. una linea nello spazio;
   c. una superficie nello spazio;
   d. una regione nello spazio.

6. La scrittura
$$\mathbf{v} = (v_1(x, y), v_2(x, y), v_3(x, y))$$
rappresenta una funzione $\mathbf{v} : \mathbb{R}^n \to \mathbb{R}^m$ con n, m uguali rispettivamente a ...

7. Si disegni il dominio piano E descritto in coordinate polari da:
$$E = \left\{(\rho, \theta) : 1 < \rho < 2,\ 0 < \theta < \frac{\pi}{2}\right\}.$$

8. Si descriva in coordinate polari un settore circolare pari a $\frac{1}{8}$ di cerchio, di raggio 2.

9. Si descriva in coordinate cilindriche un cilindro pieno, di raggio 2 e altezza 3.

10. Si disegni l'oggetto geometrico descritto in coordinate sferiche da:
$$E = \left\{(\rho, \theta, \varphi) : 2 \le \rho \le 3,\ 0 \le \theta \le \pi,\ 0 \le \varphi \le \frac{\pi}{2}\right\}.$$

2 LIMITI, CONTINUITÀ E DIFFERENZIABILITÀ PER FUNZIONI $\mathbf{f} : \mathbb{R}^n \to \mathbb{R}^m$

Vogliamo ora definire i concetti di limite, continuità, differenziabilità per una funzione di più variabili a valori vettoriali. Si uniscono qui due fatti: le nozioni di limite, continuità, differenziabilità per funzioni reali di più variabili (studiate nel capitolo 3) e il fatto che, in generale, per studiare una funzione a valori vettoriali si può ragionare componente per componente (come visto nel capitolo 2). Di conseguenza questo studio non introdurrà nuove difficoltà sostanziali rispetto a quanto già visto nel capitolo 3.

© 978-88-08-06485-1

DEFINIZIONE 4.1 Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$ definita almeno in un intorno del punto $x_0 \in \mathbb{R}^n$, salvo al più x_0 stesso e sia $L \in \mathbb{R}^m$. Si dice che:

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

se

$$\lim_{x \to x_0} |f(x) - L| = 0$$

dove quest'ultimo è il limite di una funzione *reale* di più variabili, che è stato definito nel capitolo 3 (paragrafo 2.1)

Se $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, possiamo scrivere:

$$f(x) = (f_1(x), \ldots, f_m(x)), \text{ con } f_i : A \subseteq \mathbb{R}^n \to \mathbb{R} \text{ per } i = 1, 2, \ldots, m.$$

Le funzioni f_i si diranno *componenti di* f e sono funzioni *reali di più variabili*. Analogamente a quanto visto nel capitolo 2, si dimostra che il limite si calcola componente per componente, ossia:

$$\lim_{x \to x_0} (f_1(x), \ldots, f_m(x)) = \left(\lim_{x \to x_0} f_1(x), \ldots, \lim_{x \to x_0} f_m(x) \right).$$

Così pure, definito in maniera naturale il concetto di *funzione continua* (in un punto o un aperto), si dimostra che la funzione f è continua se e solo se lo sono tutte le sue componenti. Limiti e continuità si riconducono quindi alle analoghe nozioni per le funzioni di più variabili a valori reali, studiate nel capitolo 3.

Per quanto riguarda la *differenziabilità*:

DEFINIZIONE 4.2 Diremo che f è differenziabile in x_0 se tutte le componenti f_i lo sono.

Esplicitamente, questo significa che, per $i = 1, 2, \ldots, m$, valgono le relazioni:

$$f_i(x_0 + h) - f_i(x_0) = \sum_{j=i}^{n} \frac{\partial f_i}{\partial x_j}(x_0) h_j + o(|h|) \qquad \text{per } h \to 0.$$

Si possono riscrivere queste m relazioni in forma compatta usando una *notazione matriciale*. Rappresentiamo la funzione f e i punti di $\mathbb{R}^n$ come vettori colonna delle loro componenti:

$$f = \begin{pmatrix} f_1 \\ f_2 \\ \ldots \\ f_m \end{pmatrix} \qquad x_0 = \begin{pmatrix} x_1 \\ x_2 \\ \ldots \\ x_n \end{pmatrix} \qquad h = \begin{pmatrix} h_1 \\ h_2 \\ \ldots \\ h_n \end{pmatrix}$$

e introduciamo la *matrice Jacobiana* di f, che ha per righe i gradienti delle componenti di f (calcolati nel punto x_0):

$$Df(x_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \ldots & \frac{\partial f_1}{\partial x_n} \\ \frac{\partial f_2}{\partial x_1} & \ldots & \ldots & \frac{\partial f_2}{\partial x_n} \\ \ldots & \ldots & \ldots & \ldots \\ \frac{\partial f_m}{\partial x_1} & \ldots & \ldots & \frac{\partial f_m}{\partial x_n} \end{pmatrix} (x_0)$$

© 978-88-08-06485-1

La matrice Jacobiana ha m righe e n colonne. La differenziabilità di f si esprime allora con la relazione:

$$f(x_0 + h) - f(x_0) = Df(x_0)h + o(|h|) \qquad \text{per } h \to 0 \tag{2.1}$$

dove $Df(x_0)h$ denota il prodotto matriciale (righe per colonne) tra la matrice Jacobiana di f calcolata in x_0 e il vettore incremento h e il simbolo $o(|h|)$ indica ora un *vettore* che, diviso per $|h|$, tende a zero per $h \to 0$.

Il *differenziale di* f (calcolato in x_0) è la funzione lineare $df(x_0) : \mathbb{R}^n \to \mathbb{R}^m$ definita da:

$$df(x_0) : h \mapsto Df(x_0)h.$$

Ragionando componente per componente e applicando la condizione sufficiente per la differenziabilità delle funzioni a valori reali[1], si vede subito che:

TEOREMA 4.1 *Condizione sufficiente affinché una funzione* $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, *con A aperto, risulti differenziabile in A è che tutti gli elementi della sua matrice Jacobiana siano funzioni continue in A.*

Questo criterio è quello che applicheremo di solito per garantire la differenziabilità di una funzione di più variabili a valori vettoriali, senza dover verificare direttamente la definizione.

Anche nel caso vettoriale, inoltre, è vero che se f è differenziabile, a maggior ragione è derivabile e continua (ma non vale il viceversa).

Naturalmente le funzioni $f : \mathbb{R}^n \to \mathbb{R}$ e le funzioni $f : \mathbb{R} \to \mathbb{R}^m$ rientrano come casi particolari nella definizione appena data di differenziabilità. Con le notazioni che abbiamo introdotto, per una funzione $f : \mathbb{R}^n \to \mathbb{R}$ la matrice Jacobiana è semplicemente il gradiente di f, scritto come *vettore riga*, mentre per una funzione $f : \mathbb{R} \to \mathbb{R}^m$ la matrice Jacobiana è semplicemente il vettore derivato di f, scritto come *vettore colonna*. (Non c'è niente di assoluto nella convenzione assunta su come scrivere righe e colonne della matrice Jacobiana, tuttavia è comodo fissare una convenzione e attenersi sempre a quella).

La nozione di matrice Jacobiana permette di enunciare il teorema di differenziazione delle funzioni composte in un'unica forma generale, che comprende i casi particolari già trattati nel capitolo 3, paragrafo 4.5 e altri ancora.

TEOREMA 4.2 *Siano* $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^m$, $g : B \subseteq \mathbb{R}^m \to \mathbb{R}^k$ *e supponiamo che sia ben definita almeno in un intorno C di* $x_0 \in A$ *la funzione composta* $g \circ f : C \subseteq \mathbb{R}^n \to \mathbb{R}^k$. *Se* f *è differenziabile in* x_0 *e* g *è differenziabile in* $y_0 = f(x_0)$, *anche* $g \circ f$ *è differenziabile in* x_0 *e la sua matrice Jacobiana si ottiene come prodotto (matriciale) delle matrici Jacobiane di* f *e* g, *calcolate nei punti* x_0 *e* $f(x_0)$, *rispettivamente:*

$$D(g \circ f)(x_0) = Dg(f(x_0))Df(x_0) \tag{2.2}$$

Operativamente, la formula si usa così. Posto $z = g \circ f$ e $y = f$,

$$\frac{\partial z_j}{\partial x_i}(x_0) = \sum_{h=1}^{m} \frac{\partial z_j}{\partial y_h}(f(x_0)) \frac{\partial y_h}{\partial x_i}(x_0) \tag{2.3}$$

per ogni $i = 1, 2, \ldots n$, $j = 1, 2, \ldots, k$.

[1] v. capitolo 3, paragrafo 4.3.

© 978-88-08-06485-1

Per esercizio, il lettore è invitato a ritrovare i risultati provati nel capitolo 3, paragrafo 4.5, come casi particolari di questo teorema.

La dimostrazione di questo teorema usa le stesse tecniche viste nella dimostrazione dei suoi casi particolari che abbiamo appena citato e perciò verrà omessa.

Nei prossimi paragrafi ci occuperemo del calcolo differenziale per due classi di funzioni introdotte nel paragrafo 1: le superfici in forma parametrica e le trasformazioni di coordinate in $\mathbb{R}^n$. Generalizzando il concetto di superficie in forma parametrica ci occuperemo poi di varietà k-dimensionali in $\mathbb{R}^n$, argomento connesso alla generalizzazione del teorema di Dini sulle funzioni definite implicitamente. Legato allo studio delle trasformazioni di coordinate nello spazio, invece, è il problema di fornire delle condizioni per l'invertibilità almeno locale di una trasformazione. Infine, potremo riprendere il tema dell'ottimizzazione, trattando il caso in cui il problema è soggetto a vincoli, espressi da equazioni o disequazioni, argomento che richiederà alcuni degli strumenti sviluppati in questo capitolo.

3 SUPERFICI REGOLARI IN FORMA PARAMETRICA

Abbiamo visto che una *superficie in forma parametrica* è una funzione:

$$\mathbf{r} : A \subseteq \mathbb{R}^2 \to \mathbb{R}^3$$

con $\mathbf{r} = (x, y, z)$ e

$$\begin{cases} x = x(u, v) \\ y = y(u, v) \\ z = z(u, v) \end{cases} \qquad (u, v) \in A$$

dove A è un'opportuna regione del piano $\mathbb{R}^2$.

Vogliamo ora capire come si legge, dalla parametrizzazione della superficie, il fatto che essa sia regolare (per esempio, dotata in ogni punto di piano tangente). In analogia a quanto visto per le superfici che sono grafico di una funzione di più variabili, sarà opportuno anzitutto richiedere che la funzione r sia differenziabile (nel senso visto nel paragrafo precedente).

Ora, ragioniamo nel modo seguente. Se fissiamo uno dei parametri, per esempio v, e facciamo variare u, otteniamo una curva disegnata sulla superficie. Più precisamente, al variare del valore fissato $v = v_0$ otterremo una famiglia di curve; fissando (in ogni modo possibile) $u = u_0$ otterremo un'altra famiglia di curve sulla superficie. Queste linee si chiamano *linee coordinate*: per ogni punto della superficie passa esattamente una curva di ciascuna delle due famiglie.

Per esempio, nella rappresentazione parametrica della sfera (Esempio 1.1), le linee $\varphi = \varphi_0$ sono i paralleli, quelle $\theta = \theta_0$ sono i meridiani; per ogni punto della sfera (a eccezione dei poli, che come vedremo sono punti in cui la parametrizzazione è *irregolare*) passano esattamente un meridiano e un parallelo.

Tornando a una generica superficie $\mathbf{r} = \mathbf{r}(u, v)$, le linee coordinate hanno equazioni:

$$\mathbf{r} = \mathbf{r}(u, v_0)$$

e

$$\mathbf{r} = \mathbf{r}(u_0, v).$$

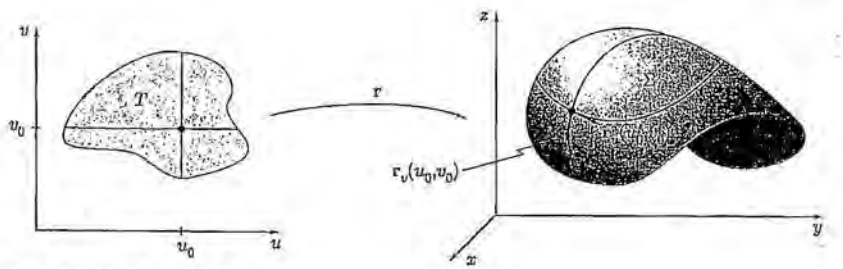


Figura 4.8. Linee coordinate.

Il vettore tangente a una linea coordinata è dato, rispettivamente, da:

$$\mathbf{r}_u(u_0, v_0) = (x_u(u_0, v_0), y_u(u_0, v_0), z_u(u_0, v_0))$$

e da

$$\mathbf{r}_v(u_0, v_0) = (x_v(u_0, v_0), y_v(u_0, v_0), z_v(u_0, v_0))$$

(dove abbiamo posto $x_u = \frac{\partial x}{\partial u}$, ecc.). Poiché stiamo assumendo r differenziabile, questi due vettori sono certamente ben definiti. Il piano tangente, da un punto di vista geometrico, dovrebbe essere l'unico piano che contiene entrambi questi vettori. Tale piano risulta ben definito se i due vettori sono entrambi non nulli e inoltre non sono paralleli (diversamente, infiniti piani soddisfano la richiesta). Queste condizioni si esprimono sinteticamente chiedendo che il prodotto vettoriale dei due vettori non si annulli:

$$\mathbf{r}_u(u_0, v_0) \times \mathbf{r}_v(u_0, v_0) = \det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ x_u(u_0, v_0) & y_u(u_0, v_0) & z_u(u_0, v_0) \\ x_v(u_0, v_0) & y_v(u_0, v_0) & z_v(u_0, v_0) \end{pmatrix} \neq 0 \tag{3.1}$$

Questa condizione è algebricamente equivalente al fatto che la matrice Jacobiana di r abbia caratteristica due (cioè massima).

Arriviamo così alla seguente:

DEFINIZIONE 4.3 Una *superficie parametrizzata* da $\mathbf{r} = \mathbf{r}(u, v)$, $\mathbf{r} : A \to \mathbb{R}^3$, $A \subseteq \mathbb{R}^2$, si dice *regolare* se r è differenziabile in A e inoltre la matrice Jacobiana di r ha caratteristica due in ogni punto di A. Se in qualche punto di A le condizioni vengono violate, chiameremo *punti singolari della superficie* i punti corrispondenti.

Esempio

Scriviamo la matrice Jacobiana della sfera di raggio R e centro nell'origine. Indicando con $\mathbf{r} = \mathbf{r}(\varphi, \theta)$ la sua rappresentazione parametrica (v. Esempio 1.1), si ha:

$$D\mathbf{r}(\varphi, \theta) = \begin{pmatrix} R\cos\varphi\cos\theta & -R\sin\varphi\sin\theta \\ R\cos\varphi\sin\theta & R\sin\varphi\cos\theta \\ -R\sin\varphi & 0 \end{pmatrix}.$$

I tre determinanti 2×2 che si ottengono sono:

$$R^2 \sin\varphi\cos\varphi \qquad R^2 \sin^2\varphi \sin\theta \qquad R^2 \sin^2\varphi\cos\theta$$

Dagli ultimi due si vede che per $\sin\varphi \neq 0$ il punto è regolare (perché $\sin\theta$ e $\cos\theta$ non si annullano contemporaneamente); invece, per $\sin\varphi = 0$ si annullano tutti e tre i determinanti e il punto è singolare. Ciò accade per $\varphi = 0$ e $\varphi = \pi$, cioè (per il significato di φ) nei poli della sfera: i poli sono dunque punti singolari; al di fuori di questi, la superficie è regolare. Ovviamente la singolarità non dipende dalla sfera intesa come oggetto geometrico (i poli sono punti come gli altri), ma dalla parametrizzazione scelta. Tuttavia, e questo dipende proprio dalla "forma" della sfera, si può dimostrare che non esiste alcuna parametrizzazione globale della sfera per la quale tutti i punti siano regolari.

Il vettore (3.1), la cui esistenza e non annullamento definisce il concetto di regolarità della superficie, ha un altro significato geometrico: per le proprietà del prodotto vettoriale, esso è ortogonale al piano che contiene i vettori $\mathbf{r}_u$, $\mathbf{r}_v$, cioè al piano tangente. Tale vettore può quindi dirsi *normale* alla superficie. Il suo versore prende il nome di *versore normale*:

$$\mathbf{n} = \frac{\mathbf{r}_u \times \mathbf{r}_v}{|\mathbf{r}_u \times \mathbf{r}_v|}.$$

Si noti che il *verso* di $\mathbf{n}$ dipende dall'ordine in cui consideriamo i parametri (u, v) e quindi dall'ordine in cui effettuiamo il prodotto vettoriale. Per esempio, per la sfera parametrizzata come nell'esempio precedente si ha:

$$\mathbf{n} = \frac{\mathbf{i}R^2\sin^2\varphi\cos\theta + \mathbf{j}R^2\sin^2\varphi\sin\theta + \mathbf{k}R^2\sin\varphi\cos\varphi}{R^2\sin\varphi} =$$

$$= \mathbf{i}\sin\varphi\cos\theta + \mathbf{j}\sin\varphi\sin\theta + \mathbf{k}\cos\varphi$$

(Notare che per la sfera centrata nell'origine è $\mathbf{n} = \mathbf{r}/|\mathbf{r}|$). Si osserva che questo versore punta verso l'*esterno* della sfera, in ogni punto: si dirà perciò *versore normale uscente*, o *normale esterna* alla sfera. Se avessimo considerato θ come primo parametro e φ come secondo, avremmo ottenuto un versore normale *entrante*, o interno, anziché uscente. Si dice *orientazione di una superficie* una scelta coerente di uno dei due versi del versore normale in ogni punto della superficie, quando ciò è possibile. Qualcosa di più su questo problema sarà detto nel capitolo 6, paragrafo 4.1.

Scriviamo ora l'equazione del *piano tangente* alla superficie. Se $\mathbf{x}(x, y, z)$ è il punto mobile sul piano tangente e $\mathbf{x}_0$ il punto in cui vogliamo calcolare il piano tangente, il vettore $(\mathbf{x} - \mathbf{x}_0)$ dovrà essere ortogonale a $\mathbf{n}$, quindi anche a $\mathbf{r}_u \times \mathbf{r}_v$. L'equazione del piano tangente è dunque:

$$(\mathbf{x} - \mathbf{x}_0) \cdot (\mathbf{r}_u \times \mathbf{r}_v) = 0$$

ossia:

$$\det \begin{pmatrix} \mathbf{x} - \mathbf{x}_0 \\ \mathbf{r}_u \\ \mathbf{r}_v \end{pmatrix} = 0.$$

Superfici cartesiane (grafico di funzioni di due variabili)
Il grafico di $z = f(x, y)$, $(x, y) \in A \subset \mathbb{R}^2$ è una superficie che si può scrivere in forma parametrica nel modo seguente:

$$\begin{cases} x = u \\ y = v \\ z = f(u, v) \end{cases} \qquad (u, v) \in A.$$

Perciò il vettore normale si calcola come segue:

$$\det \begin{pmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ 1 & 0 & f_u(u_0, v_0) \\ 0 & 1 & f_v(u_0, v_0) \end{pmatrix} = -\mathbf{i}f_u - \mathbf{j}f_v + \mathbf{k}$$

$$\mathbf{n} = \frac{-\mathbf{i}f_u - \mathbf{j}f_v + \mathbf{k}}{\sqrt{1 + |\nabla f|^2}}. \tag{3.2}$$

Quindi l'orientazione indotta dalla superficie è con la normale verso l'alto.

Si noti che se f è differenziabile in A, il suo grafico è sempre una superficie regolare (la matrice Jacobiana ha sempre un minore 2×2 con determinante uguale a 1).

Nella formula (3.2) solitamente le variabili u, v si continuano a chiamare x, y, come nell'esempio seguente.

Esempio

Il *paraboloide ellittico*

$$z = x^2 + y^2$$

ha versore normale:

$$\mathbf{n} = \frac{-2x\mathbf{i} - 2y\mathbf{j} + \mathbf{k}}{\sqrt{1 + 4(x^2 + y^2)}}.$$

Superfici di rotazione
Molte superfici comuni si possono ottenere facendo ruotare una curva γ, detta *generatrice*, attorno a un asse; come vedremo, questo è anche un metodo per scriverne effettivamente le equazioni parametriche.

In un riferimento (x, y, z), sia z l'asse attorno a cui vogliamo far ruotare una curva γ, inizialmente assegnata nel piano x, z in forma parametrica:

$$\begin{cases} x = x(t) \\ z = z(t) \end{cases} \qquad t \in I.$$

In tal caso la superficie che si ottiene con una rotazione completa della curva γ attorno all'asse z è (fig. 4.9):

$$\begin{cases} x = x(t)\cos\theta \\ y = x(t)\sin\theta \\ z = z(t) \end{cases} \qquad t \in I, \theta \in [0, 2\pi). \tag{3.3}$$

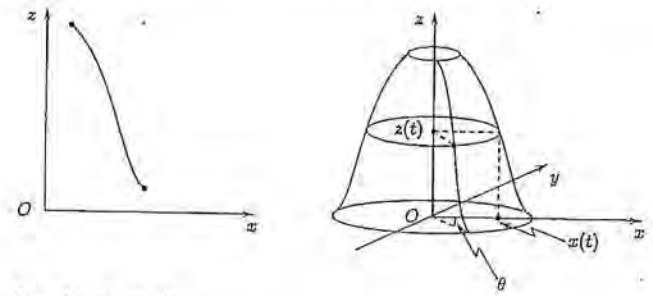


Figura 4.9. Superficie di rotazione.

Esempi

La sfera. È generata dalla rotazione intorno all'asse z della semicirconferenza assegnata nel piano x, z:

$$\begin{cases} x = R\sin\varphi \\ z = R\cos\varphi \end{cases} \quad \varphi \in [0, \pi].$$

Perciò, in base alla (3.3), la sfera (fig. 4.10) ha le equazioni note:

$$\begin{cases} x = R\sin\varphi\cos\theta \\ y = R\sin\varphi\sin\theta \\ z = R\cos\varphi \end{cases} \quad \varphi \in [0, \pi], \theta \in [0, 2\pi).$$

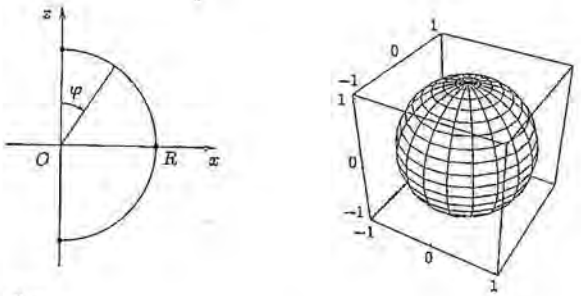


Figura 4.10. La sfera.

Il toro. È generato dalla rotazione intorno all'asse z della circonferenza, posta nel piano x, z, di centro $(R, 0)$ e raggio $r < R$:

$$\begin{cases} x = R + r\cos\varphi \\ z = r\sin\varphi \end{cases} \quad \varphi \in [0, 2\pi).$$

Perciò, in base alla (3.3), il toro (fig. 4.11) ha equazioni:

$$\begin{cases} x = (R + r\sin\varphi)\cos\theta \\ y = (R + r\sin\varphi)\sin\theta \\ z = r\cos\varphi \end{cases} \quad \varphi \in [0, 2\pi), \theta \in [0, 2\pi).$$

© 978-88-08-06485-1

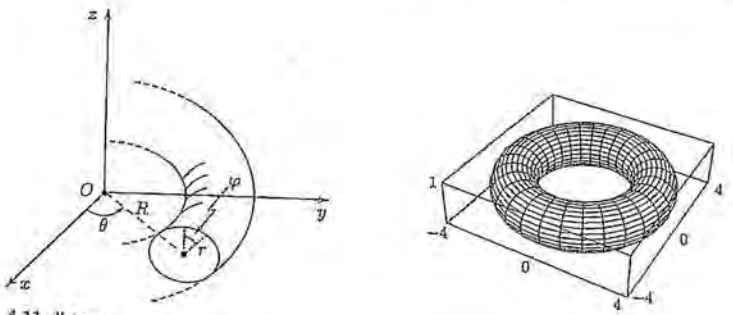


Figura 4.11. Il toro.

Il cono. Un cono a due falde è generato dalla rotazione intorno all'asse z di una retta, nel piano x, z, passante per l'origine: $z = mx, x \in \mathbb{R}$. In forma parametrica, la retta ha equazioni:

$$\begin{cases} x = t \\ z = mt \end{cases} \quad t \in \mathbb{R},$$

perciò, per la (3.3), il cono (fig. 4.12) ha equazioni:

$$\begin{cases} x = t\cos\theta \\ y = t\sin\theta \\ z = mt \end{cases} \quad t \in \mathbb{R}, \theta \in [0, 2\pi).$$

Il cono a una falda ha le stesse equazioni, ma $t \geq 0, \theta \in [0, 2\pi)$.

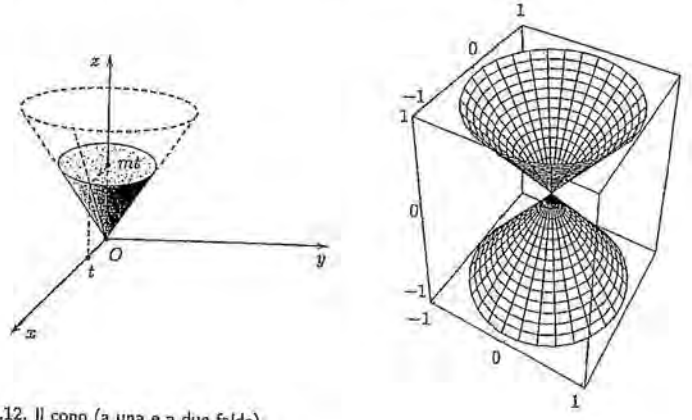


Figura 4.12. Il cono (a una e a due falde).

Il cilindro. Un cilindro di raggio r è generato dalla rotazione intorno all'asse z della retta (assegnata) posta nel piano x, z:

$$\begin{cases} x = r \\ z = t \end{cases} \quad t \in \mathbb{R},$$

 © 978-88-08-06485-1

perciò, per la (3.3), il cilindro (fig. 4.13) ha equazioni:

$$\begin{cases} x = r\cos\theta \\ y = r\sin\theta \\ z = t \end{cases} \qquad t \in \mathbb{R}, \theta \in [0, 2\pi).$$

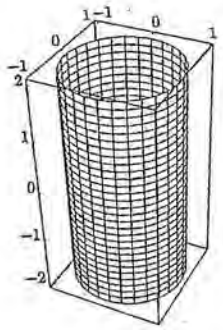


Figura 4.13. Il cilindro.

L'ellissoide di rotazione. È generato dalla rotazione intorno all'asse z di una semiellisse definita nel piano x, z dalle equazioni:

$$\begin{cases} x = a\sin\varphi \\ z = b\cos\varphi \end{cases} \qquad \varphi \in [0, \pi].$$

Perciò (fig. 4.14), per le (3.3), ha equazioni:

$$\begin{cases} x = a\sin\varphi\cos\theta \\ y = a\sin\varphi\sin\theta \\ z = b\cos\varphi \end{cases} \qquad \varphi \in [0, \pi], \theta \in [0, 2\pi).$$

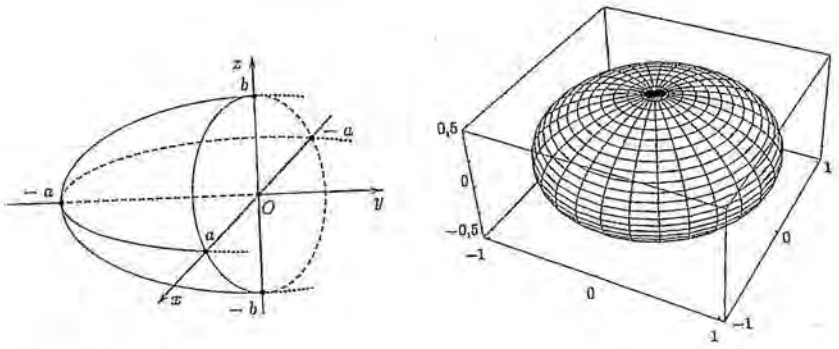


Figura 4.14. Ellissoide di rotazione.

L'ellissoide di rotazione è un particolare ellissoide, che ha due assi su tre uguali; è immediato generalizzare le equazioni precedenti all'ellissoide (generico) di semiassi a, b, c (che non è più una superficie di rotazione):

$$\begin{cases} x = a\sin\varphi\cos\theta \\ y = b\sin\varphi\sin\theta \\ z = c\cos\varphi \end{cases} \qquad \varphi \in [0, \pi], \theta \in [0, 2\pi).$$

Esercizi

147 Calcolare il versore normale sulla superficie del paraboloide iperbolico $z = xy$.

148 Scrivere la matrice Jacobiana del toro; verificare che la superficie è regolare in ogni punto.

149 Scrivere la matrice Jacobiana del cono a due falde; verificare che il vertice è l'unico punto singolare.

150 Scrivere la matrice Jacobiana di un ellissoide di semiassi a, b, c; verificare che i poli sono gli unici punti singolari; scrivere il versore normale e il piano tangente sulla superficie; il versore normale punta verso l'interno o l'esterno?

151 Scrivere le equazioni parametriche della superficie ottenuta facendo ruotare attorno all'asse z la curva $x = e^{-|z|}, z \in \mathbb{R}$, posta nel piano xz. Scriverne la matrice Jacobiana e determinare gli eventuali punti singolari. Fare un disegno della superficie per convincersi del risultato trovato. Fare poi lo stesso per la curva $z = e^{-|x|}$, $x \in \mathbb{R}$.

152 Scrivere le equazioni parametriche della superficie ottenuta facendo ruotare attorno all'asse z la curva $x = z(1-z)$, $z \in [0, 1]$, posta nel piano xz. Scriverne la matrice Jacobiana e determinare gli eventuali punti singolari. Fare un disegno della superficie per convincersi del risultato trovato.

153 Scrivere le equazioni parametriche della superficie che si ottiene facendo ruotare attorno all'asse y la curva grafico della funzione:

$$y = e^{-x^2}, \; x \in [0, 2]$$

Calcolarne poi la matrice Jacobiana, determinando gli eventuali punti singolari.

154 Nel piano xz, si consideri l'arco di curva parametrizzato:

$$\begin{cases} x = t(1+t) \\ z = \sin(\pi t) \end{cases} \qquad t \in [0, 1].$$

Si scrivano le equazioni parametriche della superficie generata dalla rotazione della curva attorno all'asse z. Si scriva la matrice Jacobiana della superficie e si dica se questa è regolare, specificando in caso contrario quali e quanti sono i suoi punti singolari.

155 Scrivere la matrice Jacobiana e il versore normale per la superficie grafico della funzione $f(x, y) = x^2y^3$.

156 Scrivere la matrice Jacobiana per la superficie:

$$\begin{cases} x = \sin^2\varphi\cos\theta \\ y = \sin^2\varphi\sin\theta \\ z = \cos^2\varphi \end{cases} \qquad \varphi \in [0, \pi], \theta \in [0, 2\pi].$$

Descrivere poi l'insieme dei punti singolari della superficie. Quanti sono?

L'insieme in cui variano i parametri naturalmente è fondamentale per precisare la superficie parametrizzata. Tenendo conto di questo, scrivere la rappresentazione parametrica delle seguenti porzioni di superficie:

La superficie laterale di un cilindro di raggio r e altezza h.

La porzione di superficie sferica di raggio 4 compresa tra un piano equatoriale e un piano parallelo a questo posto a distanza 2.

La superficie laterale di un cono di altezza h e raggio di base r.

La superficie laterale di un tronco di cono avente il raggio della base maggiore uguale a 2, il raggio della base minore uguale a 1, le basi poste tra loro a distanza 1.

4 VARIETÀ k-DIMENSIONALI IN $\mathbb{R}^n$ E FUNZIONI DEFINITE IMPLICITAMENTE

Le curve e le superfici regolari nello spazio tridimensionale sono casi particolari di un concetto geometrico molto più generale: nello spazio $\mathbb{R}^n$ (di dimensione qualunque) possiamo considerare quei sottoinsiemi che si possono vedere come oggetti geometrici curvi, regolari, di dimensione k, per qualsiasi valore di k compreso tra 1 (che è il caso delle curve in $\mathbb{R}^n$, già considerato) e $n-1$ (che sarà il caso delle *ipersuperfici* in $\mathbb{R}^n$). Questi oggetti si chiameranno *k-varietà*, o *varietà k-dimensionali*, di $\mathbb{R}^n$. Così come una superficie in $\mathbb{R}^3$, per esempio la sfera, può essere definita sia in forma parametrica,

$$\begin{cases} x = R\sin\varphi\cos\theta \\ y = R\sin\varphi\sin\theta \\ z = R\cos\varphi \end{cases} \qquad \varphi \in [0,\pi]\,,\ \theta \in [0,2\pi)$$

ossia come funzione $\mathrm{r} : \mathbb{R}^2 \to \mathbb{R}^3$, sia in forma implicita,

$$x^2 + y^2 + z^2 - R^2 = 0,$$

ossia mediante un'equazione in tre variabili, allo stesso modo una varietà k-dimensionale si potrà definire sia in forma parametrica, mediante un'opportuna funzione $\mathrm{r} : \mathbb{R}^k \to \mathbb{R}^n$ $(1 \le k \le n-1)$, sia in forma implicita, mediante un sistema di $n-k$ equazioni in n variabili. Dedicheremo il paragrafo 4.1 alla definizione delle varietà k-dimensionali in forma parametrica e il paragrafo 4.3 alla definizione delle varietà k-dimensionali in forma implicita. Strettamente legata a quest'ultimo concetto, come vedremo, è una naturale generalizzazione del teorema della funzione implicita al caso di sistemi di equazioni: discuteremo di questo nel paragrafo 4.2, preliminarmente alla trattazione delle varietà in forma implicita.

4.1 Varietà k-dimensionali in $\mathbb{R}^n$ in forma parametrica

DEFINIZIONE 4.4 Si dice *varietà regolare k-dimensionale in forma parametrica, immersa in* $\mathbb{R}^n$ (con $1 \le k \le n-1$), una funzione $\mathrm{r} : \Omega \to \mathbb{R}^n$, con Ω aperto di $\mathbb{R}^k$, tale che $\mathrm{r} \in C^1(\Omega)$ e la matrice Jacobiana di r ha rango k in ogni punto di Ω.

Notiamo che per $k = 1$ ritroviamo la definizione di curva regolare in $\mathbb{R}^n$, mentre per $k = 2$ e $n = 3$ ritroviamo la definizione di superficie regolare parametrizzata in $\mathbb{R}^3$. Nel caso particolare $k = n-1$ (e $n > 3$) la varietà si dice *ipersuperficie* in $\mathbb{R}^n$.

Dal punto di vista dell'intuizione geometrica, la varietà k-dimensionale è l'*immagine* della funzione r in $\mathbb{R}^n$, più che la funzione stessa (così come quando pensiamo a una curva pensiamo al suo sostegno in $\mathbb{R}^n$, più che alla funzione che la definisce). Coerentemente con questa intuizione, vedremo l'aperto Ω come lo spazio dei parametri e diremo per esempio che un punto $\mathrm{r} = \mathrm{r}(u_1, u_2, \ldots, u_k)$ di $\mathbb{R}^n$ *appartiene* alla varietà, o *giace sulla* varietà. Possiamo anche pensare alla varietà k-dimensionale come a un oggetto geometrico, in uno spazio n dimensionale, su cui una particella si può muovere con k gradi di libertà.

Il significato geometrico della richiesta di rango k della matrice Jacobiana è allora il seguente.

Per ogni punto $(\overline{u}_1, \overline{u}_2, \ldots, \overline{u}_k) \in \Omega$, consideriamo il corrispondente punto $\overline{\mathrm{x}} = \mathrm{r}(\overline{u}_1, \overline{u}_2, \ldots, \overline{u}_k)$ sulla varietà. Consideriamo ora le k curve (in $\mathbb{R}^n$), tutte passanti per $\overline{\mathrm{x}}$ e tracciate sulla varietà, di equazioni parametriche

$$\mathrm{r}_i(t) = \mathrm{r}\left(\overline{u}_1, \overline{u}_2, \ldots, \underset{i}{t}, \ldots, \overline{u}_k\right), \quad i = 1, 2, \ldots, k$$

definite per t in un intorno di $\overline{u}_i$, rispettivamente. La richiesta di rango massimo della matrice Jacobiana di r significa che i k vettori tangenti in $\overline{\mathrm{x}}$ alle k curve ora descritte sono tra loro linearmente indipendenti. Questo significa che queste k famiglie di curve costituiscono un buon sistema di "linee coordinate" tracciate sulla varietà.

Esempio

In $\mathbb{R}^4$, la varietà 3-dimensionale definita dalle equazioni parametriche:

$$\begin{cases} x_1 = R\cos\phi_1 \\ x_2 = R\sin\phi_1\cos\phi_2 \\ x_3 = R\sin\phi_1\sin\phi_2\cos\phi_3 \\ x_4 = R\sin\phi_1\sin\phi_2\sin\phi_3 \end{cases}$$

dove $\phi_1, \phi_2 \in [0,\pi]$, $\phi_3 \in [0, 2\pi)$, rappresenta l'*ipersfera di raggio R e centro l'origine*. Infatti è un'ipersuperficie in $\mathbb{R}^4$ e i suoi punti soddisfano la relazione

$$x_1^2 + x_2^2 + x_3^2 + x_4^2 = R^2,$$

perciò è il luogo dei punti avente distanza R dall'origine.

Verifichiamo la regolarità calcolando la matrice Jacobiana:

$$\begin{pmatrix} -R\sin\phi_1 & 0 & 0 \\ R\cos\phi_1\cos\phi_2 & -R\sin\phi_1\sin\phi_2 & 0 \\ R\cos\phi_1\sin\phi_2\cos\phi_3 & R\sin\phi_1\cos\phi_2\cos\phi_3 & -R\sin\phi_1\sin\phi_2\sin\phi_3 \\ R\cos\phi_1\sin\phi_2\sin\phi_3 & R\sin\phi_1\cos\phi_2\sin\phi_3 & R\sin\phi_1\sin\phi_2\cos\phi_3 \end{pmatrix}.$$

Ora, il determinante della matrice 3×3 corrispondente alle righe 1,2,3 è:

$$-R^3\sin^3\phi_1\sin^2\phi_2\sin\phi_3;$$

mentre il determinante della matrice 3×3 corrispondente alle righe 1,2,4 è:

$$R^3\sin^3\phi_1\sin^2\phi_2\cos\phi_3.$$

 © 978-88-08-06485-1

Il confronto tra i due mostra che se $\sin\phi_1 \sin\phi_2 \neq 0$ la caratteristica della matrice è 3. D'altro canto, se $\sin\phi_1 = 0$ la seconda e la terza colonna della matrice sono nulle e se $\sin\phi_2 = 0$ la terza colonna della matrice è nulla. In conclusione: la caratteristica della matrice Jacobiana è massima se e solo se

$$\sin\phi_1 \sin\phi_2 \neq 0.$$

I punti singolari dell'ipersuperficie sono dunque quelli che corrispondono a $\sin\phi_1 = 0$, cioè

$$(\pm R, 0, 0, 0)$$

(ovvero i due poli dell'ipersfera), e quelli che corrispondono a $\sin\phi_2 = 0$, cioè

$$(R\cos\phi_1, \pm R\sin\phi_1, 0, 0) \text{ per } \phi_1 \in [0, \pi],$$

che rappresentano complessivamente i punti della circonferenza massima $x_1^2 + x_2^2 = R^2$ sul piano $x_3 = x_4 = 0$.

Scopriamo così che, in 4 dimensioni, l'ipersfera in forma parametrica ha un'intera circonferenza di punti singolari, e non solo i due poli come in 3 dimensioni.

4.2 Funzioni implicite definite da sistemi di equazioni

Abbiamo visto (capitolo 3, paragrafo 8) che un'equazione

$$f(x_1, x_2, \ldots, x_n, y) = 0$$

sotto ipotesi ragionevoli definisce implicitamente una funzione

$$y = g(x_1, x_2, \ldots, x_n).$$

Più in generale, potremmo avere un sistema di m equazioni in $n+m$ variabili e la domanda naturale è se in questo caso riusciamo a "risolvere il sistema" esplicitando m variabili (ossia tante quante sono le equazioni a nostra disposizione) in funzione delle rimanenti n. Nel caso di un sistema di equazioni lineari, questo problema è ben noto in algebra lineare e la sua soluzione è fornita dal teorema di Rouché-Capelli: se la matrice del sistema ha rango massimo (cioè m) si riescono effettivamente a esplicitare m variabili in funzione delle altre n. Nel caso non lineare, il meglio che si può sperare, in generale, è che sia possibile far questo *localmente*, ossia nell'intorno di un punto in cui si sa che il sistema è soddisfatto; l'analogo dell'ipotesi di rango massimo della matrice del sistema lineare sarà un'ipotesi di rango massimo *sulla matrice del sistema linearizzato*, ossia *sulla matrice Jacobiana del sistema non lineare*.

Arriviamo così a formulare una versione del teorema del Dini per sistemi di equazioni. Fissiamo anzitutto un po' di notazioni compatte. Scriviamo un sistema di m equazioni in $n+m$ variabili nella forma vettoriale:

$$\mathbf{f}(\mathbf{x}, \mathbf{y}) = 0$$

dove $\mathbf{f} : A \subset \mathbb{R}^{n+m} \to \mathbb{R}^m$, $\mathbf{x} \in \mathbb{R}^n$, $\mathbf{y} \in \mathbb{R}^m$. Denotiamo con $\mathbf{D}_\mathbf{y}\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)$ la matrice Jacobiana della funzione

$$\mathbf{y} \longmapsto \mathbf{f}(\mathbf{x}_0, \mathbf{y}),$$

calcolata nel punto $\mathbf{y}_0$, ossia la matrice $m \times m$ di elementi $\frac{\partial f_i}{\partial y_j}(\mathbf{x}_0, \mathbf{y}_0)$ $(i, j = 1, 2, \ldots, m)$ e con $\mathbf{D}_\mathbf{x}\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0)$ la matrice Jacobiana della funzione

$$\mathbf{x} \longmapsto \mathbf{f}(\mathbf{x}, \mathbf{y}_0),$$

© 978-88-08-06485-1

calcolata nel punto $\mathbf{x}_0$, ossia la matrice $m \times n$ di elementi $\frac{\partial f_i}{\partial x_j}(\mathbf{x}_0, \mathbf{y}_0)$ $(i = 1, 2, \ldots, m; j = 1, 2, \ldots, n)$.

Teorema 4.3 (di Dini, della funzione implicita: caso generale) *Sia A un aperto di $\mathbb{R}^{n+m}$, $\mathbf{f} : A \to \mathbb{R}^m$, $\mathbf{f} \in C^1(A)$ e supponiamo che nel punto $(\mathbf{x}_0, \mathbf{y}_0) \in A$ sia*

$$\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0) = 0; \quad \det \mathbf{D}_\mathbf{y}\mathbf{f}(\mathbf{x}_0, \mathbf{y}_0) \neq 0.$$

Allora esistono un intorno $U \subset \mathbb{R}^n$ di $\mathbf{x}_0$ e un'unica funzione $\mathbf{g} : U \to \mathbb{R}^m$, $\mathbf{g} \in C^1(U)$, tale che, per ogni $\mathbf{x} \in U$,

$$\mathbf{f}(\mathbf{x}, \mathbf{g}(\mathbf{x})) = 0$$

$$\mathbf{D}\mathbf{g}(\mathbf{x}) = -\mathbf{D}_\mathbf{y}\mathbf{f}(\mathbf{x}, \mathbf{g}(\mathbf{x}))^{-1}\mathbf{D}_\mathbf{x}\mathbf{f}(\mathbf{x}, \mathbf{g}(\mathbf{x})). \tag{4.1}$$

Si noti che il calcolo della matrice Jacobiana della funzione implicita richiede il calcolo della matrice inversa della Jacobiana $\mathbf{D}_\mathbf{y}\mathbf{f}$, è quindi più laborioso rispetto al caso scalare. Anche questo teorema, come il suo analogo scalare, ha un significato eminentemente teorico: non aiuta a scrivere esplicitamente la funzione implicita, ma a dimostrare che essa esiste, è unica e ha una certa regolarità.

Esempio

Consideriamo il sistema di due equazioni in tre variabili:

$$\begin{cases} x^2 + y^2 = 1 \\ x^2 + \frac{y^2}{4} + \frac{z^2}{4} = 1 \end{cases}$$

Il significato geometrico del sistema è: trovare l'intersezione tra il cilindro circolare che ha asse sull'asse z e raggio 1 (prima equazione) e l'ellissoide che ha centro l'origine, assi sugli assi coordinati, di lunghezze 1, 2, 2 rispettivamente (seconda equazione). Notiamo che soddisfano il sistema, per esempio, i punti:

$$(1, 0, 0) \text{ e } \left(0, 1, \sqrt{3}\right).$$

Ci aspettiamo che l'intersezione delle due superfici consista di una o più curve, più o meno regolari. Rappresentando il sistema in forma compatta come $\mathbf{f}(x, y, z) = \mathbf{0}$, scriviamo la matrice Jacobiana:

$$\mathbf{D}\mathbf{f}(x, y, z) = \begin{pmatrix} 2x & 2y & 0 \\ 2x & \frac{y}{2} & \frac{z}{2} \end{pmatrix}.$$

Si ha:

$$\mathbf{D}\mathbf{f}(1, 0, 0) = \begin{pmatrix} 2 & 0 & 0 \\ 2 & 0 & 0 \end{pmatrix}, \text{ di rango 1;}$$

$$\mathbf{D}\mathbf{f}\left(0, 1, \sqrt{3}\right) = \begin{pmatrix} 0 & 2 & 0 \\ 0 & 1 & \frac{\sqrt{3}}{2} \end{pmatrix}, \text{ di rango 2.}$$

Cominciamo dal punto $A\left(0, 1, \sqrt{3}\right)$. In base al teorema 4.3, in un suo intorno è possibile esplicitare y e z in funzione di x, rappresentando così la linea di intersezione tra ellissoide

e cilindro come curva regolare parametrizzata. In questo caso è anche possibile risolvere esplicitamente il sistema, ottenendo:

$$\begin{cases} y = \sqrt{1-x^2} \\ z = \sqrt{3\left(1-x^2\right)} \end{cases}$$

per (x, y, z) in un intorno di $(0, 1, \sqrt{3})$. Invece nel punto $B = (1, 0, 0)$ il teorema non è applicabile: l'intersezione tra le due superfici in un intorno di quel punto non è univocamente rappresentabile come curva regolare. Infatti questa intersezione consiste di due linee regolari che si intersecano proprio nel punto $B\,(1, 0, 0)$ (fig. 4.15). Come mai il procedimento usato sopra per risolvere manualmente il sistema nell'intorno di $(0, 1, \sqrt{3})$ fallisce in un intorno di $(1, 0, 0)$?

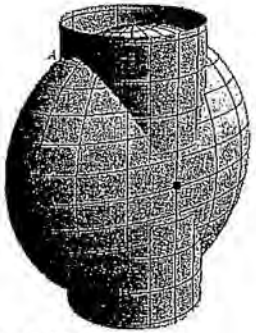


Figura 4.15. Intersezione tra cilindro ed ellissoide.

4.3 Varietà k-dimensionali in $\mathbb{R}^n$ in forma implicita

Proseguendo il discorso iniziato nel paragrafo 4.1, diamo ora la definizione di varietà k-dimensionale in $\mathbb{R}^n$, in forma cartesiana implicita. Il teorema della funzione implicita per i sistemi, appena dimostrato, consentirà di mettere in evidenza il nesso tra questo concetto e quello di varietà k-dimensionale in forma parametrica, introdotto nel paragrafo 4.1.

DEFINIZIONE 4.5 Si dice *varietà k-dimensionale in forma implicita, immersa in* $\mathbb{R}^n$, un sottoinsieme (non vuoto) di $\mathbb{R}^n$ del tipo:

$$M = \{\mathbf{x} \in \Omega : \mathbf{f}(\mathbf{x}) = 0\},$$

dove Ω è un aperto di $\mathbb{R}^n$, $\mathbf{f} : \Omega \to \mathbb{R}^{n-k}$, $\mathbf{f} \in C^1(\Omega)$ e il rango di $D\mathbf{f}$ è uguale a $n-k$ in ogni punto di Ω. La varietà si dirà *di classe* C^m (per qualche intero $m \geq 1$) se, inoltre, $\mathbf{f} \in C^m(\Omega)$; si dirà *di classe* C^∞ se è di classe C^m per ogni intero m.

Esempi

Si noti che se $k = n-1$ la funzione f ha valori reali; in questo caso M è definita da un'unica equazione in n variabili: parliamo allora di *ipersuperficie in forma cartesiana implicita*; nel caso $n = 3$, si ha semplicemente una *superficie* in forma cartesiana implicita, come la sfera in $\mathbb{R}^3$:

$$x^2 + y^2 + z^2 - R^2 = 0.$$

Se $n = 3$ e $k = 1$, abbiamo due equazioni in tre variabili che, se la condizione di rango è soddisfatta, individuerà una *curva* in forma cartesiana implicita (vista come intersezione di due superfici cartesiane), come nell'Esempio 4.2.

Se M è una varietà k-dimensionale in forma implicita e $\mathbf{x}_0$ è un punto qualsiasi di M, cioè tale che $\mathbf{f}(\mathbf{x}_0) = 0$, l'ipotesi sul rango di $D\mathbf{f}$ consente di applicare il teorema della funzione implicita per i sistemi (dimostrato nel paragrafo 4.2) ed affermare che, posto

$$\mathbf{x}_0 = (\mathbf{u}_0, \mathbf{y}_0), \mathbf{u}_0 \in \mathbb{R}^k, \mathbf{y}_0 \in \mathbb{R}^{n-k},$$

si può garantire l'esistenza di un intorno $U \subset \mathbb{R}^k$ di $\mathbf{u}_0$ e un'unica funzione $\mathbf{g} : U \to \mathbb{R}^{n-k}$, tale che $\mathbf{g} \in C^1(U)$ e

$$\mathbf{f}(\mathbf{u}, \mathbf{g}(\mathbf{u})) = 0 \text{ per ogni } \mathbf{u} \in U.$$

Questo significa che la varietà M ha coordinate esprimibili in funzione di esattamente k parametri $u_1, u_2, \ldots, u_k$. In altre parole, localmente, la varietà M si può vedere come varietà k-dimensionale in forma parametrica, descritta da equazioni del tipo:

$$\begin{cases} x_1 = u_1 \\ x_2 = u_2 \\ \ldots \\ x_k = u_k \\ x_{k+1} = g_1(u_1, u_2, \ldots, u_k) \\ \ldots \\ x_n = g_{n-k}(u_1, u_2, \ldots, u_k) \end{cases} \quad \text{per } (u_1, u_2, \ldots, u_k) \in U.$$

Esempi

Varietà delle configurazioni di un corpo rigido. Vogliamo descrivere lo spazio delle configurazioni possibili di un corpo rigido. La posizione di un corpo rigido nello spazio è individuata quando sono assegnate le posizioni di 3 suoi punti, purché non allineati. Dunque, fissati 3 punti del corpo rigido $\mathbf{p}_1, \mathbf{p}_2, \mathbf{p}_3$, non allineati, di coordinate

$$\mathbf{p}_i = (x_i, y_i, z_i),\ i = 1, 2, 3,$$

ogni posizione del corpo sarà esprimibile mediante le 9 coordinate dei tre punti $\mathbf{p}_i$, ossia: ogni posizione del corpo rigido è individuabile con un punto di $\mathbb{R}^9$. Non ogni punto di $\mathbb{R}^9$ individua però una posizione del corpo rigido: infatti, le coordinate dei 3 punti sono sottoposti al vincolo di avere mutue distanze fissate. In altre parole, proprio perché il corpo è rigido si ha:

$$\begin{aligned} |\mathbf{p}_1 - \mathbf{p}_2| &= a \\ |\mathbf{p}_2 - \mathbf{p}_3| &= b \\ |\mathbf{p}_1 - \mathbf{p}_3| &= c \end{aligned}$$

con a, b, c opportune costanti. Imponendo questi vincoli esplicitamente, si scrivono le tre equazioni:

$$\begin{cases} (x_1 - x_2)^2 + (y_1 - y_2)^2 + (z_1 - z_2)^2 = a^2 \\ (x_2 - x_3)^2 + (y_2 - y_3)^2 + (z_2 - z_3)^2 = b^2 \\ (x_1 - x_3)^2 + (y_1 - y_3)^2 + (z_1 - z_3)^2 = c^2. \end{cases}$$

Questo sistema può scriversi in forma vettoriale come un'equazione $f(x) = 0$ con $f : \mathbb{R}^9 \to \mathbb{R}^3$ (f è infinitamente derivabile). Si noti che ad ogni 9-upla di numeri reali che soddisfa queste tre equazioni corrisponde una configurazione del corpo rigido. Dunque lo spazio delle configurazioni del corpo rigido è identificabile con il sottoinsieme M di $\mathbb{R}^9$ tale che $f(x) = 0$.

Se si scrive la matrice Jacobiana di f si può constatare che, dall'ipotesi che p_1, p_2, p_3 siano punti non allineati, segue che la matrice ha in ogni punto rango massimo, cioè 3. (Si lascia questa verifica al lettore). Ne concludiamo che lo spazio M delle configurazioni di un corpo rigido è una varietà 6-dimensionale (in forma implicita), immersa in $\mathbb{R}^9$, di classe C^∞. Con linguaggio fisico, diciamo che il corpo rigido ha 6 gradi di libertà.

Esercizi

1° Si consideri il sistema:

$$\begin{cases} x + \log y + 5z - 10 = 0 \\ 2x + y^2 + 3z^3 - 25 = 0, \end{cases}$$

soddisfatto da $p_0 = (0, 1, 2)$. Mediante il teorema della funzione implicita per i sistemi, decidere quale coppia (o quali coppie) di variabili è esplicitabile in funzione della terza variabile, in un intorno di p_0. In ciascuno di questi casi, si calcoli il vettore derivato della funzione implicita, nel punto corrispondente a p_0.

2° Sia $F : \mathbb{R}^4 \to \mathbb{R}^2$ definita da:

$$F(x_1, x_2, y_1, y_2) = ((x_2 + 3)\, y_1 - \operatorname{tg}(y_1 + y_2) + 2x_1,\ \sin(y_1 + y_2) + 3x_2 - x_1(y_2 + 3)).$$

Dimostrare che $F(x_1, x_2, y_1, y_2) = 0$ definisce implicitamente

$$\begin{cases} y_1 = y_1(x_1, x_2) \\ y_2 = y_2(x_1, x_2) \end{cases}$$

in un intorno di $(0, 0, 0, 0)$.

§ 5 TRASFORMAZIONI DI COORDINATE E LORO INVERSIONE

5.1 Il teorema della funzione inversa

Abbiamo visto nel paragrafo 1.2 che una funzione $f : \mathbb{R}^n \to \mathbb{R}^n$ può vedersi come una trasformazione di coordinate nello spazio $\mathbb{R}^n$. Ci interessa stabilire delle condizioni sotto le quali questa trasformazione è invertibile (requisito necessario per poter utilizzare questa funzione come trasformazione di coordinate).

Sappiamo dall'algebra lineare che una funzione *lineare* $f : \mathbb{R}^n \to \mathbb{R}^n$ si rappresenta nella forma

$$f(x) = Ax,$$

per una certa matrice A, $n \times n$, e risulta invertibile se e solo se $\det A \neq 0$ (teorema di Cramer). In questo caso, anzi, è invertibile globalmente.

Nel caso di una funzione generica (ossia non lineare) $f : A \to \mathbb{R}^n$, tale che $f \in C^1(A)$ per qualche aperto $A \subseteq \mathbb{R}^n$, non potremo sperare in generale di avere un risultato di invertibilità globale, come vedremo con gli esempi. Ci si può aspettare però che, quando il differenziale di f in x_0 (che è la migliore approssimazione lineare

di f in un intorno di x_0) è invertibile, cioè quando la matrice Jacobiana di f calcolata in x_0 è invertibile, f risulti invertibile almeno in un intorno di x_0. È questo il contenuto del prossimo teorema, che come vedremo segue dal teorema sulla funzione implicita che abbiamo enunciato nel paragrafo 4.2.

Ricordiamo prima che:

Definizione 4.6 Una funzione $f : A \to B$ (A, B aperti di $\mathbb{R}^n$) si dice *invertibile in* A se è *iniettiva*, ossia se per ogni $x_1, x_2 \in A$,

$$f(x_1) = f(x_2) \Longrightarrow x_1 = x_2;$$

f si dice *suriettiva* su B se per ogni $y \in B$ esiste $x \in A$ tale che $f(x) = y$; si dice *biiettiva*, o *biunivoca* tra A e B se è iniettiva e suriettiva, ossia se per ogni $y \in B$ esiste uno e un solo $x \in A$ tale che $f(x) = y$.

In tal caso, la corrispondenza che associa ad ogni $y \in B$ questo elemento $x \in A$ univocamente determinato si dice *funzione inversa* di f. ∎

Notiamo che se scegliamo $B = f(A)$, la funzione f risulta automaticamente suriettiva; perciò, se f è iniettiva, la funzione inversa può sempre essere definita (con dominio $f(A)$).

Teorema 4.4 *Sia $f : A \subseteq \mathbb{R}^n \to \mathbb{R}^n$ con A aperto, tale che $f \in C^1(A)$. Supponiamo che per un dato punto $x_0 \in A$ sia*

$$\det Df(x_0) \neq 0. \tag{5.1}$$

Allora esistono un intorno U di x_0 e un intorno V di $f(x_0)$ tra i quali la funzione f è biunivoca; inoltre, detta $g : V \to U$ la corrispondenza inversa; si ha che $g \in C^1(V)$ e

$$Dg(f(x)) = Df(x)^{-1} \quad \textit{per ogni } x \in U.$$

Dimostrazione. Consideriamo la funzione

$$F : A \times \mathbb{R}^n \to \mathbb{R}^n$$

definita da

$$F(x, y) = f(x) - y.$$

Risulta $F \in C^1(A \times \mathbb{R}^n)$; posto $y_0 = f(x_0)$, si ha:

$$F(x_0, y_0) = 0 \text{ e } \det D_x F(x_0, y_0) \neq 0.$$

Perciò, per il teorema di Dini per i sistemi (teorema 4.3), esiste un intorno $V \subset \mathbb{R}^n$ di y_0 ed esiste un'unica funzione $g : V \to \mathbb{R}^n$ tale che $g \in C^1(V)$ e

$$F(g(y), y) = 0 \text{ per ogni } y \in V,$$

e cioè

$$f(g(y)) = y \text{ per ogni } y \in V.$$

Inoltre, per ogni $y \in V$,

$$Dg(y) = -D_x F(g(y), y)^{-1} D_y F(g(y), y) = Df(g(y))^{-1}. \tag{5.2}$$

 © 978-88-08-06485-1

Poniamo $U = g(V)$. L'insieme U contiene $x_0 = g(y_0)$ (per l'unicità di g). Proviamo che:

1. $f(U) \subseteq V$. Infatti, se $x \in U$, per definizione di U esiste $y \in V$ tale che $g(y) = x$ e poiché

$$f(x) = f(g(y)) = y$$

si ha $f(x) \in V$.

2. f è iniettiva in U. Infatti se $x_1, x_2 \in U$ e $f(x_1) = f(x_2)$, allora esistono $y_1, y_2 \in V$ tali che $g(y_i) = x_i$ per $i = 1, 2$. D'altro canto

$$f(x_1) = f(x_2) \Rightarrow f(g(y_1)) = f(g(y_2)) \Rightarrow$$
$$\Rightarrow y_1 = y_2 \Rightarrow g(y_1) = g(y_2) \Rightarrow x_1 = x_2.$$

3. f è suriettiva su V. Infatti se $y \in V$, posto $x = g(y)$ si ha che $x \in U$ per definizione di U e $f(x) = f(g(y)) = y$.

Concludiamo che la funzione f è biunivoca tra U e V, g è la funzione inversa di f ristretta ad U e vale la (5.2).

Osserviamo tuttavia che nulla garantisce che U sia un insieme *aperto* e quindi che sia un intorno di x_0 (mentre per definizione V è un intorno di y_0). La dimostrazione di questo fatto in effetti è un po' delicata e la omettiamo. Fatta salva questa precisazione, la dimostrazione è completa. ◊

Esempio

Si consideri la funzione $f : \mathbb{R}^2 \to \mathbb{R}^2$ data da:

$$f(x, y) = (e^x \cos y, e^x \sin y).$$

Chiaramente è $f \in C^1(\mathbb{R}^2)$. Calcoliamo la matrice Jacobiana:

$$Df(x, y) = \begin{pmatrix} e^x \cos y & -e^x \sin y \\ e^x \sin y & e^x \cos y \end{pmatrix}.$$

Si ha:

$$\det Df(x, y) = e^{2x} \neq 0 \text{ per ogni } (x, y) \in \mathbb{R}^2.$$

Il teorema di invertibilità locale garantisce quindi che per ogni punto $(x_0, y_0) \in \mathbb{R}^2$ esiste un intorno U di (x_0, y_0) in cui f è invertibile. Tuttavia, questo non consente di concludere che f sia invertibile globalmente; infatti, per esempio:

$$f(x, y + 2\pi) = f(x, y) \text{ per ogni } (x, y) \in \mathbb{R}^2,$$

perciò la funzione non è iniettiva in tutto $\mathbb{R}^2$.

OSSERVAZIONE (INVERTIBILITÀ LOCALE E GLOBALE) Riflettiamo ora sul caso unidimensionale. Se $f : (a, b) \to \mathbb{R}$, con $f \in C^1(a, b)$ e $f'(x) \neq 0$ per ogni $x \in (a, b)$, la funzione f' ha segno costante (se cambiasse di segno, per il teorema degli zeri dovrebbe anche annullarsi in almeno un punto, essendo una funzione continua); ma allora f è strettamente monotona in (a, b) e quindi è invertibile in tutto (a, b). In altre parole:

> nel caso unidimensionale, se le ipotesi del teorema di invertibilità locale sono soddisfatte in ogni punto di un intervallo aperto, la funzione è invertibile anche globalmente in quell'intervallo.

 © 978-88-08-06485-1

Invece (come mostra l'esempio precedente), in dimensione maggiore di 1, anche se le ipotesi del teorema di invertibilità locale sono soddisfatte in ogni punto del dominio, la funzione può non essere invertibile globalmente; tutto ciò che si può affermare in questo caso è che ogni punto ha un intorno in cui la funzione è invertibile, il che si esprime anche dicendo che la funzione in ogni punto è *localmente invertibile*.

Il teorema e l'osservazione precedente giustificano la prossima:

DEFINIZIONE 4.7 Sia A un aperto di $\mathbb{R}^n$. Una trasformazione di coordinate $f : A \to \mathbb{R}^n$ si dice *diffeomorfismo* (o anche diffeomorfismo globale) se $f \in C^1(A)$, f è globalmente invertibile in A e la sua funzione inversa $g : f(A) \to A$ è C^1 nel suo dominio $f(A)$; si dice *diffeomorfismo locale* se $f \in C^1(A)$ e ogni punto $x_0 \in A$ ha un intorno $U \subset A$ in cui f è invertibile, con inversa C^1. ▩

Il teorema di inversione locale garantisce che una trasformazione di coordinate $f : A \to \mathbb{R}^n$ tale che $f \in C^1(A)$ e $\det Df(x) \neq 0$ per ogni $x \in A$ è un diffeomorfismo locale in A. Se queste condizioni sono soddisfatte con l'eccezione di qualche punto di A, questi punti si diranno *punti singolari* della trasformazione.

Non ci sono condizioni altrettanto semplici, invece, che garantiscano che f sia un diffeomorfismo globale.

Riprendiamo ora gli esempi di trasformazioni di coordinate nel piano e nello spazio che abbiamo introdotto nel paragrafo 1.2 e verifichiamone la regolarità. Come vedremo, sotto certe limitazioni esse risultano effettivamente diffeomorfismi globali.

Coordinate polari nel piano

La trasformazione

$$\begin{cases} x = \rho \cos\theta \\ y = \rho \sin\theta \end{cases} \qquad \text{con } \rho \in [0, +\infty),\ \theta \in [0, 2\pi)$$

ha matrice Jacobiana:

$$\begin{pmatrix} \cos\theta & -\rho \sin\theta \\ \sin\theta & \rho \cos\theta \end{pmatrix}$$

con determinante ρ. La trasformazione è regolare, ad eccezione di $\rho = 0$, cioè l'origine nel piano xy. In effetti in questo punto viene a cadere la biunivocità locale: qualunque coppia $(0, \theta)$ nel piano $\rho\theta$ corrisponde al punto $(0, 0)$ nel piano xy. Se restringiamo la trasformazione

$$T(\rho, \theta) = (\rho \cos\theta, \rho \sin\theta)$$

all'aperto $A = (0, +\infty) \times (0, 2\pi)$, otteniamo un diffeomorfismo globale tra A e il piano $\mathbb{R}^2$ privato della semiretta $y = 0, x \geq 0$.

Coordinate cilindriche in $\mathbb{R}^3$

$$\begin{cases} x = \rho \cos\theta \\ y = \rho \sin\theta \\ z = t \end{cases} \qquad \text{con } \rho \in [0, +\infty),\ \theta \in [0, 2\pi),\ t \in \mathbb{R}.$$

 © 978-88-08-06485-1

Questa trasformazione di coordinate si può vedere come una funzione $\mathrm{f} : \mathbb{R}^3 \to \mathbb{R}^3$, con matrice Jacobiana

$$\begin{pmatrix} \cos\theta & -\rho\sin\theta & 0 \\ \sin\theta & \rho\cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

che ha determinante ρ. La trasformazione è regolare tranne che per $\rho = 0$, che questa volta corrisponde a una retta nello spazio xyz: l'asse z. In quei punti cade la biunivocità locale. Se restringiamo la trasformazione f all'aperto

$$A = (0, +\infty) \times (0, 2\pi) \times \mathbb{R},$$

otteniamo un diffeomorfismo globale tra A e lo spazio $\mathbb{R}^3$ privato del semipiano $y = 0, x \geq 0$.

Coordinate sferiche in $\mathbb{R}^3$

$$\begin{cases} x = \rho\sin\varphi\cos\theta \\ y = \rho\sin\varphi\sin\theta \\ z = \rho\cos\varphi \end{cases} \qquad \text{con } \rho > 0,\ \varphi \in [0, \pi],\ \theta \in [0, 2\pi).$$

Questa trasformazione di coordinate si può vedere come una funzione $\mathrm{f} : \mathbb{R}^3 \to \mathbb{R}^3$, con matrice Jacobiana:

$$\begin{pmatrix} \sin\varphi\cos\theta & \rho\cos\varphi\cos\theta & -\rho\sin\varphi\sin\theta \\ \sin\varphi\sin\theta & \rho\cos\varphi\sin\theta & \rho\sin\varphi\cos\theta \\ \cos\varphi & -\rho\sin\varphi & 0 \end{pmatrix}$$

il cui determinante è $\rho^2\sin\varphi$. La trasformazione è regolare tranne per $\rho = 0$ (origine nello spazio xyz) e per $\varphi = 0$, $\varphi = \pi$, che rappresentano i due semiassi z nello spazio xyz. In conclusione, sono singolari tutti i punti dell'asse z, nello spazio xyz. Se restringiamo la trasformazione f all'aperto:

$$A = (0, +\infty) \times (0, \pi) \times (0, 2\pi),$$

otteniamo un diffeomorfismo globale tra A e lo spazio $\mathbb{R}^3$ privato del semipiano $y = 0$, $x \geq 0$.

Invarianza della dimensione

Concludiamo questo paragrafo con la seguente osservazione. Il lettore potrebbe essersi chiesto come mai il problema dell'invertibilità locale di una funzione sia stato posto per trasformazioni di coordinate $\mathrm{f} : \mathbb{R}^n \to \mathbb{R}^n$ e non, più in generale, per funzioni $\mathrm{f} : \mathbb{R}^n \to \mathbb{R}^m$ con n e m diversi tra loro. È abbastanza intuitivo che non possano esserci funzioni regolari e biunivoche tra due spazi di dimensione diversa, ma si può dare una formulazione precisa a questo asserto e dimostrarlo? La risposta è affermativa ma non banale. Vale il seguente:

TEOREMA 4.5 (DI BROWER) *Se A è un aperto di $\mathbb{R}^n$ e $\mathrm{f} : A \to \mathbb{R}^n$ è continua e iniettiva in A, allora $\mathrm{f}(A)$ è un aperto di $\mathbb{R}^n$, f è biunivoca tra A e $\mathrm{f}(A)$ e f ha inversa continua.*

L'ipotesi di iniettività nel teorema precedente è fondamentale. Si osservi infatti che una funzione continua non trasforma aperti in aperti (né chiusi in chiusi), come mostra il seguente semplice:

Esempio

Sia $f : \mathbb{R} \to \mathbb{R}$ definita da:

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

Questa funzione è continua; non è iniettiva (perché è pari); $A = \mathbb{R}$ è aperto, tuttavia $f(\mathbb{R}) = (0, 1]$ non è aperto.

Non è difficile dimostrare (il lettore è invitato a provarci!) che dal teorema precedente segue anche il prossimo:

COROLLARIO 4.6 *Se A è un aperto di $\mathbb{R}^n$, B un aperto di $\mathbb{R}^m$ e $\mathrm{f} : A \to B$ è una funzione continua, biunivoca tra A e B, con inversa continua, allora $n = m$.*

Questo è il motivo per cui *non ci si pone neppure* il problema dell'invertibilità di una funzione regolare tra due spazi di dimensione diversa.

Si osservi che il teorema di Brower (come il suo corollario) riguarda funzioni *continue*, non necessariamente regolari; si tratta di un teorema di natura essenzialmente *topologica* (e non di calcolo differenziale), dalla dimostrazione profonda. Notiamo anche che, da questo teorema, il punto che nella nostra dimostrazione del teorema di invertibilità locale non è stato provato (ossia il fatto che $U = \mathrm{g}(V)$ sia aperto) segue subito.

5.2 Trasformazione di operatori differenziali

Come si trasforma un operatore differenziale eseguendo una trasformazione di coordinate? Sia, per esempio,

$$\begin{cases} x = g(u, v) \\ y = h(u, v) \end{cases}$$

una trasformazione regolare di coordinate nel piano e sia $f(x, y)$ una funzione differenziabile. Sotto l'azione della trasformazione di coordinate, f diventerà funzione di u, v e, per il teorema sul differenziale di una funzione composta, si può scrivere:

$$\frac{\partial f}{\partial u} = \frac{\partial f}{\partial x} \cdot \frac{\partial g}{\partial u} + \frac{\partial f}{\partial y} \cdot \frac{\partial h}{\partial u}$$

$$\frac{\partial f}{\partial v} = \frac{\partial f}{\partial x} \cdot \frac{\partial g}{\partial v} + \frac{\partial f}{\partial y} \cdot \frac{\partial h}{\partial v}$$

L'utilità delle trasformazioni di coordinate nello studio degli operatori differenziali è illustrato dal prossimo esempio e da alcuni degli esercizi successivi.

© 978-88-08-06485-1

Esempio

Consideriamo l'*equazione della corda vibrante*:

$$u_{tt} - c^2 u_{xx} = 0$$

con c costante. Si consideri ora la trasformazione di coordinate nel piano:

$$\begin{cases} \alpha = x + ct \\ \beta = x - ct. \end{cases}$$

Vogliamo esprimere in funzione di α, β gli operatori differenziali $\partial_x, \partial_t, \partial_{xx}, \partial_{tt}$ e riscrivere quindi l'equazione della corda vibrante rispetto alle variabili α, β. Ragionando come sopra si ha:

$$u_x = u_\alpha \cdot \frac{\partial \alpha}{\partial x} + u_\beta \cdot \frac{\partial \beta}{\partial x} = u_\alpha + u_\beta$$

$$u_t = u_\alpha \cdot \frac{\partial \alpha}{\partial t} + u_\beta \cdot \frac{\partial \beta}{\partial t} = c u_\alpha - c u_\beta$$

$$u_{xx} = \frac{\partial}{\partial x}(u_\alpha + u_\beta) = (\frac{\partial}{\partial \alpha} + \frac{\partial}{\partial \beta})(u_\alpha + u_\beta) = u_{\alpha\alpha} + 2u_{\alpha\beta} + u_{\beta\beta}$$

$$u_{tt} = \frac{\partial}{\partial t}(c u_\alpha - c u_\beta) = c^2(\frac{\partial}{\partial \alpha} - \frac{\partial}{\partial \beta})(u_\alpha - u_\beta) = c^2(u_{\alpha\alpha} - 2u_{\alpha\beta} + u_{\beta\beta}).$$

Di conseguenza:

$$u_{tt} - c^2 u_{xx} = c^2(u_{\alpha\alpha} - 2u_{\alpha\beta} + u_{\beta\beta}) - c^2(u_{\alpha\alpha} + 2u_{\alpha\beta} + u_{\beta\beta}) = -4u_{\alpha\beta}.$$

Perciò l'equazione della corda vibrante nelle nuove variabili α, β si scrive:

$$u_{\alpha\beta} = 0.$$

Quest'equazione può essere risolta facilmente. Infatti

$$\frac{\partial}{\partial \alpha}(\frac{\partial u}{\partial \beta}) = 0$$

significa che $\frac{\partial u}{\partial \beta}$ non dipende da α, cioè è una funzione di β soltanto:

$$\frac{\partial u}{\partial \beta} = f_1(\beta)$$

con f_1 funzione arbitraria. Integrando rispetto a β si trova:

$$u = f(\beta) + g(\alpha)$$

dove $f(\beta)$ è una primitiva di $f_1(\beta)$ (perciò, essendo f_1 arbitraria, è una funzione arbitraria) e $g(\alpha)$ è la "costante di integrazione", che non dipende da β, perciò è funzione (arbitraria) di α. Sostituendo infine ad α, β le loro espressioni in funzione di x, t, si trova:

$$u(x,t) = f(x - ct) + g(x + ct).$$

Questa rappresenta la più generale soluzione dell'equazione della corda vibrante, con f, g arbitrarie funzioni (due volte derivabili). La u così ottenuta è dunque la sovrapposizione di due onde, dette *onde progressive*, che viaggiano a velocità c in versi opposti, come mostra la figura 4.16.

Questo procedimento risolutivo dell'equazione della corda vibrante è dovuto a D'Alembert e risale al 1749.

© 978-88-08-06485-1

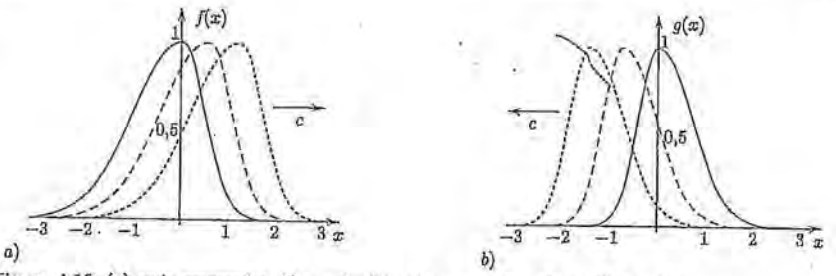


Figura 4.16. (a) onda progressiva $f(x - ct)$; (b) onda progressiva $g(x + ct)$ $(c > 0)$.

Esercizi

1. Verificare la regolarità della seguente trasformazione di coordinate, individuando gli eventuali punti singolari:

$$\begin{cases} x = uv \\ y = u^2 - v^2. \end{cases}$$

2. Si consideri la generica trasformazione di coordinate lineare affine nel piano:

$$\begin{cases} x = au + bv + c \\ y = du + ev + f \end{cases}$$

con a, b, c, d, e, f costanti. Scrivere sotto quali condizioni sui coefficienti la trasformazione è regolare.

3. Mostrare che l'espressione in coordinate polari ρ, θ dell'operatore differenziale

$$\Delta f = \frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2}$$

è data da:

$$\frac{\partial^2 f}{\partial \rho^2} + \frac{1}{\rho^2}\frac{\partial^2 f}{\partial \theta^2} + \frac{1}{\rho}\frac{\partial f}{\partial \rho}.$$

4. Osservando la relazione esistente tra coordinate polari nel piano e coordinate cilindriche nello spazio e osservando il risultato dell'esercizio precedente, riscrivere, senza eseguire calcoli, l'operatore di Laplace

$$\frac{\partial^2 f}{\partial x^2} + \frac{\partial^2 f}{\partial y^2} + \frac{\partial^2 f}{\partial z^2}$$

in coordinate cilindriche.

6 OTTIMIZZAZIONE. ESTREMI VINCOLATI

6.1 Vincoli di uguaglianza e moltiplicatori di Lagrange. Funzioni di due variabili

In molti problemi concreti di ottimizzazione, le variabili indipendenti sono soggette a vincoli di vario tipo. Alcuni esempi sono già stati dati nel capitolo 3, paragrafo 6.1. Introduciamo qui altri due esempi, tratti dalla meccanica classica e dalla microeconomia elementare, che dovrebbero dare un'ulteriore idea dell'importanza di questo tipo di problemi.

 © 978-88-08-06485-1

Esempi

Esempio *Equilibrio vincolato.* Supponiamo che un punto materiale sia vincolato *senza attrito* a una linea γ di equazione $g(x,y)=0$ e su di esso agisca una forza piana $\mathbf{F}$ conservativa[2], ossia esiste una funzione V (detta potenziale) tale che $\mathbf{F}=\nabla V$. Si vogliono determinare i punti di equilibrio stabile. Si dimostra che questo è equivalente a determinare i punti di minimo dell'energia potenziale $E=-V$, soggetti al vincolo $g(x,y)=0$; scriveremo sinteticamente

$$\begin{cases} \min E \\ \text{sub} \\ g(x,y)=0. \end{cases}$$

Esempio *Produzione sotto vincolo di budget.* Consideriamo un sistema microeconomico in cui i fattori di produzione siano il capitale K e il lavoro L. Sia $Y=Q(K,L)$ la funzione di produzione, dove K, L (positivi) sono misurati in unità di valuta corrente, per esempio in *euro*, mentre Y, quantità prodotta, è misurata in relazione al bene prodotto.

Un problema economicamente interessante è massimizzare Y soddisfacendo un vincolo sui fattori produttivi. È ragionevole, infatti, che l'azienda disponga di un assegnato budget $b>0$, con cui acquistare i servizi di capitale e lavoro. Se un'unità di capitale costa p_{cap} e una di lavoro costa p_{lav}, la spesa per l'acquisto della combinazione di fattori (K,L) risulta $Kp_{cap}+Lp_{lav}$.

Supponiamo che la spesa non possa eccedere il budget. Ci troviamo allora di fronte al problema di determinare quali coppie (K^*,L^*) massimizzano Y nel rispetto della condizione di budget $Kp_{cap}+Lp_{lav}\le b$.

È piuttosto ovvio che non conviene sottoutilizzare il budget, per cui una coppia ottima (K^*,L^*), che cioè massimizzi Y, rispetterà la condizione di budget: $K^*p_{cap}+L^*p_{lav}=b$. Il problema che interessa risolvere è quindi il seguente

$$\begin{cases} \max Y \\ \text{sub} \\ Kp_{cap}+Lp_{lav}=b. \end{cases}$$

Guidati dagli esempi precedenti, formalizziamo un problema di estremo vincolato, incominciando dal caso in cui la funzione *obiettivo* f (da ottimizzare) dipenda da due variabili legate da un'equazione di *vincolo*. Per semplicità ci limitiamo a considerare funzioni definite in tutto il piano.

Problema Date due funzioni $f=f(x,y)$ e $g=g(x,y)$, dotate di derivate parziali continue in $\mathbb{R}^2$, si vogliono determinare gli estremi di f sotto la condizione di vincolo $g(x,y)=b$, $b\in\mathbb{R}$.

Come sopra, scriviamo sinteticamente

$$\begin{cases} \max f \\ \text{sub} \\ g(x,y)=b \end{cases} \quad \text{oppure} \quad \begin{cases} \min f \\ \text{sub} \\ g(x,y)=b \end{cases}$$

secondo che si tratti di un problema di massimo o di minimo vincolato, rispettivamente.

[2]Questo concetto sarà illustrato in dettaglio nel capitolo 6. Per comprendere questo esempio, comunque, è sufficiente quanto spiegato qui.

© 978-88-08-06485-1

Il significato geometrico del problema (ad es. di *massimo* vincolato) è illustrato in figura 4.17. Si noti che, come esemplifica la figura riportata, un punto di massimo per $f(x,y)$ soggetto al vincolo $g(x,y)=b$ non è necessariamente un punto di massimo, neppure locale, per $f(x,y)$ come funzione in $\mathbb{R}^2$: non si tratta quindi semplicemente di determinare gli estremi liberi e poi vedere se qualcuno di essi cade sulla curva descritta dal vincolo, ma di massimizzare (minimizzare) la *restrizione di f al vincolo* (che è una funzione *diversa* da f).

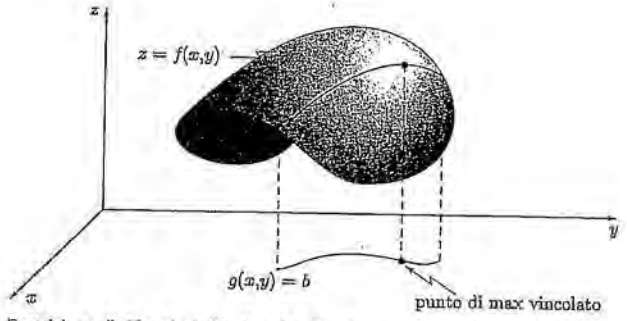


Figura 4.17. Restrizione di $f(x,y)$ al vincolo $g(x,y)=b$.

La situazione più favorevole è quella in cui l'equazione di vincolo $g(x,y)=b$ definisce esplicitamente $y=y(x)$ o $x=x(y)$ oppure, più in generale, definisce una curva γ di equazioni parametriche $x=x(t), y=y(t), t\in I$, dove I è un intervallo contenuto in $\mathbb{R}$.

Il problema è allora ricondotto alla ricerca degli estremi della funzione reale di variabile reale

$$\phi(t)=f(x(t),y(t))$$

nell'intervallo I. In questo caso si dice che *il vincolo è esplicitabile.*

Esempio

Esempio *Vincolo esplicitabile.* Riprendiamo il problema del massimo prodotto scegliendo come funzione di produzione la seguente (detta di *Cobb-Douglas*):

$$Y=aK^\alpha L^{1-\alpha} \qquad a>0 \quad 0<\alpha<1.$$

Il problema è dunque

$$\begin{cases} \max K^\alpha L^{1-\alpha} \\ \text{sub} \\ Kp_{\text{cap}}+Lp_{\text{lav}}=b. \end{cases} \qquad (b>0)$$

Ci si riconduce senza indugio a un problema di massimo libero per una funzione di una variabile, ricavando dal vincolo:

$$K=\frac{b-Lp_{\text{lav}}}{p_{\text{cap}}}$$

e riscrivendo Y come funzione della sola variabile L:

$$Y=h(L)=a\left(\frac{b-Lp_{\text{lav}}}{p_{\text{cap}}}\right)^\alpha L^{1-\alpha}=\frac{a}{p_{\text{cap}}^\alpha}(b-Lp_{\text{lav}})^\alpha L^{1-\alpha}.$$

 © 978-88-08-06485-1

Si tratta perciò di massimizzare la funzione $h(L)$ nell'intervallo $0 < L < b/p_{\text{lav}}$. Uguagliando a zero h', si trova l'equazione:

$$-\alpha p_{\text{lav}}(b - Lp_{\text{lav}})^{\alpha-1}L^{1-\alpha} + (1-\alpha)(b - Lp_{\text{lav}})^{\alpha}L^{-\alpha} = 0$$

ossia, semplificando,

$$-\alpha p_{\text{lav}}L + (1-\alpha)(b - Lp_{\text{lav}}) = 0$$

che dà, per L e successivamente per K, i valori di ottimo:

$$L^* = \frac{b(1-\alpha)}{p_{\text{lav}}}, \quad K^* = \frac{b\alpha}{p_{\text{cap}}}.$$

Questo è effettivamente un punto di massimo, come si vede facilmente dal cambiamento di segno di h' in L^*.

Al lettore non sarà sfuggito che il metodo seguito è praticabile solo in presenza di vincoli amatoriali, che si lasciano esplicitare rispetto all'una o all'altra variabile. Ciò accade raramente, soprattutto nel caso che tratteremo più avanti, in cui le variabili sono più di due, in presenza di più di un vincolo.

Presentiamo quindi un metodo più generale, noto come *metodo dei moltiplicatori di Lagrange*, che funziona quando il vincolo è rappresentato da una curva regolare assegnata in qualsiasi forma: parametrica, cartesiana esplicita o più in generale implicita.

Per illustrare il metodo, facciamo un'analogia con la strada percorsa nel caso dell'ottimizzazione libera: in quel caso, abbiamo anzitutto cercato di stabilire una *condizione necessaria, del prim'ordine*, affinché un punto (x^*, y^*) fosse di estremo. Tale condizione, espressa dal teorema di Fermat, indica che in un punto di estremo il gradiente della funzione obiettivo si deve annullare. Ciò è ragionevole poiché, in un problema di ottimizzazione libera, dal punto (x^*, y^*) ci si può muovere in ogni direzione pur di rimanere in un suo intorno. In altri termini, il teorema di Fermat afferma che le derivate lungo *ogni direzione ammissibile* si devono annullare. In un problema vincolato, tuttavia, se ci si muove da un punto di estremo lungo le direzioni parallele agli assi, in generale non si rispetta il vincolo e perciò le variazioni della funzione obiettivo lungo tali direzioni sono *non ammissibili*.

Quale direzione è ammissibile? Quali derivate direzionali devono annullarsi?

Sia (x^*, y^*) punto di estremo per la funzione f sotto il vincolo $g(x, y) = b$. Supponiamo che vicino a (x^*, y^*) il vincolo sia descritto da un arco di curva regolare, con retta tangente in (x^*, y^*) nella direzione del versore v. Ragioniamo come i pionieri del calcolo infinitesimale: se ci muoviamo in un intorno infinitesimale del punto (x^*, y^*), l'unica direzione ammissibile lungo cui muoversi è proprio la retta tangente, per cui la *derivata direzionale* di f nella direzione v dovrà annullarsi e cioè:

$$D_{\mathrm{v}} f(x^*, y^*) = 0.$$

Per la formula del gradiente, ciò significa:

$$\nabla f(x^*, y^*) \cdot \mathrm{v} = 0 \tag{6.1}$$

che esprime l'ortogonalità tra i vettori $\nabla f(x^*, y^*)$ e v.

Questo è l'analogo del teorema di Fermat. Nel prossimo teorema, formalizziamo il ragionamento euristico appena fatto e traduciamo la (6.1) in forma più operativa.

© 978-88-08-06485-1

TEOREMA 4.7 (MOLTIPLICATORI DI LAGRANGE) *Siano $f, g \in C^1(\mathbb{R}^2)$ e (x^*, y^*) punto di estremo vincolato per f sotto il vincolo*

$$g(x, y) = b.$$

Se (x^*, y^*) è regolare per il vincolo, cioè $\nabla g(x^*, y^*) \neq (0, 0)$, allora esiste $\lambda^* \in \mathbb{R}$ (detto moltiplicatore di Lagrange*) tale che:*

$$\nabla f(x^*, y^*) = \lambda^* \nabla g(x^*, y^*). \tag{6.2}$$

DIMOSTRAZIONE. Poiché (x^*, y^*) è punto di estremo vincolato, si ha in particolare $g(x^*, y^*) = b$. Poiché inoltre $\nabla g(x^*, y^*) \neq (0, 0)$, il teorema di Dini implica che in un intorno di (x^*, y^*) il vincolo $g(x, y) = b$ definisce un arco di curva regolare. Sia $x = x(t)$, $y = y(t)$ una parametrizzazione di tale arco, con $x(0) = x^*$ e $y(0) = y^*$. Il vettore

$$\mathrm{v} = (x'(0), y'(0))$$

è non nullo e tangente al vincolo.

Consideriamo la restrizione di f sull'arco di curva, data dalla funzione $\varphi(t) = f(x(t), y(t))$. Poiché $f \in C^1(\mathbb{R}^2)$ e l'arco di curva è regolare, la funzione φ è derivabile in un intorno di $t = 0$ e

$$\varphi'(t) = f_x(x(t), y(t))\ x'(t) + f_y(x(t), y(t))\ y'(t).$$

Ora, il fatto che f abbia un estremo vincolato in (x^*, y^*) equivale ad affermare che φ ha un minimo (libero!) in $t = 0$. Deve dunque essere $\varphi'(0) = 0$ ovvero

$$f_x(x^*, y^*)\ x'(0) + f_y(x^*, y^*)\ y'(0) = \nabla f(x^*, y^*) \cdot \mathrm{v} = 0.$$

Ricordando ora che anche il vettore $\nabla g(x^*, y^*)$ è ortogonale a v (vedi l'osservazione fatta sull'ortogonalità del gradiente con le curve di livello, capitolo 3, paragrafo 4.5), concludiamo che i *due vettori $\nabla f(x^*, y^*)$ e $\nabla g(x^*, y^*)$ devono essere paralleli, ossia che deve esistere un numero λ^** tale che

$$\nabla f(x^*, y^*) = \lambda^* \nabla g(x^*, y^*).$$ ◆

Ribadiamo che la (6.2) esprime il fatto che se (x^*, y^*) verifica le ipotesi del teorema, *allora la derivata di f lungo la tangente al vincolo si deve annullare* ed in tal caso diciamo che (x^*, y^*) è *punto critico vincolato*. È questa la corretta generalizzazione del teorema di Fermat al caso degli estremi vincolati.

OSSERVAZIONE Introducendo la funzione $\mathcal{L} = \mathcal{L}(x, y, \lambda)$, detta *Lagrangiana*, definita da:

$$\mathcal{L}(x, y, \lambda) = f(x, y) - \lambda[g(x, y) - b],$$

il teorema afferma che se (x^*, y^*) *è punto di estremo vincolato, allora esiste λ^* tale che il punto (x^*, y^*, λ^*) sia punto critico libero per $\mathcal{L}$. Infatti i punti critici di $\mathcal{L}$ sono soluzioni del sistema*:

$$\begin{cases} \mathcal{L}_x = f_x - \lambda g_x = 0 \\ \mathcal{L}_y = f_y - \lambda g_y = 0 \\ \mathcal{L}_\lambda = b - g = 0. \end{cases} \tag{6.3}$$

Le prime due equazioni coincidono con la (6.2) mentre la terza esprime la condizione di vincolo.

La teoria sviluppata indica il seguente modo di procedere, noto come *metodo dei moltiplicatori di Lagrange*.

 © 978-88-08-06485-1

a) Si isolano gli eventuali punti non regolari dell'insieme $g(x,y) = b$, che vanno esaminati a parte;

b) si cercano i punti critici liberi della Lagrangiana e cioè le soluzioni del sistema (6.3);

c) si determina la natura dei punti critici. A questo proposito risulta spesso utile il teorema di Weierstrass, come si vede nel seguente esempio.

Esempio

Si vogliono determinare gli estremi della funzione $f(x,y) = xy$, soggetti al vincolo[3]

$$g(x,y) = x^2 - xy + y^2 - 1 = 0. \tag{6.4}$$

Le funzioni f e g sono di classe $C^1(\mathbb{R}^2)$ e non vi sono punti singolari in $g(x,y) = 0$. La Lagrangiana è:

$$\mathcal{L}(x,y,\lambda) = xy - \lambda(x^2 - xy + y^2 - 1).$$

I punti critici sono soluzioni del seguente sistema:

$$\begin{cases} f_x(x,y) - \lambda g_x(x,y) = & y - 2\lambda x + \lambda y & = 0 \\ f_y(x,y) - \lambda g_y(x,y) = & x - 2\lambda y + \lambda x & = 0 \\ g(x,y) = & x^2 - xy + y^2 - 1 & = 0 \end{cases}$$

che ha come soluzioni:

$$(1,1,1);\quad (-1,-1,1);\quad \frac{1}{\sqrt{3}}(1,-1,-\frac{1}{\sqrt{3}});\quad \frac{1}{\sqrt{3}}(-1,1,-\frac{1}{\sqrt{3}}).$$

Dunque vi sono quattro punti critici vincolati: $p_0 = (1,1)$, $p_1 = (-1,-1)$, $p_2 = \frac{1}{\sqrt{3}}(1,-1)$, $p_3 = \frac{1}{\sqrt{3}}(-1,1)$.

Ora, essendo $E_0 = \{(x,y) \in \mathbb{R}^2 : g(x,y) = 0\}$ un insieme chiuso e limitato ed f continua, il teorema di Weierstrass[4] assicura l'esistenza di un punto di massimo e uno di minimo. Poiché $f(1,1) = f(-1,-1) = 1$ e $f(-1/\sqrt{3}, 1/\sqrt{3}) = f(1/\sqrt{3}, -1/\sqrt{3}) = -1/3$ si deduce che p_0 e p_1 sono punti di massimo globali e p_2, p_3 di minimo globali.

Il teorema 4.7 ha un'interessante interpretazione geometrica.

Introduciamo la famiglia degli insiemi di livello di f, definiti dall'equazione $f(x,y) = c$, $(c \in \mathbb{R})$.

Osserviamo che il vincolo corrisponde all'insieme di livello b di g. Sia ora (x^*, y^*) punto critico vincolato e $f(x^*, y^*) = c^*$ il livello critico di f corrispondente. Essendo per ipotesi $\nabla g(x_0, y_0) \neq (0,0)$, dalla (6.2) si deduce che:

- se $\lambda^* = 0$ allora $\nabla f(x_0, y_0) = (0,0)$ e il punto (x^*, y^*) è singolare per la linea di livello critica;
- se invece $\lambda^* \neq 0$, allora la linea di livello critica è tangente al vincolo nel punto (x^*, y^*).

[3] Anche se in questo caso il vincolo si presta a una rappresentazione parametrica globale, è forse più semplice ricorrere al metodo dei moltiplicatori di Lagrange.

[4] v. capitolo 3, paragrafo 3.2.

 © 978-88-08-06485-1

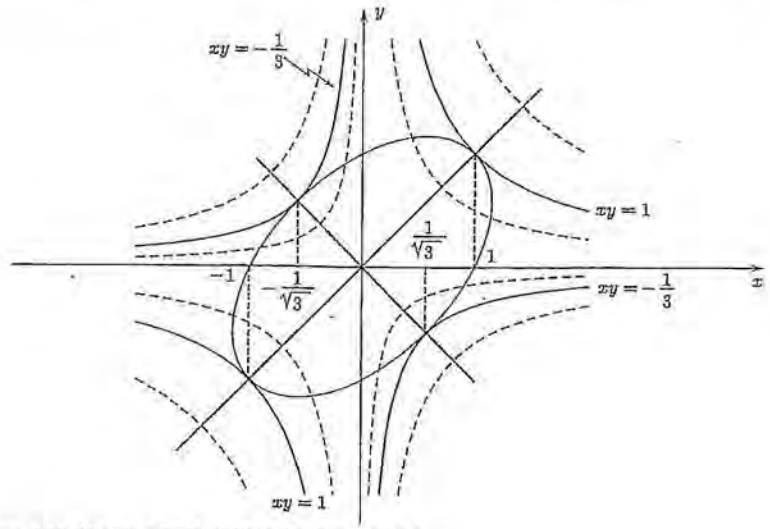


Figura 4.18. Vincolo e livelli critici di f nell'esempio 6.4.

La seconda alternativa è illustrata in figura 4.18, in riferimento al problema dell'esempio 6.4. Il vincolo (6.4) definisce un'ellisse centrata nell'origine, con assi nelle direzioni $y = x$ e $y = -x$ e diametri lunghi rispettivamente $2\sqrt{2}$ e $2\sqrt{2/3}$. Le linee di livello di $f(x,y) = xy = c$ sono iperboli equilatere di equazione $xy = c$ (due rami per ogni valore di c). I livelli $c = 1$ e $c = -1/3$, corrispondenti agli estremi condizionati, sono tangenti all'ellisse nei punti critici.

Significato economico del moltiplicatore. Prezzo ombra.

Il moltiplicatore λ^* non è un semplice elemento accessorio, ma ha un significato notevole, che vogliamo segnalare.

Un'informazione importante in un problema di ottimizzazione vincolata è la misura della sensibilità del valore estremo della funzione obiettivo (massimo o minimo) rispetto a una variazione di b che appare nell'equazione di vincolo. Pensando, per esempio, al caso della produzione ottima, b rappresenta il vincolo di budget ed è di indubbio interesse stabilire di quanto varia la produzione massima variando tale vincolo.

Supponiamo che, per ogni b in un intervallo $I = (b_1, b_2)$ si trovi un valore ottimo, per esempio di massimo, per la funzione f, sotto la condizione $g(x,y) = b$.

Per sottolineare la dipendenza da b, scriviamo:

$$M(b) = f(x^*(b), y^*(b)) \qquad \text{per il massimo vincolato di } f.$$

Vogliamo dimostrare che:

$$\lambda^*(b) = M'(b). \tag{6.5}$$

A tale scopo, scriviamo le condizioni del primo ordine:

$$f_x(x^*(b), y^*(b)) = \lambda^*(b) g_x(x^*(b), y^*(b)) \tag{6.6}$$

$$f_y(x^*(b), y^*(b)) = \lambda^*(b) g_y(x^*(b), y^*(b)) \tag{6.7}$$

$$g(x^*(b), y^*(b)) = b \tag{6.8}$$

per ogni b nell'intervallo I.

Deriviamo ora rispetto a b la (6.8); si trova:

$$g_x(x^*(b), y^*(b))\frac{dx^*}{db} + g_y(x^*(b), y^*(b))\frac{dy^*}{db} = 1.$$

Moltiplichiamo per $\lambda^*(b)$ entrambi i membri di questa equazione e usiamo le (6.6) e (6.7); si ottiene:

$$f_x(x^*(b), y^*(b))\frac{dx^*}{db} + f_y(x^*(b), y^*(b))\frac{dy^*}{db} = \lambda^*(b). \tag{6.9}$$

D'altra parte, abbiamo anche:

$$M'(b) = f_x(x^*(b), y^*(b))\frac{dx^*}{db} + f_y(x^*(b), y^*(b))\frac{dy^*}{db}$$

per cui dalla (6.9) si deduce la formula:

$$M'(b) = \lambda^*(b) .$$

o, in termini di differenziali,

$$dM = \lambda^* db. \tag{6.10}$$

Il moltiplicatore $\lambda^*(b)$ rappresenta dunque la *velocità di variazione del valore ottimo rispetto a b*.

In termini economici, la (6.10) indica di quanto aumenta la produzione ottima in conseguenza di un aumento Δb di budget: $\Delta M(b) \approx \lambda^* \Delta b$. Se $\Delta b = 1$ euro, λ^* rappresenta il valore del corrispondente aumento di quantità prodotta in condizioni di ottimalità e prende il nome di *prezzo ombra*.

Esempi

Esempio Ritorniamo al problema meccanico di *equilibrio vincolato* (v. Esempio 6.1). Usando il metodo dei moltiplicatori, si ottiene che nei punti di equilibrio (anche quelli non stabili, in realtà), deve esistere λ^* tale che:

$$\mathbf{F} = \nabla V = \lambda^* \nabla g. \tag{6.11}$$

La (6.11) indica che $\mathbf{F}$ deve essere diretta come la reazione vincolare, cioè ortogonale al vincolo, cosa che ha un evidente significato fisico: se la forza che agisce sul punto materiale ha una componente non nulla tangente al vincolo, direzione in cui il moto è possibile, la forza metterà in movimento il punto, che quindi non potrà essere in equilibrio.

Esempio Nel caso del *problema di produzione* (v. Esempio 6.2), nel punto di ottimo (K^*, L^*) si deve avere:

$$\nabla Q = \lambda^* \nabla g$$

equivalente a

$$\begin{cases} Q_K = \lambda^* g_K \\ Q_L = \lambda^* g_L. \end{cases}$$

Se $g_L \neq 0$, si deduce la relazione:

$$\frac{Q_K}{Q_L} = \frac{g_K}{g_L} = \lambda^*.$$

Per esempio, se $Q(K, L) = aK^\alpha L^{1-\alpha}$ $(a > 0, \quad 0 < \alpha < 1)$ e il vincolo di bilancio è $g(K, L) = Kp_{\text{cap}} + Lp_{\text{lav}} = b$, nel punto di ottimo si deve avere:

$$\frac{\alpha L^*}{(1-\alpha)K^*} = \frac{p_{\text{cap}}}{p_{\text{lav}}} = \lambda^*$$

che regola le quantità ottime di lavoro e capitale e il prezzo ombra della risorsa prodotta in funzione dei prezzi dei fattori di produzione.

Esercizi

Esercizio Determinare gli estremi di

$$f(x, y) = x^2 + 3y$$

con il vincolo $g(x, y) = \frac{x^2}{4} + \frac{y^2}{9} - 1 = 0$ sia esplicitando il vincolo, sia usando il metodo dei moltiplicatori di Lagrange.

Esercizio Determinare tutti i punti di massimo e minimo assoluti della funzione

$$f(x, y) = x^2 + y^2$$

soggetta al vincolo $x^{3/2} + y^{3/2} = 1$.

Esercizio Determinare tutti i punti di massimo e minimo assoluti della funzione

$$f(x, y) = x^2 y + \frac{1}{5} y^5$$

soggetta al vincolo $x^2 + y^4 = 1$.

Esercizio Un anello metallico si muove senza attrito lungo un filo piano γ di equazione $y = (x-1)^3$, sotto l'azione della forza (piana) centrale:

$$\mathbf{F}(x, y) = -\frac{x}{\sqrt{x^2+y^2}}\mathbf{i} - \frac{y}{\sqrt{x^2+y^2}}\mathbf{j}.$$

Quali sono i punti di equilibrio stabile per l'anello? (Suggerimento: seguendo il metodo visto nell'Esempio 6.1, cercare il minimo vincolato dell'energia potenziale E, con $\mathbf{F} = -\nabla E$; per far questo, individuare prima la funzione E).

Esercizio Un garage offre un servizio di automobili e autocarri, con funzione di produzione

$$Q = 15K^{1/3}L^{1/3}.$$

© 978-88-08-06485-1

Il costo unitario del lavoro è 6 euro, quello del capitale è 3 euro. Il budget previsto è 450 euro.

Usare il metodo dei moltiplicatori di Lagrange per massimizzare Q, col budget previsto.

Valutare se, con questi costi, il proprietario può pensare di aumentare il budget per avere un aumento del massimo di produzione del 10%.

Suggerimento: ricordare che $\Delta Q_{\max} \approx \lambda^* \Delta b$.

6.2 Moltiplicatori di Lagrange. Il caso generale

Consideriamo ora problemi di ottimizzazione vincolata, per funzioni di n variabili ($n \geq 2$) con un numero m di vincoli di uguaglianza.

Siano date $m+1$ funzioni reali di n variabili $f, g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ tutte di classe $C^1(\mathbb{R}^n)$. Si vogliono determinare gli estremi di f quando le variabili sono soggette alle m condizioni di vincolo:

$$\text{(6.12)} \qquad \begin{cases} g_1(x_1, x_2, \ldots, x_n) = b_1 \\ g_2(x_1, x_2, \ldots, x_n) = b_2 \\ \vdots \\ g_m(x_1, x_2, \ldots, x_n) = b_m. \end{cases}$$

Naturalmente, se i vincoli sono "indipendenti" la formulazione del problema è ragionevole solo se m, il numero di equazioni del vincolo, è *minore* di n, il numero di variabili. Già se fosse $m = n$, il sistema (6.12) è soddisfatto, in generale, al massimo da un numero finito di punti.

Se $\mathrm{x} = (x_1, x_2, \ldots, x_n)$, $\mathrm{b} = (b_1, b_2, \ldots, b_n)$ e raggruppiamo $g_1, g_2, \ldots, g_m$ nel vettore g, allora il problema di ottimizzazione vincolata si può compattare nella seguente formulazione:

$$\begin{cases} \max f \\ \text{sub} \\ \mathrm{g}(\mathrm{x}) = \mathrm{b}. \end{cases} \quad \text{oppure} \quad \begin{cases} \min f \\ \text{sub} \\ \mathrm{g}(\mathrm{x}) = \mathrm{b}. \end{cases}$$

Come già nel caso più semplice, almeno in linea di principio, la situazione più favorevole è quella in cui dal sistema (6.12) si possono ricavare esplicitamente m variabili in funzione delle altre $m-n$, per esempio:

$$\text{(6.13)} \qquad \begin{cases} x_1 = h_1(x_{m+1}, \ldots, x_n) \\ x_2 = h_2(x_{m+1}, \ldots, x_n) \\ \vdots \\ x_m = h_m(x_{m+1}, \ldots, x_n). \end{cases}$$

Ciò significa che (v. paragrafo 4.1) il sistema di vincoli definisce una *varietà di dimensione $n-m$ in forma parametrica* e il problema vincolato si riduce all'ottimizzazione libera della funzione:

$$F(x_{m+1}, \ldots, x_n) = f(h_1(x_{m+1}, \ldots, x_n), \ldots, h_m(x_{m+1}, \ldots, x_n), x_{m+1}, \ldots, x_n)$$

di $n-m$ variabili. Si capisce però che questa strada è il più delle volte impraticabile, per cui passiamo alla generalizzazione del metodo dei moltiplicatori di Lagrange.

© 978-88-08-06485-1

Ricordiamo che se $\mathrm{g}(\mathrm{x}^*) = \mathrm{b}$, x^* si dice punto *regolare* per g se la matrice Jacobiana

$$\mathrm{Dg} = \begin{pmatrix} D_{x_1} g_1 & D_{x_2} g_1 & \cdots & D_{x_n} g_1 \\ D_{x_1} g_2 & D_{x_2} g_2 & \cdots & D_{x_n} g_2 \\ \vdots & \vdots & \ddots & \vdots \\ D_{x_1} g_m & D_{x_2} g_m & \cdots & D_{x_n} g_m \end{pmatrix}$$

ha rango m in x^*. Equivalentemente, i vettori $\nabla g_j(\mathrm{x}^*)$, $j = 1, \ldots, m$, sono *linearmente indipendenti*. Esiste dunque un minore *non nullo* di ordine m estratto da $\mathrm{Dg}(\mathrm{x}^*)$. Se per esempio tale minore è

$$\det \begin{pmatrix} D_{x_1} g_1 & D_{x_2} g_1 & \cdots & D_{x_m} g_1 \\ D_{x_1} g_2 & D_{x_2} g_2 & \cdots & D_{x_m} g_2 \\ \vdots & \vdots & \ddots & \vdots \\ D_{x_1} g_m & D_{x_2} g_m & \cdots & D_{x_m} g_m \end{pmatrix},$$

il teorema sulle funzioni implicite (v. paragrafo 4.2) permette la rappresentazione (6.13) del sistema di vincoli, almeno in un intorno del punto x^*.

Dato dunque un punto regolare x^*, consideriamo un curva regolare in $\mathbb{R}^n$, assegnata da un vettore

$$\mathrm{r}(t) = (x_1(t), \ldots, x_n(t)),$$

che stia sulla varietà e passi per il punto x. Precisamente, sia:

(*i*) r definita in un intorno I di $t = 0$ e $\mathrm{r}(0) = \mathrm{x}^*$;

(*ii*) $\mathrm{r}'(t) \neq 0$ per ogni $t \in I$;

(*iii*) per ogni $t \in I$,

$$\mathrm{g}(\mathrm{r}(t)) = \mathrm{b}$$

o, più esplicitamente,

$$\text{(6.14)} \qquad g_j(x_1(t), \ldots, x_n(t)) = b_j \quad \text{per ogni } j = 1, \ldots, m.$$

Una curva con le proprietà indicate si costruisce facilmente, scegliendo prima $x_{m+1}(t)$, $\ldots, x_n(t)$ derivabili con continuità, in modo che $x_j(0) = x_j^*$ per $j = m+1, \ldots, n$ e che per almeno un indice j si abbia $x_j'(0) \neq 0$. Si pone poi:

$$x_j(t) = h_j(x_{m+1}(t), \ldots, x_n(t)) \quad \text{per ogni } j = 1, \ldots, m.$$

Vi sono almeno $n-m$ curve di questo tipo tali che i loro vettori velocità $\mathrm{v} = \mathrm{r}'(0)$ siano linearmente indipendenti (perché?). Osserviamo inoltre che, derivando la (6.14) e calcolando in $t = 0$, si ottiene:

$$\nabla g_j(\mathrm{r}(0)) \cdot \mathrm{v} = 0 \quad \text{per ogni } j = 1, \ldots, m$$

per cui v è *ortogonale a ognuno dei gradienti* $\nabla g_j(\mathrm{r}(0))$. Per questo motivo, le direzioni v si dicono *tangenziali ai vincoli*.

Possiamo ora generalizzare il metodo dei moltiplicatori al nostro caso.

© 978-88-08-06485-1

TEOREMA 4.8 (MOLTIPLICATORI DI LAGRANGE, CASO GENERALE) *Siano* $f, g_1, g_2, \dots, g_m \in C^1(\mathbb{R}^n)$, $\mathrm{g} = (g_1, g_2, \dots, g_m)$ *e sia* x^* *punto di estremo vincolato per* f *rispetto al vincolo*

$$\mathrm{g}(\mathrm{x}) = \mathrm{b}. \tag{6.15}$$

Se x^* *è un punto regolare per* g, *cioè se il rango di* $\mathrm{Dg}(\mathrm{x}^*)$ *è* m, *allora esistono* m *numeri reali* $\lambda_1^*, \dots, \lambda_m^*$, *detti moltiplicatori di Lagrange, tali che*

$$\nabla f(\mathrm{x}^*) = \sum_{j=1}^{m} \lambda_j^* \nabla g_j(\mathrm{x}^*). \tag{6.16}$$

DIMOSTRAZIONE. Sia r una curva soddisfacente le condizioni $(i), (ii), (iii)$ indicate prima dell'enunciato del teorema. Consideriamo la funzione

$$\varphi(t) = f(\mathrm{r}(t)),$$

restrizione di f a tale curva. Essendo $f \in C^1(\mathbb{R}^n)$ ed r regolare, φ risulta derivabile e

$$\varphi'(t) = \nabla f(\mathrm{r}(t)) \cdot \mathrm{r}'(t).$$

Poiché la curva soddisfa il vincolo e x^* è di estremo per f, segue che $t = 0$ è punto di estremo locale per φ e quindi, ponendo $\mathrm{v} = \mathrm{r}'(0)$, si ha:

$$\nabla f(\mathrm{x}^*) \cdot \mathrm{v} = 0. \tag{6.17}$$

Questa relazione vale per ogni curva regolare soddisfacente le condizioni $(i), (ii), (iii)$. D'altra parte, per la discussione che precede l'enunciato del teorema, sappiamo che v è ortogonale a *ognuno degli* m vettori $\nabla g_j(\mathrm{x}^*)$. Indicando con V lo spazio vettoriale generato dagli m vettori $\nabla g_j(\mathrm{x}^*)$ e con $V^\perp$ lo spazio ortogonale[5] a V, la (6.17) si può esprimere dicendo che:

$$\nabla f(\mathrm{x}^*) \text{ è ortogonale a v per ogni } \mathrm{v} \in V^\perp,$$

ossia $\nabla f(\mathrm{x}^*) \in (V^\perp)^\perp = V$.

Poiché V è generato dai vettori $\nabla g_j(\mathrm{x}^*)$, questo implica la (6.16), ossia $\nabla f(\mathrm{x}^*)$ si può scrivere come combinazione lineare dei vettori $\nabla g_j(\mathrm{x}^*)$ per $j = 1, 2, \dots, m$. ◆

Osserviamo che la (6.16) corrisponde a n equazioni scalari. Introduciamo la funzione Lagrangiana

$$\mathcal{L}(x_1, \dots x_n, \lambda_1, \dots, \lambda_m) = f(x_1, \dots x_n) - \sum_{j=1}^{m} \lambda_j [g_j(x_1, \dots x_n) - b_j]$$

dipendente da $n + m$ variabili reali. Il teorema 4.8 afferma che, se x^* è un punto regolare, allora esiste un vettore di moltiplicatori

$$\lambda^* = (\lambda_1^*, \dots, \lambda_m^*)$$

[5] Ricordiamo che, in $\mathbb{R}^n$, l'insieme dei vettori ortogonali a un sottospazio vettoriale V costituisce a sua volta uno spazio vettoriale, detto spazio ortogonale a V e indicato con $V^\perp$. Inoltre, vale la relazione $(V^\perp)^\perp = V$.

© 978-88-08-06485-1

tale che $(\mathrm{x}^*, \lambda^*)$ è un punto critico libero per la Lagrangiana. Infatti i punti critici liberi per la Lagrangiana si trovano risolvendo il seguente sistema di $n+m$ equazioni in $n+m$ incognite:

$$\begin{cases} D_{x_1}\mathcal{L} = D_{x_1} f - \sum_{j=1}^{m} \lambda_j D_{x_1} g_j = 0 \\ \vdots \\ D_{x_n}\mathcal{L} = D_{xn} f - \sum_{j=1}^{m} \lambda_j D_{x_n} g_j = 0. \\ D_{\lambda_1}\mathcal{L} = b_1 - g_1 = 0 \\ \vdots \\ D_{\lambda_m}\mathcal{L} = b_m - g_m = 0. \end{cases} \tag{6.18}$$

Le prime n equazioni equivalgono alla (6.16) mentre le ultime m esprimono le condizioni di vincolo.

Per risolvere un problema di ottimizzazione vincolata col metodo dei moltiplicatori di Lagrange, dopo aver isolato i punti singolari per il vincolo, si determinano i punti critici vincolati risolvendo il sistema (6.18).

Si cerca poi di determinare la natura dei punti critici mediante considerazioni legate al problema specifico. Esistono anche condizioni sufficienti ma sono, in generale, di difficile uso, per cui preferiamo soprassedere.

Esempi

Un problema isoperimetrico. Applichiamo il metodo dei moltiplicatori per massimizzare la funzione $V(x, y, z) = xyz$, sotto la condizione:

$$g(x, y, z) = xy + xz + yz = A \qquad A > 0. \tag{6.19}$$

Abbiamo incontrato questo problema nel capitolo 3, paragrafo 6.1: si tratta di massimizzare il volume di un parallelepipedo di lati x, y, z sotto il vincolo che la superficie laterale sia assegnata (uguale a $2A$). Dato il significato delle variabili, ci limiteremo ad ammettere solo $x \geq 0, y \geq 0, z \geq 0$.

Poiché $\nabla g = (y + z, x + z, x + y)$ si annulla solo nell'origine, la (6.19) rappresenta una superficie priva di punti singolari. Cerchiamo i punti critici risolvendo il sistema (6.18):

$$\begin{cases} V_x - \lambda g_x = yz - \lambda(y + z) = 0 \\ V_y - \lambda g_y = xz - \lambda(x + z) = 0 \\ V_z - \lambda g_z = xy - \lambda(x + y) = 0 \\ g = xy + xz + yz = A. \end{cases}$$

Si vede subito che deve essere $\lambda \neq 0$, essendo $A > 0$. Si deduce allora che deve anche essere $x \neq 0$, $y \neq 0$ e $z \neq 0$. Moltiplicando la prima equazione per x, la seconda per y, sottraendo membro a membro e semplificando per λ si trova:

$$x(y + z) = y(x + z),$$

ossia $x = y$.

Analogamente, moltiplicando la terza equazione per z, la seconda per y e sottraendo membro a membro e semplificando per λ si trova $z = y$. Quindi $x = y = z$. Usando la quarta equazione si trova l'unico punto a coordinate positive

$$x^* = y^* = z^* = \frac{\sqrt{A}}{2} \quad \text{e} \quad \lambda^* = \frac{\sqrt{A}}{8}$$

con

$$V\left(\frac{\sqrt{A}}{2},\frac{\sqrt{A}}{2},\frac{\sqrt{A}}{2}\right)=\frac{A^{3/2}}{8}.$$

Come facciamo a controllare che questo è il massimo volume ammissibile? Basterebbe assicurarsi che il massimo esista, avendo trovato un solo punto critico vincolato (notiamo che il minimo è zero, assunto sul bordo del dominio, $x=0$, o $y=0$, o $z=0$). Non si può però usare direttamente il teorema di Weierstrass in quanto il vincolo rappresenta una superficie chiusa ma illimitata.

Possiamo però pensare di ridurci a una parte di superficie contenuta nella sfera $B_R=\{x^2+y^2+z^2\le R^2\}$ con un raggio abbastanza grande. Se infatti, per esempio, si avesse $x>R$, poiché dall'equazione di vincolo si ottiene $xy<A$ e $xz<A$, che implicano $y<A/x$ e $z<A/x$, si avrebbe, per il volume:

$$V(x,y,z)=xyz<x\frac{A}{x}\frac{A}{x}<\frac{A^2}{R}.$$

Per $R>8\sqrt{A}$ si ha

$$\frac{A^2}{R}<\frac{A^{3/2}}{8}$$

e quindi, fuori dalla sfera B_R il volume è sempre minore di $A^{3/2}/8$. D'altra parte, in B_R, il massimo vincolato esiste per il teorema di Weierstrass e questo non può che essere $A^{3/2}/8$.

Ellissoidi e autovalori. In questo esempio, lo studio dei punti su un ellissoide a distanza massima e minima dal centro conduce alla comprensione di un importante significato degli autovalori di una matrice simmetrica positiva del terz'ordine.

Sia $\mathrm{A}=(a_{ij})$ una matrice quadrata, simmetrica e di ordine 3, definita positiva. Vogliamo determinare gli estremi di

$$f(\mathrm{x})=|\mathrm{x}|^2=\sum_{i=1}^{3}x_i^2$$

sotto il vincolo

$$g(\mathrm{x})=\langle \mathrm{Ax},\mathrm{x}\rangle=\sum_{i,j=1}^{3}a_{ij}x_ix_j=1$$

che definisce un ellissoide con centro nell'origine.

Essendo f continua e l'ellissoide un insieme chiuso e limitato in $\mathbb{R}^3$, massimo e minimo globali vincolati esistono.

Osserviamo che sull'ellissoide non vi sono punti singolari. Infatti:

$$g_{x_j}(\mathrm{x})=2\sum_{i=1}^{3}a_{ij}x_i \quad \text{e perciò} \quad \nabla g(\mathrm{x})=2\mathrm{Ax}.$$

Poiché $\det \mathrm{A}\neq 0$, $\nabla g(\mathrm{x})$ si annulla solo nell'origine, che non appartiene all'ellissoide.

Cerchiamo i punti critici vincolati. La Lagrangiana del problema è:

$$\mathcal{L}(\mathrm{x},\lambda)=|\mathrm{x}|^2-\lambda\left[\langle \mathrm{Ax},\mathrm{x}\rangle-1\right].$$

Il sistema (6.18) consiste, oltre al vincolo, dell'equazione (3 equazioni scalari)

$$\nabla_{\mathrm{x}}\mathcal{L}(\mathrm{x},\lambda)=2\mathrm{x}-2\lambda\mathrm{Ax}=0$$

ovvero, essendo senz'altro $\lambda\neq 0$ (altrimenti $\mathrm{x}=0$, che non soddisfa il vincolo),

$$\mathrm{Ax}=\frac{1}{\lambda}\mathrm{x}. \tag{6.20}$$

La (6.20) indica che *i punti critici vincolati sono gli autovettori di* A *che si trovano sull'ellissoide e i moltiplicatori di Lagrange sono i reciproci degli autovalori.*

Siano $0<\mu_1\le\mu_2\le\mu_3$ i reciproci degli autovalori di A e $\mathrm{w}_1,\mathrm{w}_2,\mathrm{w}_3$ tre autovettori corrispondenti che soddisfano il vincolo. Dalla (6.20) si ha:

$$\mathrm{Aw}_j=\frac{1}{\mu_j}\mathrm{w}_j \qquad j=1,2,3.$$

Moltiplicando scalarmente per w_j, otteniamo:

$$1=\langle \mathrm{Aw}_j,\mathrm{w}_j\rangle=\frac{1}{\mu_j}|\mathrm{w}_j|^2 \qquad j=1,2,3$$

che, in termini di f, si legge:

$$f(\mathrm{w}_j)=|\mathrm{w}_j|^2=\mu_j \qquad j=1,2,3.$$

Poiché già sappiamo che massimo e minimo vincolato di f esistono, possiamo trarre le seguenti importanti conclusioni, più significative se i tre autovalori sono distinti tra loro: il massimo di f è μ_3, il massimo moltiplicatore; il minimo di f è μ_1, il minimo moltiplicatore. Essendo f il quadrato della distanza dall'origine, deduciamo che $\sqrt{\mu_3}$ e $\sqrt{\mu_1}$ sono, rispettivamente, le lunghezze dei semiassi maggiore e minore dell'ellissoide.

Il moltiplicatore intermedio μ_2, se distinto dagli altri due, corrisponde a un *punto di sella vincolato*: partendo da w_2 e muovendosi verso w_3, f aumenta; muovendosi verso w_1, f diminuisce. La sua radice quadrata è la lunghezza del semiasse intermedio.

Esercizi

Trovare gli estremi vincolati di $f(x_1,\dots,x_n)=-\sum_{j=1}^{n}x_j\log x_j$ sotto la condizione $\sum_{j=1}^{n}x_j=1$.

Determinare i punti della superficie $z^2-xy=1$ più vicini all'origine.

Determinare gli estremi di $f(x,y,z)=(x+y+z)^2$ sotto la condizione $x^2+2y^2+3z^2=1$.

a) Determinare il minimo di $f(x_1,\dots,x_n)=\sum_{j=1}^{n}x_j$, $n\ge 2$, $x_j\ge 0$, sotto la condizione

$$x_1x_2\cdots x_n=1.$$

b) Dedurre la seguente disuguaglianza tra media geometrica e media aritmetica:

$$\sqrt{a_1a_2\cdots a_n}\le\frac{a_1+a_2+\cdots+a_n}{n} \qquad a_j\ge 0, \text{ per ogni } j=1,\dots,n.$$

Dimostrare che, tra tutti i triangoli di perimetro assegnato, quello con area massima è equilatero. Impostare il problema come problema di massimo vincolato, chiamando x,y,z i lati, $2p$ il perimetro (assegnato) ed esprimendo l'area mediante la formula di Erone:

$$A(x,y,z)=\sqrt{p(p-x)(p-y)(p-z)}.$$

6.3 Vincoli di disuguaglianza e teorema di Kuhn-Tucker

Considerare solo vincoli di uguaglianza costituisce a volte una seria limitazione. Specialmente nell'analisi economica ci si imbatte frequentemente nel problema di ottimizzare una funzione $f(\mathrm{x})$ quando la variabile x è soggetta a uno o più vincoli del tipo $g_j(\mathrm{x}) \leq b_j$, oltre alla condizione di non negatività delle sue componenti, cioè $x_j \geq 0$, $j = 1, \ldots, n$. È questo un tipico problema di programmazione matematica, alla quale sono dedicati numerosi testi specialistici. Per questa ragione, ci limitiamo ad accennare, senza dimostrazione, al risultato più importante della teoria, il teorema di Kuhn-Tucker, riferendoci, per fissare le idee, al problema di massimo seguente.

Problema Siano date $m+1$ funzioni reali di n variabili $f, g_1, g_2, \ldots, g_m : \mathbb{R}^n \to \mathbb{R}$ tutte di classe $C^1(\mathbb{R}^n)$. Si vuole massimizzare f quando le variabili sono soggette alla condizione di non negatività e alle m condizioni di vincolo:

$$\begin{cases} g_1(x_1, x_2, \ldots, x_n) \leq b_1 \\ g_2(x_1, x_2, \ldots, x_n) \leq b_2 \\ \vdots \\ g_m(x_1, x_2, \ldots, x_n) \leq b_m \end{cases}$$

Diciamo che un punto di massimo per f sotto i vincoli indicati *è soluzione del problema* 6.3.

Qui non c'è nessuna restrizione sul numero m di vincoli. Naturalmente supporremo che il sistema di vincoli non produca ... l'insieme vuoto!

Osservando che $\max f = -\min(-f)$, è chiaro che i risultati per il problema di massimo possono essere facilmente adattati ai problemi di minimo. Gli estremi si intendono globali.

Anche in questo caso abbiamo bisogno di introdurre condizioni di regolarità dei vincoli. In programmazione ve ne sono di vari tipi e si chiamano più propriamente *condizioni di qualificazione dei vincoli*. Precisamente, indichiamo con $J(\mathrm{x}^*)$ l'insieme degli indici $j \in \{1, \ldots, m\}$ tali che $g_j(\mathrm{x}^*) = b_j$. I vincoli corrispondenti si dicono *attivi in* x^*.

La più semplice tra le condizioni di qualificazione dei vincoli è la seguente.

DEFINIZIONE 4.8 Diciamo che i vincoli sono qualificati in x^* se i vettori $\nabla g_j(\mathrm{x}^*)$, per $j \in J(\mathrm{x}^*)$, sono linearmente indipendenti.

La presenza dei vincoli di non negatività delle variabili, cioè : $x_j \geq 0$, $j = 1, \ldots, n$, non altera la qualificazione o meno dei vincoli in un punto. Il seguente teorema è fondamentale in programmazione.

TEOREMA 4.9 (DI KUHN-TUCKER) *Sia* x^* *soluzione del problema* 6.3. *Se i vincoli sono qualificati in* x^*, *esiste un vettore di moltiplicatori*

$$\lambda^* = (\lambda_1^*, \ldots, \lambda_m^*) \in \mathbb{R}^m$$

(detti moltiplicatori di Kuhn-Tucker) tale che x^* *e* λ^* *sono soluzioni del seguente*

sistema di condizioni di ottimalità:

$$\text{(6.21)} \qquad \begin{cases} D_{x_k} f(\mathrm{x}^*) - \sum_{j=1}^m \lambda_j^* D_{x_k} g_j(\mathrm{x}^*) \leq 0, \\ x_k^* \geq 0, \qquad\qquad \textit{per ogni } k = 1, \ldots, n \\ x_k^* \left[D_{x_k} f(\mathrm{x}^*) - \sum_{j=1}^m \lambda_j^* D_{x_k} g_j(\mathrm{x}^*) \right] = 0 \end{cases}$$

$$\text{(6.22)} \qquad \begin{cases} b_i - g_i(\mathrm{x}^*) \geq 0, \\ \lambda_i^* \geq 0, \qquad\qquad \textit{per ogni } i = 1, \ldots, m \\ \lambda_i^* [b_i - g_i(\mathrm{x}^*)] = 0. \end{cases}$$

Si noti che le prime n condizioni (6.21) indicano che, se per un indice k si ha $x_k^* > 0$, allora deve essere:

$$D_{x_k} f(\mathrm{x}^*) = \sum_{j=1}^m \lambda_j^* D_{x_k} g_j(\mathrm{x}^*).$$

Dalle altre m condizioni (6.22) si deduce che, se un moltiplicatore λ_i^* è positivo, allora deve essere $b_i - g_i(\mathrm{x}^*) = 0$ e cioè questo vincolo è attivo. Ne segue che nelle prime n condizioni intervengono solo i vincoli attivi in x^*.

Introducendo la funzione Lagrangiana:

$$\mathcal{L}(\mathrm{x}^*, \lambda^*) = f(x_1, \ldots x_n) - \sum_{j=1}^m \lambda_j \left[g_j(x_1, \ldots x_n) - b_j \right],$$

le condizioni di ottimalità si possono scrivere nella forma più concisa seguente:

$$\begin{cases} \dfrac{\partial \mathcal{L}}{\partial x_k} \leq 0, \; x_k \geq 0, \; x_k \dfrac{\partial \mathcal{L}}{\partial x_k} = 0 \qquad k = 1, \ldots, n \\ \dfrac{\partial \mathcal{L}}{\partial \lambda_i} \geq 0, \; \lambda_i \geq 0, \; \lambda_i \dfrac{\partial \mathcal{L}}{\partial \lambda_i} = 0 \qquad i = 1, \ldots, n. \end{cases}$$

Infine, sottolineiamo che, se il problema fosse di minimo, cambierebbe solo il segno dei moltiplicatori, mentre la altre condizioni rimangono inalterate.

Esempio

Si voglia minimizzare la funzione $f(x_1, x_2) = (1 - x_1)^3 - x_2 - 1$, sotto i vincoli di non negatività delle variabili:

$$x_1 \geq 0 \text{ e } x_2 \geq 0.$$

Le condizioni di ottimalità sono:

$$\begin{cases} f_{x_1} = -3(1 - x_1)^2 \leq 0, \; x_1 \geq 0, \; x_1 \left[-3(1 - x_1)^2 \right] = 0 \\ f_{x_2} = -1 \leq 0, \qquad\qquad x_2 \geq 0, \; x_2 [-1] = 0. \end{cases}$$

Si trova subito $x_2 = 0$ e poi $x_1 = 0$ oppure $x_1 = 1$. Le condizioni di Kuhn-Tucker selezionano dunque i due punti $(0,0)$ e $(0,1)$ con $f(0,0) = 0$ e $f(0,1) = -1$. Osservando che $f \leq 0$ in tutto il primo quadrante, si deduce che $(0,0)$ è punto di minimo globale.

5 Calcolo integrale per funzioni di più variabili

Nel capitolo 2 abbiamo dato significato alla nozione di integrale di una funzione (continua) lungo un arco di curva (integrale di linea di prima specie): questo integrale generalizza in un certo senso la nozione di integrale definito per funzioni reali di una variabile reale. Vedremo ora altre generalizzazioni del concetto di integrale definito: anzitutto, in questo capitolo, introdurremo l'integrale per funzioni reali di più variabili, (tipicamente, due o tre), quando cioè l'insieme di integrazione ha dimensione maggiore di uno. Successivamente, nel capitolo 6, studieremo gli integrali (bidimensionali) estesi a una *superficie*, anziché a una regione del piano; questo ci condurrà al concetto di *flusso di un campo vettoriale* e a due importanti risultati di calcolo differenziale e integrale per funzioni di più variabili: i teoremi della divergenza e del rotore.

1 INTEGRALI DOPPI

1.1 Integrale di una funzione limitata definita su un rettangolo

Se $f : [a, b] \to \mathbb{R}$ è una funzione continua, ricordiamo che l'integrale

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

può vedersi come un *limite di somme di Cauchy-Riemann*:

$$\int_a^b f(x)dx = \lim_{n\to\infty} s_n \equiv \lim_{n\to\infty} \sum_{k=1}^{n} \frac{b-a}{n} f(t_k)$$

dove supponiamo di aver diviso $[a, b]$ in n intervallini di ampiezza $(b-a)/n$, t_k è un punto qualsiasi del k-esimo intervallino e il limite è fatto per $n \to \infty$, cioè all'infittirsi della partizione. Più in generale[1], per ogni funzione $f : [a, b] \to \mathbb{R}$ *limitata* è possibile definire le somme di Cauchy-Riemann s_n; si dice che la funzione f è integrabile quando esiste finito il limite della successione s_n e inoltre tale limite non dipende da come si

[1] v. volume 1, capitolo 6, paragrafo 2.

 © 978-88-08-06485-1

sono scelti i punti t_k, nei rispettivi intervallini, a ogni passo della costruzione iterativa. Si dimostra poi che le funzioni continue su $[a,b]$ risultano effettivamente integrabili.

L'idea di *integrale doppio per una funzione reale di due variabili* è una naturale estensione di questa definizione, almeno nel caso più semplice di una funzione definita su un rettangolo: se $f : [a,b] \times [c,d] \to \mathbb{R}$ è una funzione limitata, consideriamo una partizione di $[a,b]$ in n intervallini di ampiezza $(b-a)/n$, mediante $n+1$ punti x_h equispaziati:

$$x_h = a + h\frac{b-a}{n}, \text{ per } h = 0,1,2,\dots,n;$$
$$x_0 = a < x_1 < \cdots < x_n = b$$

e una analoga partizione di $[c,d]$ in n intervallini di ampiezza $(d-c)/n$ mediante $n+1$ punti y_k:

$$y_k = c + k\frac{d-c}{n}, \text{ per } k = 0,1,2,\dots,n;$$
$$y_0 = c < y_1 < \cdots < y_n = d.$$

Queste due partizioni inducono una partizione del rettangolo $[a,b] \times [c,d]$ in n^2 rettangolini

$$I_{hk} = [x_{h-1}, x_h] \times [y_{k-1}, y_k], \text{ per } h,k = 1,2,\dots,n,$$

ciascuno di area[2]

$$|I_{hk}| = \frac{(b-a)(d-c)}{n^2}.$$

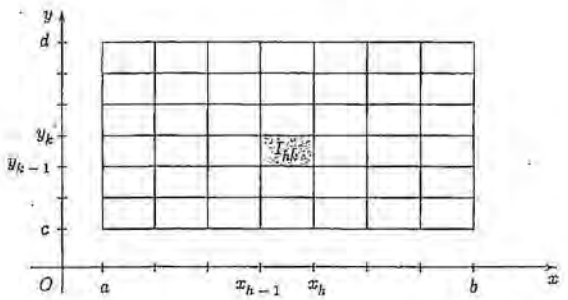


Figura 5.1.

Sia $\mathrm{p}_{hk} = (x_{hk}, y_{hk}) \in I_{hk}$ un punto qualsiasi scelto in ciascun rettangolino e consideriamo la somma di Cauchy-Riemann

$$s_n = \sum_{h,k=1}^{n} |I_{hk}| f(\mathrm{p}_{hk}). \tag{1.1}$$

[2] Qui e nel seguito indicheremo col simbolo $|A|$ l'area dell'insieme $A \subset \mathbb{R}^2$.

© 978-88-08-06485-1

Osserviamo il significato geometrico di s_n (per semplicità, nel caso $f \geq 0$): ogni addendo è il prodotto dell'area del rettangolo I_{hk} per il valore $f(\mathrm{p}_{hk})$, quindi rappresenta il volume del parallelepipedo di base I_{hk} e altezza $f(\mathrm{p}_{hk})$; la somma s_n è quindi il volume di una certa regione tridimensionale che, all'infittirsi della partizione, ci aspettiamo approssimi sempre meglio la regione tridimensionale che sta sotto il grafico di f (e sopra il piano xy). Possiamo ora dare la seguente:

Definizione 5.1 Diremo che la funzione $f : [a,b] \times [c,d] \to \mathbb{R}$, limitata, è *integrabile nel rettangolo* $R = [a,b] \times [c,d]$ se esiste finito $\lim_{n\to\infty} s_n$, con s_n come in (1.1) e se inoltre tale limite non dipende da come si sono scelti i punti p_{hk} nei rispettivi rettangolini I_{hk}, a ogni passo della costruzione. In tal caso questo limite si dirà *integrale doppio* di f sul rettangolo R e sarà indicato col simbolo:

$$\iint_R f(x,y)\,dx\,dy = \lim_{n\to\infty} \sum_{h,k=1}^{n} |I_{hk}| f(\mathrm{p}_{hk}). \tag{1.2}$$

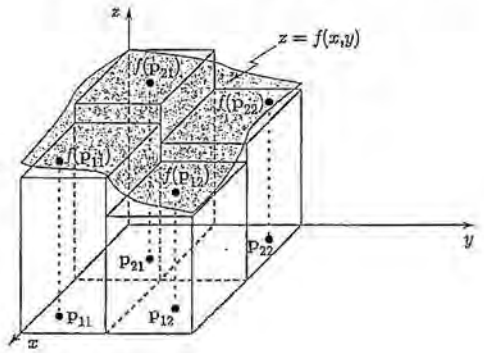


Figura 5.2.

Osserviamo che nel simbolo di integrale doppio le variabili d'integrazione x, y sono *variabili mute*; questo significa che, per esempio, si ha:

$$\iint_R f(x,y)\,dx\,dy = \iint_R f(x,v)\,dx\,dv = \iint_R f(u,v)\,du\,dv.$$

Come per le funzioni di una variabile, la definizione data lascia aperto il problema di capire quali funzioni risultino effettivamente integrabili.

Il fatto che, senza qualche ipotesi di continuità, una funzione possa non essere integrabile, si vede facilmente con esempi del seguente tipo:

Esempio

Una funzione non integrabile Sia $f : [0,1] \times [0,1] \to \mathbb{R}$ così definita:

$$f(x,y) = \begin{cases} 1 \text{ se } x \in \mathbb{Q} \\ 0 \text{ se } x \in \mathbb{R} \setminus \mathbb{Q}. \end{cases}$$

È facile convincersi che questa funzione non è integrabile; infatti ad ogni passo della costruzione, è possibile scegliere i punti p_{hk} in modo che la somma di Cauchy-Riemann s_n valga

 © 978-88-08-06485-1

0 oppure 1 (o altri valori intemedi): il limite di s_n dunque dipende dalla scelta dei punti, perciò l'integrale non esiste.

Vale invece il prossimo risultato positivo:

TEOREMA 5.1 *Se $f : [a,b] \times [c,d] \to \mathbb{R}$ è continua, allora è integrabile.*

La dimostrazione di questo teorema è delicata e viene qui omessa; essa ricalca quella del teorema analogo per funzioni di una variabile[3], basandosi in particolare sulla proprietà di *uniforme continuità* di una funzione continua su un insieme chiuso e limitato[4].

Per una funzione continua e non negativa su un rettangolo, l'integrale doppio ha il significato geometrico di volume della regione tridimensionale compresa tra il piano xy e il grafico della funzione; più precisamente, l'integrale doppio può costituire una *definizione rigorosa* di volume della regione in questione. Si noti, infatti, che non è affatto elementare, dal punto di vista geometrico, definire il volume di una regione tridimensionale a contorni curvilinei: questo richiede, comunque si proceda, qualche tipo di procedimento infinito; il calcolo integrale offre una possibile strada in questa direzione.

A questo punto si apre il problema del calcolo effettivo degli integrali doppi. Anziché cercare di generalizzare il teorema fondamentale del calcolo integrale, si rivela più utile un procedimento di *riduzione degli integrali doppi a integrali iterati*, cioè al calcolo di successivi integrali in una variabile.

Per "indovinare" una formula utile, lasciamoci guidare dal significato geometrico dell'integrale doppio come volume sotteso al grafico di $f(x,y)$.

Sia $R = [a,b] \times [c,d]$; tracciamo un piano verticale parallelo all'asse x; questo taglierà il solido di cui vogliamo calcolare il volume lungo una regione piana di area $A(y)$.

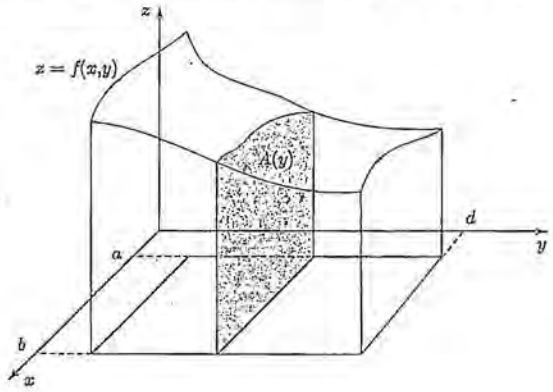


Figura 5.3.

[3] v. volume 1, Capitolo 6, Complementi.
[4] v. capitolo 3, Complementi, paragrafo 9.1.

© 978-88-08-06485-1

La "fetta" del solido che ha per base la striscia $(y, y+dy)$ avrà volume $A(y)dy$. Il volume totale si può ricostruire facendo variare y tra c e d e poi "sommando" i contributi delle singole fette. La somma totale non è altro che:

$$V = \int_c^d A(y)\,dy.$$

A sua volta, come si calcola l'area $A(y)$?

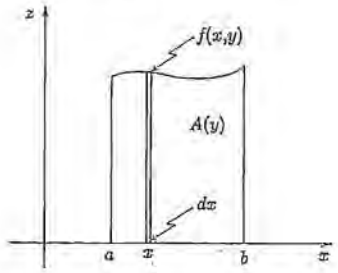


Figura 5.4.

Possiamo calcolarla per mezzo dell'integrale unidimensionale della funzione $x \mapsto f(x,y)$ con y fissato:

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

Ne segue allora

$$V = \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy$$

Tutto ciò è intuitivo e può essere enunciato in un risultato preciso:

TEOREMA 5.2 (DI RIDUZIONE, PER UN RETTANGOLO) *Se $f : [a,b] \times [c,d] \to \mathbb{R}$ è continua, allora il suo integrale doppio si può calcolare come integrale iterato, al modo seguente:*

$$\begin{aligned} \int\int_{[a,b]\times[c,d]} f(x,y)\,dxdy &= \int_a^b \left(\int_c^d f(x,y)\,dy \right) dx = \\ &= \int_c^d \left(\int_a^b f(x,y)\,dx \right) dy. \end{aligned} \tag{1.3}$$

Il teorema precedente si potrebbe enunciare e dimostrare per una funzione integrabile qualsiasi, con qualche precisazione, ma per il momento ci limitiamo al caso di funzioni continue. Per la dimostrazione, premettiamo il seguente:

LEMMA 5.3 *Sia $f : [a,b] \times [c,d] \to \mathbb{R}$ continua. Allora la funzione di una variabile:*

$$\varphi(x) = \int_c^d f(x,y)\,dy$$

 © 978-88-08-06485-1

è continua in $[a, b]$. *Analogamente, la funzione di una variabile*

$$\psi(y) = \int_a^b f(x, y)\, dx$$

è continua in $[c, d]$.

Questo lemma afferma una proprietà piuttosto intuitiva; tralasciamo tuttavia la sua dimostrazione, che si basa (come il teorema di integrabilità delle funzioni continue) sull'uniforme continuità di f sul rettangolo R.

DIMOSTRAZIONE DEL TEOREMA 5.2. Proviamo la prima uguaglianza in (1.3). In modo analogo si può provare che l'integrale doppio è uguale all'altro integrale iterato.

Vogliamo applicare la definizione di integrale, dunque consideriamo una partizione del rettangolo $[a, b] \times [c, d]$ in n^2 rettangolini I_{kh}. (Useremo le stesse notazioni introdotte nella definizione di integrale).

Consideriamo la funzione di una variabile:

$$\varphi(x) = \int_c^d f(x, y)\, dy,$$

continua in $[a, b]$ per il Lemma precedente, e quindi integrabile[5] in $[a, b]$. Vogliamo mostrare che:

$$\int_a^b \varphi(x)\, dx = \int\int_{[a,b]\times[c,d]} f(x, y)\, dxdy,$$

che è appunto la tesi. Possiamo scrivere:

$$\int_a^b \varphi(x)\, dx = \sum_{h=1}^{n} \int_{x_{h-1}}^{x_h} \varphi(x)\, dx.$$

Ora applichiamo a φ, su ciascun intervallino $[x_{h-1}, x_h]$, il Teorema della media[6]: per ogni h esiste un punto

$$x_h^* \in [x_{h-1}, x_h]$$

tale che:

$$\int_{x_{h-1}}^{x_h} \varphi(x)\, dx = \varphi(x_h^*)(x_h - x_{h-1}).$$

Dunque:

$$\int_a^b \varphi(x)\, dx = \sum_{h=1}^{n} \varphi(x_h^*) \frac{(b-a)}{n} = \sum_{h=1}^{n} \frac{(b-a)}{n} \int_c^d f(x_h^*, y)\, dy. \tag{1.4}$$

Consideriamo ciascun integrale:

$$\int_c^d f(x_h^*, y)\, dy = \sum_{k=1}^{n} \int_{y_{k-1}}^{y_k} f(x_h^*, y)\, dy.$$

Per ogni k fissato, la funzione di una variabile

$$y \longmapsto f(x_h^*, y)$$

[5] v. volume 1, capitolo 6, paragrafo 2.2.
[6] v. volume 1, capitolo 6, paragrafo 3, teorema 6.5.

© 978-88-08-06485-1

è continua in $[c, d]$. Applicando a ciascuna di queste funzioni ancora il teorema della media su ciascun intervallino $[y_{k-1}, y_k]$, otteniamo che esiste, per ogni h e per ogni k, un punto

$$y_{hk}^* \in [y_{k-1}, y_k]$$

tale che:

$$\int_{y_{k-1}}^{y_k} f(x_h^*, y)\, dy = f(x_h^*, y_{hk}^*)(y_k - y_{k-1}).$$

Pertanto:

$$\int_c^d f(x_h^*, y)\, dy = \sum_{k=1}^{n} \frac{(d-c)}{n} f(x_h^*, y_{hk}^*)$$

e, sostituendo nella (1.4),

$$\int_a^b \left(\int_c^d f(x, y)\, dy \right) dx = \int_a^b \varphi(x)\, dx =$$

$$= \sum_{h=1}^{n} \frac{(b-a)}{n} \sum_{k=1}^{n} \frac{(d-c)}{n} f(x_h^*, y_{hk}^*) = \sum_{k,h=1}^{n} |I_{hk}|\, f(\mathrm{p}_{hk})$$

con $\mathrm{p}_{hk} = (x_h^*, y_{hk}^*)$. Abbiamo quindi dimostrato che per ogni n esiste una somma di Cauchy-Riemann

$$s_n = \sum_{k,h=1}^{n} |I_{hk}|\, f(\mathrm{p}_{hk}),$$

relativa all'integrale doppio $\int\int_{[a,b]\times[c,d]} f(x, y)\, dxdy$, tale che:

$$s_n = \int_a^b \left(\int_c^d f(x, y)\, dy \right) dx.$$

Poiché già sappiamo che la funzione f è integrabile (in quanto continua), passando al limite per $n \to +\infty$ si ha che:

$$\lim_{n\to\infty} s_n = \int\int_{[a,b]\times[c,d]} f(x, y)\, dxdy,$$

dunque integrale doppio e integrale iterato coincidono. ∎

La tecnica di iterazione permette di ricondurre il calcolo di un integrale doppio a quello di due integrali semplici (cioè in una variabile), uno in sequenza all'altro. Si osservi che mentre si calcola il *primo integrale*, cioè quello più *interno*, la variabile diversa da quella rispetto a cui si integra va trattata come una *costante*.

Esempio

Si voglia calcolare:

$$\int\int_{[0,1]\times[0,2]} xe^{xy}\, dxdy.$$

La funzione è continua sul rettangolo; calcoliamo l'integrale doppio come integrale iterato al modo seguente:

$$\int_0^1 \left(\int_0^2 xe^{xy}\, dy \right) dx.$$

 © 978-88-08-06485-1

Questo significa che prima calcoliamo l'integrale interno nella variabile y, pensando x come parametro fissato:

$$\int_0^2 xe^{xy}dy = [e^{xy}]_0^2 = e^{2x} - 1;$$

quindi sostituiamo il risultato ottenuto (che è una funzione della sola x) nell'integrale esterno, calcolando:

$$\int_0^1 (e^{2x} - 1)\,dx = \left[\frac{e^{2x}}{2} - x\right]_0^1 = \frac{e^2}{2} - 2.$$

Questo è il valore dell'integrale doppio. Avremmo potuto procedere anche nell'altro ordine, calcolando:

$$\int_0^2 \left(\int_0^1 xe^{xy}dx\right) dy.$$

Lo studente è invitato a rifare il calcolo a questo modo. Si trova ovviamente lo stesso risultato, ma non necessariamente con la stessa facilità!

OSSERVAZIONE (FUNZIONI A VARIABILI SEPARATE) Quando si integra su un rettangolo una funzione del tipo $f(x)g(y)$, l'integrale doppio si calcola come prodotto di due integrali unidimensionali. Infatti:

$$\begin{aligned}\iint_{[a,b]\times[c,d]} f(x)g(y)dx\,dy &= \int_a^b \left(\int_c^d f(x)g(y)dy\right) dx = \\ &= \int_a^b f(x)\left(\int_c^d g(y)dy\right) dx = \\ &= \left(\int_c^d g(y)dy\right)\left(\int_a^b f(x)dx\right)\end{aligned}$$

Per esempio

$$\iint_{[0,1]\times[0,2]} x^3e^y dx\,dy = \left(\int_0^1 x^3dx\right)\left(\int_0^2 e^y dy\right) = \frac{1}{4}(e^2 - 1).$$

1.2 Funzioni integrabili su domini non rettangolari. Insiemi semplici, regolari, misurabili

Naturalmente il caso in cui la funzione integranda sia definita su un rettangolo è un caso particolare, che abbiamo scelto come punto di partenza. Se f è definita su un insieme $\Omega \subset \mathbb{R}^2$, limitato, come si può definire l'integrale di f? Un'idea naturale è quella di considerare un rettangolo $[a,b] \times [c,d]$ che contiene Ω, definire f in tutto il rettangolo ponendola per definizione uguale a zero fuori di Ω e integrare sul rettangolo la funzione così ottenuta (applicando la definizione precedente). Se tale integrale esiste, diremo che f è integrabile in Ω. Detto più formalmente:

DEFINIZIONE 5.2 Sia $f : \Omega \to \mathbb{R}$, con $\Omega \subset \mathbb{R}^2$ limitato, sia $R = [a,b] \times [c,d]$ un rettangolo contenente Ω (fig. 5.5) e sia:

$\widetilde{f} : R \to \mathbb{R}$ così definita:

$$\widetilde{f}(x,y) = \begin{cases} f(x,y) & \text{se } (x,y) \in \Omega \\ 0 & \text{se } (x,y) \in R \setminus \Omega. \end{cases}$$

© 978-88-08-06485-1

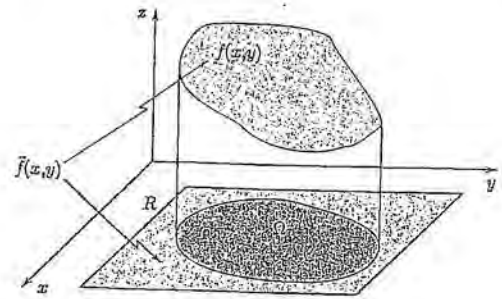


Figura 5.5.

Se $\widetilde{f}$ risulta integrabile in R (in base alla definizione 5.1), diremo che f è integrabile in Ω e porremo

$$\iint_\Omega f(x,y)\,dxdy = \iint_R \widetilde{f}(x,y)\,dxdy.$$

OSSERVAZIONE Nella definizione precedente il rettangolo R non è individuato univocamente; tuttavia si può dimostrare che l'integrabilità della funzione e il valore dell'integrale non dipendono dalla scelta di R.

Si noti che, anche se f è continua in Ω, in generale $\widetilde{f}$ risulta discontinua in R, perché presenta un "salto" lungo il bordo di Ω. Pertanto senza fare qualche ipotesi sul dominio Ω non è ovvia neppure l'integrabilità delle funzioni continue e limitate in Ω. Si consideri in proposito il prossimo:

Esempio

Sia Ω l'insieme dei punti del quadrato $[0,1] \times [0,1]$ con coordinata x razionale, cioè

$$\Omega = \{(x,y) \in [0,1] \times [0,1] : x \in \mathbb{Q}\}.$$

Vogliamo controllare se esiste

$$\iint_\Omega 1\,dx\,dy. \tag{1.5}$$

Si noti che l'integranda è costante in Ω! Se estendiamo la funzione integranda a tutto il quadrato $[0,1] \times [0,1]$ ponendola uguale a zero fuori da Ω, otteniamo la stessa funzione considerata nell'Esempio 1.1, che già sappiamo non essere integrabile. Pertanto *l'integrale* (1.5) *non esiste.*

Si capisce che questo fatto è dovuto alle cattive proprietà del dominio Ω, non certo della funzione integranda (costante!)

Vogliamo quindi individuare delle condizioni sul dominio Ω che siano facili da esprimersi e sufficienti a garantire l'integrabilità di una funzione continua e limitata in Ω.

Restringiamo il discorso a funzioni definite su *particolari* insiemi, che saranno comunque più che sufficienti per la maggior parte delle applicazioni.

 © 978-88-08-06485-1

DEFINIZIONE 5.3 (INSIEMI SEMPLICI, REGOLARI) Un insieme $E \subset \mathbb{R}^2$ si dice *y-semplice* se è del tipo

$$E = \{(x,y) \in \mathbb{R}^2 : x \in [a,b], g_1(x) \le y \le g_2(x)\}$$

con $g_1, g_2 : [a,b] \to \mathbb{R}$ funzioni continue. Questo significa che tagliando E con una retta parallela all'asse y si ottiene sempre un segmento e che questo segmento varia con continuità al variare della retta.

Analogamente, E si dirà *x-semplice* se è del tipo:

$$E = \{(x,y) \in \mathbb{R}^2 : y \in [c,d], h_1(y) \le x \le h_2(y)\}$$

con $h_1, h_2 : [c,d] \to \mathbb{R}$ funzioni continue.

E si dirà *semplice* se è x-semplice o y-semplice; si dirà *regolare* se è unione di un numero finito di insiemi semplici.

Per gli obiettivi di questo corso, non ci interesserà mai considerare insiemi che *non siano* regolari.

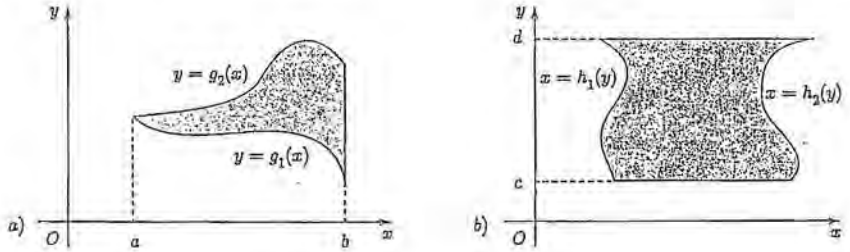


Figura 5.6. a) Regione semplice relativamente all'asse y; b) regione semplice relativamente all'asse x.

ESEMPI

L'insieme:

$$E = \{(x,y) \in \mathbb{R}^2 : 0 \le x \le 1, x(1-x) \le y \le x(2-x^2)\}$$

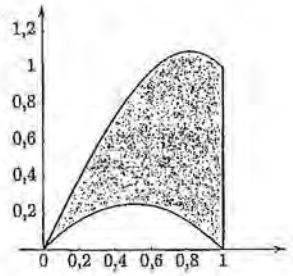


Figura 5.7.

© 978-88-08-06485-1

è (per definizione) y-semplice; disegnandolo (fig. 5.7, vediamo che non è anche x-semplice, in quanto certe rette orizzontali tagliano E in un insieme diverso da un segmento.

L'insieme di figura 5.8

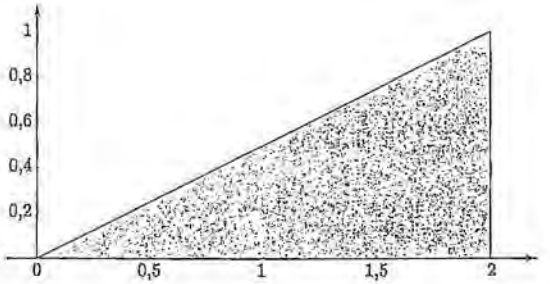


Figura 5.8.

è sia x-semplice sia y-semplice. Si può rappresentare (y-semplice) come:

$$\left\{(x,y) \in \mathbb{R}^2 : 0 \le x \le 2, 0 \le y \le \frac{x}{2}\right\}$$

oppure (x-semplice) come:

$$\{(x,y) \in \mathbb{R}^2 : 0 \le y \le 1, 2y \le x \le 2\}$$

Una corona circolare non è un insieme semplice, ma è regolare, potendosi decomporre per esempio come unione di due insiemi x-semplici o y-semplici.

L'insieme Ω dell'Esempio 1.3 *non è regolare*: è unione di un numero *infinito* di segmenti verticali, ognuno dei quali è una regione y-semplice.

Si noti che un insieme semplice è *necessariamente chiuso e limitato*, come si vede subito dalla definizione; di conseguenza anche un insieme regolare, essendo unione di un numero finito di insiemi semplici, è chiuso e limitato. In particolare una funzione continua su un dominio regolare ha massimo e minimo (per il teorema di Weierstrass) ed è perciò limitata.

Veniamo ora al risultato di integrabilità che ci aspettiamo:

TEOREMA 5.4 *Sia $\Omega \subset \mathbb{R}^2$ un dominio regolare e $f : \Omega \to \mathbb{R}$ continua. Allora f è integrabile in Ω.*

Si noti che questo risultato contiene come caso particolare il Teorema 5.1, visto che un rettangolo è un insieme regolare; non dimostreremo questo risultato.

La nozione di integrale doppio permette non solo di dar senso al concetto di volume della regione tridimensionale che sta sotto il grafico di una funzione, ma anche di quello di *area di una regione piana limitata*:

 © 978-88-08-06485-1

DEFINIZIONE 5.4 (INSIEME MISURABILE) Un insieme limitato $\Omega \subset \mathbb{R}^2$ si dirà *misurabile*[7] se la funzione costante 1 è integrabile in Ω. In tal caso, chiameremo *misura (o area)* di Ω, e indicheremo col simbolo $|\Omega|$, il numero:

$$|\Omega| = \iint_\Omega 1 \, dx\, dy.$$

Per il Teorema 5.4, *ogni insieme regolare è misurabile* (perché la funzione costante uguale ad 1 in Ω è continua!). Esistono tuttavia insiemi non misurabili, come l'insieme Ω descritto nell'Esempio 1.3.

Si rifletta in particolare sul significato del concetto di *insieme di misura nulla*, che sarà importante nello sviluppo della teoria dell'integrazione. In base alla definizione di integrale e di misura, è immediato convincersi della prossima:

PROPOSIZIONE 5.1 (CARATTERIZZAZIONE DEGLI INSIEMI DI MISURA NULLA) *Un insieme limitato $\Omega \subset \mathbb{R}^2$ è misurabile e ha misura nulla se e solo se vale la seguente proprietà:*

detto R un rettangolo contenente Ω, considerata la suddivisione di Ω in n^2 rettangolini uguali (come nella definizione di integrale) e detta A_n la somma delle aree dei rettangolini che hanno intersezione non vuota con Ω, si ha che $A_n \to 0$ per $n \to \infty$.

Notiamo esplicitamente che se la proprietà appena enunciata vale per un certo rettangolo R contenente Ω, allora vale per qualsiasi rettangolo R contenente Ω.

In parole povere, un insieme di misura nulla può essere ricoperto da rettangolini di area complessiva piccola quanto si vuole. Quali sono, in pratica, gli insiemi di misura nulla? Vale per esempio il seguente risultato:

PROPOSIZIONE 5.2 *Sia $g : [a, b] \to \mathbb{R}$ una funzione continua. Allora il grafico di g è un insieme di misura nulla.*

In particolare, quindi, un segmento ha misura nulla. Si faccia attenzione al significato di questo enunciato: la misura di cui si parla è l'*area*: certamente il grafico di una funzione continua o un segmento possono avere *lunghezza* positiva, ma la loro *area* è zero. Possiamo anche osservare quanto segue: l'integrale doppio è definito in $\mathbb{R}^2$, spazio di dimensione 2, e può essere utililizzato per definire l'*area*, che è una misura bidimensionale, dei sottoinsiemi di $\mathbb{R}^2$. Analogamente vedremo che mediante l'integrale triplo (in $\mathbb{R}^3$, spazio di dimensione 3) potremo definire il volume, che è una misura tridimensionale.

Vale anche la prossima:

PROPOSIZIONE 5.3 *L'unione di un numero finito di insiemi di misura nulla ha misura nulla.*

[7] Più precisamente, misurabile secondo Peano-Jordan. Esistono infatti definizioni diverse di misurabilità, in Analisi.

© 978-88-08-06485-1

Dalle due proposizioni precedenti segue subito il:

COROLLARIO 5.5 *Il bordo di un insieme regolare ha misura nulla.*

L'interesse di questi risultati sugli insiemi di misura nulla, in relazione al calcolo integrale, sta per esempio nel prossimo teorema, che estende in modo significativo la classe delle funzioni integrabili:

TEOREMA 5.6 *Sia $\Omega \subset \mathbb{R}^2$ un dominio regolare e $f : \Omega \to \mathbb{R}^2$ una funzione limitata e continua a eccezione di un insieme di misura nulla di punti di discontinuità. Allora f è integrabile in Ω.*

Anche l'applicabilità del teorema di riduzione sul rettangolo si estende a funzioni che presentano per esempio qualche linea di punti di discontinuità:

TEOREMA 5.7 (DI RIDUZIONE, FUNZIONI DISCONTINUE) *Sia $f : [a, b]\times[c, d] \to \mathbb{R}$ una funzione limitata, continua salvo un insieme di misura nulla di punti di discontinuità. Allora il suo integrale doppio si può calcolare come integrale iterato, cioè vale ancora la* (1.3).

1.3 Proprietà elementari dell'integrale doppio

Prima di concentrarci sul calcolo effettivo degli integrali, enunciamo ora le proprietà elementari dell'integrale doppio, analoghe a quelle dell'integrale in una variabile.

Siano $\Omega \subset \mathbb{R}^2$ un insieme limitato e misurabile, f, g funzioni integrabili in Ω, $c \in \mathbb{R}$ una costante. (Si noti la generalità delle ipotesi su f, su g e su Ω). Allora valgono le seguenti proprietà.

a) *Linearità dell'integrale:* le funzioni $f + g, cf$ sono anch'esse integrabili e

$$\iint_\Omega [f(x, y) + g(x, y)]\, dx\, dy = \iint_\Omega f(x, y)\, dx\, dy + \iint_\Omega g(x, y)\, dx\, dy,$$
$$\iint_\Omega cf(x, y)\, dx\, dy = c \iint_\Omega f(x, y)\, dx\, dy.$$

In particolare:

$$\iint_\Omega c\, dx\, dy = c\,|\Omega|\,.$$

b) *Positività e monotonia dell'integrale rispetto all'integranda:*

$$f \geq 0 \text{ in } \Omega \Rightarrow \iint_\Omega f(x, y)\, dx\, dy \geq 0;$$
$$f \geq g \text{ in } \Omega \Rightarrow \iint_\Omega f(x, y)\, dx\, dy \geq \iint_\Omega g(x, y)\, dx\, dy;$$
$$\left|\iint_\Omega f(x, y) dx\, dy\right| \leq \iint_\Omega |f(x, y)| dx\, dy \leq c|\Omega|$$

se c è una costante per cui $|f(x, y)| \leq c$ in Ω.

c) *Monotonia dell'integrale rispetto al dominio di integrazione*: se Ω' è un insieme limitato e misurabile contenuto in Ω, allora f risulta integrabile anche in Ω'; se inoltre $f(x) \geq 0$ in Ω risulta:

$$\iint_{\Omega} f(x,y)\,dx\,dy \geq \iint_{\Omega'} f(x,y)\,dx\,dy.$$

d) *Additività dell'integrale rispetto al dominio di integrazione*: se Ω_1, Ω_2 sono domini regolari, $\Omega_1 \cap \Omega_2$ ha misura nulla e f è integrabile in $\Omega_1 \cup \Omega_2$, allora:

$$\iint_{\Omega_1 \cup \Omega_2} f(x,y)\,dx\,dy = \iint_{\Omega_1} f(x,y)\,dx\,dy + \iint_{\Omega_2} f(x,y)\,dx\,dy. \tag{1.6}$$

e) *Proprietà di annullamento*:

$$|\Omega| = 0 \Rightarrow \iint_{\Omega} f(x,y)dx\,dy = 0.$$

Le proprietà a) e b) si possono provare esattamente come le analoghe proprietà dell'integrale di funzioni di una variabile[8], a partire dalla definizione di integrale come limite di somme; le proprietà c) e d) sono un po' delicate da dimostrare (non lo facciamo); la e) segue subito dall'ultima delle b).

Vediamo ora qualche altra proprietà dell'integrale che vale più specificamente per *funzioni continue* su domini opportuni.

TEOREMA 5.8 *Siano $\Omega \subset \mathbb{R}^2$ un insieme limitato e misurabile e $f : \Omega \to \mathbb{R}$ continua e limitata. Allora valgono le seguenti proprietà.*

f) Se Ω è aperto, valgono le proprietà di annullamento*:*

f1) $|\Omega| > 0, f(x,y) \geq 0, \iint_{\Omega} f(x,y)dx\,dy = 0 \Rightarrow f(x,y) \equiv 0$ in Ω;

f2) se per ogni intorno circolare $D \subset \Omega$ si ha:

$$\iint_{D} f(x,y)\,dx\,dy = 0,$$

allora necessariamente $f(x,y) \equiv 0$ in Ω.

g) Se Ω è connesso, vale il teorema della media*: esiste un punto $(x_0, y_0) \in \Omega$ tale che:*

$$\frac{1}{|\Omega|} \iint_{\Omega} f(x,y)\,dx\,dy = f(x_0, y_0). \tag{1.7}$$

DIMOSTRAZIONE.

f1). Ragioniamo per assurdo. Se esiste un punto di Ω in cui f non si annulla, per esempio $f(x_0, y_0) = \delta > 0$, per il teorema di permanenza del segno (f è continua) esiste un intorno circolare D di (x_0, y_0) ($D \subset \Omega$) in cui è $f \geq \delta/2$ e allora per il punto c) del teorema precedente:

$$0 = \iint_D f(x,y)\,dxdy \geq \iint_D f(x,y)\,dxdy \geq \frac{\delta}{2}|D|,$$

assurdo.

f2) Si dimostra in modo simile alla precedente: se esiste un punto in Ω in cui f non si annulla, per esempio $f(x_0, y_0) = \delta > 0$, esiste un intorno circolare D di (x_0, y_0) ($D \subset \Omega$) in cui è $f \geq \delta/2$ e allora

$$0 = \iint_D f(x,y)\,dxdy \geq \frac{\delta}{2}|D|,$$

assurdo.

g). Per il teorema di Weierstrass, f ha massimo M e minimo m in Ω; dalle disuguaglianze $m \leq f(x,y) \leq M$ per ogni $(x,y) \in \Omega$ si ricava:

$$m \leq \frac{1}{|\Omega|} \iint_{\Omega} f(x,y)\,dxdy \leq M.$$

Pertanto la media integrale di f, $\frac{1}{|\Omega|} \int\int_{\Omega} f(x,y)\,dxdy$, è un valore compreso tra il minimo e il massimo di f in Ω; poiché f è continua e Ω connesso, il teorema dei valori intermedi[9] garantisce l'esistenza di un punto $(x_0, y_0) \in \Omega$ per cui vale la (1.7). ◊

[8] v. volume 1, capitolo 6, paragrafo 3.

1.4 Calcolo degli integrali doppi: metodo di riduzione

Ora che abbiamo stabilito le prime proprietà dell'integrale ed enunciato delle condizioni sufficienti per l'integrabilità, si pone il problema del calcolo operativo degli integrali doppi su domini non rettangolari. La tecnica di riduzione illustrata per gli integrali su domini rettangolari si estende in modo naturale a domini più generali, in particolare a *domini semplici*, sfruttando la rappresentazione analitica di questi domini.

TEOREMA 5.9 (DI RIDUZIONE, PER DOMINI SEMPLICI) *Sia $f : \Omega \to \mathbb{R}$ continua e sia Ω un dominio x-semplice, ossia*

$$\Omega = \{(x,y) \in \mathbb{R}^2 : y \in [c,d], h_1(y) \leq x \leq h_2(y)\} \tag{1.8}$$

con $h_1, h_2 : [c,d] \to \mathbb{R}$ funzioni continue. Allora l'integrale doppio di f si può calcolare come integrale iterato *nel modo seguente:*

$$\iint_{\Omega} f(x,y)\,dx\,dy = \int_c^d \left(\int_{h_1(y)}^{h_2(y)} f(x,y)\,dx \right) dy. \tag{1.9}$$

Se invece Ω è y-semplice,

$$\Omega = \{(x,y) \in \mathbb{R}^2 : x \in [a,b], g_1(x) \leq y \leq g_2(x)\}$$

[9] v. capitolo 3, paragrafo 3.2.

© 978-88-08-06485-1

con $g_1, g_2 : [a,b] \to \mathbb{R}$ funzioni continue, vale la formula:

$$\iint_\Omega f(x,y)\,dx\,dy = \int_a^b \left(\int_{g_1(x)}^{g_2(x)} f(x,y)\,dy \right) dx.$$

Se Ω è sia x-semplice che y-semplice, valgono entrambe le formule.

Il significato geometrico della formula (1.9) è molto simile a quello già illustrato nel caso del teorema di riduzione per gli integrali su un rettangolo. Se Ω è un dominio x-semplice, tracciamo un piano verticale parallelo all'asse x; questo taglierà il solido di cui vogliamo calcolare il volume lungo una regione piana di area $A(y)$. Quest'area si calcola per mezzo dell'integrale unidimensionale della funzione $x \mapsto f(x,y)$ con y fissato:

$$A(y) = \int_{h_1(y)}^{h_2(y)} f(x,y)\,dx,$$

dove l'unica differenza rispetto al caso del rettangolo è che ora l'intervallo d'integrazione rispetto a x varia al variare di y.

Il volume totale del solido si ricostruisce poi integrando queste aree $A(y)$ per $y \in [c,d]$.

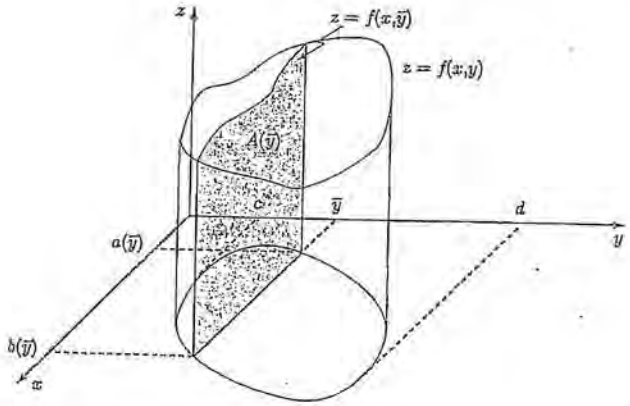


Figura 5.9.

DIMOSTRAZIONE. Sia $f : \Omega \to \mathbb{R}$ continua e sia R un rettangolo contenente Ω. Sia $\widetilde{f}$ la funzione definita su tutto R prolungando f a zero fuori da Ω. Per il teorema 5.1 f è integrabile in Ω, ossia $\widetilde{f}$ è integrabile in R. Inoltre, per il teorema 5.7, per la funzione $\widetilde{f}$ vale il teorema di riduzione sul rettangolo R. Infatti, $\widetilde{f}$ è discontinua solo su $\partial\Omega$, che è un insieme di misura nulla, per il Corollario 5.5. Scriviamo esplicitamente cosa significa questo, nell'ipotesi che Ω sia un dominio x-semplice come in (1.8):

$$\iint_\Omega f(x,y)\,dx\,dy = \iint_R \widetilde{f}(x,y)\,dx\,dy = \int_c^d \left(\int_a^b \widetilde{f}(x,y)\,dx \right) dy. \tag{1.10}$$

© 978-88-08-06485-1

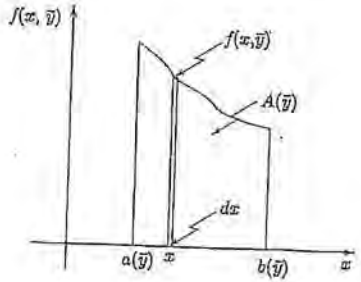


Figura 5.10.

Poiché $\widetilde{f}$ coincide con f in Ω ed è zero al di fuori di Ω, per ogni $y \in [c,d]$ si ha:

$$\widetilde{f}(x,y) = \begin{cases} f(x,y) & \text{se } h_1(y) \le x \le h_2(y) \\ 0 & \text{altrimenti.} \end{cases}$$

Di conseguenza:

$$\int_a^b \widetilde{f}(x,y)\,dx = \int_{h_1(y)}^{h_2(y)} f(x,y)\,dx,$$

che sostituita nella (1.10) dà la tesi.

Esempio

Si calcoli:

$$\iint_T xy\,dx\,dy$$

dove T è il triangolo con vertici in $(0,0)$, $(1,0)$ e $(1,1)$.

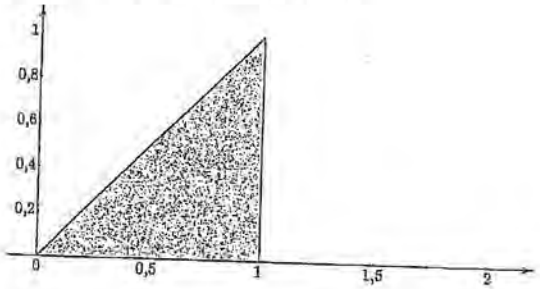


Figura 5.11.

Possiamo rappresentare T come insieme y-semplice:

$$T = \{(x,y) \in \mathbb{R}^2 : 0 < x < 1, 0 < y < x\}$$

Quindi, per calcolare l'integrale iterato, si blocca x tra 0 e 1 e si integra in y tra 0 e x; poi si integra in x tra 0 e 1:

$$\int_0^1 \left(\int_0^x xy\,dy \right) dx = \int_0^1 \left[\frac{1}{2} x y^2 \right]_0^x dx = \int_0^1 \frac{1}{2} x^3\,dx = \left[\frac{1}{8} x^4 \right]_0^1 = \frac{1}{8}.$$

 © 978-88-08-06485-1

Se avessimo rappresentato T invece come insieme x-semplice, cioè:

$$T = \{(x,y) \in \mathbb{R}^2 : 0 < y < 1, y < x < 1\},$$

l'integrale si sarebbe calcolato come segue: si blocca y tra 0 e 1 e si integra in x tra y e 1; quindi si integra in y tra 0 e 1:

$$\int_0^1 \left(\int_y^1 xy\,dx \right) dy = \int_0^1 \left[\frac{1}{2}x^2 y \right]_y^1 dy = \int_0^1 \frac{1}{2}(1-y^2)y\,dy = \left[\frac{1}{4}y^2 - \frac{1}{8}y^4 \right]_0^1 = \frac{1}{8}.$$

OSSERVAZIONE Se Ω è un dominio regolare, unione di un numero finito di domini semplici Ω_k che si intersecano a due a due solo lungo il bordo, risulta, per l'additività dell'integrale (v. (1.6)):

$$\iint_\Omega f(x,y)\,dx\,dy = \sum_{k=1}^{n} \iint_{\Omega_k} f(x,y)\,dx\,dy,$$

dove ognuno degli integrali a secondo membro si può calcolare come opportuno integrale iterato.

Calcolo di baricentri e momenti d'inerzia

Supponiamo che una lamina piana di densità superficiale $\rho(x,y)$ e massa totale M occupi la regione D. Il *baricentro* della lamina è per definizione il punto di coordinate:

$$\overline{x} = \frac{1}{M}\iint_D x\rho(x,y)dx\,dy \qquad \overline{y} = \frac{1}{M}\iint_D y\rho(x,y)dx\,dy$$

In particolare, se la lamina è omogenea (cioè ρ è costante e quindi $M = \rho|D|$), il baricentro si dice *centroide* e ha un significato puramente geometrico; infatti:

$$\overline{x} = \frac{1}{|D|}\iint_D x\,dx\,dy \qquad \overline{y} = \frac{1}{|D|}\iint_D y\,dx\,dy\,.$$

Le coordinate del centroide, cioè, dipendono solo dalla forma della lamina e non dalla densità (costante).

Il *momento d'inerzia* della lamina rispetto a un asse perpendicolare al piano (x,y) è invece la quantità:

$$I = \iint_D \delta^2(x,y)\rho(x,y)dx\,dy$$

dove $\delta(x,y)$ è la distanza di (x,y) dall'asse. Se il corpo è omogeneo e ha massa totale M, si ha:

$$I = \frac{M}{|D|}\iint_D \delta^2(x,y)\,dx\,dy.$$

Alcuni esempi e osservazioni saranno utili per mostrare le situazioni più tipiche che si presentano nel calcolo degli integrali iterati.

© 978-88-08-06485-1

Esempi

5.10 Calcolare il baricentro di una lamina triangolare di vertici $(0,0), (1,0), (0,1)$:
a) nell'ipotesi che il triangolo sia omogeneo;
b) nell'ipotesi che la densità superficiale sia $\rho(x,y) = c(x+1)$, con c costante.

a) Si ha $|T| = \frac{1}{2}$ e per simmetria sarà $\overline{x} = \overline{y}$. Abbiamo:

$$\overline{x} = 2\int_0^1 x\,dx \int_0^{1-x} dy = 2\int_0^1 x(1-x)dx = 2\left[\frac{x^2}{2} - \frac{x^3}{3} \right]_0^1 = \frac{1}{3}.$$

b) Calcoliamo prima la massa totale:

$$M = \int_0^1 dx \int_0^{1-x} c(1+x)dy = \int_0^1 c(1+x)(1-x)dx = \left[c\left(x - \frac{x^3}{3} \right) \right]_0^1 = \frac{2}{3}c.$$

Ora:

$$\overline{x} = \frac{1}{M}\int_0^1 x\,c(1+x)dx \int_0^{1-x} dy = \frac{3}{2c}\int_0^1 cx(1-x^2)dx = \frac{3}{2}\left[\frac{x^2}{2} - \frac{x^4}{4} \right]_0^1 = \frac{3}{8}$$

$$\overline{y} = \frac{1}{M}\int_0^1 c(1+x)dx \int_0^{1-x} y\,dy = \frac{3}{2c}\int_0^1 c\frac{(1+x)(1-x)^2}{2}dx =$$

$$= \frac{3}{4}\left[\frac{x^4}{4} - \frac{x^3}{3} - \frac{x^2}{2} + x \right]_0^1 = \frac{5}{16}.$$

5.11 Sia D come in figura 5.12:

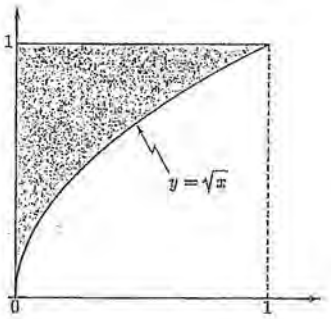


Figura 5.12.

(La curva disegnata è il grafico di $y = \sqrt{x}$ per $x \in [0,1]$). Calcolare:

$$I = \iint_D \sin(y^3)dx\,dy.$$

Rappresentando D come dominio y-semplice, si ha:

$$I = \int_0^1 dx \int_{\sqrt{x}}^1 \sin(y^3)dy.$$

 © 978-88-08-06485-1

Ora, nell'integrale interno abbiamo una funzione che non ha primitiva elementare, per cui il calcolo *esatto* dell'integrale per questa via appare impraticabile. Naturalmente l'integrale esiste, perché la funzione è continua e il dominio regolare, e può essere calcolato numericamente con l'approssimazione voluta. Proviamo, in alternativa, a rappresentare D come dominio x-semplice:

$$I = \int_0^1 \sin(y^3)dy \int_0^{y^2} dx = \int_0^1 y^2 \sin(y^3)dy =$$
$$= \left[-\frac{1}{3}\cos(y^3)\right]_0^1 = \frac{1}{3}(1-\cos 1) \simeq 0{,}153$$

In questo caso l'integrale interno è facilmente calcolabile, come pure è facilmente determinabile la primitiva nell'integrale esterno. La conclusione è che, quando entrambe le iterazioni sono possibili (perché il dominio è semplice rispetto a entrambi gli assi), *la scelta dell'ordine di iterazione* può essere importante.

Esempio Calcolare:

$$\iint_{x^2+y^2<4} \sqrt{4-x^2-y^2}\,dx\,dy.$$

Si pensi al significato geometrico: l'integrale rappresenta il volume della semisfera $z = \sqrt{4-x^2-y^2}$, di raggio 2; perciò vale:

$$\frac{1}{2}\cdot\frac{4}{3}\pi\cdot 2^3 = \frac{16}{3}\pi$$

(senza bisogno di ricorrere al calcolo integrale!)

Esempio Calcolare:

$$\iint_{[-1,1]\times[-1,1]} (x+\sin y)\,dx\,dy.$$

L'integrale, per linearità, è somma di due integrali, di x e di $\sin y$. Ognuna di queste funzioni è dispari nella propria variabile e perciò ha integrale nullo su $[-1,1]$: ne segue che l'integrale doppio è pure nullo (senza bisogno di calcoli!).

Esempio Calcolare:

$$I = \iint_D xy\,dx\,dy$$

con D come in figura 5.13.

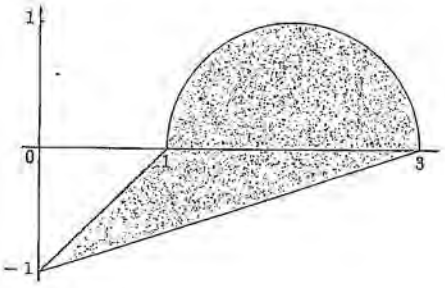


Figura 5.13.

© 978-88-08-06485-1

La rappresentazione analitica di D in questo caso è un po' complessa; la cosa più naturale è scrivere D come unione dei due insiemi seguenti:

$$D_1 = \left\{(x,y) : 0 \le x \le 1, \frac{x}{3} - 1 \le y \le x - 1\right\};$$
$$D_2 = \left\{(x,y) : 1 \le x \le 3, \frac{x}{3} - 1 \le y \le \sqrt{1-(x-2)^2}\right\}.$$

Conseguentemente, l'integrale si spezza nella somma di due integrali, ognuno dei quali si calcola come integrale iterato:

$$\iint_{D_1} xy\,dx\,dy = \int_0^1 \left(\int_{\frac{x}{3}-1}^{x-1} xy\,dy\right) dx = \int_0^1 x\left[\frac{y^2}{2}\right]_{\frac{x}{3}-1}^{x-1} dx =$$
$$= \int_0^1 \left(\frac{4}{9}x^3 - \frac{2}{3}x^2\right) dx = (\dots) - \frac{1}{9}.$$

$$\iint_{D_2} xy\,dx\,dy = \int_1^3 \left(\int_{\frac{x}{3}-1}^{\sqrt{1-(x-2)^2}} xy\,dy\right) dx = \int_1^3 \left(x\left[\frac{y^2}{2}\right]_{\frac{x}{3}-1}^{\sqrt{1-(x-2)^2}}\right) dx =$$
$$= \int_1^3 \left(-\frac{5}{9}x^3 + \frac{7}{3}x^2 - 2x\right) dx = (\dots) = \frac{10}{9}.$$

Si conclude che:

$$I = -\frac{1}{9} + \frac{10}{9} = 1.$$

Esercizi

Calcolare i seguenti integrali doppi come integrali iterati opportuni e/o mediante opportune considerazioni di simmetria.

1. $$\iint_{[0,1]\times[0,\frac{\pi}{2}]} y\sin(xy)\,dx\,dy$$

2. $$\iint_T (x+\sin y)\,dx\,dy$$
con $T = \{(x,y) : 0 < x < 1;\ 0 < y < 1-x\}$.

3. $$\iint_{x^2+y^2<2} xy^3\,dx\,dy$$

4. $$\iint_{[0,1]\times[0,1)} \frac{xy}{x^2+y^2}\,dx\,dy$$
In questo caso la funzione non è continua fino al bordo del quadrato – nell'origine è discontinua – tuttavia è limitata (perché?) e integrabile.

 © 978-88-08-06485-1

$$\iint_{[0,1]\times[0,1]} x^y \, dx\, dy$$

Vale la stessa osservazione dell'esercizio precedente. Si provi qui a calcolare l'integrale iterando nei due ordini diversi.

$$\iint_{[0,1]\times[0,\pi]} x \sin y \, dx\, dy$$

$$\iint_{\{(x,y):0\le x\le 1, 0\le y\le x\}} x \sin y \, dx\, dy$$

$$\iint_T (x+5)\, dx\, dy$$

con $T = \{(x,y) : |x| \le 1, 0 \le y \le 2 - 2|x|\}$. (Non occorre calcolare integrali!!)

$$\iint_Q y \, dx\, dy$$

dove Q è il quadrilatero di vertici $(0,1), (1,2), (2,1), (2,0)$.

Calcolare il centroide del quarto di cerchio:

$$x^2 + y^2 < R^2, \quad x > 0, \quad y > 0.$$

Calcolare il centroide del quadrilatero dell'esercizio 9.

Calcolare il baricentro della lamina rettangolare $[0,4] \times [0,3]$, di densità superficiale $\rho(x,y) = 1 + x^2 + y^2$.

Calcolare il momento d'inerzia di una lamina circolare omogenea di massa M e raggio R rispetto all'asse passante per il centro e perpendicolare al cerchio.

Calcolare l'area del dominio:

$$D = \{(x,y) : 0 \le y \le 1, y^2 - 1 \le x \le \sqrt{1-y^2}\}.$$

Calcolare il volume della regione sotto il grafico del paraboloide:

$$V = \{(x,y,z) : z < 1 - x^2 - y^2; \quad x^2 + y^2 < 1\}.$$

Calcolare:

$$\iint_R x e^{\sqrt{y}} \, dx\, dy$$

dove $R = \{(x,y) : 0 \le x \le 1; 0 \le y \le x^2\}$.

Calcolare:

$$\iint_{|x|+|y|\le 1} x^2 y^2 \, dx\, dy$$

© 978-88-08-06485-1

Calcolare:

$$\iint_D |x - y| \, dx\, dy$$

con

$$D = \{(x,y) : 0 \le x \le 1; x^2 \le y \le 1\}.$$

1.5 Calcolo degli integrali doppi: cambiamento di variabili

Nel calcolo degli integrali di funzioni di una variabile (volume 1, capitolo 6), abbiamo visto che il metodo di *integrazione per sostituzione* risulta essere uno strumento potente e flessibile. Vale la formula:

$$\int_a^b f(x)\, dx = \int_{\phi^{-1}(a)}^{\phi^{-1}(b)} f(\phi(t))\, \phi'(t)\, dt \tag{1.11}$$

dove $x = \phi(t)$ è una funzione derivabile e monotona da $\phi^{-1}([a,b])$ ad $[a,b]$.

Negli integrali di funzioni di due variabili (e, come vedremo in seguito, in quelli di n variabili), il procedimento analogo consiste nell'effettuare una *trasformazione di coordinate* nel piano: dovendo calcolare

$$\iint_D f(x,y)\, dx\, dy,$$

si opera una trasformazione di coordinate $\mathrm{T} : D' \to D$, $(x,y) = \mathrm{T}(u,v)$ con

$$\begin{cases} x = g(u,v) \\ y = h(u,v) \end{cases}$$

che porta a esprimere l'integranda in funzione di u, v,

$$f(g(u,v), h(u,v)),$$

mentre il dominio D viene trasformato in D'.

Il problema è: come si trasforma l'integrale doppio? In altre parole, qual è l'analogo bidimensionale della formula

$$dx = \phi'(t)\, dt, \tag{1.12}$$

che si utilizza nella sostituzione (1.11)?

Per capirlo, possiamo pensare alla (1.12) come a una formula che descrive come cambia l'elemento infinitesimo di lunghezza, sotto l'azione della sostituzione $x = \phi(t)$. Allora, in due variabili, la domanda diventa: come cambia, sotto l'azione della trasformazione di coordinate

$$\begin{cases} x = g(u,v) \\ y = h(u,v), \end{cases}$$

l'elemento infinitesimo d'area $dxdy$?

Per rispondere, ragioniamo nel piano (u,v). Il "rettangolo infinitesimo" compreso tra $u, u+du$ e $v, v+dv$ viene trasformato in un "parallelogramma infinitesimo" che ha per vertici i quattro punti:

$$\begin{array}{ll} (g(u,v+dv), h(u,v+dv)) & (g(u+du,v+dv), h(u+du,v+dv)) \\ (g(u,v), h(u,v)) & (g(u+du,v), h(u+du,v)) \end{array}$$

 © 978-88-08-06485-1

Due lati adiacenti del parallelogramma sono quindi individuati dai vettori

$$(g(u+du,v),h(u+du,v))-(g(u,v),h(u,v))$$

e

$$(g(u,v+dv),h(u,v+dv))-(g(u,v),h(u,v)),$$

uguali in prima approssimazione a

$$(g_u du, h_u du)$$

e

$$(g_v dv, h_v dv)$$

rispettivamente. L'area del parallelogramma trasformato è quindi (in prima approssimazione) uguale al modulo del determinante della matrice

$$\begin{pmatrix} g_u du & h_u du \\ g_v dv & h_v dv \end{pmatrix}.$$

Ciò suggerisce la seguente *formula di trasformazione dell'elemento infinitesimo d'area*

$$dx\,dy = \left|\det \begin{pmatrix} g_u & h_u \\ g_v & h_v \end{pmatrix}\right| du\,dv.$$

Si osservi che la matrice scritta non è altro che la matrice Jacobiana della trasformazione di coordinate (più precisamente, è la sua trasposta, ma il determinante non cambia per trasposizione).

Il precedente ragionamento euristico è confermato dal seguente teorema:

TEOREMA 5.10 (FORMULA DI CAMBIAMENTO DI VARIABILI) *Sia $D \subset \mathbb{R}^2$ un dominio regolare, $f : D \to \mathbb{R}$ una funzione continua e $\mathbf{T} : D' \to D$, $(x,y) = \mathbf{T}(u,v)$, con*

$$\begin{cases} x = g(u,v) \\ y = h(u,v) \end{cases} \tag{1.13}$$

una trasformazione di coordinate, o più precisamente, un diffeomorfismo globale[10] *che porta D' in D.*

Allora

$$\iint_D f(x,y)\,dx\,dy = \iint_{D'} f(g(u,v),h(u,v))\,|\det \mathbf{DT}(u,v)|\,du\,dv$$

dove

$$\mathbf{DT}(u,v) = \begin{pmatrix} g_u & h_u \\ g_v & h_v \end{pmatrix}$$

indica la matrice Jacobiana della trasformazione (1.13).

La dimostrazione di questo teorema è lunga e delicata e verrà omessa.

[10] v. capitolo 4, paragrafo 5.

© 978-88-08-06485-1

OSSERVAZIONE Negli integrali di funzioni di una variabile, il criterio per scegliere una sostituzione consiste unicamente nel considerare la semplificazione che essa induce sulla funzione integranda. In due (o più) variabili, invece, una trasformazione di coordinate può avere come scopo sia semplificare la forma della funzione integranda, sia semplificare la *forma del dominio* (per esempio, ottenendo nelle nuove variabili un dominio di tipo rettangolare); meglio ancora quando si ottengono entrambe le cose contemporaneamente. Gli esempi illustreranno quest'idea.

ESEMPI

Coordinate polari. Passando in coordinate polari

$$\begin{cases} x = \rho\cos\theta \\ y = \rho\sin\theta \end{cases}$$

si ha:

$$dx\,dy = \rho\,d\rho\,d\theta,$$

formula che può anche essere giustificata direttamente (sul piano intuitivo) nel seguente modo: il "rettangolo infinitesimo" nel piano ρ,θ individuato dai valori $\rho, \rho+d\rho, \theta, \theta+d\theta$ viene trasformato in un settore di corona circolare, nel piano x,y, che, in prima approssimazione, equivale a un rettangolo di lati, rispettivamente, $d\rho$ e $\rho\,d\theta$ (fig. 5.14) e quindi area $\rho\,d\rho\,d\theta$:

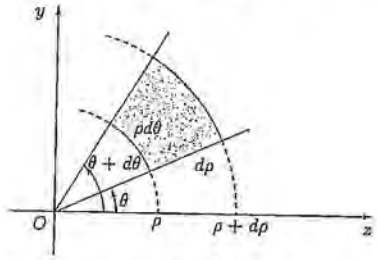


Figura 5.14. L'area della regione ombreggiata è circa $\rho\,d\rho\,d\theta$.

L'integrale doppio si trasforma pertanto nel modo seguente:

$$\iint_D f(x,y)\,dx\,dy = \iint_{D'} f(\rho\cos\theta,\rho\sin\theta)\rho\,d\rho\,d\theta$$

dove D' è il dominio D espresso nelle nuove variabili (ρ,θ).

Questo cambiamento di variabili è utile al calcolo in molti casi in cui il dominio D ha qualche simmetria di tipo radiale, per cui si esprime facilmente in coordinate polari. Per esempio un cerchio, una corona circolare, un settore circolare, in coordinate polari diventano rettangoli, poiché sono espressi da relazioni del tipo $a \le \rho \le b$, $c \le \theta \le d$. Questo rende in generale più semplice il calcolo dell'integrale iterato, perché l'integrale interno ha estremi fissi anziché variabili. Se anche la funzione integranda ha qualche simmetria radiale (per esempio è radiale, oppure è omogenea) le cose sono ancora più semplici.

Calcoliamo

$$\iint_D \frac{xy^2}{x^2+y^2}\,dx\,dy$$

 © 978-88-08-06485-1

con $D = \{(x,y) : x^2+y^2 < 4, x > 0, y > 0\}$.

Si tratta di integrare sul quarto di cerchio espresso, in coordinate polari, da: $\rho < 2; 0 < \theta < \pi/2$. Passando in coordinate polari, l'integrale diventa:

$$\int_0^{\frac{\pi}{2}} \left(\int_0^2 \frac{\rho^3 \cos\theta \sin^2\theta}{\rho^2} \rho \, d\rho \right) d\theta = \left(\int_0^{\frac{\pi}{2}} \cos\theta \sin^2\theta \, d\theta \right) \left(\int_0^2 \rho^2 \, d\rho \right) =$$
$$= \left[\frac{\sin^3\theta}{3} \right]_0^{\pi/2} \left[\frac{\rho^3}{3} \right]_0^2 = \frac{1}{3} \cdot \frac{8}{3} = \frac{8}{9}.$$

Si noti che, in questo esempio, in coordinate polari il dominio diventa un rettangolo e la funzione integranda si fattorizza nel prodotto $f(\rho)g(\theta)$; di conseguenza l'integrale doppio si fattorizza nel prodotto di due integrali unidimensionali.

Esempi

1.16 Calcoliamo:

$$I_R = \iint_{x^2+y^2<R^2} e^{-(x^2+y^2)} \, dx \, dy.$$

Passando in coordinate polari, si ha:

$$I_R = \int_0^{2\pi} \left(\int_0^R e^{-\rho^2} \rho \, d\rho \right) d\theta = 2\pi \left[-\frac{1}{2} e^{-\rho^2} \right]_0^R = \pi(1 - e^{-R^2}).$$

In questo caso il dominio è radiale, la funzione integranda è radiale (indipendente da θ), perciò l'integrale doppio in coordinate polari si riduce a un solo integrale unidimensionale (in ρ).

1.17 *Volume dell'ellissoide.* Calcoliamo il volume dell'ellissoide di semiassi a, b, c. Poiché l'ellissoide ha equazione:

$$\frac{x^2}{a^2} + \frac{y^2}{b^2} + \frac{z^2}{c^2} = 1,$$

la metà superiore dell'ellissoide è il grafico della funzione:

$$z = c\sqrt{1 - \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} \right)} \quad \text{per } \frac{x^2}{a^2} + \frac{y^2}{b^2} < 1.$$

Il volume cercato è quindi il doppio dell'integrale di questa funzione sul dominio $\frac{x^2}{a^2} + \frac{y^2}{b^2} < 1$:

$$V = 2 \iint_{\frac{x^2}{a^2}+\frac{y^2}{b^2}<1} c\sqrt{1 - \left(\frac{x^2}{a^2} + \frac{y^2}{b^2} \right)} \, dx \, dy$$

Questo integrale può essere facilmente calcolato mediante un cambiamento di variabili simile a quello polare, ma di tipo ellittico:

$$\begin{cases} x = a\rho\cos\theta \\ y = b\rho\sin\theta \end{cases}$$

che ha determinante jacobiano:

$$\begin{vmatrix} a\cos\theta & b\sin\theta \\ -a\rho\sin\theta & b\rho\cos\theta \end{vmatrix} = ab\rho.$$

© 978-88-08-06485-1

Quindi l'elemento d'area è $dx\,dy = ab\rho \, d\rho \, d\theta$ e l'integrale è:

$$V = 2\int_0^{2\pi} \left(\int_0^1 c\sqrt{1-\rho^2} ab\rho \, d\rho \right) d\theta =$$
$$= 4\pi abc \int_0^1 \rho\sqrt{1-\rho^2} \, d\rho = 4\pi abc \left[-\frac{1}{3}(1-\rho^2)^{3/2} \right]_0^1 = \frac{4}{3}\pi abc.$$

Il risultato trovato è una formula che generalizza in modo naturale la formula del volume della sfera, $\frac{4}{3}\pi r^3$.

Esercizi

1.18 Svolgere nuovamente gli esercizi 3, 10, 15, usando le coordinate polari (il procedimento risulta semplificato).

1.19 Sia D il quarto di corona circolare con centro nell'origine, raggi 1 e 2 e angolo variabile tra $\frac{\pi}{4}$ e $\frac{3}{4}\pi$. Calcolare:

$$\iint_D x^2 y \, dx \, dy.$$

1.20 Calcolare:

$$\iint_D \frac{x^3 y}{x^2+y^2} dx \, dy,$$

dove D è l'intersezione del settore circolare $\{\frac{\pi}{6} < \theta < \frac{\pi}{3}\}$ con la corona circolare con centro nell'origine e raggi 1 e 2.

1.21 Calcolare:

$$I = \iint_D \frac{xy}{x^2+y^2} dx \, dy,$$

dove D è il dominio (interno all'arco di spirale, vedi figura 5.15) espresso in coordinate polari da: $\{\rho < \theta, 0 < \theta < \frac{3}{2}\pi\}$.

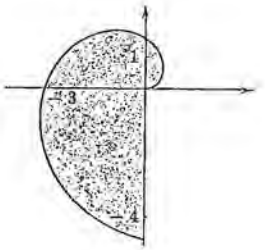


Figura 5.15.

(Si noti che in questo caso il dominio espresso in coordinate polari non è un rettangolo; di conseguenza l'integrale doppio in polari diventa un effettivo integrale iterato e non semplicemente il prodotto di due integrali).

1.22 Calcolare:

$$\iint_{[0,1]\times[0,1]} \frac{x^3 y}{x^2+y^2} dx \, dy \quad \text{e} \quad \iint_{[-1,1]\times[-1,1]} \frac{x^3 y}{x^2+y^2} dx \, dy.$$

 © 978-88-08-06485-1

25 Calcolare:

$$\iint_{x^2+4y^2<1} y^2\,dx\,dy$$

con un opportuno cambiamento di variabili.

26 Calcolare il baricentro di una lamina omogenea a forma di un quarto di corona circolare: $\{(\rho,\theta) : r < \rho < R; 0 < \theta < \frac{\pi}{2}\}$.

27 Calcolare il momento d'inerzia di una lamina omogenea di massa M a forma di corona circolare di raggi r, R, rispetto all'asse passante per il centro e perpendicolare al piano della corona.

28 Calcolare il momento d'inerzia di un disco omogeneo di massa M e raggio R, rispetto a un asse perpendicolare al suo piano e passante per un punto posto sulla circonferenza.
Suggerimento: l'equazione polare della circonferenza di raggio R passante per l'origine è: $\rho = 2R\cos\theta$, $\theta \in [-\frac{\pi}{2}, \frac{\pi}{2}]$.

29 Calcolare:

$$\iint_D \frac{xy}{\sqrt{x^2+y^2}}\,dx\,dy$$

dove D è la regione descritta in coordinate polari da $\{(\rho,\theta) : \pi < \theta < 2\pi, \rho < \theta\}$.

2 INTEGRALI DOPPI GENERALIZZATI

Come per l'integrale unidimensionale, anche per gli integrali doppi è utile generalizzare la nozione di integrale comprendendo anche i casi in cui il dominio di integrazione e/o la funzione integranda siano illimitate. L'integrale generalizzato può essere definito come un opportuno limite di integrali doppi usuali. Non diamo definizioni astratte, ma mostriamo attraverso alcuni esempi come si utilizza questa nozione. Quando la funzione da integrare ha *segno costante* sul dominio di integrazione, non ci sono troppi problemi.

Esempio

2.15 *Integrale della Gaussiana.* Calcoliamo

$$I = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}\,dx\,dy$$

Si può calcolare questo integrale come:

$$\lim_{R\to+\infty} \iint_{x^2+y^2<R^2} e^{-(x^2+y^2)}\,dx\,dy = \lim_{R\to+\infty} I_R$$

dove l'integrale I_R è stato calcolato nell'Esempio 1.16 e vale $\pi(1-e^{-R^2})$. Per $R \to +\infty$, otteniamo $I = \pi$. L'integrale generalizzato proposto, dunque, converge e il suo valore è π.

Osserviamo, incidentalmente, che questo permette di calcolare l'integrale unidimensionale notevole:

$$J = \int_{\mathbb{R}} e^{-t^2}\,dt$$

© 978-88-08-06485-1

Infatti,

$$I = \iint_{\mathbb{R}^2} e^{-(x^2+y^2)}\,dx\,dy = \int_{\mathbb{R}} e^{-x^2}\,dx \int_{\mathbb{R}} e^{-y^2}\,dy = J^2$$

da cui

$$J = \sqrt{I}$$

e quindi

$$\boxed{\int_{\mathbb{R}} e^{-t^2}\,dt = \sqrt{\pi}.}$$

Il risultato non è banale, perché la funzione e^{-t^2} non ha una primitiva elementare ed è rilevante, perché l'integrale della densità gaussiana e^{-t^2} gioca un ruolo importante in probabilità, statistica e varie altre discipline. Nei Complementi in fondo al capitolo mostreremo un'applicazione di questo integrale notevole.

Esempi

2.16 Calcoliamo:

$$I = \iint_{x^2+y^2<1} \frac{1}{(\sqrt{x^2+y^2})^\alpha}\,dx\,dy \quad \text{per} \quad \alpha > 0.$$

La funzione è illimitata. Si può calcolare:

$$I = \lim_{r\to0+} \int_{r<x^2+y^2<1} \frac{1}{(\sqrt{x^2+y^2})^\alpha}\,dx\,dy = \lim_{r\to0+} 2\pi \int_r^1 \frac{1}{\rho^\alpha}\rho\,d\rho =$$
$$= (\text{per } \alpha \neq 2) \quad \lim_{r\to0+} \frac{2\pi}{2-\alpha}(1-r^{2-\alpha}) =$$
$$= \begin{cases} \frac{2\pi}{2-\alpha} & \text{se } \alpha < 2 \\ +\infty & \text{se } \alpha > 2. \end{cases}$$

Se invece $\alpha = 2$, si ottiene:

$$I = \lim_{r\to0+} -2\pi\log r = +\infty.$$

Pertanto: l'integrale converge se $\alpha < 2$, diverge se $\alpha \geq 2$.

2.17 Un calcolo analogo mostra che:

$$\iint_{x^2+y^2>1} \frac{1}{(\sqrt{x^2+y^2})^\alpha}\,dx\,dy$$

converge se e solo se $\alpha > 2$.

La convergenza o meno di un integrale doppio generalizzato può essere studiata osservando la convergenza o meno dell'integrale iterato corrispondente, *purché la funzione integranda abbia segno costante nell'insieme di integrazione, o per lo meno cambi di segno solo un numero finito di volte*. In caso contrario, si possono presentare casi in cui il valore dell'integrale iterato non coincide con quello dell'integrale doppio.

 © 978-88-08-06485-1

Esempio

Studiare la convergenza dell'integrale doppio:

$$\iint_D e^{-xy}\,dx\,dy$$

con $D = \{(x,y) : x > 0, y > 0\}$, oppure $D = \{(x,y) : x > 1, y > 1\}$.

Il dominio è illimitato, l'integranda è positiva; calcoliamo l'integrale iterato generalizzato:

$$\int_a^{+\infty}\left(\int_a^{+\infty} e^{-xy}dx\right)dy = \int_a^{+\infty}\left[\frac{e^{-xy}}{-y}\right]_a^{+\infty} dy =$$

$$= \int_a^{+\infty}\frac{e^{-ay}}{y}dy = \begin{cases} \text{se } a=0 & \displaystyle\int_0^{+\infty}\frac{dy}{y} \quad \text{che diverge;} \\ \text{se } a=1 & \displaystyle\int_1^{+\infty}\frac{e^{-y}}{y}dy \quad \text{che converge.} \end{cases}$$

Nel secondo caso, non si sa calcolare il valore esatto dell'integrale.

3 IL CALCOLO DEGLI INTEGRALI TRIPLI

Tutto quanto abbiamo detto a proposito degli integrali doppi si generalizza a dimensione superiore, cioè per funzioni reali di più variabili. Occupiamoci qui brevemente degli integrali tripli; nei Complementi faremo qualche cenno agli integrali di funzioni di n variabili con n qualsiasi.

Nella definizione i rettangoli $[a,b]\times[c,d]$ sono sostituiti dai parallelepipedi $[a_1,b_1]\times[a_2,b_2]\times[a_3,b_3]$, dopo di che tutto procede inalterato e si arriva a definire l'integrale triplo di una funzione continua definita su un parallelepipedo.

Per quanto riguarda le funzioni definite su domini più generali, anche in questo caso, se il dominio di integrazione ha una rappresentazione analitica abbastanza semplice, si dimostra che una funzione continua risulta integrabile e il suo integrale si può calcolare mediante opportuni integrali iterati. Precisamente, si incontrano due situazioni tipiche:

Integrazione "per fili"

Sia Ω un dominio di $\mathbb{R}^3$ che si può rappresentare analiticamente nella forma:

$$\Omega = \{(x,y,z) : g_1(x,y) \le z \le g_2(x,y), (x,y) \in D\} \tag{3.1}$$

dove, a sua volta, D è un dominio regolare del piano e $g_1, g_2 : D \to \mathbb{R}$ sono continue.

Allora, se $f : \Omega \to \mathbb{R}$ è una funzione continua, f è integrabile in Ω e l'integrale si può calcolare mediante la formula (detta "di integrazione per fili"):

$$\iiint_\Omega f(x,y,z)\,dx\,dy\,dz = \iint_D\left(\int_{g_1(x,y)}^{g_2(x,y)} f(x,y,z)\,dz\right)dx\,dy \tag{3.2}$$

Geometricamente, la (3.2) significa che prima (nell'integrale più interno) si integra sul "filo" $z \in (g_1(x,y), g_2(x,y))$ e poi si fa variare (x,y) nel dominio D.

© 978-88-08-06485-1

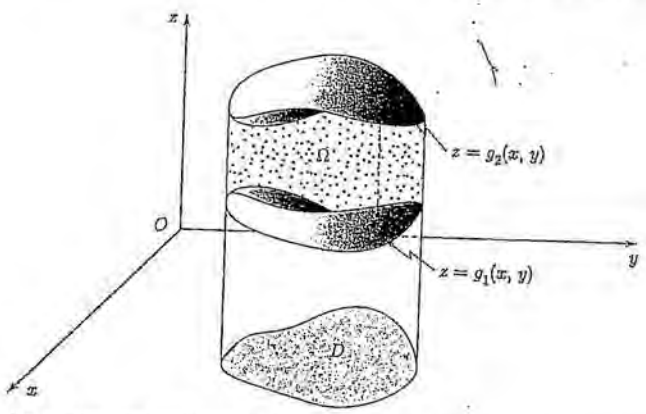


Figura 5.16. Integrazione *per fili*: si fissa (x,y) e si integra prima sul segmento tratteggiato.

Esempio

Un dominio che si rappresenta facilmente nella forma richiesta dal teorema precedente è la semisfera superiore con centro nell'origine e raggio R:

$$\Omega_1 = \left\{(x,y,z) : 0 \le z \le \sqrt{R^2-x^2-y^2}, x^2+y^2 \le R^2\right\}$$

Per esempio:

$$\iiint_{\Omega_1} x^2 z\,dx\,dy\,dz = \iint_{x^2+y^2\le R^2}\left(\int_0^{\sqrt{R^2-x^2-y^2}} x^2 z\,dz\right)dx\,dy =$$

$$= \iint_{x^2+y^2\le R^2} x^2\left[\frac{1}{2}z^2\right]_0^{\sqrt{R^2-x^2-y^2}} dx\,dy =$$

$$= \iint_{x^2+y^2\le R^2}\frac{1}{2}x^2\left(R^2-x^2-y^2\right)dx\,dy.$$

A questo punto l'integrale è ridotto a un integrale doppio; in questo caso, il dominio circolare suggerisce di usare le coordinate polari, ottenendo:

$$\int_0^{2\pi}\left(\int_0^R \frac{1}{2}\rho^2\cos^2\theta\left(R^2-\rho^2\right)\rho\,d\rho\right)d\theta = \frac{1}{2}\left(\int_0^{2\pi}\cos^2\theta\,d\theta\right)\left(\int_0^R\left(\rho^3R^2-\rho^5\right)d\rho\right) =$$

$$= \frac{\pi}{2}\left[\frac{\rho^4}{4}R^2-\frac{\rho^6}{6}\right]_0^R = \frac{\pi}{24}R^6.$$

Integrazione "per strati"

Supponiamo ora che Ω sia un dominio di $\mathbb{R}^3$ rappresentabile nella forma:

$$\Omega = \{(x,y,z) : h_1 \le z \le h_2, (x,y) \in \Omega(z)\} \tag{3.3}$$

dove, a sua volta, per ogni $z \in [h_1,h_2]$, $\Omega(z)$ è un dominio regolare del piano.

 © 978-88-08-06485-1

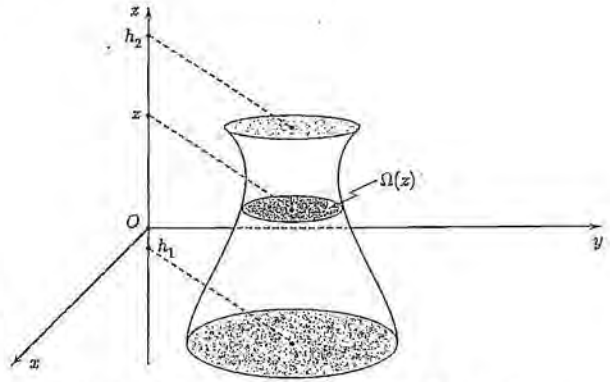


Figura 5.17. Integrazione per *strati*: si fissa z e si integra prima sulla sezione $\Omega(z)$.

Allora, se $f : \Omega \to \mathbb{R}$ è una funzione continua, f è integrabile in Ω e l'integrale si può calcolare mediante la formula (detta "di integrazione per strati"):

$$\iiint_\Omega f(x,y,z)\,dx\,dy\,dz = \int_{h_1}^{h_2} \left(\iint_{\Omega(z)} f(x,y,z)\,dx\,dy \right) dz.$$

In questo caso si integra prima sulla sezione ("strato") $\Omega(\overline{z}) = \Omega \cap \{z = \overline{z}\}$ e poi si fa variare la quota $\overline{z}$ tra le quote minima e massima di z in Ω.

Esempio

Un dominio che si rappresenta bene nella forma richiesta dall'integrazione per strati è il cono. Sia:

$$\Omega_2 = \left\{ (x,y,z) : x^2 + y^2 \le \left(\frac{R}{h} z\right)^2, 0 \le z \le h \right\}$$

il cono circolare di altezza h, raggio R, vertice nell'origine e asse lungo l'asse z. Per esempio:

$$\iiint_{\Omega_2} (x^2+y^2)\,dx\,dy\,dz = \int_0^h \left(\iint_{x^2+y^2 \le \left(\frac{R}{h}z\right)^2} (x^2+y^2)\,dx\,dy \right) dz =$$

(calcolando l'integrale interno in coordinate polari)

$$= \int_0^h \left(\int_0^{2\pi} \left(\int_0^{\frac{R}{h}z} \rho^3\,d\rho \right) d\theta \right) dz = 2\pi \int_0^h \frac{R^4}{4h^4} z^4\,dz = 2\pi \frac{R^4}{4h^4}\frac{h^5}{5} = \frac{\pi}{10} R^4 h.$$

Nel seguito chiameremo *domini regolari di* $\mathbb{R}^3$ gli insiemi che si possono esprimere come unione di un numero finito di insiemi dei tipi (3.1) e (3.3).

Cambio di variabili negli integrali tripli

Anche per il calcolo degli integrali tripli, può essere utile eseguire particolari trasformazioni di coordinate. Vale in proposito un teorema perfettamente analogo a quello enunciato per gli integrali doppi:

© 978-88-08-06485-1

TEOREMA 5.11 (FORMULA DI CAMBIAMENTO DI VARIABILI NEGLI INTEGRALI TRIPLI) *Sia* $D \subset \mathbb{R}^3$ *un dominio regolare,* $f : D \to \mathbb{R}$ *una funzione continua e* $\mathbf{T} : D' \to D$ *un diffeomorfismo globale, con* $(x,y,z) = \mathbf{T}(u,v,w)$,

$$\begin{cases} x = x(u,v,w) \\ y = y(u,v,w) \\ z = z(u,v,w). \end{cases} \tag{3.4}$$

Allora

$$\iiint_D f(x,y,z)\,dx\,dy\,dz =$$

$$= \iiint_{D'} f(x(u,v,w), y(u,v,w), z(u,v,w)) \,|\det \mathbf{DT}(u,v,w)|\,du\,dv\,dw$$

dove $\mathbf{DT}(u,v,w)$ *indica la matrice Jacobiana della trasformazione* (3.4).

In pratica, il teorema precedente viene applicato osservando che l'*elemento di volume* si trasforma secondo la legge:

$$dx\,dy\,dz = |\det \mathbf{DT}(u,v,w)|\,du\,dv\,dw.$$

Per esempio, per le *coordinate sferiche* (vedi cap. 4, par. 5) abbiamo (fig. 5.18a):

$$\begin{cases} x = \rho \sin\varphi \cos\theta \\ y = \rho \sin\varphi \sin\theta \\ z = \rho \cos\varphi \end{cases} \quad \text{con } \rho > 0,\ \varphi \in [0,\pi],\ \theta \in [0,2\pi)$$

$$dx\,dy\,dz = \rho^2 \sin\varphi\,d\rho\,d\varphi\,d\theta.$$

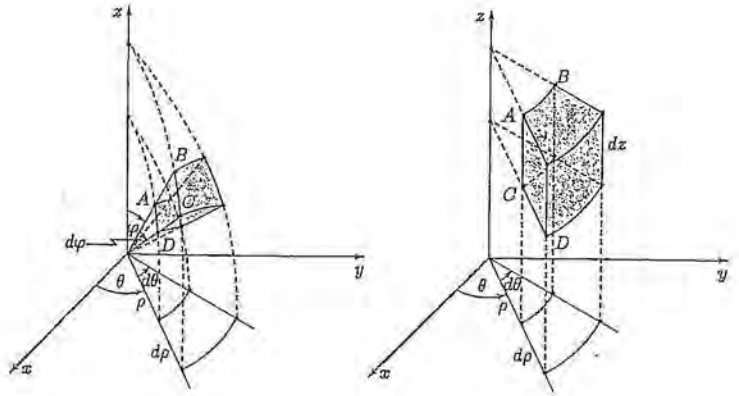


Figura 5.18. a) Elemento di volume in coordinate sferiche: $AB = d\rho$, $AD = \rho d\varphi$, $AC \simeq \rho \sin\varphi d\theta$. b) Elemento di volume in coordinate cilindriche: $AB = \rho d\theta$, $AC = dz$, $CD = d\rho$.

 © 978-88-08-06485-1

Per le *coordinate cilindriche* (fig. 5.18b):

$$\begin{cases} x = \rho\cos\theta \\ y = \rho\sin\theta \\ z = t \end{cases} \qquad \text{con } \rho \in [0,+\infty),\ \theta \in [0,2\pi),\ t \in \mathbb{R}$$

$$dx\,dy\,dz = \rho\,d\rho\,d\theta\,dt.$$

Interpretazioni geometriche e fisiche degli integrali tripli

Volume: se Ω è una regione tridimensionale, il suo volume $|\Omega|$ si può definire e quindi calcolare, come l'integrale triplo (se esiste):

$$|\Omega| = \iiint_\Omega dx\,dy\,dz.$$

Massa totale: se Ω è un corpo tridimensionale di densità $\rho(x,y,z)$, la sua massa totale M è uguale a:

$$M = \iiint_\Omega \rho(x,y,z)\,dx\,dy\,dz.$$

Baricentro: se Ω è un corpo tridimensionale di densità $\rho(x,y,z)$, il suo baricentro ha coordinate $(\overline{x},\overline{y},\overline{z})$ date dagli integrali:

$$\overline{x} = \frac{1}{M}\iiint_\Omega x\rho(x,y,z)\,dx\,dy\,dz \quad \text{ecc.}$$

(con M massa totale, calcolata come sopra). Se la densità è costante, il baricentro si dirà centroide e le formule si semplificano come segue:

$$\overline{x} = \frac{1}{|\Omega|}\iiint_\Omega x\,dx\,dy\,dz \quad \text{ecc.}$$

Momento d'inerzia: se Ω è un corpo tridimensionale di densità $\rho(x,y,z)$, il suo momento d'inerzia I rispetto a un certo asse è dato dall'integrale:

$$I = \iiint_\Omega \delta^2(x,y,z)\,\rho(x,y,z)\,dx\,dy\,dz$$

dove $\delta(x,y,z)$ è la distanza del punto (x,y,z) variabile dall'asse in questione. Per esempio, rispetto all'asse z si ha:

$$I = \iiint_\Omega (x^2+y^2)\,\rho(x,y,z)\,dx\,dy\,dz$$

e, se la densità è costante,

$$I = \frac{M}{|\Omega|}\iiint_\Omega (x^2+y^2)\,dx\,dy\,dz.$$

© 978-88-08-06485-1

Esempi

Calcoliamo:

$$I = \iiint_{[0,2]\times[0,1]\times[0,3]} x^2y^3z\,dx\,dy\,dz.$$

Si ha:

$$I = \int_0^2 x^2dx\int_0^1 y^3dy\int_0^3 zdz = \left[\frac{x^3}{3}\right]_0^2\left[\frac{y^4}{4}\right]_0^1\left[\frac{z^2}{2}\right]_0^3 = \frac{8}{3}\cdot\frac{1}{4}\cdot\frac{9}{2} = 3.$$

Sia D il cono definito da $x^2+y^2 \le (1-z)^2$, per $0 \le z \le 1$, e supponiamo che abbia densità $\rho(x,y,z) = z$. Calcoliamo la massa totale.

Rappresentiamo il dominio D nella forma:

$$D = \{(x,y,z) : 0 < z < 1-\sqrt{x^2+y^2},\ x^2+y^2 < 1\}.$$

Perciò:

$$\begin{aligned} M &= \iint_{x^2+y^2<1} dx\,dy\int_0^{1-\sqrt{x^2+y^2}} z\,dz = \\ &= \iint_{x^2+y^2<1} \frac{1}{2}(1-\sqrt{x^2+y^2})^2dx\,dy = \text{(in polari)} \\ &= 2\pi\int_0^1 \frac{1}{2}(1-\rho)^2\rho\,d\rho = \pi\int_0^1(\rho^3-2\rho^2+\rho)d\rho = \pi\left[\frac{\rho^4}{4}-\frac{2}{3}\rho^3+\frac{\rho^2}{2}\right]_0^1 = \frac{\pi}{12}. \end{aligned}$$

Calcoliamo il baricentro del cono dell'esempio precedente. Per simmetria, il baricentro starà sull'asse z, e quindi $\overline{x} = \overline{y} = 0$. Integriamo per fili:

$$\begin{aligned} \overline{z} &= \frac{1}{M}\iint_{x^2+y^2<1} dx\,dy\int_0^{1-\sqrt{x^2+y^2}} z^2\,dz = \\ &= \frac{12}{\pi}\cdot 2\pi\iint_{x^2+y^2<1}\frac{1}{3}(1-\sqrt{x^2+y^2})^3dx\,dy = \\ &= 8\int_0^1(1-\rho)^3\rho\,d\rho = \text{(con calcoli di routine...)} = \frac{2}{5}. \end{aligned}$$

Calcoliamo il momento d'inerzia, rispetto al suo asse, del cono omogeneo di massa M, raggio R e altezza h. Rappresentiamo il cono come: $x^2+y^2 \le (\frac{Rt}{h})^2$, $0 \le t \le h$ (fig. 5.19). Il momento d'inerzia è, integrando per strati:

$$\begin{aligned} I &= \frac{M}{\frac{1}{3}\pi R^2h}\int_0^h dt\iint_{x^2+y^2\le(\frac{Rt}{h})^2}(x^2+y^2)dx\,dy = \\ &= \frac{3M}{\pi R^2h}\int_0^h dt\left(2\pi\int_0^{\frac{Rt}{h}}\rho^3d\rho\right) = \frac{6M}{R^2h}\int_0^h\frac{1}{4}\left(\frac{Rt}{h}\right)^4dt = \\ &= \frac{3MR^2}{2h^5}\int_0^h t^4dt = \frac{3}{10}MR^2. \end{aligned}$$

 © 978-88-08-06485-1

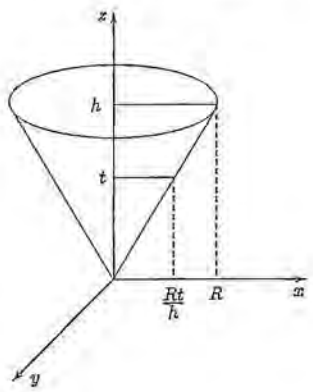


Figura 5.19.

4 DERIVAZIONE SOTTO IL SEGNO DI INTEGRALE

Nella deduzione di molti modelli matematici delle scienze applicate, ci si trova di fronte al problema di derivare integrali dipendenti da un parametro. Un esempio tipico è il seguente.

Esempio

Sia $\rho(x,y,z,t)$ la densità di un mezzo continuo in movimento, nel punto (x,y,z) all'istante t. La massa totale racchiusa all'istante t in una certa regione D dello spazio si può ottenere integrando in D la densità:

$$M(t) = \iiint_D \rho(x,y,z,t)\,dx\,dy\,dz.$$

Ci chiediamo ora con quale velocità varia, nel tempo, la massa $M(t)$ contenuta nella regione D. Per rispondere calcoliamo:

$$\frac{dM}{dt} = \frac{d}{dt}\iiint_D \rho(x,y,z,t)\,dx\,dy\,dz. \tag{4.1}$$

Ecco il problema: come si calcola la derivata, rispetto al "parametro" t, di questo integrale?

Ragioniamo così. L'integrale è un limite di somme; la derivata di una somma è la somma delle derivate; se anche l'operazione di limite può essere scambiata con quella di derivata, la derivata rispetto a t si può "portare sotto il segno di integrale". Si ha allora:

$$\frac{d}{dt}\iiint_D \rho(x,y,z,t)\,dx\,dy\,dz = \iiint_D \frac{\partial\rho}{\partial t}(x,y,z,t)\,dx\,dy\,dz. \tag{4.2}$$

Si osservi che la derivata rispetto a t sotto il segno di integrale è una derivata *parziale*, perché ρ dipende anche da altre variabili.

Sotto ipotesi ragionevoli, la (4.2) si può dimostrare. Vale il seguente:

TEOREMA 5.12 *Sia $\rho : D\times[a,b]\to\mathbb{R}$, con D dominio regolare di $\mathbb{R}^3$. Indicando con (x,t) i punti di $D\times[a,b]$, supponiamo che ρ e ρ_t siano continue in $D\times[a,b]$; allora per ogni $t\in[a,b]$ vale la (4.2).*

© 978-88-08-06485-1

Lo stesso enunciato vale, a maggior ragione, per integrali doppi o semplici.

Le ipotesi di questo teorema possono essere indebolite, comprendendo in particolare anche il caso in cui gli integrali coinvolti siano generalizzati; questo richiede però una certa cautela. Rinunciando a dare un enunciato preciso in questo caso, diciamo solo che quando si applica la (4.2) *bisogna sempre controllare che l'integrale a secondo membro sia convergente.*

Esempio

Mostriamo come la derivazione sotto il segno di integrale sia talvolta un metodo di calcolo di integrali definiti (a partire da altri già noti). Calcoliamo l'integrale (generalizzato) unidimensionale, dipendente dal parametro $a>0$:

$$\int_{-\infty}^{+\infty}\frac{dx}{x^2+a^2} = \left[\frac{1}{a}\operatorname{arctg}\frac{x}{a}\right]_{-\infty}^{+\infty} = \frac{\pi}{a}.$$

Deriviamo ora rispetto ad a ambo i membri dell'identità trovata:

$$-\frac{\pi}{a^2} = \frac{d}{da}\int_{-\infty}^{+\infty}\frac{dx}{x^2+a^2} = \int_{-\infty}^{+\infty}\frac{\partial}{\partial a}\left(\frac{1}{x^2+a^2}\right)dx = \int_{-\infty}^{+\infty} -\frac{2a}{(x^2+a^2)^2}dx.$$

Abbiamo quindi così calcolato l'integrale:

$$\int_{-\infty}^{+\infty}\frac{1}{(x^2+a^2)^2}dx = \frac{\pi}{2a^3},\ \text{per ogni } a>0.$$

Il calcolo dello stesso integrale determinando la primitiva di $\frac{1}{(x^2+a^2)^2}$ sarebbe stato più laborioso.

Esercizi

Stabilire se converge o meno l'integrale doppio generalizzato:

$$\iint_D x\,dx\,dy$$

con $D=\{(x,y) : x>0, 0<y<e^{-x}\}$.

(Notare che l'insieme di definizione è illimitato e la funzione integranda è *illimitata all'infinito*; sembrerebbe non esserci speranza di convergenza...).

Calcolare l'integrale doppio generalizzato:

$$\iint_{x^2+y^2\le 4}\log(x^2+y^2)dx\,dy.$$

Una sfera di raggio R e centro $(0,0,0)$ ha densità $f(x,y,z) = a\rho + b$, con $\rho = \sqrt{x^2+y^2+z^2}$. Calcolare la massa totale della sfera.

Suggerimento: Scrivere l'integrale triplo e introdurre le coordinate sferiche.

Calcolare massa totale e baricentro di un oggetto semisferico di raggio R e centro nell'origine, nell'ipotesi che la densità sia $f(x,y,z) = a\rho+b$ e la semisfera sia quella superiore.

Stabilire per quali $\alpha > 0$ converge l'integrale triplo generalizzato:

$$\iiint_{x^2+y^2+z^2<1} \frac{1}{(\sqrt{x^2+y^2+z^2})^\alpha} dx\,dy\,dz.$$

Rispondere poi alla stessa domanda per il dominio illimitato $x^2+y^2+z^2>1$.

Calcolare la massa totale del cilindro

$$x^2+y^2<R^2; \quad 0 \le z \le h,$$

supponendo che la sua densità sia espressa, in coordinate cilindriche, da $f(\rho,\theta,z) = a\rho+bz+c$.

Calcolare il momento d'inerzia di una sfera omogenea di raggio R e massa M rispetto a un asse passante per il centro.

Calcolare il momento d'inerzia di un cilindro omogeneo di raggio R, altezza h e massa M rispetto al suo asse.

Calcolare l'integrale triplo:

$$\iiint_D \frac{1}{(y+1)^3} dx\,dy\,dz$$

dove

$$D = \{(x,y,z) : 0<x<2, 0<z<1, 0<y<x+z\}.$$

Nella teoria del potenziale gioca un ruolo importante la funzione:

$$g(x,y,z) = \iiint_{x'^2+y'^2+z'^2<R^2} \frac{f(x',y',z')}{\sqrt{(x-x')^2+(y-y')^2+(z-z')^2}} dx'\,dy'\,dz'$$

definita a partire da una funzione f, continua su tutto il dominio di integrazione. Calcolare g_x, g_{xx} derivando sotto il segno di integrale. Rendersi conto che la seconda derivata non può essere calcolata in questo modo, in quanto porta a un integrale divergente nell'intorno di (x,y,z), mentre la derivata prima porta a un integrale convergente.

5 COMPLEMENTI

5.1 La funzione Gamma di Eulero

Ci occupiamo ora brevemente di un'importante *funzione speciale*, nota come *funzione Gamma di Eulero*. Questa funzione compare in molte questioni matematiche, (in particolare nell'espressione esatta del valore di certe costanti notevoli o di certi integrali) e in molte questioni fisiche e probabilistiche. Vedremo un'applicazione di questo tipo nel prossimo paragrafo di questi complementi, quando calcoleremo il volume della sfera n-dimensionale.

Poniamo, per $x>0$,

$$\Gamma(x) = \int_0^{+\infty} t^{x-1}e^{-t}dt. \tag{5.1}$$

L'integrale nella (5.1) è l'integrale generalizzato della funzione reale di variabile reale $f(t) = t^{x-1}e^{-t}$, che dipende dal parametro x; la funzione f è continua in $(0,\infty)$; per $t \to +\infty$ tende a zero molto rapidamente e pertanto risulta integrabile; per $t \to 0^+$ si ha:

$$f(t) \sim t^{x-1},$$

che risulta integrabile purché l'esponente $x-1$ sia >-1, ossia[11] per $x>0$. Quindi l'integrale in (5.1) è ben definito per ogni $x>0$ e il suo valore dipende dal parametro x. Pertanto è corretto dire che $\Gamma(x)$ è una funzione, definita in $(0,+\infty)$ a valori reali[12]; Γ viene detta appunto funzione Gamma di Eulero.

Il prossimo teorema riassume alcune proprietà importanti di questa funzione.

TEOREMA 5.13
(a) La funzione Γ di Eulero è derivabile infinite volte in $(0,+\infty)$.
(b) Vale la formula di ricorrenza:

$$\Gamma(x) = (x-1)\Gamma(x-1) \text{ per ogni } x>1.$$

(c) Per ogni n intero positivo è:

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

(d) Vale l'identità:

$$\Gamma(x) = 2\int_0^{+\infty} u^{2x-1}e^{-u^2}du.$$

Le proprietà (a) e (c) ci dicono che la funzione Γ si può vedere come *una funzione regolare* (infinitamente derivabile) *che interpola i valori del fattoriale*, cioè che assume i valori $(x-1)!$ quando il suo argomento x è intero, ma è definita su tutta la semiretta $x>0$. Non stupisce di conseguenza che, alla pari dell'espressione $n!$, la funzione Gamma compaia in tante formule matematiche notevoli.

DIMOSTRAZIONE. (a) Nel paragrafo 4 abbiamo discusso il problema della derivazione sotto il segno di integrale. Anche se, nel caso degli integrali generalizzati, non abbiamo enunciato un risultato preciso, possiamo osservare che

$$\frac{d^k}{dx^k}\left[t^{x-1}e^{-t}\right] = t^{x-1}e^{-t}(\log t)^k$$

e che l'integrale generalizzato

$$\int_0^{+\infty} t^{x-1}e^{-t}(\log t)^k\,dt$$

converge per ogni k. Si può dimostrare (non lo facciamo) che la derivata k-esima della funzione Γ si può effettivamente calcolare derivando sotto il segno di integrale; pertanto Γ è derivabile infinite volte.

(b) Integriamo per parti nell'integrale generalizzato:

$$\Gamma(x) = \int_0^{+\infty} \underbrace{\left(t^{x-1}\right)}_{f}\underbrace{\left(e^{-t}\right)}_{g'}dt = \left[-t^{x-1}e^{-t}\right]_0^{+\infty} + \int_0^{+\infty}(x-1)t^{x-2}e^{-t}dt =$$
$$= 0 + (x-1)\Gamma(x-1),$$

dove si è inteso (coerentemente alla definizione di integrale generalizzato):

$$\left[-t^{x-1}e^{-t}\right]_0^{+\infty} = \lim_{r\to+\infty}\left(-r^{x-1}e^{-r}\right) - \lim_{\varepsilon\to 0}\left(-\varepsilon^{x-1}e^{-\varepsilon}\right) = 0+0$$

perché per ipotesi $x>1$.

[11] Si vedano i criteri di integrabilità studiati per gli integrali generalizzati in una variabile nel volume 1, capitolo 6, paragrafo 8.

[12] Nella matematica superiore si definisce la funzione Γ anche se x è complesso, purché con parte reale positiva; noi ci limiteremo al caso di x *reale* e positivo.

 © 978-88-08-06485-1

(c) Dalla (b) abbiamo:

$$\Gamma(n) = (n-1)\,\Gamma(n-1) \tag{5.2}$$

mentre, direttamente dalla definizione, osserviamo che:

$$\Gamma(1) = 1. \tag{5.3}$$

Per ricorrenza, dalle (5.2)-(5.3) segue che $\Gamma(n) = (n-1)!$

(d) È sufficiente calcolare l'integrale che definisce $\Gamma(x)$ col cambio di variabili $t = u^2$. ♦

Funzione Gamma sui semiinteri

Sono utili talvolta (vedremo un esempio nel prossimo paragrafo) i valori che la $\Gamma(x)$ assume per x *semiintero*, ossia quando $x = n/2$ con n intero positivo.

Se n è *pari*, $n = 2k$, sappiamo già che:

$$\Gamma\left(\frac{n}{2}\right) = \Gamma(k) = (k-1)! = \left(\frac{n}{2}-1\right)!$$

Occupiamoci quindi del caso n *dispari*, cioè $n = 2k+1$ con k intero positivo. Il caso più semplice è $k = 0$: dalla (d) del teorema precedente si ha:

$$\Gamma\left(\frac{1}{2}\right) = 2\int_0^{+\infty} u^{2\cdot\frac{1}{2}-1} e^{-u^2}\,du = 2\int_0^{+\infty} e^{-u^2}\,du = \int_{\mathbb{R}} e^{-u^2}\,du.$$

Abbiamo calcolato questo integrale nell'Esempio 2.1 del paragrafo 2, provando che vale $\sqrt{\pi}$. Quindi:

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}.$$

Più in generale, usando più volte la formula di ricorrenza per la Gamma (punto (b) del teorema) abbiamo, per $k \geq 1$:

$$\begin{aligned}
\Gamma\left(\frac{n}{2}\right) &= \Gamma\left(k+\frac{1}{2}\right) = \left(k-1+\frac{1}{2}\right)\Gamma\left(k-1+\frac{1}{2}\right) = \\
&= \left(k-1+\frac{1}{2}\right)\left(k-2+\frac{1}{2}\right)\Gamma\left(k-2+\frac{1}{2}\right) = \dots \\
&= \left(k-1+\frac{1}{2}\right)\left(k-2+\frac{1}{2}\right)\cdots\left(2+\frac{1}{2}\right)\left(1+\frac{1}{2}\right)\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right) = \\
&= \frac{(2k-1)\,(2k-3)\,(2k-5)\cdots 5\cdot 3}{2^{k-1}}\sqrt{\pi} = \\
&= \frac{(2k-1)!!}{2^k}\sqrt{\pi},
\end{aligned}$$

dove il simbolo $(2k-1)!!$ ("semifattoriale") indica per definizione il prodotto di tutti gli interi dispari da $(2k-1)$ in giù:

$$(2k-1)!! = (2k-1)\,(2k-3)\,(2k-5)\cdots 5\cdot 3.$$

© 978-88-08-06485-1

Possiamo quindi scrivere:

$$\begin{aligned}
&\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}, \\
&\Gamma\left(\frac{n}{2}\right) = \frac{(n-2)!!}{2^{\frac{n-1}{2}}}\sqrt{\pi} \quad \text{per } n \text{ dispari} \geq 3, \\
&\Gamma\left(\frac{n}{2}\right) = \left(\frac{n-2}{2}\right)! \quad \text{per } n \text{ pari.}
\end{aligned} \tag{5.4}$$

5.2 Definizioni e proprietà elementari degli integrali in $\mathbb{R}^n$

La definizione di integrale per una funzione di n variabili a valori reali, limitata e definita su un insieme limitato, generalizza in modo naturale quella di integrale doppio e triplo e non la ripeteremo in dettaglio: schematicamente, si considerano funzioni definite su un (iper)parallelepipedo (che possiamo chiamare anche n-scatola),

$$Q = [a_1, b_1] \times [a_2, b_2] \times \cdots \times [a_n, b_n],$$

si suddivide ogni intervallino in k parti uguali (e quindi l'n-scatola iniziale in k^n n-scatole uguali), si definisce in modo naturale una successione $\{s_k\}$ di somme di Cauchy-Riemann di $f : Q \to \mathbb{R}$ e si dà la definizione di funzione integrabile e di integrale. Successivamente, si estende la definizione a funzioni definite su un insieme limitato $\Omega \subset \mathbb{R}^n$, al solito modo.

Osserviamo solo che, dal punto di vista notazionale, un integrale in n variabili viene solitamente indicato con un unico simbolo di integrale, al modo seguente:

$$\int_\Omega f(x)\,dx$$

dove quindi il simbolo $\int_\Omega$ indica già l'integrale multiplo.

Valgono inalterate le proprietà formali dell'integrale (linearità, monotonia...), così come il fatto che le funzioni continue su un dominio regolare risultino integrabili.

La *tecnica di riduzione* degli integrali multipli a integrali in dimensione inferiore si adatta in maniera naturale, ma è complicata dalle numerose possibilità che ora sostituiscono le sole due forme di integrazione "per fili" e "per strati" che si presentavano per gli integrali tripli.

Il *teorema sul cambio di variabili* negli integrali multipli vale in maniera analoga e anche in questo caso il modulo del determinante jacobiano della trasformazione compare nell'espressione dell'elemento infinitesimo di volume.

In questo contesto, vogliamo ora fare qualche osservazione sull'uso delle *coordinate polari* in $\mathbb{R}^n$ e su alcune questioni collegate.

Coordinate polari in $\mathbb{R}^n$

La trasformazione di coordinate che lega coordinate euclidee e coordinate polari (dette anche ipersferiche) in $\mathbb{R}^n$ è data da:

$$\begin{cases}
x_1 = \rho\cos\phi_1 \\
x_2 = \rho\sin\phi_1\cos\phi_2 \\
x_3 = \rho\sin\phi_1\sin\phi_2\cos\phi_3 \\
\dots \\
x_{n-1} = \rho\sin\phi_1\dots\sin\phi_{n-2}\cos\phi_{n-1} \\
x_n = \rho\sin\phi_1\dots\sin\phi_{n-2}\sin\phi_{n-1}
\end{cases} \tag{5.5}$$

dove $\rho = |x|$, l'ultimo angolo ϕ_{n-1} varia in $[0, 2\pi)$, mentre tutti gli altri variano in $[0, \pi]$.

 © 978-88-08-06485-1

Si può dimostrare che l'elemento infinitesimo di volume ha la forma:

$$dx_1 dx_2 \dots dx_n = \rho^{n-1} \sin^{n-2}\phi_1 \sin^{n-3}\phi_2 \dots \sin\phi_{n-1} d\rho\, d\phi_1\, d\phi_2 \dots d\phi_{n-1}.$$

Per certi scopi, è sufficiente ricordare che l'elemento infinitesimo ha la forma:

$$dx_1 dx_2 \dots dx_n = \rho^{n-1} d\rho \cdot \omega(\phi_1, \phi_2, \dots, \phi_{n-1})\, d\phi_1\, d\phi_2 \dots d\phi_{n-1}$$

per un'opportuna funzione ω. (Quest'ultimo fatto si capisce facilmente osservando la struttura delle equazioni (5.5) e quindi la forma che assumerà la matrice Jacobiana).

Integrale di una funzione radiale

Calcoliamo l'integrale di una funzione radiale (cioè dipendente solo da $|x|$) su un dominio a simmetria radiale, per esempio una sfera. Si ha:

$$\int_{\{x\in\mathbb{R}^n:|x|<R\}} f(|x|)\,dx = \int_0^R \rho^{n-1} f(\rho)\,d\rho \int_\Sigma \omega(\phi_1, \phi_2, \dots, \phi_{n-1})\, d\phi_1\, d\phi_2 \dots d\phi_{n-1}$$

(dove abbiamo indicato sinteticamente con $\int_\Sigma$ l'integrale $n-1$ dimensionale nell'insieme di variabilità di tutte le variabili angolari). Ora l'integrale interno è una costante ω_n (dipendente dalla dimensione n, ma indipendente dalla funzione radiale f) e l'identità precedente può essere riscritta nella forma più semplice:

$$\int_{\{x\in\mathbb{R}^n:|x|<R\}} f(|x|)\,dx = \omega_n \int_0^R \rho^{n-1} f(\rho)\,d\rho. \tag{5.6}$$

La costante ω_n ha il significato geometrico di *area della superficie sferica di raggio unitario* e ora calcoleremo esplicitamente il suo valore.

Volume e superficie laterale della sfera in $\mathbb{R}^n$

Per calcolare ω_n, utilizziamo ancora la (5.6), nel caso di una particolare funzione radiale di cui, per altra via, conosciamo l'integrale: la gaussiana $e^{-|x|^2}$. Applicando la (5.6) per $R \to \infty$:

$$\int_{\mathbb{R}^n} e^{-|x|^2} dx = \omega_n \int_0^{+\infty} \rho^{n-1} e^{-\rho^2} d\rho. \tag{5.7}$$

Il primo membro della (5.7) si può trattare con lo stesso ragionamento fatto nel caso $n = 2$ nel paragrafo 2:

$$\begin{aligned}\int_{\mathbb{R}^n} e^{-|x|^2} dx &= \int_{\mathbb{R}^n} e^{-x_1^2} e^{-x_2^2} \dots e^{-x_n^2} dx_1 dx_2 \dots dx_n = \\ &= \int_{\mathbb{R}} e^{-x_1^2} dx_1 \int_{\mathbb{R}} e^{-x_2^2} dx_2 \cdot \dots \cdot \int_{\mathbb{R}} e^{-x_n^2} dx_n = \\ &= \left(\int_{\mathbb{R}} e^{-t^2} dt\right)^n = \left(\sqrt{\pi}\right)^n = \pi^{n/2}.\end{aligned}$$

Confrontando il secondo membro della (5.7) con la (d) del teorema 5.13 si vede che:

$$\Gamma\left(\frac{n}{2}\right) = 2\int_0^{+\infty} \rho^{n-1} e^{-\rho^2} d\rho,$$

© 978-88-08-06485-1

dove Γ è la funzione Gamma di Eulero, perciò:

$$\pi^{n/2} = \frac{\omega_n}{2}\Gamma\left(\frac{n}{2}\right)$$

da cui ricaviamo:

$$\omega_n = \frac{2\pi^{n/2}}{\Gamma\left(\frac{n}{2}\right)}.$$

Ricordiamo che nelle (5.4) abbiamo calcolato esplicitamente i valori $\Gamma\left(\frac{n}{2}\right)$ (funzione Gamma ad argomento semiintero).

Il calcolo di ω_n è utile anche per quello di un'altra costante notevole: il *volume della sfera di raggio R in $\mathbb{R}^n$*. Indicando con

$$B_R = \{x \in \mathbb{R}^n : |x| < R\},$$

si ha infatti, passando in coordinate polari e applicando la (5.6):

$$\mathrm{Vol}(B_R) = \int_{B_R} dx = \omega_n \int_0^R \rho^{n-1} d\rho = \omega_n \frac{R^n}{n} = \frac{2\pi^{n/2}}{n\Gamma\left(\frac{n}{2}\right)} R^n.$$

L'ultima espressione trovata si può riscrivere in vari modi, osservando per esempio che

$$\frac{2\pi^{n/2}}{n\Gamma\left(\frac{n}{2}\right)} = \frac{\pi^{n/2}}{\frac{n}{2}\Gamma\left(\frac{n}{2}\right)} = \frac{\pi^{n/2}}{\Gamma\left(\frac{n}{2}+1\right)}.$$

In definitiva otteniamo, tenendo conto anche delle (5.4), che:

$$\mathrm{Vol}(B_R) = \frac{\pi^{n/2}}{\frac{n}{2}\Gamma\left(\frac{n}{2}\right)} R^n = \begin{cases} \dfrac{\pi^{n/2}}{\left(\frac{n}{2}\right)!} R^n & \text{per } n \text{ pari} \\ \dfrac{2^{\frac{n+1}{2}} \pi^{\frac{n-1}{2}}}{n!!} R^n & \text{per } n \text{ dispari.} \end{cases}$$

A titolo di esempio:

per $n = \dots$	$\mathrm{vol}(B_R) = \dots$	$\mathrm{area}(\partial B_R)$
2	πR^2	$2\pi R$
3	$\frac{4}{3}\pi R^3$	$4\pi R^2$
4	$\frac{\pi^2}{2} R^4$	$2\pi^2 R^3$
5	$\frac{8}{15}\pi^2 R^5$	$\frac{8}{3}\pi^2 R^4$

Osserviamo, che per ogni dimensione n, il numero π (elevato a un'opportuna potenza intera) compare nella formula che assegna il volume della sfera. Notiamo che, curiosamente, per R fissato e $n \to \infty$, $\mathrm{Vol}(B_R) \to 0$. È anche utile osservare che, in ogni dimensione n:

$$\mathrm{area}(\partial B_R) = \frac{d}{dR}\left(\mathrm{vol}(B_R)\right).$$

L'area della superficie sferica (non inaspettatamente) è la derivata rispetto al raggio del volume della stessa sfera.

6 Campi vèttoriali

1 CAMPI VETTORIALI E INTEGRALI DI LINEA DI SECONDA SPECIE

1.1 Linee di campo

Come abbiamo visto nel capitolo 4, paragrafo 1.3, un *campo vettoriale* è una funzione che a ogni punto dello spazio fisico e in ogni istante assegna un vettore, che può avere il significato di *velocità*, di *forza* agente nel punto assegnato (campo gravitazionale, campo elettrostatico...) o altri ancora. Se il campo *non* varia nel tempo si dice *stazionario*.

Volendo tenere distinti tempo e spazio, un campo vettoriale è rappresentabile con una funzione $\mathbf{F} : I \times \Omega \to \mathbb{R}^m$, dove $I \subseteq \mathbb{R}$ è un intervallo di tempo e $\Omega \subseteq \mathbb{R}^n$ è un insieme di punti dello spazio. Si noti come gli spazi $\mathbb{R}^n$ ed $\mathbb{R}^m$ siano "pensati" in due modi differenti: il primo come insieme di punti dello spazio, il secondo come insieme di vettori.

In questo contesto, una funzione $f : I \times \Omega \to \mathbb{R}$ viene anche detta *campo scalare*. Così, per esempio, si può dire che nella descrizione termodinamica di un gas, la temperatura $\theta = \theta(t, x, y, z)$ e la pressione $p = p(t, x, y, z)$ sono campi scalari, mentre la velocità delle particelle $\mathbf{v} = \mathbf{v}(t, x, y, z)$ è un campo vettoriale.

Salvo avviso contrario lavoreremo prevalentemente con campi stazionari in 3 dimensioni. La riduzione a $n = 2$ non dovrebbe comportare difficoltà al lettore. Tutto quanto diremo per campi vettoriali $\mathbf{F} : \Omega \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ vale, salvo avviso contrario, per campi $\mathbf{F} : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^n$ con n qualsiasi.

Per un campo stazionario $\mathbf{F} : \Omega \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ useremo una delle due notazioni:

$$\mathbf{F}(x, y, z) = F_1(x, y, z)\mathbf{i} + F_2(x, y, z)\mathbf{j} + F_3(x, y, z)\mathbf{k}$$

oppure

$$\mathbf{F}(x, y, z) = (F_1(x, y, z), F_2(x, y, z), F_3(x, y, z)),$$

dove le componenti $F_i : \Omega \subseteq \mathbb{R}^3 \to \mathbb{R}$ sono campi scalari. In generale assumeremo che Ω sia un *aperto connesso* ed $\mathbf{F} \in C^1(\Omega)$, cioè abbia componenti derivabili con continuità in Ω.

Definizione 6.1 Dato un campo vettoriale $\mathbf{F} : \Omega \subseteq \mathbb{R}^3 \to \mathbb{R}^3$, con $\mathbf{F} \in C^1(\Omega)$, chiameremo *linea di campo* di $\mathbf{F}$ una qualsiasi curva regolare tangente in ogni punto

a $\mathbf{F}$. Se $\mathbf{F}$ rappresenta un campo di forza, le linee di campo si dicono anche *linee di forza*; se è un campo di velocità, si dicono *linee di flusso*.

Sia $\mathrm{r} = \mathrm{r}(t) = (x(t), y(t), z(t))$ una linea del campo $\mathbf{F}$; il fatto che nel punto $\mathrm{r}(t)$ il campo $\mathbf{F}$ sia tangente alla linea significa che il vettore $\mathbf{F}(\mathrm{r}(t))$ è parallelo al vettore $\mathrm{r}'(t)$; esiste dunque una funzione (scalare) $\lambda(t)$ tale che

$$\mathrm{r}'(t) = \lambda(t)\mathbf{F}(\mathrm{r}(t)).$$

Per determinare le linee di campo occorre dunque risolvere il *sistema di equazioni differenziali del primo ordine*

$$\begin{cases} x' = \lambda F_1(x, y, z) \\ y' = \lambda F_2(x, y, z) \\ z' = \lambda F_3(x, y, z) \end{cases}$$

equivalente a

$$\frac{dx}{F_1(x, y, z)} = \frac{dy}{F_2(x, y, z)} = \frac{dz}{F_3(x, y, z)}.$$

Imponendo una "condizione iniziale", ossia il passaggio della curva per un punto (x_0, y_0, z_0), il sistema avrà in generale un'unica soluzione (locale), purché il campo sia ben definito, regolare e non nullo[1]. Come conseguenza si ottiene che, in generale, *per ogni punto dello spazio passa una e una sola linea di campo*. Le linee di campo costituiscono dunque una famiglia di linee, a due a due prive di intersezioni.

Esempio

Consideriamo il campo gravitazionale generato da una massa puntiforme posta nell'origine:

$$\mathbf{F}(x, y, z) = -k\frac{x\mathrm{i} + y\mathrm{j} + z\mathrm{k}}{\rho^3} \qquad k > 0,\ \rho = \sqrt{x^2 + y^2 + z^2}.$$

Per determinare le linee di campo occorre risolvere il sistema:

$$\frac{dx}{x} = \frac{dy}{y} = \frac{dz}{z}$$

ovvero

$$\begin{cases} \log|x| = \log|y| + a \\ \log|z| = \log|y| + b \end{cases}$$

e quindi

$$\begin{cases} x = k_1 y \\ z = k_2 y. \end{cases}$$

Nell'esempio precedente, le linee di campo sono semplicemente le semirette con estremo nell'origine. Più precisamente, data la natura del campo $\mathbf{F}$, esse *sono attratte dall'origine*, che perciò si dice *pozzo* per le linee del campo. Un punto da cui le linee

[1] Questo seguirà dai risultati generali sul problema di Cauchy per sistemi di equazioni differenziali ordinarie, che discuteremo nel capitolo 8, paragrafo 2.

di campo *nascono* si dice invece *sorgente*. Pozzi e sorgenti sono *punti singolari* per il campo, ovvero punti in cui il campo non è definito, oppure non è regolare, oppure si annulla (altrimenti, per quanto notato in precedenza, per quei punti dovrebbe passare un'unica linea del campo).

Esempi

Consideriamo il campo piano di velocità $\mathrm{v} = (-y, x)$. Le linee di campo si ottengono risolvendo l'equazione

$$\frac{dx}{-y} = \frac{dy}{x},$$

ossia

$$x\,dx = -y\,dy$$

che dà

$$\frac{1}{2}x^2 = -\frac{1}{2}y^2 + c,$$

ossia

$$x^2 + y^2 = 2c.$$

Perciò le linee di campo sono circonferenze concentriche; per ogni punto del piano ne passa una e una sola, a eccezione dell'origine, in cui la linea degenera in un punto: si noti che il campo si annulla proprio in questo punto. Questo punto singolare non è né un pozzo né una sorgente.

Sia $U : \mathbb{R}^2 \to \mathbb{R}$ un campo scalare (differenziabile) e sia $\mathbf{F} = \nabla U$. Allora $\mathbf{F} : \mathbb{R}^2 \to \mathbb{R}^2$ è un campo vettoriale. Le linee di campo hanno la direzione di ∇U che, lo ricordiamo, è perpendicolare in ogni punto alle linee di livello di U. Le linee di campo di $\mathbf{F}$ sono ortogonali alle curve di livello del campo scalare U.

1.2 Gradiente, rotore e divergenza

Oltre all'operatore gradiente e all'operatore di Laplace, definiti su campi scalari rispettivamente dalle formule

$$\text{grad} = \nabla = \mathrm{i}\frac{\partial}{\partial x} + \mathrm{j}\frac{\partial}{\partial y} + \mathrm{k}\frac{\partial}{\partial z},$$

e

$$\Delta = \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2},$$

altri due operatori differenziali di estrema importanza sia teorica che applicativa sono *il rotore* e *la divergenza*. Cominciamo a introdurli insieme ad alcune formule di uso frequente, riservandoci di ritornare sulla loro interpretazione fisica più avanti.

L'operatore rotore

Quest'operatore agisce su campi vettoriali tridimensionali.

© 978-88-08-06485-1

Definizione 6.2 Sia $\mathbf{F} : \Omega \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ un campo vettoriale di classe $C^1(\Omega)$. Si definisce *rotore* di $\mathbf{F}$ il campo:

$$\mathrm{rot}\mathbf{F} = \nabla \times \mathbf{F} = \left(\frac{\partial F_3}{\partial y} - \frac{\partial F_2}{\partial z}\right)\mathbf{i} + \left(\frac{\partial F_1}{\partial z} - \frac{\partial F_3}{\partial x}\right)\mathbf{j} + \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\mathbf{k}$$
$$= \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ F_1 & F_2 & F_3 \end{vmatrix}.$$

Il simbolo $\nabla \times \mathbf{F}$ rappresenta il prodotto vettoriale formale tra l'operatore ∇ e il campo $\mathbf{F}$.

Se $\nabla \times \mathbf{F} = 0$, il campo $\mathbf{F}$ si dice *irrotazionale*. Nel caso particolare di un campo piano $\mathbf{F}(x,y) = \mathbf{i}F_1(x,y) + \mathbf{j}F_2(x,y)$, si ottiene

$$\nabla \times \mathbf{F} = \left(\frac{\partial F_2}{\partial x} - \frac{\partial F_1}{\partial y}\right)\mathbf{k}.$$

In questo caso, quindi, $\nabla \times \mathbf{F}$ è perpendicolare al piano in cui "vive" il campo $\mathbf{F}$; si noti che il concetto di rotore ha senso solo se ambientato nello spazio $\mathbb{R}^3$.

Il vettore $\nabla \times \mathbf{F}$, calcolato in un punto p, fornisce una misura di quanto il campo $\mathbf{F}$ "ruoti" attorno al punto p.

Esempio

Consideriamo il campo di velocità di un fluido in moto rotatorio uniforme attorno all'asse z. Se $\Omega = \Omega\mathbf{k}$ è il vettore velocità angolare e $\mathbf{r}(x,y) = x\mathbf{i} + y\mathbf{j}$ è il vettore posizione del punto rispetto all'asse di rotazione, la velocità nel punto (x,y) è data da:

$$\mathbf{v} = \Omega \times \mathbf{r} = -y\Omega\,\mathbf{i} + x\Omega\,\mathbf{j}.$$

Il rotore del campo di velocità è allora:

$$\nabla \times \mathbf{v} = \left(\frac{\partial}{\partial x}(x\Omega) - \frac{\partial}{\partial y}(-y\Omega)\right)\mathbf{k} = 2\Omega\mathbf{k}$$

ovvero

$$\Omega = \frac{1}{2}\nabla \times \mathbf{v}$$

Il vettore velocità angolare è quindi $\frac{1}{2}$ del rotore della velocità. Generalizzando questo esempio, se $\mathbf{v}$ è il campo di velocità in un mezzo continuo (per esempio un fluido), il vettore $\frac{1}{2}\nabla \times \mathbf{v}$ ha il significato di velocità angolare locale del moto; è un indice della vorticosità locale del moto (ma si veda l'esercizio 1).

L'operatore divergenza

Un altro operatore differenziale che agisce sui campi vettoriali è l'operatore divergenza.

Definizione 6.3 Sia $\mathbf{F} : \Omega \subseteq \mathbb{R}^n \to \mathbb{R}^n$ un campo vettoriale $C^1(\Omega)$, $\mathbf{F} = (F_1, F_2, \ldots, F_n)$. Si definisce *divergenza* di $\mathbf{F}$ il campo scalare:

$$\mathrm{div}\mathbf{F} = \nabla \cdot \mathbf{F} = \sum_{i=1}^{n} \frac{\partial F_i}{\partial x_i}.$$

© 978-88-08-06485-1

Per esempio, se $\mathbf{F} = F_1(x,y,z)\mathbf{i} + F_2(x,y,z)\mathbf{j} + F_3(x,y,z)\mathbf{k}$ è un campo vettoriale in $\mathbb{R}^3$,

$$\nabla \cdot \mathbf{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}$$

e $\nabla \cdot \mathbf{F}$ risulta un campo scalare, cioè una funzione reale di 3 variabili. Il simbolo $\nabla \cdot \mathbf{F}$ denota il prodotto scalare formale dell'operatore ∇ e il campo $\mathbf{F}$.

Anche la divergenza di un campo, come il rotore, ha un significato fisico, che si potrà comprendere più avanti, quando incontreremo il teorema della divergenza (v. paragrafo 5). Per ora possiamo dire, senza darne una giustificazione, che la divergenza di un campo segnala punto per punto l'intensità delle sorgenti (o pozzi) del campo stesso: per esempio, la divergenza del campo gravitazionale segnala la densità di materia nel punto (sorgente del campo), la divergenza del campo "corrente di calore" segnala la presenza di sorgenti (o pozzi) di calore in un mezzo ecc.

Un campo a divergenza nulla in Ω si dice *solenoidale* in Ω.

Esempio

Calcoliamo la divergenza del campo gravitazionale generato da una massa puntiforme posta nell'origine; abbiamo visto che il campo è:

$$\mathbf{F}(x,y,z) = -k\frac{x\mathbf{i} + y\mathbf{j} + z\mathbf{k}}{\rho^3}.$$

Calcoliamo, per $\rho \neq 0$:

$$\nabla \cdot \mathbf{F} = -k\left\{\partial_x\left(\frac{x}{\rho^3}\right) + \partial_y\left(\frac{y}{\rho^3}\right) + \partial_z\left(\frac{z}{\rho^3}\right)\right\} =$$
$$= -k\left\{\frac{\rho^3 - 3x^2\rho}{\rho^6} + \frac{\rho^3 - 3y^2\rho}{\rho^6} + \frac{\rho^3 - 3z^2\rho}{\rho^6}\right\} =$$
$$= -\frac{k}{\rho^6}\left\{3\rho^3 - 3\left(x^2 + y^2 + z^2\right)\rho\right\} = 0.$$

Pertanto, fuori dall'unico punto in cui c'è una massa, la divergenza del campo gravitazionale è nulla. Nel punto in cui si trova la massa puntiforme, la divergenza non è definita, perché il campo stesso non è definito (il modulo del campo tende a ∞ nell'origine). Si può concludere che il campo è solenoidale fuori dall'origine.

Identità differenziali

Vogliamo ora segnalare alcune importanti relazioni tra gli operatori gradiente, rotore, divergenza. Va sempre tenuto presente che:

- *l'operatore gradiente trasforma un campo scalare in un campo vettoriale;*
- *l'operatore rotore trasforma un campo vettoriale (tridimensionale) in un altro campo vettoriale (tridimensionale);*
- *l'operatore divergenza trasforma un campo vettoriale in un campo scalare.*

Ciò significa che se $\mathbf{F}$ è un campo vettoriale e u è un campo scalare, entrambi di classe $C^2(\Omega)$, si possono calcolare le seguenti composizioni di operatori:

$$\mathrm{div}\,\mathrm{rot}\,\mathbf{F} \qquad \mathrm{rot}\,\mathrm{grad}\,u \qquad \mathrm{div}\,\mathrm{grad}\,u \qquad \mathrm{grad}\,\mathrm{div}\,\mathbf{F}$$

mentre non hanno senso, per esempio, le scritture $\mathrm{rot}\,\mathrm{div}\,\mathbf{F}$, $\mathrm{grad}\,\mathrm{rot}\,\mathbf{F}$ ecc.

Alcune di queste composizioni hanno un risultato notevole:

PROPOSIZIONE 6.1 *Siano $u, \mathbf{F}$, rispettivamente, un campo scalare e un campo vettoriale (tridimensionale) di classe $C^2(\Omega)$. Allora:*

$$\nabla \times (\nabla u) = rot\ grad\ u = 0 \tag{1.1}$$

$$\nabla \cdot (\nabla \times \mathbf{F}) = div\ rot\ \mathbf{F} = 0 \tag{1.2}$$

$$\nabla \cdot (\nabla u) = div\ grad\ u = \Delta u. \tag{1.3}$$

La (1.3) vale per $u : \mathbb{R}^n \to \mathbb{R}$ con n qualunque.

DIMOSTRAZIONE.

(??). Sia $\mathbf{F} = \nabla u$. Allora:

$$\nabla \times \mathbf{F} = \left(\frac{\partial^2 u}{\partial y \partial z} - \frac{\partial^2 u}{\partial z \partial y}\right)\mathbf{i} + \left(\frac{\partial^2 u}{\partial z \partial x} - \frac{\partial^2 u}{\partial x \partial z}\right)\mathbf{j} + \left(\frac{\partial^2 u}{\partial y \partial x} - \frac{\partial^2 u}{\partial x \partial y}\right)\mathbf{k} = 0$$

per il teorema di Schwarz sull'uguaglianza delle derivate seconde miste.

(1.2).

$$\text{div rot } \mathbf{F} = \nabla \cdot (\nabla \times \mathbf{F}) = \begin{vmatrix} \partial_x & \partial_y & \partial_z \\ \partial_x & \partial_y & \partial_z \\ F_1 & F_2 & F_3 \end{vmatrix} = 0$$

ancora per l'uguaglianza delle derivate seconde miste.

(1.3).

$$\nabla \cdot (\nabla u) = \sum_{i=1}^{n} \frac{\partial}{\partial x_i}\left(\frac{\partial u}{\partial x_i}\right) = \sum_{i=1}^{n} \frac{\partial^2 u}{\partial x_i^2} = \Delta u. \qquad \blacklozenge$$

Vi sono molte altre identità differenziali che riguardano l'azione di questi operatori su *due* campi scalari o vettoriali. Ne scriviamo qualcuna:

$$\begin{aligned} \nabla \cdot (f\,\mathbf{G}) &= \nabla f \cdot \mathbf{G} + f\,(\nabla \cdot \mathbf{G}) \\ \nabla \cdot (f \nabla g) &= f \Delta g + \nabla f \cdot \nabla g \\ \nabla \cdot (f\,\nabla \times \mathbf{G}) &= \nabla f \cdot (\nabla \times \mathbf{G}) \\ \nabla \times (f\mathbf{G}) &= \nabla f \times \mathbf{G} + f\,(\nabla \times \mathbf{G}) \\ \nabla \times (f \nabla g) &= \nabla f \times \nabla g \\ \nabla \times (\nabla \times \mathbf{G}) &= \nabla\,(\nabla \cdot \mathbf{G}) - \Delta \mathbf{G} \end{aligned} \tag{1.4}$$

Qui f, g sono campi scalari, $\mathbf{G}$ è un campo vettoriale, tutti i campi sono derivabili quante volte occorre. Il lettore è invitato a verificare le (1.4) per esercizio[2]. Tutte le identità seguono applicando le definizioni degli operatori differenziali, la regola di derivazione del prodotto e le identità scritte nella Proposizione 6.1. Per il rotore, è utile usare la definizione come determinante formale.

[2]Nell'ultima delle (1.4), il simbolo $\Delta\mathbf{G}$ indica il vettore ottenuto applicando l'operatore Δ a ciascuna componente di $\mathbf{G}$.

1.3 Integrale di linea di un campo vettoriale. Lavoro e circuitazione.

Un concetto fisico fondamentale legato ai campi vettoriali, soprattutto nella loro interpretazione come campi di forza, è quello di lavoro. Il *lavoro elementare compiuto da una forza* $\mathbf{F}$ dovuto a uno spostamento infinitesimo $d\mathbf{r} = dx\mathbf{i} + dy\mathbf{j} + dz\mathbf{k}$ del suo punto di applicazione è, per definizione,

$$dL = \mathbf{F} \cdot d\mathbf{r} = F_1 dx + F_2 dy + F_3 dz.$$

Questa scrittura formale acquista un significato compiuto nella seguente:

DEFINIZIONE 6.4 (LAVORO DI UN CAMPO VETTORIALE) Sia γ *un arco di curva*, regolare a tratti, parametrizzata da $\mathbf{r}(t) = i x(t) + j y(t) + k z(t)$, $t \in [a, b]$. Definiamo *integrale di linea (o lavoro)* di $\mathbf{F}$ *lungo* γ l'integrale:

$$\begin{aligned} &\int_\gamma \mathbf{F} \cdot d\mathbf{r} = \int_a^b \mathbf{F}(\mathbf{r}(t)) \cdot \mathbf{r}'(t)\,dt = \\ &= \int_a^b \left[F_1(x(t), y(t), z(t))x'(t) + F_2(x(t), y(t), z(t))y'(t) + F_3(x(t), y(t), z(t))z'(t)\right] dt \end{aligned}$$

La quantità $\mathbf{F} \cdot d\mathbf{r}$ si può riscrivere anche nella forma $\mathbf{F} \cdot \mathbf{T}\,dt$, dove $\mathbf{T}$ è il versore tangente alla curva. ▩

L'integrale $\int_\gamma \mathbf{F} \cdot d\mathbf{r}$ rappresenta pertanto il lavoro totale compiuto da $\mathbf{F}$ per spostare il suo punto di applicazione da $\mathbf{r}(a)$ a $\mathbf{r}(b)$ lungo γ. Altre interpretazioni altrettanto significative sono comunque possibili, a seconda della natura del campo $\mathbf{F}$. Se la curva γ è semplice e chiusa, si usa il simbolo $\oint_\gamma \mathbf{F} \cdot d\mathbf{r}$ e si chiama *circuitazione del campo* $\mathbf{F}$.

Esempio

Consideriamo il campo di velocità piano:

$$\mathbf{v} = -y\mathbf{i} + x\mathbf{j}.$$

Calcoliamo:

a) la circuitazione di $\mathbf{v}$ lungo la circonferenza di raggio 1 e centro l'origine, percorsa in verso antiorario;

b) l'integrale di linea lungo l'arco di spirale $\rho = \theta$, per $\theta \in [0, 2\pi]$.

a) Parametrizziamo la circonferenza con

$$\mathbf{r}(t) = (\cos t, \sin t),\ t \in [0, 2\pi].$$

Pertanto

$$\mathbf{r}'(t) = (-\sin t, \cos t)$$

e la circuitazione è

$$\oint_\gamma \mathbf{v} \cdot d\mathbf{r} = \int_0^{2\pi} \left[-y(t)x'(t) + x(t)y'(t)\right] dt =$$

$$\int_0^{2\pi} \left[-\sin t(-\sin t) + \cos t \cdot \cos t\right] dt = \int_0^{2\pi} dt = 2\pi.$$

© 978-88-08-06485-1

b) Parametrizziamo l'arco di spirale con

$$\mathrm{r}(\theta) = (\theta\cos\theta, \theta\sin\theta).$$

Perciò

$$\mathrm{r}'(\theta) = (\cos\theta - \theta\sin\theta, \sin\theta + \theta\cos\theta)$$

e

$$\int_\gamma \mathrm{v}\cdot d\mathrm{r} = \int_0^{2\pi} [-\theta\sin\theta(\cos\theta - \theta\sin\theta) + \theta\cos\theta(\sin\theta + \theta\cos\theta)]\, d\theta$$
$$= \int_0^{2\pi} \theta^2 d\theta = \frac{8\pi^3}{3}.$$

Essendo riconducibile a un integrale di Riemann, l'integrale di linea di un campo vettoriale possiede le proprietà di linearità rispetto all'integranda e di additività rispetto all'unione di curve, perfettamente analoghe a quelle dell'integrale di linea introdotto nel capitolo 2, paragrafo 5. Tuttavia, tra i due integrali sussiste un'importante differenza e per questo motivo vengono denominati, rispettivamente, *integrale di linea di prima e di seconda specie*. Confrontiamo i due integrali:
– l'integrale di linea di *prima* specie di una funzione continua f, lungo una curva γ parametrizzata da $\mathrm{r}(t) = (x(t), y(t), z(t))$, $t \in [a, b]$, è (cap 2, par 5):

$$\int_a^b f(\mathrm{r}(t))\,|\mathrm{r}'(t)|\, dt$$

– l'integrale di linea di *seconda* specie, lungo γ, del campo $\mathbf{F}$, definito sopra, è:

$$\int_a^b \mathbf{F}(\mathrm{r}(t))\cdot\mathrm{r}'(t)\, dt$$

Si osservi il diverso modo in cui il vettore tangente alla curva è coinvolto nell'integrale. Questa differenza ha una importante conseguenza:
– l'integrale di prima specie è invariante per cambiamenti di parametrizzazione della curva, anche quando la nuova parametrizzazione ne cambia l'orientazione: il risultato cioè non cambia cambiando il verso di percorrenza della curva;
– l'integrale di seconda specie, invece, cambia di segno se si cambia l'orientazione sulla curva, mentre continua a essere invariante per cambiamenti di parametro che non alterino l'orientazione: *il segno dell'integrale, dunque, dipende dal verso di percorrenza della curva.*[3]

È corretto quindi parlare di lavoro di un campo lungo una curva orientata (per esempio "lungo una circonferenza percorsa in senso antiorario"). Si ricordi che la parametrizzazione della curva induce un'orientazione: per cambiarla, occorre cambiare parametrizzazione. Per esempio, la circonferenza unitaria percorsa in senso orario ha la parametrizzazione:

$$\mathrm{r}(t) = (\cos t, -\sin t),\ t \in [0, 2\pi].$$

Naturalmente un campo vettoriale potrebbe avere qualche componente nulla, o qualche componente che è sempre ortogonale alla curva su cui si integra, nel qual caso l'integrale $\int_\gamma \mathbf{F}\cdot d\mathbf{r}$ si riduce ad avere solo uno o due dei tre addendi $\int_\gamma F_1 dx + F_2 dy + F_3 dz$. Più in generale, per un qualsiasi campo scalare $F : \mathbb{R}^3 \to \mathbb{R}$ definito almeno nei punti della curva regolare γ, possiamo dare direttamente la definizione di

$$\int_\gamma F dx = \int_a^b F(\mathrm{r}(t))\, x'(t)\, dt,$$

e analogamente per $\int_\gamma F dy$ e $\int_\gamma F dz$.

[3] Quando nell'integrale $\int_a^b f(x)\, dx$ introduciamo la convenzione

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

stiamo in realtà considerando l'integrale di Riemann come un integrale di linea di seconda specie, lungo l'asse x orientato.

1.4 Campi conservativi e potenziali

L'esempio 1.6 mostra che, in generale, la circuitazione di un campo vettoriale è non nulla. Come vedremo, ciò equivale ad affermare che l'integrale di linea lungo un arco di curva di estremi p_0 e p_1 *non dipende solo dagli estremi* bensì anche dal cammino percorso per congiungerli. Vi sono tuttavia speciali campi vettoriali, detti *conservativi*, per i quali la circuitazione è nulla. Precisamente, abbiamo la:

DEFINIZIONE 6.5 Un campo vettoriale $\mathbf{F} : \Omega \subseteq \mathbb{R}^3 \to \mathbb{R}^3$ si dice *conservativo* in Ω se $\mathbf{F} \in C^1(\Omega)$ ed esiste una funzione $U : \Omega \to \mathbb{R}$, detta *potenziale di* $\mathbf{F}$, tale che $U \in C^2(\Omega)$ e $\mathbf{F} = \nabla U$ in Ω, cioè

$$F_1 = \frac{\partial U}{\partial x},\ F_2 = \frac{\partial U}{\partial y},\ F_3 = \frac{\partial U}{\partial z}.$$

Lo scalare $E_{pot} = -U$ rappresenta *l'energia potenziale* associata al campo. Coerentemente col suo significato fisico, il potenziale U, se esiste, è univocamente definito nell'aperto Ω, a meno di una costante additiva, purché Ω sia connesso.

L'aggettivo *conservativo* deriva dal fatto che, se una particella è in moto lungo una curva di equazione $\mathrm{r} = \mathrm{r}(t)$ sotto l'azione di una forza data da $\mathbf{F} = \nabla U$, allora *l'energia meccanica totale E_{tot} della particella si conserva durante il moto*. Infatti, si ha:

$$E_{tot} = E_{cin} + E_{pot} = \frac{1}{2}mv^2 - U$$

dove m è la massa della particella e $v = v(t) = |\mathrm{r}'(t)|$ la sua velocità scalare. Scriviamo $E_{cin} = \frac{1}{2}m\mathrm{r}'\cdot\mathrm{r}'$ e ricordiamo che, per la seconda legge della dinamica, $\mathbf{F} = m\mathrm{a} = m\mathrm{r}''$; differenziando E_{tot}, si ha:

$$dE_{tot} = m\mathrm{r}'\cdot d\mathrm{r}' - dU = m\mathrm{r}'\cdot\mathrm{r}''dt - \nabla U\cdot\mathrm{r}'dt = \mathbf{F}\cdot\mathrm{r}'dt - \mathbf{F}\cdot\mathrm{r}'dt = 0$$

e quindi E_{tot} non varia.

Molti campi di interesse fisico sono conservativi e, quando questo si verifica, il campo ha notevoli proprietà che ne facilitano lo studio. Incontreremo un po' alla volta queste proprietà.

© 978-88-08-06485-1

Esempio

1.7. Il campo gravitazionale di una massa puntiforme (vedi esempio 1.1) è conservativo nello spazio privato dell'origine. Infatti,

$$-k\frac{(x,y,z)}{\rho^3} = \nabla\left(\frac{k}{\rho}\right) \qquad (\rho = \sqrt{x^2+y^2+z^2})$$

come si verifica subito. Il potenziale gravitazionale è dunque:

$$U(x,y,z) = \frac{k}{\rho}.$$

Si noti che le linee di campo (rette uscenti dall'origine) sono normali in ogni punto alle *superfici equipotenziali*, cioè alle superfici $U = cost.$, che in questo caso sono sfere centrate nell'origine.

Vediamo come si calcola il lavoro per un campo conservativo. In tutto il seguito assumiamo sempre che Ω sia un *aperto connesso.*

LEMMA 6.1. *Sia $\mathbf{F} = \nabla U$ un campo conservativo in Ω e sia γ una curva regolare a tratti, orientata e contenuta in Ω, parametrizzata da $\mathrm{r} = \mathrm{r}(t)$, $t \in [a,b]$. Siano $\mathrm{p} = \mathrm{r}(a)$, $\mathrm{q} = \mathrm{r}(b)$ gli estremi della curva. Il lavoro di $\mathbf{F}$ lungo γ è dato da:*

$$\int_\gamma \mathbf{F}\cdot d\mathbf{r} = U(\mathrm{q}) - U(\mathrm{p}).$$

Il lavoro, quindi, dipende solo dagli estremi del cammino e non dalla forma del medesimo. In particolare, la circuitazione di un campo conservativo è nulla.

DIMOSTRAZIONE. Se $\mathrm{r}(t) = (x(t), y(t), z(t))$, $t \in [a,b]$, è regolare, si ha:

$$\int_\gamma \mathbf{F}\cdot d\mathbf{r} = \int_a^b \nabla U(\mathrm{r}(t))\cdot \mathrm{r}'(t)\ dt = \tag{1.5}$$
$$= \int_a^b \frac{d}{dt}U(\mathrm{r}(t))\ dt = U(\mathrm{r}(b)) - U(\mathrm{r}(a)) = U(\mathrm{q}) - U(\mathrm{p}).$$

Se r è C^1 a tratti, allora si può sudddividere $[a,b]$ nell'unione di N intervalli $[t_j, t_{j+1}]$ con $j = 1,\dots N$, in modo che $t_1 = a$ e $t_{N+1} = b$ e che $\mathrm{r} \in C^1[t_j, t_{j+1}]$. Applicando la (1.5) a ognuno degli intervalli $[t_j, t_{j+1}]$ e sommando i singoli contributi, si trova:

$$\int_\gamma \mathbf{F}\cdot d\mathbf{r} = \sum_{j=1}^N \int_{t_j}^{t_{j+1}} \mathbf{F}\cdot d\mathbf{r} = \sum_{j=1}^N [U(\mathrm{r}(t_{j+1})) - U(\mathrm{r}(t_j))] = U(\mathrm{q}) - U(\mathrm{p})$$

che conclude la dimostrazione. ◊

In realtà, la condizione che la circuitazione sia nulla lungo *ogni curva semplice e chiusa contenuta in* Ω caratterizza i campi conservativi. Vale infatti il seguente importante teorema.

© 978-88-08-06485-1

TEOREMA 6.2. *Sia $\mathbf{F} \in C^1(\Omega)$. Le seguenti tre affermazioni sono equivalenti:*

(a) Per ogni coppia di curve regolari a tratti γ_1, γ_2 contenute in Ω e aventi stesso punto iniziale e stesso punto finale (e per il resto disgiunte),

$$\int_{\gamma_1} \mathbf{F}\cdot d\mathbf{r} = \int_{\gamma_2} \mathbf{F}\cdot d\mathbf{r}.$$

(b) Per ogni curva chiusa γ, semplice regolare a tratti e contenuta in Ω,

$$\oint_\gamma \mathbf{F}\cdot d\mathbf{r} = 0.$$

(c) $\mathbf{F}$ è conservativo.

Confrontando l'enunciato di questo teorema col Lemma precedente si nota che: l'annullarsi del lavoro lungo qualunque curva *semplice*, chiusa e regolare a tratti implica la conservatività del campo, che a sua volta implica l'annullarsi del lavoro lungo qualunque curva chiusa e regolare a tratti (anche non semplice).

DIMOSTRAZIONE. L'implicazione $(c) \Rightarrow (b)$ è contenuta nel lemma 6.1. Facciamo vedere che $(a) \Rightarrow (c)$ e che $(b) \Rightarrow (a)$.

$(a) \Rightarrow (c)$. Siano $\mathrm{p}_0 = (x_0, y_0, z_0)$ e $\mathrm{p} = (x, y, z)$ due punti in Ω e $\gamma \subset \Omega$ una curva orientata di estremi p_0 e p. Pensando p_0 fissato, per l'ipotesi (a), l'integrale curvilineo lungo γ dipende solo da p e quindi

$$\int_\gamma \mathbf{F}\cdot d\mathbf{r} = U(x,y,z).$$

Vogliamo dimostrare che

$$\frac{\partial U}{\partial x} = F_1 \text{ in } \Omega.$$

Congiungiamo p al punto $(x+\Delta x, y, z)$ con un segmento rettilineo Γ di equazioni parametriche

$$x(t) = x + t\Delta x,\ y(t) = y,\ z(t) = z, \qquad t \in [0,1],$$

come mostrato in figura 6.1.

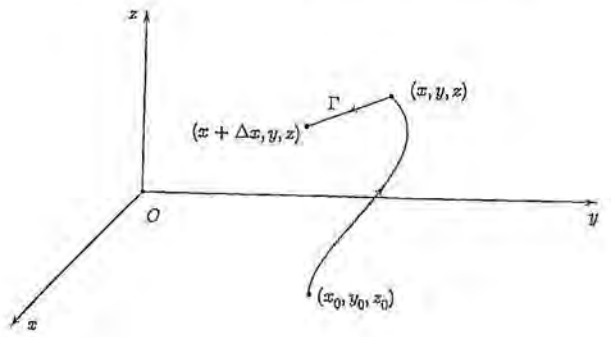


Figura 6.1.

© 978-88-08-06485-1

Essendo $x'(t) = \Delta x$, $y'(t) = z'(t) = 0$, usando le proprietà di additività dell'integrale di linea, possiamo scrivere:

$$U(x+\Delta x, y, z) = U(x,y,z) + \int_\Gamma \mathbf{F}\cdot d\mathbf{r} = U(x,y,z) + \int_0^1 F_1(x+t\Delta x, y, z)\,\Delta x\,dt.$$

Ne segue che

$$\frac{U(x+\Delta x, y, z) - U(x,y,z)}{\Delta x} = \int_0^1 F_1(x+t\Delta x, y, z)\,dt.$$

Essendo F_1 continua, passando al limite sotto il segno di integrale per $\Delta x \to 0$, si ottiene $\frac{\partial U}{\partial x} = F_1$.

In modo analogo si dimostra che:

$$\frac{\partial U}{\partial y} = F_2 \text{ e } \frac{\partial U}{\partial z} = F_3 \text{ in } \Omega,$$

per cui U è una funzione potenziale per il campo $\mathbf{F}$ che risulta pertanto conservativo.

$(b) \Rightarrow (a)$ Siano γ_1, γ_2 curve regolari a tratti contenute in Ω e aventi stesso punto iniziale e stesso punto finale, per il resto disgiunte. Cambiamo l'orientazione di γ_2 e chiamiamo γ_2^* la nuova curva. Sappiamo che

$$\int_{\gamma_2^*} \mathbf{F}\cdot d\mathbf{r} = -\int_{\gamma_2} \mathbf{F}\cdot d\mathbf{r}.$$

Definiamo poi la curva semplice e chiusa $\gamma = \gamma_1 \cup \gamma_2^*$, che risulta regolare a tratti. Usando (b) e l'additività dell'integrale di linea, si ha:

$$0 = \oint_\gamma \mathbf{F}\cdot d\mathbf{r} = \int_{\gamma_1} \mathbf{F}\cdot d\mathbf{r} + \int_{\gamma_2^*} \mathbf{F}\cdot d\mathbf{r} = \int_{\gamma_1} \mathbf{F}\cdot d\mathbf{r} - \int_{\gamma_2} \mathbf{F}\cdot d\mathbf{r}$$

da cui (a).

OSSERVAZIONE Sottolineiamo che la dimostrazione di $(a) \Rightarrow (c)$ indica un modo per costruire una funzione potenziale di un campo $\mathbf{F}$, qualora questo risulti conservativo. La ricetta generale è la seguente: si fissa un punto p_0 in Ω e lo si congiunge con una curva γ qualsiasi, purché contenuta in Ω, con un punto generico $\mathrm{p} = (x,y,z)$. Un potenziale, in particolare quello che si annulla in p_0, è dato da:

$$U(x,y,z) = \int_\gamma \mathbf{F}\cdot d\mathbf{r}.$$

Un metodo più diretto per costruire un potenziale sarà indicato più avanti, in alcuni esempi.

Il teorema implica che, se si riesce a trovare una curva semplice e chiusa $\gamma \subset \Omega$ tale che

$$\oint_\gamma \mathbf{F}\cdot d\mathbf{r} \neq 0,$$

allora $\mathbf{F}$ *non è conservativo*. Il seguente esempio in 2 dimensioni, che riprenderemo anche in seguito, è particolarmente significativo.

© 978-88-08-06485-1

Esempio

Siano $\Omega = \mathbb{R}^2 \setminus \{0,0\}$ e

$$\mathbf{F}(x,y) = -\frac{y}{x^2+y^2}\mathrm{i} + \frac{x}{x^2+y^2}\mathrm{j}.$$

Calcoliamo la circuitazione di $\mathbf{F}$ lungo la circonferenza unitaria, centrata nell'origine, di equazioni parametriche $\mathrm{r}(t) = (\cos t, \sin t)$, $t \in [0, 2\pi]$.

Si ha, essendo $x(t)^2 + y(t)^2 = 1$, $x'(t) = -\sin t$, $y'(t) = \cos t$:

$$\oint_\gamma \mathbf{F}\cdot d\mathbf{r} = \int_0^{2\pi} [-\sin t(-\sin t) + \cos t(\cos t)]\,dt = 2\pi.$$

Concludiamo che $\mathbf{F}$ *non è conservativo in* Ω.

1.5 Campi irrotazionali. Insiemi semplicemente connessi

Un campo avente rotore nullo in tutto il suo dominio si dice *irrotazionale*. Tra campi irrotazionali e campi conservativi esiste uno stretto legame, che tuttavia va illustrato con precisione.

Innanzitutto, se $\mathbf{F}$ è conservativo e U è un suo potenziale, si ha $\mathbf{F} = \nabla U$ e allora [cfr. (1.1)] $\nabla \times \mathbf{F} = 0$. Pertanto:

PROPOSIZIONE 6.2 *Se* $\mathbf{F}$ *è conservativo in* Ω, *allora è irrotazionale in* Ω.

Sorge spontanea la domanda: *è vero il viceversa?* Cioè: *se* $\nabla \times \mathbf{F} = 0$ *in* Ω, *allora* $\mathbf{F}$ *è conservativo in* Ω?

In generale la risposta è negativa. Riprendiamo infatti l'esempio 1.8. Abbiamo constatato che

$$\mathbf{F}(x,y) = -\frac{y}{x^2+y^2}\mathrm{i} + \frac{x}{x^2+y^2}\mathrm{j}$$

non è conservativo in $\mathbb{R}^2 \setminus \{0,0\}$. Tuttavia, si ha:

$$F_1(x,y) = -\frac{y}{x^2+y^2} \text{ e } F_2(x,y) = \frac{x}{x^2+y^2}$$

ed essendo un campo piano, la condizione di irrotazionalità equivale a:

$$\frac{\partial F_1}{\partial y} = \frac{\partial F_2}{\partial x} \quad \text{in } \mathbb{R}^2 \setminus \{0,0\}.$$

Abbiamo:

$$\frac{\partial}{\partial y}\Big(-\frac{y}{x^2+y^2}\Big) = \frac{y^2-x^2}{(x^2+y^2)^2} = \frac{\partial}{\partial x}\Big(\frac{x}{x^2+y^2}\Big)$$

e perciò $\nabla \times \mathbf{F} = \mathbf{0}$ in $\mathbb{R}^2 \setminus \{0,0\}$.

Pertanto, la condizione di irrotazionalità *non* è sufficiente a garantire che $\mathbf{F}$ sia conservativo. La ragione di questo fatto è legata alle *proprietà topologiche* dell'aperto Ω in cui il campo è irrotazionale.

La nozione chiave è quella di *aperto semplicemente connesso*, di cui daremo solo una definizione intuitiva, ma sufficiente ai nostri scopi. Ricordiamo che un aperto Ω si

© 978-88-08-06485-1

dice *connesso* (vedi capitolo 3, paragrafo 3.1) se, presi due punti di Ω, esiste un arco di curva continuo che congiunge questi due punti ed è interamente contenuto in Ω.

DEFINIZIONE 6.6 L'aperto Ω si dice *semplicemente connesso* se è connesso e inoltre soddisfa la condizione seguente: *ogni curva semplice, chiusa e interamente contenuta in Ω, può essere ridotta mediante una deformazione continua a un unico punto, senza mai uscire da Ω.*

Vediamo alcuni esempi tipici.

Esempio

Nel piano:

- sono semplicemente connessi: cerchi, ellissi, poligoni, semipiani e tutte le figure piane che si possono ottenere da queste deformandole con continuità; il piano stesso o il piano privato di una semiretta.
- non sono semplicemente connessi: il piano o un cerchio (ellisse, poligono) privato di un punto; una corona circolare; ogni insieme che presenta "un buco".

Nello spazio:

- sono semplicemente connessi: sfere, ellissoidi, poliedri convessi e tutti i solidi che si possono ottenere da questi deformandoli con continuità; una corona sferica, un semispazio, lo spazio privato di un numero finito di punti.
- non sono semplicemente connessi: il toro, la sfera privata di un diametro, lo spazio privato di una retta.

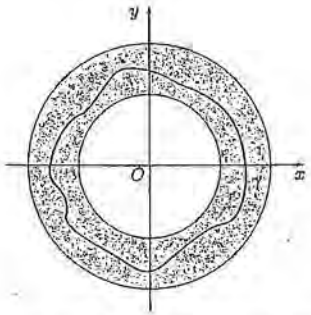


Figura 6.2. La corona circolare non è semplicemente connessa: la curva γ non si può contrarre a un punto senza uscire dalla corona.

Chiarita questa nozione, possiamo enunciare un risultato profondo, che non dimostreremo:

TEOREMA 6.3 *Sia $\mathbf{F} \in C^1(\Omega)$, irrotazionale in Ω. Se Ω è semplicemente connesso, allora $\mathbf{F}$ è* conservativo *in Ω.*

© 978-88-08-06485-1

Naturalmente la condizione topologica su Ω è solo sufficiente. Vi sono campi conservativi anche in insiemi non semplicemente connessi (si veda l'esempio 1.11).

OSSERVAZIONE Sia $\nabla \times \mathbf{F} = \mathbf{0}$ in un aperto Ω qualunque (anche non connesso!). Ogni punto di Ω ha un intorno sferico B_R contenuto in Ω e B_R è semplicemente connesso. Si può sempre applicare il teorema in B_R e dedurne che in questo intorno (dunque localmente) il campo è conservativo. Si dice allora che il campo $\mathbf{F}$ è *localmente conservativo* e che quindi esiste un *potenziale locale*.

Vediamo una serie di esempi che illustrano i concetti presentati finora.

Esempi

Controlliamo se il campo

$$\mathbf{F} = -\mathrm{i}y + \mathrm{j}x$$

è conservativo o almeno localmente conservativo nel piano. Calcoliamone poi la circuitazione lungo l'ellisse γ

$$\begin{cases} x = a\cos t \\ y = b\sin t \end{cases} \quad t \in [0, 2\pi].$$

Calcoliamo $\nabla \times \mathbf{F}$. Poiché:

$$\partial_x F_2 - \partial_y F_1 = 1 - (-1) = 2$$

il campo non è irrotazionale in alcuna regione del piano. Pertanto non è conservativo e non c'è ragione di aspettarsi che il lavoro lungo un cammino chiuso sia nullo (naturalmente, potrebbe anche esserlo per qualche cammino speciale): occorre dunque calcolarlo in base alla definizione. Essendo

$$\mathbf{r}'(t) = (-a\sin t, b\cos t),$$

si trova:

$$\oint_\gamma \mathbf{F}\cdot d\mathbf{r} = \int_0^{2\pi} [-b\sin t\cdot(-a\sin t) + a\cos t\cdot b\cos t]\, dt = 2\pi ab.$$

In $\Omega = \mathbb{R}^2 \setminus \{0,0\}$ consideriamo il campo vettoriale:

$$\mathbf{F} = \frac{x}{x^2+y^2}\mathrm{i} + \frac{y}{x^2+y^2}\mathrm{j}.$$

Controlliamo prima se il campo è conservativo in Ω e poi calcoliamo:

a) la circuitazione lungo l'ellisse di equazione $3x^2 - xy + 10y^2 = 1$, percorsa in senso antiorario;

b) il lavoro lungo l'arco di parabola γ di equazione $y = 1 + x^2, x \in [0,2]$;

c) il lavoro lungo un qualsiasi arco di una qualsiasi circonferenza centrata in $(0,0)$.

Calcoliamo innanzitutto $\nabla \times \mathbf{F}$. Abbiamo:

$$\partial_x F_2 - \partial_y F_1 = \frac{-2xy}{(x^2+y^2)^2} + \frac{2xy}{x^2+y^2} \equiv 0$$

per cui $\nabla \times \mathbf{F} = \mathbf{0}$ in Ω ed $\mathbf{F}$ è localmente conservativo. Per calcolare un potenziale $U = U(x,y)$ osserviamo che deve essere:

$$U_x = \frac{x}{x^2+y^2}, \text{ perciò } U = \int \frac{x}{x^2+y^2}dx = \frac{1}{2}\log(x^2+y^2) + c_1(y)$$

© 978-88-08-06485-1

poiché la "costante d'integrazione" è una quantità costante rispetto a x (la variabile d'integrazione) e quindi è funzione della sola y. Analogamente:

$$U_y = \frac{y}{x^2+y^2}, \text{ perciò } U = \int \frac{y}{x^2+y^2}dy = \frac{1}{2}\log(x^2+y^2) + c_2(x).$$

Scegliendo $c_1(y) = c_2(x) = 0$ si trova il potenziale

$$U(x,y) = \frac{1}{2}\log(x^2+y^2)$$

che soddisfa entrambe le richieste in tutto Ω.

Perciò $\mathbf{F}$ è effettivamente conservativo in tale insieme, anche se Ω non è semplicemente connesso. Di conseguenza:

a) la circuitazione lungo qualunque curva chiusa, in particolare lungo l'ellisse indicata, è nulla;

b) poiché l'arco di parabola ha estremi $\mathrm{p} = (0,1)$; $\mathrm{q} = (2,5)$, il lavoro è dato da:

$$L = \int_\gamma \mathbf{F}\cdot d\mathbf{r} = U(\mathrm{q}) - U(\mathrm{p}) = \frac{1}{2}\log 29;$$

c) il lavoro lungo qualsiasi arco della circonferenza di centro $(0,0)$ e raggio R è nullo, perché gli estremi di questo arco hanno lo stesso potenziale $\frac{1}{2}\log(R^2)$. Questo è un fatto generale: se un campo è conservativo, è nullo non solo il lavoro lungo un cammino chiuso, ma anche il lavoro lungo un cammino che ha i suoi estremi sulla stessa linea equipotenziale.

ESEMPIO Calcoliamo il lavoro del campo

$$\mathbf{F} = -\frac{y}{x^2+y^2}\mathrm{i} + \frac{x}{x^2+y^2}\mathrm{j}$$

a) sulla circonferenza unitaria di centro $(1,2)$, in senso antiorario;

b) sulla curva raffigurata nella figura seguente.

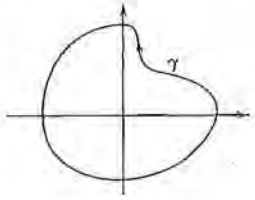


Figura 6.3.

a) Abbiamo visto (esempio 1.8) che questo campo è irrotazionale in $\Omega = \mathbb{R}^2 \setminus \{0,0\}$, non è conservativo in tutto Ω, ma lo è in qualsiasi aperto semplicemente connesso $A \subset \Omega$.

Per esempio, $\mathbf{F}$ è conservativo nel piano privato di una qualsiasi semiretta uscente dall'origine e poiché la circonferenza unitaria di centro $(1,2)$ è contenuta nel piano privato della semiretta $\{y = 0, x > 0\}$, il lavoro lungo questa curva è nullo.

[4] Un potenziale nei semipiani $x > 0$ e $x < 0$ è dato da

$$U(x,y) = \operatorname{arctg}\frac{y}{x},$$

come si controlla facilmente.

© 978-88-08-06485-1

b) Infine, affermiamo che la circuitazione lungo la curva γ in figura è uguale alla circuitazione lungo la circonferenza unitaria γ_0 centrata nell'origine e perciò uguale a 2π, come calcolato nell'esempio 1.8.

Per spiegare questa affermazione si osservi la figura 6.4. La curva Γ disegnata non circonda l'origine e infatti è contenuta nel piano privato del semiasse positivo delle ascisse, che costituisce una regione semplicemente connessa. Il campo è conservativo in questa regione e quindi la circuitazione lungo la curva è nulla. D'altro canto, ci si convince facilmente che

$$0 = \oint_\Gamma \mathbf{F}\cdot d\mathbf{r} = \oint_\gamma \mathbf{F}\cdot d\mathbf{r} - \oint_{\gamma_0} \mathbf{F}\cdot d\mathbf{r}$$

poiché, per via della diversa orientazione, i due segmenti rettilinei danno invece contributi uguali e contrari al calcolo dell'integrale e quindi sono ininfluenti. Ne segue che il lavoro lungo γ è uguale al lavoro lungo γ_0.

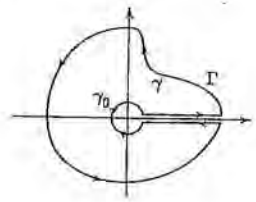


Figura 6.4.

OSSERVAZIONE In generale, se $\mathbf{F}$ è un campo irrotazionale in un aperto Ω qualsiasi (anche non semplicemente connesso), diciamo che due cammini contenuti in Ω sono *equivalenti* se si possono deformare uno nell'altro con continuità *senza mai uscire da* Ω. In tal caso, con un'argomentazione simile a quella usata nell'ultimo esempio, si vede che il lavoro di $\mathbf{F}$ lungo i due cammini è lo stesso.

ESEMPIO

ESEMPIO Nell'aperto $\mathbb{R}^3 \setminus \{z = 0\}$ consideriamo il campo vettoriale:

$$\mathbf{F} = (xy - \sin z)\mathrm{i} + (\frac{1}{2}x^2 - \frac{e^y}{z})\mathrm{j} + (\frac{e^y}{z^2} - x\cos z)\mathrm{k}.$$

Controlliamo se è irrotazionale. Abbiamo:

$$\begin{cases} \partial_y F_1 - \partial_x F_2 = x - x \equiv 0 \\ \partial_z F_1 - \partial_x F_3 = -\cos z - (-\cos z) \equiv 0 \\ \partial_z F_2 - \partial_y F_3 = \frac{e^y}{z^2} - \frac{e^y}{z^2} \equiv 0 \end{cases}$$

perciò $\nabla \times \mathbf{F} = 0$ e quindi $\mathbf{F}$ è localmente conservativo.

Cerchiamo un potenziale locale. Questa volta procediamo imponendo una dopo l'altra le condizioni[5]:

$$U_x = xy - \sin z,\ U_y = \frac{1}{2}x^2 - \frac{e^y}{z},\ U_z = \frac{e^y}{z^2} - x\cos z. \tag{1.6}$$

[5] Si può controllare che questo metodo coincide con quello indicato nell'osservazione alla fine del paragrafo 1.4 quando si scelga come curva γ di integrazione una spezzata di tre segmenti paralleli agli assi coordinati.

© 978-88-08-06485-1

Dalla prima ricaviamo

$$U(x,y,z) = \int (xy - \sin z)\,dx = \frac{1}{2}x^2 y - x\sin z + a(y,z) \tag{1.7}$$

con a da determinarsi. Essendo $U_y = \frac{1}{2}x^2 + a_y(y,z)$, dalla seconda delle (1.6) abbiamo:

$$a_y(y,z) = -\frac{e^y}{z}$$

da cui

$$a(y,z) = -\frac{e^y}{z} + b(z).$$

Finora abbiamo dunque:

$$U(x,y,z) = \frac{1}{2}x^2 y - x\sin z - \frac{e^y}{z} + b(z),$$

con b da determinarsi.

Essendo $U_z = -x\cos z + \frac{e^y}{z^2} + b'(z)$, imponendo la terza delle (1.6) troviamo:

$$b'(z) = 0$$

da cui $b(z) = k$, costante. In conclusione, i potenziali locali hanno la forma

$$U(x,y,z) = \frac{1}{2}x^2 y - x\sin z - \frac{e^y}{z} + k.$$

Il campo è effettivamente conservativo, separatamente, in ciascuno dei due semispazi $z > 0$, $z < 0$, poiché questi sono semplicemente connessi.

1.6 Campi solenoidali e potenziale vettore

Abbiamo visto la relazione esistente tra campi *irrotazionali* e campi *conservativi*: se $\nabla\times\mathbf{F} = \mathbf{0}$ in un aperto Ω, $\mathbf{F}$ è localmente conservativo e cioè si può scrivere localmente come $\mathbf{F} = \nabla U$; lo scalare U è un potenziale di $\mathbf{F}$ ed è definito univocamente a meno di costanti additive. Se Ω è semplicemente connesso, allora $\mathbf{F}$ è anche conservativo globalmente, cioè U è un potenziale in tutto Ω.

L'importanza dei campi conservativi risiede nell'invarianza dei loro integrali di linea rispetto a deformazioni continue del cammino di integrazione, che costituisce una versione del principio di conservazione dell'energia meccanica.

Questioni analoghe si presentano esplorando la relazione che sussiste tra i campi *solenoidali*, l'equazione $\nabla\cdot\mathbf{F} = 0$ e i campi che si possono esprimere come *rotore di un altro campo vettoriale*, cioè per i quali esiste $\mathbf{V}$ tale che $\nabla\times\mathbf{V} = \mathbf{F}$. Il campo $\mathbf{V}$ prende il nome di *potenziale vettore di* $\mathbf{F}$.

Osserviamo subito che:

PROPOSIZIONE 6.3 *Sia* $\mathbf{V}\in C^2(\Omega)$ *un* potenziale vettore di $\mathbf{F}$ *in* Ω. *Allora* $\nabla\cdot\mathbf{F} = 0$ *in* Ω.

DIMOSTRAZIONE. Essendo $\mathbf{F} = \nabla\times\mathbf{V}$ si ha [cfr. (1.2)] $\nabla\cdot\mathbf{F} = \nabla\cdot(\nabla\times\mathbf{V}) = 0$. ♣

In altri termini, *solo i campi solenoidali possono ammettere potenziale vettore.* È poi importante osservare che, se $\mathbf{V}$ è un *potenziale vettore di* $\mathbf{F}$ in Ω e ψ è una funzione scalare regolare in Ω, allora $\mathbf{V} + \nabla\psi$ è ancora un potenziale vettore di $\mathbf{F}$ in Ω, essendo $\nabla\times(\nabla\psi) = \mathbf{0}$. Dunque un potenziale vettore, se esiste, non è unico. Questo grado di libertà nella determinazione di $\mathbf{V}$ risulta molto utile nelle applicazioni, specialmente all'elettromagnetismo.

© 978-88-08-06485-1

Si presenta spontanea la domanda: *è vero che se* $\nabla\cdot\mathbf{F} = 0$ *in* Ω, *allora* $\mathbf{F}$ *ha un potenziale vettore in* Ω?

La risposta è *positiva a livello locale*, nel senso che *esiste un potenziale vettore in ogni intorno sferico* $B\subset\Omega$ *di ogni punto di* Ω, ma è ancora negativa a livello globale. Precisamente, abbiamo il seguente teorema.

TEOREMA 6.4 *Sia* $\mathbf{F}\in C^1(\Omega)$ *tale che* $\nabla\cdot\mathbf{F} = 0$ *in* Ω. *Allora esiste localmente un potenziale vettore* $\mathbf{V}$ *di* $\mathbf{F}$. *In particolare, si può scegliere* $\mathbf{V}$ *con una componente nulla; per esempio, se* $B\subset\Omega$ *è un intorno sferico del punto* (x_0,y_0,z_0), *si può scegliere*

$$\mathbf{V}(x,y,z) = V_2(x,y,z)\,\mathrm{j} + V_3(x,y,z)\,\mathrm{k}$$

dove

$$V_2(x,y,z) = \int_{x_0}^{x} F_3(s,y,z)\,ds \tag{1.8}$$

e

$$V_3(x,y,z) = -\int_{x_0}^{x} F_2(s,y,z)\,ds + \int_{y_0}^{y} F_1(x_0,s,z)\,ds. \tag{1.9}$$

DIMOSTRAZIONE. Ci limitiamo a controllare che (1.8) e (1.9) definiscono effettivamente un potenziale vettore di $\mathbf{F}$. Abbiamo

$$\nabla\times\mathbf{V} = \left(\frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z}\right)\mathrm{i} - \frac{\partial V_3}{\partial x}\mathrm{j} + \frac{\partial V_2}{\partial x}\mathrm{k}$$

e quindi, usando (1.8) e (1.9), troviamo:

$$\frac{\partial V_3}{\partial y} - \frac{\partial V_2}{\partial z} = -\int_{x_0}^{x}\frac{\partial F_2}{\partial y}(s,y,z)\,ds + F_1(x_0,y,z) - \int_{x_0}^{x}\frac{\partial F_3}{\partial z}(s,y,z)\,ds =$$

$$= F_1(x_0,y,z) - \int_{x_0}^{x}\left\{\frac{\partial F_2}{\partial y}(s,y,z) + \frac{\partial F_3}{\partial z}(s,y,z)\right\}ds =$$

essendo, per l'ipotesi $\nabla\cdot\mathbf{F} = 0$, $\dfrac{\partial F_2}{\partial y} + \dfrac{\partial F_3}{\partial z} = -\dfrac{\partial F_1}{\partial x}$

$$= F_1(x_0,y,z) + \int_{x_0}^{x}\frac{\partial F_1}{\partial x}(s,y,z)\,ds = F_1(x,y,z).$$

Inoltre,

$$-\frac{\partial V_3}{\partial x} = F_2(x,y,z) \quad\text{e}\quad \frac{\partial V_2}{\partial x} = F_3(x,y,z).$$

Ne segue che $\nabla\times\mathbf{V} = \mathbf{F}$. ♣

© 978-88-08-06485-1

Esempio

1.14 Sia $\mathbf{F} = xz\mathbf{i} - yz\mathbf{j} - 2\mathbf{k}$. Si può controllare che $\nabla \cdot \mathbf{F} = 0$, perciò $\mathbf{F}$ possiede localmente un potenziale vettore. Per calcolarlo, scegliamo $(x_0, y_0, z_0) = (0, 0, 0)$. Le (1.8) e (1.9) danno:

$$V_1 = 0,\ V_2(x, y, z) = -\int_0^x 2ds = -2x,\ V_3(x, y, z) = \int_0^x yz ds = xyz$$

(notando che $F_1(0, y, z) = 0$) e quindi $\mathbf{V}(x, y, z) = -2x\mathbf{j} - xyz\mathbf{k}$.

Vediamo due esempi che illustrano qualche utilizzo dei concetti precedenti.

Esempi

1.15 *Magnetostatica.* Indichiamo con $\mathbf{j} = \mathbf{j}(\mathbf{x})$, $\mathbf{x} = (x_1, x_2, x_3)$, il campo densità di corrente dovuto a una determinata distribuzione di cariche nello spazio. Il campo magnetico $\mathbf{B}$ da esso generato è dato, in unità standard, dalla formula

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\mathbf{R}^3} \mathbf{j}(\mathbf{y}) \times \frac{\mathbf{x} - \mathbf{y}}{|\mathbf{x} - \mathbf{y}|^3} d\mathbf{x},$$

dove μ_0 è la *permeabilità magnetica* nel vuoto. Abbiamo visto nell'esempio 1.5 (cambiando un poco le notazioni) che, se fissiamo $\mathbf{y}$, il campo

$$\mathbf{F}(\mathbf{x}) = \frac{\mathbf{x} - \mathbf{y}}{|\mathbf{x} - \mathbf{y}|^3}$$

è solenoidale e pertanto anche $\mathbf{B}(\mathbf{x})$ lo è, come si vede calcolando $\nabla \cdot \mathbf{B}$ derivando sotto il segno di integrale rispetto a $\mathbf{x}$. Cerchiamo un potenziale vettore. Scriviamo:

$$\frac{\mathbf{x} - \mathbf{y}}{|\mathbf{x} - \mathbf{y}|^3} = -\nabla\left(\frac{1}{|\mathbf{x} - \mathbf{y}|}\right),$$

dove le derivate sono rispetto a $\mathbf{x}$. Usando ora l'identità $\mathbf{G} \times \nabla f = -\nabla \times (f\mathbf{G}) + f(\nabla \times \mathbf{G})$ (vedi la quarta delle formule (1.4)), possiamo scrivere[6]

$$\mathbf{j}(\mathbf{y}) \times \frac{\mathbf{x} - \mathbf{y}}{|\mathbf{x} - \mathbf{y}|^3} = \nabla \times \frac{\mathbf{j}(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|}$$

e, successivamente,

$$\mathbf{B}(\mathbf{x}) = \frac{\mu_0}{4\pi} \nabla \times \iiint_{\mathbf{R}^3} \frac{\mathbf{j}(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} d\mathbf{x}.$$

Il vettore (che si chiama *potenziale newtoniano di* $\mu_0\mathbf{j}$)

$$\mathbf{A}(\mathbf{x}) = \frac{\mu_0}{4\pi} \iiint_{\mathbf{R}^3} \frac{\mathbf{j}(\mathbf{y})}{|\mathbf{x} - \mathbf{y}|} d\mathbf{x}$$

è dunque un potenziale vettore di $\mathbf{B}$. Si può dimostrare (ma la prova non è elementare) che $\Delta\mathbf{A} = -\mu_0\mathbf{j}$.

1.16 *Equazioni di Maxwell.* Le equazioni di Maxwell, nel vuoto e in assenza di sorgenti, sono:

$$\nabla \cdot \mathbf{B} = 0$$

[6] Ricordare che le derivate sono fatte rispetto a $\mathbf{x}$, perciò $\nabla \times \mathbf{j}(\mathbf{y}) = 0$.

© 978-88-08-06485-1

$$\nabla \cdot \mathbf{E} = 0 \tag{1.10}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{1.11}$$

$$\nabla \times \mathbf{B} = \frac{1}{c_0^2} \frac{\partial \mathbf{E}}{\partial t} \tag{1.12}$$

dove $\mathbf{E}$ indica il campo elettrico e c_0 la velocità della luce nel vuoto. Qui tutti i campi sono variabili nel tempo.

Essendo solenoidale, $\mathbf{B}$ ha un potenziale vettore (locale) $\mathbf{A}$, $\nabla \times \mathbf{A} = \mathbf{B}$, per cui possiamo riscrivere la (1.11) nella forma seguente:

$$\nabla \times \left(\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}\right) = 0$$

che mostra l'irrotazionalità del vettore $\mathbf{E} + \frac{\partial \mathbf{A}}{\partial t}$. Esiste quindi un potenziale scalare φ tale che

$$\mathbf{E} = \nabla\varphi - \frac{\partial \mathbf{A}}{\partial t}.$$

Vediamo dunque che $\mathbf{B}$ ed $\mathbf{E}$ possono essere espressi in termini del potenziale vettore $\mathbf{A}$ e del potenziale scalare φ.

Sostituendo nella (1.10), otteniamo:

$$\Delta\varphi - \frac{\partial}{\partial t} \nabla \cdot \mathbf{A} = 0.$$

Per avere un'equazione nelle sole φ e $\mathbf{A}$, usiamo il grado di libertà nella scelta del potenziale vettore scegliendo $\mathbf{A}' = \mathbf{A} + \nabla g$, $g = g(\mathbf{x}, t)$, in modo che valga la seguente *condizione di Lorentz*:

$$\nabla \cdot \mathbf{A}' - \frac{1}{c_0^2} \frac{\partial \varphi}{\partial t} = 0.$$

Con questa scelta del potenziale vettore, la (1.10) diventa:

$$\Delta\varphi - \frac{1}{c_0^2} \frac{\partial^2 \varphi}{\partial t^2} = 0$$

che è l'*equazione delle onde* in dimensione 3. Infine, la (1.12) diventa:

$$\nabla \times (\nabla \times \mathbf{A}') = \frac{1}{c_0^2} \frac{\partial}{\partial t} \left(\nabla\varphi - \frac{\partial \mathbf{A}'}{\partial t}\right)$$

ossia, per l'identità $\nabla \times (\nabla \times \mathbf{G}) = \nabla(\nabla \cdot \mathbf{G}) - \Delta\mathbf{G}$,

$$-\Delta\mathbf{A}' + \nabla\left(\nabla \cdot \mathbf{A}' - \frac{1}{c_0^2} \frac{\partial \varphi}{\partial t}\right) + \frac{1}{c_0^2} \frac{\partial^2 \mathbf{A}'}{\partial^2 t} = 0$$

per cui anche $\mathbf{A}'$ è soluzione dell'equazione vettoriale delle onde

$$\Delta\mathbf{A}' - \frac{1}{c_0^2} \frac{\partial^2 \mathbf{A}'}{\partial^2 t} = 0.$$

Abbiamo dimostrato che un campo $\mathbf{F}$ solenoidale in un aperto Ω ammette sempre *localmente* un potenziale vettore $\mathbf{V}$. D'altra parte, proprio il campo (vedi esempio 1.5)

$$\mathbf{F}(\mathbf{x}) = \frac{\mathbf{x}}{\rho^3} \qquad (\rho = |\mathbf{x}|)$$

che è solenoidale in $\Omega = \mathbb{R}^3 \setminus \{0\}$, *non ammette potenziale vettore in tutto* Ω. Dimostreremo questo fatto più avanti, usando il teorema del rotore. Lo stesso campo è invece conservativo in $\mathbb{R}^3 \setminus \{0\}$, come si è visto nell'Esempio 1.7.

Per garantire che $\nabla \cdot \mathbf{F} = 0$ diventi anche sufficiente per l'esistenza di un potenziale vettore in tutto Ω intervengono ancora le proprietà topologiche di Ω. Non è più sufficiente che Ω sia *semplicemente connesso:* infatti $\mathbb{R}^3 \setminus \{0\}$ lo è, ma abbiamo appena visto che non basta. Una condizione topologica sufficiente è la seguente: Ω *è connesso e per ogni superficie* $\Sigma \subset \Omega$, *chiusa e regolare, frontiera di un dominio tridimensionale* D, *si ha* $D \subset \Omega$. Diciamo allora che Ω è *fortemente connesso*. Vale allora il seguente risultato che ci limitiamo a enunciare.

TEOREMA 6.5 *Sia* $\mathbf{F} \in C^1(\Omega)$ *tale che* $\nabla \cdot \mathbf{F} = 0$ *in* Ω. *Se* Ω *è fortemente connesso allora esiste* $\mathbf{V} \in C^2(\Omega)$ *tale che* $\nabla \times \mathbf{V} = \mathbf{F}$ *in tutto* Ω.

1.7 Il linguaggio delle forme differenziali

I concetti di *lavoro di un campo vettoriale, campo conservativo, campo irrotazionale,* possono essere espressi equivalentemente usando un diverso linguaggio, quello delle *forme differenziali*. La teoria delle forme differenziali è in realtà una teoria matematica profonda e complessa, che permette di dimostrare risultati molto generali; qui ci interessa soltanto segnalare allo studente una terminologia alternativa per indicare concetti che già conosce. Il motivo è che entrambi i linguaggi sono molto usati nella letteratura scientifica, e vanno pertanto conosciuti. Si può dire che il linguaggio dei campi vettoriali corrisponde più al punto di vista fisico, mentre quello delle forme differenziali corrisponde più al punto di vista matematico.

Linguaggio dei campi vettoriali	Linguaggio delle forme differenziali
Sia $\mathbf{F} = \mathbf{i}F_1 + \mathbf{j}F_2 + \mathbf{k}F_3$	
Lavoro elementare di $\mathbf{F}$ $dL = F_1\,dx + F_2\,dy + F_3\,dz$	Forma differenziale $\omega = F_1\,dx + F_2\,dy + F_3\,dz$
Lavoro del campo $\mathbf{F}$ sul cammino γ $\int_\gamma \mathbf{F}\cdot d\mathbf{r} = \int_a^b \mathbf{F}(\mathbf{r}(t))\cdot \mathbf{r}'(t)\,dt$	Integrale della forma differenziale ω sul cammino γ $\int_\gamma \omega = \int_a^b \mathbf{F}(\mathbf{r}(t))\cdot \mathbf{r}'(t)\,dt$
$\mathbf{F}$ è conservativo $\exists U : \mathbf{F} = \nabla U$	ω è esatta $\exists f : \omega = df = f_x\,dx + f_y\,dy + f_z\,dz$
$\mathbf{F}$ è irrotazionale $\operatorname{rot}\mathbf{F} = 0$	ω è chiusa $\operatorname{rot}\mathbf{F} = 0$
"Un campo conservativo è irrotazionale."	"Una forma differenziale esatta è chiusa."
"Un campo irrotazionale in un aperto semplicemente connesso è conservativo."	"Una forma differenziale chiusa in un aperto semplicemente connesso è esatta."

Esercizi

1 Sia $\mathbf{v}(x,y) = y\mathbf{i}$ il campo di velocità di un fluido in moto stazionario nel piano x,y.
a) Determinare le linee di campo e dedurne che il fluido sembra *non* ruotare.
b) Calcolare il rotore di $\mathbf{v}$. Interpretare il risultato tenendo conto del punto a).

2 Il potenziale gravitazionale generato da una sfera omogenea di raggio R è:

$$U(x,y,z) = \begin{cases} \frac{4}{3}\pi R^3 \frac{\delta}{r} & \text{se } r \geq R \\ 2\pi\delta(R^2 - \frac{r^2}{3}) & \text{se } r < R \end{cases}$$

dove δ è una costante (la densità della sfera) e $r = \sqrt{x^2+y^2+z^2}$. Dopo aver calcolato il campo gravitazionale $\mathbf{F} = \nabla U$, si calcoli la divergenza del campo, verificando che essa è nulla nei punti dello spazio privi di materia e costante nella regione dello spazio in cui vi è una densità costante e positiva di materia.

3 Come l'esercizio precedente, nel caso del potenziale gravitazionale generato da uno strato sferico di raggio interno R_1 e raggio esterno R_2:

$$U(x,y,z) = \begin{cases} \frac{4}{3}\pi\delta(R_2^3 - R_1^3)\frac{1}{r} & \text{se } r > R_2 \\ 2\pi\delta\left(R_2^2 - \frac{r^2}{3}\right) - \frac{4}{3}\pi\delta\frac{R_1^3}{r} & \text{se } R_1 < r \leq R_2 \\ 2\pi\delta(R_2^2 - R_1^2) & \text{se } r \leq R_1 \end{cases}$$

4 Dimostrare tutte le identità differenziali (1.4).

5 Dimostrare che il lavoro di un campo vettoriale non cambia se si riparametrizza la curva mediante un cambiamento di parametro regolare che conserva il verso di percorrenza, mentre cambia di segno se viene cambiato il verso di percorrenza.

6 Dato il campo vettoriale piano $\mathbf{F} = x\mathbf{i} - y\mathbf{j}$:
verificare che è solenoidale in tutto il piano;
cercare un potenziale vettore di $\mathbf{F}$.
Suggerimento: cercare un potenziale ψ indipendente dalla variabile z.

7 Calcolare il lavoro del campo $\mathbf{F} = -y\mathbf{i} + x^2\mathbf{j}$ lungo il triangolo di vertici $(0,0)$, $(1,0)$, $(0,1)$ percorso in senso antiorario.

8 Sia $\mathbf{F} = y\mathbf{i} + x\mathbf{j}$. Determinare le linee integrali del campo e disegnarne alcune. Verificare che il campo è conservativo in $\mathbb{R}^2$ e calcolarne il potenziale. Determinare le linee equipotenziali e disegnarne alcune. Calcolare il lavoro lungo il segmento $y = 2x, x \in (-1,3)$.

9 Stabilire se il campo

$$\mathbf{F} = \frac{y}{z}\mathbf{i} + \frac{x}{z}\mathbf{j} - \frac{xy}{z^2}\mathbf{k}$$

è conservativo nel semispazio $z > 0$, determinando in caso affermativo un potenziale. Il campo è solenoidale nel semispazio $z > 0$?

10 Se il potenziale elettrostatico $U(x,y,z)$ di una carica puntiforme posta nell'origine anziché essere $U = k/r$ fosse $U = k/r^\alpha$ per qualche esponente $\alpha > 0$, $r = \sqrt{x^2+y^2+z^2}$, il campo ∇U sarebbe ancora conservativo (per definizione!). Sarebbe ancora solenoidale, fuori dall'origine? (*Questo argomento può essere utilizzato per dimostrare sperimentalmente che l'esponente "vero" è esattamente, e non solo approssimativamente,* 1).

© 978-88-08-06485-1

Si consideri il campo vettoriale

$$\mathbf{F} = 3x^2y^2z\,\mathbf{i} + (2x^3yz + z)\mathbf{j} + (x^3y^2 + y)\mathbf{k}$$

e l'arco di curva

$$\Gamma: \begin{cases} x = t \\ y = t^2 \\ z = t^3 \end{cases} \qquad t \in [0,1]\,.$$

a) Si calcoli il lavoro di $\mathbf{F}$ lungo Γ, applicando la definizione (cioè calcolando un opportuno integrale).

b) Si verifichi che il campo $\mathbf{F}$ è irrotazionale in tutto lo spazio $\mathbb{R}^3$ e se ne deduca che è conservativo.

c) Si determini un potenziale $U(x,y,z)$ di $\mathbf{F}$ in $\mathbb{R}^3$.

d) Si calcoli il lavoro di $\mathbf{F}$ lungo l'arco di curva Γ' di equazioni parametriche

$$\begin{cases} x = t(2 - t^2) \\ y = \sin(\frac{\pi}{2}t) \\ z = t(3 - 2t) \end{cases} \qquad t \in [0,1]$$

utilizzando il potenziale calcolato al punto precedente.

e) Si calcoli la divergenza di $\mathbf{F}$ e si dica se esiste un insieme aperto di $\mathbb{R}^3$ in cui il campo è solenoidale.

Si consideri il campo vettoriale

$$\mathbf{F} = e^y\mathbf{i} + (xe^y + z\cos(yz))\mathbf{j} + (y\cos(yz))\mathbf{k}.$$

a) Si dica se il campo è irrotazionale nel suo insieme di definizione.

b) Si dica se il campo è conservativo nel suo insieme di definizione, determinando in caso affermativo un potenziale.

c) Si calcoli il lavoro del campo lungo la semicirconferenza che giace sul semipiano $z = 0, y > 0$ e ha per estremi i punti $(0,0,0), (1, \log 2, 0)$.

Applicando l'identità $\nabla \times (\nabla \times \mathbf{G}) = \nabla(\nabla \cdot \mathbf{G}) - \Delta\mathbf{G}$ alle equazioni di Maxwell (1.11) e (1.12), verificare che anche i campi elettrico e magnetico sono soluzioni dell'equazione vettoriale delle onde (si veda l'esempio 1.16).

2 FORMULA DI GAUSS-GREEN NEL PIANO

In questo paragrafo ci occupiamo della relazione tra integrali doppi e integrali di linea. I risultati di questo paragrafo sono casi particolari dei teoremi della divergenza e del rotore, che tratteremo nei paragrafi 5 e 6.

Cominciamo a considerare domini D *semplici* rispetto a entrambi gli assi[7] (fig. 6.5) e un campo vettoriale $\mathbf{F}(x,y) = P(x,y)\mathbf{i} + Q(x,y)\mathbf{j}$ di classe $C^1(D)$, cioè continuo insieme alle sue derivate del prim'ordine fino alla frontiera di D. Diciamo che il bordo di D, ∂D, è *orientato positivamente* quando su di esso è fissato il verso di percorrenza *antiorario*. In tale caso, descrivendo la frontiera nel verso positivo, si lasciano i punti di D a sinistra. Nel caso opposto, ∂D è orientato negativamente. Per sottolineare l'orientazione useremo i simboli $\partial^+ D$ o $\partial^- D$, rispettivamente. Vale il seguente lemma.

[7] Il concetto di dominio semplice è stato definito nel capitolo 5, paragrafo 1.2. A ogni modo, questo concetto viene richiamato esplicitamente nell'enunciato del prossimo lemma.

© 978-88-08-06485-1

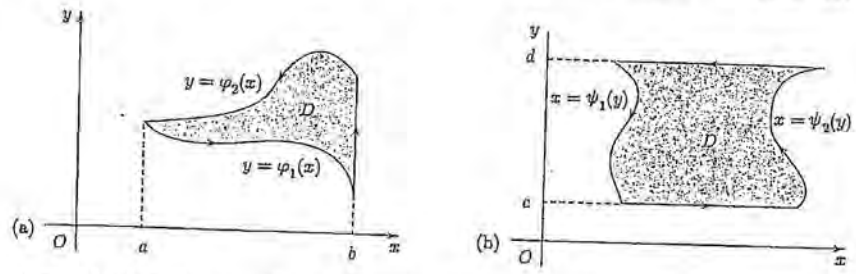


Figura 6.5. a) Dominio y-semplice con orientazione positiva. b) Dominio x-semplice con orientazione positiva.

LEMMA 6.6 *Sia* $\mathbf{F} \in C^1(D)$.

a) Se $D = \{(x,y) \in \mathbb{R}^2;\ a \le x \le b,\ \varphi_1(x) \le y \le \varphi_2(x)\}$ con φ_1, φ_2 continue in $[a,b]$, allora

$$\iint_D P_y\,dx\,dy = -\int_{\partial^+ D} P\,dx. \tag{2.1}$$

b) Se $D = \{(x,y) \in \mathbb{R}^2;\ c \le y \le d;\ \psi_1(y) \le x \le \psi_2(y)\}$ con ψ_1 e ψ_2 continue in $[c,d]$, allora

$$\iint_D Q_x\,dx\,dy = \int_{\partial^+ D} Q\,dy. \tag{2.2}$$

DIMOSTRAZIONE. Le dimostrazioni di a) e b) sono analoghe. Proviamo la a).

Si ha, ricordando il teorema di riduzione per integrali doppi:

$$\iint_D P_y\,dx\,dy = \int_a^b dx \int_{\varphi_1(x)}^{\varphi_2(x)} P_y\,dy = \int_a^b [P(x,\varphi_2(x)) - P(x,\varphi_1(x))]\,dx \tag{2.3}$$

D'altra parte, essendo $dx = 0$ sui tratti rettilinei del bordo:

$$\begin{aligned} \int_{\partial^+ D} P\,dx &= \int_a^b P(x,\varphi_1(x))dx + \int_b^a P(x,\varphi_2(x))dx \\ &= \int_a^b [P(x,\varphi_1(x)) - P(x,\varphi_2(x))] \end{aligned} \tag{2.4}$$

Dalle (2.3) e (2.4) si ottiene la (2.1). ♦

Il seguente teorema è la versione più semplice del teorema di Gauss-Green[8] nel piano.

TEOREMA 6.7 (FORMULA DI GAUSS-GREEN) *Sia D un dominio limitato in $\mathbb{R}^2$ che sia semplice rispetto a entrambi gli assi. Se $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} \in C^1(D)$, allora vale la formula*

$$\iint_D (Q_x - P_y)dx\,dy = \int_{\partial^+ D} P\,dx + Q\,dy. \tag{2.5}$$

[8] George Green, matematico inglese (1793-1841).

© 978-88-08-06485-1

DIMOSTRAZIONE. Nelle ipotesi del teorema valgono entrambe le (2.1) e (2.2) del lemma. Sottraendo membro a membro la (2.1) dalla (2.2) si ottiene la (2.5). ◆

La formula (2.5) vale per domini molto più generali di quelli considerati finora. Per esempio, domini D esprimibili come unione di un numero finito di domini semplici rispetto a entrambi gli assi, a due a due disgiunti. Chiameremo questi domini *s-decomponibili*[9].

L'orientazione positiva di ∂D si ottiene orientando le singole curve che compongono il bordo in modo tale che, percorrendole, si lasci il dominio alla propria sinistra.

Il caso tipico è quello in cui ∂D è unione di una curva γ e di N curve $\gamma_1, \gamma_2, \dots, \gamma_N$ interne a γ, come in figura 6.6.

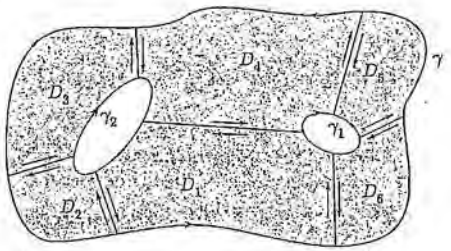


Figura 6.6. Dominio *s*-decomponibile per il teorema di Gauss-Green.

L'orientazione positiva di ∂D richiede il senso antiorario per γ e quello orario per $\gamma_1, \dots, \gamma_N$.

Se $D = \bigcup_{j=1}^{k} D_j$, D_j semplice rispetto a entrambi gli assi, per ogni $j = 1, \dots, k$, usando la (2.5) su ciascun D_j, si ha:

$$\text{(2.6)} \qquad \iint_D (Q_x - P_y)dx\,dy = \sum_{j=1}^{k} \iint_{D_j} (Q_x - P_y)dx\,dy = \sum_{j=1}^{k} \int_{\partial^+ D_j} P\,dx + Q\,dy.$$

Osserviamo ora che gli archi di curva che fanno parte del bordo di due domini adiacenti compaiono due volte nella somma a destra della (2.6) con orientazione opposta e perciò si elidono (fig. 6). Rimangono gli archi di curva che compongono $\partial^+ D$. Dunque

$$\sum_{j=1}^{k} \int_{\partial^+ D_j} P\,dx + Q\,dy = \int_{\partial^+ D} P\,dx + Q\,dy,$$

da cui la (2.5).

[9] La *s* sta per *semplicemente*. Si noti che questi domini sono più particolari dei domini regolari che abbiamo definito, nel capitolo 5, come quei domini unione di un numero finito di domini semplici rispetto ad *almeno un* asse.

© 978-88-08-06485-1

Si noti che nel caso di domini come quello in figura 6.6 la (2.6) si può scrivere più semplicemente nella forma seguente:

$$\text{(2.7)} \qquad \iint_D (Q_x - P_y)dx\,dy = \oint_\gamma P\,dx + Q\,dy + \oint_{\gamma_1} P\,dx + Q\,dy + \oint_{\gamma_2} P\,dx + Q\,dy$$

con ovvio significato dei simboli.

Esempio

Sia D il dominio indicato in figura 6.7 e $\mathbf{F}$ il campo vettoriale di componenti $P(x,y) = y^2$ e $Q(x,y) = x^2$. Si voglia calcolare $\int_{\partial^+ D} y^2\,dx + x^2 dy$. Usando il teorema di Gauss-Green si ha:

$$\int_{\partial^+ D} y^2\,dx + x^2 dy = \iint_D (2x - 2y)dx\,dy = 0$$

essendo D simmetrico rispetto alla bisettrice, mentre $2x - 2y$ cambia segno scambiando x con y.

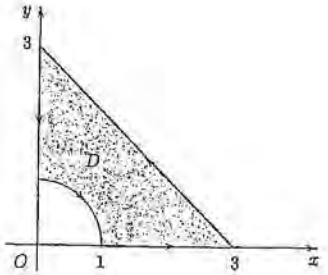


Figura 6.7. Dominio *s*-decomponibile per il teorema di Gauss-Green.

Calcolo di aree mediante integrali curvilinei

Le formule (2.1), (2.2) o (2.5) possono essere usate per calcolare l'area di un dominio D, ammissibile per il teorema 6.7, mediante un integrale curvilineo esteso a ∂D. Infatti, scegliendo $P(x,y) = y$ nella (2.1), otteniamo:

$$\text{(2.8)} \qquad \text{area}(D) = \iint_D dx\,dy = -\int_{\partial^+ D} y\,dx.$$

Scegliendo $Q(x,y) = x$ nella (2.2), otteniamo:

$$\text{(2.9)} \qquad \text{area}(D) = \iint_D dx\,dy = \int_{\partial^+ D} x\,dy.$$

Infine, sommando le (2.8) e (2.9) e dividendo poi per due, si ricava:

$$\text{(2.10)} \qquad \text{area}(D) = \frac{1}{2}\int_{\partial^+ D} x\,dy - y\,dx.$$

© 978-88-08-06485-1

Esempio

Sia D l'interno dell'ellisse di equazione $x^2/a^2 + y^2/b^2 = 1$. L'orientazione positiva del bordo si ottiene con la parametrizzazione

$$\begin{cases} x(\theta) = a\cos\theta \\ y(\theta) = b\sin\theta \end{cases} \quad \theta \in [0, 2\pi].$$

Dalla (2.10):

$$\text{area}\,(D) = \frac{1}{2}\int_0^{2\pi} \left[ab(\cos\theta)^2 + ab(\sin\theta)^2\right] d\theta = \pi ab.$$

Esercizi

14 Usare il teorema di Gauss-Green per valutare l'integrale:

$$\oint_{\gamma^+} (y^2\,dx + x\,dy)$$

dove:
(i) γ è il bordo del quadrato di vertici $(0,0)$, $(1,0)$, $(1,1)$, $(0,1)$;
(ii) γ è l'ellisse $x^2/a^2 + y^2/b^2 = 1$.

15 Calcolare l'area racchiusa dalla curva piana (astroide) di equazioni parametriche:

$$\begin{cases} x = \cos^3 t \\ y = \sin^3 t \end{cases} \quad t \in [0, 2\pi]$$

utilizzando la formula (2.10).

16 Sia γ una curva chiusa semplice e regolare di equazione polare $\rho = f(\theta)$, $\theta \in [\theta_0, \theta_1]$. Se γ è la frontiera di D, dimostrare che:

$$\text{area}(D) = \frac{1}{2}\int_{\theta_1}^{\theta_2} [f(\theta)]^2\,d\theta.$$

3 AREA E INTEGRALI DI SUPERFICIE

3.1 Area di una superficie

Come abbiamo visto, una volta poste le basi del concetto di integrale doppio, è abbastanza naturale generalizzare quest'idea a dimensioni superiori. Altrettanto naturale è generalizzare il discorso in un'altra direzione: integrare una funzione su un oggetto geometrico bidimensionale "curvo" anziché piatto, cioè su una *superficie*, anziché in un insieme del piano.

Le possibili applicazioni dovrebbero essere evidenti: anzitutto, il calcolo dell'area di una superficie; quello della massa totale dislocata su una superficie, che si può vedere come l'integrale (esteso a tutta la superficie) della funzione densità superficiale; analogo discorso vale per il calcolo della carica elettrica totale su una superficie metallica a partire dalla funzione densità superficiale di carica.

© 978-88-08-06485-1

Alla definizione di area di una superficie Σ si potrebbe pervenire con un procedimento simile a quello usato per definire la lunghezza di una curva e cioè: approssimare la superficie con una superficie poliedrica "appoggiata" sulla superficie, per la quale la nozione di area è quella della geometria elementare, per poi definire area di $\sum$ il limite delle aree delle superfici poliedriche approssimanti, al tendere a zero del massimo diametro delle facce. Tale procedimento è tuttavia piuttosto delicato e pertanto daremo una definizione più diretta di elemento di area, con una giustificazione intuitiva.

Consideriamo una superficie regolare Σ descritta da tre equazioni parametriche

$$\begin{cases} x = x(u,v) \\ y = y(u,v) \\ z = z(u,v) \end{cases} \qquad (u,v) \in T \subset \mathbb{R}^2$$

o dall'equazione vettoriale

$$\mathrm{r}\,(u,v) = x(u,v)\mathrm{i} + y\,(u,v)\,\mathrm{j} + z\,(u,v)\,\mathrm{k} \quad (u,v) \in T \subset \mathbb{R}^2;$$

supponiamo che la superficie sia *semplice*, cioè tale che a ogni punto (u,v) *interno* a T corrisponda un solo punto su Σ. In generale T è un dominio s-decomponibile, secondo la definizione data nel paragrafo 2.

Sappiamo che il vettore $\mathrm{r}_u \times \mathrm{r}_v$ è un vettore normale al piano tangente alla superficie. D'altra parte, usando le proprietà del prodotto vettoriale (vedi capitolo 2, paragrafo 1), possiamo affermare che il modulo del vettore $\mathrm{r}_u \times \mathrm{r}_v$ rappresenta l'area del parallelogramma individuato dai due vettori. Quest'area uguaglia in prima approssimazione l'area del "parallelogramma curvilineo" disegnato sulla superficie e compreso tra le quattro linee coordinate

$$u = u_0,\, u = u_0 + du,\, v = v_0,\, v = v_0 + dv$$

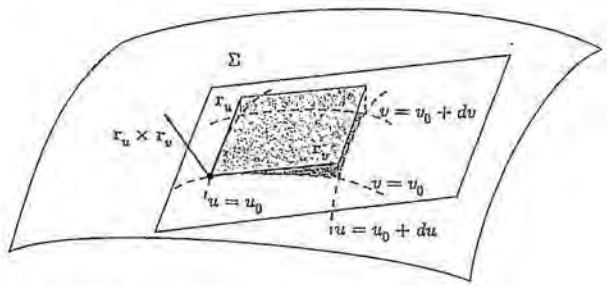


Figura 6.8. I lati del parallelogramma curvilineo sono approssimati a meno di infinitesimi di ordine superiore a du e dv da quelli del parallelogramma rettilineo individuato dai vettori $\mathrm{r}_u du$ e $\mathrm{r}_v dv$.

È allora ragionevole chiamare *elemento d'area sulla superficie* la quantità

$$dS = |\mathrm{r}_u \times \mathrm{r}_v|\,du\,dv$$

e dare la seguente:

© 978-88-08-06485-1

DEFINIZIONE 6.7 L'area di Σ è assegnata dalla formula

$$a(\Sigma) = \iint_\Sigma dS = \iint_T |\mathrm{r}_u \times \mathrm{r}_v| \, du\, dv.$$

Si può controllare che $a(\Sigma)$ *non dipende* dalla parametrizzazione scelta per Σ, come del resto è ragionevole che sia, trattandosi di un'area. Si osservi anche che, per definizione di superficie regolare parametrizzata, l'elemento d'area non si annulla mai.

Esempio

L'elemento d'area sulla superficie della sfera di raggio R è:

$$dS = R^2 \sin\varphi \, d\varphi \, d\theta \qquad (\varphi, \theta) \in [0, \pi] \times [0, 2\pi]$$

e quindi

$$a(\Sigma) = \iint_\Sigma dS = \int_0^{2\pi} d\theta \int_0^{\pi} R^2 \sin\varphi \, d\varphi = 4\pi R^2.$$

Area di una superficie cartesiana

Per una superficie cartesiana, cioè che è grafico di una funzione di due variabili $z = g(x, y)$, $(x, y) \in T$, si ha:

$$\mathrm{r}(x, y) = x\mathrm{i} + y\mathrm{j} + g(x, y)\,\mathrm{k}$$

e

$$\mathrm{r}_x \times \mathrm{r}_y = -g_x\mathrm{i} - g_y\mathrm{j} + \mathrm{k}, \quad |\mathrm{r}_x \times \mathrm{r}_y| = \sqrt{1 + |\nabla g|^2};$$

allora la formula per l'area diventa:

$$a(\Sigma) = \iint_T \sqrt{1 + |\nabla g|^2}\, dx\, dy.$$

Esempi

Consideriamo il *paraboloide ellittico*

$$z = \frac{x^2 + y^2}{2},$$

con $(x, y) \in T$ dove

$$T = \{(x, y) : x^2 + y^2 < 8\}.$$

L'elemento d'area è

$$dS = \sqrt{1 + (x^2 + y^2)}\, dx\, dy$$

e quindi, passando a coordinate polari, si trova:

$$a(\Sigma) = \int_0^{2\pi} d\theta \int_0^{\sqrt{8}} \sqrt{1 + \rho^2}\, \rho\, d\rho =$$
$$= 2\pi \left[\frac{1}{3}\left(1 + \rho^2\right)^{3/2}\right]_0^{\sqrt{8}} = \frac{52}{3}\pi.$$

© 978-88-08-06485-1

Calcoliamo l'*area della superficie del toro*, definito dall'equazione parametrica:

$$\mathrm{r}(\varphi, \theta) = (R + r\cos\varphi)\cos\theta\,\mathrm{i} + (R + r\cos\varphi)\sin\theta\,\mathrm{j} + r\sin\varphi\,\mathrm{k},$$

con $r < R$, $\varphi \in [0, 2\pi]$, $\theta \in [0, 2\pi]$. Si ha:

$$\mathrm{r}_\varphi = -r\sin\varphi\cos\theta\,\mathrm{i} - r\sin\varphi\sin\theta\,\mathrm{j} + r\cos\varphi\,\mathrm{k}$$
$$\mathrm{r}_\theta = -(R + r\cos\varphi)\sin\theta\,\mathrm{i} + (R + r\cos\varphi)\cos\theta\,\mathrm{j}$$

per cui

$$\mathrm{r}_\varphi \times \mathrm{r}_\theta = \begin{vmatrix} \mathrm{i} & \mathrm{j} & \mathrm{k} \\ -r\sin\varphi\cos\theta & -r\sin\varphi\sin\theta & r\cos\varphi \\ -(R + r\cos\varphi)\sin\theta & (R + r\cos\varphi)\cos\theta & 0 \end{vmatrix} =$$

$$= -r(R + r\cos\varphi)\cos\varphi\cos\theta\,\mathrm{i} - r(R + r\cos\varphi)\cos\varphi\sin\theta\,\mathrm{j} +$$
$$+ r\sin\varphi(R + r\cos\varphi)\mathrm{k}$$

e il suo elemento d'area risulta:

$$dS = r(R + r\cos\varphi)d\varphi\, d\theta.$$

Perciò l'area della superficie è:

$$\iint_\Sigma dS = \int_0^{2\pi} \left(\int_0^{2\pi} r(R + r\cos\varphi)d\varphi\right) d\theta =$$
$$= 2\pi\left(2\pi rR + r^2 \int_0^{2\pi} \cos\varphi d\varphi\right) = 4\pi^2 rR,$$

risultato che si può scrivere anche nella forma più espressiva $(2\pi r)\cdot(2\pi R)$: l'area della superficie del toro è il prodotto delle lunghezze di due circonferenze di raggi r, R.

Area di una superficie di rotazione

Calcoliamo l'elemento d'area per le superfici di rotazione della forma[10]

$$\begin{cases} x = x(t)\cos\theta \\ y = x(t)\sin\theta \\ z = z(t) \end{cases} \qquad t \in I \subset \mathbb{R},\ \theta \in [0, 2\pi).$$

Poiché

$$\mathrm{r}_t \times \mathrm{r}_\theta = \det\begin{pmatrix} \mathrm{i} & \mathrm{j} & \mathrm{k} \\ x'(t)\cos\theta & x'(t)\sin\theta & z'(t) \\ -x(t)\sin\theta & x(t)\cos\theta & 0 \end{pmatrix} =$$
$$= -x(t)z'(t)\cos\theta\,\mathrm{i} - x(t)z'(t)\sin\theta\,\mathrm{j} + x(t)x'(t)\,\mathrm{k},$$

l'elemento d'area risulta:

$$dS = |x(t)|\sqrt{x'(t)^2 + z'(t)^2}\, dt\, d\theta.$$

[10] v. capitolo 4, paragrafo 3.

Le formule precedenti vanno ovviamente modificate se l'asse di rotazione è l'asse x, o y, anziché l'asse z.

Esempio

Calcoliamo l'area della superficie generata dalla rotazione attorno all'asse z della curva $x = z^3$, per $z \in [0,1]$:

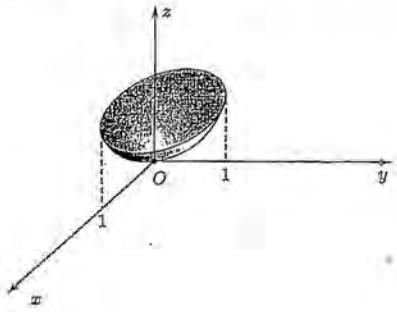


Figura 6.9.

Ponendo $x(t) = t^3, z(t) = t$, si ha:

$$dS = t^3\sqrt{1+9t^4}\,dt\,d\theta$$

e l'area della superficie è:

$$\int_0^{2\pi} d\theta \int_0^1 t^3\sqrt{1+9t^4}\,dt = 2\pi\left[\frac{1}{54}(1+9t^4)^{3/2}\right]_0^1 = \frac{\pi}{27}\left(10\sqrt{10}-1\right).$$

Elemento d'area di superfici in forma implicita

Per le superfici Σ definite da equazioni del tipo

$$F(x,y,z) = 0$$

con, per esempio, $F_z \neq 0$, per il teorema delle funzioni implicite[11] si ha localmente:

$$z = f(x,y);$$
$$\frac{\partial f}{\partial x} = -\frac{F_x}{F_z}; \qquad \frac{\partial f}{\partial y} = -\frac{F_y}{F_z}$$

e pertanto l'elemento d'area su Σ è dato da:

$$dS = \sqrt{1+\left(\frac{F_x}{F_z}\right)^2+\left(\frac{F_y}{F_z}\right)^2}\,dxdy = \frac{|\nabla F|}{|F_z|}dxdy.$$

[11] v. capitolo 3, paragrafo 8, teorema 3.26.

3.2 Integrale di superficie di una funzione continua

Se F è una funzione continua, definita in una regione dello spazio contenente la superficie Σ, la sua restrizione sulla superficie è $\tilde{F}(u,v) = F(x(u,v), y(u,v), z(u,v))$ ed è naturale definire l'integrale di F su Σ mediante la formula:

$$\iint_\Sigma F dS = \iint_T \tilde{F}(u,v)\,|\mathrm{r}_u \times \mathrm{r}_v|\,du\,dv.$$

L'integrale di superficie di F viene dunque espresso mediante un integrale doppio nel dominio piano T, che si calcola con le tecniche studiate nel capitolo 5.

Esempi

Sia Σ la semisfera superiore di raggio 2 e centro $(0,0,0)$; calcoliamo l'integrale di superficie:

$$I = \iint_\Sigma x^2 z\,dS.$$

Le equazioni parametriche della semisfera sono:

$$\begin{cases} x(\varphi,\theta) = 2\sin\varphi\cos\theta \\ y(\varphi,\theta) = 2\sin\varphi\sin\theta \\ z(\varphi,\theta) = 2\cos\varphi \end{cases} \quad \text{con} \quad \varphi \in [0,\tfrac{\pi}{2}],\ \theta \in [0,2\pi],$$

mentre $dS = 4\sin\varphi\,d\varphi\,d\theta$. Allora:

$$I = \int_0^{2\pi}\left(\int_0^{\frac{\pi}{2}}(2\sin\varphi\cos\theta)^2 2\cos\varphi 4\sin\varphi\,d\varphi\right)d\theta =$$
$$= 32\left(\int_0^{2\pi}\cos^2\theta d\theta\right)\left(\int_0^{\frac{\pi}{2}}\sin^3\varphi\cos\varphi\,d\varphi\right) =$$
$$= 32\pi\left[\frac{\sin^4\varphi}{4}\right]_0^{\pi/2} = 8\pi.$$

Integrale di una superficie cartesiana. Se Σ è il grafico di $z = g(x,y)$, $(x,y) \in T$, l'integrale di superficie di $F = F(x,y,z)$ in questo caso è:

$$\iint_T F(x,y,g(x,y))\sqrt{1+|\nabla g(x,y)|^2}dx\,dy.$$

Per esempio, l'integrale dell'esempio 3.5 poteva essere calcolato anche per questa via: la semisfera superiore ha equazione cartesiana:

$$z = \sqrt{4-(x^2+y^2)}.$$

Essendo

$$\nabla g(x,y) = \left(-\frac{x}{\sqrt{4-(x^2+y^2)}}, -\frac{y}{\sqrt{4-(x^2+y^2)}}\right),$$

l'elemento d'area diventa:

$$dS = \sqrt{1+\frac{x^2+y^2}{4-(x^2+y^2)}}dx\,dy = \frac{2dx\,dy}{\sqrt{4-(x^2+y^2)}}$$

© 978-88-08-06485-1

e l'integrale di superficie è:

$$\iint_{\{x^2+y^2<4\}} x^2\sqrt{4-(x^2+y^2)}\,\frac{2dx\,dy}{\sqrt{4-(x^2+y^2)}} =$$
$$= 2\iint_{\{x^2+y^2<4\}} x^2\,dx\,dy = 2\int_0^{2\pi}\cos^2\theta\,d\theta\int_0^2\rho^3 d\rho = 8\pi.$$

Calcolo di baricentri e momenti d'inerzia

Nello spazio tridimensionale, sia Σ una superficie materiale non omogenea di *densità superficiale* ρ e sia $\mathrm{r}: T \to \mathbb{R}^3$ una parametrizzazione regolare di Σ,

$$\mathrm{r}(u,v) = (x(u,v), y(u,v), z(u,v)).$$

Allora la *massa totale di* Σ è data da:

$$m = \iint_\Sigma \rho\,dS = \iint_T \rho(\mathrm{r}(u,v))\,|\mathrm{r}_u \times \mathrm{r}_v|\,du\,dv.$$

Se il corpo è *omogeneo* (ρ = costante) si ritrova:

$$m = \rho \cdot a(\Sigma).$$

Il *baricentro* di Σ è il punto $B \equiv (\overline{x}, \overline{y}, \overline{z})$, dove

$$\begin{cases} \overline{x} = \dfrac{1}{m}\displaystyle\iint_\Sigma x\rho\,dS = \frac{1}{m}\iint_T x(u,v)\,\rho(\mathrm{r}(u,v))\,|\mathrm{r}_u \times \mathrm{r}_v|\,du\,dv \\ \overline{y} = \dfrac{1}{m}\displaystyle\iint_\Sigma y\rho\,dS = \frac{1}{m}\iint_T y(u,v)\,\rho(\mathrm{r}(u,v))\,|\mathrm{r}_u \times \mathrm{r}_v|\,du\,dv \\ \overline{z} = \dfrac{1}{m}\displaystyle\iint_\Sigma z\rho\,dS = \frac{1}{m}\iint_T z(u,v)\,\rho(\mathrm{r}(u,v))\,|\mathrm{r}_u \times \mathrm{r}_v|\,du\,dv \end{cases}$$

Se la superficie è omogenea (ρ = costante), il baricentro si dice *centroide* e ha coordinate:

$$\overline{x} = \frac{\rho}{m}\iint_\Sigma x\,dS = \frac{1}{a(\Sigma)}\iint_T x(u,v)\,|\mathrm{r}_u \times \mathrm{r}_v|\,du\,dv = \frac{\displaystyle\iint_T x(u,v)\,|\mathrm{r}_u \times \mathrm{r}_v|\,du\,dv}{\displaystyle\iint_T |\mathrm{r}_u \times \mathrm{r}_v|\,du\,dv},$$

con analoghe espressioni per $\overline{y}, \overline{z}$.

Calcoliamo ora il *momento d'inerzia di* Σ *rispetto a un asse fissato*. Se $\delta(x,y,z)$ indica la distanza di (x,y,z) da quest'asse, si ha:

$$I = \iint_\Sigma \delta^2\rho\,dS = \iint_T \delta^2(\mathrm{r}(u,v))\rho(\mathrm{r}(u,v))\,|\mathrm{r}_u \times \mathrm{r}_v|\,du\,dv.$$

Se il corpo è omogeneo,

$$I = \frac{m}{a(\Sigma)}\iint_\Sigma \delta^2\,dS = m\,\frac{\displaystyle\iint_T \delta^2(\mathrm{r}(u,v))\,|\mathrm{r}_u \times \mathrm{r}_v|\,du\,dv}{\displaystyle\iint_T |\mathrm{r}_u \times \mathrm{r}_v|\,du\,dv}.$$

Osserviamo che, anche in questo caso, il momento d'inerzia ha le dimensioni di *massa* × [*lunghezza*]^2.

© 978-88-08-06485-1

Esercizi

Calcolare l'area della porzione di paraboloide $z = x^2 + y^2$ per $x^2 + y^2 < r^2$.

Calcolare la massa totale di una superficie sferica di raggio R la cui densità è funzione della latitudine, al modo seguente: $f(\varphi,\theta) = k\varphi(\pi - \varphi)$, con k costante.

Calcolare l'area della superficie generata dalla rotazione attorno all'asse z della curva $z = x^2, x \in [1,2]$.

Calcolare il baricentro e il momento d'inerzia della superficie laterale di un cono omogeneo di massa M, raggio R e altezza h. Fare lo stesso poi per la superficie completa del cono (aggiungere cioè anche la base).

Calcolare il momento d'inerzia di una superficie sferica omogenea di raggio R e massa M, rispetto a un asse passante per il centro.

Sia Σ la semisfera superiore di centro l'origine e raggio 1. Calcolare:

$$\iint_\Sigma (x^2 + z)dS.$$

Calcolare l'area della porzione di superficie di equazioni parametriche:

$$\begin{cases} x = u\cos t \\ y = u\sin t \\ z = t \end{cases} \qquad u \in [0,1],\ t \in [0, 2\pi].$$

Con l'aiuto di un computer, tracciare il grafico della superficie dell'esercizio precedente.

Calcolare l'integrale di superficie:

$$\iint_\Sigma \frac{x^2 - z}{\sqrt{1 + 4(x^2+y^2)}}\,dS,$$

dove Σ è la porzione di grafico della funzione $z = x^2 - y^2$ per $(x,y) \in T$ e

$$T = \{(x,y) : x^2 + 4y^2 \le 4\}.$$

4 INTEGRALE DI SUPERFICIE DI UN CAMPO VETTORIALE. FLUSSO

Consideriamo un campo vettoriale $\mathbf{F}$ e una superficie Σ, trasversale alle linee di campo. L'integrale su Σ della componente di $\mathbf{F}$ normale a Σ rappresenta *il flusso di* $\mathbf{F}$ *attraverso* Σ e gioca un ruolo importante nella formulazione di molte leggi fisiche, nonché in numerose applicazioni. Questo integrale è dunque un integrale di superficie di una funzione scalare, ma per definirlo correttamente occorrono ingredienti un po' diversi rispetto a prima e precisamente:

- una superficie regolare Σ;
- un campo vettoriale $\mathbf{F}$, definito almeno nei punti della superficie;
- una scelta coerente del versore normale $\mathbf{n}$ su *tutta* la superficie.

L'ultimo punto conduce alle nozioni di *orientazione* e di *bordo* di una superficie.

© 978-88-08-06485-1

4.1 Superfici orientate. Bordo di una superficie. Superfici regolari a pezzi

Sia Σ una superficie regolare parametrizzata da

$$\mathrm{r}(u,v) = x(u,v)\mathrm{i} + y(u,v)\mathrm{j} + z(u,v)\mathrm{k} \quad (u,v) \in T \subset \mathbb{R}^2,$$

dove T è un dominio s-decomponibile. I versori

$$\mathrm{n} = \mathrm{vers}\,(\mathrm{r}_u \times \mathrm{r}_v)$$

e $-\mathrm{n}$ sono entrambi normali alla superficie Σ e la scelta di uno dei due corrisponde intuitivamente a privilegiare uno dei "due lati" di Σ. Per esempio, possiamo pensare al "lato interno" oppure al "lato esterno" di una superficie sferica. Più precisamente, la scelta di uno dei due versori normali è legata al concetto di *orientazione* di Σ.

DEFINIZIONE 6.8 Una superficie regolare Σ si dice *orientabile* se, per ogni curva continua chiusa che giace sulla superficie, parametrizzata da $\mathrm{r} : [a,b] \to \Sigma$, si ha che $\mathrm{n}(\mathrm{r}(b)) = \mathrm{n}(\mathrm{r}(a))$.

In altre parole, se seguiamo il versore normale alla superficie lungo una curva chiusa che giace sulla superficie, dopo un giro completo il versore normale risulta ancora puntato nello stesso verso.

La scelta di n o di $-\mathrm{n}$ come versore normale determina l'*orientazione* di Σ; si dice *superficie orientata* una superficie su cui sia stata scelta una delle due orientazioni possibili. ■

Esempi di superfici orientabili sono le superfici cartesiane. Per una superficie di equazione $z = g(x,y)$, si ha:

$$\mathrm{n} = \frac{-g_x\mathrm{i} - g_y\mathrm{j} + \mathrm{k}}{\sqrt{1+|\nabla g|^2}},$$

e questo versore normale risulta essere rivolto sempre *verso l'alto* (cioè forma un angolo acuto col versore dell'asse z), mentre $-\mathrm{n}$ punta sempre *verso il basso*.

Invece, per una superficie come la sfera o il toro, il versore n risulta puntare sempre *verso l'interno* o sempre verso l'esterno della superficie, a seconda di quale dei due parametri chiamiamo u e quale v: questo, naturalmente, è arbitrario. In questo senso siamo noi che *scegliamo un'orientazione sulla superficie*. Si parla in questo caso di *normale interna* o *esterna* (o anche *entrante* o *uscente*), rispettivamente.

Tutti questi sono esempi in cui una scelta coerente di un'orientazione su tutta la superficie è *possibile*. Non sempre è così: esistono *superfici non orientabili*, come il *nastro di Moebius* (fig. 6.10): comunque si scelga l'orientazione del versore n in un punto, muovendolo con continuità lungo una linea meridiana, possiamo ritrovarci nello stesso punto con orientazione opposta (si veda l'esercizio 36).

Veniamo ora al *bordo* di una superficie orientata. Si tratta di una nozione piuttosto delicata, per cui, rinunciando alla massima generalità, preferiamo concentrarci su alcuni casi particolari.

Supponiamo che T sia s-decomponibile; questo in particolare implica che T è chiuso (perché unione di un numero finito di insiemi chiusi).

© 978-88-08-06485-1

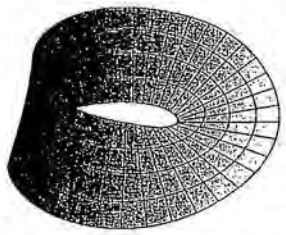


Figura 6.10.

a) Se $\mathrm{r} = \mathrm{r}(u,v)$ realizza una corrispondenza *biunivoca* tra T e Σ, questo garantisce che Σ non presenti autointersezioni. Inoltre, poiché T è s-decomponibile, la frontiera di T è unione di una curva γ_0 e di N curve $\gamma_1, \dots, \gamma_N$ interne a γ_0, (fig. 6.11). Il *bordo di* Σ, che indichiamo con $\partial\Sigma$, è allora dato dall'unione delle immagini $\Gamma_0, \Gamma_1, \dots, \Gamma_N$ delle curve $\gamma_0, \gamma_1, \dots, \gamma_N$

Un'orientazione di Σ, ottenuta scegliendone uno dei due lati, induce un'orientazione, che chiameremo *positiva*, su ciascuna delle curve che ne compongono il bordo: *si sceglie il verso di percorrenza di ogni* Γ_j, $j = 0,1,\dots,N$, *in modo che, percorrendole rimanendo sul lato di* Σ *scelto come positivo, si lasciano i punti di* Σ *alla sinistra* (fig. 6.11). Quando il bordo di una superficie è orientato positivamente, useremo il simbolo $\partial^+\Sigma$.

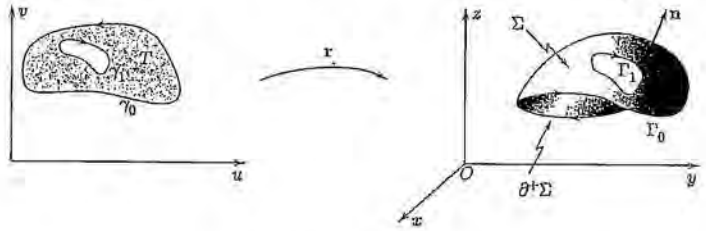


Figura 6.11. Superficie con bordo orientato positivamente.

b) Se $\mathrm{r} = \mathrm{r}(u,v)$ è *biunivoca* solo all'interno di T, il *bordo* di Σ è contenuto nell'unione delle immagini $\Gamma_0, \Gamma_1, \dots, \Gamma_N$ delle curve $\gamma_0, \gamma_1, \dots, \gamma_N$, ma alcune delle curve Γ_j possono non far parte del bordo.

Per la semisfera dell'esempio 3.5 si ha $T = [0, \frac{\pi}{2}] \times [0, 2\pi]$ e l'immagine del bordo di T è il cerchio equatoriale.

Per il toro, $T = [0, 2\pi] \times [0, 2\pi]$ e l'immagine di ∂T è l'unione del cerchio Γ_1 nel piano $z = 0$, di raggio $(R+r)$ e centrato nell'origine, di equazione

$$\mathrm{r}(0,\theta) = \mathrm{r}(2\pi,\theta) = (R+r)\cos\theta\,\mathrm{i} + (R+r)\sin\theta\,\mathrm{j}$$

e del cerchio nel piano $x = R$, di raggio r e centrato nel punto $(1,0,0)$, di equazione

$$\mathrm{r}(\varphi,0) = \mathrm{r}(\varphi,2\pi) = (R + r\cos\varphi)\,\mathrm{i} + r\sin\varphi\,\mathrm{k}.$$

© 978-88-08-06485-1

Chiaramente, queste curve non costituiscono il bordo del toro, che infatti è una superficie *priva di bordo*.

Le superfici prive di bordo si dicono *chiuse* e alcune di esse si incontrano comunemente. Esempi tipici sono la sfera, l'ellissoide, il cilindro illimitato, che dividono lo spazio in due regioni disgiunte (interno ed esterno) e risultano pertanto orientabili (scegliendo la normale entrante o quella uscente).

Esistono tuttavia superfici *chiuse e non orientabili*, come quella un poco "patologica" illustrata in figura 6.12, che prende il nome di *otre di Klein*.

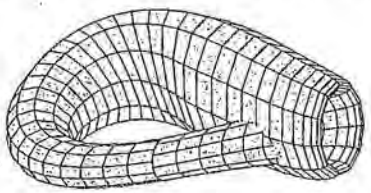


Figura 6.12. L'*otre di Klein*. È una superficie chiusa ma non si può parlare di *interno* o *esterno* alla superficie.

Nel seguito avremo bisogno di considerare superfici *regolari a pezzi*, come un parallelepipedo, che lungo gli spigoli non è dotato di versore normale: che cosa significa in questo caso "scegliere coerentemente n"? Che cos'è il bordo?

DEFINIZIONE 6.9 Σ è una superficie *regolare a pezzi* se esiste un numero finito di curve regolari a tratti, dette spigoli, $\sigma_1, \ldots, \sigma_N$, contenute in Σ, che suddividono Σ in un numero finito N_0 di superfici regolari Σ_j, dette facce. ∎

Il *bordo* di una superficie regolare a pezzi è costituito dall'unione dei bordi delle facce Σ_j con esclusione degli spigoli che appartengono al bordo di due facce adiacenti.

Se una superficie è regolare a pezzi, mediante una deformazione continua può essere trasformata in una superficie regolare. Per esempio, un parallelepipedo può essere trasformato in una superficie regolare con una piccola deformazione continua che ne smussi gli spigoli. Diremo che una superficie regolare a pezzi è orientabile quando la superficie regolare che si ottiene con una deformazione continua della superficie di partenza è orientabile, nel senso della definizione 6.8.

Se una superficie regolare a pezzi e orientabile è *chiusa*, è sufficiente parlare di "normale entrante" o "normale uscente", per individuare un'orientazione coerente sulle varie "facce". Non dovrebbe essere difficile per il lettore capire come orientare

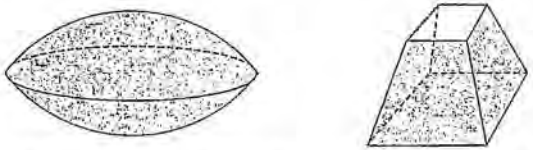


Figura 6.13. Esempi di superfici C^∞ a pezzi: guscio formato da due calotte sferiche (superficie chiusa) e superficie laterale di un tronco di piramide ($\delta\Sigma$ è l'unione dei bordi di due quadrilateri, basi del tronco).

© 978-88-08-06485-1

coerentemente le facce di una superficie regolare a pezzi, non chiusa (naturalmente, nei casi in cui ciò sia possibile).

4.2 Flusso

Veniamo ora alla definizione di flusso di un campo vettoriale attraverso una superficie orientata.

In *fluidodinamica* il flusso di un fluido, (per esempio l'acqua in un canale) rappresenta la *portata*. Più precisamente, sia $\mathbf{v}$ il campo di velocità dell'acqua; ci chiediamo quale sia il volume d'acqua che attraversa nell'unità di tempo una certa superficie ideale. Questa quantità è detta *portata*. Rappresentiamo un elemento infinitesimo di superficie, dS, e il suo versore normale, $\mathbf{n}$; consideriamo il prodotto scalare di $\mathbf{v}$ con $\mathbf{n}$, che misura l'avanzamento istantaneo dell'acqua in quel punto, rispetto alla superficie: se moltiplichiamo questo numero per l'area infinitesima dS, otterremo il volume infinitesimo d'acqua che fluisce attraverso la superficie nel punto, nell'unità di tempo:

$$\mathbf{v} \cdot \mathbf{n}\, dS = \frac{dV}{dt}.$$

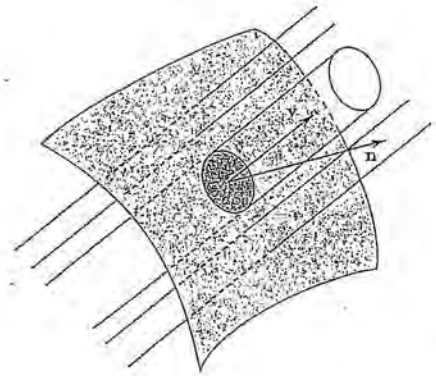


Figura 6.14.

Se vogliamo ottenere il *flusso totale del campo* $\mathbf{v}$ *attraverso la superficie orientata* Σ, basta integrare questi flussi elementari, ottenendo l'integrale di superficie:

$$\Phi = \iint_\Sigma \mathbf{v} \cdot \mathbf{n}\, dS$$

Altri ambiti in cui la nozione di flusso si presenta naturalmente sono i seguenti. In *elettrodinamica*: se $\mathbf{j}$ è il vettore densità di corrente elettrica, il suo flusso attraverso Σ dà la corrente che transita per quella sezione; in *termodinamica*: se $\mathbf{j}$ è il vettore densità di flusso di calore, il suo flusso attraverso Σ dà la quantità di calore che attraversa Σ nell'unità di tempo.

© 978-88-08-06485-1

In generale:

Definizione 6.10 Dati una superficie regolare orientata Σ con versore normale $\mathbf{n}$ e un campo vettoriale $\mathbf{F}$ di classe C^1 in un intorno di Σ, si definisce *flusso del vettore* $\mathbf{F}$ attraverso Σ nella direzione e verso di $\mathbf{n}$ l'integrale

$$\Phi(\mathbf{F};\Sigma) = \iint_{\Sigma} \mathbf{F}\cdot\mathbf{n}\,dS.$$

Per il calcolo dell'integrale, se

$$\mathbf{r}(u,v) = x(u,v)\mathbf{i} + y(u,v)\mathbf{j} + z(u,v)\mathbf{k} \quad (u,v)\in T\subset\mathbb{R}^2$$

è la parametrizzazione usata per Σ, essendo

$$\mathbf{n} = \frac{\mathbf{r}_u\times\mathbf{r}_v}{|\mathbf{r}_u\times\mathbf{r}_v|} \quad \text{e} \quad dS = |\mathbf{r}_u\times\mathbf{r}_v|\,du dv,$$

si ottiene, posto $\tilde{\mathbf{F}}(u,v) = \mathbf{F}(x(u,v), y(u,v), z(u,v))$:

$$\Phi(\mathbf{F};\Sigma) = \iint_T \tilde{\mathbf{F}}\cdot(\mathbf{r}_u\times\mathbf{r}_v)\,du dv.$$

Notiamo che *se si cambia l'orientazione sulla superficie* (ovvero il segno del vettore $(\mathbf{r}_u\times\mathbf{r}_v)$), *il flusso cambia segno.*

Per una *superficie cartesiana*, di equazione $z = g(x,y)$ con $g(x,y)\in T$, se si sceglie la normale *verso l'alto*, si ha:

$$\mathbf{n} = \frac{-g_x\mathbf{i} - g_y\mathbf{j} + \mathbf{k}}{\sqrt{1+|\nabla g|^2}}, \quad dS = \sqrt{1+|\nabla g|^2}\,dx dy$$

e quindi si trova, posto $\tilde{\mathbf{F}}(x,y) = \mathbf{F}(x,y,g(x,y))$:

$$\Phi(\mathbf{F};\Sigma) = \iint_T \left[-\tilde{F}_1 g_x - \tilde{F}_2 g_y + \tilde{F}_3\right] dx dy.$$

Esempio

Calcoliamo il flusso uscente dalla superficie sferica di raggio R e centro l'origine, per il campo $\mathbf{F}(x,y) = x\mathbf{i} + y\mathbf{k}$.

Scriviamo la superficie in forma cartesiana, come unione delle due semisfere S^+, S^-.

La semisfera superiore S^+ è definita da

$$z = \sqrt{R^2-(x^2+y^2)}, \text{ per } x^2+y^2 < R^2$$

con versore uscente $\mathbf{n}$ che punta *verso l'alto*. Abbiamo:

$$-g_x\mathbf{i} - g_y\mathbf{j} + \mathbf{k} = \frac{x}{\sqrt{R^2-(x^2+y^2)}}\mathbf{i} + \frac{y}{\sqrt{R^2-(x^2+y^2)}}\mathbf{j} + \mathbf{k}$$

e quindi:

$$\begin{aligned}\Phi(\mathbf{F},S^+) &= \iint_{\{x^2+y^2<R^2\}} \left(\frac{x^2}{\sqrt{R^2-(x^2+y^2)}} + y\right) dx\,dy = \text{(per simmetria)}\\ &= \iint_{\{x^2+y^2<R^2\}} \frac{x^2}{\sqrt{R^2-(x^2+y^2)}}\,dx\,dy = \\ &= \int_0^{2\pi} (\cos\theta)^2\,d\theta \int_0^R \frac{\rho^3}{\sqrt{R^2-\rho^2}}\,d\rho = \pi\int_0^R \frac{\rho^3}{\sqrt{R^2-\rho^2}}\,d\rho.\end{aligned}$$

L'ultimo integrale si calcola per parti:

$$\begin{aligned}\int_0^R \rho^2\frac{\rho}{\sqrt{R^2-\rho^2}}\,d\rho &= \left[-\rho^2\sqrt{R^2-\rho^2}\right]_0^R + \int_0^R 2\rho\sqrt{R^2-\rho^2}\,d\rho = \\ &= \left[-\frac{2}{3}(R^2-\rho^2)^{3/2}\right]_0^R = \frac{2}{3}R^3.\end{aligned}$$

Quindi $\Phi(\mathbf{F},S^+) = \frac{2}{3}\pi R^3$.

La semisfera inferiore S^- è definita da

$$z = -\sqrt{R^2-(x^2+y^2)}, \quad \text{per } x^2+y^2 < R^2$$

con versore uscente $\mathbf{n}$ che punta verso il basso (attenzione a questo punto: la scelta di $\mathbf{n}$ sull'intera superficie dev'essere coerente), perciò abbiamo:

$$g_x\mathbf{i} + g_y\mathbf{j} - \mathbf{k} = \frac{x}{\sqrt{R^2-(x^2+y^2)}}\mathbf{i} + \frac{y}{\sqrt{R^2-(x^2+y^2)}}\mathbf{j} - \mathbf{k}$$

e quindi, usando i calcoli precedenti:

$$\begin{aligned}\Phi(\mathbf{F},S^-) &= \iint_{\{x^2+y^2<R^2\}} \left(\frac{x^2}{\sqrt{R^2-(x^2+y^2)}} - y\right) dx\,dy = \text{(per simmetria)}\\ &= \iint_{\{x^2+y^2<R^2\}} \frac{x^2}{\sqrt{R^2-(x^2+y^2)}}\,dx\,dy = \frac{2}{3}\pi R^3.\end{aligned}$$

Il flusso totale uscente è dunque uguale a $\frac{4}{3}\pi R^3$.

Invitiamo il lettore a eseguire il calcolo del flusso usando la rappresentazione parametrica della sfera.

Vediamo ora un altro esempio, tecnicamente più semplice e di interesse applicativo.

Esempio

Sia

$$\mathbf{E} = \kappa q\frac{\mathbf{r}}{|\mathbf{r}|^3},$$

con $\mathbf{r}(x,y,z) = x\mathbf{i}+y\mathbf{j}+z\mathbf{k}$ e κ costante, il campo elettrostatico $\mathbf{E}$ generato da una carica puntiforme q posta nell'origine. Calcoliamo il flusso di $\mathbf{E}$ uscente dalla sfera Σ_R con centro nell'origine e raggio R.

La normale esterna alla sfera è

$$\mathbf{n} = \frac{\mathbf{r}}{|\mathbf{r}|},$$

 © 978-88-08-06485-1

perciò, osservando che $\Sigma_R = \{(x,y,z) : |\mathbf{r}(x,y,z)| = R\}$, il flusso è dato da:

$$\Phi(\mathbf{E},\Sigma_R) = \iint_{\{|\mathbf{r}|=R\}} \kappa q \frac{\mathbf{r}}{|\mathbf{r}|^3}\cdot\frac{\mathbf{r}}{|\mathbf{r}|}dS = \iint_{\{|\mathbf{r}|=R\}} \kappa q \frac{|\mathbf{r}|^2}{|\mathbf{r}|^4}dS = \frac{\kappa q}{R^2}\iint_{\{|\mathbf{r}|=R\}} dS$$
$$= \frac{\kappa q}{R^2}\cdot 4\pi R^2 = (4\pi\kappa)\, q.$$

Pertanto, il flusso è proporzionale alla carica q e la costante non dipende dal raggio della sfera. Il risultato ottenuto è semplice ma significativo e permette di dimostrare un risultato molto più generale, noto come *teorema di Gauss dell'elettrostatica: data nello spazio una distribuzione continua di carica, il flusso del campo elettrico uscente da una qualsiasi superficie chiusa Σ è pari a $4\pi\kappa\, q_{tot}$, dove q_{tot} è la carica totale racchiusa nella regione delimitata da Σ.*

5 TEOREMA DELLA DIVERGENZA

Il teorema della divergenza ha svolto un ruolo fondamentale per gli sviluppi classici e moderni dell'Analisi Matematica e delle scienze applicate.

Da un punto di vista matematico, questo risultato riunisce molti dei concetti sviluppati fin qui: derivate parziali, superfici, integrali di superficie, campi vettoriali, flusso di un campo vettoriale.

Da un punto di vista applicativo, questo teorema si trova nella "cassetta degli attrezzi" di pressoché ogni branca della fisica matematica: in particolare, è alla base della deduzione matematica di molte equazioni fondamentali della fisica e della possibilità di trasformare relazioni di tipo integrale in altre di tipo differenziale o viceversa. Tutto ciò potrà essere illustrato solo in minima parte in questo testo, ma proseguendo i suoi studi, il lettore avrà certamente modo di riscontrare quanto qui affermato.

TEOREMA 6.8 (DELLA DIVERGENZA, O DI GAUSS) *Sia $D \subset \mathbb{R}^3$ un dominio limitato, semplice rispetto a tutti e tre gli assi cartesiani, la cui frontiera è una superficie regolare a pezzi e orientabile; indichiamo con $\mathbf{n}_e$ il versore normale esterno a ∂D e sia $\mathbf{F} = F_1\mathbf{i} + F_2\mathbf{j} + F_3\mathbf{k}$ un campo vettoriale $C^1(D)$. Allora vale la formula:*

$$\iiint_D \nabla\cdot\mathbf{F}\, dxdydz = \iint_{\partial D} \mathbf{F}\cdot\mathbf{n}_e\, dS \tag{5.1}$$

dove, ricordiamo,

$$\nabla\cdot\mathbf{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z}.$$

Il teorema afferma che *il flusso di un campo vettoriale uscente da una superficie chiusa uguaglia l'integrale della divergenza del campo nella regione racchiusa dalla superficie stessa.*

Volendo dare subito un'interpretazione di questo teorema, possiamo dire che la (5.1) rappresenta un'*equazione di bilancio*: la divergenza di un campo vettoriale è una misura dell'intensità delle sorgenti (o dei pozzi, a seconda del segno) del campo; usando, per esempio, il linguaggio della meccanica dei fluidi, l'integrale della divergenza del campo velocità è il bilancio complessivo di quanta acqua zampilla dalle sorgenti, meno quanta acqua viene inghiottita nei pozzi, nella regione racchiusa dalla superficie, nell'unità di tempo. D'altro canto, il flusso della velocità attraverso la superficie chiusa rappresenta il saldo totale di quanta acqua esce dalla regione racchiusa dalla superficie, meno quanta acqua vi entra, nell'unità di tempo: è naturale che le due quantità debbano coincidere (la differenza tra ciò che esce e ciò che entra deve essere stata prodotta all'interno).

ESEMPIO

Calcolare il flusso uscente dalla superficie sferica Σ_R di raggio R e centro nell'origine, per il campo $\mathbf{F} = x\mathbf{i} + y\mathbf{k}$.

È quanto abbiamo già fatto nell'esempio 4.1. Calcoliamo lo stesso flusso utilizzando stavolta il teorema della divergenza. Abbiamo:

$$\nabla\cdot\mathbf{F} = \frac{\partial F_1}{\partial x} + \frac{\partial F_2}{\partial y} + \frac{\partial F_3}{\partial z} = 1 + 0 + 0 = 1$$

e quindi

$$\Phi(\mathbf{F};\Sigma_R) = \iiint_{\{x^2+y^2+z^2<R^2\}} 1\, dx\, dy\, dz = \text{Volume della sfera} = \frac{4}{3}\pi R^3.$$

Per dimostrare il teorema, cominciamo a stabilire il seguente lemma valido in un dominio *semplice rispetto all'asse z*, ovvero

$$D = \{(x,y,z) : \varphi_1(x,y) < z < \varphi_2(x,y),\ (x,y) \in B \subset \mathbb{R}^2\}$$

dove B è un aperto limitato e semplicemente connesso del piano e φ_1, φ_2 sono funzioni di classe $C^1(\overline{B})$.

LEMMA 6.9 *Nelle ipotesi precedenti, vale la formula:*

$$\iiint_D (F_3)_z\, dxdydz = \iint_{\partial D} F_3\, \mathbf{n}_e\cdot\mathbf{k}\, dS. \tag{5.2}$$

DIMOSTRAZIONE. Consideriamo l'integrale a sinistra. Dalla formula di riduzione degli integrali tripli, si ha:

$$\iiint_D (F_3)_z\, dxdydz = \iint_B dxdy \int_{\varphi_1(x,y)}^{\varphi_2(x,y)} (F_3)_z\, dz =$$
$$= \iint_B [F_3(x,y,\varphi_2(x,y)) - F_3(x,y,\varphi_1(x,y))]\, dxdy. \tag{5.3}$$

Per l'integrale a destra, osserviamo che ∂D è l'unione di tre superfici:

1. Σ_1, di equazione $z = \varphi_1(x,y)$ con normale esterna ed elemento d'area dati da:

$$\mathbf{n}_e = \frac{\varphi_{1x}\mathbf{i} + \varphi_{1y}\mathbf{j} - \mathbf{k}}{\sqrt{1+|\nabla\varphi_1|^2}}, \qquad dS = \sqrt{1+|\nabla\varphi_1|^2}\, dxdy$$

e quindi, su Σ_1, si ha $\mathbf{n}_e\cdot\mathbf{k}\, dS = -dxdy$.

© 978-88-08-06485-1

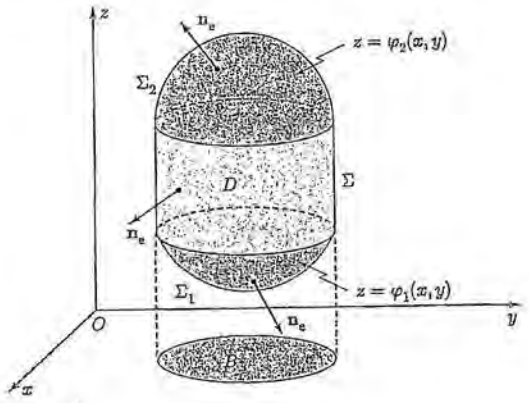


Figura 6.15. Dominio semplice rispetto all'asse z.

2. Σ_2, di equazione $z = \varphi_2(x,y)$ con normale esterna ed elemento d'area dati da:

$$\mathbf{n}_e = \frac{-\varphi_{2x}\mathbf{i} - \varphi_{2y}\mathbf{j} + \mathbf{k}}{\sqrt{1+|\nabla\varphi_2|^2}}, \qquad dS = \sqrt{1+|\nabla\varphi_2|^2}\,dxdy$$

e quindi, su Σ_2, si ha $\mathbf{n}_e \cdot \mathbf{k}\, dS = dxdy$.

3. Infine, Σ, la parte laterale dove $\mathbf{n}_e \cdot \mathbf{k} = 0$.

Deduciamo che:

$$\iint_{\partial D} F_3\, \mathbf{n}_e \cdot \mathbf{k}\, dS = \iint_{\Sigma_2} F_3\, \mathbf{n}_e \cdot \mathbf{k}\, dS + \iint_{\Sigma_1} F_3\, \mathbf{n}_e \cdot \mathbf{k}\, dS = \\ = \iint_B [F_3(x,y,\varphi_2(x,y)) - F_3(x,y,\varphi_1(x,y))]\, dxdy$$

e dalla (5.3) segue la tesi.

In modo analogo si può dimostrare che, se D è *semplice rispetto all'asse* y, allora:

$$\iiint_D (F_2)_y\, dxdydz = \iint_{\partial D} F_2\, \mathbf{n}_e \cdot \mathbf{j}\, dS, \tag{5.4}$$

mentre, se è *semplice rispetto all'asse* x, allora:

$$\iiint_D (F_1)_x\, dxdydz = \iint_{\partial D} F_1\, \mathbf{n}_e \cdot \mathbf{1}\, dS. \tag{5.5}$$

Sommando le (5.2), (5.4), (5.5) otteniamo il teorema della divergenza.

OSSERVAZIONE (SIGNIFICATO DELL'OPERATORE DIVERGENZA) Applichiamo il teorema della divergenza a una sfera $B_r(\mathrm{p}_0)$, con centro in un punto p_0 e raggio r

© 978-88-08-06485-1

arbitrariamente piccolo, e dividiamo ambo i membri per il volume della sfera (che indichiamo con $|B_r(\mathrm{p}_0)|$); abbiamo:

$$\frac{1}{|B_r(\mathrm{p}_0)|}\iiint_{\{B_r(\mathrm{p}_0)\}} \nabla\cdot\mathbf{F}\, dx\,dy\,dz = \frac{1}{|B_r(\mathrm{p}_0)|}\iint_{\partial B_r(\mathrm{p}_0)} \mathbf{F}\cdot\mathbf{n}_e\, dS.$$

Facciamo ora tendere a zero r. Il primo membro, per il teorema della media, tende a $\mathrm{div}\mathbf{F}(\mathrm{p}_0)$ e quindi:

$$\nabla\cdot\mathbf{F}(\mathrm{p}_0) = \lim_{r\to 0}\frac{1}{|B_r(\mathrm{p}_0)|}\iint_{\partial B_r(\mathrm{p}_0)} \mathbf{F}\cdot\mathbf{n}_e\, dS.$$

Questa formula indica che *la divergenza di un campo è la densità di flusso del campo uscente dal punto per unità di volume*. È questo il significato rigoroso delle affermazioni un po' vaghe fatte in precedenza sulle relazioni tra divergenza di un campo e "pozzi e sorgenti" del medesimo.

Il teorema della divergenza, in realtà, è valido per una classe di domini molto più vasta di quelli semplici rispetto a una terna di assi cartesiani. Con una tecnica che ormai dovrebbe essere familiare allo studente, si può estendere la validità del teorema a domini D limitati, che siano inoltre decomponibili in un numero finito di sottodomini semplici rispetto ai tre assi e la cui frontiera sia costituita dall'unione di un numero finito di superfici Σ_j regolari a pezzi, orientabili, chiuse e disgiunte. Un caso tipico è quello di una "scatola" D privata di N sferette come in figura 6.16.

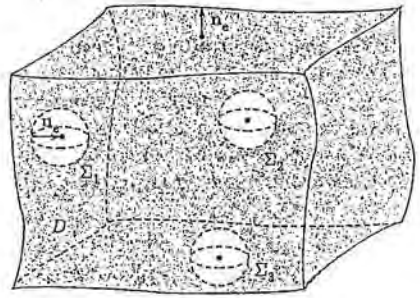


Figura 6.16. Dominio ammissibile per il teorema della divergenza.

L'equazione di Poisson per il potenziale elettrostatico

Come visto nell'esempio 4.2, il teorema di Gauss dell'elettrostatica afferma che il flusso del campo elettrico $\mathbf{E}$ uscente da una superficie chiusa $\Sigma = \partial D$ uguaglia $4\pi\kappa q$, dove q è la carica totale in D:

$$\iint_{\partial D} \mathbf{E}\cdot\mathbf{n}_e\, dS = 4\pi\kappa q_{\mathrm{tot}}.$$

Vogliamo trasformare questa uguaglianza tra due quantità "globali" (cioè che coinvolgono tutto D) in un'*uguaglianza puntuale* tra due funzioni. Otterremo un'equazione differenziale. Il ragionamento è tipico:

© 978-88-08-06485-1

1) Si applica l'identità precedente a una *generica* regione D.

2) Si riscrive la carica totale contenuta in D come l'integrale su D della densità di carica ρ:

$$q_{\text{tot}} = \iiint_D \rho \, dx\, dy\, dz$$

e quindi

$$\iint_{\partial D} \mathbf{E} \cdot \mathbf{n}_e \, dS = 4\pi\kappa \iiint_D \rho \, dx\, dy\, dz.$$

3) Per ottenere un'uguaglianza tra due integrali di volume, si trasforma il primo membro usando il teorema della divergenza; il risultato è:

$$\iiint_D \nabla \cdot \mathbf{E} \, dx\, dy\, dz = 4\pi\kappa \iiint_D \rho \, dx\, dy\, dz.$$

4) Poiché l'identità precedente si può riscrivere nella forma

$$\iiint_D [\nabla \cdot \mathbf{E} - 4\pi\kappa\rho] \, dx\, dy\, dz = 0$$

e questo deve valere per qualunque regione D (comunque piccola), l'unica possibilità è che l'integranda stessa sia zero, ossia valga l'identità:

$$\nabla \cdot \mathbf{E} = 4\pi\kappa\rho.$$

Questa è l'*equazione differenziale* che traduce, punto per punto, il contenuto del teorema di Gauss dell'elettrostatica. È una delle quattro fondamentali equazioni di Maxwell, che reggono la teoria dell'elettromagnetismo.

5) Ricordiamo ora che il campo elettrostatico è conservativo, cioè si può scrivere $\mathbf{E} = \nabla U$ dove U è il potenziale elettrostatico; perciò $\operatorname{div}\mathbf{E} = \operatorname{div}\nabla U = \Delta U$. Si ottiene quindi che *il potenziale elettrostatico soddisfa l'equazione seguente, detta di Poisson:*

$$\Delta U = 4\pi\kappa\rho$$

dove $4\pi k$ è una costante e ρ è la densità di carica.

Equazione di continuità

In questo paragrafo è significativo usare campi scalari e vettoriali *non stazionari*. Sia $\mathbf{v} = \mathbf{v}(t, x, y, z)$ la velocità di un fluido, $\rho = \rho(t, x, y, z)$ la sua densità e D un elemento di volume occupato dal fluido. Se dS è una piccola porzione di superficie contenuta in ∂D con normale $\mathbf{n}$, abbiamo già visto che la massa di fluido che attraversa quella sezione nell'unità di tempo sarà $\rho\mathbf{v} \cdot \mathbf{n}\, dS$. Integrando su ∂D, si ottiene il flusso netto uscente di $\rho\mathbf{v}$ attraverso ∂D. Per il teorema della divergenza, vale l'identità:

$$\iiint_D \nabla \cdot (\rho\mathbf{v}) \, dx\, dy\, dz = \iint_{\partial D} \rho\mathbf{v} \cdot \mathbf{n}_e \, dS$$

D'altra parte la quantità di fluido uscita da D nell'unità di tempo è anche uguale a $-\frac{dm}{dt}$ dove m è la massa totale del fluido contenuto in D (nell'ipotesi che, all'interno della regione D non ci sia "creazione" né "distruzione" di massa). La massa totale m si può calcolare come integrale in D della densità:

$$m = \iiint_D \rho \, dx\, dy\, dz.$$

Perciò, se si assume di poter derivare sotto il segno di integrale[12], si ha:

$$-\frac{dm}{dt} = -\iiint_D \frac{\partial\rho}{\partial t} dx\, dy\, dz.$$

Abbiamo dunque:

$$\iiint_D \left[\frac{\partial\rho}{\partial t} + \nabla \cdot (\rho\mathbf{v})\right] dx\, dy\, dz = 0$$

e, per l'arbitrarietà di D, si deduce

$$\frac{\partial\rho}{\partial t} + \nabla \cdot (\rho\mathbf{v}) = 0$$

che si chiama *equazione di continuità* ed esprime il principio di conservazione della massa.

Se interpretiamo ρ come densità di carica elettrica e $\mathbf{v}$ come velocità della carica, l'equazione di continuità si rilegge come principio di conservazione della carica elettrica, che è una legge dell'elettrodinamica.

Qualche altra conseguenza e applicazione del teorema della divergenza sarà discussa negli esercizi.

6 TEOREMA DEL ROTORE

Introduciamo ora un altro risultato fondamentale con profonde implicazioni.

TEOREMA 6.10 (DEL ROTORE, O DI STOKES) *Sia Σ una superficie regolare orientabile, orientata con il versore normale $\mathbf{n}$, dotata di bordo $\partial^+\Sigma$ orientato positivamente*[13]. *Supponiamo inoltre che $\partial^+\Sigma$ sia una curva regolare, o l'unione di più curve regolari, e sia T il versore tangente a $\partial^+\Sigma$. Se $\mathbf{F} = P\mathbf{i} + Q\mathbf{j} + R\mathbf{k}$ è un campo vettoriale regolare definito in un intorno di Σ, allora vale la formula:*

$$\iint_\Sigma (\nabla \times \mathbf{F}) \cdot \mathbf{n}\, dS = \oint_{\partial+\Sigma} \mathbf{F} \cdot \mathbf{T} \, ds \tag{6.1}$$

Detto in parole: *il flusso del rotore di un campo vettoriale attraverso Σ uguaglia la circuitazione del campo lungo il bordo della superficie stessa, se orientato positivamente.*

Si noti che per poter calcolare il rotore di $\mathbf{F}$ nei punti di Σ è necessario che il campo sia definito in una regione tridimensionale che contiene la superficie al suo interno. Il lavoro di un campo lungo un cammino orientato è stato definito nel paragrafo 1.3. Il

[12] Ciò è giustificato se la densità del fluido è una funzione regolare.

[13] Si veda il paragrafo 4.1 per le definizioni.

© 978-88-08-06485-1

 © 978-88-08-06485-1

teorema vale anche per superfici regolari a pezzi, il cui bordo sia costituito da curve regolari a pezzi; risultano dunque ammissibili anche superfici come un tronco di cono o di piramide.

Il teorema vale anche nel caso particolare di una superficie orientabile chiusa, quindi priva di bordo: in questo caso il secondo membro della (6.1) sarà nullo.

Omettiamo la dimostrazione del teorema. Ci limitiamo a osservare che, scrivendo

$$\mathbf{n} = \cos(\mathbf{n},\mathbf{i})\,\mathbf{i} + \cos(\mathbf{n},\mathbf{j})\,\mathbf{j} + \cos(\mathbf{n},\mathbf{k})\,\mathbf{k}$$

e osservando che

$$\mathbf{T}\,ds = dx\mathbf{i} + dy\mathbf{j} + dz\mathbf{k},$$

la (6.1) si può scrivere nella forma più esplicita seguente:

$$(6.2)\quad \iint_\Sigma \{(R_y - Q_z)\cos(\mathbf{n},\mathbf{i}) + (P_z - R_x)\cos(\mathbf{n},\mathbf{j}) + (Q_x - P_y)\cos(\mathbf{n},\mathbf{k})\}\,dS =$$

$$= \oint_{\partial^+\Sigma} P\,dx + Q\,dy + R\,dz.$$

Nel caso particolare in cui la superficie Σ coincide con un dominio D del piano xy (con frontiera regolare) ed $\mathbf{F}$ sia un campo vettoriale piano,

$$\mathbf{F}(x,y) = P(x,y)\,\mathbf{i} + Q(x,y)\,\mathbf{j},$$

si ha $\mathbf{n} = \mathbf{k}$, $(\nabla \times \mathbf{F})\cdot\mathbf{n} = (Q_x - P_y)$, e la (6.2) si riscrive:

$$\iint_D (Q_x - P_y)\,dx\,dy = \oint_{\partial^+ D} P\,dx + Q\,dy,$$

ossia ritroviamo il *teorema di Gauss-Green*, provato nel paragrafo 2, di cui quindi il teorema di Stokes risulta essere una generalizzazione.

Esempi

6.1 Calcoliamo il flusso del rotore di $\mathbf{F} = x\mathbf{i} + y\mathbf{k}$, attraverso la superficie Σ_R (quarto di sfera) orientata verso l'alto:

$$\begin{cases} x = R\sin\varphi\cos\theta \\ y = R\sin\varphi\sin\theta \\ z = R\cos\varphi \end{cases} \quad \text{con} \quad \varphi \in \left[0, \frac{\pi}{2}\right] \quad \theta \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right].$$

Utilizziamo sia il calcolo diretto che il calcolo, alternativo, mediante la circuitazione del campo sul bordo.

Abbiamo:

$$\nabla \times \mathbf{F} = \begin{vmatrix} \mathbf{i} & \mathbf{j} & \mathbf{k} \\ \partial_x & \partial_y & \partial_z \\ x & 0 & y \end{vmatrix} = \mathbf{i}$$

$$\mathbf{n} = \mathbf{i}\sin\varphi\cos\theta + \mathbf{j}\sin\varphi\sin\theta + \mathbf{k}\cos\varphi$$

$$\Phi(\mathbf{F},\Sigma_R) = \int_{-\pi/2}^{\pi/2} d\theta \int_0^{\pi/2} \sin\varphi\cos\theta R^2 \sin\varphi\,d\varphi =$$

$$= R^2 \int_{-\pi/2}^{\pi/2} \cos\theta\,d\theta \int_0^{\pi/2} \sin^2\varphi\,d\varphi = R^2 \cdot 2 \cdot \frac{\pi}{4} = \frac{\pi}{2}R^2.$$

© 978-88-08-06485-1

Calcoliamo, in alternativa, la circuitazione del campo lungo il bordo di Σ_R; il bordo è unione di due semicirconferenze percorse in verso opportuno. La prima ha equazioni parametriche

$$\begin{cases} x(t) = R\cos t \\ y(t) = R\sin t \\ z(t) = 0 \end{cases} \qquad t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

e dà il contributo:

$$L_1 = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} [R\cos t(-R\sin t) + (R\sin t)\cdot 0]dt = 0 \text{ (per simmetria)}.$$

La seconda è

$$\begin{cases} x(t) = 0 \\ y(t) = R\cos t \\ z(t) = R\sin t \end{cases} \qquad t \in \left[-\frac{\pi}{2}, \frac{\pi}{2}\right]$$

e dà il contributo:

$$L_2 = \int_{-\frac{\pi}{2}}^{\frac{\pi}{2}} [0 + R\cos t(R\cos t)]dt = \frac{\pi}{2}R^2.$$

I due risultati coincidono, come previsto, in base al teorema di Stokes.

6.2 Come applicazione del teorema di Stokes, dimostriamo che il campo $\mathbf{F}(\mathbf{x}) = \mathbf{x}/|\mathbf{x}|$, solenoidale in $\Omega = \mathbb{R}^3 \setminus \{0\}$ non può ammettere potenziale vettore in *tutto* Ω. Supponiamo per assurdo che esista $\mathbf{V} \in C^2(\Omega)$ tale che $\nabla \times \mathbf{V} = \mathbf{F}$ in Ω. Per il teorema di Stokes, se Σ è una superficie regolare, orientata, abbiamo

$$(6.3)\qquad \iint_\Sigma \mathbf{F}\cdot\mathbf{n}\,dS = \oint_{\partial^+\Sigma} \mathbf{V}\cdot\mathbf{T}\,ds.$$

Scegliamo come Σ_r la superficie sferica di raggio unitario con centro nell'origine, privata della calotta di raggio r e centro nel polo nord $(0,0,1)$, come mostrato in figura 6.17.

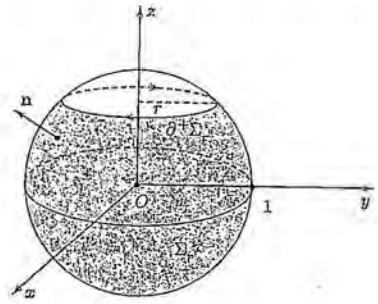


Figura 6.17.

Su Σ_r scegliamo $\mathbf{n} = \mathbf{n}_e = x\mathbf{i} + y\mathbf{j} + z\mathbf{k}$. Di conseguenza $\partial^+\Sigma_r$ è costituito dal bordo della calotta percorso in senso *orario*.

Essendo $|\mathbf{x}| = 1$ su Σ, si ha $\mathbf{F}\cdot\mathbf{n} = 1$ e pertanto

$$\iint_{\Sigma_r} \mathbf{F}\cdot\mathbf{n}\,dS = a(\Sigma_r)$$

 © 978-88-08-06485-1

e quindi, dalla (6.3) si ha:

$$a(\Sigma_r) = \oint_{\partial^+\Sigma_r} \mathrm{V}\cdot\mathrm{T}\,ds. \tag{6.4}$$

Facciamo tendere r a zero. Allora l'area della calotta tende a zero cosicché $a(\Sigma_r) \to 4\pi$.

D'altra parte, se M è il massimo di $|\mathrm{V}|$ sul bordo della sfera, possiamo scrivere:

$$\left|\oint_{\partial^+\Sigma_r} \mathrm{V}\cdot\mathrm{T}\,ds\right| \le M\left|\oint_{\partial^+\Sigma_r} ds\right| = M\,l(\partial\Sigma_r).$$

Ma se $r \to 0$, $l(\partial\Sigma_r) \to 0$ e si ottiene così una contraddizione con la (6.4).

La *legge di Neumann dell'elettromagnetismo* afferma che un campo magnetico B, variabile nel tempo, induce un campo elettrico E, in base alla legge:

$$-\frac{d\Phi_{\mathrm{B}}}{dt} = \oint_C \mathrm{E}\cdot\mathrm{T}\,ds$$

dove Φ_{B} è il flusso di B attraverso una superficie Σ orientata e $C = \partial^+\Sigma$. Così formulata, la legge coinvolge grandezze globali (il flusso attraverso un'intera superficie, il lavoro del campo elettrico lungo un cammino). Si vuole dare una formulazione alternativa di questa legge, che sia vera punto per punto. Per il teorema di Stokes, il secondo membro uguaglia

$$\iint_\Sigma (\nabla\times\mathrm{E})\cdot\mathrm{n}\,dS.$$

Il primo membro si può invece riscrivere come

$$-\frac{d}{dt}\iint_\Sigma \mathrm{B}\cdot\mathrm{n}\,dS = -\iint_\Sigma \frac{\partial\mathrm{B}}{\partial t}\cdot\mathrm{n}\,dS.$$

Dall'uguaglianza

$$-\iint_\Sigma \frac{\partial\mathrm{B}}{\partial t}\cdot\mathrm{n}\,dS = \iint_\Sigma (\nabla\times\mathrm{E})\cdot\mathrm{n}\,dS$$

che deve valere per ogni superficie Σ, segue l'uguaglianza delle funzioni integrande:

$$-\frac{\partial\mathrm{B}}{\partial t} = \nabla\times\mathrm{E}.$$

Questa è una delle quattro equazioni di Maxwell dell'elettromagnetismo; è un'equazione differenziale che lega, punto per punto, i campi vettoriali B, E. Il teorema di Stokes è servito per dedurla a partire dalla legge di Neumann.

OSSERVAZIONE (SIGNIFICATO DELL'OPERATORE ROTORE) Fissiamo un punto p_0 nello spazio, un versore n spiccato da p_0. Sia poi $D_r(\mathrm{p}_0)$ un disco di centro p_0, raggio r e versore normale n. Applichiamo a $D_r(\mathrm{p}_0)$ la (6.1) e dividiamo per l'area del disco. Otteniamo:

$$\frac{1}{\pi r^2}\iint_{D_r(\mathrm{p}_0)} (\nabla\times\mathrm{F})\cdot\mathrm{n}\,dS = \frac{1}{\pi r^2}\oint_{\partial^+ D_r(\mathrm{p}_0)} \mathrm{F}\cdot\mathrm{T}\,ds$$

Facendo tendere r a zero, il primo membro tende, per il teorema della media, a $[(\nabla\times\mathrm{F})\cdot\mathrm{n}](\mathrm{p}_0)$. Otteniamo quindi:

$$[(\nabla\times\mathrm{F})\cdot\mathrm{n}](\mathrm{p}_0) = \lim_{r\to 0}\frac{1}{\pi r^2}\oint_{\partial^+ D_r(\mathrm{p}_0)} \mathrm{F}\cdot\mathrm{T}\,ds$$

© 978-88-08-06485-1

ossia: la componente del rotore di un campo rispetto al versore n in un punto p di una superficie è uguale alla densità superficiale di circuitazione del campo attorno a p, nel piano perpendicolare a n.

Campi irrotazionali

Se un campo F è irrotazionale in una certa regione dello spazio, il flusso del rotore di F lungo una qualsiasi superficie Σ regolare e con bordo è nullo. Ritroviamo così il risultato che la circuitazione di F lungo $\partial\Sigma$ è nulla.

Il teorema di Stokes implica qualcosa in più. Se, per esempio, F è un campo irrotazionale, ma non conservativo, nel piano privato dell'origine e Σ è una corona circolare centrata nell'origine, il teorema implica che il lavoro di F lungo il bordo orientato della corona circolare è nullo. Questo significa che il lavoro di F sulla circonferenza interna è uguale a quello sulla circonferenza esterna (se sono percorse nello stesso verso).

ESERCIZI

Sia

$$\mathrm{E} = kq\frac{\mathrm{r}}{|\mathrm{r}|^3}$$

con $\mathrm{r} = (x, y, z)$, k, q costanti; sia D_R il disco di centro $(0, a, 0)$ e raggio R, disposto perpendicolarmente all'asse y. Calcolare il flusso di E attraverso D_R, orientato nel verso delle y crescenti.

Suggerimenti: parametrizzare D_R con le equazioni

$$\begin{cases} x = \rho\cos\theta \\ y = a \\ z = \rho\sin\theta \end{cases} \qquad \rho\in[0,R] \quad \theta\in[0,2\pi)$$

e osservare che $|\mathrm{r}| = \sqrt{\rho^2 + a^2}$.

Sia Σ la superficie (chiusa) del cilindro V descritto da $x^2 + y^2 \le R$ e $0 \le z \le h$. Calcolare il flusso uscente da V del campo vettoriale

$$\mathrm{v} = x\mathrm{i} - y\mathrm{j} + z\mathrm{k}$$

a) direttamente dalla definizione di flusso;
b) utilizzando il teorema della divergenza.

Sia Σ la superficie della semisfera $x^2 + y^2 + z^2 = R^2$, $z \ge 0$, orientata verso l'alto, e sia

$$\mathrm{F} = y\mathrm{i} + z\mathrm{j} - x\mathrm{k}.$$

Calcolare il flusso di $\mathrm{rot}\,\mathrm{F}$ attraverso Σ:
a) direttamente dalla definizione;
b) utilizzando il teorema del rotore.

Sia

$$\mathrm{F} = \frac{-y}{x^2+y^2}\mathrm{i} + \frac{x}{x^2+y^2}\mathrm{j}.$$

 © 978-88-08-06485-1

Utilizzando il teorema di Stokes, mostrare che il lavoro del campo $\mathbf{F}$ lungo l'ellisse con centro nell'origine e semiassi 3 e 4 e lungo la circonferenza unitaria con centro nell'origine, percorse entrambe in senso antiorario, è lo stesso. Calcolare poi tale lavoro.

Il campo $\mathbf{B}$ ha la proprietà che il suo flusso attraverso una qualsiasi superficie chiusa è nullo. Quanto vale $\nabla \cdot \mathbf{B}$? (Questa proprietà è vera per il campo magnetico).

Quanto vale il flusso di $\nabla \times \mathbf{F}$ attraverso una superficie chiusa? Rispondere riscrivendo il flusso del rotore:

a) mediante il teorema del rotore;

b) mediante il teorema della divergenza (e ricordando un'opportuna identità).

Questo esercizio richiede di imitare il metodo seguito nell'esempio 6.3, per ricavare ora una formulazione differenziale (equazione di Maxwell) della *legge di Ampère generalizzata*. Siano: $\mathbf{B}$ il campo magnetico, $\mathbf{E}$ il campo elettrico, $\mathbf{j}$ la densità di flusso di corrente elettrica, I l'intensità di corrente elettrica. Il punto di partenza è l'identità:

$$\oint_{\partial+\Sigma} \mathbf{B}\cdot\mathbf{T}\,ds = \mu_0 I + \mu_0\varepsilon_0 \frac{d\Phi(\mathbf{E},\Sigma)}{dt}$$

dove: μ_0, ε_0 sono costanti; Σ è una generica superficie regolare con bordo, orientata con il versore normale $\mathbf{n}$ e infine

$$I = \Phi(\mathbf{j},\Sigma).$$

Utilizzando tutti questi fatti e il teorema di Stokes, ricavare dall'equazione di partenza un'equazione differenziale (senza più integrali) che coinvolga $\mathbf{E}, \mathbf{B}, \mathbf{j}$. Tutti questi sono campi vettoriali variabili nel tempo.

(Questo esercizio richiede di utilizzare le tecniche del paragrafo 5). Sia $u = (t, x, y, z)$ la temperatura di un corpo D che non possiede all'interno sorgenti o pozzi di calore, $\mathbf{j}$ la densità di flusso di calore, $\rho = \rho(x, y, z)$ la densità del corpo. L'equazione di bilancio dell'energia fornisce:

$$\iint_{\Sigma} \mathbf{j}\cdot\mathbf{n}_e\,dS = -\iiint_D c_v\rho\frac{\partial u}{\partial t}dx\,dy\,dz$$

dove c_v è il calore specifico a volume costante. La legge di Fourier della conduzione del calore afferma che

$$\mathbf{j} = -k\nabla u$$

con $k > 0$, costante di diffusività. Utilizzando questi fatti, il teorema della divergenza e il fatto che la relazione precedente debba valere per qualsiasi regione D dello spazio, dedurne un'equazione differenziale a cui è soggetta la temperatura u (equazione di diffusione del calore).

Siano f, g due funzioni reali di tre variabili; applicando il teorema della divergenza al campo $f\nabla g$ su un dominio regolare D, dimostrare che:

$$\iiint_D f\Delta g\,dx\,dy\,dz + \iiint_D \nabla f\cdot\nabla g\,dx\,dy\,dz = \iint_{\partial D} f\,\nabla g\cdot\mathbf{n}_e\,dS.$$

Dedurre poi da questa identità il seguente risultato: se f è soluzione del problema

$$\begin{cases}\Delta f = 0 & \text{in } D\\ f = 0 & \text{su } \partial D,\end{cases}$$

allora f è identicamente nulla in D.

© 978-88-08-06485-1

Siano f, g due funzioni reali di tre variabili; applicando il teorema della divergenza al campo $fg\mathbf{i}$ su un dominio regolare D, dimostrare la seguente *formula di integrazione per parti in più dimensioni*:

$$\iiint_D f\frac{\partial g}{\partial x}\,dx\,dy\,dz = \iint_{\partial D} fg\,n_x\,dS - \iiint_D g\frac{\partial f}{\partial x}\,dx\,dy\,dz$$

dove n_x è la prima componente del versore $\mathbf{n}_e$. Formule analoghe valgono per le derivate rispetto a y, z.

Il nastro di Moebius può essere parametrizzato dalle equazioni:

$$\begin{cases} x(t,\theta) = \left(r + t\sin\dfrac{\theta}{2}\right)\cos\theta \\ y(t,\theta) = \left(r + t\sin\dfrac{\theta}{2}\right)\sin\theta \qquad t\in[-l,l] \quad \theta\in[0,2\pi) \\ z(t,\theta) = t\cos\dfrac{\theta}{2} \end{cases}$$

dove r, l sono fissati, con $r > l$.

Assumiamo che $r = 3, l = 1$. Calcolare il vettore normale $\mathbf{r}_t \wedge \mathbf{r}_\theta$ per $t = 0$ e θ che varia tra 0 e 2π , ossia lungo la circonferenza equatoriale di raggio r sul piano $z = 0$. Constatare che il vettore normale non ritorna, dopo un giro completo, al valore iniziale. Dedurre che la superficie non è orientabile.

Si consideri il campo vettoriale tridimensionale:

$$\mathbf{F} = \mathbf{i}(z - y) + \mathbf{j}x(1 + z^2) + \mathbf{k}xy.$$

a) Calcolare la circuitazione di $\mathbf{F}$ lungo l'ellisse che giace nel piano $z = 0$ e ha equazione

$$\frac{x^2}{4} + \frac{y^2}{9} = 1$$

percorsa in senso antiorario.

b) Calcolare $\nabla \times \mathbf{F}$.

c) Scrivere esplicitamente l'integrale doppio che assegna il flusso di $\nabla \times \mathbf{F}$ attraverso la superficie

$$z = 1 - \left(\frac{x^2}{4} + \frac{y^2}{9}\right) \text{ per } \frac{x^2}{4} + \frac{y^2}{9} < 1$$

orientata con la normale verso l'alto. (Non si richiede di calcolare l'integrale doppio, ma solo di scriverlo, esplicitamente).

d) Si può dire quanto vale il flusso di cui al punto precedente, senza calcolare l'integrale doppio corrispondente?

7 Serie di potenze e serie di Fourier

1 SERIE DI FUNZIONI E CONVERGENZA TOTALE

Supponiamo che su un certo intervallo $I \subset \mathbb{R}$ siano definite le funzioni

$$f_n : I \to \mathbb{R} \quad \text{per} \quad n = 1, 2, 3, \ldots$$

Per ogni $x \in I$ si può considerare la successione di numeri reali $\{f_n(x)\}_{n=1}^{\infty}$; a partire da questa successione si può poi considerare la serie

$$\sum_{n=1}^{\infty} f_n(x)$$

che, per ciascun $x \in I$ fissato, può essere una serie convergente, divergente o irregolare. Esprimeremo questo fatto dicendo che la *serie di funzioni* $\sum_{n=1}^{\infty} f_n$ converge, diverge o è irregolare in un certo punto $x \in I$. Particolarmente interessante è il caso in cui la serie converge in tutti i punti di un intervallo $I^* \subseteq I$; in tal caso la somma della serie è una nuova funzione f, definita in I^*. Diremo allora che la serie di funzioni *converge puntualmente* in I^*.

Abbiamo già incontrato alcuni esempi di serie di funzioni:
– la *serie geometrica*,

$$\sum_{n=0}^{\infty} x^n = \frac{1}{1-x} \quad \text{per } -1 < x < 1 \tag{1.1}$$

– la *serie esponenziale*,

$$\sum_{n=0}^{\infty} \frac{x^n}{n!} = e^x \text{ per ogni } x \in \mathbb{R},$$

e altre serie di Taylor di funzioni derivabili infinite volte.[1] Si tratta in tutti i casi di serie della forma

$$\sum_{n=0}^{\infty} a_n x^n,$$

in cui le funzioni $f_n(x)$ hanno la forma $a_n x^n$ e perciò si chiamano *serie di potenze*.

[1] v. volume 1, capitolo 5, paragrafo 2.1.

 © 978-88-08-06485-1

In questo capitolo introdurremo le nozioni fondamentali su queste serie e studieremo un'altra classe importante di serie di funzioni: le *serie di Fourier*. Prima però conviene dare qualche risultato di carattere generale sulle serie di funzioni.

Un problema tipico che si pone per le serie di funzioni è: quali sono le proprietà che la somma f della serie "eredita" dai termini f_n? Per esempio:

Se le f_n sono continue, anche f lo è?

Se le f_n sono derivabili, anche f lo è? E si può calcolare f' "derivando la serie termine a termine", cioè estendendo alla "somma infinita" la linearità della derivata? Questo significa scambiare le operazioni di somma e di derivazione:

$$\left[\sum_{n=1}^{\infty} f_n(x)\right]' = \sum_{n=1}^{\infty} f_n'(x).$$

Analogamente, si può calcolare l'integrale di $f(x)$ "integrando termine a termine", cioè estendendo alla "somma infinita" la linearità dell'integrale? Questo significa scambiare le operazioni di somma e di integrazione:

$$\int_a^b \left(\sum_{n=1}^{\infty} f_n(x)\right) dx = \sum_{n=1}^{\infty} \left(\int_a^b f_n(x)dx\right).$$

Non ci si può in generale aspettare che la convergenza della serie $\sum_{n=1}^{\infty} f_n(x)$, senza altre ipotesi, garantisca la risposta affermativa alle domande precedenti. Illustriamo questa problematica mediante alcuni esempi.

ESEMPI

Derivabilità termine a termine della serie geometrica. Consideriamo la serie geometrica (1.1). Sappiamo che per la sua somma parziale n-esima vale l'identità elementare:

$$\sum_{k=0}^{n} x^k = \frac{x^{n+1}-1}{x-1}.$$

Derivando ambo i membri di quest'uguaglianza (il primo membro ora è una somma finita) otteniamo:

$$\sum_{k=1}^{n} kx^{k-1} = \frac{(n+1)x^n(x-1)-(x^{n+1}-1)}{(x-1)^2} = \frac{nx^{n+1}-(n+1)x^n+1}{(x-1)^2},$$

che per $|x|<1$ e $n\to\infty$ dà, essendo $nx^{n+1}\to 0$ e $(n+1)x^n\to 0$

$$\sum_{k=1}^{\infty} kx^{k-1} = \frac{1}{(x-1)^2}.$$

Poiché d'altro canto la somma trovata $\frac{1}{(x-1)^2}$ non è altro che la derivata di $\frac{1}{1-x}$, ne deduciamo che in questo caso particolare è vero che

$$\left(\sum_{k=0}^{\infty} x^k\right)' = \sum_{k=1}^{\infty} kx^{k-1}$$

ossia la serie geometrica si può derivare termine a termine, per $|x|<1$.

© 978-88-08-06485-1

Si noti che il ragionamento di questo esempio non si può ripetere in generale, in quanto si basa sulla conoscenza di un'espressione semplice ed esplicita per la somma parziale n-esima della serie.

Una serie non derivabile termine a termine. Consideriamo la serie

$$\sum_{n=1}^{\infty} \frac{\sin(3^n x)}{2^n}.$$

Si vede subito che la serie converge in ogni punto x, perché converge assolutamente:

$$\left|\frac{\sin(3^n x)}{2^n}\right| \le \frac{1}{2^n}.$$

Osserviamo che ciascuna funzione $f_n(x) = \frac{\sin(3^n x)}{2^n}$ è infinitamente derivabile in tutto $\mathbb{R}$. Tuttavia la serie delle derivate,

$$\sum_{n=1}^{\infty} \frac{3^n \cos(3^n x)}{2^n},$$

non è convergente. Per esempio, in $x=0$ dà $\sum_{n=1}^{\infty}\left(\frac{3}{2}\right)^n$, che diverge. Quindi, questa particolare serie non è certamente derivabile termine a termine. In realtà, non sappiamo nemmeno dire, a questo punto del discorso, se la somma della serie sia una funzione derivabile o meno[2].

Una serie non integrabile termine a termine. Sia $\{x_n\}_{n=1}^{\infty}$ una qualsiasi enumerazione dei numeri razionali[3] dell'intervallo $[0,1]$ e definiamo

$$f_n : [0,1] \to \mathbb{R}$$
$$f_n(x) = \begin{cases} 1 & \text{se } x = x_n \\ 0 & \text{altrimenti.} \end{cases}$$

Notiamo che ogni funzione f_n è diversa da zero in un solo punto di $[0,1]$, quindi è continua salvo un punto, ed è integrabile, con integrale zero. La serie

$$\sum_{n=1}^{\infty} f_n(x)$$

evidentemente converge alla *funzione di Dirichlet*[4]

$$f(x) = \begin{cases} 1 & \text{se } x \in \mathbb{Q} \\ 0 & \text{altrimenti.} \end{cases}$$

[2] Una funzione di questo tipo in effetti fu utilizzata da Weierstrass nel 1872 per esibire un esempio di funzione continua su un intervallo ma *in nessun punto* derivabile. L'esempio di Weierstrass è il seguente:

$$\sum_{n=1}^{\infty} \frac{\cos(3^n \pi x)}{2^n}.$$

[3] Per esempio, possiamo disporre in successione i razionali di $[0,1]$ in questo modo:

$$0, 1, \frac{1}{2}, \frac{1}{3}, \frac{2}{3}, \frac{1}{4}, \frac{3}{4}, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, \frac{1}{6}, \dots$$

ordinandoli cioè secondo l'ordine crescente del denominatore, quindi del numeratore, ed evitando di ripetere frazioni equivalenti a quelle già scritte.

[4] già incontrata nel volume 1, capitolo 6, paragrafo 2.2.

 © 978-88-08-06485-1

Questa funzione f è discontinua in tutti i punti di $[0,1]$ e non è integrabile! In particolare, l'uguaglianza

$$\sum_{n=1}^{\infty}\int_0^1 f_n(x)\,dx = \int_0^1 \left(\sum_{n=1}^{\infty} f_n(x)\right) dx \text{ non vale,}$$

in quanto il primo membro vale zero (ogni integrale è nullo) mentre l'integrale a secondo membro non esiste.

Esiste però un modo abbastanza semplice per "rafforzare" la convergenza in modo che tutto funzioni bene. Lo esprimiamo attraverso la definizione di *convergenza totale*.

DEFINIZIONE 7.1 Sia $I \subseteq \mathbb{R}$ un intervallo e siano $f_n : I \to \mathbb{R}$, $n = 1, 2, 3, \ldots$ Diremo che la serie $\sum_{n=1}^{\infty} f_n$ *converge totalmente* in I se esiste una successione $\{a_n\}_{n=1}^{\infty}$ di numeri reali positivi tali che:

i) $|f_n(x)| \le a_n$ per ogni $x \in I$, $n \in \mathbb{N}$;

ii) $\sum_{n=1}^{\infty} a_n$ converge.

Osserviamo che la convergenza totale della serie di funzioni $\sum_{n=1}^{\infty} f_n$ in I implica la convergenza assoluta, e quindi la convergenza semplice[5], della serie numerica $\sum_{n=1}^{\infty} f_n(x)$ per ogni $x \in I$. Pertanto risulta ben definita la funzione, somma della serie,

$$f : I \to \mathbb{R}$$
$$f(x) = \sum_{n=1}^{\infty} f_n(x).$$

La convergenza totale è un'ipotesi piuttosto forte. Rivediamo da questo punto di vista i precedenti esempi.

Esempio

Il lettore verifichi le seguenti affermazioni.

(a) La serie geometrica (v. Esempio 1.1) converge totalmente in $[-\delta, \delta]$ per ogni $\delta < 1$, ma non converge totalmente in $(-1, 1)$.

(b) La serie dell'Esempio 1.2 converge totalmente, mentre la serie delle derivate non è neppure convergente.

(c) La serie dell'Esempio 1.3 non converge totalmente.

Vediamo alcuni teoremi generali che mostrano le buone proprietà che si possono dedurre dall'ipotesi di convergenza totale.

[5] Ricordiamo che, per una serie numerica con termini a segno variabile, la convergenza assoluta implica la convergenza semplice (v. volume 1, capitolo 5, Teorema 5.3).

© 978-88-08-06485-1

TEOREMA 7.1 (CONTINUITÀ DELLA SOMMA) *Sia $I \subseteq \mathbb{R}$ un intervallo, siano $f_n : I \to \mathbb{R}$ continue in $x_0 \in I$ per $n = 1, 2, 3, \ldots$ e supponiamo che la serie $\sum_{n=1}^{\infty} f_n$ converga totalmente in I. Allora la somma della serie, $f(x) = \sum_{n=1}^{\infty} f_n(x)$, è una funzione continua in x_0. In particolare, se le f_n sono continue in I, anche f è continua in I.*

DIMOSTRAZIONE. Sia $x_0 \in I$ e proviamo che $f(x) \to f(x_0)$ per $x \to x_0$, ossia che, per ogni $\varepsilon > 0$ esiste $\delta > 0$ tale che

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| < \varepsilon.$$

Scriviamo perciò:

$$f(x) - f(x_0) = \sum_{n=1}^{\infty} f_n(x) - \sum_{n=1}^{\infty} f_n(x_0) =$$

poiché le due serie sono convergenti

$$= \sum_{n=1}^{\infty} [f_n(x) - f_n(x_0)] = \sum_{n=1}^{N} [f_n(x) - f_n(x_0)] + \sum_{n=N+1}^{\infty} [f_n(x) - f_n(x_0)]$$
$$\equiv A_N(x) + B_N(x)$$

per un certo intero N che fisseremo in seguito. L'ipotesi di convergenza totale implica:

$$|B_N(x)| \le \sum_{n=N+1}^{\infty} |f_n(x) - f_n(x_0)| \le \sum_{n=N+1}^{\infty} (|f_n(x)| + |f_n(x_0)|) \le 2 \sum_{n=N+1}^{\infty} a_n$$

(con a_n come nella definizione 7.1). Ricordiamo ora che *il resto di una serie convergente è infinitesimo*[6]. Questo significa che, fissato $\varepsilon > 0$, esiste un N_0 tale che:

$$\sum_{n=N_0+1}^{\infty} a_n < \varepsilon \text{ e quindi } |B_{N_0}(x)| < 2\varepsilon.$$

Fissiamo ora $N = N_0$. Allora:

$$|A_{N_0}(x)| \le \sum_{n=1}^{N_0} |f_n(x) - f_n(x_0)|.$$

D'altro canto, ciascuna delle funzioni f_n per $n = 1, 2, \ldots, N_0$ per ipotesi è continua in x_0, perciò $A_{N_0}(x) \to 0$ per $x \to x_0$. In particolare esiste $\delta > 0$ tale che:

$$|x - x_0| < \delta \Rightarrow |A_{N_0}(x)| < \varepsilon.$$

In conclusione, abbiamo trovato $\delta > 0$ tale che:

$$|x - x_0| < \delta \Rightarrow |f(x) - f(x_0)| \le |A_{N_0}(x)| + |B_{N_0}(x)| < 3\varepsilon$$

e la tesi è dimostrata. ∎

[6] v. volume 1, capitolo 5, Teorema 5.2.

 © 978-88-08-06485-1

TEOREMA 7.2 (DERIVABILITÀ TERMINE A TERMINE) *Sia $I \subseteq \mathbb{R}$ un intervallo, siano $f_n : I \to \mathbb{R}$ derivabili in I per $n = 1, 2, 3, \ldots$ e supponiamo che:*

(i) *la serie $\sum_{n=1}^{\infty} f_n(x)$ converga per ogni $x \in I$;*

(ii) *la serie $\sum_{n=1}^{\infty} f_n'(x)$ converga totalmente in I:*

$$\sum_{n=1}^{\infty} |f_n'(x)| \leq \sum_{n=1}^{\infty} a_n < \infty \text{ per ogni } x \in I.$$

Allora, detta $f(x)$ la somma della serie $\sum_{n=1}^{\infty} f_n(x)$, si ha che f è derivabile in I e

$$f'(x) = \left(\sum_{n=1}^{\infty} f_n(x)\right)' = \sum_{n=1}^{\infty} f_n'(x) \text{ per ogni } x \in I.$$

DIMOSTRAZIONE. Sia $x \in I$ e mostriamo che

$$\lim_{h \to 0} \frac{f(x+h) - f(x)}{h} = \sum_{n=1}^{\infty} f_n'(x).$$

Consideriamo quindi:

$$\frac{f(x+h) - f(x)}{h} - \sum_{n=1}^{\infty} f_n'(x) =$$

poiché le serie convergono

$$= \sum_{n=1}^{\infty} \left[\frac{f_n(x+h) - f_n(x)}{h} - f_n'(x)\right] =$$

$$= \sum_{n=1}^{N} \left[\frac{f_n(x+h) - f_n(x)}{h} - f_n'(x)\right] + \sum_{n=N+1}^{\infty} \left[\frac{f_n(x+h) - f_n(x)}{h} - f_n'(x)\right] \equiv$$

$$\equiv A_N + B_N$$

per un certo intero N da scegliersi in seguito[7]. Consideriamo prima B_N. Applicando il teorema di Lagrange a ogni addendo della serie, otteniamo che, per ogni n, esiste un numero $\theta_n \in (0,1)$ per cui si ha:

$$B_N = \sum_{n=N+1}^{\infty} \left[f_n'(x + \theta_n h) - f_n'(x)\right].$$

Sfruttando ora l'ipotesi di convergenza totale della serie delle derivate, scriviamo:

$$|B_N| \leq \sum_{n=N+1}^{\infty} |f_n'(x + \theta_n h) - f_n'(x)| \leq 2 \sum_{n=N+1}^{\infty} a_n.$$

[7] In realtà A_N e B_N dipendono anche da x e h, ma per non appesantire la scrittura eviteremo di indicarle con $A_N(x,h)$, $B_N(x,h)$.

© 978-88-08-06485-1

Poiché il resto di una serie convergente è infinitesimo, fissato $\varepsilon > 0$ esiste un N_0 tale che:

$$|B_{N_0}| \leq 2 \sum_{n=N_0+1}^{\infty} a_n < \varepsilon.$$

Fissiamo $N = N_0$ e occupiamoci ora di A_{N_0}. Poiché f_n è derivabile in x, possiamo scrivere:

$$f_n(x+h) - f_n(x) = f_n'(x)\, h + \omega_n(h)$$

con $\frac{\omega_n(h)}{h} \to 0$ per $h \to 0$. (Abbiamo applicato il teorema di linearizzazione delle funzioni derivabili[8]). Sostituendo questa identità a ogni addendo di A_{N_0}, abbiamo:

$$A_{N_0} = \sum_{n=1}^{N_0} \left[\frac{\omega_n(h)}{h}\right].$$

Dunque A_{N_0} è somma di un numero finito di funzioni che tendono a zero per $h \to 0$, pertanto, fissato $\varepsilon > 0$, esiste $\delta > 0$ tale che

$$|A_{N_0}| < \varepsilon \text{ per } |h| < \delta.$$

Abbiamo quindi provato che, per ogni $\varepsilon > 0$, esiste $\delta > 0$ tale che:

$$\left|\frac{f(x+h) - f(x)}{h} - \sum_{n=1}^{\infty} f_n'(x)\right| \leq |A_{N_0}| + |B_{N_0}| < 2\varepsilon \text{ se } |h| < \delta,$$

che è la tesi. ∎

TEOREMA 7.3 (INTEGRABILITÀ TERMINE A TERMINE) *Siano $f_n : [a,b] \to \mathbb{R}$ continue per $n = 1, 2, 3, \ldots$ e supponiamo che la serie $\sum_{n=1}^{\infty} f_n$ converga totalmente in $[a,b]$ a una funzione f (continua, per il teorema 7.1). Allora la serie è integrabile termine a termine, cioè:*

$$\int_a^b f(x)\,dx = \int_a^b \left(\sum_{n=1}^{\infty} f_n(x)\right) dx = \sum_{n=1}^{\infty} \left(\int_a^b f_n(x)\,dx\right). \tag{1.2}$$

DIMOSTRAZIONE. Provare la (1.2) equivale a mostrare che:

$$A_n \equiv \int_a^b f(x)\,dx - \sum_{k=1}^{n} \left(\int_a^b f_k(x)\,dx\right) \to 0 \text{ per } n \to \infty. \tag{1.3}$$

Poiché l'integrale è lineare su somme finite, possiamo scrivere

$$A_n = \int_a^b \left[f(x) - \sum_{k=1}^{n} f_k(x)\right] dx = \int_a^b \left[\sum_{k=n+1}^{\infty} f_k(x)\right] dx.$$

Ora, dall'ipotesi di convergenza totale abbiamo:

$$\left|\sum_{k=n+1}^{\infty} f_k(x)\right| \leq \sum_{k=n+1}^{\infty} |f_k(x)| \leq \sum_{k=n+1}^{\infty} a_k$$

[8] v. volume 1, capitolo 4, paragrafo 7.1.

(con a_k come nella definizione 7.1). Integrando, otteniamo:

$$|A_n| \le \int_a^b \left| \sum_{k=n+1}^{\infty} f_k(x) \right| dx \le (b-a) \sum_{k=n+1}^{\infty} a_k ,$$

che tende a zero per $n \to \infty$ (perché è il resto di una serie convergente); abbiamo quindi provato la (1.3). ∎

2 SERIE DI POTENZE

2.1 Proprietà fondamentali delle serie di potenze

DEFINIZIONE 7.2 Si dice *serie di potenze di centro* $x_0 \in \mathbb{R}$ una serie di funzioni del tipo:

$$\sum_{n=0}^{\infty} a_n (x-x_0)^n \tag{2.1}$$

dove $\{a_n\}_{n=0}^{\infty}$ è una successione a valori reali e x è variabile reale. Gli a_n si dicono *coefficienti della serie di potenze.*

Per le serie di potenze, l'insieme degli x per cui si ha convergenza è sempre un intervallo, detto intervallo di convergenza; la metà della lunghezza di questo intervallo si dice raggio di convergenza. Precisamente:

TEOREMA 7.4 (RAGGIO DI CONVERGENZA) *Consideriamo la serie* (2.1) *e supponiamo che esista il limite*

$$l = \lim_{n\to\infty} \sqrt[n]{|a_n|}. \quad \textit{Poniamo} \quad R = \begin{cases} 1/l & \text{se } l \neq 0, \infty \\ +\infty & \text{se } l = 0 \\ 0 & \text{se } l = \infty. \end{cases}$$

Allora:
se $|x-x_0| < R$, *la serie* (2.1) *converge (assolutamente);*
se $|x-x_0| > R$, *la serie* (2.1) *non converge.*

OSSERVAZIONE Notiamo che nel caso $|x-x_0| > R$ la serie, a seconda dei casi, può sia divergere che essere irregolare.

Se $R = 0$, la serie converge solo per $x = x_0$; se $R = \infty$, la serie converge per ogni x reale.

Per $x = x_0 \pm R$ (gli estremi dell'intervallo di convergenza), il teorema non afferma nulla e in effetti la serie può convergere, divergere o essere irregolare, a seconda dei casi; ci occuperemo successivamente di questo problema.

DIMOSTRAZIONE. Sia $|x-x_0| < R$ e mostriamo che la serie a termini positivi $\sum_n |a_n (x-x_0)^n|$ converge, sfruttando il criterio della radice per le serie a termini non negativi[9]:

$$\lim_{n\to\infty} \sqrt[n]{|a_n (x-x_0)^n|} = |x-x_0| \cdot \lim_{n\to\infty} \sqrt[n]{|a_n|} = \frac{|x-x_0|}{R} < 1,$$

perciò la serie (2.1) converge (assolutamente).

[9] v. volume 1, capitolo 5, paragrafo 1.2.

Sia ora $|x-x_0| > R$. Ragionando come sopra, si ottiene che:

$$\lim_{n\to\infty} \sqrt[n]{|a_n (x-x_0)^n|} = \frac{|x-x_0|}{R} > 1.$$

Per quanto osservato nella dimostrazione del criterio della radice, questo implica che $|a_n (x-x_0)^n| \to \infty$; dunque anche $a_n (x-x_0)^n$ non tende a zero e la serie $\sum a_n (x-x_0)^n$ non converge. ∎

Il teorema precedente suggerisce la seguente:

DEFINIZIONE 7.3 Data una serie di potenze (2.1), si dice *raggio di convergenza della serie* un numero $R \in [0, \infty]$ con le proprietà espresse dal teorema precedente, ossia tale che, se $|x-x_0| < R$, la serie (2.1) converge e, se $|x-x_0| > R$, la serie (2.1) non converge.

Si noti che il limite $\lim_{n\to\infty} \sqrt[n]{|a_n|}$ potrebbe non esistere, perciò dal teorema precedente non segue automaticamente che esista sempre un numero con queste proprietà; tuttavia, si può dimostrare[10] che è proprio così: *per ogni serie di potenze esiste uno (e un solo) raggio di convergenza* $R \in [0, \infty]$.

Per il calcolo del raggio di convergenza di una serie di potenze si può utilizzare anche il seguente criterio, alternativo al precedente:

TEOREMA 7.5 (CRITERIO DEL RAPPORTO PER LE SERIE DI POTENZE) *Consideriamo la serie* (2.1) *e supponiamo che esista, finito o infinito, il limite*

$$\lim_{n\to\infty} \left| \frac{a_{n+1}}{a_n} \right| = l.$$

Allora il raggio di convergenza della serie di potenze è

$$R = \frac{1}{l}$$

(con la solita convenzione $1/0 = \infty, 1/\infty = 0$*).*

La dimostrazione è analoga a quella del teorema precedente: si applica questa volta il criterio del rapporto per le serie numeriche a termini positivi[11] alla serie $\sum_n |a_n (x-x_0)^n|$ e si ragiona di conseguenza.

Esempio

Calcoliamo il raggio di convergenza delle seguenti serie:

a. $\sum_{n=0}^{\infty} (-1)^n \frac{n}{2^n} (x-3)^n$ b. $\sum_{n=0}^{\infty} \frac{n+1}{n!} x^n$ c. $\sum_{n=0}^{\infty} n! (x-1)^n$

[10] Per farlo occorrerebbe però introdurre il concetto di *limite superiore di una successione.* Preferiamo non entrare in questi dettagli.

[11] v. ancora volume 1, capitolo 5, paragrafo 1.2.

 © 978-88-08-06485-1

a. È una serie di potenze di centro $x_0 = 3$. Abbiamo:

$$\lim_{n\to\infty} \sqrt[n]{\left|(-1)^n \frac{n}{2^n}\right|} = \lim_{n\to\infty} \frac{\sqrt[n]{n}}{2} = \frac{1}{2}$$

ricordando che $\sqrt[n]{n} \to 1$ (vedi volume 1, capitolo 3, Esempio 1.21). Dunque $R = 2$: la serie converge per $|x-3| < 2$, cioè per $x \in (1,5)$.

b. È una serie di potenze di centro 0. Data la presenza dei fattoriali, qui conviene invece usare il criterio del rapporto. Si ha:

$$\lim_{n\to\infty} \left|\frac{\frac{n+2}{(n+1)!}}{\frac{n+1}{n!}}\right| = \lim_{n\to\infty} \frac{n+2}{(n+1)^2} = 0$$

perciò $R = \infty$: la serie converge per ogni $x \in \mathbb{R}$.

c. È una serie di potenze di centro $x_0 = 1$. Ancora col criterio del rapporto, abbiamo:

$$\lim_{n\to\infty} \left|\frac{(n+1)!}{n!}\right| = \lim_{n\to\infty} (n+1) = +\infty.$$

Perciò $R = 0$: la serie converge solo per $x = 1$.

OSSERVAZIONE *Raggio di convergenza di una serie lacunare.* Consideriamo la serie di potenze

$$\sum_{n=0}^{\infty} \frac{x^{2n}}{3^n}.$$

Come si vede dal confronto con la serie geometrica, questa converge a $1/\left(1-\frac{x^2}{3}\right)$ per $|x| < \sqrt{3}$. Notiamo che i coefficienti della serie sono:

$$a_n = \begin{cases} \dfrac{1}{3^{n/2}} & \text{per } n \text{ pari} \\ 0 & \text{per } n \text{ dispari,} \end{cases}$$

per cui formalmente non esiste né

$$\lim_{n\to\infty} \sqrt[n]{|a_n|}, \quad \text{né} \quad \lim_{n\to\infty} \left|\frac{a_{n+1}}{a_n}\right|.$$

Una serie di potenze che abbia infiniti coefficienti a_n nulli si dice *serie lacunare*. Si può dimostrare che, se esiste il limite l della sottosuccessione di $\sqrt[n]{|a_n|}$ costituita dai soli termini non nulli, il numero $R = 1/l$ è ancora il raggio di convergenza della serie (come nell'esempio che abbiamo appena fatto).

Nel caso particolare di serie contenenti solo potenze di esponente pari o dispari, la cosa più semplice è vedere la serie come una serie completa di potenze di x^2, che converge per $x^2 < R$ e quindi $|x| < \sqrt{R}$. Nell'esempio precedente:

$$\sum_{n=0}^{\infty} \frac{x^{2n}}{3^n} = \sum_{n=0}^{\infty} \frac{t^n}{3^n} \text{ con } t = x^2, \text{ converge per } |t| < 3, \text{ cioè } |x| < \sqrt{3}.$$

Il seguente teorema raccoglie le principali proprietà delle serie di potenze, che, come vedremo, seguono facilmente combinando i risultati generali sulle serie di funzioni totalmente convergenti (v. paragrafo 1) con le proprietà appena viste che riguardano il raggio di convergenza di una serie di potenze.

© 978-88-08-06485-1

TEOREMA 7.6 (PROPRIETÀ DELLE SERIE DI POTENZE) *Consideriamo la serie* (2.1) *e supponiamo che il suo raggio di convergenza R sia positivo (finito o infinito). Allora:*

1. *Per ogni numero $r \in (0,R)$, la serie* (2.1) *converge totalmente nell'intervallo $[x_0 - r, x_0 + r]$.*

2. *La somma della serie* (2.1) *è una funzione continua nell'intervallo $(x_0 - R, x_0 + R)$.*

3. *La somma della serie* (2.1) *è una funzione derivabile nell'intervallo $(x_0 - R, x_0 + R)$ e la serie può essere derivata termine a termine:*

$$\frac{d}{dx}\left(\sum_{n=0}^{\infty} a_n (x-x_0)^n\right) = \sum_{n=1}^{\infty} n a_n (x-x_0)^{n-1}.$$

Inoltre, la serie derivata è anch'essa una serie di potenze di centro x_0, con lo stesso raggio di convergenza R di quella di partenza.

4. *Il processo descritto al punto precedente può essere iterato indefinitamente; pertanto la somma di una serie di potenze è una funzione derivabile infinite volte nell'intervallo $(x_0 - R, x_0 + R)$.*

5. *La somma della serie* (2.1) *ammette primitiva nell'intervallo $(x_0 - R, x_0 + R)$, che può essere calcolata termine a termine:*

$$\int \sum_{n=0}^{\infty} a_n (x-x_0)^n \, dx = \sum_{n=0}^{\infty} \frac{a_n}{n+1} (x-x_0)^{n+1} + c.$$

Inoltre, la primitiva è anch'essa esprimibile come serie di potenze di centro x_0, con lo stesso raggio di convergenza R di quella di partenza.

6. *Per ogni intervallo $[a,b] \subset (x_0 - R, x_0 + R)$, la somma della serie* (2.1) *è integrabile in $[a,b]$ e si può integrare termine a termine:*

$$\int_a^b \sum_{n=0}^{\infty} a_n (x-x_0)^n \, dx = \sum_{n=0}^{\infty} \int_a^b a_n (x-x_0)^n \, dx.$$

DIMOSTRAZIONE.

1. Sia $r \in (0,R)$. Dalla dimostrazione del teorema sul raggio di convergenza, si vede che la serie

$$\sum_{n=0}^{\infty} |a_n| r^n$$

converge. Quindi per $x \in [x_0 - r, x_0 + r]$ si ha

$$|a_n (x-x_0)^n| \le |a_n| r^n,$$

e la serie (2.1) converge totalmente.

© 978-88-08-06485-1

2. Poiché le funzioni
$$f_n(x) = a_n (x - x_0)^n$$
sono continue in $\mathbb{R}$, per il punto 1 e il Teorema 7.1 si ha che la somma della serie di potenze è una funzione f continua in $[x_0 - r, x_0 + r]$ per ogni $r < R$ e perciò è continua in $(x_0 - R, x_0 + R)$.

3. Consideriamo la serie delle derivate:
$$\sum_{n=0}^{\infty} f_n'(x) = \sum_{n=1}^{\infty} n a_n (x - x_0)^{n-1} = \sum_{n=0}^{\infty} (n+1) a_{n+1} (x - x_0)^n$$
(si noti che ora la somma parte da 1 perché la derivata del termine costante a_0 è zero). Calcoliamo il raggio di convergenza della serie derivata, supponendo che quello della serie di partenza sia[12]
$$R = \frac{1}{\lim_{n\to\infty} \sqrt[n]{|a_n|}}.$$
Per la serie derivata, dobbiamo quindi calcolare il limite di
$$\frac{1}{\sqrt[n]{|(n+1)\,a_{n+1}|}} \sim \frac{1}{\sqrt[n]{|a_{n+1}|}} = \left(\frac{1}{\sqrt[n+1]{|a_{n+1}|}}\right)^{\frac{n+1}{n}} \sim R^{\frac{n+1}{n}} \to R.$$
Dunque, la serie derivata è un'altra serie di potenze avente lo stesso raggio di convergenza di quella di partenza. In particolare, sarà totalmente convergente in $[x_0 - r, x_0 + r]$ per ogni $r < R$, perciò la somma della serie di potenze è una funzione derivabile e si può derivare termine a termine la serie, in ogni intervallo $[x_0 - r, x_0 + r]$ (per $r < R$) e quindi in $(x_0 - R, x_0 + R)$.

4. Il punto precedente può essere ora applicato alla somma della serie derivata e il processo si può iterare indefinitamente, da cui l'asserto.

5. Consideriamo la serie di potenze
$$\sum_{n=0}^{\infty} \frac{a_n}{n+1} (x - x_0)^{n+1} + c = \sum_{n=1}^{\infty} \frac{a_{n+1}}{n} (x - x_0)^n + c$$
e ragioniamo come nella dimostrazione del punto 3: ne deduciamo che questa ha lo stesso raggio di convergenza della serie di partenza e, ancora per il punto 3, è derivabile termine a termine; d'altro canto, la sua serie derivata è la serie di partenza, perciò questa serie ne è la primitiva.

6. Per il punto 1, la serie di partenza converge totalmente in $[x_0 - r, x_0 + r]$ per ogni $r < R$. Allora, per il Teorema 7.3, la serie è integrabile termine a termine in ogni intervallo $[a, b]$ contenuto in $[x_0 - r, x_0 + r]$ per qualche $r < R$, ossia in ogni intervallo $[a, b]$ contenuto in $(x_0 - R, x_0 + R)$. ♣

Esempio

I prossimi esempi mostrano come la possibilità di integrare e derivare termine a termine le serie sia uno strumento potente che permette talvolta di calcolare la somma di una serie o di sviluppare in serie una funzione data.

[12] In questa dimostrazione supporremo che sia possibile calcolare il raggio di convergenza mediante il criterio della radice. La nostra conclusione vale comunque in generale.

© 978-88-08-06485-1

Esempio Vogliamo sviluppare in serie di potenze la funzione
$$f(x) = \operatorname{arctg} x.$$
Naturalmente potremmo calcolare le derivate della funzione in 0 e scriverne la serie di Taylor, ma possiamo anche procedere in altro modo. Sfruttiamo il fatto che $f'(x) = 1/(1+x^2)$ e questa funzione si può sviluppare in serie di potenze utilizzando la serie geometrica
$$\sum_{n=0}^{\infty} t^n = \frac{1}{1-t} \quad \text{con } t = -x^2.$$
Si trova:
$$\sum_{n=0}^{\infty} (-1)^n x^{2n} = \frac{1}{1+x^2}$$
convergente per $|x| < 1$. Integrando termine a termine in $[0, x]$, si trova:
$$\operatorname{arctg} x = \sum_{n=0}^{\infty} (-1)^n \int_0^x t^{2n} dt = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1}.$$

Esempio Con lo stesso metodo possiamo ottenere lo sviluppo in serie di $\log(1+x)$. Utilizzando la serie geometrica, scriviamo, per $|t| < 1$,
$$\frac{1}{1+t} = \sum_{n=0}^{\infty} (-1)^n t^n$$
e integrando termine a termine in $[0, x]$ si ha, per $|x| < 1$,
$$\log(1+x) = \sum_{n=0}^{\infty} (-1)^n \int_0^x t^n dt = \sum_{n=0}^{\infty} (-1)^n \frac{x^{n+1}}{n+1} = \sum_{n=1}^{\infty} (-1)^{n-1} \frac{x^n}{n}.$$

Le proprietà appena viste delle serie di potenze rivelano la forza e il limite, al tempo stesso, di questo strumento di approssimazione. La forza sta nella facilità d'uso: una serie di potenze si può derivare e antiderivare termine a termine "finché si vuole" e l'intervallo di convergenza non cambia in queste operazioni. La limitazione sta nel fatto che questo strumento è adatto solo ad approssimare funzioni estremamente regolari: per quanto detto, se volessimo approssimare una funzione discontinua, o continua ma non derivabile, per esempio, non potremmo mai sperare di scriverla come somma di una serie di potenze, perché queste serie rappresentano solo funzioni infinitamente derivabili.

Ecco il motivo per cui in matematica si studiano anche altri tipi di serie di funzioni, sempre costruite a partire da funzioni elementari molto regolari, ma tali da poter avere come somma anche funzioni meno regolari: queste serie avranno però la proprietà di convergere "meno facilmente" delle serie di potenze. Avremo quindi una teoria più flessibile, quanto alle applicazioni, che pagheremo con una maggior difficoltà nell'indagare la convergenza di tali serie; tale è, per esempio, la teoria delle serie di Fourier.

 © 978-88-08-06485-1

Comportamento di una serie di potenze agli estremi dell'intervallo di convergenza
Terminiamo questa discussione sui concetti di base delle serie di potenze discutendo il problema della convergenza negli estremi dell'intervallo.

Esempio

(i) La serie geometrica

$$\sum_{n=0}^{\infty} x^n$$

ha raggio di convergenza $R = 1$. Osserviamo che per $x = 1$ la serie diverge; per $x = -1$ è irregolare. In entrambi i casi, non converge.

(ii) La serie di Taylor della funzione $\log(1+x)$ è

$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n}$$

(v. Esempio 2.3) e, vista come serie di potenze, ha raggio $R = 1$. Osserviamo che per $x = 1$ la serie converge (criterio di Leibniz per le serie a segni alterni), mentre per $x = -1$ la serie diverge (serie armonica).

Come si vede, agli estremi dell'intervallo la serie può convergere, divergere, oscillare. Nel caso converga, ci si può chiedere che relazione ci sia tra la somma di quella serie nell'estremo dell'intervallo e all'interno. Per esempio, per la serie di Taylor di $\log(1+x)$ ciò significa chiedersi se vale l'uguaglianza

$$\sum_{n=1}^{\infty} (-1)^{n+1} \frac{1}{n} = \log 2$$

(che si ottiene per $x = 1$, estremo dell'intervallo di convergenza). Risponde a questa domanda il prossimo teorema, che non dimostriamo:

TEOREMA 7.7 (DI ABEL) *Sia $R \in (0, \infty)$ il raggio di convergenza di una serie di potenze $\sum_{n=0}^{\infty} a_n x^n$ e supponiamo che*

$$\sum_{n=0}^{\infty} a_n R^n \text{ converga.}$$

Allora

$$\sum_{n=0}^{\infty} a_n R^n = \lim_{x \to R^-} \left(\sum_{n=0}^{\infty} a_n x^n \right).$$

Analogamente, se converge

$$\sum_{n=0}^{\infty} a_n (-R)^n,$$

allora

$$\sum_{n=0}^{\infty} a_n (-R)^n = \lim_{x \to (-R)^+} \left(\sum_{n=0}^{\infty} a_n x^n \right).$$

© 978-88-08-06485-1

Esempio

(i) $$\sum_{n=1}^{\infty} \frac{(-1)^{n+1}}{n} = \lim_{x \to 1^-} \left(\sum_{n=1}^{\infty} \frac{(-1)^{n+1} x^n}{n} \right) = \lim_{x \to 1^-} (\log(1+x)) = \log 2.$$

(ii) Nell'Esempio 2.2 abbiamo calcolato lo sviluppo

$$\operatorname{arctg} x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{2n+1} \quad \text{per } |x| < 1.$$

D'altro canto, l'ultima serie converge anche per $x = 1$ (criterio di Leibniz), quindi, in base al teorema di Abel, possiamo calcolare:

$$\sum_{n=0}^{\infty} \frac{(-1)^n}{2n+1} = \lim_{x \to 1^-} \operatorname{arctg} x = \frac{\pi}{4},$$

risultato interessante perché fornisce una rappresentazione di π come serie a segni alterni[13]:

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

2.2 Serie di Taylor e serie di potenze

Vogliamo ora confrontare la nozione di *serie di Taylor*, introdotta nello studio del calcolo differenziale per funzioni di una variabile[14], con quella di *serie di potenze*, introdotta in questo paragrafo.

Se $f : \mathbb{R} \to \mathbb{R}$ è una funzione derivabile infinite volte in $x = 0$, possiamo scrivere il suo sviluppo di MacLaurin con resto secondo Lagrange:

$$f(x) = \sum_{k=0}^{n} \frac{f^{(k)}(0)}{k!} x^k + E_n(x),$$

per ogni x in un intorno I di 0, con

$$E_n(x) = \frac{f^{(n+1)}(c)}{(n+1)!} x^{n+1},$$

per un opportuno punto $c \in (0, x)$, dipendente da x e da n. Se si riesce a dimostrare che, per ogni $x \in I$, $E_n(x) \to 0$ per $n \to \infty$, allora f si può rappresentare in *serie di MacLaurin*:

$$f(x) = \sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k \qquad \forall x \in I.$$

[13] Ricordiamo che una serie a segni alterni consente di controllare la somma della serie tra due somme parziali successive. Per esempio:

$$4 \sum_{k=0}^{99\,999} \frac{(-1)^k}{2k+1} < \pi < 4 \sum_{k=0}^{100\,000} \frac{(-1)^k}{2k+1},$$

cioè

$$3{,}14158 < \pi < 3{,}14160.$$

[14] v. volume 1, capitolo 5, paragrafo 2.

 © 978-88-08-06485-1

In tal caso diciamo che f è *sviluppabile in serie di MacLaurin*. In particolare, f si può vedere come somma di una serie di potenze, da cui si deduce anche che l'intervallo I è necessariamente del tipo $(-R, R)$ (eventualmente comprendendo anche uno o due estremi).

Si noti che la verifica della condizione $E_n(x) \to 0$ per $n \to \infty$ richiede qualche informazione sul comportamento delle derivate di f *in un intorno* di $x = 0$ e non solo in quel punto. Come vedremo con alcuni esempi, dalla sola conoscenza delle derivate di f nel punto si può concludere poco, in generale.

Rovesciamo ora, in un certo senso, il discorso. Consideriamo una serie di potenze

$$\sum_{k=0}^{\infty} a_k x^k \tag{2.2}$$

e supponiamo che abbia raggio di convergenza $R > 0$. Allora, nell'intervallo $(-R, R)$, la somma della serie è una funzione $f(x)$ infinitamente derivabile. Chiediamoci: questa funzione f è sviluppabile in serie di MacLaurin? La risposta è: sì e la sua serie di MacLaurin è la (2.2) stessa. In altre parole, se poniamo

$$f(x) = \sum_{k=0}^{\infty} a_k x^k \qquad \forall x \in (-R, R)$$

e calcoliamo le derivate di f nell'origine, troviamo:

$$f^{(k)}(0) = a_k \cdot k!$$

perciò

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k = \sum_{k=0}^{\infty} a_k x^k.$$

Questi ragionamenti mostrano che *la classe delle funzioni sviluppabili in serie di MacLaurin coincide con la classe delle serie di potenze di centro nell'origine, aventi raggio di convergenza positivo.*

Le due nozioni di serie di potenze e serie di MacLaurin rappresentano dunque due punti di vista diversi sullo stesso oggetto: dal primo punto di vista, si parte da una funzione $f(x)$, assegnata in qualche forma analitica, e se ne calcola la serie di MacLaurin; dal secondo punto di vista, si parte dalla serie di potenze e la si riconosce come serie di MacLaurin di una certa funzione.

In modo del tutto analogo, le *funzioni sviluppabili in serie di Taylor di centro* x_0 per x in un intorno di x_0

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(x_0)}{k!} (x - x_0)^k$$

sono tutte e sole le *serie di potenze centrate in* x_0

$$\sum_{k=0}^{\infty} a_k (x - x_0)^k$$

con raggio di convergenza $R > 0$. Queste funzioni si dicono anche *analitiche*. Più precisamente:

© 978-88-08-06485-1

Definizione 7.4 Una funzione si dice *analitica* in un intervallo (a, b) se per ogni $x_0 \in (a, b)$ la funzione è sviluppabile in serie di potenze (ovvero in serie di Taylor) di centro x_0, con raggio di convergenza positivo.

Le funzioni trascendenti elementari

$$e^x, \quad \sin x, \quad \cos x, \quad \mathrm{Sh}x, \quad \mathrm{Ch}x$$

sono esempi di funzioni analitiche in tutto $\mathbb{R}$; le funzioni

$$\log(1 + x), \quad (1 + x)^\alpha$$

sono invece analitiche[15] in $(-1, +\infty)$.

Ricordiamo ancora una volta i loro sviluppi in $x_0 = 0$:

$$e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!} \qquad \text{per ogni } x \in \mathbb{R}$$

$$\sin x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n+1}}{(2n+1)!} \qquad \text{per ogni } x \in \mathbb{R}$$

$$\cos x = \sum_{n=0}^{\infty} (-1)^n \frac{x^{2n}}{(2n)!} \qquad \text{per ogni } x \in \mathbb{R}$$

$$\mathrm{Sh}x = \sum_{n=0}^{\infty} \frac{x^{2n+1}}{(2n+1)!} \qquad \text{per ogni } x \in \mathbb{R}$$

$$\mathrm{Ch}x = \sum_{n=0}^{\infty} \frac{x^{2n}}{(2n)!} \qquad \text{per ogni } x \in \mathbb{R}$$

$$\log(1+x) = \sum_{n=1}^{\infty} (-1)^{n+1} \frac{x^n}{n} \qquad \text{per } |x| < 1$$

$$(1+x)^\alpha = \sum_{n=0}^{\infty} \binom{\alpha}{n} x^n \qquad \text{per } |x| < 1,$$

$$\text{con } \binom{\alpha}{n} = \frac{\alpha(\alpha-1)(\alpha-2)\cdots(\alpha-n+1)}{n!}$$

Funzioni C^∞, ma non analitiche

Non tutte le funzioni derivabili infinite volte in un punto sono analitiche in un intorno di quel punto.

Infatti, ragionando, per fissare le idee, nell'origine, una volta calcolate le derivate $f^{(k)}(0)$ per ogni $k \geq 0$, possiamo sempre scrivere formalmente la serie di MacLaurin

[15] Ricordiamo che, in base alla definizione, questo significa che per ogni $x_0 \in (-1, +\infty)$ la funzione è sviluppabile in serie di potenze di centro x_0 e raggio positivo; naturalmente non esiste una *singola* serie di potenze che converga in tutto $(-1, +\infty)$. Infatti, l'insieme di convergenza di una serie di potenze può essere un segmento o l'intera retta, ma non una semiretta.

© 978-88-08-06485-1

di f,

$$\sum_{k=0}^{\infty} \frac{f^{(k)}(0)}{k!} x^k,$$

ma:

(1) non è detto che questa serie di potenze abbia raggio di convergenza positivo;

(2) non è detto che, se questa serie ha raggio di convergenza positivo, la sua somma coincida con f.

Esempio

2.16 *Una funzione f la cui serie di Taylor non converge a f* Sia

$$f(x) = \begin{cases} e^{-\frac{1}{x^2}} & \text{per } x \neq 0 \\ 0 & \text{per } x = 0. \end{cases}$$

Il lettore è invitato a verificare che f è derivabile infinite volte in 0 e

$$f^{(k)}(0) = 0 \quad \text{per ogni } k.$$

Dunque, la serie di MacLaurin di f è la serie identicamente nulla, che ovviamente converge, ma non rappresenta f. La funzione f è quindi infinitamente derivabile in tutto $\mathbb{R}$, ma non è analitica in nessun intorno dell'origine. Questo mostra (2).

OSSERVAZIONE L'esistenza di esempi come il precedente mostra che il solo fatto di saper calcolare tutte le derivate di una funzione in un punto e di ottenere una serie di Taylor con raggio di convergenza positivo ancora non garantisce che la serie di Taylor rappresenti la funzione di partenza. Per garantire questo fatto, occorre qualche informazione sulle derivate della funzione in un intorno del punto stesso, come già osservato.

Esempio

2.17 *Una funzione f la cui serie di Taylor non converge* Sia

$$g(x) = \int_0^{+\infty} \frac{e^{-t}}{1+tx} dt.$$

La funzione[16] g è definita come integrale generalizzato dipendente dal parametro x. Si vede subito che per ogni $x \geq 0$ l'integrale generalizzato converge, perciò risulta ben definita $g : [0, +\infty) \to \mathbb{R}$. Notiamo che

$$g(0) = \int_0^{+\infty} e^{-t} dt = 1.$$

[16] chiamata talvolta *funzione di Stieltjes*.

© 978-88-08-06485-1

Si può dimostrare che g è infinitamente derivabile sotto il segno di integrale, perciò:

$$g'(x) = \int_0^{+\infty} -t\frac{e^{-t}}{(1+tx)^2} dt; \qquad g'(0) = \int_0^{+\infty} -te^{-t} dt = -1;$$

$$g''(x) = \int_0^{+\infty} 2t^2\frac{e^{-t}}{(1+tx)^3} dt; \qquad g''(0) = 2\int_0^{+\infty} t^2 e^{-t} dt = 2 \cdot 2;$$

$$g'''(x) = \int_0^{+\infty} -3 \cdot 2t^3\frac{e^{-t}}{(1+tx)^4} dt; \quad g'''(0) = -3 \cdot 2\int_0^{+\infty} t^3 e^{-t} dt = -3! \cdot 3!$$

e in generale

$$g^{(k)}(0) = (-1)^k (k!)^2,$$

per cui la serie di MacLaurin di g è

$$\sum_{k=0}^{\infty} (-1)^k k! x^k,$$

che ha raggio di convergenza nullo. Notiamo che la funzione g è definita solo per $x \geq 0$, perciò le sue derivate in zero sono in realtà derivate destre e non possiamo dire che g sia C^∞ in un intorno dell'origine. Se però ora definiamo

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

otteniamo una funzione $f \in C^\infty(\mathbb{R})$, la cui serie di MacLaurin è

$$\sum_{k=0}^{\infty} (-1)^k k! x^{2k},$$

che ha ancora raggio di convergenza nullo. Di nuovo, troviamo una funzione $C^\infty(\mathbb{R})$ che non è analitica in nessun intorno dell'origine, questa volta perché la serie di MacLaurin non converge (mentre nell'esempio precedente, converge a una funzione diversa da f). Questo esempio illustra la situazione (1).

Per quanto questo esempio possa apparire artificioso, funzioni come g appaiono naturalmente in diversi problemi fisici (per esempio di meccanica quantistica), in cui una certa grandezza è espressa da un integrale dipendente da un parametro, che converge solo per valori del parametro maggiori o uguali a una certa soglia.

Esercizi

2.18 Provare che la serie

$$\sum_{n=0}^{\infty} \left(\frac{x}{1+x^2}\right)^n$$

converge totalmente in $\mathbb{R}$. Qual è la sua somma? *Suggerimento*: dimostrare che $\left|\frac{x}{1+x^2}\right| \leq \frac{1}{2}$.

2.19 Dimostrare che la serie

$$\sum_{n=0}^{\infty} \frac{\sin(n!x)}{2^n}$$

rappresenta una funzione continua su tutto $\mathbb{R}$. La serie si può derivare termine a termine?

Si consideri la serie di funzioni

$$\sum_{n=0}^{\infty} e^{-n^2 x}.$$

a) Determinare l'insieme I dei punti in cui converge puntualmente.

b) Stabilire se la serie converge totalmente in I oppure, in caso contrario, stabilire in quali intervalli contenuti in I la serie converge totalmente.

c) Stabilire se la somma $f(x)$ della serie è una funzione continua in I oppure in quale sottointervallo è continua.

d) Lo stesso per la derivabilità di f.

Determinare il raggio di convergenza delle seguenti serie di potenze:

$$\sum_{n=0}^{\infty} \frac{n^3 x^n}{2^n}; \quad \sum_{n=1}^{\infty} n^n x^n; \quad \sum_{n=1}^{\infty} \frac{(-1)^n x^n}{n};$$
$$\sum_{n=1}^{\infty} \frac{3^n}{2^{2n}} x^n; \quad \sum_{n=1}^{\infty} 2^n x^n; \quad \sum_{n=1}^{\infty} \frac{n^5 + 2^n}{3^n + (-1)^n 2^n} x^{2n}.$$

Per ciascuna delle serie precedenti per cui il raggio di convergenza è positivo e finito, si controlli il comportamento della serie negli estremi dell'intervallo di convergenza, stabilendo se essa converge, diverge o oscilla.

Esempio

Dimostrare le seguenti identità, valide per $|x| < 1$:

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

Suggerimento: per cominciare, derivare termine a termine la serie geometrica (come già visto nell'Esempio 1.1); per provare la seconda identità occorre anche utilizzare la prima. Da questo esercizio dovrebbe risultare chiaro come, con un procedimento iterativo e molta pazienza, si può calcolare la somma di $\sum_{n=1}^{\infty} n^k x^n$ per qualsiasi intero positivo k.

Ricavare lo sviluppo in serie di potenze della funzione *seno integrale*, definita da:

$$\mathrm{Si}(x) = \int_0^x \frac{\sin t}{t} dt.$$

Suggerimento: dalla serie di potenze di $\sin t$ scrivere quella dell'integranda, quindi integrare termine a termine.

Sfruttando la serie di MacLaurin della funzione

$$f(x) = (1+x)^\alpha,$$

si scriva lo sviluppo in serie di

$$g(x) = \frac{1}{\sqrt{1-x^2}} \quad \text{per } |x| < 1.$$

Integrando poi termine a termine questa serie di potenze in $[0, x]$ per $|x| < 1$, si ottenga lo sviluppo di MacLaurin della funzione $h(x) = \arcsin x$.

Valutando infine la serie di $\arcsin x$ in $x = 1/\sqrt{2}$, si ricavi uno sviluppo in serie del numero π.

Sfruttando opportunamente le serie di MacLaurin delle funzioni trascendenti elementari, si calcoli la somma delle seguenti serie numeriche:

$$\sum_{n=1}^{\infty} \frac{(-1)^n 4^n}{n!}; \quad \sum_{n=1}^{\infty} \frac{5^n}{2n}; \quad \sum_{n=0}^{\infty} \frac{(-1)^n 9^n}{(2n+1)!}.$$

3 SERIE TRIGONOMETRICHE E SERIE DI FOURIER

3.1 Polinomi trigonometrici e serie trigonometriche

Le funzioni $\sin x$, $\cos x$ sono periodiche di periodo 2π (in breve: 2π-periodiche); le funzioni $\sin 3x$, $\cos 3x$ sono invece $\frac{2\pi}{3}$-periodiche; in generale $\sin nx$, $\cos nx$ sono funzioni $\frac{2\pi}{n}$-periodiche. Una combinazione lineare di alcune di queste funzioni, per esempio

$$2\sin x + \frac{1}{2}\sin 3x - 4\cos 7x,$$

è ancora (almeno) 2π-periodica.

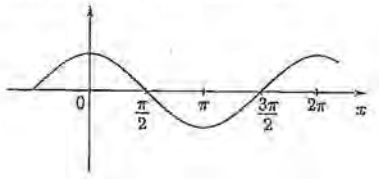


Figura 7.1. $x \longmapsto \cos x$.

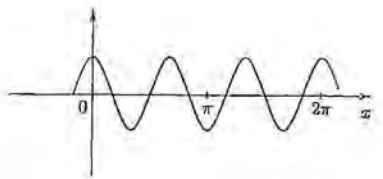


Figura 7.2. $x \longmapsto \cos 3x$.

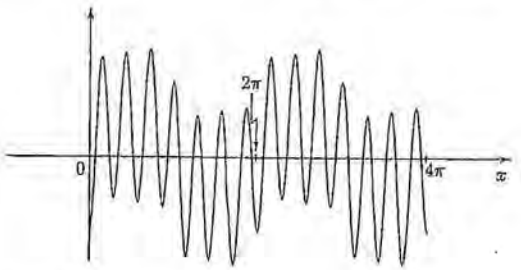


Figura 7.3. $x \longmapsto 2\sin x + \frac{1}{2}\sin 3x - 4\cos 7x$.

Anche il grafico di figura 7.4 è quello di una funzione 2π-periodica. Nessuno si aspetterebbe però che anch'essa risulti dalla sovrapposizione di funzioni del tipo $\sin nx$ e $\cos nx$. Proprio questa invece fu la straordinaria intuizione di Fourier: che "ogni

 © 978-88-08-06485-1

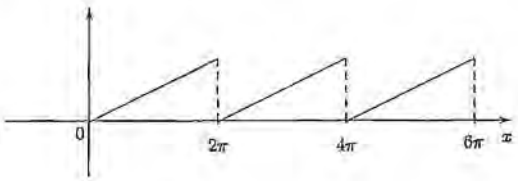


Figura 7.4.

funzione" 2π-periodica potesse essere rappresentata da una somma, eventualmente infinita, di funzioni periodiche elementari del tipo

$$\sin x, \cos x, \sin 2x, \cos 2x, \ldots, \sin nx, \cos nx, \ldots$$

più eventualmente la funzione costante. Precisiamo il discorso.

DEFINIZIONE 7.5 Si dice *polinomio trigonometrico di ordine* n una funzione del tipo:

$$P_n(x) = a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx), \tag{3.1}$$

dove a_k, b_k sono numeri reali o complessi, assegnati.

Si dice *serie trigonometrica* l'analoga espressione dove la somma finita sia sostituita da una serie:

$$a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx). \tag{3.2}$$

Queste serie risultano di grandissima utilità in svariate applicazioni, tra cui l'elaborazione o la compressione di segnali periodici di vario tipo, o di immagini, e la risoluzione esplicita di equazioni alle derivate parziali in domini spaziali di geometria piuttosto semplice.

Sono polinomi trigonometrici, per esempio:

$$P_3(x) = 2 + \cos x - 2\cos 3x + \sqrt{2}\sin 2x, \qquad P_{10}(x) = \sum_{k=1}^{10} \frac{\sin kx}{k} \qquad \text{ecc.}$$

Sono serie trigonometriche:

$$\sum_{k=1}^{\infty} \frac{\sin kx}{k}, \qquad \sum_{k=1}^{\infty} \left\{ \frac{\cos kx}{k^2+1} + (-1)^k \frac{\sin kx}{k!} \right\} \qquad \text{ecc.}$$

La definizione data di serie trigonometrica è puramente formale: come per ogni serie di funzioni, non è detto che converga né, in caso affermativo, che la sua somma sia una funzione con buone proprietà (paragonabili a quelle dei polinomi trigonometrici, che sono infinitamente derivabili). L'unica cosa per ora ovvia è che, se una serie trigonometrica (3.2) converge, la sua somma è una funzione periodica di periodo 2π.

© 978-88-08-06485-1

Convergenza delle serie trigonometriche

Consideriamo una serie trigonometrica (3.2). Se i coefficienti a_k, b_k non tendono a zero, certamente la serie non converge. Se tendono a zero abbastanza velocemente, in modo che le serie $\sum |a_k|, \sum |b_k|$ convergano, per il Teorema 7.1 la serie (3.2) converge totalmente in $\mathbb{R}$, a una funzione continua; infatti:

$$|a_k \cos kx + b_k \sin kx| \le |a_k| + |b_k|.$$

Quando questa condizione non è verificata, perché a_k e b_k tendono a zero troppo lentamente, per studiare la convergenza della serie può essere utile il seguente criterio, che dimostreremo nei Complementi:

PROPOSIZIONE 7.1 *Sia* $\{a_n\}$ *una successione a valori reali positivi che tende monotonamente a zero. Allora:*

(a) *Le serie trigonometriche*

$$\sum_{n=1}^{\infty} a_n \cos nx, \qquad \sum_{n=1}^{\infty} a_n \sin nx$$

convergono per ogni $x \in (0, 2\pi)$. *Negli estremi* $0, 2\pi$ *la serie di coseni può convergere o no, mentre la serie di seni è identicamente nulla.*

(b) *Le serie trigonometriche*

$$\sum_{n=1}^{\infty} (-1)^n a_n \cos nx, \qquad \sum_{n=1}^{\infty} (-1)^n a_n \sin nx$$

convergono per ogni $x \in [0, 2\pi]$, $x \neq \pi$. *In* $x = \pi$ *la serie di coseni può convergere o no, mentre la serie di seni è identicamente nulla.*

Si noti che, anche in caso di convergenza, non è garantito che la somma della serie sia una funzione continua.

Esempio

Le serie trigonometriche

$$\sum_{n=1}^{\infty} \frac{\sin nx}{n}, \quad \sum_{n=1}^{\infty} \frac{\cos nx}{n}, \quad \sum_{n=1}^{\infty} \frac{\sin nx}{\sqrt{n}}, \quad \sum_{n=1}^{\infty} \frac{\cos nx}{\sqrt{n}} \quad \text{ecc.}$$

convergono per ogni $x \in (0, 2\pi)$, in base alla proposizione. Si noti che non sarebbe possibile dimostrarne la convergenza in base al criterio della convergenza totale, perché le serie $\sum_{n=1}^{\infty} 1/n, \sum_{n=1}^{\infty} 1/\sqrt{n}$ divergono.

Osserviamo che cosa accade in $x = 0$: le serie di seni sono identicamente nulle; le serie di coseni in questo caso divergono. Analogamente le serie

$$\sum_{n=1}^{\infty} (-1)^n \frac{\sin nx}{n}, \quad \sum_{n=1}^{\infty} (-1)^n \frac{\cos nx}{\sqrt{n}} \quad \text{ecc.}$$

convergono per $x \in [0, 2\pi]$, $x \neq \pi$. In $x = \pi$ la serie di seni è identicamente nulla, quella di coseni diverge.

 © 978-88-08-06485-1

Derivazione di una serie trigonometrica

Osserviamo ora che la serie delle derivate della (3.2) è:

$$\sum_{k=1}^{\infty} \{-ka_k \sin(kx) + kb_k \cos(kx)\}, \tag{3.3}$$

che è ancora una serie trigonometrica, avente coefficienti

$$\begin{cases} a'_k = kb_k \\ b'_k = -ka_k. \end{cases}$$

Si vede subito che questa serie converge *più difficilmente* di quella di partenza: i suoi coefficienti sono più grandi in valore assoluto. In generale, ci aspettiamo che la serie (3.2) possa convergere, ma non essere derivabile termine a termine; lo sarà, per esempio, se i suoi coefficienti tendono a zero così rapidamente che anche $\sum |ka_k|, \sum |kb_k|$ convergono.

Da queste poche considerazioni ci possiamo aspettare che una serie trigonometrica convergente possa rappresentare funzioni regolari, oppure funzioni continue ma non derivabili, o anche funzioni discontinue. Le serie trigonometriche si presentano quindi come uno strumento più adatto delle serie di potenze a rappresentare funzioni anche non regolari. L'altra faccia della medaglia è che in generale sarà più delicato garantire la convergenza e la possibilità di operare su di esse, per esempio, derivando termine a termine.

Osserviamo ora i grafici delle somme parziali di alcune serie trigonometriche con diversa rapidità di convergenza, per esemplificare le proprietà appena stabilite. Si noti che, più lenta è la convergenza, più "tremolante" è il grafico; il che suggerisce che la serie delle derivate non converge.





$$\sum_{k=1}^{\infty} \frac{\sin kx}{k^3}$$

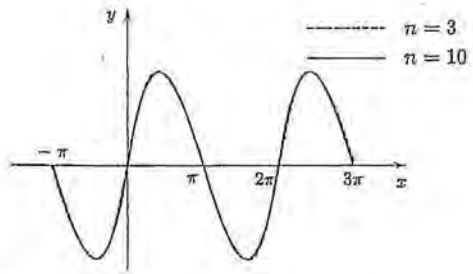


Figura 7.5. Grafico delle somme parziali per $n = 3, n = 10$.

Il termine generale va a zero rapidamente; sia la serie $\sum_{k=0}^{\infty} \frac{\sin kx}{k^3}$ che la serie delle derivate convergono totalmente (in particolare, la serie si può derivare termine a termine); la somma parziale per $n = 10$ rappresenta una funzione molto "liscia" ed è quasi indistinguibile da quella per $n = 3$.



$$\sum_{k=1}^{\infty} \frac{\sin kx}{k}$$

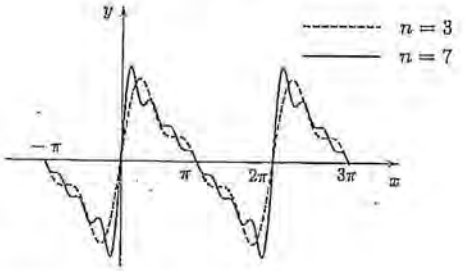


Figura 7.6. Grafico delle somme parziali per $n = 3, n = 7$.

Il termine generale tende a zero lentamente, la serie converge (per la proposizione 7.1, perché $1/k$ è positivo, decrescente e tende a zero) ma non totalmente; la somma potrebbe essere una funzione discontinua (vedremo in seguito che è proprio così); la serie delle derivate, poi, diverge.

Nel seguito di questo capitolo ci occuperemo del problema di determinare una serie trigonometrica che converga, in un senso opportuno, a una *assegnata* funzione. Questo significa poter approssimare la funzione con un polinomio trigonometrico (somma parziale della serie). Per capire come si individua un polinomio trigonometrico candidato naturale ad approssimare una funzione assegnata e per dimostrare che le cose vanno effettivamente così, si rivela molto utile un punto di vista geometrico, che sfrutta la geometria degli spazi vettoriali con prodotto scalare e in particolare l'idea di ortogonalità. A questo argomento dedichiamo quindi ora una breve parentesi, prima di entrare nel cuore del problema.

3.2 Richiami sugli spazi vettoriali con prodotto scalare

Sia V uno spazio vettoriale reale, di dimensione finita o infinita. Diciamo che V è dotato di un *prodotto scalare*, o *prodotto interno*, se esiste un'operazione che associa a ogni coppia di vettori $u, v \in V$ un numero reale, indicato col simbolo $\langle u, v \rangle$, che gode delle seguenti proprietà:

$$\begin{aligned} \langle u, v \rangle &= \langle v, u \rangle \\ \langle u + v, w \rangle &= \langle u, w \rangle + \langle v, w \rangle \\ \langle \lambda u, v \rangle &= \lambda \langle u, v \rangle \text{ per ogni } \lambda \in \mathbb{R} \\ \langle u, u \rangle &\geq 0 \text{ e } \langle u, u \rangle = 0 \text{ se e solo se } u = 0. \end{aligned}$$

© 978-88-08-06485-1

Se poniamo, per ogni $u \in V$,

$$\|u\| = \sqrt{\langle u, u \rangle},$$

si dimostra che $\|\cdot\|$ risulta essere una *norma*, cioè soddisfa le seguenti proprietà:

$$\|u\| \geq 0 \text{ e } \|u\| = 0 \text{ se e solo se } u = 0;$$
$$\|\lambda u\| = |\lambda| \, \|u\| ;$$
$$\|u + v\| \leq \|u\| + \|v\|$$

per ogni $u, v \in V, \lambda \in \mathbb{R}$.

Esempi

Negli spazi $\mathbb{R}^n$ il prodotto scalare usuale è definito da

$$\langle u, v \rangle = \sum_{i=1}^{n} u_i v_i \quad \text{se } u = (u_1, u_2, \ldots, u_n),\ v = (v_1, v_2, \ldots, v_n),$$

e la norma che esso induce è il modulo del vettore:

$$\|u\| = \sqrt{\sum_{i=1}^{n} u_i^2} = |u|.$$

Nello spazio $C[0, 2\pi]$, se definiamo

$$\langle f, g \rangle = \int_0^{2\pi} f(t)\, g(t)\, dt,$$

si verifica che $\langle \cdot, \cdot \rangle$ è un prodotto scalare, che induce la norma:

$$\|f\| = \left(\int_0^{2\pi} f^2(t)\, dt \right)^{1/2}.$$

In uno spazio vettoriale con prodotto interno, diciamo che due vettori sono *ortogonali* quando il loro prodotto scalare è nullo. In simboli:

$$u \perp v \iff \langle u, v \rangle = 0.$$

Vale il

TEOREMA 7.8 (DI PITAGORA) *Se $u_1, u_2, \ldots, u_n \in V$ sono a due a due ortogonali, allora*

$$\left\| \sum_{i=1}^{n} u_i \right\|^2 = \sum_{i=1}^{n} \|u_i\|^2.$$

DIMOSTRAZIONE. Per definizione di norma e bilinearità del prodotto scalare:

$$\left\| \sum_{i=1}^{n} u_i \right\|^2 = \left\langle \sum_{i=1}^{n} u_i, \sum_{j=1}^{n} u_j \right\rangle = \sum_{i,j=1}^{n} \langle u_i, u_j \rangle$$

poiché i vettori sono ortogonali a due a due, $\langle u_i, u_j \rangle = 0$ per $i \neq j$, quindi:

$$= \sum_{i=1}^{n} \langle u_i, u_i \rangle = \sum_{i=1}^{n} \|u_i\|^2. \qquad \blacklozenge$$

© 978-88-08-06485-1

Vale la seguente:

PROPOSIZIONE 7.2 *Se V_0 è un sottospazio di dimensione finita di V, V_0 possiede sempre una* base ortonormale, *cioè una base*

$$e_1, e_2, \ldots, e_n$$

tale che i vettori e_i siano a due a due ortogonali e ciascuno di norma unitaria. Sinteticamente, si scrive[17]*:*

$$\langle e_i, e_j \rangle = \delta_{ij} \equiv \begin{cases} 1 & \text{se } i = j \\ 0 & \text{se } i \neq j. \end{cases}$$

L'esistenza di una base ortonormale in un sottospazio di dimensione finita si può dimostrare mediante il classico *metodo di ortonormalizzazione di Gram-Schmidt*. Nel caso che ci interesserà nel seguito, tuttavia, sapremo esibire esplicitamente tale base, senza bisogno di ricorrere ad argomenti astratti. Perciò non dimostriamo questa proposizione.

Supponiamo ora di avere un sottospazio V_0 finito dimensionale di V, un elemento $u \in V$ e di voler trovare l'elemento $u_0 \in V_0$ che meglio approssima u (nella distanza della norma di V). Geometricamente, è evidente che tale elemento u_0 sarà la proiezione ortogonale di u su V_0. Questo fatto può essere enunciato con precisione e dimostrato facilmente:

TEOREMA 7.9 (DELLA PROIEZIONE) *Nelle ipotesi precedenti, detta $\{e_1, e_2, \ldots, e_n\}$ una base ortonormale di V_0, la proiezione di u su V_0 è il vettore*

$$u_0 = \sum_{i=1}^{n} \langle u, e_i \rangle \, e_i$$

e ha le seguenti proprietà:

(i) *$u - u_0$ è ortogonale a ogni elemento di V_0;*

(ii) *$\|u_0\|^2 = \sum_{i=1}^{n} |\langle u, e_i \rangle|^2 \leq \|u\|^2$;*

(iii) *u_0 è l'elemento di V_0 che rende minima la quantità $\|u - v\|$ al variare di v in V_0.*

DIMOSTRAZIONE. (i) Proviamo che $u_0 \perp e_j$ per $j = 1, 2, \ldots, n$. Per linearità, seguirà che u_0 è ortogonale a tutte le combinazioni lineari di $e_1, e_2, \ldots, e_n$, cioè a tutti gli elementi di V_0. Si ha:

$$\langle u - u_0, e_j \rangle = \left\langle u - \sum_{i=1}^{n} \langle u, e_i \rangle \, e_i, e_j \right\rangle = \langle u, e_j \rangle - \sum_{i=1}^{n} \langle u, e_i \rangle \, \langle e_i, e_j \rangle =$$
$$= \langle u, e_j \rangle - \langle u, e_j \rangle = 0$$

perché $\langle e_i, e_j \rangle = \delta_{ij}$.

[17] utilizzando il simbolo di Kronecker δ_{ij}, definito come sopra.

(ii) Scriviamo $u = (u - u_0) + u_0$. Per il punto (i) è $(u - u_0) \perp u_0$, perciò per il teorema di Pitagora si ha

$$\|u\|^2 = \|u - u_0\|^2 + \|u_0\|^2 \geq \|u_0\|^2 .$$

D'altro canto, per l'ortonormalità dei vettori e_i e il teorema di Pitagora si ha:

$$\|u_0\|^2 = \left\| \sum_{i=1}^{n} \langle u, e_i \rangle e_i \right\|^2 = \sum_{i=1}^{n} |\langle u, e_i \rangle|^2 .$$

(iii) Se $v \in V_0$, scriviamo $u - v = (u - u_0) + (u_0 - v)$. Poiché $(u_0 - v) \in V_0$, $(u - u_0) \perp (u_0 - v)$ e ancora il teorema di Pitagora dà

$$\|u - v\|^2 = \|u - u_0\|^2 + \|u_0 - v\|^2 \geq \|u - u_0\|^2$$

che mostra appunto che u_0 rende minima la distanza da u, tra tutti gli elementi di V_0. ∎

3.3 Coefficienti e serie di Fourier di una funzione. Approssimazione in media quadratica

Una serie trigonometrica (3.2), se converge, ha per somma una funzione periodica di periodo 2π. Quindi le funzioni che vorremo sviluppare in serie trigonometrica saranno funzioni definite su tutto $\mathbb{R}$ e periodiche di periodo 2π. Una funzione di questo genere è individuata ovviamente dai suoi valori in un qualunque intervallo di periodicità, cioè di ampiezza 2π. In altri termini, è sufficiente assegnare una funzione "base", f, definita per esempio in $[0, 2\pi]$, e pensare al suo prolungamento periodico, ottenuto riproducendo copie di f sugli intervalli $[2k\pi, 2(k+1)\pi]$, $k \neq 0$.

In alternativa, potremmo avere una funzione che originariamente è definita solo su un intervallo, per esempio $[0, 2\pi]$, e volerla approssimare mediante polinomi trigonometrici. Implicitamente, questo significa che stiamo pensando di prolungare la funzione a tutto $\mathbb{R}$, estendendola in modo 2π-periodico al di fuori dell'intervallo originario; la funzione che approssimeremo sarà in questo caso la funzione periodizzata.

In questo paragrafo ci porremo da questo secondo punto di vista (ricordando che, comunque, è del tutto equivalente al primo). Ci chiediamo quindi:

> data una funzione $f : [0, 2\pi] \to \mathbb{R}$, sotto quali condizioni si può vedere f come somma di una serie trigonometrica e come si determinano i coefficienti a_k, b_k di tale serie?

Consideriamo lo spazio vettoriale V costituito dalle funzioni $f : [0, 2\pi] \to \mathbb{R}$ integrabili[18] (quindi in particolare limitate), munito del prodotto scalare

$$\langle f, g \rangle = \int_0^{2\pi} f(t)\, g(t)\, dt,$$

e della relativa norma:

$$\|f\| = \sqrt{\langle f, f \rangle} = \left(\int_0^{2\pi} f^2(t)\, dt \right)^{1/2} .$$

[18] in senso ordinario, non in senso generalizzato. (Si veda il volume 1, capitolo 6).

Ricordiamo che la norma a sua volta induce una nozione di *distanza* tra le funzioni:

$$d(f, g) = \|f - g\| = \left(\int_0^{2\pi} [f(t) - g(t)]^2\, dt \right)^{1/2} .$$

Lo spazio V contiene in particolare $C[0, 2\pi]$.

Rispetto a questa struttura di spazio vettoriale con prodotto scalare, le funzioni $\cos kx$, $\sin kx$ risultano avere proprietà notevoli:

PROPOSIZIONE 7.3 *Le funzioni trigonometriche* $\sin kx$, $\cos kx$ *soddisfano le seguenti relazioni integrali (dette relazioni di* ortogonalità*), per ogni* $k, h = 1, 2, 3 \ldots$:

$$\int_0^{2\pi} (\sin kx)^2\, dx = \int_0^{2\pi} (\cos kx)^2\, dx = \pi$$

$$\int_0^{2\pi} \sin kx \sin hx\, dx = \int_0^{2\pi} \cos kx \cos hx\, dx = 0 \quad \text{se } h \neq k$$

$$\int_0^{2\pi} \sin kx \cos hx\, dx = \int_0^{2\pi} \sin kx\, dx = \int_0^{2\pi} \cos hx\, dx = 0.$$

La verifica di queste relazioni si basa sui metodi elementari di integrazione[19].

Il risultato di questa proposizione è facile da ricordare: il prodotto di due funzioni qualsiasi (diverse) tra quelle del sistema trigonometrico $\sin(kx)$, $\cos(hx)$ ha integrale nullo su un periodo; invece, il quadrato di una di queste ha per integrale la metà della lunghezza del periodo. Fa eccezione la costante 1 (formalmente, $\cos(0x)$), che ovviamente ha integrale 2π.

Il contenuto della proposizione precedente si può parafrasare dicendo che le funzioni:

$$\frac{1}{\sqrt{2\pi}}, \quad \frac{\cos kx}{\sqrt{\pi}}, \quad \frac{\sin kx}{\sqrt{\pi}} \qquad (k = 1, 2, \ldots) \tag{3.4}$$

costituiscono un *sistema ortonormale* nello spazio vettoriale V, ovvero queste funzioni sono "vettori" a due a due ortogonali, ciascuno di norma 1.

Un polinomio trigonometrico di grado $\leq n$ è una combinazione lineare dei $2n + 1$ vettori (3.4) per $k = 1, 2, \ldots, n$, ovvero è un generico elemento del sottospazio vettoriale V_n generato da questi $2n + 1$ vettori, che costituiscono una *base ortonormale* di V_n.

È immediato a questo punto interpretare il teorema della proiezione, discusso nel paragrafo precedente nel contesto di uno spazio vettoriale astratto, in termini del problema di approssimazione che ci interessa.

> Data una qualsiasi funzione $f \in V$, il polinomio trigonometrico di grado $\leq n$ che meglio approssima f, ovvero che rende minima la distanza $d(s_n, f)$ sopra definita, è la proiezione $S_n f$ di f su V_n.

[19] v. volume 1, capitolo 6, paragrafo 5.4.

© 978-88-08-06485-1

Poiché conosciamo una base ortonormale di V_n, è facile calcolare tale proiezione ortogonale. Se indichiamo sinteticamente con e_i i vettori della base di V_n $(i = 0, 1, 2, \dots 2n)$, tale proiezione sarà:

$$S_n f = \sum_{i=0}^{2n} \langle f, e_i \rangle \, e_i$$

Esplicitamente:

$$S_n f\,(x) = \left(\int_0^{2\pi} \frac{1}{\sqrt{2\pi}} f\,(t)\; dt\right) \frac{1}{\sqrt{2\pi}} + \\ + \sum_{k=1}^{n} \left[\left(\int_0^{2\pi} \frac{\cos kt}{\sqrt{\pi}} f\,(t)\; dt\right) \frac{\cos kx}{\sqrt{\pi}} + \left(\int_0^{2\pi} \frac{\sin kt}{\sqrt{\pi}} f\,(t)\; dt\right) \frac{\sin kx}{\sqrt{\pi}}\right] =$$

$$= \frac{1}{2} a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) \tag{3.5}$$

avendo posto:

$$a_k = \frac{1}{\pi} \int_0^{2\pi} f\,(t) \cos kt \; dt \quad \text{per } k = 0, 1, 2, \dots$$

$$b_k = \frac{1}{\pi} \int_0^{2\pi} f\,(t) \sin kt \; dt \quad \text{per } k = 1, 2, \dots \tag{3.6}$$

I coefficienti a_k, b_k, definiti dalle (3.6), si dicono *coefficienti di Fourier di* f; il polinomio trigonometrico $S_n f\,(x)$ si dice n-esima somma di Fourier di f e la serie trigonometrica

$$\frac{1}{2} a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx)$$

si dice *serie di Fourier di* f. Si scrive anche:

$$f \sim \frac{1}{2} a_0 + \sum_{k=1}^{\infty} (a_k \cos kx + b_k \sin kx) \tag{3.7}$$

per indicare che la serie trigonometrica a secondo membro è la serie di Fourier associata a f.

Il prossimo teorema raccoglie ciò che sappiamo finora su $S_n f$ in base al teorema della proiezione.

TEOREMA 7.10 *Sia $f \in V$, siano a_k, b_k i suoi coefficienti di Fourier assegnati da (3.6) e $S_n f\,(x)$ l'n-esima somma di Fourier di f, data da (3.5). Allora:*

© 978-88-08-06485-1

1. *La norma di $S_n f$ si calcola dai coefficienti di Fourier al seguente modo:*

$$\|S_n f\|^2 = \pi \left[\frac{a_0^2}{2} + \sum_{k=1}^{n} (a_k^2 + b_k^2)\right]. \tag{3.8}$$

2. *Vale la disuguaglianza di Bessel:*

$$\|S_n f\|^2 \leq \|f\|^2.$$

3. *La serie numerica*

$$\sum_{k=1}^{\infty} (a_k^2 + b_k^2)$$

converge.

4. *Vale il Lemma di Riemann-Lebesgue: per $k \to \infty$*

$$a_k \to 0; \quad b_k \to 0.$$

5. *$S_n f$ è, tra i polinomi trigonometrici di grado $\leq n$, quello che rende minima la distanza da f, cioè la quantità $\|f - p_n\|$ (al variare di p_n in V_n).*

6. *Lo scarto quadratico medio tra f e $S_n f$ si calcola al seguente modo:*

$$\int_0^{2\pi} [f(x) - S_n f(x)]^2 dx = \int_0^{2\pi} f(x)^2 dx - \pi \left[\frac{a_0^2}{2} + \sum_{k=1}^{n} (a_k^2 + b_k^2)\right].$$

DIMOSTRAZIONE.

1. Segue dal teorema di Pitagora, sfruttando l'ortogonalità delle funzioni trigonometriche e la loro norma:

$$\|S_n f\|^2 = \left\| \frac{1}{2} a_0 + \sum_{k=1}^{n} (a_k \cos kx + b_k \sin kx) \right\|^2 = \\ = \left(\frac{1}{2} a_0\right)^2 \|1\|^2 + \sum_{k=1}^{n} \left(a_k^2 \|\cos kx\|^2 + b_k^2 \|\sin kx\|^2\right) = \\ = \frac{a_0^2}{4} \cdot 2\pi + \pi \sum_{k=1}^{n} (a_k^2 + b_k^2),$$

che è la (3.8).

2. È il punto (ii) del teorema 7.9, nella nostra situazione.

3. Segue dai due punti precedenti, passando al limite per $n \to \infty$. (Ricordare che una serie a termini non negativi con somme parziali limitate è convergente[20]).

4. Segue dal punto 3, perché il termine generale di una serie convergente è infinitesimo.

[20] v. volume 1, capitolo 5, paragrafo 1.2.

5. È il punto (iii) del teorema 7.9, nella nostra situazione.

6. $$\|f - S_n f\|^2 = \langle f - S_n f, f - S_n f\rangle = \langle f, f\rangle - 2\langle f, S_n f\rangle + \langle S_n f, S_n f\rangle =$$
$$= \|f\|^2 - 2\langle f, S_n f\rangle + \|S_n f\|^2 .$$

Ora osserviamo che, poiché $f - S_n f$ è ortogonale a tutti gli elementi di V_n, quindi in particolare a $S_n f$:

$$0 = \langle f - S_n f, S_n f\rangle = \langle f, S_n f\rangle - \langle S_n f, S_n f\rangle$$

da cui $\langle f, S_n f\rangle = \langle S_n f, S_n f\rangle = \|S_n f\|^2$. Perciò:

$$\|f - S_n f\|^2 = \|f\|^2 - 2\langle f, S_n f\rangle + \|S_n f\|^2 =$$
$$= \|f\|^2 - 2\|S_n f\|^2 + \|S_n f\|^2 = \|f\|^2 - \|S_n f\|^2$$

che è la tesi, tenendo conto del punto 1. ♦

Il teorema precedente descrive esaurientemente le proprietà della somma parziale $S_n f(x)$, per ogni n fissato. Ora, l'intuizione geometrica dice che, al crescere di n, la somma parziale $S_n f(x)$ della serie di Fourier approssima sempre meglio la funzione f (che viene ricostruita attraverso un numero crescente di sue componenti ortogonali). La domanda è: questa approssimazione migliora indefinitamente, fino a potersi affermare che la serie di Fourier associata a f rappresenta f? Ossia, per $n \to \infty$, è vero che $S_n f \to f$ in qualche senso? Notiamo che il teorema precedente non contiene questa affermazione. Vale tuttavia il seguente risultato, che non dimostreremo:

TEOREMA 7.11 (CONVERGENZA DELLE SERIE DI FOURIER IN NORMA QUADRATICA) *Sia $f \in V$, siano a_k, b_k i suoi coefficienti di Fourier assegnati da (3.6) e $S_n f(x)$ l'n-esima somma di Fourier di f, data da (3.5). Allora:*

1.

$$\|f - S_n f\| \to 0 \text{ per } n \to \infty. \tag{3.9}$$

Più esplicitamente, questo significa che, per $n \to \infty$,

$$\int_0^{2\pi} [f(x) - S_n f(x)]^2 dx = \int_0^{2\pi} f(x)^2 dx - \pi\left[\frac{a_0^2}{2} + \sum_{k=1}^{n}(a_k^2 + b_k^2)\right] \to 0. \tag{3.10}$$

2. Vale l'uguaglianza di Parseval:

$$\int_0^{2\pi} f(x)^2 dx = \pi\left[\frac{a_0^2}{2} + \sum_{k=1}^{\infty}(a_k^2 + b_k^2)\right]. \tag{3.11}$$

Si osservi che l'uguaglianza di Parseval segue dalla (3.10), mentre l'uguaglianza nella (3.10) è il punto 6 del teorema 7.10; il teorema dipende quindi unicamente dalla relazione (3.9) di convergenza.

Il contenuto del teorema si può esprimere dicendo che $S_n f \to f$ *in media quadratica*, o anche che *la funzione f è sviluppabile in serie di Fourier in media quadratica*; questo significa che f si può ottenere come somma della sua serie di Fourier, pur di intepretare la convergenza della serie nel senso appena spiegato. In particolare, non è detto che $f(x)$ sia uguale per *ogni* x alla somma della serie: ci occuperemò nel paragrafo 3.6 del problema della *convergenza puntuale*.

Enunciamo anche il seguente risultato, che costituisce un passo importante della dimostrazione del teorema precedente e ha anche un significato indipendente:

TEOREMA 7.12 *Sia $f : [0, 2\pi] \to \mathbb{R}$ limitata e integrabile e supponiamo che i coefficienti di Fourier di f siano tutti nulli. Allora f è identicamente nulla, tranne al più su un insieme di misura nulla*[21]. *Se f è anche continua, si può invece affermare che f è zero in tutti i punti.*

Per linearità, questo implica anche che se due funzioni limitate e integrabili hanno gli stessi coefficienti di Fourier, allora coincidono, salvo eventualmente un insieme di misura nulla; se inoltre le funzioni sono continue, coincidono in tutto $[0, 2\pi]$.

Questo risultato dice che i coefficienti di Fourier *individuano univocamente* la funzione di partenza, anche se ancora non sappiamo se la serie di Fourier converge o no, e in quale senso. Il teorema precedente viene citato come risultato di *completezza del sistema trigonometrico*; infatti, si può interpretare anche dicendo che il sistema trigonometrico è sufficiente a "catturare tutte le informazioni di una funzione". È chiaro inoltre che se la successione $S_n f$ tende a una funzione g, per esempio in media quadratica, i coefficienti di Fourier della funzione limite saranno uguali a quelli di f, con cui è costruita la successione $S_n f$. In base al teorema 7.12, questo in particolare implica che $g = f$, ossia: *se la serie di Fourier di f converge, allora converge proprio a f*. (Si noti che questa affermazione non è ovvia: come abbiamo visto nel paragrafo 2.2, al contrario, la *serie di MacLaurin* di f talvolta converge a una funzione diversa da f). Pertanto, per dimostrare il teorema 7.11 è sufficiente provare che la serie $S_n f$ converge in norma a una certa funzione $g \in V$. Questo tuttavia va al di là della strumentazione matematica in nostro possesso in questo corso, mentre diventa un risultato relativamente semplice nel contesto della teoria dell'*integrale di Lebesgue*.

3.4 Esempi e osservazioni sul calcolo dei coefficienti di Fourier

Premettiamo il seguente

LEMMA 7.13 (INTEGRALI SU PERIODI EQUIVALENTI PER FUNZIONI PERIODICHE) *Se $f : \mathbb{R} \to \mathbb{R}$ è una funzione periodica di periodo $T > 0$, allora per ogni $a \in \mathbb{R}$ è:*

$$\int_a^{a+T} f(x)\,dx = \int_0^T f(x)\,dx.$$

[21] Ricordiamo che un insieme $E \subset \mathbb{R}$ ha misura nulla se per ogni $\varepsilon > 0$ esiste una successione di intervalli $\{I_n\} \subset \mathbb{R}$ la cui unione contiene E e la cui lunghezza complessiva è $\sum_{n=1}^{\infty} |I_n| < \varepsilon$. Questa nozione è analoga a quella di insieme di misura nulla nel piano, che abbiamo definito nel contesto della teoria degli integrali doppi (v. capitolo 5, paragrafo 1.2).

 © 978-88-08-06485-1

DIMOSTRAZIONE. Supponiamo prima $a \in (0,T)$. Allora:

$$\int_a^{a+T} f(x)\,dx = \int_a^T f(x)\,dx + \int_T^{a+T} f(x)\,dx =$$

eseguendo nel secondo integrale la sostituzione $x = T + t$

$$= \int_a^T f(x)\,dx + \int_0^a f(T+t)\,dt =$$

poiché f è T-periodica

$$= \int_a^T f(x)\,dx + \int_0^a f(t)\,dx = \int_0^T f(x)\,dx,$$

per l'additività dell'integrale rispetto all'intervallo di integrazione.

Sia ora a qualsiasi. Per un certo intero k, sarà $a \in [kT, (k+1)T)$. Eseguendo nell'integrale la sostituzione $x = t + kT$, si ha:

$$\int_a^{a+T} f(x)\,dx = \int_{a-kT}^{a-kT+T} f(t+kT)\,dt =$$

per la periodicità di f

$$\int_{a-kT}^{a-kT+T} f(t)\,dt.$$

D'altro canto $a - kT \in [0,T)$, perciò per quanto sopra dimostrato questo integrale uguaglia:

$$\int_0^T f(t)\,dt. \qquad \blacklozenge$$

Il lemma implica che se f è una funzione definita su tutto $\mathbb{R}$ e 2π-periodica, i suoi coefficienti di Fourier si possono calcolare anche mediante gli integrali:

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos kx\,dx, \quad b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin kx\,dx$$

o su qualunque altro intervallo di ampiezza 2π. Questo fatto è utile anche in vista della prossima:

OSSERVAZIONE (COEFFICIENTI DI FOURIER DI FUNZIONI PARI O DISPARI) Supponiamo che la funzione $f(x)$ che vogliamo sviluppare in serie di Fourier sia 2π-periodica e simmetrica (pari o dispari), allora le funzioni $f(x)\cos hx$ e $f(x)\sin hx$ sono anch'esse simmetriche. Il calcolo dei coefficienti di Fourier si semplifica conseguentemente:
se f è dispari,

$$\begin{aligned} a_h &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\cos hx\,dx = 0; \\ b_h &= \frac{1}{\pi}\int_{-\pi}^{\pi} f(x)\sin hx\,dx = \frac{2}{\pi}\int_0^{\pi} f(x)\sin hx\,dx. \end{aligned} \tag{3.12}$$

Analogamente, se f è pari,

$$a_h = \frac{2}{\pi}\int_0^{\pi} f(x)\cos hx\,dx \qquad b_h = 0; \tag{3.13}$$

© 978-88-08-06485-1

Esempio

Calcoliamo i coefficienti di Fourier dell'*onda quadra*:

$$f(x) = \begin{cases} A & \text{per } 0 < x < \pi \\ -A & \text{per } -\pi < x < 0 \end{cases} \qquad \text{prolungata periodicamente.}$$

Per definizione si ha:

$$a_k = \frac{1}{\pi}\int_{-\pi}^{\pi} A\,\mathrm{sgn}(x)\cos kx\,dx = 0$$

(integrale di una funzione dispari su un intervallo simmetrico);

$$b_k = \frac{1}{\pi}\int_{-\pi}^{\pi} A\,sgn(x)\sin kx\,dx = \frac{2}{\pi}\int_0^{\pi} A\sin kx\,dx =$$

$$= \frac{2A}{\pi}\left[-\frac{\cos kx}{k}\right]_0^{\pi} = \frac{2A}{\pi k}(1-(-1)^k) = \begin{cases} 0 & \text{per } k = 2n \\ \dfrac{4A}{\pi(2n+1)} & \text{per } k = 2n+1 \end{cases}$$

Dunque:

$$f(x) \sim \sum_{n=0}^{\infty} \frac{4A}{\pi(2n+1)}\sin(2n+1)x.$$

A titolo d'esempio, tracciamo il grafico di f insieme a quello della somma n-esima della sua serie di Fourier, per qualche valore di n:

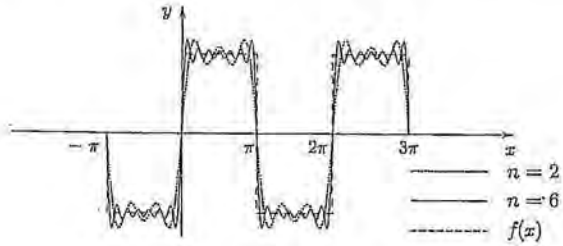


Figura 7.7. Grafico di $f(x)$ e di $s_n(x) = \sum_{k=0}^{n} \frac{4A}{\pi(2k+1)}\sin(2k+1)x$ per $n = 2$ e $n = 6$; qui $A = 1$.

Per evidenziare il fatto che l'approssimazione è nel senso dello scarto quadratico, osserviamo il grafico di $[f(x) - S_n f(x)]^2$ per $n = 6$. Si nota che l'integrale di questa funzione (cioè E_n) è molto piccolo (fig. 7.8).

Periodi diversi da 2π

Naturalmente, l'intervallo $[0, 2\pi]$ non ha niente di speciale rispetto a qualunque altro intervallo della retta. Una funzione definita, per esempio, sull'intervallo $[0,T]$ si può periodizzare definendola su tutto $\mathbb{R}$ come funzione periodica di periodo T. Il sistema di funzioni trigonometriche con periodo (multiplo di) T è

$$1; \quad \sin\left(\frac{2\pi kx}{T}\right); \quad \cos\left(\frac{2\pi kx}{T}\right) \qquad k = 1, 2, 3\ldots \tag{3.14}$$

 © 978-88-08-06485-1

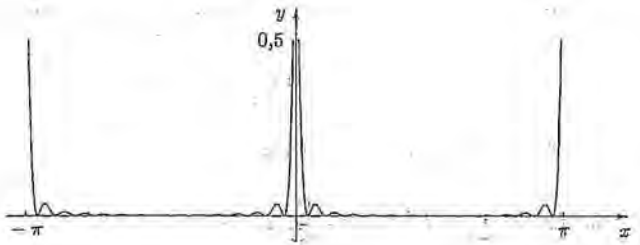


Figura 7.8.

I risultati della proposizione 7.3 rimangono veri nella forma sintetica in cui sono stati enunciati: l'integrale su $(0,T)$ di due funzioni diverse, tra le (3.14), è nullo; l'integrale su $(0,T)$ del quadrato di una di esse vale $T/2$ (tranne per la costante 1 che ha integrale T).

Le formule (3.6) per il calcolo dei coefficienti di Fourier vanno sostituite corrispondentemente da:

$$a_k = \frac{2}{T}\int_0^T f(x)\cos(\omega k x)\,dx \text{ per ogni } k \geq 0 \tag{3.15}$$

$$b_k = \frac{2}{T}\int_0^T f(x)\sin(\omega k x)\,dx \text{ per ogni } k \geq 1 \tag{3.16}$$

dove si è posto come di consueto $\omega = 2\pi/T$; la serie di Fourier associata a f ha ora la forma:

$$f(x) \sim \frac{a_0}{2} + \sum_{k=1}^{\infty} a_k \cos(\omega k x) + b_k \sin(\omega k x).$$

L'uguaglianza di Parseval (3.11) assume la forma:

$$\int_0^T f(x)^2 dx = \frac{T}{2}\left[\frac{a_0^2}{2} + \sum_{k=1}^{\infty}(a_k^2 + b_k^2)\right] \tag{3.17}$$

Esempio

Sia

$$f(x) = \text{mant}(x) = x - [x]$$

(dove $[x]$ è la parte intera). La funzione mantissa di x si può definire anche come:

$$f(x) = \begin{cases} x \text{ per } x \in [0,1] \\ \text{prolungata in } \mathbb{R} \text{ periodica di periodo 1.} \end{cases}$$

© 978-88-08-06485-1

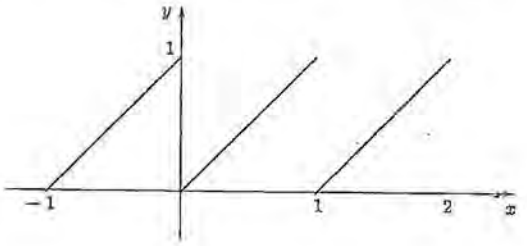


Figura 7.9.

Calcoliamo i suoi coefficienti di Fourier su $[0,1]$. Con semplici integrazioni per parti si ottiene:

$$a_h = 2\int_0^1 x\cos(2\pi h x)\,dx = 0 \text{ per } h > 0;\ a_0 = 1$$

$$b_h = 2\int_0^1 x\sin(2\pi h x)\,dx = -\frac{1}{\pi h}.$$

Quindi:

$$x \sim \frac{1}{2} - \frac{1}{\pi}\sum_{h=1}^{\infty}\frac{\sin 2h\pi x}{h} \text{ in } [0,1].$$

Il grafico di $f(x)$ insieme alla somma parziale della sua serie di Fourier, fino a $n = 6$, è rappresentato in figura 7.10.

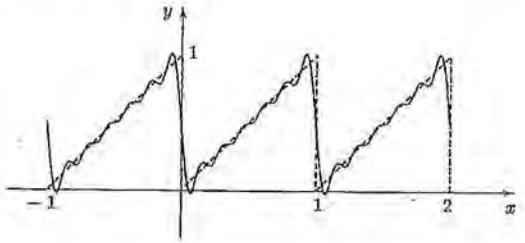


Figura 7.10.

3.5 Forma esponenziale complessa delle serie di Fourier

L'identità di Eulero

$$e^{inx} = \cos nx + i\sin nx$$

suggerisce che una serie di Fourier possa esprimersi in forma più compatta utilizzando gli esponenziali complessi e^{inx}, anziché le funzioni trigonometriche $\cos nx, \sin nx$. Effettivamente è così e questo può essere utile da vari punti di vista, sia teorici sia pratici.

Per capire che forma assumano le formule precedenti in questo contesto, consideriamo ora lo spazio vettoriale V^* delle funzioni $f : [0, 2\pi] \to \mathbb{C}$ integrabili[22]. Per $f, g \in V^*$, definiamo

$$\langle f, g \rangle = \int_0^{2\pi} f(t)\,\overline{g(t)}dt$$

(dove $\overline{g(t)}$ indica il complesso coniugato di $g(t)$). Lo spazio V^* risulta uno spazio vettoriale *complesso* munito di prodotto scalare; questo, per definizione, significa che $\langle \cdot, \cdot \rangle$ soddisfa ora le seguenti proprietà[23], come si verifica immediatamente:

$$\langle u, v \rangle = \overline{\langle v, u \rangle},$$

che in particolare implica che $\langle u, u \rangle$ sia reale;

$$\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle,$$
$$\langle \lambda u, v \rangle = \lambda \langle u, v \rangle \text{ per ogni } \lambda \in \mathbb{C}$$

che, insieme alla prima, implica anche

$$\langle u, \lambda v \rangle = \overline{\lambda} \langle u, v \rangle \text{ per ogni } \lambda \in \mathbb{C};$$

e infine

$$\langle u, u \rangle \geq 0 \text{ e } \langle u, u \rangle = 0 \text{ se e solo se } u = 0.$$

L'ultima proprietà implica in particolare che si può ancora definire una norma ponendo:

$$\|u\| = \sqrt{\langle u, u \rangle}.$$

Tornando allo spazio di funzioni V^*, abbiamo:

$$\|f\| = \left(\int_0^{2\pi} f(t)\,\overline{f(t)}dt \right)^{1/2} = \left(\int_0^{2\pi} |f(t)|^2\,dt \right)^{1/2}.$$

Consideriamo ora le funzioni esponenziali complesse e^{inx} per $n \in \mathbb{Z}$ (quindi ora n può essere anche un intero negativo). Ricordando che $\overline{(e^{imx})} = e^{-imx}$, si vede che queste soddisfano le seguenti relazioni di ortogonalità, per $n, m \in \mathbb{Z}$:

$$\langle e^{inx}, e^{imx} \rangle = \int_0^{2\pi} e^{inx} e^{-imx} dx =$$
$$= \int_0^{2\pi} \cos(n-m)x\,dx + i \int_0^{2\pi} \sin(n-m)x\,dx = \begin{cases} 2\pi & \text{se } n = m \\ 0 & \text{se } n \neq m. \end{cases}$$

[22] Ossia tali che le loro parti reale e immaginaria siano integrabili. Se $f = f_1 + if_2$ con $f_1, f_2 : [0, 2\pi] \to \mathbb{R}$, porremo per definizione:

$$\int_0^{2\pi} f(t)\,dt = \int_0^{2\pi} f_1(t)\,dt + i \int_0^{2\pi} f_2(t)\,dt.$$

[23] Queste proprietà sono leggermente diverse da quelle discusse nel paragrafo 3.3 per gli spazi vettoriali reali muniti di prodotto scalare e si riducono a quelle nel caso in cui le funzioni e gli scalari che utilizziamo siano reali.

Indicando ora con V_n^* (con $n \geq 1$) lo spazio vettoriale che ha per base ortonormale

$$\frac{e^{ikx}}{\sqrt{2\pi}} \text{ per } k = 0, \pm 1, \pm 2, \ldots, \pm n$$

vediamo che la proiezione di f su V_n^* è data da:

$$S_n f(x) = \sum_{k=-n}^{n} \left\langle f, \frac{e^{-ikx}}{\sqrt{2\pi}} \right\rangle \frac{e^{ikx}}{\sqrt{2\pi}} =$$
$$= \sum_{k=-n}^{n} \left(\frac{1}{2\pi} \int_0^{2\pi} f(t)\,e^{-ikt}dt \right) e^{ikx} = \sum_{k=-n}^{n} \widehat{f}(k)\,e^{ikx},$$

avendo posto

$$\widehat{f}(k) = \frac{1}{2\pi} \int_0^{2\pi} f(t)\,e^{-ikt}dt \text{ per } k \in \mathbb{Z}.$$

La serie di Fourier di f in forma complessa è perciò

$$f \sim \sum_{k=-\infty}^{+\infty} \widehat{f}(k)\,e^{ikx} \tag{3.18}$$

Con le notazioni complesse abbiamo quindi un'unica successione $\widehat{f}(k)$ di coefficienti di Fourier di f, anziché due successioni a_k, b_k. Questa successione, in compenso, è indiciata sugli interi relativi anziché sui soli interi positivi.

Si noti che, per definizione, la somma n-esima di una serie del tipo $\sum_{k=-\infty}^{+\infty}$ (detta *serie bilatera*) è la somma $\sum_{k=-n}^{n}$, per n intero positivo.

Notiamo anche che i coefficienti di Fourier $\widehat{f}(k)$ assumeranno in generale valori complessi, *anche se la funzione di partenza è reale*. Confrontando la (3.18) con le (3.6) si vede subito che valgono le seguenti relazioni, che permettono di passare dai coefficienti calcolati in forma trigonometrica a quelli calcolati in forma esponenziale complessa e viceversa:

per $k = 1, 2, 3 \ldots$.

$\widehat{f}(k) = \frac{1}{2}(a_k - ib_k)$	$a_k = \widehat{f}(k) + \widehat{f}(-k)$
$\widehat{f}(-k) = \frac{1}{2}(a_k + ib_k)$	$b_k = i(\widehat{f}(k) - \widehat{f}(-k))$
$\widehat{f}(0) = \frac{1}{2}a_0$	$a_0 = 2\widehat{f}(0)$.

Per esempio, per una funzione f a valori reali e simmetrica pari, risulta $b_k = 0$ e a_k reale, per cui $\widehat{f}(k)$ sono reali; per una funzione f a valori reali e simmetrica dispari, $\widehat{f}(k)$ risulteranno immaginari puri; per una funzione reale non simmetrica, $\widehat{f}(k)$ saranno in generale complessi.

Scriviamo esplicitamente la forma che assumono con notazioni complesse le relazioni notevoli che abbiamo visto nei teoremi 7.10, 7.11:

$$\|S_n f\|^2 = 2\pi \sum_{k=-n}^{n} \left|\widehat{f}(k)\right|^2 \leq \|f\|^2 \text{ (disuguaglianza di Bessel)}$$

$$\|f\|^2 = 2\pi \sum_{k=-\infty}^{\infty} \left|\widehat{f}(k)\right|^2 \text{ (uguaglianza di Parseval)}$$

$$\widehat{f}(n) \to 0 \text{ per } n \to \pm\infty \text{ (lemma di Riemann-Lebesgue).}$$

Anche per le serie di Fourier in forma esponenziale, tutto può essere adattato a un intervallo diverso da $[0, 2\pi]$.

Riportiamo, per comodità del lettore, la forma assunta dalle precedenti relazioni per funzioni T-periodiche, ovvero definite su $[0, T]$ e periodizzate.

Sistema ortonormale adattato a $[0, T]$:

$$\frac{e^{i\omega kx}}{\sqrt{T}} \text{ per } k \in \mathbb{Z}, \text{ con } \omega = 2\pi/T.$$

Coefficienti di Fourier:

$$\widehat{f}(k) = \frac{1}{T}\int_0^T f(t)\, e^{-i\omega kt} dt \text{ per } k \in \mathbb{Z}.$$

Serie di Fourier:

$$f \sim \sum_{k=-\infty}^{+\infty} \widehat{f}(k)\, e^{i\omega kx}.$$

Uguaglianza di Parseval:

$$\|f\|^2 = T \sum_{k=-\infty}^{\infty} \left|\widehat{f}(k)\right|^2.$$

3.6 Convergenza puntuale delle serie di Fourier

Abbiamo visto che la successione $S_n f$ delle somme di Fourier di f converge a f in media quadratica, ossia in un opportuno senso integrale. È importante stabilire se la convergenza si abbia anche in senso puntuale. In questo paragrafo enunceremo un paio di risultati positivi sulla convergenza puntuale delle serie di Fourier. Vedremo che, sotto ipotesi di sufficiente regolarità di $f(x)$, si potrà garantire che la serie di Fourier converge, e converge proprio a $f(x)$; inoltre, sotto qualche ipotesi ulteriore, sarà possibile anche derivare la serie termine a termine.

Per enunciare un primo classico risultato di convergenza puntuale delle serie di Fourier, premettiamo la seguente:

DEFINIZIONE 7.6 Si dice che $f : [0, T] \to \mathbb{R}$ è *regolare a tratti* se $f : [0, T] \to \mathbb{R}$ è limitata in $[0, T]$ e l'intervallo si può decomporre in un numero finito di intervallini su ciascuno dei quali la funzione è continua e derivabile; inoltre agli estremi di ciascun intervallino esistono finiti i limiti sia di $f(x)$ che di $f'(x)$. ▩

In particolare, le ipotesi di questo teorema ammettono che f possieda un numero finito di punti di discontinuità a salto e di punti angolosi, mentre non sono ammessi asintoti verticali né punti a tangente verticale, come cuspidi. Ne segue in particolare che:

la funzione f è integrabile, perciò i coefficienti di Fourier sono ben definiti;
in ogni punto x_0 esistono $\lim_{x\to x_0^+} f(x) \equiv f(x_0^+)$ e $\lim_{x\to x_0^-} f(x) \equiv f(x_0^-)$.

TEOREMA 7.14 (CONVERGENZA PUNTUALE, FUNZIONI REGOLARI A TRATTI) *Se $f : [0, T] \to \mathbb{R}$ è regolare a tratti, allora la serie di Fourier converge in ogni punto $x_0 \in (0, T)$ alla media dei due limiti destro e sinistro:*

$$\frac{a_0}{2} + \sum_{k=1}^{\infty} \{a_k \cos(\omega k x_0) + b_k \sin(\omega k x_0)\} = \frac{f(x_0^+) + f(x_0^-)}{2}$$

con $\omega = 2\pi/T$; nei due estremi dell'intervallo 0 e T la serie converge invece a $\frac{f(0^+)+f(T^-)}{2}$.

In particolare, in ogni punto di $(0, T)$ in cui f è continua, la serie di Fourier converge a $f(x)$; negli estremi questo è vero solo se $f(0) = f(T)$.

DIMOSTRAZIONE. Per semplificare le notazioni, supponiamo che il periodo sia $T = 1$. Inoltre, per la dimostrazione è più comodo usare la scrittura complessa della serie di Fourier. Perciò, posto

$$\widehat{f}(k) = \int_{-1/2}^{1/2} f(x)\, e^{-2\pi i k x} dx \text{ per } k \in \mathbb{Z},$$

la nostra tesi è:

$$S_n f(x_0) \equiv \sum_{k=-n}^{n} \widehat{f}(k)\, e^{2\pi i k x_0} \to \frac{f(x_0^+) + f(x_0^-)}{2} \text{ per } n \to +\infty.$$

Eseguendo eventualmente una traslazione sull'asse x e una sull'asse y, possiamo ricondurci al caso in cui

$$x_0 = 0 \text{ e } f(0^+) = -f(0^-). \tag{3.19}$$

Sotto queste ipotesi, la tesi si riscrive allora come:

$$S_n f(0) \to 0 \text{ per } n \to +\infty.$$

Si ha:

$$S_n f(0) = \sum_{k=-n}^{n} \widehat{f}(k) = \sum_{k=-n}^{n} \int_{-1/2}^{1/2} f(x)\, e^{-2\pi i k x} dx. \tag{3.20}$$

Osservando che

$$\sum_{k=-n}^{n} e^{2\pi i k x} = \sum_{k=-n}^{n} e^{-2\pi i k x},$$

la (3.20) si può riscrivere come

$$S_n f(0) = \frac{1}{2}\left\{\int_{-1/2}^{1/2} f(x) \sum_{k=-n}^{n} e^{-2\pi i k x} dx + \int_{-1/2}^{1/2} f(x) \sum_{k=-n}^{n} e^{2\pi i k x} dx\right\} =$$

eseguendo nel secondo integrale il cambio di variabile $x \mapsto -x$

$$= \frac{1}{2}\left\{\int_{-1/2}^{1/2} f(x) \sum_{k=-n}^{n} e^{-2\pi i k x} dx + \int_{-1/2}^{1/2} f(-x) \sum_{k=-n}^{n} e^{-2\pi i k x} dx\right\} =$$

$$= \int_{-1/2}^{1/2} \frac{f(x)+f(-x)}{2} \sum_{k=-n}^{n} e^{-2\pi i k x} dx =$$

$$= \sum_{k=-n}^{n} \int_{-1/2}^{1/2} \frac{f(x)+f(-x)}{2} e^{-2\pi i k x} dx = \sum_{k=-n}^{n} \widehat{g}(k) = S_n g(0),$$

avendo posto

$$g(x) = \frac{f(x)+f(-x)}{2}.$$

Notiamo che la funzione g è anch'essa regolare a tratti, inoltre nell'origine è continua (a differenza di f) e vale zero, per la (3.19). Proveremo ora che:

$$S_n g(0) \to 0 \quad \text{per } n \to +\infty. \tag{3.21}$$

Definiamo la funzione:

$$h(x) = \frac{g(x)}{e^{2\pi i x} - 1}.$$

La funzione h è certamente regolare a tratti in $\left[-\frac{1}{2}, \frac{1}{2}\right]$, tranne al più il punto $x = 0$, che ora studiamo. Per $x \to 0$,

$$h(x) \sim \frac{g(x)}{2\pi i x} \to \begin{cases} \dfrac{g'(0^+)}{2\pi i} & \text{per } x \to 0^+ \\ \dfrac{g'(0^-)}{2\pi i} & \text{per } x \to 0^- \end{cases}$$

dove abbiamo applicato il teorema di De L'Hospital e l'ipotesi di regolarità a tratti di f (quindi di g), che assicura l'esistenza dei limiti finiti $g'(0^+)$, $g'(0^-)$. Poiché h è regolare a tratti fuori dall'origine e limitata anche in un intorno di $x = 0$, in particolare è limitata e integrabile in $\left[-\frac{1}{2}, \frac{1}{2}\right]$ (si noti che invece potrebbe non essere regolare a tratti in tutto $\left[-\frac{1}{2}, \frac{1}{2}\right]$). Per il lemma di Riemann-Lebesgue (v. teorema 7.10) sappiamo allora che

$$\widehat{h}(k) \to 0 \text{ per } k \to \pm\infty. \tag{3.22}$$

Sfruttando questa informazione, dimostriamo ora la (3.21). Esprimiamo anzitutto i coefficienti $\widehat{g}(k)$ in termini di $\widehat{h}(k)$

$$\widehat{g}(k) = \int_{-1/2}^{1/2} g(x) e^{-2\pi i k x} dx = \int_{-1/2}^{1/2} h(x) \left(e^{2\pi i x} - 1\right) e^{-2\pi i k x} dx =$$

$$= \int_{-1/2}^{1/2} h(x) e^{-2\pi i (k-1) x} dx - \int_{-1/2}^{1/2} h(x) e^{-2\pi i k x} dx =$$

$$= \widehat{h}(k-1) - \widehat{h}(k).$$

Perciò

$$S_n g(0) = \sum_{k=-n}^{n} \widehat{g}(k) = \sum_{k=-n}^{n} \left[\widehat{h}(k-1) - \widehat{h}(k)\right] =$$

$$= \left[\widehat{h}(-n-1) - \widehat{h}(-n)\right] + \left[\widehat{h}(-n) - \widehat{h}(-n+1)\right] + \cdots + \left[\widehat{h}(n-1) - \widehat{h}(n)\right] =$$

(si tratta di una *somma telescopica*[24], in cui gli addendi si cancellano due a due)

$$= \widehat{h}(-n-1) - \widehat{h}(n) \to 0 \text{ per } n \to +\infty, \text{ per (3.22).}$$

Questo conclude la dimostrazione. ∎

Un altro criterio di convergenza puntuale, che non dimostriamo, è il seguente:

TEOREMA 7.15 (CONVERGENZA PUNTUALE, FUNZIONI MONOTONE A TRATTI) *Sia $f : [0,T] \to \mathbb{R}$ limitata in $[0,T]$ e monotona a tratti (ossia l'intervallo si può decomporre in un numero finito di intervallini su ciascuno dei quali la funzione è crescente oppure decrescente). Allora vale la stessa conclusione del teorema precedente.*

Dai due teoremi precedenti segue in particolare che:

Se f è continua e monotona in tutto $[0,T]$ oppure continua in tutto $[0,T]$ e regolare a tratti, la serie di Fourier converge puntualmente a f in ogni punto di $(0,T)$, converge anche agli estremi se vale la condizione di raccordo $f(0) = f(T)$ (che ha il significato di garantire la continuità in $\mathbb{R}$ della funzione periodizzata).

Il risultato appena enunciato può essere raffinato in vari modi, indebolendone le ipotesi; segnaliamo tuttavia che *la sola continuità di f (anche unita alla condizione di raccordo $f(0) = f(T)$) è insufficiente a garantire la convergenza puntuale della serie di Fourier in tutto l'intervallo.* Segnaliamo in proposito che esistono esempi (non facili da costruire) di funzioni continue la cui serie di Fourier diverge in un punto, o anche in molti punti; ciò che si può garantire in positivo, in conseguenza di teoremi profondi di analisi armonica, è che, data una funzione f continua e periodica in $\mathbb{R}$, l'insieme dei punti in cui la serie di Fourier di f non converge *ha misura nulla*[25].

Derivabilità termine a termine delle serie di Fourier

Occupiamoci ora del problema di derivare termine a termine una serie di Fourier. L'importanza di poter effettuare questa operazione sarà mostrata, per esempio, nelle applicazioni alle equazioni differenziali, che discuteremo nel paragrafo 3.8. Sia

$$f(x) \sim \frac{a_0}{2} + \sum_{h=1}^{\infty} \{a_h \cos hx + b_h \sin hx\}$$

Abbiamo visto nel paragrafo 3.1 che, se

$$\sum_{h=1}^{\infty} h|a_h|, \qquad \sum_{h=1}^{\infty} h|b_h| < \infty$$

allora sia la serie di f che la serie derivata convergono totalmente, perciò la serie si può derivare termine a termine. Questo è stato stabilito da un'analisi diretta dei coefficienti

[24] v. volume 1, capitolo 5, paragrafo 1.1, Esempio 1.3.

[25] Ovvero, come già ricordato, per ogni $\varepsilon > 0$ tale insieme può essere ricoperto dall'unione di una successione I_n di intervalli di lunghezza complessiva minore di ε.

della serie. Spesso, invece, è utile poter affermare *a priori*, in base alle proprietà di regolarità di f, il fatto che la sua serie di Fourier si possa derivare termine a termine.

TEOREMA 7.16 (DERIVAZIONE TERMINE A TERMINE DELLE SERIE DI FOURIER) *Sia $f : [0,T] \to \mathbb{R}$, di classe $C^1[0,T]$ e sia f' regolare a tratti (per esempio, $f \in C^2[0,T]$); inoltre, sia $f(0) = f(T)$. Allora la serie di Fourier di f si può derivare termine a termine in $(0,T)$; se inoltre è anche $f'(0) = f'(T)$, la serie di Fourier di f si può derivare termine a termine in tutto $[0,T]$.*

Si noti che le condizioni di raccordo $f(0) = f(T)$ e $f'(0) = f'(T)$, unite all'ipotesi $f \in C^1[0,T]$ implicano che la *periodizzata di* f sia $C^1(\mathbb{R})$. Si presti attenzione a questo fatto generale:

Se una funzione f regolare in $[0,2\pi]$ viene periodizzata a tutto $\mathbb{R}$, non è detto che la sua periodizzata sia regolare in $\mathbb{R}$.

Ciò che occorre garantire per poter affermare l'ultima cosa sono appunto le condizioni di raccordo agli estremi (su f e sulla sua derivata). Il teorema di derivabilità termine a termine richiede che la periodizzata di f (e non solo f) sia sufficientemente regolare.

Grosso modo, per derivare termine a termine la serie di Fourier di f, è necessario che f possieda *due* derivate e che inoltre soddisfi almeno la condizione di raccordo continuo $f(0) = f(T)$.

DIMOSTRAZIONE. Applichiamo il teorema di convergenza puntuale delle serie di Fourier alla funzione f'; questo è lecito perché f' è regolare a tratti per ipotesi. Poiché f' è anche continua per ipotesi (perché $f \in C^1[0,T]$), la serie di Fourier di f' convergerà a $f'(x)$ per ogni $x \in (0,T)$:

$$f'(x) = \frac{a_0'}{2} + \sum_{k=1}^{\infty} \left(a_k' \cos k\omega x + b_k' \sin k\omega x\right) \tag{3.23}$$

con $\omega = 2\pi/T$, dove abbiamo indicato con a_k', b_k' i coefficienti di Fourier di f'. Calcoliamo ora questi coefficienti, per esprimerli in funzione di quelli di f. Sfruttando la condizione di raccordo $f(T) = f(0)$, si ha:

$$a_0' = \frac{2}{T}\int_0^T f'(x)\,dx = \frac{2}{T}\left[f(T) - f(0)\right] = 0$$

$$a_k' = \frac{2}{T}\int_0^T f'(x)\cos k\omega x\,dx = \frac{2}{T}\left\{\left[f(x)\cos k\omega x\right]_0^T + k\omega\int_0^T f(x)\sin k\omega x\,dx\right\} =$$

$$= \frac{2}{T}\left\{\left[f(T) - f(0)\right] + k\omega\int_0^T f(x)\sin k\omega x\,dx\right\} = k\omega b_k$$

$$b_k' = (\text{con calcoli analoghi}) = -k\omega a_k.$$

Perciò possiamo riscrivere lo sviluppo di f' nella forma:

$$f'(x) = \sum_{k=1}^{\infty}\left(k\omega b_k \cos k\omega x - k\omega a_k \sin k\omega x\right) =$$

$$= \sum_{k=1}^{\infty}\left(a_k \cos k\omega x + b_k \sin k\omega x\right)'$$

e questo significa appunto che la serie di Fourier di f si può derivare termine a termine in $(0,T)$. Osserviamo ora che, se vale anche l'ipotesi $f'(0) = f'(T)$, in base al teorema 7.15 (applicato a f') lo sviluppo (3.23) vale anche per $x = 0$ e $x = T$ e di conseguenza la nostra conclusione vale anche in questi punti. ∎

Si noti, nella dimostrazione precedente, l'importanza della condizione di raccordo $f(T) = f(0)$: se questa non è soddisfatta, non può valere la derivazione termine a termine.

Esempi

Consideriamo la funzione e^x su $[0,2\pi]$. Per il teorema 7.15 (o 7.14), la sua serie converge a e^x per ogni $x \in (0,2\pi)$; per $x = 0, 2\pi$, la serie converge alla media dei valori di f nei due estremi: $\frac{1}{2}(e^{2\pi}+1)$. Il teorema 7.16 non è applicabile perché non vale la condizione di raccordo $f(0) = f(2\pi)$.

Calcoliamo ora i coefficienti di Fourier:

$$a_h = \frac{1}{\pi}\int_0^{2\pi} e^x \cos hx\,dx.$$

$$\int_0^{2\pi} e^x \cos hx\,dx = (\text{per parti}) \quad \left[e^x \cos hx\right]_0^{2\pi} + \int_0^{2\pi} e^x h \sin hx\,dx =$$

$$= e^{2\pi} - 1 + h\{\left[e^x \sin hx\right]_0^{2\pi} - \int_0^{2\pi} e^x h \cos hx\,dx\} =$$

$$= e^{2\pi} - 1 - h^2\int_0^{2\pi} e^x \cos hx\,dx$$

da cui otteniamo:

$$\int_0^{2\pi} e^x \cos hx\,dx = \frac{e^{2\pi}-1}{1+h^2} \quad \text{e quindi} \quad a_h = \frac{1}{\pi}\cdot\frac{e^{2\pi}-1}{1+h^2}.$$

Con calcoli simili, si trova:

$$b_h = \frac{h}{\pi}\cdot\frac{1-e^{2\pi}}{1+h^2}.$$

Quindi:

$$e^x \sim \frac{e^{2\pi}-1}{2\pi} + \sum_{h=1}^{\infty}\left(\frac{1}{\pi}\cdot\frac{e^{2\pi}-1}{1+h^2}\cos hx + \frac{h}{\pi}\cdot\frac{1-e^{2\pi}}{1+h^2}\sin hx\right).$$

Notiamo che la serie dei coseni converge totalmente, mentre la serie dei seni converge solo semplicemente; quest'ultima, poi, certamente non può essere derivata termine a termine. (Lo studente si renda conto del perché di ogni affermazione fatta).

Facciamo tracciare da un computer il grafico della somma parziale fino a $n = 5$, insieme al grafico di e^x, rappresentati in figura 7.11.

Si nota che la massima imprecisione si ha agli estremi dell'intervallo: questo è naturale perché la funzione e^x non soddisfa la condizione $f(0) = f(2\pi)$, mentre la sua serie di Fourier sì. In altre parole, la serie di Fourier non può convergere totalmente, altrimenti rappresenterebbe una funzione continua su $\mathbb{R}$, mentre il prolungamento periodico della funzione base uguale a e^x su $[0,2\pi]$ è discontinua in tutti i punti $2k\pi$.

Riprendiamo l'esempio 3.7: abbiamo calcolato lo sviluppo di $\mathrm{mant}(x)$ su $[0,1]$. Per il teorema 7.15 (o 7.14), la serie converge a $\mathrm{mant}(x)$ per ogni $x \in (0,1)$; per $x = 0, 1$, la serie

 © 978-88-08-06485-1

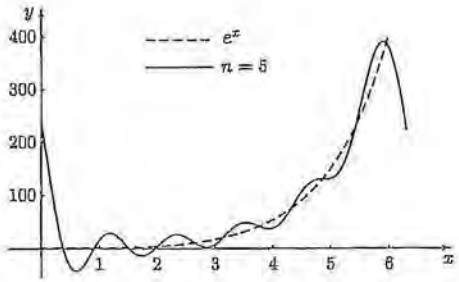


Figura 7.11.

converge alla media dei valori di f nei due estremi: $\frac{1}{2}$. Il grafico (fig. 7.10) conferma questi fatti. Di nuovo, il teorema 7.16 non è applicabile.

ESEMPIO Studiamo la sviluppabilità in serie di Fourier di $f(x) = x^2$ in $[-\pi, \pi]$.

La funzione è $C^2([-\pi, \pi])$, soddisfa la condizione $f(-\pi) = f(\pi)$, mentre non soddisfa la condizione $f'(-\pi) = f'(\pi)$. Perciò: la serie di Fourier di f converge a x^2 in tutto $[-\pi, \pi]$; la serie si può derivare termine a termine in tutto $(-\pi, \pi)$.

Data la simmetria pari di f, si ha:

$$b_h = 0$$

$$\frac{a_0}{2} = \frac{\pi^2}{3}$$

$$a_h = \frac{2}{\pi}\int_0^\pi x^2 \cos hx\, dx = \text{ (con due integrazioni per parti)}$$

$$= \frac{2}{\pi} \cdot \frac{2\pi}{h^2} \cos h\pi = (-1)^h \frac{4}{h^2}.$$

Quindi:

$$x^2 = \frac{\pi^2}{3} + \sum_{h=1}^{\infty} (-1)^h \frac{4}{h^2} \cos hx \quad \text{per ogni } x \in [-\pi, \pi] \tag{3.24}$$

Osserviamo il grafico di x^2 su $[-\pi, \pi]$ insieme alla somma parziale per $n = 3$ della sua serie di Fourier:

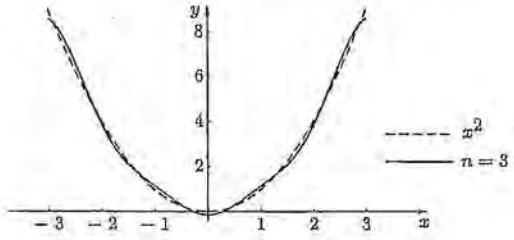


Figura 7.12.

© 978-88-08-06485-1

Come si vede, già per $n = 3$ l'approssimazione è buona.

Se nella (3.24), che vale per ogni $x \in [-\pi, \pi]$, sostituiamo $x = 0$, otteniamo:

$$0 = \frac{\pi^2}{3} + \sum_{h=1}^{\infty} (-1)^h \frac{4}{h^2}$$

da cui leggiamo la somma della serie numerica:

$$\sum_{h=1}^{\infty} (-1)^h \frac{1}{h^2} = -\frac{\pi^2}{12}. \tag{3.25}$$

Se invece sostituiamo $x = \pi$, otteniamo (poiché $\cos h\pi = (-1)^h$):

$$\pi^2 = \frac{\pi^2}{3} + \sum_{h=1}^{\infty} \frac{4}{h^2}$$

ossia:

$$\sum_{h=1}^{\infty} \frac{1}{h^2} = \frac{\pi^2}{6}. \tag{3.26}$$

Se deriviamo termine a termine la (3.24), come è lecito fare per ogni $x \in (-\pi, \pi)$, otteniamo:

$$2x = -\sum_{h=1}^{\infty} (-1)^h \frac{4}{h} \sin hx$$

ossia

$$x = -2\sum_{h=1}^{\infty} (-1)^h \frac{1}{h} \sin hx \quad \text{per ogni } x \in (-\pi, \pi). \tag{3.27}$$

Abbiamo quindi ottenuto lo sviluppo di Fourier di un'altra funzione. Osserviamo il grafico di $f'(x) = 2x$ insieme alla somma parziale per $n = 3$ della serie di Fourier derivata termine a termine:

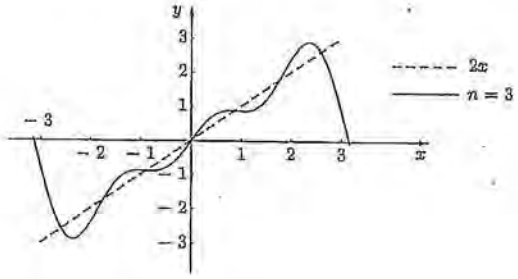


Figura 7.13.

OSSERVAZIONE Le identità (3.25)-(3.26) trovate sono un esempio significativo di serie numeriche la cui somma si riesce a calcolare in modo esatto vedendo quelle serie come serie di Fourier di funzioni note, calcolate in un punto fissato. È questo un fenomeno abbastanza tipico: le serie numeriche che si sanno sommare in modo esatto sono quasi sempre serie di funzioni calcolate in un punto particolare, come già visto a proposito delle serie di Taylor.

Velocità di convergenza a zero dei coefficienti di Fourier

Il fatto di poter derivare termine a termine una serie di Fourier ci dà informazioni sull'ordine di infinitesimo dei coefficienti di Fourier.

Se f è solo integrabile, sappiamo che $a_n, b_n \to 0$ (lemma di Riemann-Lebesgue, v. teorema 7.10, punto 4).

Se f è $C^1(\mathbb{R})$ (e T-periodica), allora f' è integrabile in $[0,T]$, quindi la serie di Fourier di f' converge in media quadratica. Ma, come abbiamo visto nella dimostrazione del teorema 7.16,

$$f'(x) \sim \sum_{k=1}^{\infty} (-k\omega a_k \sin k\omega x + k\omega b_k \cos k\omega x),$$

perciò il teorema 7.10 implica che:

$$\sum_{k=1}^{\infty} k^2(a_k^2 + b_k^2) < \infty$$

e in particolare $k^2(a_k^2 + b_k^2) \to 0$, da cui si legge che

$$a_k = o\left(\frac{1}{k}\right), \qquad b_k = o\left(\frac{1}{k}\right) \qquad \text{per } k \to \infty.$$

L'affermazione appena fatta si può raffinare: se f è continua in $\mathbb{R}$, T-periodica e regolare a tratti, è ancora vero che f' è integrabile in $[0,T]$ e vale ancora il passaggio di integrazione per parti necessario a mostrare che i coefficienti di Fourier di f' sono $k\omega b_k$ e $-k\omega a_k$. Di conseguenza rimane vera la nostra conclusione.

Questo ragionamento si può inoltre iterare, arrivando al seguente:

TEOREMA 7.17 *Sia f una funzione T-periodica. Per un certo intero $s \geq 1$ sia $f \in C^{s-1}(\mathbb{R})$ e $f^{(s-1)}$ sia regolare a tratti in $[0,T]$. Allora*

$$a_k = o\left(\frac{1}{k^s}\right) \qquad b_k = o\left(\frac{1}{k^s}\right) \qquad \textit{per } k \to \infty.$$

In sostanza: *maggiore è la regolarità di f, maggiore sarà la rapidità di convergenza a zero dei suoi coefficienti di Fourier.*

3.7 Alcune interpretazioni fisiche

Mostriamo ora brevemente alcune applicazioni e interpretazioni fisiche dei concetti fin qui introdotti sulle serie di Fourier.

Serie di Fourier e analisi armonica di un segnale periodico

L'operazione di scomporre una funzione periodica f di periodo T in somma (finita o infinita) di funzioni periodiche semplici del tipo $\cos n\omega t$ e $\sin n\omega t$ (con $\omega = 2\pi/T$) costituisce l'*analisi armonica* della funzione f: il termine

$$a_1 \cos \omega t + b_1 \sin \omega t$$

si chiama *armonica fondamentale*; il termine $a_n \cos n\omega t + b_n \sin n\omega t$ $(n = 2, 3, \ldots)$ si chiama *armonica n-esima*. Questa terminologia è derivata dall'acustica. Si noti che il primo coefficiente

$$\frac{a_0}{2} = \frac{1}{T}\int_0^T f(t)\,dt$$

è il *valore medio* di $f(t)$.

Quando un diapason, colpito, si mette a vibrare, le sue vibrazioni possono essere descritte da una funzione periodica semplice come

$$x(t) = a\cos\omega t \quad \text{oppure} \quad a\sin\omega t$$

($x(t)$ misura lo spostamento, al tempo t, di un corno del diapason dalla sua posizione di riposo).

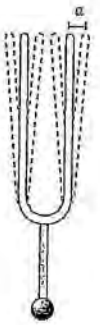


Figura 7.14.

Il numero positivo a misura l'*ampiezza* dell'oscillazione (spostamento massimo) e viene percepito dall'orecchio come *intensità* o *volume di suono*; $1/T = \omega/2\pi$ è la *frequenza* tipica del diapason: essa viene percepita dall'orecchio come *altezza* del suono. Il diapason ha la proprietà di emettere suoni *puri*, cioè suoni di una determinata frequenza.

In generale i suoni emessi dagli strumenti musicali (per esempio, una corda del pianoforte) non sono puri, ma risultano dalla sovrapposizione di diverse armoniche. Ciò è dovuto al fatto che una corda fissata agli estremi può oscillare in diversi modi, cioè con diverse frequenze: queste frequenze (teoricamente infinite) sono multipli di una *frequenza fondamentale* propria della corda (la corda di un *do* può vibrare, oltre che con la frequenza propria di quella nota, anche con frequenza doppia, corrispondente al *do* dell'ottava superiore, o tripla, corrispondente al *sol* dell'ottava superiore e così via). La vibrazione effettiva di un dato punto della corda risulta dalla sovrapposizione delle vibrazioni corrispondenti ai singoli modi: essa è perciò descritta da una somma (teoricamente infinita) di funzioni periodiche semplici del tipo:

$$a_n \cos(\omega n t) \quad \text{oppure} \quad b_n \sin(\omega n t), \qquad \omega = \frac{2\pi}{T}$$

(cioè da una serie di Fourier). Ci si può chiedere perché sia necessario sovrapporre vibrazioni fondamentali di tipo sia sinusoidale che cosinusoidale. L'identità trigono-

metrica[26]

$$a_n \cos(\omega nt) + b_n \sin(\omega nt) = \sqrt{a_n^2 + b_n^2}\cos(\omega nt - \varphi_n)$$
$$\text{con } \varphi_n = \operatorname{arctg}\frac{b_n}{a_n}$$

mostra che in realtà ogni armonica $a_n \cos(\omega nt)+b_n \sin(\omega nt)$ si può esprimere come una sola cosinusoide (o sinusoide), a patto però di inserire in essa una *fase* φ_n. L'ampiezza complessiva di questa armonica risulta quindi $\sqrt{a_n^2 + b_n^2}$.

In altre parole, possiamo anche dire: una vibrazione periodica si può vedere come sovrapposizione di vibrazioni elementari tutte di tipo cosinusoidale, che hanno frequenze multiple di una frequenza fondamentale, ampiezze diverse e sono inoltre sfasate tra loro.

La stessa terminologia (ampiezza, frequenza, armoniche...) viene utilizzata al di fuori dell'ambito acustico, nell'analisi di segnali (periodici) di vario tipo, per esempio elettromagnetici.

Il fatto che i coefficienti di Fourier tendano a zero per $n \to \infty$ (lemma di Riemann-Lebesgue) corrisponde al fatto che, in pratica (cioè dal punto di vista sia fisico che dell'approssimazione numerica), solo le prime armoniche sono rilevanti nella descrizione di un segnale periodico: quelle successive hanno ampiezza talmente piccola da essere praticamente trascurabili. Più numerose sono le armoniche presenti (in modo rilevante) in un segnale, più il segnale potrà avere una forma complicata e irregolare, per la presenza di molte armoniche di alta frequenza, che oscillano rapidamente. Viceversa, la presenza di pochi termini nello sviluppo di Fourier significa la presenza delle sole basse frequenze, perciò un segnale lentamente variabile e dalla forma "liscia".

Abbiamo anche visto che, più regolare è una funzione periodica (cioè maggiore il numero di derivate che possiede su tutto $\mathbb{R}$), più rapida è la convergenza a zero dei suoi coefficienti di Fourier. Per quanto detto, questo fatto significa in pratica che è minore il numero di termini effettivamente presenti nello sviluppo. Quindi:

funzione molto regolare	$\Longrightarrow$	poche armoniche significative nello sviluppo di Fourier
funzione poco regolare	$\Longrightarrow$	molte armoniche di alta frequenza nello sviluppo di Fourier.

Energia di un segnale. Interpretazione fisica dell'uguaglianza di Parseval

Supponiamo che $f(t)$ rappresenti l'ampiezza all'istante t di un segnale T-periodico nel tempo (segnale acustico, o elettromagnetico, o altro); la sua energia è allora data da:

$$E(f) = \frac{1}{T}\int_0^T f(t)^2 dt$$

cioè dal valore efficace di f in un periodo. Se f viene sviluppata in serie di Fourier, ogni singola armonica possiede la sua energia: precisamente, chiamiamo *energia dell'armonica di ordine* n (≥ 1) l'espressione:

$$E_n = \frac{1}{T}\int_0^T [a_n \cos n\omega t + b_n \sin n\omega t]^2 dt =$$

$$= \frac{1}{T}\int_0^T [a_n^2(\cos n\omega t)^2 + b_n^2(\sin n\omega t)^2 + 2a_n b_n \cos n\omega t \sin n\omega t]dt =$$

(per la proposizione 7.3)

$$= \frac{a_n^2 + b_n^2}{2}.$$

Si noti che, per $n = 0$, risulta:

$$E_0 = \frac{a_0^2}{4} = \left(\frac{1}{T}\int_0^T f(t)\,dt\right)^2.$$

L'uguaglianza di Parseval (3.17) si riscrive così:

$$E(f) = \sum_{n=0}^{\infty} E_n$$

e ha dunque un'interpretazione fisica notevole: *l'energia del segnale si ottiene come somma delle energie delle singole armoniche* (compresa l'energia dell'armonica zero, che è il quadrato del valor medio).

3.8 Applicazioni alle equazioni differenziali della fisica matematica. Metodo di separazione delle variabili

L'equazione della corda vibrante

Accenniamo, per cominciare, a uno dei problemi da cui storicamente è nato l'utilizzo delle serie di Fourier: lo studio dell'*equazione della corda vibrante*.

Consideriamo una corda vibrante illimitata disposta lungo l'asse x; sia $y(t, x)$ l'altezza al tempo t del punto che, a riposo, occupa la posizione x; questa funzione soddisfa l'equazione a derivate parziali

$$\frac{\partial^2 y}{\partial t^2} = a^2 \frac{\partial^2 y}{\partial x^2} \quad (\text{con } a \text{ costante}).$$

All'equazione precedente vanno accoppiate le condizioni iniziali che precisano posizione e velocità iniziali della corda:

$$\begin{cases} y(0, x) = \varphi(x) \\ \partial_t y(0, x) = \psi(x) \end{cases}$$

con φ, ψ assegnate.

Nel 1749 D'Alembert mostrò che l'equazione precedente è soddisfatta da

$$y(t, x) = \frac{f(x + at) + g(x - at)}{2}$$

per qualsiasi coppia di funzioni f, g due volte derivabili[27].

[26] v. volume 1, capitolo 2, paragrafo 3.4.

[27] Nel capitolo 4, paragrafo 5.2, abbiamo mostrato in dettaglio questo fatto, in un esempio.

 © 978-88-08-06485-1

Nel 1753 Daniel Bernoulli propose invece la seguente soluzione per l'equazione delle onde, questa volta per una corda fissata agli estremi, $x = \pm\pi$:

$$y(t,x) = \sum_{k=1}^{\infty} (c_k \sin kx \cos akt + d_k \sin kx \sin akt)\,. \tag{3.28}$$

Si osservi che ogni somma parziale della serie soddisfa l'equazione differenziale, per qualsiasi scelta dei coefficienti c_k, d_k:

$$\frac{\partial^2}{\partial t^2}\left(\sum_{k=1}^{n} c_k \sin kx \cos akt\right) - a^2 \frac{\partial^2}{\partial x^2}\left(\sum_{k=1}^{n} c_k \sin kx \cos akt\right) =$$
$$= \sum_{k=1}^{n} c_k \left[\sin kx \frac{\partial^2}{\partial t^2}(\cos akt) - a^2 \cos akt \frac{\partial^2}{\partial x^2}(\sin kx)\right] = 0$$

(e analogamente per l'addendo contenente d_k). Inoltre, $y(t, \pm\pi) = 0$ (condizione che esprime il fatto che gli estremi della corda sono fissati); infine,

$$y(0,x) = \sum_{k=1}^{\infty} c_k \sin kx = \varphi(x) \tag{3.29}$$
$$\partial_t y(0,x) = \sum_{k=1}^{\infty} d_k ak \sin kx = \psi(x)$$

perciò i dati iniziali φ, ψ devono potersi scrivere in questo modo.

Ciascun termine della serie (3.28) ha un significato fisico: rappresenta una vibrazione stazionaria, o *modo normale* di vibrazione della corda. L'ipotesi di Bernoulli è quindi che qualsiasi vibrazione reale possa vedersi come sovrapposizione di infinite vibrazioni stazionarie. Ma questo implicherebbe che, assegnata comunque la configurazione iniziale $\varphi(x)$ della corda, si possa scrivere $\varphi(x)$ come serie di seni (3.29) (e analogamente per la ψ). A questo punto, storicamente, si aprì una lunga diatriba tra i sostenitori dell'uno e dell'altro metodo di soluzione, a proposito di quale metodo fosse più generale, sotto quali ipotesi rappresentasse la soluzione e perfino su cosa si dovesse intendere per *funzione*. La prima risposta rigorosa a queste domande fu data nel 1829 da Dirichlet, che precisò un insieme di condizioni sufficienti affinché una funzione data sia sviluppabile in serie trigonometrica; contestualmente, Dirichlet diede la definizione moderna di funzione come corrispondenza univoca tra due insiemi numerici, a prescindere dalla *forma analitica* della legge che realizza questa corrispondenza. Dirichlet chiarì le ipotesi precise di un teorema che era stato anticipato (senza prova rigorosa) da Fourier, a cui si attribuisce usualmente il merito di essere stato il primo a intuire la portata generale del metodo degli sviluppi in serie trigonometrica, che da lui presero il nome [28].

Come si vede, c'è quindi una ragione fisica per ritenere che una funzione (ragionevole) si possa sviluppare in serie di seni; inoltre, è necessario determinare i coefficienti di questo sviluppo per poter scrivere la soluzione del problema che interessa.

[28] Fourier introdusse questi sviluppi in serie nello studio di un altro problema fisico-matematico: la diffusione del calore, problema a cui accenneremo più avanti in questo stesso paragrafo.

© 978-88-08-06485-1

Mostriamo ora, con un po' più di dettaglio analitico, un altro paio di esempi classici di applicazioni delle serie di Fourier alla risoluzione di problemi ai limiti per equazioni differenziali a derivate parziali, importanti per la fisica matematica. Avremo modo così di illustrare il *metodo di separazione delle variabili* e l'utilizzo congiunto di molti strumenti e tecniche studiate in tutto il corso.

Problema di Dirichlet sul cerchio per l'equazione di Laplace

Consideriamo il seguente problema: determinare una funzione $u(x,y)$ che risolve:

$$\begin{cases} u_{xx} + u_{yy} = 0 & \text{per } x^2 + y^2 < 1 \\ u(x,y) = f(x,y) & \text{per } x^2 + y^2 = 1. \end{cases} \tag{3.30}$$

In altre parole, vogliamo determinare una funzione u che soddisfa l'equazione di Laplace $u_{xx} + u_{yy} = 0$ all'interno del cerchio unitario e sul bordo del cerchio assuma il dato f assegnato. Si tratta di un problema che può avere vari significati fisici, come la determinazione della configurazione a riposo di una membrana elastica circolare fissata al bordo a una certa sagoma, o la determinazione della temperatura stazionaria di una lastra circolare isolata, nota la temperatura sul bordo.

Riscriviamo anzitutto il problema in coordinate polari. Si tratta allora di cercare una funzione $u(\rho,\theta)$ che risolve la seguente equazione[29] e condizione al contorno

$$\begin{cases} \dfrac{\partial^2 u}{\partial \rho^2} + \dfrac{1}{\rho^2}\dfrac{\partial^2 u}{\partial \theta^2} + \dfrac{1}{\rho}\dfrac{\partial u}{\partial \rho} = 0 & \text{per } \rho \in [0,1); \theta \in [0, 2\pi) \\ u(1,\theta) = f(\theta) & \text{per } \theta \in [0, 2\pi). \end{cases}$$

Applichiamo il metodo di separazione delle variabili, che consiste nel cercare soluzioni a variabili separate, del tipo

$$u(\rho,\theta) = R(\rho)T(\theta)\,.$$

Sostituendo questa funzione nell'equazione differenziale, abbiamo:

$$R''(\rho)T(\theta) + \frac{1}{\rho^2}R(\rho)T''(\theta) + \frac{1}{\rho}R'(\rho)T(\theta) = 0,$$

che riscriviamo nella forma:

$$\rho^2 \frac{R''(\rho)}{R(\rho)} + \rho \frac{R'(\rho)}{R(\rho)} = -\frac{T''(\theta)}{T(\theta)}.$$

Ora, poiché il primo membro è una funzione della sola ρ e il secondo è funzione della sola θ, l'unica possibilità è che ciascuno dei due membri sia costante. Poniamo perciò, per una certa costante λ,

$$\begin{cases} -\dfrac{T''(\theta)}{T(\theta)} = \lambda \\ \rho^2 \dfrac{R''(\rho)}{R(\rho)} + \rho \dfrac{R'(\rho)}{R(\rho)} = \lambda \end{cases}$$

[29] Nel capitolo 4, paragrafo 5.2, abbiamo infatti mostrato come si scrive l'operatore di Laplace in coordinate polari.

 © 978-88-08-06485-1

ossia

$$\begin{cases} -T''(\theta) = \lambda T(\theta) \\ \rho^2 R''(\rho) + \rho R'(\rho) = \lambda R(\rho). \end{cases}$$

La prima è un'equazione differenziale lineare del second'ordine a coefficienti costanti, che quindi può avere soluzioni trigonometriche o esponenziali. Dovendo essere però $T(\theta)$ una funzione 2π-periodica, l'unica possibilità è che abbia soluzioni trigonometriche del tipo

$$T(\theta) = a\cos k\theta + b\sin k\theta,$$

il che significa che $\lambda = k^2$ per qualche intero $k = 0, 1, 2, \ldots$ Sostituendo questo valore nell'equazione per R, abbiamo:

$$\rho^2 R''(\rho) + \rho R'(\rho) = k^2 R(\rho) \text{ per } \rho \in [0,1). \tag{3.31}$$

Questa è un'equazione lineare omogenea a coefficienti non costanti; tuttavia, è di un tipo particolare che sappiamo risolvere. Si tratta infatti di un'*equazione di Eulero*[30], di cui possiamo cercare soluzioni del tipo

$$R(\rho) = \rho^\alpha,$$

per $\alpha \in \mathbb{R}$ da determinarsi. Sostituendo nella (3.31), si trova:

$$\begin{aligned} \rho^2 \alpha(\alpha-1)\rho^{\alpha-2} + \rho\alpha\rho^{\alpha-1} &= k^2\rho^\alpha \\ \alpha(\alpha-1) + \alpha &= k^2 \\ \alpha &= \pm k \end{aligned}$$

che dà le due soluzioni

$$R(\rho) = \rho^k, \quad R(\rho) = \rho^{-k}, \quad \text{per } k \neq 0,$$

mentre per $k = 0$ la (3.31) ha le due soluzioni

$$R(\rho) = 1, \quad R(\rho) = \log\rho.$$

Le soluzioni $R(\rho) = \rho^{-k}$, $R(\rho) = \log\rho$ non sono accettabili perché illimitate per $\rho \to 0$. Complessivamente otteniamo perciò le seguenti soluzioni a variabili separate:

$$u_k(\rho,\theta) = \rho^k \{a_k \cos k\theta + b_k \sin k\theta\}, \text{ per } k = 0, 1, 2, 3 \ldots$$

Queste sono infinite soluzioni dell'equazione differenziale nel cerchio. Notiamo che u_k non soddisfa in generale il dato al bordo. Osserviamo però che, essendo l'equazione differenziale lineare e omogenea, ogni combinazione lineare di soluzioni è ancora soluzione. Per ottenere una soluzione che assuma anche il dato al bordo, si considera allora la serie infinita

$$u(\rho,\theta) = \sum_{k=0}^{\infty} \rho^k \{a_k \cos k\theta + b_k \sin k\theta\}$$

[30] v. capitolo 1, esercizio 42.

© 978-88-08-06485-1

e si cercano coefficienti a_k, b_k per cui sia vero che $u(1,\theta) = f(\theta)$. Imponendo questa condizione, si ha:

$$\sum_{k=0}^{\infty} \{a_k \cos k\theta + b_k \sin k\theta\} = f(\theta),$$

che mostra che *i coefficienti a_k, b_k vanno scelti come coefficienti di Fourier di f*. In altre parole, la soluzione del problema (3.30) è assegnata da:

$$u(\rho,\theta) = \frac{a_0}{2} + \sum_{k=1}^{\infty} \rho^k \{a_k \cos k\theta + b_k \sin k\theta\}, \text{ con}$$

$$a_k = \frac{1}{\pi}\int_0^{2\pi} f(\theta)\cos k\theta\, d\theta$$

$$b_k = \frac{1}{\pi}\int_0^{2\pi} f(\theta)\sin k\theta\, d\theta.$$

Tutto questo è finora piuttosto formale: per rendere rigorosi tutti i passaggi effettuati e poter quindi affermare un risultato preciso sulla soluzione del problema di Dirichlet, occorre fare opportune ipotesi di regolarità su f. Tuttavia qui ci interessava principalmente dare un'idea di questa tecnica risolutiva, perciò non ci soffermiamo su questi aspetti. Quest'osservazione vale anche per il prossimo esempio dell'equazione del calore.

Equazione del calore sul segmento

Consideriamo il problema di determinare la temperatura $u(t,x)$ di una sbarra omogenea, rappresentata dal segmento $[0,\pi]$, all'istante t e nel punto x, supponendo che la temperatura agli estremi sia costantemente tenuta uguale a zero, che la temperatura iniziale sia una funzione $u_0(x)$ nota e che non ci siano sorgenti o pozzi di calore all'interno della sbarra. Sotto queste ipotesi, u risolve il problema:

$$\begin{cases} u_t - \kappa u_{xx} = 0 & \text{per } x \in (0,\pi), t > 0 \\ u(0,t) = u(\pi,t) = 0 & \text{per } t > 0 \\ u(x,0) = u_0(x) & \text{per } x \in [0,\pi] \end{cases} \tag{3.32}$$

dove la costante positiva κ è un coefficiente di conducibilità (espresso in opportune unità di misura). Proprio questo è il problema che portò Fourier a studiare gli sviluppi in serie trigonometriche.

Possiamo affrontarlo mediante il metodo di separazione delle variabili, illustrato nell'esempio precedente.

Cerchiamo dunque soluzioni a variabili separate, del tipo

$$u(x,t) = X(x)T(t),$$

che sostituite nell'equazione danno:

$$\begin{aligned} X(x)T'(t) - \kappa X''(x)T(t) &= 0 \\ \frac{T'(t)}{\kappa T(t)} &= \frac{X''(x)}{X(x)} = \lambda \end{aligned}$$

 © 978-88-08-06485-1

per un'opportuna costante λ (si veda il ragionamento che abbiamo fatto nell'esempio precedente).

Ora, l'equazione

$$X''(x) = \lambda X(x) \text{ per } x \in (0, \pi)$$

insieme alle condizioni $X(0) = X(\pi)$, necessarie per avere una soluzione u nulla agli estremi, come richiesto dal problema, implica:

$$X(x) = \sin kx \text{ per } k = 1, 2, 3 \ldots \quad \text{e}$$
$$\lambda = -k^2$$

che, sostituita nell'equazione in t, dà:

$$T'(t) = -k^2 \kappa T(t)$$

ossia

$$T(t) = c_k e^{-k^2 \kappa t}.$$

Ora cerchiamo una soluzione u che sia sovrapposizione delle infinite soluzioni a variabili separate, con coefficienti c_k da determinarsi in modo che u soddisfi anche la condizione iniziale. Si trova:

$$u(x,t) = \sum_{k=1}^{\infty} c_k e^{-k^2 \kappa t} \sin kx \tag{3.33}$$

con

$$u(x,0) = \sum_{k=1}^{\infty} c_k \sin kx = u_0(x).$$

I coefficienti c_k sono dunque determinati dal fatto che devono assegnare lo sviluppo di Fourier di f in serie di soli seni. Per chiarire quest'ultimo punto, si noti che il sistema di funzioni $\cos kx, \sin kx$ è il sistema ortogonale relativo a un intervallo di ampiezza 2π, mentre la nostra funzione è definita solo su $[0, \pi]$. Possiamo allora definirla su tutto $[-\pi, \pi]$ facendo una riflessione dispari, ossia ponendo $f(-x) = -f(x)$ per $x \in [0, \pi]$. Questo fa sì che la funzione così prolungata abbia uno sviluppo di Fourier in cui compaiono solo le funzioni $\sin kx$ (come desiderato), sviluppo che, valendo in tutto $[-\pi, \pi]$, varrà in particolare in $[0, \pi]$. Tenendo conto della formula per il calcolo dei coefficienti di Fourier di una funzione dispari, si ha allora:

$$c_k = \frac{2}{\pi} \int_0^{\pi} u_0(x) \sin kx dx,$$

che sostituiti nella (3.33) danno la soluzione cercata del problema (3.32).

Esempio

3.11 La soluzione del problema (3.32), con $u_0(x) = x(\pi - x)$ e $\kappa = 1$, si trova calcolando i coefficienti di Fourier

$$c_k = \frac{2}{\pi} \int_0^{\pi} x(\pi - x) \sin kx dx = \text{con calcoli di routine}$$
$$= \frac{4}{\pi k^3} \left(1 - (-1)^k\right) = \begin{cases} \frac{8}{\pi k^3} & \text{per } k \text{ dispari} \\ 0 & \text{per } k \text{ pari} \end{cases}$$

© 978-88-08-06485-1

e, quindi, sostituendoli nella formula risolutiva (3.33):

$$u(x,t) = \sum_{k=0}^{\infty} \frac{8}{\pi (2k+1)^3} e^{-(2k+1)^2 t} \sin kx.$$

Un grafico della somma dei primi termini di questa serie è dato in figura 7.15.

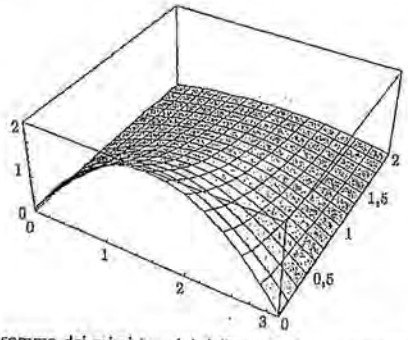


Figura 7.15. Grafico della somma dei primi termini della serie che rappresenta la soluzione del problema dell'Esempio 3.11.

Sviluppo in serie di seni o di coseni

L'ultimo esempio mostra che talvolta nelle applicazioni è utile sviluppare una funzione su un intervallo in serie di soli seni (o coseni), anche se a priori la funzione non ha le proprietà di simmetria che giustificano l'annullarsi dei coefficienti dei coseni (o seni, rispettivamente). Vogliamo ora puntualizzare in generale come si procede in casi come questo.

Supponiamo di avere una funzione $f(x)$ definita a priori solo sull'intervallo $[0, T]$. Possiamo definire f su tutto $[-T, T]$ facendone una riflessione pari, cioè definendo

$$\widetilde{f}(x) = \begin{cases} f(x) & \text{se } x \in [0, T] \\ f(-x) & \text{se } x \in [-T, 0] \end{cases}$$

e ottenendo così una funzione pari su $[-T, T]$; questa funzione può essere quindi prolungata su tutto $\mathbb{R}$ come funzione $2T$-periodica; i coefficienti di Fourier di $\widetilde{f}$ saranno dunque dati da:

$$a_k = \frac{1}{T} \int_{-T}^{T} \widetilde{f}(x) \cos \omega kx dx = \frac{2}{T} \int_0^T f(x) \cos \omega kx dx, \text{ con } \omega = \frac{2\pi}{T},$$
$$b_k = 0.$$

Si osservi che la funzione $\widetilde{f}$ è continua in 0 solo se $f(0) = 0$; in compenso, $\widetilde{f}$ soddisfa sempre la condizione di raccordo $\widetilde{f}(T) = \widetilde{f}(-T)$, proprio perché è pari.

Se la serie di Fourier di $\widetilde{f}$ converge, in particolare in $[0, T]$ rappresenterà f: ciò significa che abbiamo trovato un *metodo per sviluppare una funzione definita su un intervallo $[0, T]$ in serie di soli coseni, se lo preferiamo, anziché in serie di seni e coseni.*

 © 978-88-08-06485-1

Analogamente, la funzione f definita su $[0,T]$ potrebbe essere prolungata su tutto $[-T,T]$ effettuando una riflessione dispari:

$$\widetilde{f}(x) = \begin{cases} f(x) & \text{se } x \in [0,T] \\ -f(-x) & \text{se } x \in [-T,0]. \end{cases}$$

La serie di Fourier di $\widetilde{f}$ in $[-T,T]$ avrà coefficienti:

$$a_k = 0,$$
$$b_k = \frac{1}{T}\int_{-T}^{T} \widetilde{f}(x)\sin\omega kx\,dx = \frac{2}{T}\int_0^T f(x)\sin\omega kx\,dx, \text{ con } \omega = \frac{2\pi}{T},$$

ossia sarà una serie di soli seni, che ristretta all'intervallo $[0,T]$ rappresenta f.

Le prossime figure illustrano il significato grafico dell'operazione di riflessione pari o dispari e successiva estensione periodica.

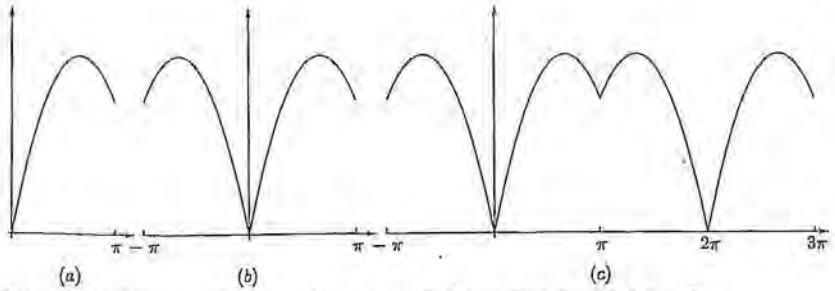


Figura 7.16. a) Funzione di partenza. b) Riflessa pari. c) Estensione 2π-periodica su $\mathbb{R}$.

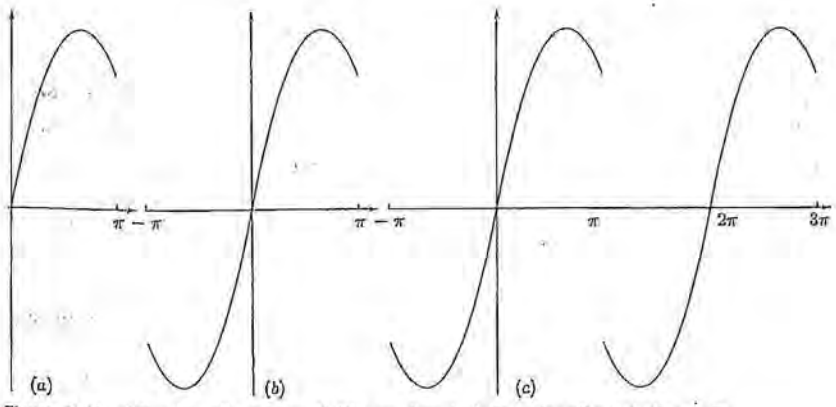


Figura 7.17. a) Funzione di partenza. b) Riflessa dispari. c) Estensione 2π-periodica su $\mathbb{R}$.

Esempio

Abbiamo calcolato in precedenza (esempio 3.8) lo sviluppo di Fourier della funzione e^x su $[0,2\pi]$, trovando:

$$e^x \sim \frac{e^{2\pi}-1}{2\pi} + \sum_{h=1}^{\infty}\left\{\frac{1}{\pi}\cdot\frac{e^{2\pi}-1}{1+h^2}\cos hx + \frac{h}{\pi}\cdot\frac{1-e^{2\pi}}{1+h^2}\sin hx\right\} \tag{3.34}$$

La stessa funzione potrebbe essere rappresentata in serie di soli seni o soli coseni. Per esempio, calcoliamo lo sviluppo di e^x in serie di soli seni, su $[0,2\pi]$. Questo significa pensare di fare una riflessione dispari di e^x in $[-2\pi,2\pi]$ e prolungare per periodicità la funzione così ottenuta. Il procedimento che abbiamo appena descritto per l'intervallo $[0,\pi]$ va adattato all'intervallo $[0,2\pi]$. In sostanza, consideriamo $[0,2\pi]$ come metà di $[-2\pi,2\pi]$, che ora è pensato come periodo fondamentale; le funzioni trigonometriche adattate a questo intervallo sono:

$$\cos\left(\frac{hx}{2}\right), \qquad \sin\left(\frac{hx}{2}\right).$$

Avremo:

$$a_h = 0;\ b_h = \frac{1}{\pi}\int_0^{2\pi} e^x \sin\left(\frac{hx}{2}\right)dx = \frac{2h}{\pi}\left(\frac{1-e^{2\pi}(-1)^h}{4+h^2}\right)$$

e quindi

$$e^x \sim \sum_{h=1}^{\infty}\frac{2h}{\pi}\left(\frac{1-e^{2\pi}(-1)^h}{4+h^2}\right)\sin\left(\frac{hx}{2}\right). \tag{3.35}$$

Si confrontino (3.34) e (3.35): sono due diversi sviluppi di e^x validi su $(0,2\pi)$. Il primo è uno sviluppo in serie di $\sin hx, \cos hx$; il secondo è uno sviluppo in serie di $\sin\frac{hx}{2}$. Rappresentiamo la funzione e^x su $(0,2\pi)$, insieme alla somma parziale fino a $n = 10$ del suo sviluppo in serie di soli seni:

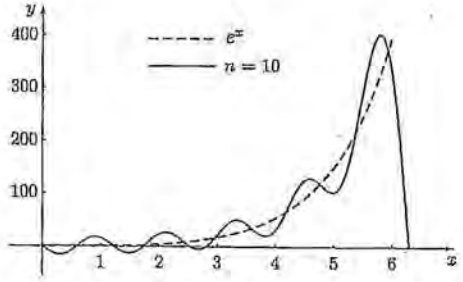


Figura 7.18.

Si confronti col grafico dello sviluppo in seni e coseni, riportato nell'esempio 3.8.

Esercizi

Stabilire per ciascuna delle seguenti serie trigonometriche se non converge, converge puntualmente, converge totalmente in $[0,2\pi]$ o in un sottointervallo:

$$\sum_{k=1}^{\infty}\frac{\cos kx}{2^k}, \qquad \sum_{k=1}^{\infty}\frac{\sin kx}{\sqrt{k}}, \qquad \sum_{k=1}^{\infty}\frac{(-1)^k+k}{1+k^2}\cos kx.$$

 © 978-88-08-06485-1

Sfruttando i risultati che abbiamo esposto, dire per quali $\alpha \in \mathbb{R}$ si può affermare che la serie:

$$\sum_{k=1}^{\infty} \frac{\sin kx}{k^{\alpha}}$$

a) converge puntualmente in $[0, 2\pi]$;
b) rappresenta una funzione continua in $\mathbb{R}$;
c) è derivabile termine a termine.

Sfruttando le identità trigonometriche

$$(\cos x)^2 = \tfrac{1+\cos 2x}{2} \qquad \cos\alpha\cos\beta = \frac{\cos(\alpha+\beta)+\cos(\alpha-\beta)}{2}$$

$$(\sin x)^2 = \tfrac{1-\cos 2x}{2} \qquad \sin\alpha\sin\beta = \frac{\cos(\alpha-\beta)-\cos(\alpha+\beta)}{2}$$

$$\sin x\cos x = \tfrac{1}{2}\sin 2x \qquad \sin\alpha\cos\beta = \frac{\sin(\alpha+\beta)+\sin(\alpha-\beta)}{2}$$

è possibile riscrivere sotto forma di polinomio trigonometrico qualsiasi funzione del tipo

$$(\sin x)^r(\cos x)^s, \quad \text{con } r, s \text{ interi} \geq 0.$$

Sfruttando questa osservazione (quindi *senza calcolare integrali*!), scrivere gli sviluppi di Fourier su $[0, 2\pi]$ delle funzioni:

(a) $\cos^2 x \sin x$ (b) $(\cos x)^4$ (c) $(\sin x)^3$
(d) $\sin x(\cos x)^3$ (e) $(\sin x)^3(\cos x)^2$.

Sfruttando opportunamente lo sviluppo di Fourier (3.27), calcolare la somma della serie numerica:

$$\sum_{n=1}^{\infty} \frac{(-1)^n}{2n+1}.$$

Si consideri la funzione $f(x) = |x|$ su $[-\pi, \pi]$. In base ai risultati visti, si può affermare che la serie di Fourier di f converge a $|x|$? Calcolare quindi lo sviluppo di Fourier di $f(x)$ e verificare che effettivamente la serie converge totalmente. Infine, porre $x = 0$ nello sviluppo in serie trovato e ottenere la somma di una serie numerica notevole.

Sia

$$f(x) = \begin{cases} 0 \text{ in } [0, \pi] \\ 1 \text{ in } (\pi, 2\pi) \end{cases}$$

periodizzata di periodo 2π. In base ai risultati visti, dire che cosa si può affermare a priori sulla serie di Fourier di f. Calcolare poi lo sviluppo di Fourier su tale intervallo. Con l'ausilio di un computer, tracciare il grafico di $f(x)$ e delle somme parziali della sua serie di Fourier per i primi valori di n.

Sia $f(x) = x^2$ in $[0, 2\pi]$. In base ai risultati visti, che cosa si può affermare a priori sulla serie di Fourier di f su tale intervallo? Calcolare poi lo sviluppo di Fourier. Con l'ausilio di un computer, tracciare il grafico di $f(x)$ e delle somme parziali della sua serie di Fourier per i primi valori di n.

Suggerimento: si confronti con l'Esempio 3.10, notando che la funzione f nei due casi è *diversa*: per evidenziarlo, disegnare il grafico delle funzioni che prolungano periodicamente f nei due casi.

© 978-88-08-06485-1

Sia $f(x) = x^2(2\pi - x)^2$ in $[0, 2\pi]$. In base ai risultati visti, che cosa si può affermare a priori sulla serie di Fourier di f su tale intervallo e in particolare sulla sua derivabilità termine a termine e sulla rapidità di convergenza a zero dei coefficienti di Fourier?

Sia

$$f(x) = \begin{cases} 0 & \text{in } [0, \pi] \\ x - \pi & \text{in } (\pi, 2\pi). \end{cases}$$

In base ai risultati visti, che cosa si può affermare a priori sulla serie di Fourier di f? Calcolare poi lo sviluppo di Fourier su tale intervallo. Con l'ausilio di un computer, tracciare il grafico di $f(x)$ e delle somme parziali della sua serie di Fourier per i primi valori di n.

Sviluppare in serie di Fourier il prolungamento periodico della funzione x^4 per $x \in [-\pi, \pi]$. Dedurre la formula

$$\sum_{n=1}^{\infty} \frac{1}{n^4} = \frac{\pi^4}{90}.$$

Tracciare il grafico delle seguenti funzioni sull'intervallo $[-2\pi, 4\pi]$:
a) $f(x) = e^{-x}$ su $[0, 2\pi]$, prolungata 2π-periodica.
b) $f(x) = e^{-x}$ su $[-\pi, \pi]$, prolungata 2π-periodica.
c) $f(x) = e^{-x}$ su $[0, \pi]$, riflessa pari su $[-\pi, \pi]$ e poi prolungata 2π-periodica.
d) $f(x) = e^{-x}$ su $[0, \pi]$, riflessa dispari su $[-\pi, \pi]$ e poi prolungata 2π-periodica.

Calcolare i coefficienti di Fourier delle seguenti funzioni e scrivere esplicitamente la corrispondente serie di Fourier, prestando attenzione a come la funzione è definita e a qual è il suo periodo:

(a) $f(x) = x^2$ per $x \in [0, 1]$, periodizzata di periodo $T = 1$.
(b) $f(x) = x^2$ per $x \in [0, 1]$, riflessa dispari in $[-1, 0]$ e periodizzata di periodo $T = 2$.
(c) $f(x) = x^3$ per $x \in [0, 1]$, riflessa pari in $[-1, 0]$ e periodizzata di periodo $T = 2$.

Dimostrare il seguente fatto, che può risultare utile nel calcolo dei coefficienti di Fourier di una funzione definita a tratti.

Se $f, g : [0, T] \to \mathbb{R}$ sono funzioni C^1 a tratti in $[0, T]$ e continue in $[0, T]$ (quindi possono avere punti angolosi ma non punti di discontinuità), allora vale la "solita" formula di integrazione per parti:

$$\int_0^T f(x)\, g'(x)\, dx = (fg)(T) - (fg)(0) - \int_0^T f'(x)\, g(x)\, dx.$$

Suggerimento: mostrare prima che per una funzione di questo tipo vale il teorema fondamentale del calcolo integrale $\int_0^T f'(x)\, dx = f(T) - f(0)$, riscrivendo il primo integrale come somma di integrali sugli intervalli in cui f è C^1 e applicando a ogni addendo il teorema fondamentale del calcolo (per funzioni C^1).

Utilizzare poi il risultato precedente per calcolare i coefficienti di Fourier della funzione definita sull'intervallo $(-3, 3)$ come in figura 7.19 e periodizzata di periodo 6.

© 978-88-08-06485-1

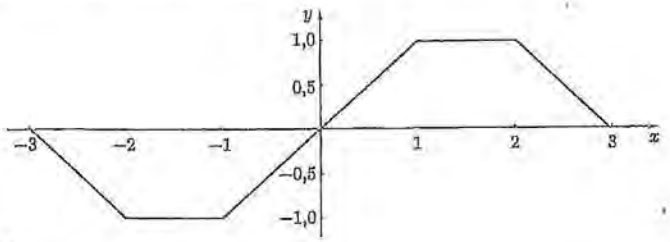


Figura 7.19.

Calcolare i coefficienti di Fourier in forma complessa e scrivere esplicitamente la corrispondente serie di Fourier, prestando attenzione a come la funzione è definita e a qual è il suo periodo.

(a) $f(x) = x$ per $x \in [-\pi, \pi]$, periodizzata di periodo $T = 2\pi$.
(b) $f(x) = x$ per $x \in [0, 2\pi]$, periodizzata di periodo $T = 2\pi$.
(c) $f(x) = |x|$ per $x \in [-1, 1]$, periodizzata di periodo $T = 2$.

Calcolare i seguenti sviluppi in serie di Fourier:
(a) $f(x) = x$ su $[0, 2\pi]$, prolungata 2π-periodica.
(b) $f(x) = x$ su $[-\pi, \pi]$, prolungata 2π-periodica.
(c) $f(x) = x$ su $[0, \pi]$, in serie di soli coseni.

Scrivere lo sviluppo di $f(x) = \cos x$ in $[0, \pi]$, in serie di soli seni.

Facendo riferimento all'esempio dell'equazione del calore, illustrato nel paragrafo 3.8, si consideri la serie di funzioni:

$$u(x,t) = \sum_{k=1}^{\infty} c_k e^{-k^2 t} \sin kx, \quad \text{per } t > 0, x \in [0, \pi].$$

Sfruttando il teorema 7.2, si provi che, se $\{c_k\}$ è una successione qualunque, purché limitata, di coefficienti reali, è possibile calcolare le derivate parziali u_t e u_{xx}, per $t > 0, x \in [0, \pi]$, derivando la serie termine a termine. Mostrare che, sotto queste ipotesi, u risolve, per $t > 0, x \in [0, \pi]$, l'*equazione del calore* $u_t - u_{xx} = 0$, con le condizioni agli estremi $u(0,t) = 0 = u(\pi, t)$.

Facendo riferimento all'esempio dell'equazione di Laplace nel cerchio (v. paragrafo 3.8), scrivere esplicitamente la soluzione del problema:

$$\begin{cases} \Delta u = 0 & \text{per } \rho < 1 \\ u(1, \theta) = \cos^2 \theta & \text{per } \theta \in [0, 2\pi). \end{cases}$$

Facciamo riferimento ancora all'esempio dell'equazione di Laplace nel cerchio (v. paragrafo 3.8). Le funzioni di due variabili che, in coordinate polari, si scrivono

$$\rho^k \cos k\theta, \quad \rho^k \sin k\theta, \quad \text{per } k = 1, 2, 3, \ldots$$

si dicono armoniche elementari nel piano. Sfruttando la forma esponenziale dei numeri complessi, si verifichi che queste funzioni sono la parte reale e immaginaria di z^k, se $z = x + iy$. Si sfrutti questa relazione per scrivere le armoniche elementari come funzioni di x, y, per i primi valori di k. (Per esempio, $k = 2, 3, 4$). Si verifichi esplicitamente che queste funzioni risolvono l'equazione di Laplace $u_{xx} + u_{yy} = 0$ in $\mathbb{R}^2$.

© 978-88-08-06485-1

4 COMPLEMENTI

4.1 Il metodo di Frobenius per la soluzione delle equazioni differenziali

Illustriamo ora, su un esempio specifico che ha particolare interesse nelle applicazioni, un metodo di risoluzione per certe classi di equazioni differenziali lineari a coefficienti non costanti che utilizza gli sviluppi in serie di potenze e prende il nome di *metodo di Frobenius*.

Consideriamo l'equazione differenziale:

$$u''(t) + \frac{u'(t)}{t} + u(t) = 0. \tag{4.1}$$

La (4.1) si dice *equazione di Bessel*[31] ed emerge in svariati problemi di fisica matematica. Per risolverla, cerchiamo una soluzione della (4.1) che si possa scrivere come serie di potenze:

$$u(t) = \sum_{n=0}^{\infty} c_n t^n \tag{4.2}$$

dove i coefficienti c_n $(n = 0, 1, 2, \ldots)$ sono da determinarsi.

Il procedimento consiste nel calcolare le derivate di $u(t)$ derivando termine a termine la (4.2) e quindi sostituire le serie che esprimono u, u', u'' nella (4.1). Imponendo che questa risulti soddisfatta identicamente in t, troveremo opportune condizioni sui coefficienti c_n che ci permetteranno di determinarli. Abbiamo:

$$u'(t) = \sum_{n=1}^{\infty} n c_n t^{n-1}; \; u''(t) = \sum_{n=2}^{\infty} n(n-1) c_n t^{n-2}.$$

Sostituiamo nella (4.1):

$$\sum_{n=2}^{\infty} n(n-1) c_n t^{n-2} + \sum_{n=1}^{\infty} n c_n t^{n-2} + \sum_{n=0}^{\infty} c_n t^n = 0.$$

Moltiplichiamo tutto per t^2 e raggruppiamo i termini simili delle prime due serie:

$$c_1 t + \sum_{n=2}^{\infty} [n(n-1) + n] c_n t^n + \sum_{n=0}^{\infty} c_n t^{n+2} = 0.$$

Eseguiamo una traslazione di indici nella seconda serie e raggruppiamo i termini simili:

$$c_1 t + \sum_{n=2}^{\infty} \left(n^2 c_n + c_{n-2}\right) t^n = 0.$$

Affinché questa serie di potenze sia identicamente nulla devono annullarsi tutti i coefficienti, perciò:

$$c_1 = 0; \; c_n = -\frac{1}{n^2} c_{n-2}. \tag{4.3}$$

Osserviamo le (4.3): c_0 è indeterminato; $c_1 = 0$; la relazione ricorsiva

$$c_n = -\frac{c_{n-2}}{n^2}$$

implica allora che tutti i coefficienti c_n con n dispari si annullano.

[31] dal nome di Friedrich Wilhelm Bessel (1784 –1846), matematico e astronomo tedesco.

 © 978-88-08-06485-1

Quelli con n pari si determinano ricorsivamente in funzione di c_0:

$$c_2 = -\frac{1}{2^2}c_0$$
$$c_4 = \frac{1}{4^2 \cdot 2^2}c_0$$
$$c_6 = -\frac{1}{6^2 \cdot 4^2 \cdot 2^2}c_0$$
$$\dots$$

Si capisce che la forma generale dei coefficienti è:

$$c_{2k} = (-1)^k \frac{1}{2^{2k}(k!)^2}c_0$$

e quindi:

$$u(t) = c_0 \sum_{k=0}^{\infty} (-1)^k \frac{1}{2^{2k}(k!)^2} t^{2k}.$$

Il fatto che la serie di potenze sia moltiplicata per un coefficiente indeterminato è naturale: l'equazione differenziale è lineare omogenea e, se u è soluzione, anche $c_0 u$ è soluzione, per ogni $c_0 \in \mathbb{R}$. La funzione

$$J_0(t) = \sum_{k=0}^{\infty} (-1)^k \frac{1}{2^{2k}(k!)^2} t^{2k} = \sum_{k=0}^{\infty} \frac{(-1)^k}{(k!)^2} \left(\frac{t}{2}\right)^{2k} \tag{4.4}$$

si dice *funzione di Bessel di ordine zero* ed è un'importante *funzione speciale* della fisica matematica. Il lettore verifichi (per esempio calcolando il raggio di convergenza della (4.4) col criterio del rapporto) che la serie di potenze converge in tutto $\mathbb{R}$. Questa funzione soddisfa la (4.1) in tutto $\mathbb{R}$. Il grafico qualitativo di $J_0(t)$ è il seguente:

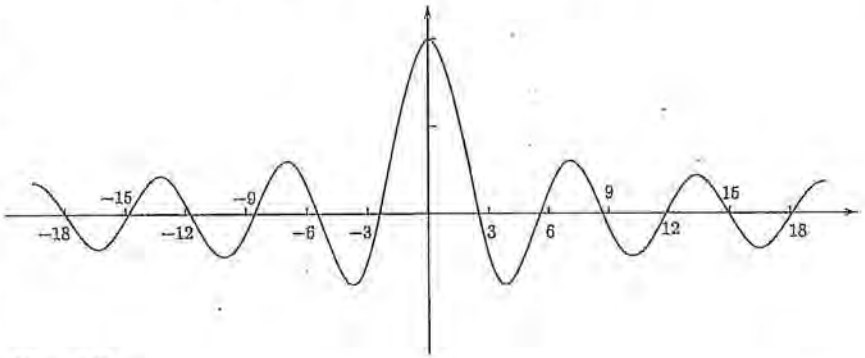


Figura 7.20.

Per ottenere l'integrale generale della (4.1) occorrerebbe determinare una seconda soluzione, indipendente dalla (4.4), ma tale seconda soluzione non ammette uno sviluppo in serie del tipo (4.2); non ce ne occuperemo.

Si rifletta sul fatto che ciò che ha reso applicabile il metodo è il fatto che i coefficienti non costanti dell'equazione differenziale fossero *potenze* di t.

© 978-88-08-06485-1

Esercizi

Equazione e polinomi di Legendre Si consideri l'equazione differenziale

$$(1-x^2)y'' - 2xy' + \lambda y = 0,$$

da risolversi per $x \in (-1,1)$, con λ parametro reale che potrà essere fissato in seguito. Si tratta dell'*equazione di Legendre*, che compare in vari problemi di fisica matematica.

(a) Si cerchi una soluzione

$$y(x) = \sum_{k=0}^{\infty} a_k x^k$$

e, utilizzando il metodo di Frobenius, si ricavi la seguente relazione iterativa sui coefficienti:

$$a_{k+2} = \frac{k(k+1) - \lambda}{(k+1)(k+2)} a_k \quad \text{per } k \geq 2.$$

(b) La relazione precedente lascia i primi due coefficienti a_0, a_1 indeterminati. Si osservi che, ponendo:

$$\begin{cases} a_0 = 1 \\ a_1 = 0 \end{cases} \quad \text{oppure} \quad \begin{cases} a_0 = 0 \\ a_1 = 1, \end{cases}$$

si ottiene una serie contenente solo potenze pari o dispari (rispettivamente). Scegliendo inoltre $\lambda = n(n+1)$ per un intero positivo n fissato (pari o dispari, rispettivamente), tale serie termina, ossia fornisce in realtà un polinomio di grado n. Se quindi per ogni intero n facciamo questa scelta di λ, a_0 e a_1, troviamo una successione di polinomi, detti *polinomi di Legendre*, soluzioni della corrispondente equazione differenziale.

(c) Scrivere esplicitamente i polinomi di Legendre $P_n(x)$ per $n = 0, 1, 2, 3, 4$.

Si consideri l'equazione differenziale (di Legendre[32])

$$(1-x^2)y'' - 2xy' + \lambda y = 0,$$

da risolversi per $x \in (-1,1)$ e valori reali del parametro λ, che si può riscrivere anche nella forma

$$\left((1-x^2)y'\right)' + \lambda y = 0.$$

Dimostrare che due soluzioni $y_\lambda(x), y_\mu(x)$, relative rispettivamente a valori λ, μ del parametro diversi tra loro, sono ortogonali nello spazio vettoriale $C^2[-1,1]$, munito del prodotto scalare

$$\langle f, g \rangle = \int_{-1}^{1} f(x)\, g(x)\, dx.$$

Suggerimento: scrivere

$$y_\mu \cdot \left[\left((1-x^2)y_\lambda'\right)' + \lambda y_\lambda\right] = 0$$
$$y_\lambda \cdot \left[\left((1-x^2)y_\mu'\right)' + \mu y_\mu\right] = 0$$

e poi...

Questo esercizio mostra che i polinomi di Legendre, calcolati nell'esercizio precedente, sono un sistema ortogonale nello spazio vettoriale qui considerato. Questo fatto è alla base di una tecnica di approssimazione di funzioni mediante polinomi sull'intervallo $(-1,1)$.

[32] Si veda l'esercizio precedente.

 © 978-88-08-06485-1

4.2 Criteri per la convergenza delle serie trigonometriche

Vogliamo ora dimostrare la Proposizione 7.1, enunciata nel paragrafo 3.1, utile a garantire la convergenza puntuale di certe serie trigonometriche che non convergono totalmente. Cominciamo perciò a richiamare il Criterio di Dirichlet, che abbiamo già enunciato nel contesto delle serie numeriche[33], e che ora dimostreremo:

TEOREMA 7.18 (CRITERIO DI DIRICHLET) *Siano $\{\alpha_n\}, \{\gamma_n\}$ due successioni tali che:*

1. *$\{\gamma_n\}$ è a valori complessi e la sua successione delle somme parziali è limitata; in altre parole, esiste una costante $K > 0$ tale che, detta $C_n = \sum_{k=0}^{n} \gamma_k$, risulta:*

$$|C_n| \leq K \ \forall n.$$

2. *$\{\alpha_n\}$ è a valori reali positivi e tende monotonamente a zero:*

$$\alpha_n > 0 \ \forall n; \quad \alpha_{n+1} \leq \alpha_n \ \forall n; \quad \lim_{n\to\infty} \alpha_n = 0.$$

Allora la serie $\sum \alpha_n \gamma_n$ è convergente.

DIMOSTRAZIONE. Mostreremo che il resto della serie $\sum \alpha_n \gamma_n$ tende a zero, ossia che:

$$\sum_{n=N}^{M} \alpha_n \gamma_n \to 0 \text{ per } N, M \to +\infty. \tag{4.5}$$

Questo è equivalente al fatto che la successione delle somme parziali di questa serie sia una *successione di Cauchy*, ossia che, posto

$$s_N = \sum_{n=0}^{N} \alpha_n \gamma_n,$$

risulti $s_M - s_N \to 0$ per $N, M \to +\infty$. A sua volta, questo fatto implica la convergenza della successione s_N (teorema di completezza[34] di $\mathbb{R}^n$), ossia la convergenza della serie $\sum \alpha_n \gamma_n$.

Per provare (4.5), scriviamo le seguenti identità tra sommatorie:

$$\begin{aligned} \sum_{n=N}^{M} \alpha_n \gamma_n &= \sum_{n=N}^{M} \alpha_n (C_n - C_{n-1}) = \sum_{n=N}^{M} \alpha_n C_n - \sum_{n=N}^{M} \alpha_n C_{n-1} = \\ &= \sum_{n=N}^{M} \alpha_n C_n - \sum_{n=N-1}^{M-1} \alpha_{n+1} C_n = \\ &= \alpha_M C_M - \alpha_N C_{N-1} + \sum_{n=N}^{M-1} (\alpha_n - \alpha_{n+1}) C_n. \end{aligned} \tag{4.6}$$

Dunque possiamo maggiorare, tenendo conto della positività e monotonia di α_n:

$$\left| \sum_{n=N}^{M} \alpha_n \gamma_n \right| \leq \alpha_M |C_M| + \alpha_N |C_{N-1}| + \sum_{n=N}^{M-1} (\alpha_n - \alpha_{n+1}) |C_n| \leq$$

[33] v. volume 1, capitolo 5, Teorema 5.7.
[34] v. capitolo 3, Complementi, paragrafo 9.1.

© 978-88-08-06485-1

per la limitatezza di C_n

$$\leq K \left\{ \alpha_M + \alpha_N + \sum_{n=N}^{M-1} (\alpha_n - \alpha_{n+1}) \right\} =$$

ora la somma è diventata una somma telescopica, in cui gli addendi si cancellano a due a due

$$= K \{\alpha_M + \alpha_N + \alpha_{M-1} - \alpha_{N+1}\} \to 0$$

per $N, M \to \infty$, perché $\alpha_n \to 0$. Ne segue l'asserto. ∎

OSSERVAZIONE Nella (4.6) abbiamo dimostrato in particolare la seguente identità, che val la pena mettere in evidenza perché risulta utile anche in altri contesti:

$$\sum_{n=N}^{M} \alpha_n \gamma_n = \alpha_M C_M - \alpha_N C_{N-1} - \sum_{n=N}^{M-1} (\alpha_{n+1} - \alpha_n) C_n. \tag{4.7}$$

La (4.7) vale per due *qualsiasi* successioni $\{\alpha_n\}_{n=0}^{\infty}, \{\gamma_n\}_{n=0}^{\infty}$ a valori reali o complessi e per ogni coppia di interi $M > N \geq 0$, con $C_n = \sum_{k=0}^{n} \gamma_k$. Essa prende il nome di *formula di sommazione per parti* e la sua analogia con la formula di integrazione per parti è evidente.

Come conseguenza del criterio di Dirichlet, possiamo ora provare la Proposizione 7.1, enunciata nel paragrafo 3.1. Ne riportiamo l'enunciato

PROPOSIZIONE 7.1 *Sia $\{a_n\}$ una successione a valori reali positivi che tende monotonamente a zero. Allora:*

(a) *Le serie trigonometriche*

$$\sum_{n=1}^{\infty} a_n \cos nx, \quad \sum_{n=1}^{\infty} a_n \sin nx$$

convergono per ogni $x \in (0, 2\pi)$. Negli estremi $0, 2\pi$ la serie di coseni può convergere o no, mentre la serie di seni è identicamente nulla.

(b) *Le serie trigonometriche*

$$\sum_{n=1}^{\infty} (-1)^n a_n \cos nx, \quad \sum_{n=1}^{\infty} (-1)^n a_n \sin nx$$

convergono per ogni $x \in [0, 2\pi], x \neq \pi$. In $x = \pi$ la serie di coseni può convergere o no, mentre la serie di seni è identicamente nulla.

Si noti che, anche in caso di convergenza, non è garantito che la somma della serie sia una funzione continua.

DIMOSTRAZIONE. (a) Applichiamo il criterio di Dirichlet alla successione a valori complessi

$$\gamma_n = e^{inx}.$$

Scrivendo:

$$C_n = \sum_{k=0}^{n} \gamma_k = \sum_{k=0}^{n} \left(e^{ix}\right)^k = \frac{1 - e^{ix(n+1)}}{1 - e^{ix}}$$

si ha

$$|C_n| = \left| \frac{1 - e^{ix(n+1)}}{1 - e^{ix}} \right| \leq \frac{2}{|1 - e^{ix}|}.$$

 © 978-88-08-06485-1

Pertanto per ogni $x \neq 2k\pi$, la successione C_n è limitata e possiamo affermare che converge la serie a termini complessi

$$\sum_{n=1}^{\infty} e^{inx} a_n = \sum_{n=1}^{\infty} a_n (\cos nx + i \sin nx) = \sum_{n=1}^{\infty} a_n \cos nx + i \sum_{n=1}^{\infty} a_n \sin nx.$$

Separando parte reale e parte immaginaria, si conclude che per ogni $x \neq 2k\pi$ convergono le due serie a termini reali

$$\sum_{n=1}^{\infty} a_n \cos nx, \qquad \sum_{n=1}^{\infty} a_n \sin nx.$$

(b) Ragionando analogamente con

$$\gamma_n = e^{in(x+\pi)}$$

si ha

$$|C_n| \leq \frac{2}{|1 - e^{i(x+\pi)}|},$$

che è limitata per $x \neq \pi + 2k\pi$ e si conclude in modo analogo, poiché

$$e^{in(x+\pi)} a_n = (-1)^n a_n (\cos nx + i \sin nx).$$ ◆

4.3 Fenomeno di Gibbs

Esaminiamo il grafico di una funzione discontinua insieme a quello della sua somma di Fourier n-esima.

Per esempio, la funzione

$$f(x) = \frac{\pi - x}{2} \text{ in } [0, 2\pi), \text{ periodizzata in } \mathbb{R},$$

ha la seguente serie di Fourier (come si può facilmente controllare):

$$\sum_{k=1}^{\infty} \frac{\sin kx}{k}.$$

Il grafico in figura 7.21 mostra la funzione f insieme alla somma parziale $S_{10}f$, su $(-\pi, \pi)$. In accordo ai teoremi sulla convergenza puntuale, in $x = 0$ (punto di discontinuità a salto per la periodizzata di f) la serie converge a 0 (che è la media tra i limiti destro e sinistro in 0, ossia $\pm\frac{\pi}{2}$). Si osserva che, vicino al punto di salto, la somma $S_n f(x)$ non solo non approssima bene la funzione (perché $S_n f$ è continua e f no) ma addirittura presenta un'oscillazione *maggiore* di quella della funzione discontinua. Più precisamente, si può dimostrare che la differenza tra il massimo assunto da $S_n f$ in prossimità del punto di salto è maggiore di circa il 9% rispetto a quello di f, mentre ci si sarebbe potuto aspettare che $S_n f$, essendo più regolare di f, ne smussasse l'andamento, presentendo un'oscillazione minore. Questo fenomeno, sgradito nelle applicazioni, si presenta costantemente nell'approssimazione di Fourier di una funzione con una discontinuità a salto; può essere previsto teoricamente e prende il nome di *fenomeno di Gibbs*. Vogliamo illustrarlo in modo quantitativamente preciso almeno su questo esempio specifico.

A questo scopo, confrontiamo la somma parziale di Fourier

$$S_n f(x) = \sum_{k=1}^{n} \frac{\sin kx}{k}$$

© 978-88-08-06485-1

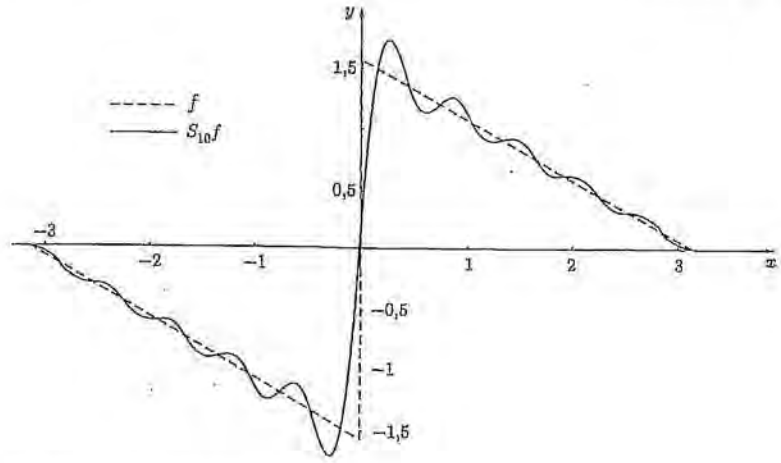


Figura 7.21. La funzione f e la sua somma $S_{10}f$.

con l'integrale definito

$$\int_0^a \frac{\sin t}{t} dt \text{ per } a > 0.$$

Dividendo l'intervallo $[0, a]$ in n parti uguali, possiamo vedere la somma n-esima di Fourier $S_n f(x)$ come una somma di Cauchy-Riemann di f in $[0, a]$:

$$S_n f(x) = \sum_{k=1}^{n} \frac{\sin kx}{k} = \left(\text{per } x = \frac{a}{n}\right)$$
$$= \sum_{k=1}^{n} \frac{\sin\left(\frac{ka}{n}\right)}{\left(\frac{ka}{n}\right)} \cdot \frac{a}{n}.$$

Perciò ci aspettiamo che

$$S_n f\left(\frac{a}{n}\right) \to \int_0^a \frac{\sin t}{t} dt \text{ per } n \to \infty,$$

e questo è vero per ogni $a > 0$. Scegliamo ora come a il punto di massimo assoluto della funzione integrale

$$g(a) = \int_0^a \frac{\sin t}{t} dt.$$

Per individuare tale a, per prima cosa tracciamo mediante il computer il grafico della funzione integrale per $a > 0$ (fig. 7.22).

Dal grafico vediamo che il punto di massimo assoluto della funzione integrale, per $a > 0$, è il punto di massimo relativo di ascissa positiva minima. Calcoliamo tale punto derivando la funzione integrale:

$$g'(a) = \frac{\sin a}{a} = 0 \text{ per } a = k\pi, \ k = 1, 2, 3 \ldots$$

 © 978-88-08-06485-1

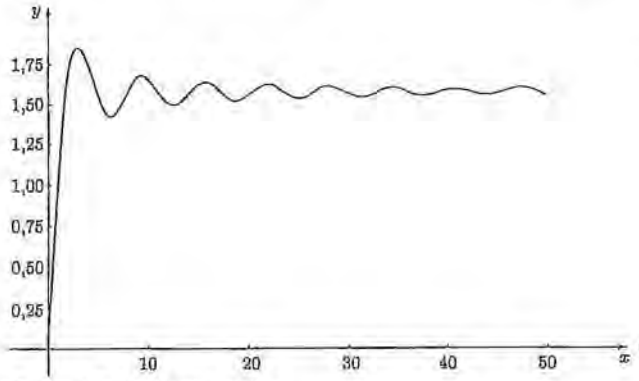


Figura 7.22. Grafico della funzione integrale g.

perciò il punto di massimo assoluto di g è $a = \pi$ e il massimo assoluto è

$$g(\pi) = \int_0^{\pi} \frac{\sin t}{t} dt \simeq 1{,}85194$$

(valore che è stato ottenuto mediante integrazione numerica). Ci aspettiamo perciò che per n grande risulti

$$S_n f\left(\frac{\pi}{n}\right) \simeq 1{,}85194.$$

Questo significa che ogni somma parziale di Fourier di f (per n abbastanza grande) assume nel punto π/n (prossimo a 0) il valore $\simeq 1.85194$, mentre

$$\lim_{x \to 0+} f(x) = \frac{\pi}{2} \simeq 1{,}5708.$$

Per simmetria, per $x \to 0^-$ si avrà un analogo fenomeno, per cui complessivamente il rapporto tra il salto di $S_n f$ e il salto di f vicino all'origine è circa

$$\frac{2 \cdot 1{,}85194}{2 \cdot 1{,}5708} = 1{,}17898,$$

che significa che $S_n f$ ha un salto del 17,8% maggiore di quello di f; la metà di questo valore, 8,9%, indica di quanto il massimo di $S_n f$ è maggiore del massimo di f vicino al salto; il valore è quello che si può prevedere teoricamente per una qualsiasi funzione regolare a tratti che abbia una discontinuità a salto, come abbiamo enunciato all'inizio nel paragrafo.

8 Teoria qualitativa di equazioni differenziali e sistemi

Le equazioni differenziali considerate nel capitolo 1 (equazioni del primo ordine lineari oppure a variabili separabili, del second'ordine lineari a coefficienti costanti) sono di tipo molto speciale e consentono di avere una rappresentazione dell'integrale generale in forma abbastanza esplicita. Questa circostanza non è però frequente nelle applicazioni; spesso le proprietà più rilevanti delle soluzioni di un problema differenziale (come l'esistenza, l'unicità, la stabilità, il comportamento asintotico ecc.) possono essere ottenute solo attraverso lo studio teorico qualitativo; queste informazioni sono fondamentali per procedere poi, con tecniche numeriche, oggi enormemente sviluppate, a una rappresentazione quantitativa.

In questo capitolo riprendiamo, da un punto di vista un po' più avanzato rispetto a quello adottato nel capitolo 1, lo studio delle equazioni differenziali: nel paragrafo 1 consideriamo le equazioni del prim'ordine, discutendo in particolare il problema di Cauchy (v. paragrafo 1.1) e trattando più estesamente le *equazioni autonome*, ossia quelle in cui nell'equazione non compare esplicitamente il tempo (v. paragrafo 1.3, dove viene introdotto il concetto di *diagramma di fase*); nel paragrafo 2 si considerano invece i *sistemi di n equazioni del prim'ordine*, a cui, come vedremo, si possono ricondurre anche le *equazioni di ordine n*; mostreremo come si estendano a questo contesto i risultati teorici fondamentali discussi nel caso delle equazioni del prim'ordine nel paragrafo 1.1. Quindi ci soffermeremo sullo studio dei sistemi autonomi, in particolare bidimensionali (paragrafo 4). Il problema della stabilità per i sistemi autonomi sarà affrontato prima, nel paragrafo 4.2, nel caso lineare (e in particolare bidimensionale) e poi più brevemente nel caso non lineare, nel paragrafo 4.3.

1 EQUAZIONI DEL PRIM'ORDINE

1.1 Problema di Cauchy

Consideriamo un'equazione del prim'ordine in forma normale, cioè del tipo[1]

$$y' = f(t, y). \tag{1.1}$$

[1] Come già nel capitolo 1, nella trattazione teorica la variabile indipendente sarà indicata sempre con t e pensata come variabile tempo.

 © 978-88-08-06485-1

Piuttosto che prendere in esame la totalità delle soluzioni della (1.1), è più conveniente puntare l'attenzione sulle soluzioni del problema di Cauchy associato alla (1.1), cioè sulle funzioni che soddisfano la condizione supplementare (detta *condizione iniziale*)

$$y(t_0) = y_0 \tag{1.2}$$

(con $y_0 \in \mathbb{R}$ assegnato) e che sono soluzioni della (1.1) almeno *localmente*, cioè in un intorno del punto in cui è assegnata la condizione iniziale[2].

Come abbiamo già affermato[3], per equazioni di questo tipo si può assicurare, sotto larghe ipotesi, che il problema di Cauchy (1.1)-(1.2) ammetta un'unica soluzione. Esaminiamo separatamente la questione dell'*esistenza* di almeno una soluzione (richiesta minima perché il problema possa descrivere l'evoluzione di un qualche fenomeno reale) e quella dell'*unicità* di questa soluzione (che mostra come l'evoluzione di un tale fenomeno sia univocamente determinata dalla sua dinamica [cioè dall'equazione (1.1) e dalla condizione iniziale (1.2)].

Enunceremo ora, senza dimostrazione, alcuni risultati fondamentali della teoria classica delle equazioni differenziali. Il primo è il seguente:

TEOREMA 8.1 (DI PEANO, DI ESISTENZA DI UNA SOLUZIONE) *Sia $D \subset \mathbb{R}^2$ un aperto e $f : D \to \mathbb{R}$ continua; sia $(t_0, y_0) \in D$. Allora il problema di Cauchy* (1.1)-(1.2) *ammette almeno una soluzione, definita in un intorno di t_0.*

Che la continuità di f sia una richiesta naturale, si vede osservando la semplice equazione

$$y' = f(t),$$

le cui soluzioni sono le primitive di f; sappiamo che se f ha, per esempio, una discontinuità a salto, non ammette alcuna primitiva[4].

D'altra parte, la sola continuità di f è sufficiente a garantire l'esistenza, ma non l'unicità, della soluzione del problema di Cauchy.

ESEMPIO

Il problema

$$\begin{cases} y' = y^{2/3} \\ y(0) = 0 \end{cases} \tag{1.3}$$

ammette senz'altro la soluzione identicamente nulla $y(t) = 0$. Ne ammette, però, almeno un'altra; infatti l'equazione, a variabili separabili, può essere scritta nella forma:

$$\frac{dy}{y^{2/3}} = dt,$$

dalla quale per integrazione otteniamo

$$3y^{1/3} = t + c.$$

[2] Una soluzione dell'equazione è dunque una funzione y definita in un intorno I di t_0, ivi derivabile e tale che risulti $y'(t) = f(t, y(t))$ per ogni $t \in I$.

[3] capitolo 1, paragrafo 2.1.

[4] v. volume 1, capitolo 6, paragrafo 4.

© 978-88-08-06485-1

La condizione $y(0) = 0$ richiede $c = 0$; la funzione

$$y(t) = \frac{t^3}{27}$$

è dunque anch'essa una soluzione del problema.

Come vedremo nel prossimo teorema, l'unicità della soluzione del problema di Cauchy può essere garantita aumentando le richieste di *regolarità* della funzione $f(t, y)$. Osserviamo che nel problema (1.3) la f è continua, ma la sua derivata $f_y = \frac{2}{3}y^{-1/3}$ è discontinua e illimitata in un intorno del punto iniziale $(t_0, y_0) = (0, 0)$. Per garantire l'unicità dovremo escludere questi fenomeni. Vale infatti il seguente risultato:

TEOREMA 8.2 (DI CAUCHY, DI ESISTENZA E UNICITÀ LOCALE) *Siano f e $\partial f/\partial y$ continue in un aperto $D \subset \mathbb{R}^2$ e sia $(t_0, y_0) \in D$. Allora esiste un intorno I di t_0 tale che il problema di Cauchy* (1.1)-(1.2) *ammette una e una sola soluzione definita in I.*

OSSERVAZIONE (CONDIZIONE DI LIPSCHITZ) La condizione $f_y \in C(D)$ può essere attenuata. Infatti, basta richiedere che il rapporto incrementale di f rispetto a y sia localmente limitato. Più precisamente, chiediamo che, per ogni insieme chiuso e limitato $K \subset D$, esista una costante L_K tale che:

$$|f(t, y_1) - f(t, y_2)| \leq L_K |y_1 - y_2| \quad \text{per ogni } (t, y_1), (t, y_2) \in K.$$

Se ciò accade diremo che f è *localmente lipschitziana*[5] *rispetto a y, uniformemente in t.*

Dunque: *se f è continua in D e localmente lipschitziana rispetto a y, uniformemente in t, vale ancora il teorema precedente.*

Osserviamo che la continuità di f_y implica la locale lipschitzianità di f. Infatti, se f_y è continua in D, per il teorema di Weierstrass essa è limitata in K; detto allora

$$L_K = \max_{(t,y) \in K} \left| \frac{\partial f}{\partial y}(t, y) \right|,$$

per il teorema di Lagrange applicato alla funzione di una variabile $y \mapsto f(t, y)$ si ha, per ogni $(t, y_1), (t, y_2) \in K$ e un opportuno $\overline{y} \in (y_1, y_2)$

$$|f(t, y_1) - f(t, y_2)| = \left| \frac{\partial f}{\partial y}(t, \overline{y})(y_1 - y_2) \right| \leq L_K |y_1 - y_2|,$$

dunque f risulta localmente lipschitziana rispetto a y, uniformemente in t. Il viceversa però non è in generale vero. Esempi standard di funzioni localmente lipschitziane (ma non derivabili) sono funzioni continue il cui grafico è regolare salvo un numero finito di punti angolosi.

[5] dal nome del matematico tedesco Rudolph Otto Sigismund Lipschitz (1832 – 1903).

 © 978-88-08-06485-1

Esempio

In base al teorema 8.2 e all'osservazione precedente, il problema di Cauchy

$$\begin{cases} y' = |y| + 1 \\ y(0) = 0 \end{cases}$$

ammette una e una sola soluzione. Essa è data da

$$y(t) = \left(e^{|t|} - 1\right) \operatorname{sgn}(t),$$

come si può verificare. Si noti che la funzione $f(t,y) = |y| + 1$ è continua e localmente lipschitziana in $\mathbb{R}^2$, uniformemente in t; tuttavia non ha derivata f_y continua in tutto $\mathbb{R}^2$.

Mostriamo come i risultati enunciati nel capitolo 1, paragrafo 2, sul problema di Cauchy per le equazioni del prim'ordine lineari o a variabili separabili, si possono ritrovare come casi particolari dei teoremi appena enunciati.

Esempio

Cominciamo dalle *equazioni lineari del prim'ordine*:

$$y' = a(t)\, y + b(t),$$

per le quali si ha:

$$f(t,y) = a(t)\, y + b(t).$$

Se i coefficienti a, b sono continui in I, la funzione f è continua e

$$\frac{\partial f}{\partial y}(t,y) = a(t)$$

è pure continua; valgono perciò le ipotesi del teorema 8.2, che garantisce esistenza e unicità locale della soluzione del problema di Cauchy. Per ritrovare completamente il risultato enunciato nel capitolo 1 (che garantiva l'esistenza di soluzione in tutto l'intervallo di continuità dei coefficienti, non solo in un intorno del punto iniziale), dovremo prima discutere un ulteriore risultato, di esistenza e unicità globale (si veda il prossimo teorema 8.3).

Consideriamo ora le *equazioni del prim'ordine a variabili separabili*:

$$y' = a(t)\, b(y),$$

per le quali si ha

$$f(t,y) = a(t)\, b(y).$$

Se a, b sono funzioni continue, f è continua e il teorema di Peano garantisce l'esistenza di soluzioni. Se $b \in C^1$, allora:

$$\frac{\partial f}{\partial y}(t,y) = a(t)\, b'(y)$$

è continua, e valgono le ipotesi del teorema di Cauchy, che garantisce anche l'unicità locale. Infine, l'ipotesi $b \in C^1$ si può indebolire chiedendo che b abbia rapporto incrementale limitato (come osservato nel teorema 1.1, capitolo 1), il che significa che $f(t,y)$ è localmente lipschitziana, uniformemente in t.

© 978-88-08-06485-1

OSSERVAZIONE (ULTERIORE REGOLARITÀ DELLA SOLUZIONE) Se $\varphi(t)$ è soluzione del problema di Cauchy (1.1)-(1.2) in un intervallo I, per definizione φ è almeno derivabile e quindi continua in I. Poiché $f(t,y)$ è continua, dall'identità

$$\varphi'(t) = f(t, \varphi(t))$$

leggiamo che φ' è anche continua (perché per il teorema di continuità della funzione composta $f(t, \varphi(t))$ è continua). Quindi, in realtà φ è di classe $C^1(I)$. Una maggior regolarità di f implica poi un'ulteriore regolarità della soluzione φ; ragionando iterativamente, si mostra che:

$$f \in C^k(D) \Longrightarrow \varphi \in C^{k+1}(I);$$

$$f \in C^\infty(D) \Longrightarrow \varphi \in C^\infty(I).$$

Intervallo massimale di esistenza della soluzione

Vogliamo ora capire che cosa si può dire in generale sull'intervallo in cui esiste la soluzione del problema di Cauchy. Introduciamo il problema col prossimo:

Esempio

Si considerino i seguenti due problemi:

$$y' + 2ty^2 = 0 \text{ con condizione}$$

(1.4) $$y(0) = 1 \quad \text{oppure}$$

(1.5) $$y(0) = -1.$$

La funzione $f(t,y)$ è $C^\infty(\mathbb{R}^2)$; perciò i due problemi ammettono ognuno una sola soluzione, che sarà di classe C^∞ in base all'osservazione precedente. Poiché l'equazione è a variabili separabili, si scrive nella forma

$$-\frac{dy}{y^2} = 2t\, dt,$$

che per integrazione dà

$$\frac{1}{y} = t^2 + c;$$

dunque, l'integrale generale è

$$y(t) = \frac{1}{t^2 + c}.$$

La condizione (1.4) implica $c = 1$, che dà la soluzione

$$y(t) = \frac{1}{t^2 + 1},$$

definita su tutta la retta e limitata; la condizione (1.5) implica $c = -1$, che dà la soluzione

$$y(t) = \frac{1}{t^2 - 1},$$

definita solo sull'intervallo $(-1, 1)$ e illimitata.

In generale, possiamo avere qualche informazione sull'intervallo in cui esiste la soluzione ragionando al modo seguente. Sia $t \mapsto \varphi(t; y_0)$ la soluzione[6] del problema di Cauchy (1.1)-(1.2) e consideriamo un rettangolo

$$R = \{(t, y) : t_0 - a \le t \le t_0 + a, y_0 - b \le y \le y_0 + b\}, \quad a, b > 0,$$

centrato nel punto iniziale (t_0, y_0) e contenuto in D. Sia

$$M = \max_{(t,y)\in R} |f(t, y)|$$

(tale massimo esiste perché f è continua sull'insieme R chiuso e limitato). L'equazione (1.1) ci dice che $f(t, y)$ è la pendenza di φ nel punto (t, y), cosicché M e $-M$ saranno la massima e minima pendenza possibile in R. Questo implica che il grafico di φ deve stare nel doppio settore delimitato dalle rette di equazione

$$y = y_0 + M(t - t_0) \quad \text{e} \quad y = y_0 - M(t - t_0).$$

Queste rette tagliano la frontiera di R nei punti di ascissa $t_0 \pm b/M$ se $b/M < a$, oppure in punti di ascissa $t_0 \pm a$ se $b/M \ge a$ (fig. 8.1 a) e b)).

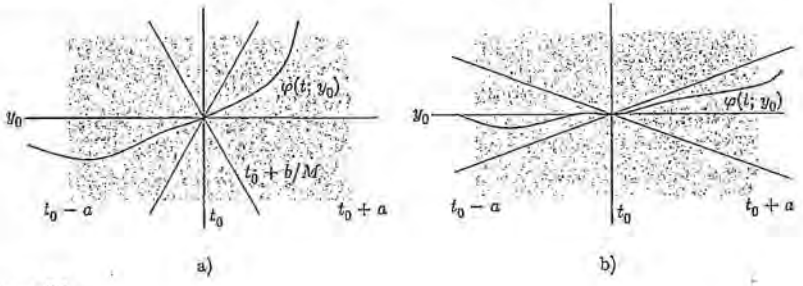


Figura 8.1.

Se poniamo

$$\delta = \min\left(a, \frac{b}{M}\right)$$

possiamo concludere che la soluzione è definita *almeno* nell'intervallo $[t_0 - \delta, t_0 + \delta]$.

L'argomento può ora essere iterato. Consideriamo il punto

$$(t_1, y_1) = (t_0 + \delta, \varphi(t_0 + \delta; y_0))$$

che appartiene alla frontiera di R e quindi all'interno di D (essendo questo un aperto); esisterà allora un rettangolo chiuso R_1 centrato in (t_1, y_1) e contenuto in D, del tipo:

$$R_1 = \{(t, y) : t_1 - a_1 \le t \le t_1 + a_1, y_1 - b_1 \le y \le y_1 + b_1\}, \quad a_1, b_1 > 0.$$

[6] L'abbiamo indicata con $\varphi(t; y_0)$, anche se è una funzione della sola variabile t, per evidenziare che dipende dalla condizione iniziale y_0.

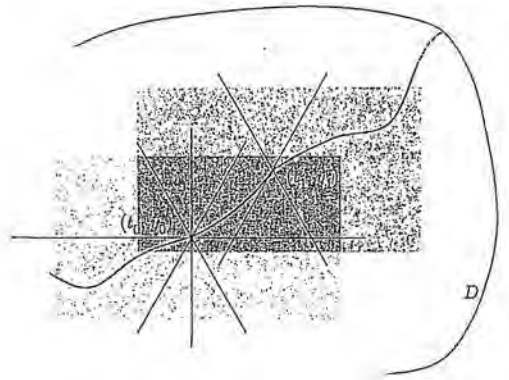


Figura 8.2.

Il teorema 8.2 ci assicura allora che il problema di Cauchy con punto iniziale (t_1, y_1) ha un'unica soluzione $\varphi(t; y_1)$ definita almeno in $I_1 = [t_1 - \delta_1, t_1 + \delta_1]$ (con δ_1 opportuno). Tale soluzione, per l'unicità, deve coincidere con $\varphi(t; y_0)$ in $I \cap I_1 = [t_1 - \delta_1, t_0 + \delta_0]$. Perciò questa soluzione è un'*estensione* della precedente $\varphi(t; y_0)$, che risulta ora definita in $[t_0 - \delta_0, t_1 + \delta_1]$ (estensione destra, figura 8.2).

Iterando il procedimento, si arriva così a definire un intervallo massimale destro $[t_0, t_{\max})$ di esistenza della soluzione $\varphi(t; y_0)$; analogamente si definisce un intervallo massimale sinistro $(t_{\min}, t_0]$, con $-\infty \le t_{\min} < t_{\max} \le +\infty$.

In conclusione:

> nelle ipotesi del teorema 8.2, la soluzione del problema di Cauchy (1.1)-(1.2) è definita in un intervallo massimale $(t_{\min}, t_{\max})$; al di fuori di questo intervallo la soluzione non esiste oppure esce dal dominio D.

Più precisamente, si può dimostrare che per $t \to t^+_{\min}$ e per $t \to t^-_{\max}$ il grafico della soluzione del problema di Cauchy si avvicina indefinitamente al bordo di D.

Si osservi, tuttavia, che anche nel caso in cui sia $D = \mathbb{R}^2$ questo non significa necessariamente che la soluzione sia definita su tutto $\mathbb{R}$: potrebbe anche tendere all'infinito per t tendente a un valore finito.

Dal punto di vista delle applicazioni, sarebbe utile avere informazioni sufficienti a garantire l'esistenza (e unicità) della soluzione in tutto un *prefissato* intervallo $[a, b]$ contenente il punto t_0. A questa esigenza risponde (parzialmente) il prossimo

TEOREMA 8.3 (DI ESISTENZA E UNICITÀ GLOBALE) *Siano f e f_y di classe $C^0(\overline{S})$, essendo $\overline{S}$ la striscia*

$$\overline{S} = [a, b] \times \mathbb{R};$$

inoltre esistano due numeri positivi h, k (in generale dipendenti da a, b) per cui risulti:

$$|f(t, y)| \le h + k|y| \quad \textit{per ogni } (t, y) \in \overline{S}. \tag{1.6}$$

Allora ogni soluzione dell'equazione (1.1) *è definita su tutto l'intervallo $[a, b]$.*

 © 978-88-08-06485-1

Si noti che le ipotesi di questo teorema implicano quelle del teorema 8.2 in $[a,b]\times[c,d]$ per ogni intervallo $[c,d]\subset\mathbb{R}$.

A sua volta, la condizione (1.6) è soddisfatta per esempio se:

i) f è limitata in $\overline{S}$:

$$|f(t,y)|\le h \quad \text{per ogni } (t,y)\in\overline{S} \tag{1.7}$$

(in questo caso vale la (1.6) con $k=0$), oppure se

ii) f_y è limitata in $\overline{S}$:

$$|f_y(t,y)|\le k \quad \text{per ogni } (t,y)\in\overline{S}. \tag{1.8}$$

Infatti, se vale la (1.8), si può scrivere:

$$|f(t,y)| = \left|f(t,0)+\int_0^y f_y(t,s)\,ds\right| \le M + k|y|$$

con

$$M=\max_{[a,b]}|f(t,0)|,$$

e quindi vale (1.6).

Tuttavia queste condizioni sono sufficienti, ma niente affatto necessarie per l'esistenza globale della soluzione del problema di Cauchy, come dimostra l'Esempio 1.5, con la condizione (1.4): in quel caso la (1.6) non vale, eppure la soluzione esiste su tutto $\mathbb{R}$.

Si osservi, infine, che per un'equazione lineare del prim'ordine (v. capitolo 1, paragrafo 2.3), si ha

$$f(t,y)=a(t)\,y+b(t),$$

perciò la (1.6) è soddisfatta se i coefficienti $a(t)$ e $b(t)$ sono continui nell'intervallo $[a,b]$. Ritroviamo quindi un risultato enunciato nel capitolo 1: *la soluzione del problema di Cauchy per un'equazione lineare del prim'ordine è definita su tutto l'intervallo in cui i coefficienti sono continui.* (Si confronti con quanto osservato nell'Esempio 1.3).

Vediamo invece un paio di esempi di applicazione del teorema precedente a equazioni non lineari:

ESEMPIO

Il problema

$$\begin{cases} y' = \cos(ty) \\ y(t_0)=y_0 \end{cases}$$

ammette un'unica soluzione (per ogni coppia di dati $(t_0,y_0)\in\mathbb{R}^2$), definita su tutta la retta e di classe C^∞. Infatti, essendo

$$|\cos(ty)|\le 1 \quad \text{per ogni } (t,y)\in\mathbb{R}^2,$$

la condizione (1.7) è soddisfatta in tutto il piano con $h=1$.

Il problema

$$\begin{cases} y' = \dfrac{y^5}{1+y^4} \\ y(0)=1 \end{cases}$$

ammette un'unica soluzione (per ogni coppia di dati $(t_0,y_0)\in\mathbb{R}^2$), definita su tutta la retta e di classe C^∞. Infatti la condizione (1.6) è soddisfatta in tutto il piano con $h=0$ e $k=1$:

$$|f(t,y)| = |y|\,\frac{y^4}{1+y^4} \le |y| \quad \text{per ogni } (t,y)\in\mathbb{R}^2.$$

Dipendenza continua dai dati

Un'altra questione molto naturale che riguarda la soluzione di un problema di Cauchy è la seguente. Supponiamo che f soddisfi le ipotesi per cui il problema (1.1)-(1.2) ha una e una sola soluzione. Sia per esempio $\varphi(t;y_0)$ la soluzione del problema con condizione iniziale $y(t_0)=y_0$, e supponiamo che sia definita in (t_0-T,t_0+T). Se ora andiamo a risolvere il problema di Cauchy per la stessa equazione, ma con una condizione iniziale di poco diversa, ossia imponendo $y(t_0)=y_1$ con y_1 "vicina" a y_0, la nuova soluzione $\varphi(t;y_1)$ sarà definita su un nuovo intervallo, diciamo (t_0-T_1,t_0+T_1); potremo garantire che le due soluzioni $\varphi(t;y_0)$ e $\varphi(t;y_1)$ si mantengano "vicine" almeno su un intervallo comune di definizione? E inoltre, se andiamo a risolvere il problema di Cauchy con un'equazione leggermente diversa, ossia $y'=f_1(t,y)$ dove f_1 è una funzione "vicina" a f, potremo garantire ancora che la nuova soluzione si mantenga vicina alla precedente? Queste domande sono estremamente naturali per varie ragioni. Anzitutto, dal punto di vista applicativo, la condizione iniziale y_0 di un sistema fisico, così come la sua legge di evoluzione f, sono sempre note con una certa approssimazione; perciò è di vitale importanza sapere che, se sbagliamo di poco a valutare quantitativamente gli ingredienti y_0, f del nostro modello, sbaglieremo conseguentemente di poco nel calcolare la soluzione.

Inoltre, il fatto che il modello matematico che facciamo di un certo fenomeno fisico garantisca le proprietà di *esistenza* della soluzione, *unicità* della soluzione, *dipendenza continua* della soluzione dagli ingredienti del problema si può vedere come indice del fatto che il nostro modello è stato ben formulato. Se qualcuna di queste proprietà venisse a cadere, significherebbe o che stiamo cercando di descrivere un fenomeno intrinsecamente molto complesso e instabile, oppure che il nostro modello matematico è inadeguato per qualche aspetto e va modificato.

Il risultato di dipendenza continua si può formulare così:

TEOREMA 8.4 *Siano $f_1(t,y)$, $f_2(t,y)$ due funzioni continue in un aperto $D\subset\mathbb{R}^2$, localmente lipschitziane rispetto a y, uniformemente in t. Siano $(t_0,y_1),(t_0,y_2)\in D$ e siano $\varphi_1(t)$, $\varphi_2(t)$, rispettivamente, le soluzioni dei problemi di Cauchy:*

$$\begin{cases} y' = f_1(t,y) \\ y(t_0)=y_1 \end{cases} \qquad \begin{cases} y' = f_2(t,y) \\ y(t_0)=y_2. \end{cases}$$

Allora, se nell'intervallo $[t_0,t_0+T]$ entrambe le soluzioni sono definite, vale la disuguaglianza:

$$|\varphi_1(t)-\varphi_2(t)| \le (|y_1-y_2|+(t-t_0)M)\,e^{L(t-t_0)} \quad \text{per ogni } t\in[t_0,t_0+T] \tag{1.9}$$

dove

$$M=\max_{(t,y)\in K}|f_1(t,y)-f_2(t,y)|,$$

L è la costante di lipschitzianità di f_1 e f_2 su K, e K è un insieme chiuso e limitato contenuto in D e contenente i grafici delle soluzioni $\varphi_1(t)$, $\varphi_2(t)$.

 © 978-88-08-06485-1

Si osservi che la stima di dipendenza continua (1.9) esprime proprio il fatto che, se y_1 è vicino a y_2 (quindi $|y_1 - y_2|$ è piccolo) e f_1 è vicina a f_2 (quindi M è piccolo), allora $\varphi_1(t)$ è vicina a $\varphi_2(t)$ in $[t_0, t_0 + T]$. In particolare, per $t = t_0$ troviamo:

$$|\varphi_1(t) - \varphi_2(t)| \leq |y_1 - y_2|,$$

che dice nuovamente che la soluzione del problema di Cauchy è unica (infatti per $y_1 = y_2$ le due soluzioni coincidono); invece per $f_1 = f_2$ (e t qualsiasi) troviamo:

$$|\varphi_1(t) - \varphi_2(t)| \leq |y_1 - y_2| e^{L(t-t_0)} \leq |y_1 - y_2| e^{LT} \text{ per ogni } t \in [t_0, t_0 + T],$$

che dice che le due soluzioni si discostano per non più di $c|y_1 - y_2|$.

Il fatto che la costante c abbia la forma e^{LT} ha un significato profondo: ci dice che le soluzioni relative a condizioni iniziali diverse, ma vicine, si mantengono "ragionevolmente vicine" per intervalli di tempo limitati, ma possono allontanarsi senza alcun controllo per tempi lunghi: infatti, la costante di dipendenza continua cresce esponenzialmente con T.

La dipendenza continua si ha quindi *per tempi finiti*, mentre per $t \to +\infty$ può accadere che soluzioni relative a condizioni iniziali vicine abbiano comportamenti radicalmente diversi[7]. È questo il motivo per cui lo studio del comportamento asintotico delle soluzioni per tempi lunghi è un capitolo importante e profondo dell'analisi delle equazioni differenziali, la cui problematica non è affatto chiusa dal sapere che la soluzione per tempi finiti dipende con continuità dai dati. Nel seguito del capitolo daremo ampio spazio proprio a questo problema.

Esercizi

Per i seguenti problemi di Cauchy, trovare le soluzioni con i loro intervalli massimali di definizione.

(a) $\begin{cases} y' = 1 + y^2 \\ y(0) = 0 \end{cases}$ (b) $\begin{cases} y' = y^2 \\ y(0) = 1 \end{cases}$ (c) $\begin{cases} y' = -y^2 \\ y(0) = 1 \end{cases}$

Per il problema

$$\begin{cases} y' = \operatorname{arctg} y - \frac{1}{t} \\ y(1) = b \quad (b > 0), \end{cases}$$

(a) discutere unicità ed esistenza locale;
(b) determinare l'intervallo massimale di esistenza;
(c) tracciare il grafico qualitativo delle soluzioni.

Dato il problema

$$\begin{cases} y' = \sqrt[3]{y} \\ y(0) = k \quad (k \in \mathbb{R}), \end{cases}$$

(a) trovare i valori di k per cui il problema ha esattamente una soluzione locale;
(b) mostrare che per gli altri valori di k il problema ha più di una soluzione.

[7] si veda l'Esercizio 8.

© 978-88-08-06485-1

Studiare esistenza e unicità di soluzione per il problema:

$$\begin{cases} y' = \frac{2}{t} y + 2t\sqrt{y} \\ y(1) = 0. \end{cases}$$

Risolvere dettagliatamente il problema di Cauchy dell'esempio 1.2.

Dimostrare la seguente affermazione fatta in precedenza nell'osservazione a pagina 421: se φ è soluzione dell'equazione $y' = f(t, y)$ e $f \in C^k(D)$, allora $\varphi \in C^{k+1}(I)$; se $f \in C^\infty(D)$, allora $\varphi \in C^\infty(I)$.

Si consideri il problema di Cauchy:

$$\begin{cases} y' = |t| \left(e^{ty} - 1\right) \\ y(0) = 2. \end{cases}$$

Senza risolverlo, si dica qual è la massima regolarità che possiamo aspettarci dalla soluzione.

(i) Si considerino i due problemi di Cauchy:

$$\begin{cases} y' = y \\ y(0) = c_i \end{cases} \quad i = 1, 2.$$

Mostrare che, per ogni $\varepsilon > 0$, esistono condizioni iniziali $c_1, c_2 \in \mathbb{R}$ tali che $|c_1 - c_2| < \varepsilon$ ma, dette φ_i le soluzioni dei rispettivi problemi, risulta:

$$\lim_{t \to +\infty} |\varphi_1(t) - \varphi_2(t)| = +\infty.$$

Pertanto, la dipendenza continua della soluzione dalle condizioni iniziali vale solo per tempi finiti.

(ii) Si provi poi lo stesso per i due problemi:

$$\begin{cases} y' = c_i y \\ y(0) = 1 \end{cases} \quad i = 1, 2.$$

Pertanto, anche la dipendenza continua dalla legge di evoluzione vale solo per tempi finiti.

1.2 Alcune classi di equazioni del prim'ordine

Mostriamo ora alcune classi di equazioni del prim'ordine, diverse da quelle già incontrate nel capitolo 1, per le quali è possibile talvolta arrivare a una soluzione esplicita.

Equazioni di Bernoulli

Sono equazioni del tipo:

$$y' = P(t) y + Q(t) y^\alpha$$

con $\alpha \in \mathbb{R}$, $\alpha \neq 0, \alpha \neq 1$ (altrimenti si tratta di un'equazione lineare, che conosciamo già). Supponiamo P, Q continue in un intervallo $I \subseteq \mathbb{R}$.

Se $\alpha > 0$, $y = 0$ è una soluzione dell'equazione.

Se $y \neq 0$, dividendo per y^α si ottiene

$$y^{-\alpha} y' = P(t) y^{1-\alpha} + Q(t).$$

 © 978-88-08-06485-1

Ponendo

$$z = y^{1-\alpha},$$

si ha:

$$z' = (1-\alpha)\, y^{-\alpha} y'$$

e perciò z è soluzione dell'equazione lineare

$$z' = (1-\alpha)\, P(t)\, z + (1-\alpha)\, Q(t).$$

Trovata z con la formula[8]

$$z(t) = ce^{-A(t)} + (1-\alpha)\, e^{-A(t)} \int Q(t)\, e^{A(t)} dt,$$

dove $A(t) = (\alpha - 1) \int P(t) dt$, si trova:

$$y(t) = z(t)^{\frac{1}{1-\alpha}}.$$

(Nel caso in cui $(1-\alpha)$ è un intero pari o un razionale con numeratore pari, c'è anche la soluzione $y = -z(t)^{\frac{1}{1-\alpha}}$).

Esempio

Integrare le seguenti equazioni di Bernoulli:

(a) $y' + y \sin t + y^{\alpha} \sin 2t = 0$ (b) $(1-t^2)\, y' - ty - \alpha t y^2 = 0$

(c) $y - y' \cos t - y^2 (\cos t)(2 - \sin t) = 0$ (d) $ty' + y(1 - ty^n) = 0 \quad n \in \mathbb{N}$.

Equazioni differenziali esatte

Sono equazioni del tipo:

$$y' = -\frac{P(t,y)}{Q(t,y)}, \tag{1.10}$$

dove P, Q sono continue in un dominio $A \subseteq \mathbb{R}^2$, $Q \neq 0$, e tali che la forma differenziale

$$\omega = P(t,y)\, dt + Q(t,y)\, dy$$

sia esatta, esista cioè $F = F(t,y)$, funzione potenziale[9], tale che:

$$\frac{\partial F}{\partial t} = P \quad \text{e} \quad \frac{\partial F}{\partial y} = Q \quad \text{in } A.$$

Ricordiamo che, se A è semplicemente connesso in $\mathbb{R}^2$ e $\partial Q/\partial t, \partial P/\partial y$ sono continue in A, ω è esatta se e solo se, in A, risulta:

$$\frac{\partial Q}{\partial t} = \frac{\partial P}{\partial y}.$$

[8] v. capitolo 1, formula (2.12).

[9] v. capitolo 6, paragrafo 1.4

© 978-88-08-06485-1

In tal caso, un potenziale F può essere calcolato con i metodi discussi nel capitolo 6, paragrafo 1.5. In particolare, possiamo scrivere:

$$F(t,y) = \int_{t_0}^{t} P(s,y)\, ds + \int_{y_0}^{y} Q(t_0,u)\, du.$$

Trovato F, l'integrale generale della (1.10) sarà assegnato, in forma implicita, dalla formula seguente:

$$F(t,y) = c.$$

In particolare, l'equazione

$$F(t,y) = 0$$

seleziona quella soluzione che soddisfa la condizione di Cauchy $y(t_0) = y_0$.

Esempio

Si voglia risolvere il problema di Cauchy:

$$\begin{cases} y' = -\dfrac{t+y}{t-3y} \\ y(0) = 1. \end{cases}$$

Qui $P(t,y) = t + y, Q(t,y) = t - 3y \neq 0$ nei due semipiani $t > 3y$ e $t < 3y$. Poiché il punto iniziale $(t_0, y_0) = (0,1)$ appartiene al semipiano $t < 3y$, scegliamo quest'ultimo come dominio A. Essendo poi

$$\frac{\partial Q}{\partial t} = 1 = \frac{\partial P}{\partial y},$$

la forma differenziale $\omega = (t+y)\, dt + (t-3y)\, dy$ è esatta in A. La soluzione è assegnata implicitamente dalla formula:

$$\int_0^t (s+y)\, ds + \int_1^y (-3u)\, du = 0$$

cioè

$$\frac{t^2}{2} + yt - \frac{3}{2}y^2 + \frac{3}{2} = 0$$

che, risolta in y, ricordando ancora che vogliamo $y(0) = 1$, dà:

$$y(t) = \frac{1}{3}\left(t + \sqrt{4t^2 + 9}\right),$$

soluzione del problema di Cauchy.

Equazioni differenziali omogenee

Sono equazioni del tipo:

$$y' = f\left(\frac{y}{t}\right) \tag{1.11}$$

o, alternativamente, $y' = g(t,y)$ con g omogenea di grado zero[10], ossia:

$$g(\lambda t, \lambda y) = g(t,y) \text{ per ogni } \lambda \in \mathbb{R}.$$

[10] v. capitolo 3, paragrafo 9.2.

 © 978-88-08-06485-1

Supponiamo f, g continue in un intervallo e in un aperto di $\mathbb{R}^2$, rispettivamente. Ponendo

$$z(t) = \frac{y(t)}{t} \quad \text{ossia} \quad y(t) = tz(t) \tag{1.12}$$

si trova $y' = z + tz'$, che, sostituita nella (1.11), dà per z l'equazione:

$$z + tz' = f(z),$$

cioè l'equazione a variabili separabili

$$z' = \frac{f(z) - z}{t}.$$

Risolta[11] l'equazione in z, si determina y dalla (1.12).

Esercizio

10 Integrare le seguenti equazioni omogenee:

$$(a)\quad y' = \frac{t^3 + y^3}{ty^2}; \qquad (b)\quad y' = \frac{y}{t + \sqrt{t^2 + y^2}}.$$

1.3 Equazioni autonome. Diagrammi di fase. Stabilità

In questo paragrafo consideriamo le equazioni *autonome*, cioè del tipo

$$y' = f(y) \tag{1.13}$$

con $f : I \to \mathbb{R}$, $I \subseteq \mathbb{R}$ intervallo; assumeremo $f \in C^1(I)$. In altre parole, f non dipende esplicitamente da t.

Notiamo che questa è una particolare equazione a variabili separabili[12]; se impostiamo la risoluzione scrivendo

$$\int \frac{dy}{f(y)} = \int dt,$$

otteniamo (indicando con g una primitiva di $1/f$):

$$g(y) = t + c$$

e quindi l'integrale generale ha la forma:

$$y(t) = h(t + c)$$

per h funzione opportuna e c costante arbitraria. Questo implica la seguente proprietà di *invarianza per traslazioni temporali*:

[11] coi metodi visti nel capitolo 1, paragrafo 2.2.
[12] v. capitolo 1, paragrafo 2.2.

© 978-88-08-06485-1

se φ è una soluzione di (1.13) in (a, b), anche $\psi(t) = \varphi(t + \tau)$, con $\tau \in \mathbb{R}$ fissato, è una soluzione in $(a - \tau, b - \tau)$.

Tenuto conto della precedente osservazione, nello studio del problema di Cauchy possiamo senz'altro assumere $t_0 = 0$; la soluzione che soddisfa la condizione $y(0) = y_0$, cioè la soluzione che "parte" da y_0, sarà indicata con $\varphi(t; y_0)$. Tale soluzione esiste senz'altro, almeno in un intorno di $t_0 = 0$, in base al teorema 8.2, o anche per quanto abbiamo dimostrato nel capitolo 1, paragrafo 4, riguardo alle equazioni a variabili separabili.

Diamo ora un'importante

DEFINIZIONE 8.1 Si dice *soluzione di equilibrio* o *stato stazionario di* (1.13) ogni (eventuale) zero di $f : \overline{y} \in \mathbb{R}$ è uno stato stazionario se $f(\overline{y}) = 0$; a $\overline{y}$ corrisponde allora la soluzione costante (soluzione di equilibrio) della (1.13)

$$y(t) = \overline{y} \quad \text{per } t \in \mathbb{R}.$$

Esempio

Consideriamo l'*equazione logistica*, studiata nell'esempio 2.6 del capitolo 1, cioè l'equazione (1.13) con

$$f(y) = \varepsilon y \left(1 - \frac{1}{k} y\right), \quad \text{con } \varepsilon, k > 0 \text{ costanti.}$$

Le soluzioni $y = 0$ e $y = k$ sono le soluzioni di equilibrio. Tutte le altre soluzioni che partono da uno stato $y_0 > 0$ (con $y_0 \neq k$) tendono a k per $t \to +\infty$. Infatti, come abbiamo visto, tali soluzioni sono date da:

$$y(t) = \frac{k y_0 e^{\varepsilon t}}{k - y_0 + y_0 e^{\varepsilon t}}. \tag{1.14}$$

Si può dire che la soluzione $y(t) = k$ "attrae" asintoticamente ogni altra soluzione. Al contrario, lo stato stazionario $y(t) = 0$ non ha questo potere di attrazione: ogni soluzione che parte da uno stato $y_0 \neq 0$, anche molto vicino a y_0, se ne allontana, al crescere di t, per tendere a k.

La discussione dell'esempio precedente è stata sviluppata sulla base di un'esplicita rappresentazione delle soluzioni dell'equazione logistica (cioè la (1.14)) ma, per le equazioni autonome, anche quando formule esplicite per le soluzioni non sono disponibili, è possibile ricavare informazioni attraverso il cosiddetto *diagramma di fase*, strumento fondamentale che ora introdurremo.

Diagramma di fase

Rappresentiamo (fig. 8.3) su un piano, che chiameremo *piano delle fasi*, il grafico di f: sull'asse delle ascisse (detto *asse delle fasi*) collochiamo i valori della *variabile di stato* y; le intersezioni del grafico di f con l'asse delle fasi ci danno gli stati di equilibrio (cfr. y_1 e y_2 in figura 8.3).

Per determinare il comportamento qualitativo delle soluzioni ragioniamo così. Indichiamo, come sopra, con $\varphi(t; y_0)$ la soluzione della (1.13) che parte dallo stato y_0;

 © 978-88-08-06485-1

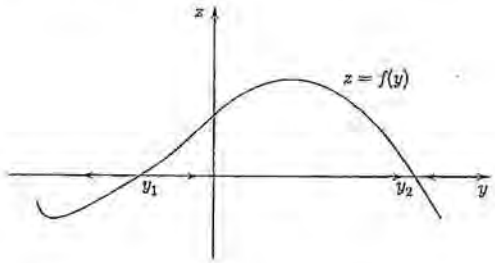


Figura 8.3.

con riferimento alla figura 8.3, se $y_0 < y_1$, essendo $f(y_0) < 0$, sarà, in base alla (1.13), $f(y_0) = \varphi'(0; y_0) < 0$; perciò φ parte con pendenza negativa e decresce lungo l'asse y allontanandosi da y_1. Se $y_1 < y_0 < y_2$ risulta invece $\varphi'(0; y_0) > 0$; perciò $\varphi(t; y_0)$ ora parte con pendenza positiva e aumenta lungo l'asse y, allontanandosi da y_1 e avvicinandosi a y_2, *senza però mai raggiungere* y_2 (se ciò avvenisse al tempo $t = \tau$, avremmo due diverse soluzioni $\psi(t) = y_2$ e $\varphi(t; y_0)$ dello stesso problema di Cauchy, con condizione $y(\tau) = y_2$). Infine, se $y_0 > y_2$, la soluzione $\varphi(t; y_0)$, essendo $\varphi'(0; y_0) < 0$, si avvicina indefinitamente a y_2. La figura 8.3, che rappresenta graficamente la precedente discussione, si chiama *diagramma di fase* per l'equazione (1.13).

Stabilità

Possiamo ora precisare la nozione di *stabilità* di un punto di equilibrio.

DEFINIZIONE 8.2 Diremo che uno stato stazionario $\overline{y}$ è:

(i) *stabile* se per ogni $\varepsilon > 0$ esiste $\delta = \delta(\varepsilon) > 0$ tale che, se $|y_0 - \overline{y}| < \delta$, allora la soluzione $\varphi(t; y_0)$ esiste per ogni $t \geq 0$ e soddisfa

$$|\varphi(t; y_0) - \overline{y}| < \varepsilon \text{ per ogni } t \geq 0;$$

in altre parole: una soluzione che parte in prossimità di $\overline{y}$ rimane sempre (cioè per ogni $t \geq 0$) vicino a $\overline{y}$.

(ii) *asintoticamente stabile* se è stabile e inoltre esiste $\delta_1 > 0$ tale che, se $|y_0 - \overline{y}| < \delta_1$, allora risulta

$$\varphi(t; y_0) \to \overline{y} \quad \text{per } t \to +\infty; \tag{1.15}$$

in questo caso il punto $\overline{y}$ viene detto un *attrattore (locale)* e il più grande intorno U di $\overline{y}$ tale che se $y_0 \in U$ vale la (1.15) si dice *bacino di attrazione* di $\overline{y}$; se il bacino di attrazione di $\overline{y}$ coincide col dominio di f, si dice che $\overline{y}$ è un *attrattore globale*.

(iii) Infine, diremo che uno stato stazionario $\overline{y}$ è *instabile* se non è stabile, cioè se esistono soluzioni che, pur partendo in prossimità di $\overline{y}$, non rimangono sempre vicine a $\overline{y}$. Naturalmente ciò non significa che tutte le soluzioni si allontanino da $\overline{y}$; se invece accade quest'ultima cosa, il punto di equilibrio si chiama *repulsore*.

© 978-88-08-06485-1

ESEMPI

Consideriamo il caso elementare dell'equazione lineare autonoma

$$y' = ay + b \text{ con } a, b \in \mathbb{R} \text{ (costanti)},$$

$a \neq 0$. Dal diagramma di fase si vede subito che l'unico punto di equilibrio $\overline{y} = -b/a$ è un attrattore globale se $a < 0$, un repulsore se $a > 0$.

Riprendiamo ancora l'Esempio 1.9 dell'equazione logistica. Si osserva che il punto $\overline{y} = 0$ è un repulsore, mentre $\overline{y} = k$ è un attrattore, con bacino di attrazione costituito dalla semiretta $(0, +\infty)$.

OSSERVAZIONE (DIPENDENZA CONTINUA E STABILITÀ) La dipendenza continua della soluzione di un problema di Cauchy dai dati iniziali, discussa nel paragrafo 1.1, ci permette di affermare che, se $\overline{y}$ è uno stato stazionario, fissato un tempo $T > 0$ e un $\varepsilon > 0$ qualsiasi, qualunque soluzione parta da una condizione iniziale y_0 sufficientemente vicina a $\overline{y}$ rimarrà fino al tempo T a distanza minore di ε da $\overline{y}$. Il problema della stabilità, tuttavia, riguarda il comportamento delle soluzioni "per tempi lunghi" (non solo $< T$), perciò non è in alcun modo risolto dal fatto di sapere che la soluzione dipende con continuità dalla condizione iniziale.

Anche se lo studio del diagramma di fase per equazioni autonome del primo ordine è generalmente abbastanza semplice, può essere utile avere un criterio generale di stabilità asintotica o di instabilità, di uso pratico immediato. Ciò si può ottenere rimpiazzando in vicinanza dei punti di equilibrio il grafico di f con la retta tangente al grafico (*metodo di linearizzazione*; ricordiamo che per ipotesi f è di classe C^1) e discutendo il caso lineare corrispondente. Vale infatti il seguente

TEOREMA 8.5 (CRITERIO DI STABILITÀ/INSTABILITÀ) *Sia I un intervallo aperto, $\overline{y} \in I$, $f \in C^1(I)$, $f(\overline{y}) = 0$, $f'(\overline{y}) \neq 0$. Allora:*

i) se $f'(\overline{y}) < 0$, $\overline{y}$ è asintoticamente stabile;

ii) se $f'(\overline{y}) > 0$, $\overline{y}$ è instabile.

DIMOSTRAZIONE. (i) Sia $f'(\overline{y}) < 0$. Poiché f' è continua, esiste un intorno $U = (\overline{y} - h, \overline{y} + h)$ di $\overline{y}$ in cui f' resta negativa: $f'(t) \leq -k < 0$ per ogni $t \in U$. Per $y_0 \in U$, consideriamo la soluzione $\varphi(t; y_0)$ e proviamo che

$$\varphi(t; y_0) \to \overline{y} \quad \text{per } t \to +\infty.$$

Supponiamo prima che $y_0 \in (\overline{y}, \overline{y} + h)$; allora $\varphi(0; y_0) > \overline{y}$. Questo implica che per ogni $t > 0$ per cui $\varphi(t; y_0)$ è definita, risulta $\varphi(t; y_0) > \overline{y}$; se così non fosse, per un certo istante $t^* > 0$ avremmo $\varphi(t^*; y_0) = \overline{y}$, e il problema di Cauchy con condizione $y(t^*) = \overline{y}$ avrebbe due soluzioni distinte (la costante $\overline{y}$ e $\varphi(t^*; y_0)$), contro l'unicità di tale soluzione.

Osserviamo che risulta anche $\varphi(t; y_0) < y_0$ per ogni $t > 0$ per cui $\varphi(t; y_0)$ è definita. Infatti, per l'argomento appena visto, per nessun $t > 0$ può essere $\varphi(t; y_0) = y_0$, dunque $\varphi(t; y_0) - y_0$ è sempre positivo o sempre negativo per $t > 0$, e poiché

$$\varphi'(0; y_0) = f(\varphi(0; y_0)) = f(y_0) < 0,$$

dev'essere $\varphi(t; y_0) < y_0$ per ogni $t > 0$ per cui $\varphi(t; y_0)$ è definita.

 © 978-88-08-06485-1

Concludiamo che per ogni $t > 0$ per cui $\varphi(t; y_0)$ è definita, risulta:

$$\overline{y} \leq \varphi(t; y_0) < y_0.$$

Questo implica che in realtà $\varphi(t; y_0)$ è definita per ogni $t > 0$ (perché $\varphi(t; y_0)$ si può sempre prendere come nuova condizione iniziale per prolungare la soluzione oltre t).

Definiamo ora

$$w(t) = \varphi(t; y_0) - y_0.$$

Si ha, ricordando che $f(\overline{y}) = 0$ e applicando a w il teorema di Lagrange,

$$w'(t) = \varphi'(t; y_0) = f(\varphi(t; y_0)) = f(\varphi(t; y_0)) - f(\overline{y}) = \\ = f'(\xi(t))[\varphi(t; y_0) - \overline{y}] = f'(\xi(t))\, w(t)$$

per qualche punto $\xi(t) \in (\overline{y}, \varphi(t; y_0))$. In particolare, $\xi(t) < \varphi(t; y_0) < y_0$, perciò

$$f'(\xi(t)) \leq -k < 0$$

e, poiché $w(t) > 0$,

$$w'(t) \leq -k w(t).$$

Integrando da 0 a t si ottiene:

$$0 < w(t) \leq (y_0 - \overline{y})\, e^{-kt} \to 0 \text{ per } t \to +\infty.$$

Perciò

$$\varphi(t; y_0) \to y_0 \text{ per } t \to +\infty.$$

Se $y_0 \in (\overline{y} - h, \overline{y})$, l'argomento è simile e la conclusione la stessa.

(ii). Se $f'(\overline{y}) > 0$, ancora per continuità di f' esiste un intorno $V = (\overline{y} - k, \overline{y} + k)$ di $\overline{y}$ in cui f' resta positiva. Per $y_0 \in V$, si considera la soluzione $\varphi(t; y_0)$; un argomento simile a quello già visto implica che per ogni $t > 0$ risulta:

$$\varphi(t; y_0) > y_0 > \overline{y}$$

pertanto $\varphi(t; y_0)$ non può tendere a $\overline{y}$ per $t \to +\infty$ e $\overline{y}$ è instabile. ◆

Esercizi

Per le due equazioni seguenti tracciare e discutere il diagramma di fase:

$$(a) \quad y' = e^y - 1; \quad (b) \quad y' = 1 - e^y.$$

Quindi risolvere esplicitamente le equazioni con condizione iniziale $y(0) = y_0$ e determinare l'intervallo massimale di definizione della soluzione in dipendenza da y_0.

Trovare gli stati stazionari dell'equazione

$$y' = y^2(y - 2)$$

e studiare la loro stabilità in base al diagramma di fase.

Trovare gli stati di equilibrio, al variare del parametro reale m, dell'equazione

$$y' = y^3 - 2y^2 + my$$

 © 978-88-08-06485-1

e studiare la loro stabilità.

Trovare gli stati stazionari dell'equazione

$$y' = \sin y$$

e studiare la loro stabilità in base al diagramma di fase.

Trovare gli stati stazionari dell'equazione

$$y' = \operatorname{tg} y$$

e studiare la loro stabilità in base al diagramma di fase. Quindi risolvere esplicitamente l'equazione con condizione iniziale $y(0) = y_0$ e determinare l'intervallo massimale di definizione della soluzione in dipendenza da y_0.

2 PROBLEMA DI CAUCHY PER SISTEMI DI n EQUAZIONI DEL PRIM'ORDINE O EQUAZIONI DI ORDINE n

Generalizziamo ora la teoria vista nel paragrafo 1.1 per le equazioni del prim'ordine al caso più significativo dei sistemi di n equazioni del prim'ordine. Come vedremo subito, anche le equazioni di ordine n si possono riscrivere come sistemi di n equazioni del prim'ordine. Perciò la teoria che svilupperemo nel seguito si può anche vedere come il quadro teorico generale che fa da sfondo alle equazioni differenziali discusse nel capitolo 1 e ad altre non lineari.

Sistemi di equazioni del prim'ordine

Supponiamo che lo stato di un sistema fisico sia descritto da n *variabili di stato* $y_1(t), y_2(t), \ldots, y_n(t)$, funzioni del tempo t, che raggrupperemo in un *vettore di stato*

$$\mathbf{y}(t) = (y_1(t), y_2(t), \ldots, y_n(t)).$$

L'evoluzione di $\mathbf{y}(t)$ sia governata da un sistema di equazioni del tipo:

$$\text{(2.1)} \qquad \begin{cases} y_1'(t) = f_1(t, y_1, y_2, \ldots, y_n) \\ \vdots \\ y_n'(t) = f_n(t, y_1, y_2, \ldots, y_n). \end{cases}$$

Scriveremo il sistema in forma compatta

$$\text{(2.2)} \qquad \mathbf{y}'(t) = \mathbf{f}(t, \mathbf{y}(t))$$

dove il vettore $\mathbf{f} = (f_1, f_2, \ldots, f_n)$ è una funzione

$$\mathbf{f} : D \subseteq \mathbb{R}^{n+1} \to \mathbb{R}^n.$$

Una soluzione di (2.1) nell'intervallo I è un vettore

$$\varphi(t) = (\varphi_1(t), \varphi_2(t), \ldots, \varphi_n(t)),$$

le cui componenti sono funzioni continue e derivabili in I, le quali simultaneamente soddisfano il sistema (2.1) (cioè lo rendono un'identità per $t \in I$).

 © 978-88-08-06485-1

Equazioni di ordine n

Osserviamo che un'*equazione differenziale di ordine n in forma normale*

$$y^{(n)}(t) = f\left(t, y, y', y'', \dots, y^{(n-1)}\right) \tag{2.3}$$

nella funzione scalare incognita $y(t)$ può sempre essere scritta nella forma di un sistema del tipo (2.1) ponendo

$$y_1 = y;\quad y_2 = y';\dots;\quad y_n = y^{(n-1)}. \tag{2.4}$$

Infatti le precedenti e la (2.3) implicano allora:

$$\begin{cases} y_1' = y_2 \\ y_2' = y_3 \\ \vdots \\ y_n' = f_n(t, y_1, y_2, \dots, y_n). \end{cases} \tag{2.5}$$

Perciò d'ora in poi svilupperemo il discorso teorico per i sistemi di n equazioni differenziali del prim'ordine, tenendo presente che i risultati riguardanti i sistemi si traducono anche in risultati riguardanti le equazioni differenziali di ordine n.

Negli esempi e negli esercizi illustreremo anche qualche caso in cui un sistema di equazioni di ordine superiore al primo può essere riscritto sotto forma di sistema di un numero sufficientemente grande di equazioni del prim'ordine.

Problema di Cauchy

Fissato un istante iniziale t_0, il *problema di Cauchy* per il sistema (2.2) consiste nel cercare una soluzione locale di (2.2) (cioè definita in un non precisato intervallo I contenente t_0) tale che sia soddisfatta la *condizione iniziale*:

$$\mathbf{y}(t_0) = \mathbf{y}_0 \tag{2.6}$$

con $\mathbf{y}_0 \in \mathbb{R}^n$ vettore assegnato. Nel caso dell'equazione di ordine n (2.3), tenuto conto delle (2.4)-(2.5), il problema di Cauchy consiste nel trovare una soluzione locale (funzione $y(t)$ derivabile n volte in I, con $t_0 \in I$, soddisfacente la (2.3)), tale che siano soddisfatte le n condizioni iniziali:

$$\begin{cases} y(t_0) = y_0 \\ y'(t_0) = y_1 \\ \vdots \\ y^{(n-1)}(t_0) = y_{n-1}. \end{cases} \tag{2.7}$$

Si faccia attenzione al fatto che $y_0, y_1, \dots, y_{n-1}$ nella (2.7) indicano costanti assegnate, mentre $y_1, y_2, \dots, y_n$ nella (2.1) indicano funzioni incognite.

Prima di continuare vediamo qualche esempio di modelli matematici, tra i più semplici e noti, che si incontrano in alcune scienze applicate e si formulano mediante equazioni differenziali o sistemi del tipo sopra illustrato.

© 978-88-08-06485-1



Modello preda-predatore di Lotka-Volterra. Negli anni immediatamente successivi al primo conflitto mondiale, nel mare Adriatico si registrò un calo di una certa specie di pesce commestibile e un aumento del suo predatore. Si trattava di capire come mai l'assenza di pesca durante la guerra avesse favorito lo sviluppo del predatore e sfavorito quello della preda. La questione fu posta al matematico italiano Vito Volterra che, nel 1924 circa, risolse il problema con un modello divenuto celebre.

Posto $x = x(t)$ = numero di prede e $y = y(t)$ = numero di predatori, Volterra formulò le seguenti ipotesi, basate sulle informazioni dei biologi: il tasso (relativo) di crescita delle prede è costante in assenza di predatori ma decresce linearmente come funzione di y; i predatori decrescono a un tasso costante in assenza di prede, ma crescono linearmente con x.

Il modello che ne deriva è il seguente:

$$\begin{cases} \dfrac{x'}{x} = a - by \\ \dfrac{y'}{y} = -c + dx \end{cases} \tag{2.8}$$

con a, b, c, d costanti positive, che assume la forma (2.2) ponendo:

$$\mathbf{y} = (x, y)$$
$$f_1(t, \mathbf{y}) = ax - bxy$$
$$f_2(t, \mathbf{y}) = -cy + dxy.$$

Si osservi che $\mathbf{f} = (f_1, f_2)$ non dipende esplicitamente da t: il sistema è autonomo. Il sistema (2.8) era stato studiato già nel 1920 dal polacco Lotka in relazione a un problema di cinetica chimica, ma l'analisi di Volterra si spinse considerevolmente oltre: vi ritorneremo in dettaglio più avanti, nel paragrafo 4, parlando appunto dei sistemi autonomi.



Dinamica unidimensionale. L'*equazione di Newton*

$$y'' = F(t, y, y') \tag{2.9}$$

regola il moto unidimensionale di un corpo materiale puntiforme. Si tratta di un'equazione differenziale del second'ordine; qui y, y', y'' hanno rispettivamente il significato di posizione, velocità e accelerazione. Il moto è completamente determinato se aggiungiamo alla (2.9) le condizioni iniziali:

$$y(0) = y_0;\quad y'(0) = y_1$$

(cioè la posizione e velocità iniziale del punto). Abbiamo già incontrato nel capitolo 1 alcuni importanti casi particolari di quest'equazione:

(i) l'equazione della caduta dei gravi nel vuoto

$$y'' = g$$

o in presenza dell'attrito dell'aria[13]

$$y'' = g - \frac{h}{m}y';$$

(ii) l'equazione dell'oscillatore armonico per le oscillazioni libere

$$y'' = -\omega^2 y,$$

[13] v. capitolo 1, esempio 2.7.

 © 978-88-08-06485-1

per le oscillazioni smorzate

$$y'' = -hy' - \omega^2 y$$

o forzate[14]

$$y'' = -hy' - \omega^2 y + f(t).$$

Vediamo ora un altro esempio di equazione di Newton:

Equazione di van der Pol. Si tratta della celebre equazione proposta nel 1920 da van der Pol, matematico olandese esperto in aspetti teorici della radioingegneria:

$$y'' = -y + \varepsilon\left(1 - y^2\right) y' \quad \text{con } \varepsilon > 0 \text{ costante.}$$

Essa si può interpretare come una perturbazione dell'oscillatore armonico (con $\omega^2 = 1$) mediante il termine forzante $\varepsilon\left(1-y^2\right)y'$ che per valori grandi di y ($|y| > 1$) è una resistenza attiva (dissipa energia) e contrasta il moto, mentre per piccoli valori di y ($|y| < 1$) rappresenta una resistenza negativa (fornisce energia), che nei circuiti elettrici si realizza come *feedback*, o controllo retroattivo. Si presenta in questo caso l'interessante fenomeno delle *oscillazioni autosostenute*. Un'analisi dell'equazione di van der Pol sarà presentata più avanti nel paragrafo 4.3, esempio 4.9.

Moto dei pianeti. Supponiamo che nell'origine del sistema di riferimento in $\mathbb{R}^3$ sia posto un corpo di massa M (sole) il cui campo gravitazionale di forza per unità di massa è dato dal vettore

$$MG\nabla\left(\frac{1}{r}\right), \text{ dove } r = \sqrt{x^2+y^2+z^2}$$

e G è la costante di gravitazione. Il moto di un corpo di massa $m \ll M$ sotto l'influenza di questo campo di forze è dato dal sistema seguente, di 3 equazioni differenziali non lineari del second'ordine:

$$\begin{cases} x'' = -MG\dfrac{x}{r^3} \\ y'' = -MG\dfrac{y}{r^3} \\ z'' = -MG\dfrac{z}{r^3} \end{cases} \tag{2.10}$$

che può essere facilmente riscritto nella forma di sistema di 6 equazioni differenziali del prim'ordine (ponendo $x = y_1, x' = y_2, y = y_3 \ldots$), e quindi rientra anch'esso nella teoria che stiamo delineando. Nel capitolo 2, paragrafo 7.2, abbiamo già incontrato il sistema (2.10), e abbiamo mostrato come da esso si possono ricavare le leggi di Keplero sul moto dei pianeti.

Risultati fondamentali sul problema di Cauchy per i sistemi

I seguenti teoremi generalizzano al caso dei sistemi (e quindi delle equazioni di ordine superiore al primo) i teoremi 8.1, 8.2, 8.3, 8.4 validi nel caso scalare (paragrafo 1.1).

TEOREMA 8.6 (DI PEANO, DI ESISTENZA DI UNA SOLUZIONE) *Sia $D \subset \mathbb{R}^{n+1}$ un aperto e $\mathrm{f} : D \to \mathbb{R}^n$ continua; sia $(t_0, \mathrm{y}_0) \in D$. Allora il problema di Cauchy*

$$\begin{cases} \mathrm{y}'(t) = \mathrm{f}(t, \mathrm{y}(t)) \\ \mathrm{y}(t_0) = \mathrm{y}_0 \end{cases} \tag{2.11}$$

ammette almeno una soluzione, definita in un intorno di t_0.

[14] v. capitolo 1, esempio 3.2. Si veda anche la discussione dettagliata di quest'equazione e del suo significato fisico, svolta nel capitolo 1, paragrafo 3.6.

© 978-88-08-06485-1

TEOREMA 8.7 (DI CAUCHY, DI ESISTENZA E UNICITÀ LOCALE) *Sia $D \subset \mathbb{R}^{n+1}$ un aperto,*

$$\mathrm{f} = (f_1, \ldots, f_n) : D \to \mathbb{R}^n$$

continua, $\frac{\partial f_i}{\partial y_k}$ $(i, k = 1, 2, \ldots, n)$ continue in D e sia $(t_0, \mathrm{y}_0) \in D$. Allora esiste un intorno I di t_0 tale che il problema di Cauchy (2.11) ammette una e una sola soluzione definita in I.

TEOREMA 8.8 (DI ESISTENZA E UNICITÀ GLOBALE) *Sia*

$$\overline{S} = [a, b] \times \mathbb{R}^n,$$
$$\mathrm{f} = (f_1, \ldots, f_n) : \overline{S} \to \mathbb{R}^n$$

continua in $\overline{S}$, $\frac{\partial f_i}{\partial y_k}$ $(i, k = 1, 2, \ldots, n)$ continue in $\overline{S}$; inoltre esistano due numeri positivi h, k (in generale dipendenti da a, b) per cui risulti

$$|\mathrm{f}(t, \mathrm{y})| \le h + k|\mathrm{y}| \quad \text{per ogni } (t, \mathrm{y}) \in \overline{S}. \tag{2.12}$$

Allora ogni soluzione dell'equazione $\mathrm{y}(t) = \mathrm{f}(t, \mathrm{y}(t))$ è definita su tutto l'intervallo $[a, b]$.

Come già osservato nel caso scalare, l'ipotesi (nei teoremi 8.7 e 8.8) che le derivate parziali rispetto alle variabili y_i siano definite e continue si può indebolire chiedendo semplicemente che $\mathrm{f}(t, \mathrm{y})$ sia *localmente lipschitziana in y, uniformemente rispetto a t*. Ciò significa che, per ogni insieme chiuso e limitato $K \subset D$, esiste una costante L_K tale che:

$$|\mathrm{f}(t, \mathrm{y}) - \mathrm{f}(t, \mathrm{z})| \le L_K |\mathrm{y} - \mathrm{z}| \quad \text{per ogni } \mathrm{y}, \mathrm{z} \in K, \quad t \in I.$$

Ancora, sotto l'ipotesi di continuità di f e delle sue derivate parziali rispetto alle y_i, la condizione (2.12), sufficiente per l'esistenza e unicità globale, è soddisfatta in particolare nei due casi semplici: quando $\mathrm{f}(t, \mathrm{y})$ è limitata in y:

$$|\mathrm{f}(t, \mathrm{y})| \le h \quad \text{per ogni } (t, \mathrm{y}) \in \overline{S} \tag{2.13}$$

e quando le sue derivate parziali rispetto alle y_i sono limitate:

$$\left|\frac{\partial f_i}{\partial y_k}(t, \mathrm{y})\right| \le k \quad \text{per ogni } (t, \mathrm{y}) \in \overline{S}, \quad i, k = 1, 2, \ldots, n. \tag{2.14}$$

TEOREMA 8.9 (DIPENDENZA CONTINUA DAI DATI) *Siano $\mathrm{f}_1(t, \mathrm{y})$, $\mathrm{f}_2(t, \mathrm{y})$ due funzioni continue in un aperto $D \subset \mathbb{R}^{n+1}$, localmente lipschitziane rispetto a y, uniformemente in t. Siano (t_0, y_1), $(t_0, \mathrm{y}_2) \in D$ e siano $\varphi_1(t)$, $\varphi_2(t)$, rispettivamente, le soluzioni dei problemi di Cauchy:*

$$\begin{cases} \mathrm{y}' = \mathrm{f}_1(t, \mathrm{y}) \\ \mathrm{y}(t_0) = \mathrm{y}_1 \end{cases} \qquad \begin{cases} \mathrm{y}' = \mathrm{f}_2(t, \mathrm{y}) \\ \mathrm{y}(t_0) = \mathrm{y}_2. \end{cases}$$

 © 978-88-08-06485-1

Allora, se nell'intervallo $[t_0, t_0+T]$ *entrambe le soluzioni sono definite, vale la disuguaglianza:*

$$|\varphi_1(t)-\varphi_2(t)| \le (|y_1-y_2|+(t-t_0)M)\,e^{L(t-t_0)} \quad \text{per ogni } t\in(t_0,t_0+T)$$

dove

$$M=\max_{(t,y)\in K}|f_1(t,y)-f_2(t,y)|,$$

L *è la costante di lipschitzianità di* f_1 *e* f_2 *su* K, *e* K *è un insieme chiuso e limitato contenuto in* D *e contenente i grafici delle soluzioni* $\varphi_1(t)$, $\varphi_2(t)$.

Un caso notevole in cui possiamo applicare i risultati precedenti è il seguente:

Esempio

Equazioni e sistemi lineari. Consideriamo un sistema lineare del tipo:

$$\begin{cases} y_1' = a_{11}(t)\,y_1 + a_{12}(t)\,y_2 + \ldots + a_{1n}(t)\,y_n + b_1(t) \\ y_2' = a_{21}(t)\,y_1 + a_{22}(t)\,y_2 + \ldots + a_{2n}(t)\,y_n + b_2(t) \\ \vdots \\ y_n' = a_{n1}(t)\,y_1 + a_{n2}(t)\,y_2 + \ldots + a_{nn}(t)\,y_n + b_n(t) \end{cases} \tag{2.15}$$

o, in forma matriciale,

$$\mathbf{y}' = \mathrm{A}(t)\,\mathbf{y} + \mathbf{b},$$

dove

$$\mathrm{A}(t) = (a_{ij}(t))_{i,j=1}^n; \quad \mathbf{y}(t) = \begin{pmatrix} y_1(t) \\ y_2(t) \\ \ldots \\ y_n(t) \end{pmatrix}; \quad \mathbf{b}(t) = \begin{pmatrix} b_1(t) \\ b_2(t) \\ \ldots \\ b_n(t) \end{pmatrix}.$$

Se a_{ij} e b_k sono funzioni di classe $C^0([a,b])$, la (2.14) è soddisfatta; infatti, in questo caso f è continua e

$$\left|\frac{\partial f_i}{\partial y_k}(t,y)\right| = |a_{ik}(t)| \le h \text{ per ogni } (t,y)\in\overline{S},$$

avendo posto

$$h=\max_{i,k=1,\ldots,n}\ \max_{t\in[a,b]}|a_{ik}(t)|.$$

Dunque, la soluzione del problema di Cauchy per il sistema lineare (2.15) esiste ed è unica in tutto $[a,b]$. Se poi le funzioni a_{ij} e b_k sono continue in tutto $\mathbb{R}$, la soluzione sarà definita in tutto $\mathbb{R}$ (perché è definita in ogni intervallo $[a,b]\subset\mathbb{R}$).

Analogamente, per un'equazione lineare di ordine n

$$y^{(n)}(t) = a_0(t)\,y + a_1(t)\,y' + \ldots + a_{n-1}(t)\,y^{(n-1)} + b(t)$$

si ricava facilmente (riscrivendo l'equazione in forma di sistema di n equazioni del prim'ordine) che ogni soluzione è definita in tutto l'intervallo di continuità dei coefficienti.

OSSERVAZIONE (CONDIZIONI DIVERSE DA QUELLE INIZIALI) Concludiamo questo paragrafo segnalando che, nello studio dei modelli matematici nelle scienze applicate, si incontrano talvolta altri tipi di condizioni aggiuntive, diverse dalle condizioni inziali (2.6); un esempio è dato dalle cosiddette *condizioni ai limiti*, condizioni cioè che vengono prescritte agli estremi di un intervallo, come per esempio nel caso seguente:

$$\begin{cases} y'' = f(t,y,y') \\ y(a) = A \\ y(b) = B. \end{cases}$$

In questo caso le condizioni chiedono alla soluzione di "passare" per i due punti (a,A), (b,B) del piano (t,y) e perciò la soluzione *deve* essere definita in tutto l'intervallo $[a,b]$. Ciò conferisce a un problema ai limiti un aspetto di *globalità* che quello ai valori iniziali, a priori, non possiede. Per i suddetti motivi, la teoria utilizza metodologie di indagine molto diverse nei due tipi di problemi; noi non ce ne occuperemo.

Esercizi

Risolvere il seguente sistema:

$$\begin{cases} x' = tx \\ y' = ty + x. \end{cases}$$

Suggerimento: la prima equazione non dipende da y: risolverla e sostituire la x così trovata nella seconda equazione.

Si consideri il seguente sistema di due equazioni del second'ordine in due funzioni incognite x_1, x_2:

$$\begin{cases} m_1 x_1'' + k_1 x_1 + k_2(x_1 - x_2) = F_1(t) \\ m_2 x_2'' + k_3 x_2 + k_2(x_2 - x_1) = F_2(t). \end{cases}$$

Significato fisico: due masse m_1, m_2 si muovono senza attrito su una retta orizzontale, avendo posizione x_1, x_2; tre molle di costanti elastiche k_1, k_2, k_3 collegano rispettivamente la prima massa a un punto fisso, le due masse tra loro, la seconda massa a un secondo punto fisso; la massa m_i è soggetta a una forza esterna F_i dipendente dal tempo.

(i) Riscrivere il sistema nella forma di un sistema di 4 equazioni del prim'ordine.

(ii) Nell'ipotesi $m_1 = m_2, k_1 = k_2 = k_3, F_1(t) \equiv F_2(t) \equiv 0$ mostrare come, operando opportunamente sul sistema di partenza, se ne può ricavare un'equazione di 4° grado nella sola incognita x_1, lineare a coefficienti costanti. Risolverla con i metodi esposti nel capitolo 1, paragrafo 4.2, quindi impostare la risoluzione dell'equazione in x_2.

3 COMPLEMENTI

3.1 Lemma di Gronwall e dipendenza continua

Dimostriamo ora il teorema 8.9 di dipendenza continua delle soluzioni dagli ingredienti del problema.

Premettiamo il seguente lemma, che esprime un'importante disuguaglianza dell'Analisi:

LEMMA 8.10 (DI GRONWALL) *Sia* $h: I\to\mathbb{R}$ *una funzione continua e non negativa, sia* $t_0\in I$ *e supponiamo che per certe due costanti* $a, b>0$ *sia:*

$$h(t) \le a + b\int_{t_0}^t h(s)\,ds \quad \text{per ogni } t\ge t_0, t\in I.$$

 © 978-88-08-06485-1

Allora

$$h(t) \leq ae^{b(t-t_0)} \quad \text{per ogni } t \geq t_0, t \in I.$$

DIMOSTRAZIONE. Sia

$$w(t) = a + b\int_{t_0}^{t} h(s)\,ds,$$

allora $w \in C^1(I)$, $w'(t) = bh(t)$, perciò

$$w'(t) \leq bw(t) \quad \text{per ogni } t \geq t_0, t \in I.$$

Integrando in (t_0, t), troviamo:

$$\int_{t_0}^{t} \frac{w'(s)}{w(s)} ds \leq b(t - t_0)$$

ossia (poiché $w(t_0) = a$)

$$w(t) \leq ae^{b(t-t_0)}$$

e, poiché per ipotesi $h(t) \leq w(t)$, la tesi è provata. ∎

Proviamo ora il teorema di dipendenza continua, di cui per comodità riportiamo l'enunciato.

TEOREMA 8.9 (DIPENDENZA CONTINUA DAI DATI) *Siano $f_1(t, y)$, $f_2(t, y)$ due funzioni continue in un aperto $D \subset \mathbb{R}^{n+1}$, localmente lipschitziane rispetto a y, uniformemente in t. Siano (t_0, y_1), $(t_0, y_2) \in D$ e siano $\varphi_1(t)$, $\varphi_2(t)$, rispettivamente, le soluzioni dei problemi di Cauchy:*

$$\begin{cases} y' = f_1(t, y) \\ y(t_0) = y_1 \end{cases} \qquad \begin{cases} y' = f_2(t, y) \\ y(t_0) = y_2. \end{cases}$$

Allora, se nell'intervallo $[t_0, t_0 + T]$ entrambe le soluzioni sono definite, vale la disuguaglianza:

$$|\varphi_1(t) - \varphi_2(t)| \leq (|y_1 - y_2| + (t - t_0)M)\, e^{L(t-t_0)} \quad \text{per ogni } t \in (t_0, t_0 + T)$$

dove

$$M = \max_{(t,y)\in K} |f_1(t, y) - f_2(t, y)|,$$

L è la costante di lipschitzianità di f_1 e f_2 su K, e K è un insieme chiuso e limitato contenuto in D e contenente i grafici delle soluzioni $\varphi_1(t)$, $\varphi_2(t)$.

DIMOSTRAZIONE. Scriviamo anzitutto[15]:

$$\varphi_i(t) = y_i + \int_{t_0}^{t} \varphi_i'(\tau)\,d\tau = y_i + \int_{t_0}^{t} f_i(\tau, \varphi_i(\tau))\,d\tau, \quad \text{per } i = 1, 2.$$

[15] abbiamo usato il teorema fondamentale del calcolo integrale in forma vettoriale: v. capitolo 2, teorema 2.2.

© 978-88-08-06485-1

Sottraendo membro a membro le due identità per $i = 1, 2$, abbiamo:

$$\varphi_1(t) - \varphi_2(t) = y_1 - y_2 + \int_{t_0}^{t} [f_1(\tau, \varphi_1(\tau)) - f_1(\tau, \varphi_2(\tau))]\,d\tau + \\ + \int_{t_0}^{t} [f_1(\tau, \varphi_2(\tau)) - f_2(\tau, \varphi_2(\tau))]\,d\tau,$$

da cui[16]:

$$|\varphi_1(t) - \varphi_2(t)| \leq |y_1 - y_2| + \int_{t_0}^{t} |f_1(\tau, \varphi_1(\tau)) - f_1(\tau, \varphi_2(\tau))|\,d\tau + \\ + \int_{t_0}^{t} |f_1(\tau, \varphi_2(\tau)) - f_2(\tau, \varphi_2(\tau))|\,d\tau$$

per definizione di M e per la lipschitzianità di f_1

$$\leq |y_1 - y_2| + \int_{t_0}^{t} L|\varphi_2(\tau) - \varphi_2(\tau)|\,d\tau + M(t - t_0).$$

Abbiamo dunque ottenuto, per ogni $t \in (t_0, t_0 + T)$,

$$|\varphi_1(t) - \varphi_2(t)| \leq |y_1 - y_2| + M(t - t_0) + L\int_{t_0}^{t} |\varphi_2(\tau) - \varphi_2(\tau)|\,d\tau$$

da cui, per il lemma di Gronwall, segue la tesi. ∎

4 SISTEMI AUTONOMI BIDIMENSIONALI

4.1 Generalità

Un sistema (2.2) si dice *autonomo* se f non dipende esplicitamente da t, cioè se è della forma:

$$y' = f(y). \tag{4.1}$$

In questo paragrafo ci occuperemo dei sistemi di questo tipo *bidimensionali*, anche se alcune definizioni date nel seguito si estendono facilmente al caso n-dimensionale. Scriveremo il sistema nella forma scalare:

$$\begin{cases} x' = f(x, y) \\ y' = g(x, y) \end{cases} \tag{4.2}$$

con $f, g \in C^1(D)$, D aperto di $\mathbb{R}^2$. Queste ipotesi garantiscono esistenza e unicità per la soluzione del problema di Cauchy con i dati iniziali:

$$\begin{cases} x(t_0) = x_0 \\ y(t_0) = y_0 \end{cases} \quad \text{con } (x_0, y_0) \in D \text{ assegnato.}$$

[16] v. Lemma 2.3, capitolo 2.

 © 978-88-08-06485-1

Curve integrali e orbite

Una soluzione di (4.2) è una funzione

$$\mathbf{r}(t) = (x(t), y(t))$$

definita su qualche intervallo $I \subseteq \mathbb{R}$, di classe $C^1(I)$ e soddisfacente il sistema (4.2). Il grafico di questa funzione è un sottoinsieme di $\mathbb{R}^3$ che rappresenta una curva di equazioni parametriche

$$\begin{cases} x = x(t) \\ y = y(t) \\ z = t \end{cases} \quad t \in I \tag{4.3}$$

e viene detta *curva integrale*. Il piano delle variabili di stato x, y è detto *piano delle fasi*; la proiezione di una curva integrale sul piano delle fasi è la curva piana orientata (in base all'orientazione di I) di equazioni

$$\begin{cases} x = x(t) \\ y = y(t) \end{cases} \quad \text{cioè } \mathbf{r}(t), \text{ con } t \in I.$$

Tali curve si chiamano *orbite* o *traiettorie* del sistema (4.2). $\mathbf{r}(t)$ è il vettore posizione (esso ci dice in quale stato si trova il sistema all'istante t) e $\mathbf{r}'(t) = (x'(t), y'(t))$ (il vettore velocità) è tangente alla traiettoria in $\mathbf{r}(t)$. Risolvere il sistema (4.2) si può interpretare così: in D è assegnato un campo di velocità che associa a ogni punto $(x, y) \in D$ un vettore

$$\mathbf{v} = (f(x, y), g(x, y));$$

si tratta di trovare le curve che sono tangenti a $\mathbf{v}$ in ogni punto.

Esempio

La figura 8.4 rappresenta il campo di velocità del sistema:

$$\begin{cases} x' = -y \\ y' = x. \end{cases} \tag{4.4}$$

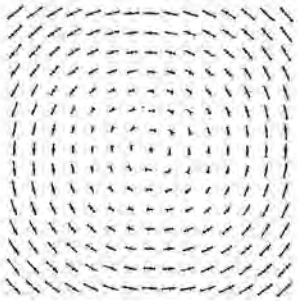


Figura 8.4. Il campo di velocità $\mathbf{v} = (-y, x)$ associato al sistema (4.4).

© 978-88-08-06485-1

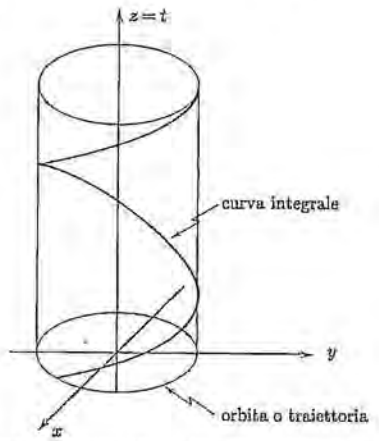


Figura 8.5. Una curva integrale del sistema (4.4) e la corrispondente orbita.

Una soluzione di (4.4) è:

$$\mathbf{r}(t) = (\cos t, \sin t) \text{ per } t \in [0, 2\pi),$$

il cui grafico è un arco di elica cilindrica e la cui orbita è la circonferenza con centro nell'origine e raggio 1 percorsa in senso antiorario.

Come abbiamo visto nel caso unidimensionale, il fatto che il sistema sia autonomo comporta che *le soluzioni sono invarianti per traslazioni temporali*, cioè se $\mathbf{r}(t) = (x(t), y(t))$ è una soluzione di (4.2) definita in $I = (a, b)$, allora

$$\mathbf{s}(t) = (x(t+\tau), y(t+\tau))$$

(per $\tau \in \mathbb{R}$ fissato) è soluzione dello stesso sistema in $(a-\tau, b-\tau)$; infatti:

$$\begin{aligned} \mathbf{s}'(t) &= (x'(t+\tau), y'(t+\tau)) = \\ &= (f(x(t+\tau), y(t+\tau)), g(x(t+\tau), y(t+\tau))) = \mathbf{v}(\mathbf{s}(t)). \end{aligned}$$

Si osservi che le due soluzioni hanno la *stessa* orbita: perciò a ogni orbita corrispondono infinite soluzioni, ognuna ottenuta da un'altra per traslazione temporale. Se, invece di una traslazione, facciamo un'inversione temporale, cioè consideriamo la funzione $\mathbf{s}(t) = \mathbf{r}(-t)$, questa risulta ancora soluzione di (4.2) in $(-b, -a)$, a cui corrisponde geometricamente ancora la stessa orbita di $\mathbf{r}(t)$, ma percorsa in verso opposto.

Una conseguenza importante della proprietà di invarianza appena ricordata è che *le orbite non possono intersecarsi*: ogni punto $\mathbf{p} \in D$ appartiene a una e una sola orbita. Infatti, se il punto $\mathbf{p}$ appartenesse a due orbite distinte, corrispondenti a due soluzioni $\mathbf{r}(t)$, $\mathbf{s}(t)$, esisterebbero due istanti t_1, t_2 per cui

$$\mathbf{r}(t_1) = \mathbf{s}(t_2) = \mathbf{p}.$$

Ma allora, posto $\tau = t_2 - t_1$, la funzione $\mathbf{w}(t) = \mathbf{s}(t+\tau)$ sarebbe una soluzione del sistema tale che:

$$\mathbf{w}(t_1) = \mathbf{s}(t_1+\tau) = \mathbf{s}(t_2) = \mathbf{p}$$

e perciò w ed r sarebbero due diverse soluzioni soddisfacenti la stessa condizione iniziale, cioè due diverse soluzioni dello stesso problema di Cauchy, in contrasto col teorema di esistenza e unicità.

Stati stazionari e cicli

Due tipi di orbite hanno un'importanza particolare: quelle che si riducono a un solo punto (sono detti *stati di equilibrio* o *stati stazionari* o *punti critici* o *singolari*; corrispondono a una soluzione costante) e i *cicli* (cioè curve semplici e chiuse che, come vedremo, corrispondono a *soluzioni periodiche*).

DEFINIZIONE 8.3 Un punto $(\overline{x},\overline{y}) \in D$ è *uno stato stazionario* del sistema (4.2) se

$$f(\overline{x},\overline{y}) = g(\overline{x},\overline{y}) = 0.$$

Un tale punto è perciò l'orbita corrispondente alla soluzione costante $\mathrm{r}(t) = (\overline{x},\overline{y})$, il cui grafico in $\mathbb{R}^3$ è la retta parallela all'asse t passante per tale punto. Se all'istante iniziale il sistema si trova in questo stato, vi rimane per sempre. Nessuna soluzione può raggiungere questo stato in un tempo finito (due diverse orbite non hanno punti in comune), ma una soluzione può tendere a un punto di equilibrio per $t \to \pm\infty$; anzi, se una soluzione $(x(t), y(t))$ ammette limite $(\overline{x},\overline{y})$ per $t \to +\infty$ (oppure per $t \to -\infty$), allora il punto $(\overline{x},\overline{y})$ è uno stato stazionario del sistema, come si deduce passando al limite in ambo i membri della (4.2).

Abbiamo già sottolineato l'importanza del concetto di stabilità. Come abbiamo fatto per il caso scalare, possiamo classificare i punti stazionari riguardo alla stabilità.

DEFINIZIONE 8.4 Indicando con $\varphi(t;\mathrm{q})$ la soluzione del sistema (4.2) che si trova all'istante $t = 0$ nel punto q, diremo che il punto di equilibrio $\overline{\mathrm{p}} = (\overline{x},\overline{y})$ è:

(i) *stabile* se, per ogni $\varepsilon > 0$ esiste $\delta = \delta(\varepsilon) > 0$ tale che, se $|\overline{\mathrm{p}} - \mathrm{q}| < \delta$, la soluzione $\varphi(t;\mathrm{q})$ esiste per ogni $t \geq 0$ e risulta

$$|\varphi(t;\mathrm{q}) - \overline{\mathrm{p}}| < \varepsilon \quad \text{per ogni } t \geq 0;$$

(ii) *asintoticamente stabile* se è stabile e inoltre esiste δ_1 tale che, se $|\overline{\mathrm{p}} - \mathrm{q}| < \delta_1$, risulta

$$\varphi(t;\mathrm{q}) \to \overline{\mathrm{p}} \quad \text{per } t \to +\infty;$$

(iii) *instabile* se non è stabile.

Se $\overline{\mathrm{p}}$ è asintoticamente stabile, si dice anche che è un *attrattore locale*: esiste un intorno U di $\overline{\mathrm{p}}$ tale che ogni orbita che ha origine in U converge a $\overline{\mathrm{p}}$; l'unione di tutti gli intorni U con questa proprietà costituisce il *bacino di attrazione* di $\overline{\mathrm{p}}$.

Di fondamentale importanza sono anche le soluzioni $\varphi(t)$ *periodiche*, cioè: esiste $T > 0$ tale che $\varphi(t+T) = \varphi(t)$ per ogni $t \in \mathbb{R}$, ma $\varphi(t+\tau) \neq \varphi(t)$ per $\tau \in (0,T)$. Allora T si chiama *periodo*; chiaramente la traiettoria descritta da $\varphi(t)$ è una curva semplice e chiusa (*ciclo*).

La cosa notevole è che, per i sistemi autonomi, vale anche il viceversa, cioè:

PROPOSIZIONE 8.1 *Se in due istanti diversi il sistema ripassa dallo stesso punto, allora l'orbita è periodica.*

DIMOSTRAZIONE. Questo dipende ancora dall'invarianza per traslazioni temporali. Sia $\varphi(t;\mathrm{y}_0)$ la soluzione del problema di Cauchy con condizione $\varphi(0;\mathrm{y}_0) = \mathrm{y}_0$; supponiamo che $\varphi(T;\mathrm{y}_0) = \varphi(0;\mathrm{y}_0)$ per un certo $T > 0$ (anzi, possiamo supporre che T sia il minimo dei numeri per cui questo accade), e proviamo che allora $\varphi(t+T;\mathrm{y}_0) = \varphi(t;\mathrm{y}_0)$ per ogni $t > 0$. Infatti, $\psi(t) = \varphi(t+T;\mathrm{y}_0)$ risolve ancora il sistema (4.1) e soddisfa la condizione:

$$\psi(0) = \varphi(T;\mathrm{y}_0) = \varphi(0;\mathrm{y}_0) = \mathrm{y}_0.$$

Per l'unicità della soluzione del problema di Cauchy, $\psi(t) = \varphi(t;\mathrm{y}_0)$, che è la tesi.

Perciò, *per i sistemi autonomi i concetti di orbita chiusa e orbita periodica coincidono.*

Per esempio, la circonferenza $x^2 + y^2 = 1$ è un ciclo per il sistema (4.4); è facile mostrare che ogni soluzione di questo sistema è 2π-periodica (Esempio 4.1). Ogni ciclo divide il piano delle fasi in due regioni, una interna e una esterna: ogni orbita che nasce (al tempo t_0) all'interno di un ciclo vi resta per sempre intrappolata; ogni orbita che nasce all'esterno resta per sempre esterna. Un'orbita può avvicinarsi indefinitamente a un ciclo, avvolgendosi a spirale attorno a esso senza mai toccarlo; in tal caso il ciclo prende il nome di *ciclo limite*.

Anche i cicli limite si possono classificare in base alla stabilità, come abbiamo fatto per gli stati stazionari.

In conclusione, si può mostrare che le traiettorie possono essere solo di tre tipi: *punti di equilibrio, cicli, curve semplici non chiuse.*

Integrali primi, o costanti di moto

Un concetto importante è quello di *integrale primo*, o *costante del moto*. Sia $E : D \to \mathbb{R}$ una funzione di classe $C^1(D)$. Essa si chiamerà integrale primo se è costante lungo le traiettorie del sistema. Più precisamente, sia $\varphi(t)$ una soluzione del sistema (4.2) (definita nel suo intervallo massimale); allora è ben definita la funzione composta $E(\varphi(t))$. Essendo $\varphi'(t) = \mathrm{f}(\varphi(t))$, si ha:

$$\frac{d}{dt}\left[E(\varphi(t))\right] = \nabla E(\varphi(t)) \cdot \varphi'(t) = \nabla E(\varphi(t)) \cdot \mathrm{f}(\varphi(t)) = (\nabla E \cdot \mathrm{f})(\varphi(t)).$$

Allora chiameremo E integrale primo o costante del moto per il sistema (4.2) se $\frac{d}{dt}[E(\varphi(t))] = 0$, cioè se e solo se ∇E risulta ortogonale al campo di velocità (f,g). L'importanza degli integrali primi è dovuta al fatto che, se un integrale primo E è noto, le sue curve di livello

$$E(x,y) = k$$

hanno le seguenti proprietà: ogni curva di livello di E è unione di orbite di (4.2) e, viceversa, ogni orbita è contenuta in qualche curva di livello. Per convincersene, basta osservare che:

(i) $\mathrm{v}(x,y)$ (cioè il vettore velocità (f,g)) è tangente all'orbita passante per (x,y);

(ii) ∇E è ortogonale alla curva di livello passante per (x,y);

(iii) ∇E è ortogonale a v.

 © 978-88-08-06485-1

Perciò v dev'essere tangente alla curva di livello nel punto (x,y) e questa dev'essere un'orbita in vicinanza di (x,y).

Esempi

Illustriamo ora i concetti introdotti in questo paragrafo con alcuni esempi. Qui ci limiteremo a studiare il comportamento delle orbite vicino ai punti di equilibrio, che di solito è ciò che più interessa nelle applicazioni.

Una descrizione qualitativamente accurata dell'andamento delle orbite di un sistema si chiama *ritratto di fase*; un completo ritratto di fase sarà descritto nel paragrafo 4.2 per i sistemi lineari, mentre l'analisi di qualche caso non lineare sarà svolta nel paragrafo 4.3.

Osserviamo preliminarmente che, nell'intorno di un punto regolare (non di equilibrio), al sistema (4.2) è associata l'equazione differenziale (ottenuta eliminando t dalle due equazioni del sistema):

$$\frac{dy}{dx}=\frac{g(x,y)}{f(x,y)} \quad \text{oppure} \quad \frac{dx}{dy}=\frac{f(x,y)}{g(x,y)}. \tag{4.5}$$

Ogni soluzione dell'equazione differenziale coincide (almeno localmente) con una traiettoria del sistema e viceversa. Si noti, che nei punti regolari in cui $g(x,y)=0$ oppure $f(x,y)=0$, la direzione del campo vettoriale (f,g), tangente all'orbita, è orizzontale o verticale, rispettivamente. Studiando i segni di f e g si può poi ripartire il piano delle fasi nelle zone di crescenza e decrescenza per x e y, determinando così anche il verso di percorrenza delle orbite.

Esempio Il sistema lineare (4.4), cioè

$$\begin{cases} x'=-y \\ y'=x, \end{cases}$$

il cui campo di velocità è rappresentato dalla figura 8.4, ha un unico punto di equilibrio: l'origine. L'equazione differenziale delle traiettorie è:

$$y'=-\frac{x}{y} \quad \text{cioè} \quad yy'=-x$$

(oppure $x'=-\frac{y}{x}$ cioè $xx'=-y$), il cui integrale generale è

$$x^2+y^2=c, \quad \text{con } c\in\mathbb{R}.$$

Tutte le orbite sono cicli, tutte le soluzioni sono periodiche. Nella figura 8.5 è rappresentata un'orbita e la corrispondente curva integrale, che è un'elica cilindrica. Per determinare il verso di percorrenza dell'orbita basta osservare, per esempio, che, essendo $x'=-y$, $x(t)$ dev'essere decrescente nel semipiano $y>0$; perciò l'orbita è percorsa in senso antiorario. Il punto di equilibrio è stabile, ma non asintoticamente stabile. Osserviamo infine che $E(x,y)=x^2+y^2$ è un integrale primo; infatti $\nabla E=(2x,2y)$ è ortogonale al vettore velocità $(-y,x)$.

Esempio Il *sistema di Lotka-Volterra*

$$\begin{cases} \frac{x'}{x}=a-by \\ \frac{y'}{y}=-c+dx \end{cases} \tag{4.6}$$

© 978-88-08-06485-1

(che abbiamo introdotto nell'esempio 2.1, a cui rimandiamo per il significato modellistico) ha l'unico punto di equilibrio $(c/d,a/b)$. Un integrale primo è dato da:

$$E(x,y)=-c\log x+dx-a\log y+by.$$

Infatti, moltiplicando la prima delle (4.6) per $(-c+dx)$ e la seconda per $(a-by)$, sottraendo si trova:

$$0=\left(-\frac{c}{x}+d\right)x'-\left(\frac{a}{y}-b\right)y'=\frac{d}{dt}[-c\log x+dx-a\log y+by]$$

e quindi

$$\frac{d}{dt}[E(x(t),y(t))]=0.$$

Dallo studio di questo integrale primo possiamo estrarre informazioni qualitative sull'andamento delle orbite (almeno in vicinanza del punto di equilibrio), come ora vedremo. Essendo

$$\frac{\partial^2 E}{\partial x^2}=\frac{c}{x^2}, \quad \frac{\partial^2 E}{\partial y^2}=\frac{a}{y^2}, \quad \frac{\partial^2 E}{\partial x\partial y}=0,$$

si ha che E è strettamente convessa[17] in $(0,+\infty)\times(0,+\infty)$. Inoltre:

$$\frac{\partial E}{\partial x}=-\frac{c}{x}+d=0 \text{ per } x=\frac{c}{d} \text{ e}$$
$$\frac{\partial E}{\partial y}=-\frac{a}{y}+b=0 \text{ per } y=\frac{a}{b},$$

quindi l'unico punto di equilibrio $\left(\frac{c}{d},\frac{a}{b}\right)$ è punto di minimo globale[18] per E. Osserviamo inoltre che $E(x,y)\to+\infty$ quando x oppure y tende a 0 oppure a $+\infty$. Si deduce allora che, per $k>E\left(\frac{c}{d},\frac{a}{b}\right)$, gli insiemi di livello

$$\{(x,y)\in(0,+\infty)\times(0,+\infty): E(x,y)=k\}$$

sono curve chiuse regolari che contengono il punto di equilibrio $\left(\frac{c}{d},\frac{a}{b}\right)$ al loro interno. L'evoluzione della coppia $x(t),y(t)$ è di conseguenza periodica con un periodo T che dipende da k. Il ritratto di fase è indicato in figura 8.6a. Lo studio dei segni di $a-by$ e $-c+dx$ indica che il verso di percorrenza sulle orbite è antiorario. In figura 8.6b è illustrato l'andamento tipico di $x=x(t)$ e $y=y(t)$.

Come nell'esempio precedente, il punto di equilibrio è stabile, ma non asintoticamente stabile.

Possiamo ora rispondere alla questione relativa alla pesca nel mare Adriatico posta nell'esempio 2.1.

Anzitutto osserviamo che, essendo $x(0)=x(T)$ e $y(0)=y(T)$, si ha:

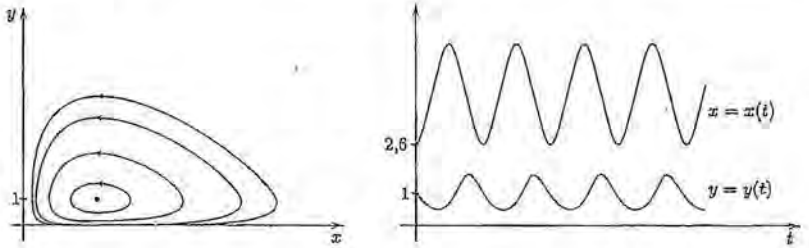
$$\int_0^T\frac{x'}{x}dt=[\log x(t)]_0^T=\log x(T)-\log x(0)=0$$

e analogamente:

$$\int_0^T\frac{y'}{y}dt=0.$$

[17] v. capitolo 3, paragrafo 7.1.
[18] v. capitolo 3, paragrafo 7.2.

 © 978-88-08-06485-1



a) b)

Figura 8.6. a) orbite del sistema Lotka-Volterra corrispondenti ai dati iniziali $x = 0{,}5,\ 0{,}7,\ 1{,}4,\ 2{,}6$ e $y = 1$. b) grafici di $x = x(t)$, $y = y(t)$, soluzioni del sistema di Lotka-Volterra con dati iniziali $x(0) = 2{,}6$ e $y(0) = 1$.

Integrando allora ambo i membri delle (4.6), si ottiene:

$$0 = \int_0^T (a - by(t))\,dt = aT - b\int_0^T y(t)\,dt;$$
$$0 = \int_0^T (-c + dx(t))\,dt = -cT + d\int_0^T x(t)\,dt,$$

ovvero

$$\frac{1}{T}\int_0^T y(t)\,dt = \frac{a}{b};\quad \frac{1}{T}\int_0^T x(t)\,dt = \frac{c}{d}. \tag{4.7}$$

Le formule (4.7) indicano che le coordinate a/b e c/d del punto di equilibrio hanno anche il significato di *numero medio di individui nell'arco di un periodo*. È interessante osservare che questi valori medi sono comuni a *tutte* le orbite.

Ora, una diminuzione della pesca:
- aumenta il tasso di crescita di x; invece di a si ha $a + \varepsilon$;
- riduce il tasso di diminuzione di y; da c si passa a $c - \eta$;
- lascia inalterati i coefficienti b, d di interazione tra le specie.

I numeri medi (4.7) diventano allora $\frac{a+\varepsilon}{b}$ per i predatori, $\frac{c-\eta}{d}$ per le prede; queste ultime registrano perciò una diminuzione media. La conclusione, nota come "principio di Volterra", è valida in molti altri casi, per modelli anche più realistici del sistema (4.6).

Integrale primo dell'energia. Consideriamo l'equazione (autonoma) di Newton

$$y'' = f(y), \tag{4.8}$$

che modellizza un *sistema conservativo a un grado di libertà*. Supponiamo che $f \in C^1(\mathbb{R})$, cosicché per ogni $t_0, \xi_0, \xi_1 \in \mathbb{R}$ il problema di Cauchy con le condizioni iniziali

$$\begin{cases} y(t_0) = \xi_0 \\ y'(t_0) = \xi_1 \end{cases}$$

ha un'unica soluzione locale. Il sistema equivalente a (4.8) ha la forma

$$\begin{cases} y_1' = y_2 \\ y_2' = f(y_1), \end{cases} \tag{4.9}$$

© 978-88-08-06485-1

Nello spazio delle variabili y_1, y_2 (spazio delle fasi), consideriamo le seguenti funzioni:

$$T(y_2) = \frac{y_2^2}{2} = \text{energia cinetica};$$
$$U(y_1) = -\int_0^{y_1} f(s)\,ds = \text{energia potenziale};$$
$$E = T + U = \text{energia meccanica}.$$

La ragione del termine *conservativo* sta nel fatto che, durante il moto, l'energia meccanica E rimane costante (si conserva). Più precisamente, se

$$\varphi(t) = (\varphi_1(t), \varphi_2(t))$$

è una traiettoria del sistema, si ha:

$$\frac{d}{dt}E(\varphi_1(t), \varphi_2(t)) = \frac{d}{dt}T(\varphi_2(t)) + \frac{d}{dt}U(\varphi_1(t)) =$$
$$= \varphi_2(t)\varphi_2'(t) - f(\varphi_1(t))\varphi_1'(t) =$$

(utilizzando le (4.9))

$$= 0.$$

Perciò E è una costante del moto, o un integrale primo del sistema. Si può mostrare che, se l'energia potenziale U si mantiene limitata inferiormente, vale a dire se esiste c_0 tale che $U(y) \geq c_0$ per ogni $y \in \mathbb{R}$, allora ogni soluzione di (4.1) esiste in tutto $\mathbb{R}$.

Grandi oscillazioni di un pendolo senza attrito. Un corpo puntiforme è sospeso all'estremità di un'asta rigida, di massa trascurabile e lunghezza l. L'asta è fissata all'altra estremità a un perno ed è libera di ruotare senza attrito in un piano verticale, soggetta alla sola forza gravitazionale (fig. 8.7). L'equazione di Newton del moto è

$$ml\varphi'' = -mg\sin\varphi$$

ossia, ponendo $k = g/l$,

$$\varphi'' = -k\sin\varphi. \tag{4.10}$$

Alla (4.10) è associato il sistema (ponendo $z_1 = \varphi, z_2 = \varphi'$)

$$\begin{cases} z_1' = z_2 \\ z_2' = -k\sin z_1. \end{cases}$$

I punti di equilibrio sono infiniti e hanno coordinate $(\nu\pi, 0), \nu \in \mathbb{Z}$. La periodicità di $\sin z_1$ indica che è sufficiente limitarsi a considerare $z_1 \in [-\pi, \pi]$.

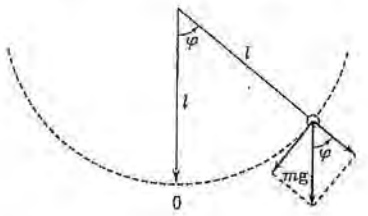


Figura 8.7.

 © 978-88-08-06485-1

Essendo il sistema conservativo, la funzione energia

$$E(z_1, z_2) = \frac{1}{2}z_2^2 + k(1 - \cos z_1)$$

è un integrale primo (Esempio 4.4).

Incidentalmente osserviamo che l'energia potenziale $U(z_1) = k(1 - \cos z_1)$ è non negativa e quindi le soluzioni sono definite su tutto $\mathbb{R}$.

Le orbite sono descritte dalla famiglia degli insiemi di livello dell'energia

$$\frac{1}{2}z_2^2 + k(1 - \cos z_1) = c \qquad c \geq 0,$$

che si può risolvere rispetto a z_2, ottenendo:

$$z_2 = \pm\sqrt{2(c - k(1 - \cos z_1))}. \tag{4.11}$$

Si distinguono allora i casi seguenti[19].

1. $c = 0$; è questo il livello minimo di energia e le (4.11) si riducono al punto di equilibrio $(0, 0)$.

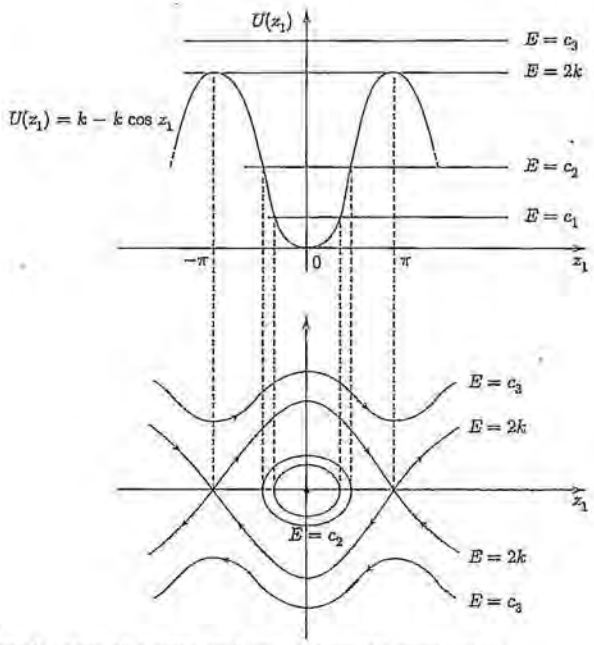


Figura 8.8. Energia potenziale e orbite per il moto del pendolo senza attrito.

[19] Il lettore è invitato a rendersi conto in dettaglio delle prossime affermazioni, studiando le funzioni (4.11) al variare dei parametri c, k.

© 978-88-08-06485-1

2. $0 < c < 2k$; ogni coppia di equazioni (4.11) individua una traiettoria chiusa attorno all'origine. Questi livelli di energia corrispondono al moto oscillatorio intorno alla posizione di equilibrio $(0, 0)$.
3. $c = 2k$; ogni coppia di equazioni (4.11) individua una curva chiusa nel piano contenente quattro orbite: i due punti di equilibrio $(\pm\pi, 0)$ e altre due orbite che li connettono, che si chiamano *eterocline* e corrispondono a un avvicinamento (in tempo infinito) del pendolo alla posizione di equilibrio verticale sopra il perno.
4. $c > 2k$; ciascuna delle (4.11) individua un'orbita non chiusa. Questi livelli di energia corrispondono al moto di rivoluzione completa del pendolo.

La configurazione generale delle orbite è indicata in figura 8.8. Il verso di percorrenza è determinato (per esempio) dal fatto che $z_1' = z_2 > 0$ per $z_2 > 0$. Quanto alla stabilità degli equilibri, dalla discussione precedente si comprende che l'origine è stabile (non asintoticamente), mentre i punti $(\pm\pi, 0)$ sono instabili (si veda anche nel paragrafo 4.3, esempio 4.8)

Esercizi

(18) Fare un esempio di sistema non autonomo per cui esistono orbite chiuse non periodiche. (Suggerimento: considerare $\varphi(t) = (\cos(t^2), \sin(t^2))$ e trovare un sistema di cui questa è soluzione). Di conseguenza, l'equivalenza tra orbite chiuse e orbite periodiche, che abbiamo dimostrato per i sistemi autonomi, non vale in generale per sistemi qualsiasi.

(19) Scrivere l'integrale primo dell'energia (v. esempio 4.4) per l'equazione dell'oscillatore armonico

$$y'' = -\omega^2 y.$$

(20) Il seguente sistema è una variante del sistema di Lotka-Volterra:

$$\begin{cases} x' = x(a - by - ex) \\ y' = y(-c + dx - fy) \end{cases}$$

con $a, b, c, d > 0$, $e, f \in \mathbb{R}$. Trovare gli stati stazionari.

(21) Mostrare che la funzione

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

è una costante del moto per il sistema

$$\begin{cases} x' = -x - 2y \\ y' = 2x + y. \end{cases}$$

(22) Si consideri il sistema autonomo bidimensionale (4.2). Mostrare che, se l'equazione differenziale delle traiettorie è esatta[20], allora esiste un integrale primo.

(23) *Un problema di farmacocinetica.* Sia $c(t)$ la concentrazione di una sostanza nel sangue di un paziente. Una dose costante c_0 viene somministrata a ogni intervallo di tempo T, in particolare al tempo $t = 0$. Si studi l'andamento di $c(t)$ nei seguenti due casi:

(i) $c'(t) = kc(t)$;

[20] v. paragrafo 1.2.

 © 978-88-08-06485-1

(ii) $c'(t) = kc(t)^2$,
con $k > 0$. In particolare:
a) tracciare il grafico di $c(t)$ in un intervallo di tempo $[0, mT]$, $m \in \mathbb{N}$;
b) studiare il comportamento asintotico di $c(t)$ per $t \to +\infty$ e dimostrare che esiste una concentrazione limite c_∞.
c) Si può dire che c_∞ è un punto d'equilibrio? È stabile? Come dipende c_∞ dalla dose c_0?

4.2 Stabilità per sistemi autonomi lineari

Consideriamo il sistema (2.2) e cerchiamo di studiare in generale i suoi punti di equilibrio. Osserviamo anzitutto che, per i sistemi autonomi, se $\overline{y}$ è punto di equilibrio, con la traslazione $z(t) = y(t) - \overline{y}$ il sistema $y' = f(y)$ si trasforma in

$$z' = f(z(t) + \overline{y}) \equiv g(z(t)).$$

Alla soluzione $y(t) = \overline{y}$ del sistema originale corrisponde la soluzione $z(t) = 0$ del nuovo sistema e viceversa. Perciò nella discussione che segue, assumeremo che 0 sia il punto di equilibrio da prendere in esame.

In questo paragrafo supporremo che il sistema (2.2) sia *bidimensionale* e *lineare*, cioè f sia rappresentata da una matrice A 2×2 a coefficienti reali[21]. In questo caso l'analisi può essere condotta facilmente in modo completo, fino a una classificazione di tutti i possibili comportamenti delle traiettorie vicino a 0.

Consideriamo dunque un sistema omogeneo bidimensionale:

$$\begin{cases} x' = ax + by \\ y' = cx + dy \end{cases}, \quad \text{con } a, b, c, d \in \mathbb{R}. \tag{4.12}$$

Sarà allora

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

e supporremo $\det A = ad - bc \neq 0$, cosicché l'origine è l'*unico* stato di equilibrio del sistema. Osserviamo che un'equazione del second'ordine lineare omogenea a coefficienti costanti[22],

$$x'' = Ax' + Bx$$

può riscriversi nella forma di un sistema (4.12), ponendo

$$\begin{cases} x' = y \\ y' = Ay + Bx. \end{cases}$$

Viceversa, al sistema (4.12) corrisponde sempre un'equazione del second'ordine a coefficienti costanti. Infatti, derivando la seconda equazione (4.12) abbiamo:

$$y'' = cx' + dy'$$

da cui, sostituendo in quest'equazione x' e y' dati dalle (4.12), otteniamo:

$$y'' = c(a + d)x + (cb + d^2)y. \tag{4.13}$$

[21] Alla fine del paragrafo accenneremo brevemente ai sistemi lineari n dimensionali.
[22] v. capitolo 1, paragrafo 3.4.

© 978-88-08-06485-1

Ricavando poi dalla seconda delle (4.12) $cx = y' - dy$ e sostituendola nella (4.13), abbiamo infine:

$$y'' = (a + d)y' - (cb - ad)y = (\mathrm{Tr}A)y' - (\det A)y. \tag{4.14}$$

Nel capitolo 1 abbiamo esaminato le *soluzioni* dell'equazione (4.14); qui analizzeremo le *orbite* del sistema corrispondente (4.12). Lo studente vedrà come, anche in questo caso elementare, l'analisi delle orbite descrive con maggiore chiarezza ed evidenza il quadro generale, che non lo studio delle soluzioni.

Per seguire la prossima discussione, conviene ricordare che, come abbiamo visto nel capitolo 1, paragrafo 3.4, le soluzioni della (4.14) sono esponenziali $e^{\lambda t}$, con λ soluzione dell'equazione caratteristica[23]:

$$\lambda^2 - (a + d)\lambda - (cb - ad) = 0$$

che si può scrivere nella forma

$$\begin{vmatrix} a - \lambda & b \\ c & d - \lambda \end{vmatrix} = (a - \lambda)(d - \lambda) - cb = 0$$

e perciò λ è un autovalore della matrice A. Se h è l'autovettore corrispondente, *la soluzione del sistema* (4.12) *sarà* $\mathrm{h}e^{\lambda t}$. Infatti, per definizione di autovalore, $Ah = \lambda h$ e quindi

$$\lambda \mathrm{h}e^{\lambda t} = A\mathrm{h}e^{\lambda t}$$

che, ponendo

$$\begin{pmatrix} x \\ y \end{pmatrix} = \mathrm{h}e^{\lambda t},$$

si riscrive come

$$\begin{pmatrix} x' \\ y' \end{pmatrix} = \lambda \mathrm{h}e^{\lambda t} = A\mathrm{h}e^{\lambda t} = A\begin{pmatrix} x \\ y \end{pmatrix},$$

pertanto $\mathrm{h}e^{\lambda t}$ è soluzione del sistema (4.12).

Esempio

L'equazione

$$y'' = \omega^2 y$$

è equivalente al sistema $y' = Ay$ con

$$A = \begin{pmatrix} 0 & 1 \\ \omega^2 & 0 \end{pmatrix} \quad \text{cioè} \quad \begin{cases} y_1' = y_2 \\ y_2' = \omega^2 y_1. \end{cases}$$

Gli autovalori di A sono $\lambda_1 = \omega$; $\lambda_2 = -\omega$ e i corrispondenti autovettori sono

$$\mathrm{h}^1 = \begin{pmatrix} 1 \\ \omega \end{pmatrix} \quad \text{e} \quad \mathrm{h}^2 = \begin{pmatrix} 1 \\ -\omega \end{pmatrix}.$$

[23] Questo è vero nel caso in cui le radici dell'equazione caratteristica siano semplici; lo studente modifichi l'osservazione nel caso della radice doppia.

 © 978-88-08-06485-1

L'integrale generale dell'equazione si scrive:

$$y(t) = c_1 e^{\omega t} + c_2 e^{-\omega t}$$

(con c_1, c_2 costanti arbitrarie), mentre l'integrale generale del sistema si scrive:

$$\mathrm{y}(t) = c_1 \mathrm{h}^1 e^{\omega t} + c_2 \mathrm{h}^2 e^{-\omega t}.$$

Consideriamo dunque il sistema (4.12), indichiamo con λ_1 e λ_2 gli autovalori di A, soluzioni dell'equazione caratteristica, e distinguiamo vari casi secondoché la matrice A possegga:

i) due autovalori λ_1, λ_2 reali e distinti;

ii) due autovalori reali coincidenti;

iii) due autovalori complessi coniugati.

Distingueremo poi dei sottocasi in corrispondenza del segno degli autovalori.

i) *Autovalori reali e distinti.* Il sistema possiede le due soluzioni (linearmente indipendenti) $\mathrm{h}^1 e^{\lambda_1 t}$, $\mathrm{h}^2 e^{\lambda_2 t}$, dove h^1 e h^2 sono autovettori (reali) di A; ogni soluzione reale sarà della forma

$$\varphi(t) = c_1 \mathrm{h}^1 e^{\lambda_1 t} + c_2 \mathrm{h}^2 e^{\lambda_2 t},$$

con c_1, c_2 costanti arbitrarie.

i1) Sia $\lambda_1 < \lambda_2 < 0$. L'origine è globalmente asintoticamente stabile. Infatti ogni orbita, partendo da un punto qualsiasi dello spazio delle fasi (x, y), si avvicina indefinitamente all'origine. Prendendo $c_2 = 0$ si hanno le due traiettorie rettilinee indicate nella figura 8.9a con γ_1^+ e γ_1^-; prendendo $c_1 = 0$ si hanno le due traiettorie γ_2^+ e

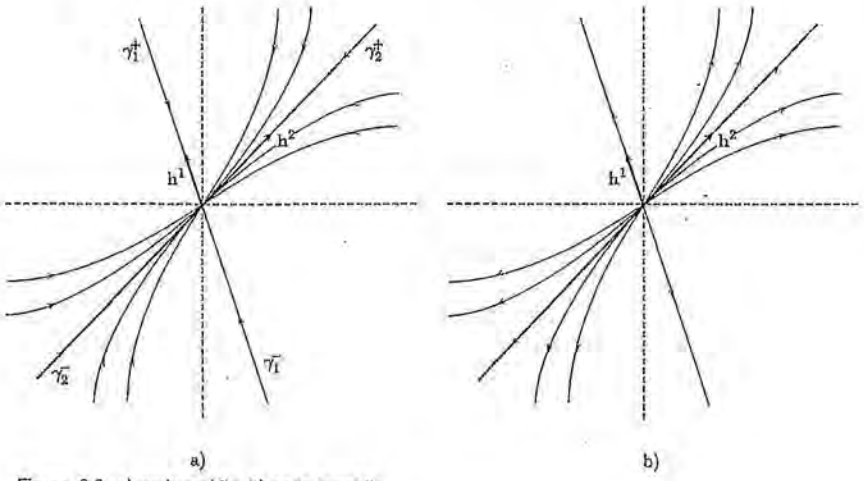


Figura 8.9. a) nodo stabile. b) nodo instabile

γ_2^-; l'andamento delle altre traiettorie è pure illustrato nella figura 10a. Si osservi che, essendo $\lambda_2 > \lambda_1$, per $t \to +\infty$ il termine $c_1 \mathrm{h}^1 e^{\lambda_1 t}$ è infinitesimo di ordine superiore rispetto all'altro $c_2 \mathrm{h}^2 e^{\lambda_2 t}$, cosicché tutte le traiettorie (tranne $\gamma_1^{\pm}$) sono asintotiche a $\gamma_2^{\pm}$.

Il punto di equilibrio prende il nome di *nodo* (stabile).

i2) Sia $\lambda_2 > \lambda_1 > 0$. L'origine è instabile. La discussione di questo caso è del tutto analoga alla precedente; le traiettorie hanno, geometricamente, lo stesso andamento, ma il punto mobile, rappresentante la soluzione, partendo da un punto qualsiasi dello spazio delle fasi, prossimo quanto si vuole all'origine, si allontana indefinitamente da essa.

Si parla in questo caso di *nodo* (instabile). (fig. 8.9b)

i3) Sia $\lambda_1 < 0 < \lambda_2$. L'origine è instabile. Per $c_2 = 0$ abbiamo le traiettorie rettilinee $\gamma_1^{\pm}$ lungo le quali il punto mobile si avvicina indefinitamente all'origine (essendo $\lambda_1 < 0$); per $c_1 = 0$ abbiamo le traiettorie $\gamma_2^{\pm}$ lungo le quali il punto si allontana indefinitamente dall'origine (essendo $\lambda_2 > 0$). Ogni altra traiettoria risulta asintotica ad una del tipo γ_1 per $t \to -\infty$ e ad una del tipo γ_2 per $t \to +\infty$; l'andamento è illustrato in figura 8.10.

L'origine si chiama, in questo caso, *colle* o *sella*.

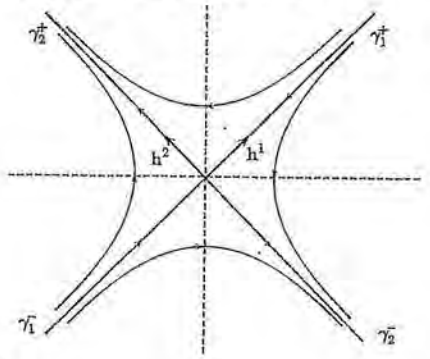


Figura 8.10. Colle o sella.

ii) *Autovalori reali e coincidenti.* Sia λ l'autovalore doppio della matrice A. La soluzione generale del sistema è della forma

$$\varphi(t) = \mathrm{c}^1 e^{\lambda t} + \mathrm{c}^2 t e^{\lambda t},$$

dove $\mathrm{c}^1, \mathrm{c}^2$ sono due *opportuni* vettori, dipendenti da due sole costanti arbitrarie[24].

[24] Infatti, per determinare c^1 e c^2 possiamo inserire φ nel sistema $\varphi' = \mathrm{A}\varphi$, ottenendo:

$$\lambda \mathrm{c}^1 e^{\lambda t} + \mathrm{c}^2 (1 + \lambda t) e^{\lambda t} = \mathrm{A}\mathrm{c}^1 e^{\lambda t} + \mathrm{A}\mathrm{c}^2 t e^{\lambda t}.$$

Uguagliando i coefficienti dei termini corrispondenti, abbiamo il sistema

$$\begin{cases} \mathrm{A}\mathrm{c}^1 = \lambda \mathrm{c}^1 + \mathrm{c}^2 \\ \mathrm{A}\mathrm{c}^2 = \lambda \mathrm{c}^2. \end{cases}$$

Risolvendolo in $\mathrm{c}^1, \mathrm{c}^2$, si trova appunto che questi vettori dipendono da due sole costanti arbitrarie.

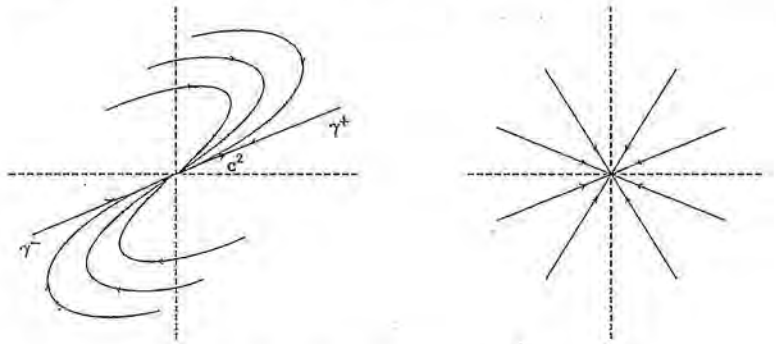


Figura 8.11. a) nodo stabile. b) nodo a stella stabile.

Se $\lambda < 0$, l'origine è globalmente asintoticamente stabile: tutte le traiettorie convergono verso l'origine e sono asintotiche alle traiettorie rettilinee di equazione

$$\varphi(t) = \mathrm{c}^2 t e^{\lambda t},$$

indicate nella figura 8.11a con γ^+ e γ^-.

Se $\lambda > 0$, tutte si allontanano dall'origine, che è instabile.

Il punto 0 si chiama ancora *nodo*.

Un caso speciale si ha se la matrice A è diagonale (e allora $a = d$ mentre $c = b = 0$); la soluzione generale è semplicemente:

$$\varphi(t) = \mathrm{c} e^{\lambda t},$$

con c vettore arbitrario. Le traiettorie sono allora tutte rettilinee (fig. 8.11b) e il punto 0 si dice anche *nodo a stella*, stabile se $\lambda < 0$, instabile se $\lambda > 0$.

iii) *Autovalori complessi coniugati.* Siano $\lambda = \alpha + i\beta$ e $\overline{\lambda} = \alpha - i\beta$ gli autovalori di A; gli autovettori corrispondenti possono essere scelti in modo che risultino anch'essi coniugati; abbiamo così la coppia di soluzioni (linearmente indipendenti)

$$\mathrm{h} e^{\alpha t + i\beta t}, \quad \overline{\mathrm{h}} e^{\alpha t - i\beta t}. \tag{4.15}$$

Conviene sostituire a queste soluzioni una coppia di soluzioni reali. Posto $\mathrm{h} = \mathrm{h}^1 + i\mathrm{h}^2$, con $\mathrm{h}^1, \mathrm{h}^2$ vettori reali, prendendo la semisomma e la semidifferenza delle due soluzioni (4.15), otteniamo le due soluzioni reali e linearmente indipendenti:

$$\varphi_1(t) = e^{\alpha t}\left(\mathrm{h}^1 \cos\beta t - \mathrm{h}^2 \sin\beta t\right);$$
$$\varphi_2(t) = e^{\alpha t}\left(\mathrm{h}^1 \sin\beta t + \mathrm{h}^2 \cos\beta t\right).$$

Quindi l'integrale generale è dato da

$$\varphi(t) = c_1\varphi_1(t) + c_2\varphi_2(t) = e^{\alpha t}\left[\left(c_1\mathrm{h}^1 + c_2\mathrm{h}^2\right)\cos\beta t + \left(c_2\mathrm{h}^1 - c_1\mathrm{h}^2\right)\sin\beta t\right]$$

con c_1, c_2 costanti reali arbitrarie. Distinguiamo ora i seguenti casi.

iii1) Se $\alpha = 0$ (autovalori immaginari puri), l'origine è stabile (ma non asintoticamente stabile); le traiettorie sono delle ellissi. Proviamo che in tal caso il sistema ha un integrale primo della forma:

$$E(x,y) = cx^2 - 2axy - by^2 = \text{costante}. \tag{4.16}$$

Infatti, essendo gli autovalori immaginari puri, risulta $a = -d$ e l'equazione differenziale delle traiettorie (4.5) si scrive

$$\frac{dy}{dx} = \frac{ax + by}{cx - ay}.$$

Essa è pertanto un'*equazione differenziale esatta*[25] e una funzione potenziale è data da $E(x,y)$ come in (4.16). L'integrale generale è quindi assegnato dalla famiglia di ellissi $E(x,y) =$ costante, le quali sono anche la famiglia delle traiettorie.

Abbiamo quindi un'infinità di orbite periodiche; l'origine, in questo caso, è un *centro* (fig. 8.12a).

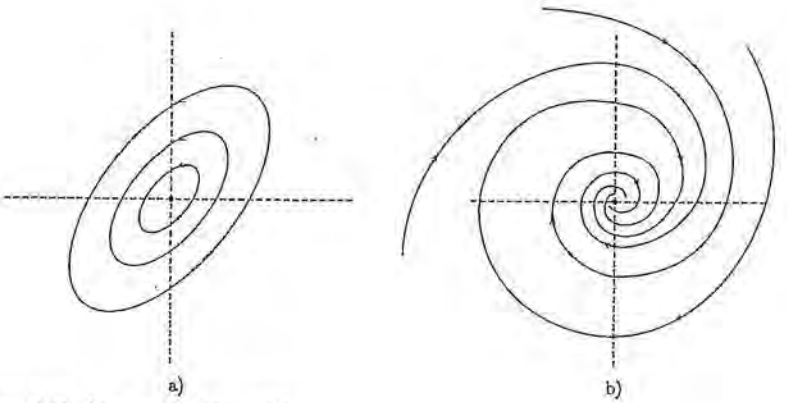


Figura 8.12. a) centro. b) fuoco stabile.

iii2) Se $\alpha < 0$, l'origine è globalmente asintoticamente stabile. Tutte le traiettorie si avvicinano indefinitamente all'origine, con andamento a spirale (fig. 8.12b). L'origine è un *fuoco* stabile.

iii3) Se $\alpha > 0$, l'andamento geometrico delle traiettorie è lo stesso che nel caso precedente, ma il punto mobile si allontana indefinitamente dall'origine. L'origine è un *fuoco* instabile.

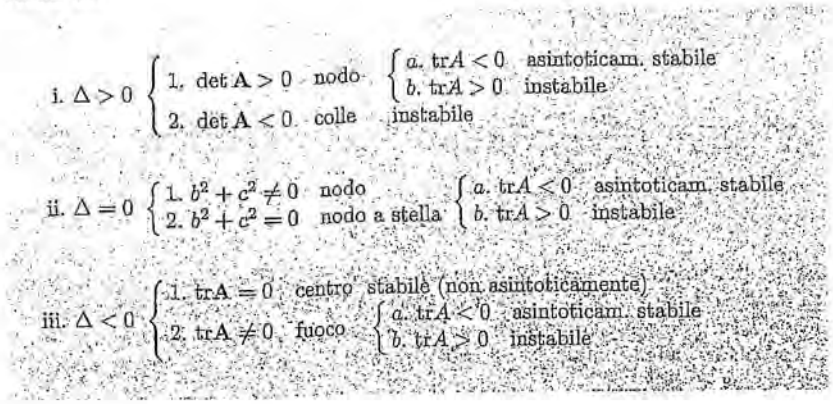
L'allievo è invitato a completare la discussione con l'esame del caso degenere, cioè quando il rango della matrice A è 1.

Possiamo riassumere i risultati nel seguente quadro relativo alla natura di 0. Poniamo $\Delta = (tr\,\mathrm{A})^2 - 4\det\mathrm{A}$.

[25] v. paragrafo 1.2.

 © 978-88-08-06485-1

Tabella 8.1.



Ricordiamo che nello schema precedente si parla sempre di stabilità per $t \to +\infty$.

Applichiamo ora quanto detto allo studio delle piccole oscillazioni di un pendolo con attrito.

Esempio

Piccole oscillazioni del pendolo con attrito Riprendiamo lo studio delle oscillazioni del pendolo (Esempio 4.5) inserendo una forza d'attrito proporzionale alla velocità angolare φ', limitatamente però a "piccole oscillazioni", in modo da poter considerare $\sin\varphi \simeq \varphi$. L'equazione diventa:

$$\varphi'' = -k\varphi - h\varphi'$$

(con $k, h > 0$), corrispondente al sistema $(z_1 = \varphi, z_2 = \varphi')$

$$\begin{cases} z_1' = z_2 \\ z_2' = -kz_1 - hz_2. \end{cases} \tag{4.17}$$

Ritroviamo così la situazione esaminata nello studio delle vibrazioni meccaniche (capitolo 1, paragrafo 3.6): mentre là si sono esaminate le soluzioni dell'equazione, qui analizzeremo le orbite del sistema corrispondente.

La matrice dei coefficienti è:

$$\mathrm{A} = \begin{pmatrix} 0 & 1 \\ -k & -h \end{pmatrix}.$$

Si ha $\mathrm{tr}\mathrm{A} = -h < 0$ e $\det \mathrm{A} = k > 0$. Distinguiamo dunque 3 casi in riferimento al quadro riassuntivo (tab. 8.1).

Sia $h^2 \geq 4k$ (attrito forte). Siamo nel caso i.1.a o ii.1.a. L'origine è un *nodo asintoticamente stabile*. Il pendolo non esegue oscillazioni, ma va direttamente alla posizione di equilibrio (che raggiunge però in un tempo infinito).

Sia $0 < h^2 < 4k$ (attrito debole). Siamo nel caso ii.2.b. L'origine è un *fuoco asintoticamente stabile*. Il pendolo esegue oscillazioni smorzate, di ampiezza sempre minore, fino a raggiungere asintoticamente, la posizione di equilibrio.

Nel caso $h = 0$, assenza di attrito, l'origine è un *centro*: il pendolo esegue indefinitamente oscillazioni sempre della medesima ampiezza.

Per sistemi lineari autonomi di ordine n lo studio del comportamento delle orbite intorno all'origine (punto di equilibrio) è certamente più complesso. Riportiamo, senza entrare nei dettagli e senza dimostrazione, il seguente risultato.

TEOREMA 8.11 (STABILITÀ PER SISTEMI n DIMENSIONALI) *Indichiamo con λ_j, per $j = 1, 2, \ldots, k$ gli autovalori distinti della matrice* A, $n \times n$.

1. *L'origine è un punto di equilibrio asintoticamente stabile se e solo se* $\mathrm{Re}\,\lambda_j < 0$ *per* $j = 1, 2, \ldots, k$. *La stabilità è globale in* $\mathbb{R}^n$.
2. *L'origine è stabile, non asintoticamente, se e solo se* $\mathrm{Re}\,\lambda_j \leq 0$ *e tutti gli autovalori con parte reale nulla sono regolari (cioè la dimensione dell'autospazio è uguale alla molteplicità dell'autovalore).*
3. *L'origine è instabile negli altri casi.*

Esercizi

Studiare il sistema (4.2) nel caso degenere, cioè quando $\det \mathrm{A} = ad - bc = 0$ (ma $\mathrm{A} \neq 0$). Si osservi che ora l'origine non è l'unico punto di equilibrio, ma c'è un'intera retta formata da punti di equilibrio. Si studi il comportamento delle orbite.

Determinare la natura del punto di equilibrio $(0, 0)$ per i seguenti sistemi:

$$(a)\ \begin{cases} x' = 2x + 3y \\ y' = -x - y \end{cases} \qquad (b)\ \begin{cases} x' = 4x - 3y \\ y' = -8x - y \end{cases}$$

$$(c)\ \begin{cases} x' = -3x + 4y \\ y' = -2x + 3y \end{cases} \qquad (d)\ \begin{cases} x' = -3x + y \\ y' = -x - y \end{cases}$$

$$(e)\ \begin{cases} x' = 3x + y \\ y' = x + 3y \end{cases} \qquad (f)\ \begin{cases} x' = -2x + 3y \\ y' = 3x - 2y \end{cases}$$

$$(g)\ \begin{cases} x' = y \\ y' = -x + 2y \end{cases} \qquad (h)\ \begin{cases} x' = x - y \\ y' = x + y \end{cases}$$

4.3 Stabilità per sistemi autonomi non lineari

Per un sistema lineare autonomo la stabilità di un punto di equilibrio è completamente risolta dal teorema 8.11, che abbiamo visto illustrato nei dettagli nel caso bidimensionale. Se il sistema non è lineare, la situazione è molto meno soddisfacente.

Un modo di procedere è cercare di sostituire il sistema originale con un sistema lineare approssimante, utilizzare la teoria lineare e cercare di trasferire le informazioni ottenute a quello non lineare.

In generale, però, l'approssimazione è efficace solo nell'intorno del punto di equilibrio considerato e perciò, anche se le informazioni sul sistema lineare hanno carattere globale, la loro validità in riferimento al sistema originale è per lo più locale. È un tipico esempio della dicotomia locale-globale associata a non lineare-lineare.

 © 978-88-08-06485-1

Il procedimento ora accennato prende il nome di *metodo di linearizzazione*.

Consideriamo un sistema autonomo

$$\mathbf{y}' = \mathbf{f}(\mathbf{y}) \tag{4.18}$$

con $\mathbf{f} \in C^1(D)$, D aperto di $\mathbb{R}^n$ e tale che $\mathbf{f}(\mathbf{0}) = \mathbf{0}$. Ipotesi fondamentale per tutto il seguito è che la matrice jacobiana di $\mathbf{f}$ in $\mathbf{0}$, $\mathrm{Df}(\mathbf{0})$, sia non singolare. Si può dimostrare che questo, tra l'altro, assicura che $\mathbf{0}$ sia un punto di equilibrio *isolato*, ossia che esiste un intorno di $\mathbf{0}$ in cui non vi sono altri punti di equilibrio del sistema.

Essendo $\mathbf{f} \in C^1(D)$, si può scrivere:

$$\mathbf{f}(\mathbf{y}) = \mathrm{Df}(\mathbf{0})\,\mathbf{y} + \mathbf{g}(\mathbf{y})$$

con $\mathbf{g}(\mathbf{y}) = o(|\mathbf{y}|)$ per $\mathbf{y} \to \mathbf{0}$. Poniamo $\mathrm{A} = \mathrm{Df}(\mathbf{0})$ e consideriamo il sistema lineare

$$\mathbf{z}' = \mathrm{A}\mathbf{z}, \tag{4.19}$$

che prende il nome di sistema linearizzato (in $\mathbf{y} = \mathbf{0}$).

Il risultato fondamentale, che ci limitiamo a enunciare, è il seguente:

TEOREMA 8.12 (DI LINEARIZZAZIONE)

(i) *Se* $\mathbf{0}$ *è asintoticamente stabile per il sistema linearizzato* (4.19), *allora lo è anche per il sistema non lineare* (4.18).

(ii) *Se esiste un autovalore di* A *con parte reale positiva, allora il punto* $\mathbf{0}$ *è instabile per il sistema non lineare* (4.18).

Illustreremo ora questo teorema con alcuni esempi bidimensionali.

Se riprendiamo in esame il sistema di Lotka-Volterra (Esempi 2.1 e 4.3) e l'equazione del pendolo (Esempio 4.5) e applichiamo il metodo di linearizzazione sopra esposto, ritroviamo i risultati che l'esame diretto delle orbite già ci aveva permesso di ottenere: il punto di equilibrio del sistema di Lotka-Volterra è stabile (non asintoticamente); nel sistema del pendolo l'origine è stabile, mentre i punti $(\pm\pi, 0)$ sono instabili (corrispondono alla posizione verticale del pendolo al di sopra del perno.

Esempio

Grandi oscillazioni di un pendolo con attrito Consideriamo l'equazione

$$\varphi'' = -k\sin\varphi - h\varphi'$$

$(h, k > 0)$ che descrive le oscillazioni di un pendolo in presenza di una forza di attrito $(h > 0)$, corrispondente al sistema $(z_1 = \varphi, z_2 = \varphi')$

$$\begin{cases} z_1' = z_2 \\ z_2' = -k\sin z_1 - hz_2. \end{cases} \tag{4.20}$$

Ricordiamo che nell'esempio 4.7 abbiamo studiato questo sistema nell'ipotesi delle piccole oscillazioni; quell'ipotesi corrispondeva a considerare $\sin\varphi \simeq \varphi$ e quindi avere un sistema lineare. Ora esaminiamo il problema nel caso di oscillazioni di ampiezza qualsiasi, che porta

© 978-88-08-06485-1

a un sistema non lineare, e lo studiamo per linearizzazione. I punti di equilibrio del sistema sono $(0,0)$, $(\pi,0)$ (gli altri punti $(k\pi, 0)$ corrispondono fisicamente ai due punti precedenti).

Esaminiamo prima l'origine. Linearizzando in $(0,0)$ il sistema (4.20), ritroviamo il sistema (4.17) studiato nell'Esempio 4.7: sappiamo che in quel caso, per $h > 0$, l'origine è asintoticamente stabile; perciò, in base al teorema 8.12, lo sarà anche nel caso non lineare in esame.

Per linearizzare l'altro punto di equilibrio, $(\pi, 0)$, bisogna prima calcolare la matrice jacobiana di

$$\mathbf{f}(\mathbf{z}) = (z_2, -k\sin z_1 - hz_2)$$

in $(\pi, 0)$. Si trova

$$\mathrm{A} = \begin{pmatrix} 0 & 1 \\ k & -h \end{pmatrix}.$$

Essendo $\det \mathrm{A} = -k < 0$ e $\mathrm{tr}\mathrm{A} = -h, \Delta = (\mathrm{tr}\mathrm{A})^2 - 4\det \mathrm{A} > 0$, siamo nel caso i.2 della tabella 8.1, che corrisponde a un colle. Pertanto esiste un autovalore con parte reale positiva e l'origine è instabile per entrambi i sistemi (linearizzato e non lineare).

In figura 8.13 sono raffigurate le traiettorie del sistema per $h^2 < 4k$ (attrito debole). L'andamento delle traiettorie in figura suggerisce che si mantenga non solo la stabilità (nel caso $h > 0$), ma anche la configurazione delle traiettorie in un intorno di 0.

Astraendo dall'esempio specifico ora considerato, e rimanendo a un livello puramente intuitivo, si può dire che se $\mathbf{0}$ è un *fuoco*, un *colle* o un *nodo* (non a stella) per il sistema linearizzato, allora il ritratto di fase in un intorno di $\mathbf{0}$ per il sistema non lineare sarà *topologicamente equivalente*, cioè qualitativamente simile. Si potrà così parlare di *fuoco*, *colle* o *nodo* anche per il *sistema non lineare*.

Ciò significa sostanzialmente che una "piccola" perturbazione sulla funzione $\mathbf{f}$ non altera localmente il ritratto di fase in questi casi. Si parla allora di *stabilità strutturale* del sistema. Come l'esempio stesso del pendolo indica, non ci si può aspettare che un centro sia strutturalmente stabile: una piccola perturbazione può distruggere la periodicità di una soluzione.

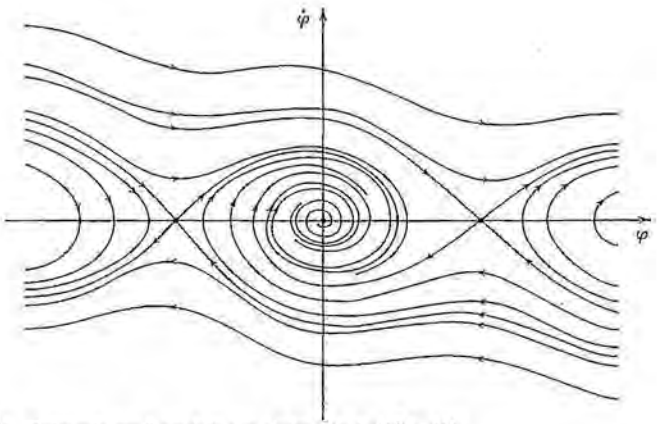


Figura 8.13. Traiettorie per il pendolo con attrito debole ($h^2 < 4k$).

 © 978-88-08-06485-1

Osserviamo infine che, mentre per i sistemi lineari la topologia dei punti di equilibrio si riduce ai sei casi elencati nella tabella 8.1, quella per i sistemi non lineari è molto più ricca; non ci occupiamo di questo argomento, rimandando il lettore interessato a testi più specialistici.

Cicli limite

Nei modelli di Lotka-Volterra e del pendolo senza attrito, esistono infinite orbite chiuse che circondano un punto di equilibrio. Può succedere che esistano una o più orbite chiuse isolate sulle quali, almeno localmente, le altre traiettorie si avvolgono a spirale. Queste orbite si chiamano *cicli limite* e tipicamente indicano la presenza di *oscillazioni autosostenute*. Per i sistemi bidimensionali si può dimostrare che un'orbita chiusa contiene *sempre* un punto di equilibrio al suo interno. Un classico esempio è l'equazione di van der Pol, che abbiamo già considerato nell'Esempio 2.3.

Esempio

Equazione di van der Pol Si tratta della seguente equazione:

$$y'' = -y + \varepsilon\left(1 - y^2\right) y' \quad \text{con } \varepsilon > 0 \text{ costante.} \tag{4.21}$$

Come già osservato nell'Esempio 2.3, quest'equazione si può interpretare come una perturbazione dell'oscillatore armonico (con $\omega^2 = 1$) mediante il termine forzante $\varepsilon\left(1 - y^2\right) y'$ che per valori grandi di y ($|y| > 1$) è una resistenza attiva (dissipa energia) e contrasta il moto, mentre per piccoli valori di y ($|y| < 1$) rappresenta una resistenza negativa (fornisce energia), che nei circuiti elettrici si realizza come *feedback*, o controllo retroattivo.

Ponendo $y_1 = y$ e $y_2 = y'$, l'equazione si riduce al sistema:

$$\begin{cases} y_1' = y_2 \\ y_2' = -y_1 + \varepsilon\left(1 - y_1^2\right) y_2 \end{cases}$$

che ha come unico punto di equilibrio l'origine. I commenti alla (4.21) fatti sopra ci portano a congetturare che le soluzioni si assestino lungo un'orbita chiusa, corrispondente a una soluzione periodica.

Usando un computer, proviamo a disegnare alcune orbite nel piano delle fasi y, y'. Il risultato è quello indicato in figura 8.14. I calcoli numerici confortano la nostra congettura.

Occorre però un'analisi teorica, che si può condurre meglio, anziché nel piano delle fasi, nel cosiddetto piano di Lienard[26], ossia delle variabili y e

$$z = y' + \varepsilon\left(\frac{y^3}{3} - y\right).$$

Infatti, osservando che

$$y'' + \varepsilon\left(y^2 - 1\right) y' = \frac{d}{dt}\left[y' + \varepsilon\left(\frac{y^3}{3} - y\right)\right],$$

la (4.21) si riduce al sistema seguente:

$$\begin{cases} y' = z - \varepsilon\left(\frac{y^3}{3} - y\right) \\ z' = -y. \end{cases} \tag{4.22}$$

[26] Alfred Lienard (1869-1958), fisico e matematico francese.

© 978-88-08-06485-1

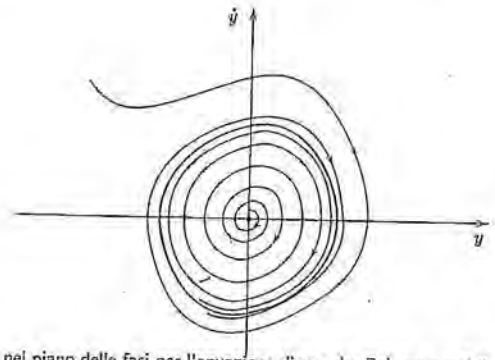


Figura 8.14. Orbita nel piano delle fasi per l'equazione di van der Pol con $\varepsilon = 0{,}2$. Si noti la presenza di un ciclo limite (linea spessa) stabile.

La corrispondenza

$$(y, y') \leftrightarrow (y, z)$$

tra punti dei due piani è chiaramente biunivoca e regolare, cosicché il ritratto di fase è qualitativamente simile nei due casi. In particolare, punti di equilibrio e orbite chiuse si corrispondono nei due piani. L'unico punto di equilibrio è $(0, 0)$. Linearizzando si trova:

$$\begin{cases} y' = \varepsilon y + z \\ z' = -y. \end{cases}$$

a cui è associata la matrice

$$\mathrm{A} = \begin{pmatrix} \varepsilon & 1 \\ -1 & 0 \end{pmatrix}.$$

Poiché $\mathrm{tr}\mathrm{A} = \varepsilon > 0, \det \mathrm{A} = 1 > 0, \Delta = (\mathrm{tr}\mathrm{A})^2 - 4 \det \mathrm{A} < 0$, siamo nel caso iii.2.b del quadro riassuntivo: l'origine è un *fuoco instabile*. In base al teorema 8.12, è un fuoco instabile anche per il sistema (4.22) e le orbite che partono in un intorno di $(0, 0)$ se ne allontanano a spirale.

Infine, si dimostra che: *esiste un unico ciclo stabile, nel senso che ogni traiettoria diversa da $(0, 0)$ tende ad avvolgersi su di esso per $t \to +\infty$.*

Dal punto di vista delle applicazioni radioingegneristiche, l'esistenza di un ciclo limite rappresenta la possibilità del sistema di oscillare con un periodo indipendente dalle condizioni iniziali.

Concludiamo il paragrafo con un teorema, che ci limitiamo a enunciare, che può essere usato, nel caso bidimensionale, per dimostrare l'esistenza di un ciclo limite.

TEOREMA 8.13 (DI BENDIXSON-POINCARÉ) *Sia dato il sistema bidimensionale autonomo:*

$$\begin{cases} x' = f(x, y) \\ y' = g(x, y) \end{cases}$$

con f, g di classe C^1 in un aperto Ω di $\mathbb{R}^2$. Sia $\Omega' \subset \Omega$ un aperto limitato non contenente punti critici e tale che ogni orbita che entri in Ω' non ne possa più uscire. Allora in Ω' esiste un ciclo limite.

 © 978-88-08-06485-1

Esercizi

26 Studiare, col metodo di linearizzazione, la natura dei punti di equilibrio del sistema:

$$\begin{cases} x' = x^2 + y \\ y' = x - y^2. \end{cases}$$

27 Classificare i punti di equilibrio per i sistemi:

$$(a)\ \begin{cases} x' = x + y^2 - 2 \\ y' = x + y \end{cases}; \quad (b)\ \begin{cases} x' = -x + x^2 y \\ y' = 2 - 2y. \end{cases}$$

28 Per il sistema

$$\begin{cases} x' = -x + 2x^3 y^2 \\ y' = -y \end{cases}$$

tracciare qualitativamente l'andamento delle traiettorie nel piano delle fasi e determinare, in particolare, il loro comportamento nell'intorno dell'origine.

29 Discutere, al variare di $k \in \mathbb{R}$, la stabilità dell'origine per il sistema:

$$\begin{cases} x' = y - xy^2 \\ y' = -x + ky - y^3. \end{cases}$$

A Trasformata di Laplace e trasformata di Fourier

In questo capitolo presentiamo, in forma molto contenuta, due strumenti importanti dell'Analisi matematica, che trovano larghissimo uso nella risoluzione di molti problemi applicati, in particolare quando questi sono formulati per mezzo di equazioni differenziali.

La teoria delle trasformate integrali supera certamente i limiti del presente corso; ma abbiamo ritenuto che possa essere utile per lo studente, che ne abbia la curiosità, avere un'idea delle notevoli possibilità offerte dall'uso di questi strumenti matematici.

1 TRASFORMATA DI LAPLACE. DEFINIZIONE ED ESEMPI

Sia f una funzione di variabile reale, nulla in $(-\infty, 0)$.

Si chiama *trasformata di Laplace* di f la funzione $\mathcal{L}[f]$ definita dalla formula seguente:

$$\mathcal{L}[f](p) = \int_0^{+\infty} e^{-px} f(x)\, dx \tag{1.1}$$

dove $p \in \mathbb{C}$.

L'applicazione

$$\mathcal{L}: \quad f \longmapsto \mathcal{L}[f]$$

si chiama *trasformazione di Laplace.*

La (1.1) richiede qualche spiegazione. Ricordiamo che, se $p = \alpha + i\beta \in \mathbb{C}$, l'integranda della (1.1) si può riscrivere, separando esplicitamente la parte reale dalla parte immaginaria, al modo seguente:

$$e^{-px} f(x) = e^{-\alpha x} \cos(\beta x) f(x) - i e^{-\alpha x} \sin(\beta x) f(x).$$

Per definizione, il simbolo $\int_0^{+\infty} e^{-px} f(x)\, dx$ indica allora il numero complesso

$$\int_0^{+\infty} e^{-\alpha x} \cos(\beta x) f(x)\, dx - i \int_0^{+\infty} e^{-\alpha x} \sin(\beta x) f(x)\, dx$$

 © 978-88-08-06485-1

purché, naturalmente, entrambi gli integrali generalizzati appena scritti convergano; a sua volta, questo è certamente vero se converge l'integrale

$$\int_0^{+\infty} e^{-\alpha x}\,|f(x)|\,dx$$

(si vedano i criteri di integrabilità su intervalli illimitati, studiati nel volume 1, capitolo 6, paragrafo 8.4). Se dunque per un certo $\alpha \in \mathbb{R}$ questo integrale converge, $\mathcal{L}[f](p)$ risulterà definita per $p = \alpha + i\beta$, $\forall \beta \in \mathbb{R}$.

Se $\mathcal{L}[f](p)$ è ben definita per almeno un valore di $p \in \mathbb{C}$, f si dice trasformabile (secondo Laplace).

La seguente proposizione indica che, se f è integrabile in ogni intervallo $[0, A]$ e non cresce più che esponenzialmente per $x \to +\infty$, allora è trasformabile. Più precisamente:

TEOREMA A.1 *Se per x abbastanza grande si può scrivere $|f(x)| \le Me^{\alpha x}$, con M ed α positivi, allora f è trasformabile, $\mathcal{L}[f]$ esiste nel semipiano complesso $\mathrm{Re}\, p > \alpha$ ed inoltre*

$$\lim_{\mathrm{Re}\, p \to +\infty} \mathcal{L}[f](p) = 0\,. \tag{1.2}$$

DIMOSTRAZIONE. Si ha $|e^{-px} f(x)| = e^{-(\mathrm{Re}\, p)x}|f(x)| < Me^{(\alpha - \mathrm{Re}\, p)x}$ per x abbastanza grande. Se $\mathrm{Re}\, p > \alpha$, la funzione $Me^{(\alpha - \mathrm{Re}\, p)x}$ è integrabile in senso generalizzato, per cui lo è anche $e^{-px} f(x)$. Inoltre, si ha

$$|\mathcal{L}[f](p)| \le \int_0^{+\infty} Me^{(\alpha - \mathrm{Re}\, p)x}\,dx = \frac{M}{\mathrm{Re}\, p - \alpha} \longrightarrow 0 \quad \text{se } \mathrm{Re}\, p \to +\infty\,. \qquad \blacklozenge$$

In generale si può dimostrare che:

se $\mathcal{L}[f](p)$ esiste per $p = p_0$, allora esiste anche nel semipiano $\{p \in \mathbb{C} : \mathrm{Re}\, p > \mathrm{Re}\, p_0\}$ e vale la (1.2).

Se α_0 è l'estremo inferiore dei numeri α tali che $\mathcal{L}[f]$ esiste per $\mathrm{Re}\, p > \alpha$, il semipiano $\mathrm{Re}\, p > \alpha_0$ si chiama *semipiano di convergenza*.

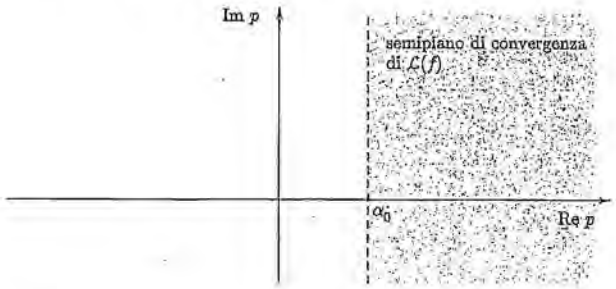


Figura A.1.

Esempi

Funzione di Heaviside.

$$H(x) = \begin{cases} 1 & x \ge 0 \\ 0 & x < 0 \end{cases}$$

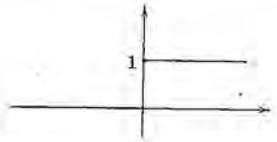


Figura A.2.

Si ha:

$$\mathcal{L}[H](p) = \int_0^{+\infty} e^{-px}\,dx = \lim_{T \to +\infty} \int_0^T e^{-px}\,dx =$$

$$= \lim_{T \to +\infty} \left[-\frac{1}{p} e^{-px} \right]_0^T = \lim_{T \to +\infty} \frac{1 - e^{-pT}}{p}\,.$$

Il limite esiste finito solo se $\mathrm{Re}\, p > 0$; dunque

$$\boxed{\mathcal{L}[H](p) = \frac{1}{p}}$$

con semipiano di convergenza $\mathrm{Re}\, p > 0$.

Impulso unitario di durata ε.

$$I_\varepsilon(x) = \begin{cases} \dfrac{1}{\varepsilon} & 0 \le x < \varepsilon \\ 0 & x \ge \varepsilon \end{cases}$$

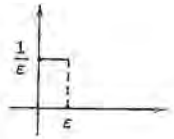


Figura A.3.

Notiamo che $\int_0^{+\infty} I_\varepsilon(x)\,dx = 1$ per ogni $\varepsilon > 0$.

Si ha

$$\mathcal{L}[I_\varepsilon](p) = \int_0^\varepsilon \frac{1}{\varepsilon} e^{-px}\,dx = \frac{1}{\varepsilon}\left[-\frac{1}{p} e^{-px} \right]_0^\varepsilon = \frac{1 - e^{-p\varepsilon}}{p\varepsilon}\,. \tag{1.3}$$

Poiché $\lim_{p \to 0} \mathcal{L}[I_\varepsilon](p) = 1$, ponendo $\mathcal{L}[I_\varepsilon](0) := 1$, si estende la definizione di $\mathcal{L}[I_\varepsilon]$ a tutto il piano complesso.

Impulso unitario istantaneo (distribuzione di Dirac). La distribuzione di Dirac nell'origine, indicata con il simbolo $\delta(x)$ è, ricordiamo, il "limite" per $\varepsilon \to 0$ dell'impulso unitario I_ε, definito nell'esempio 1.2:

$$\delta(x) = \lim_{\varepsilon \to 0} I_\varepsilon(x) \tag{1.4}$$

La definizione è solo formale in quanto da (1.4) si ricaverebbe:

$$\delta(x) = 0 \quad \text{se} \quad x \ne 0\,, \qquad \delta(0) = +\infty \tag{1.5}$$

e inoltre

$$\int_{-\infty}^{+\infty} \delta(x)\,dx = 1\,. \tag{1.6}$$

Ora, nessuna funzione dell'analisi classica può avere le proprietà (1.5) e (1.6); una definizione rigorosa trova posto nell'ambito della *teoria delle distribuzioni* di L. Schwartz ed infatti è corretto chiamare $\delta(x)$ distribuzione (anziché funzione) di Dirac.

 © 978-88-08-06485-1

Ancora una volta, ci accontentiamo dell'aspetto formale, interpretando la distribuzione di Dirac come un oggetto matematico adatto a rappresentare un impulso unitario, per la (1.6), concentrato nell'origine, per le (1.5).

La trasformata di Laplace di δ si trova facendo il limite per $\varepsilon \to 0$ delle trasformate di I_ε; dalla (1.3):

$$\mathcal{L}[\delta](p) = \lim_{\varepsilon \to 0} \frac{1 - e^{-p\varepsilon}}{p\varepsilon} = 1 .$$

Dunque:

(1.7) $$\boxed{\mathcal{L}[\delta] = 1}$$

Esponenziale.

$$f(x) = \begin{cases} e^{-ax} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

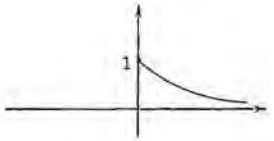


Figura A.4.

con $a \in \mathbb{C}$.

Una scrittura equivalente è $f(x) = e^{-ax} \cdot H(x)$.

Si ha:

$$\mathcal{L}[f](p) = \int_0^{+\infty} e^{-ax} e^{-px}\, dx = \lim_{T \to +\infty} \int_0^T e^{-(a+p)x}\, dx = \lim_{T \to +\infty} \frac{1}{p+a} \left\{ 1 - e^{-(a+p)T} \right\}$$

Se $\mathrm{Re}(a+p) > 0$, $e^{-(a+p)T} \longrightarrow 0$ per $T \to +\infty$ e quindi

$$\boxed{\mathcal{L}[e^{-ax} H(x)](p) = \frac{1}{p+a}}$$

con semipiano di convergenza $\mathrm{Re}\, p > -\mathrm{Re}\, a$.

2 PROPRIETÀ DELLA TRASFORMATA DI LAPLACE

Esaminiamo le relazioni tra la trasformazione di Laplace e le seguenti operazioni: combinazione lineare di funzioni, traslazioni, derivazione, integrazione, prodotto di convoluzione. Quando non sorgono confusioni scriveremo $\mathcal{L}[f]$ anziché $\mathcal{L}[f \cdot H]$ anche per funzioni f non necessariamente nulle in $(-\infty, 0)$.

Linearità

Siano f e g trasformabili con semipiano di convergenza $\mathrm{Re}\, p > \alpha$, $\mathrm{Re}\, p > \beta$, rispettivamente.

Allora, se $\lambda, \mu \in \mathbb{C}$, $\lambda f + \mu g$ è trasformabile e

(2.1) $$\boxed{\mathcal{L}[\lambda f + \mu g](p) = \lambda \mathcal{L}[f](p) + \mu \mathcal{L}[g](p)}$$

con semipiano di convergenza $\mathrm{Re}\, p > \max\{\alpha, \beta\}$.

La dimostrazione della (2.1) segue subito dalla linearità dell'integrale generalizzato.

© 978-88-08-06485-1

Esempi

Calcoliamo la trasformata di Laplace di $\sin \omega x \cdot H(x)$ e $\cos \omega x \cdot H(x)$. Ricordando le formule di Eulero e la (1.4)

$$\sin \omega x = \frac{e^{i\omega x} - e^{-i\omega x}}{2i} \qquad \cos \omega x = \frac{e^{i\omega x} + e^{-i\omega x}}{2}$$

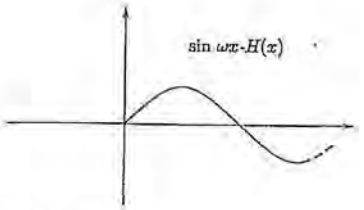


Figura A.5.

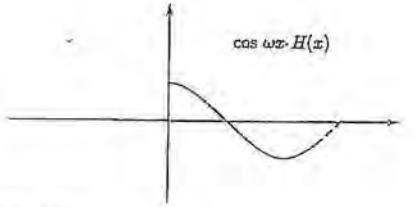


Figura A.6.

(2.2)
$$\begin{aligned} \mathcal{L}[\sin \omega x \cdot H(x)](p) &= \frac{1}{2i} \left\{ \mathcal{L}\left[e^{i\omega x} H(x) \right](p) - \mathcal{L}\left[e^{-i\omega x} H(x) \right](p) \right\} \\ &= \frac{1}{2i} \left\{ \frac{1}{p - i\omega} - \frac{1}{p + i\omega} \right\} = \frac{\omega}{p^2 + \omega^2} \end{aligned}$$

(2.3)
$$\begin{aligned} \mathcal{L}[\cos \omega x \cdot H(x)](p) &= \frac{1}{2} \left\{ \mathcal{L}[e^{i\omega x} H(x)](p) + \mathcal{L}[e^{-i\omega x} H(x)](p) \right\} \\ &= \frac{1}{2} \left\{ \frac{1}{p - i\omega} + \frac{1}{p + i\omega} \right\} = \frac{p}{p^2 + \omega^2} \end{aligned}$$

con semipiano di convergenza $\mathrm{Re}\, p > 0$.

Essendo $(\cos \omega x)^2 = (1 + \cos 2\omega x)/2$, si ha:

$$\begin{aligned} \mathcal{L}[(\cos \omega x)^2 \cdot H(x)](p) &= \frac{1}{2} \mathcal{L}[H](p) + \frac{1}{2} \mathcal{L}[\cos 2\omega x \cdot H(x)](p) \\ &= \frac{1}{2} \cdot \left\{ \frac{1}{p} + \frac{p}{p^2 + 4\omega^2} \right\} \quad (\mathrm{Re}\, p > 0) \end{aligned}$$

Poiché $(\sin \omega x)^2 = 1 - (\cos \omega x)^2$ si ha poi, sempre per $\mathrm{Re}\, p > 0$:

$$\mathcal{L}[(\sin \omega x)^2 H(x)](p) = \frac{1}{p} - \frac{1}{2} \left\{ \frac{1}{p} + \frac{p}{p^2 + 4\omega^2} \right\} = \frac{1}{2} \left\{ \frac{1}{p} - \frac{p}{p^2 + 4\omega^2} \right\}$$

 © 978-88-08-06485-1

Formule del ritardo

Sia f trasformabile con semipiano di convergenza $\mathrm{Re}\, p > \alpha$.

Definiamo, per $a > 0$:

$$g(x) = f(x-a)H(x-a) = \begin{cases} f(x-a) & x \geq a \\ 0 & x < a \end{cases}$$

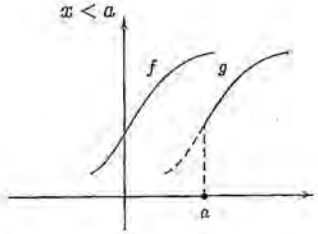


Figura A.7.

Si ha, facendo il cambio di variabile $t = x - a$:

$$\begin{aligned} \mathcal{L}[g](p) &= \int_a^{+\infty} e^{-px} f(x-a)\, dx = \\ &= \int_0^{+\infty} e^{-p(t+a)} f(t) dt = \\ &= e^{-pa} \int_0^{+\infty} e^{-pt} f(t)\, dt = e^{-pa} \mathcal{L}[f](p) \end{aligned}$$

Dunque

$$\boxed{\mathcal{L}[f(x-a)H(x-a)](p) = e^{-pa}\mathcal{L}[f](p)} \qquad (\mathrm{Re}\, p > \alpha) \tag{2.4}$$

D'altra parte, per verifica diretta si ha:

$$\boxed{\mathcal{L}[e^{ax} f(x)](p) = \mathcal{L}[f](p-a)} \qquad (\mathrm{Re}\, p > \alpha + a) \tag{2.4'}$$

Se si interpreta x come variabile temporale e si passa da una funzione $g(x)$ a $g(x-a)$, un determinato valore di g viene osservato con a unità temporali di ritardo. Per questo motivo la funzione $g(x-a)$ si chiama funzione "ritardata" e le formule (2.4)-(2.4') si chiamano *formule del ritardo*.

ESEMPIO

Trasformata di una funzione periodica. Sia f periodica di periodo T. Se si considera

$$f_0(x) = \begin{cases} f(x) & 0 \leq x < T \\ 0 & x < 0 \ \text{ o } \ x \geq T \end{cases}$$

la funzione $f(x) \cdot H(x)$ si può scrivere come somma di infinite traslazioni di f_0 nel modo seguente:

$$\begin{aligned} f(x) \cdot H(x) &= f_0(x) + f_0(x-T) + f_0(x-2T) + \ldots + f_0(x-nT) + \ldots \\ &= \sum_{n=0}^{\infty} f_0(x-nT). \end{aligned} \tag{2.5}$$

Supponendo di poter integrare termine a termine la serie in (2.5), si trova, usando la (2.4)

$$\begin{aligned} \mathcal{L}[f(x) \cdot H(x)](p) &= \sum_{n=0}^{\infty} \mathcal{L}[f_0(x-nT)](p) = \\ &= \mathcal{L}[f_0](p) \cdot \sum_{n=0}^{\infty} e^{-nTp}. \end{aligned} \tag{2.6}$$

© 978-88-08-06485-1

La serie geometrica a secondo membro della (2.6) è convergente ed ha somma $(1 - e^{-Tp})^{-1}$ se $|e^{-Tp}| = e^{-T\,\mathrm{Re}\, p} < 1$ e cioè se $\mathrm{Re}\, p > 0$.

Quindi, nel semipiano $\mathrm{Re}\, p > 0$, vale la formula

$$\mathcal{L}[f(x)H(x)](p) = \frac{\mathcal{L}[f_0](p)}{1 - e^{-Tp}}. \tag{2.7}$$

Ad esempio se $T = 1$ e $f_0(x) = x$ in $[0,1)$ si ha

$$\begin{aligned} \mathcal{L}[f_0](p) &= \int_0^1 e^{-px} x\, dx = \left[-\frac{1}{p} x e^{-px} \right]_0^1 + \frac{1}{p} \int_0^1 e^{-px}\, dx = \\ &= -\frac{e^{-p}}{p} - \frac{1}{p^2} \left[e^{-px} \right]_0^1 = -\frac{e^{-p}}{p} + \frac{1}{p^2}(1 - e^{-p}). \end{aligned}$$

Per la funzione periodica della figura A.8 si ha, dalla (2.7):

$$\mathcal{L}[f](p) = \mathcal{L}[f_0](p) \cdot (1 - e^{-p})^{-1} = \frac{1}{p(1-e^p)} + \frac{1}{p^2}.$$

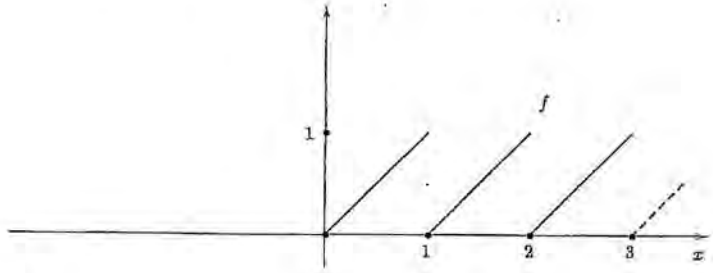


Figura A.8.

Trasformazione di una derivata

TEOREMA A.2 *Sia f derivabile con derivata continua a tratti in $[0,+\infty)$. Sia inoltre f' trasformabile nel semipiano $\mathrm{Re}\, p > \alpha$. Allora f è trasformabile nel semipiano $\mathrm{Re}\, p > \max\{\alpha, 0\}$ e vale la formula*

$$\boxed{\mathcal{L}[f'](p) = p\mathcal{L}[f](p) - f(0)} \tag{2.8}$$

DIMOSTRAZIONE. Per semplicità la facciamo sotto l'ulteriore ipotesi che $|f'(x)| \leq Me^{\alpha x}$ e che f' sia continua. Si ha:

$$\mathcal{L}[f'](p) = \lim_{T \to +\infty} \int_0^T e^{-px} f'(x)\, dx.$$

 © 978-88-08-06485-1

Integrando per parti si ricava

$$\int_0^T e^{-px} f'(x)\,dx = \left[e^{-px} f(x)\right]_0^T + p\int_0^T e^{-px} f(x)\,dx = \tag{2.9}$$
$$= e^{-pT} f(T) - f(0) + p\int_0^T e^{-px} f(x)\,dx.$$

Si mostra poi che, se $\mathrm{Re}\, p > \max\{\alpha, 0\}$, $e^{-pT} f(T) \to 0$. ◆

La (2.8) si può generalizzare se esistono le derivate di f di ordine superiore e sono trasformabili. In particolare, usando la (2.8) prima per f'' e poi per f', si trova

$$\mathcal{L}[f''](p) = p\mathcal{L}[f'](p) - f'(0) = p^2\mathcal{L}[f] - pf(0) - f'(0)$$

cioè:

$$\boxed{\mathcal{L}[f''](p) = p^2\mathcal{L}[f](p) - pf(0) - f'(0)} \tag{2.10}$$

Poiché $\lim_{\mathrm{Re}\, p \to +\infty} \mathcal{L}[f'](p) = 0$, dalla (2.8) si ricava la formula

$$\lim_{\mathrm{Re}\, p \to +\infty} p\mathcal{L}[f](p) = f(0) \tag{2.11}$$

Nel caso in cui f o f' abbiano un *salto* in $x = 0$, nelle (2.8), (2.10) e (2.11), $f(0)$ e $f'(0)$ vanno sostituiti con $f(0^+)$ e $f'(0^+)$.

Le formule (2.8) e (2.10) convertono operazioni di derivazione in operazioni di tipo algebrico; vedremo più avanti alcune applicazioni alle equazioni differenziali.

Trasformazione di una primitiva

Il teorema A.2 e la formula (2.8) si possono interpretare in termini di una funzione e una sua primitiva, in particolare quella che si annulla in $x = 0$.

Sia f continua a tratti e poniamo

$$F(x) = \int_0^x f(t)\,dt$$

Se f è trasformabile nel semipiano $\mathrm{Re}\, p > \alpha$, allora F è trasformabile nel semipiano $\mathrm{Re}\, p > \max\{\alpha, 0\}$ e vale la formula

$$\boxed{\mathcal{L}[F](p) = \frac{\mathcal{L}[f](p)}{p}} \tag{2.12}$$

Esempio

Sappiamo che $\mathcal{L}[H] = \frac{1}{p}$ (cfr. esempio 1.1).

Poiché $F_1(x) = \int_0^x H(t)\,dt = \begin{cases} x & x \geq 0 \\ 0 & x < 0 \end{cases}$ si ha:

$$\mathcal{L}[F_1](p) = \frac{\mathcal{L}[H](p)}{p} = \frac{1}{p^2} \qquad (\mathrm{Re}\, p > 0)$$

© 978-88-08-06485-1

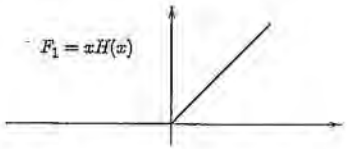


Figura A.9.

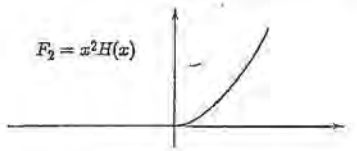


Figura A.10.

Analogamente se $F_2(x) = x^2H(x)$ si ha $F_2'(x) = 2F_1(x)$ e quindi

$$\mathcal{L}[F_2](p) = \frac{2}{p^3}. \quad (\mathrm{Re}\, p > 0)$$

Generalizzando si ottiene la formula

$$\boxed{\mathcal{L}[x^n H(x)](p) = \frac{n!}{p^{n+1}}} \quad n \text{ intero } \mathrm{Re}\, p > 0 \tag{2.13}$$

Trasformata di un prodotto di convoluzione

Siano f e g nulle in $(-\infty, 0)$. Il loro prodotto di convoluzione è definito dalla formula

$$(f * g)(x) = \int_0^x f(x-t)g(t)\,dt = \int_0^x f(t)g(x-t)\,dt.$$

Se f è trasformabile nel semipiano $\mathrm{Re}\, p > \alpha$ e $|g|$ è trasformabile nello stesso semipiano, allòra anche $f * g$ è *trasformabile in* $\mathrm{Re}\, p > \alpha$ e vale la formula

$$\boxed{\mathcal{L}[f * g](p) = \mathcal{L}[f](p) \cdot \mathcal{L}[g](p)} \tag{2.14}$$

Derivata della trasformata di Laplace

Sia

$$\mathcal{L}[f](p) = \int_0^{+\infty} e^{-px} f(x)\,dx \qquad (\mathrm{Re}\, p > \alpha)$$

Se, ad esempio, $|f(x)| \leq Me^{\alpha x}$ definitivamente per $x \to +\infty$, si può derivare rispetto a p sotto il segno di integrale, ottenendo

$$\frac{d}{dp}\mathcal{L}[f](p) = \int_0^{+\infty} -xf(x)e^{-px}\,dx$$

da cui la formula

$$\boxed{\mathcal{L}[-xf(x)](p) = \frac{d}{dp}\mathcal{L}[f](p)} \qquad (\mathrm{Re}\, p > \alpha) \tag{2.15}$$

Generalizzando, si ottiene

$$\mathcal{L}[x^2 f(x)](p) = \frac{d^2}{dp^2}\mathcal{L}[f](p)$$

 © 978-88-08-06485-1

ed infine, per ogni $n \geq 0$:

$$\mathcal{L}[(-1)^n x^n f(x)](p) = \frac{d^n}{dp^n}\mathcal{L}[f](p) \tag{2.16}$$

Esempio

Poiché $\mathcal{L}[\sin \omega x](p) = \dfrac{\omega}{p^2+\omega^2}$, si ha

$$\mathcal{L}[x \sin \omega x](p) = -\frac{d}{dp}\left(\frac{\omega}{p^2+\omega^2}\right) = \frac{2p\omega}{(p^2+\omega^2)^2}.$$

Integrale della trasformata di Laplace

Se $f(x)/x$ è trasformabile in $\mathrm{Re}\, p > \alpha$, dalla (2.15) si ha:

$$\mathcal{L}[f](p) = -\mathcal{L}\left[-x\frac{f(x)}{x}\right](p) = -\frac{d}{dp}\mathcal{L}\left[\frac{f(x)}{x}\right](p)$$

pensando p reale, integrando tra p e $+\infty$ e ricordando che $\mathcal{L}\left[\frac{f(x)}{x}\right](p) \to 0$ se $p \to +\infty$, si trova la formula

$$\mathcal{L}\left[\frac{f(x)}{x}\right](p) = \int_p^{+\infty} \mathcal{L}[f](s)\, ds \qquad (p > \alpha) \tag{2.17}$$

Raduniamo in una tabella le formule principali ottenute in questo paragrafo. Le funzioni nella colonna di sinistra *si intendono nulle per* $x < 0$.

Funzione	Trasformata	
$f(x-a)\cdot H(x-a)$	$e^{-ap}\mathcal{L}[f]$	formule del *ritardo*
$e^{ax}f(x)$	$\mathcal{L}[f](p-a)$	
$f'(x)$	$p\mathcal{L}[f] - f(0)$	formule delle derivate
$f''(x)$	$p^2\mathcal{L}[f] - pf(0) - f'(0)$	
$\int_0^x f(t)\, dt$	$\frac{1}{p}\mathcal{L}[f]$	formula della primitiva
$-xf(x)$	$\frac{d}{dp}\mathcal{L}[f]$	moltiplicazione per x^n
$(-1)^n x^n f(x)$	$\frac{d^n}{dp^n}\mathcal{L}[f]$	
$\frac{f(x)}{x}$	$\int_p^{+\infty} \mathcal{L}[f](s)ds$	divisione per x (p reale)
$\int_0^x f(x-t)g(t)dt$	$\mathcal{L}[f]\cdot\mathcal{L}[g]$	convoluzione

© 978-88-08-06485-1

Elenchiamo anche le trasformate delle funzioni più comuni (sempre intese nulle per $x < 0$).

Funzione	Trasformata	Validità
1	$\frac{1}{p}$	$\mathrm{Re}\, p > 0$
e^{-ax}	$\frac{1}{p+a}$	$\mathrm{Re}\, p > -\mathrm{Re}\, a$
x	$\frac{1}{p^2}$	$\mathrm{Re}\, p > 0$
x^n	$\frac{n!}{p^{n+1}}$	$\mathrm{Re}\, p > 0$, $n \in \mathbb{N}$
$\sin \omega x$	$\frac{\omega}{p^2+\omega^2}$	$\mathrm{Re}\, p > 0$
$\cos \omega x$	$\frac{p}{p^2+\omega^2}$	$\mathrm{Re}\, p > 0$
$\mathrm{Sh}\, ax$	$\frac{a}{p^2-a^2}$	$\mathrm{Re}\, p > \lvert\mathrm{Re}\, a\rvert$
$\mathrm{Ch}\, ax$	$\frac{p}{p^2-a^2}$	$\mathrm{Re}\, p > \lvert\mathrm{Re}\, a\rvert$
δ	1	$p \in \mathbb{C}$

Esempio

Poiché $\mathcal{L}[e^{-x}-1](p) = \dfrac{1}{p+1} - \dfrac{1}{p} = -\dfrac{1}{p(p+1)}$ $(\mathrm{Re}\, p > 0)$, abbiamo

$$\mathcal{L}\left[\frac{e^{-x}-1}{x}\right](p) = \int_p^{+\infty} -\frac{1}{s(s+1)}\, ds = \lim_{T\to+\infty}\int_p^T -\left[\frac{1}{s} - \frac{1}{s+1}\right] ds =$$

$$= \lim_{T\to+\infty}\left[-\log\frac{s}{s+1}\right]_p^T = \log\frac{p}{p+1}.$$

3 TRASFORMAZIONE INVERSA DI LAPLACE

Si può dimostrare che, data una trasformata di Laplace $F(p)$, è unica la funzione $f(x)$ tale che $\mathcal{L}[f] = F$.

L'operazione che permette di passare da F ad f si chiama *inversione* della trasformazione di Laplace ed f, indicata anche con $\mathcal{L}^{-1}[F]$, si chiama *antitrasformata* di F.

Le formule indicate nelle tabelle possono pensarsi anche come formule di antitrasformazione, se lette da destra verso sinistra.

Vediamo alcuni esempi di calcolo di antitrasformate.

 © 978-88-08-06485-1

Esempi

a) $F(p)=\dfrac{2}{p+3}$ b) $F(p)=\dfrac{4}{p^3}+\dfrac{6}{p^2+4}$.

a) $\mathcal{L}^{-1}\left[\dfrac{2}{p+3}\right](x)=2\mathcal{L}^{-1}\left[\dfrac{1}{p+3}\right](x)=2e^{-3x}\cdot H(x)$.

b) $\mathcal{L}^{-1}[F](x)=4\mathcal{L}^{-1}\left[\dfrac{1}{p^3}\right](x)+6\mathcal{L}^{-1}\left[\dfrac{1}{p^2+4}\right]=2\mathcal{L}^{-1}\left[\dfrac{2!}{p^3}\right]+3\mathcal{L}^{-1}\left[\dfrac{2}{p^2+4}\right]=$
$=(2x^2+3\sin 2x)H(x)$.

a) $F(p)=\dfrac{1}{(p-2)^2+9}$ b) $F(p)=\dfrac{p+3}{p^2+2p+5}$.

a) $\mathcal{L}^{-1}[F](x)=e^{2x}\mathcal{L}^{-1}\left[\dfrac{1}{p^2+9}\right]=\dfrac{1}{3}e^{2x}\sin 3x\cdot H(x)$.

b) $F(p)=\dfrac{p+3}{(p+1)^2+4}=\dfrac{p+1}{(p+1)^2+4}+\dfrac{2}{(p+1)^2+4}$.

e quindi

$$\mathcal{L}^{-1}[F](x)=\mathcal{L}^{-1}\left[\frac{p+1}{(p+1)^2+4}\right](x)+\mathcal{L}^{-1}\left[\frac{2}{(p+1)^2+4}\right]=$$
$$=e^{-x}\mathcal{L}^{-1}\left[\frac{p}{p^2+4}\right](x)+e^{-x}\mathcal{L}^{-1}\left[\frac{2}{p^2+4}\right]=$$
$$=e^{-x}(\cos 2x+\sin 2x)H(x).$$

$F(p)=\dfrac{1}{p(p^2+1)}$.

Poiché $\mathcal{L}^{-1}\left(\dfrac{1}{p}\right)=H(x)$ e $\mathcal{L}^{-1}\left(\dfrac{1}{p^2+1}\right)=\sin x\cdot H(x)$

$$\mathcal{L}^{-1}\left(\frac{1}{p}\cdot\frac{1}{p^2+1}\right)=\int_0^x H(t)\sin(x-t)\,dt=\int_0^x\sin(x-t)\,dt=\left[\cos(x-t)\right]_0^x=$$
$$=(1-\cos x)H(x).$$

Si poteva anche scrivere $F(p)=\dfrac{1}{p}-\dfrac{p}{p^2+1}$ e poi $\mathcal{L}^{-1}[F]=\mathcal{L}^{-1}\left[\dfrac{1}{p}\right]-\mathcal{L}^{-1}\left[\dfrac{p}{p^2+1}\right]$, pervenendo allo stesso risultato.

4 APPLICAZIONI ALLE EQUAZIONI DIFFERENZIALI. FUNZIONE DI TRASFERIMENTO DI UN SISTEMA

Consideriamo un'equazione differenziale lineare del secondo ordine a coefficienti costanti, non omogenea:

$$a_0y''(t)+a_1y'(t)+a_2y(t)=b(t)\qquad(t\geq 0)\tag{4.1}$$

© 978-88-08-06485-1

con condizioni iniziali

$$\begin{cases}y(0)=\alpha\\ y'(0)=\beta.\end{cases}\tag{4.2}$$

Applicando la trasformazione di Laplace, l'equazione (4.1) si trasforma in un'equazione algebrica. Poniamo $Y(p)=\mathcal{L}[y](p)$ e $B(p)=\mathcal{L}[b](p)$ e usiamo le formule per la trasformazione delle derivate; tenendo conto delle condizioni iniziali, si trova:

$$\mathcal{L}[y'](p)=pY(p)-\alpha\qquad\mathcal{L}[y''](p)=p^2Y(p)-p\alpha-\beta.$$

La (4.1) è allora trasformata nella seguente equazione:

$$(a_0p^2+a_1p+a_2)Y(p)=B(p)+\alpha(a_0p+a_1)+a_0\beta\tag{4.3}$$

da cui

$$Y(p)=\frac{1}{a_0p^2+a_1p+a_2}\cdot(B(p)+\alpha(a_0p+a_1)+a_0\beta).$$

La soluzione y si ricava dunque dall'antitrasformata del secondo membro. In particolare, se $\alpha=\beta=0$ si ha:

$$Y(p)=S(p)\cdot B(p)\tag{4.4}$$

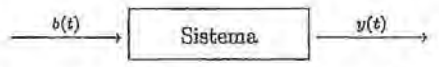
dove abbiamo posto $S(p)=\frac{1}{a_0p^2+a_1p+a_2}$, e perciò, usando la formula della convoluzione, si trova

$$y=\mathcal{L}^{-1}[B]*\mathcal{L}^{-1}[S]=b*\mathcal{L}^{-1}[S]\tag{4.5}$$

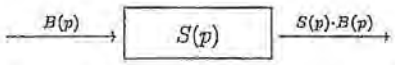
per $t\geq 0$.

Le formule (4.4) e (4.5) hanno una notevole interpretazione fisica. Supponiamo che la (4.1) modellizzi un sistema fisico caratterizzato da un *ingresso* $b=b(t)$ (ad esempio interpretabile come *forzante esterna* o *segnale d'entrata*), e da una *risposta* $y=y(t)$ (*moto* di una particella o *segnale in uscita* nelle interpretazioni precedenti).

Schematicamente:



Applicando la trasformata di Laplace, lo schema precedente viene tradotto nel seguente, in base all'equazione (4.4):



La funzione S caratterizza quindi il sistema fisico e prende il nome di *funzione di trasferimento* (o trasmittanza).

Se la funzione in ingresso è l'impulso unitario concentrato in $t=0$, cioè $b(t)=\delta$, si ha $B(p)=\mathcal{L}[\delta](p)=1$. In tal caso, quindi, $Y(p)=S(p)$ e si deduce che *la funzione di trasferimento è la trasformata della risposta all'impulso unitario*. La (4.5) indica

allora che, se le condizioni iniziali sono nulle, *la risposta alla funzione $b(t)$ in ingresso è data dalla convoluzione di $b(t)$ con la risposta all'impulso unitario*, $s = \mathcal{L}^{-1}[S]$.

Esempio

Il sistema fisico rappresentato dal circuito R, L, C indicato in figura è retto dall'equazione differenziale

$$LCu'' + RCu' + u = b(t)$$

dove u rappresenta la differenza di potenziale e b la forza elettromotrice applicata.

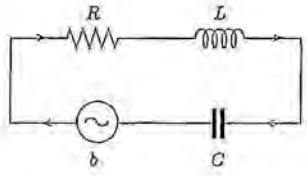


Figura A.11.

La funzione di trasferimento è data da

$$S(p) = \frac{1}{LCp^2 + RCp + 1} = \frac{\omega_0^2}{p^2 + 2m\omega_0 p + \omega_0^2}$$

dove s'è posto

$$\omega_0 = \frac{1}{\sqrt{LC}} \qquad \text{(pulsazione propria del circuito)}$$

$$m = \frac{R}{2}\sqrt{\frac{C}{L}} \qquad \text{(coefficiente di smorzamento)}$$

Il calcolo di $\mathcal{L}^{-1}[S]$ dipende dal segno di $\Delta = \omega_0^2(m^2 - 1)$. Ad esempio, nel caso $m < 1$, si ha, scrivendo $p^2 + 2m\omega_0 p + \omega_0^2 = (p + m\omega_0)^2 - (1 - m^2)\omega_0^2 = (p + m\omega_0)^2 - \Delta$,

$$s(t) = \mathcal{L}^{-1}[S](t) = \frac{\omega_0}{\sqrt{1-m^2}}\mathcal{L}^{-1}\left[\frac{\omega_0\sqrt{1-m^2}}{(p+m\omega_0)^2 + (1-m^2)\omega_0^2}\right](t) =$$

$$= \frac{\omega_0}{\sqrt{1-m^2}} e^{-m\omega_0 t} \cdot \sin(\omega_0\sqrt{1-m^2}\, t) \cdot H(t).$$

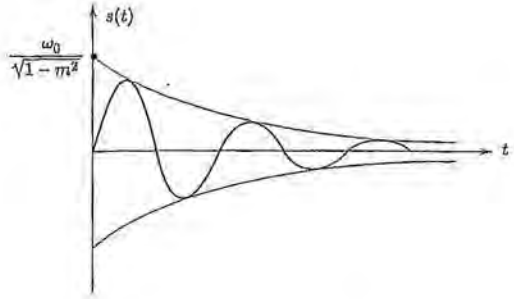


Figura A.12.

Se ora si vuole determinare la risposta del sistema ad un ingresso del tipo $b(t) = E \cdot H(t)$, con condizioni iniziali nulle, si ha dalla (4.5), per $t \geq 0$:

$$y(t) = (b * s)(t) = \frac{E\omega_0}{\sqrt{1-m^2}} \int_0^t e^{-m\omega_0 s} \sin(\omega_0\sqrt{1-m^2}\, s)\, ds =$$

$$= E - \frac{E}{\sqrt{1-m^2}} e^{-m\omega_0 t} \cos(\omega_0\sqrt{1-m^2}\, t + \varphi)$$

dove $\varphi = \operatorname{arctg} \frac{m}{\sqrt{1-m^2}}$.

Si noti che, se $E = 1$, $y'(t) = s(t)$ e cioè: la risposta all'impulso unitario è la derivata della risposta alla funzione di Heaviside.

5 TRASFORMATA DI FOURIER

La serie di Fourier, di cui si è parlato nel capitolo 7, rappresenta una funzione $g(y)$, τ-periodica, come somma di funzioni periodiche semplici del tipo (usiamo la rappresentazione complessa)

$$\hat{f}_{(k)} e^{i\omega k y} \qquad \left(\omega = \frac{2\pi}{\tau};\quad k \in \mathbb{Z}\right)$$

Le costanti (complesse) $\hat{f}(k)$ (che si calcolano con la formula (3.18) del capitolo 7 ponendo $x = \omega y$, $f(x) = f(\omega y) = g(y)$) individuano completamente la funzione $g(y)$ somma della serie; queste costanti formano il cosiddetto *spettro* di $g(y)$: i moduli $|\hat{f}(k)|$ costituiscono lo *spettro delle ampiezze*, gli argomenti di $\hat{f}(k)$ lo *spettro delle fasi*. Questo spettro si dice *discreto*, perché è costituito da una infinità numerabile di elementi (vedi tabella a pagina seguente).

Anche una funzione non periodica $f(x)$, definita su tutta la retta, può essere "analizzata" attraverso il suo "spettro", in maniera analoga a quanto si fa per le funzioni periodiche; lo strumento per fare ciò è dato dalla trasformata di Fourier.

Sia f una funzione definita su $\mathbb{R}$ con le seguenti proprietà:

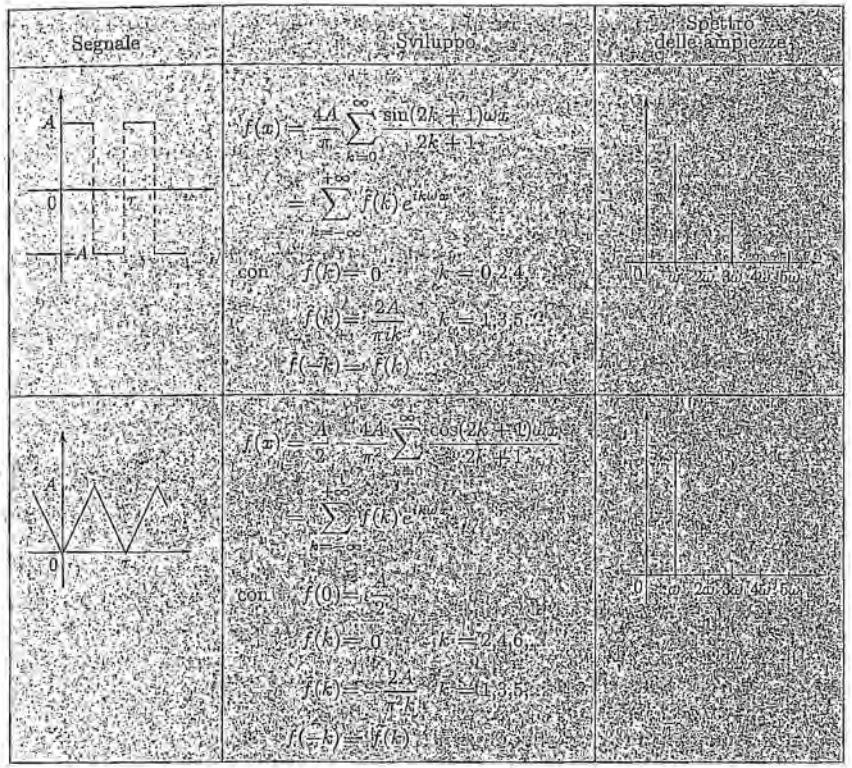
- C^1 a pezzi su ogni intervallo limitato $[-a, +a]$
- assolutamente integrabile da $-\infty$ a $+\infty$, cioè $\displaystyle\int_{-\infty}^{+\infty} |f(x)|\, dx < +\infty$

Si definisce *trasformata di Fourier* di f la funzione $\hat{f} : \mathbb{R} \longrightarrow \mathbb{C}$ assegnata dalla formula:

$$\hat{f}(\nu) = \int_{-\infty}^{+\infty} f(x) e^{-i2\pi\nu x} dx \tag{5.1}$$

La funzione $\hat{f}$ risulta *continua* su tutta la retta e inoltre si annulla all'infinito:

$$\boxed{\hat{f}(\nu) \text{ è continua su } \mathbb{R} \text{ e } \lim_{\nu \to \pm\infty} \hat{f}(\nu) = 0} \tag{5.2}$$



Il grafico della funzione $\nu \longmapsto |\hat{f}(\nu)|$ si chiama ancora *spettro delle ampiezze* di f, il grafico della funzione $\nu \longmapsto$ argomento di $\hat{f}(\nu)$, *spettro di fase*. In questo caso si parla però di *spettro continuo*, poiché il grafico di queste funzioni è costituito da una infinità continua di elementi.

Trasformata inversa

Nelle ipotesi su f ricordate all'inizio, si dimostra che, se $\hat{f}(\nu) = \int_{-\infty}^{+\infty} f(x)e^{-i2\pi\nu x}dx$, allora

$$\int_{-\infty}^{+\infty} \hat{f}(\nu)e^{i2\pi\nu x}d\nu = \frac{1}{2}[f(x+) + f(x-)] \tag{5.3}$$

e perciò, in ogni punto in cui f è continua, si ha:

$$\boxed{\hat{f}(\nu) = \int_{-\infty}^{+\infty} f(x)e^{-i2\pi\nu x}dx \iff f(x) = \int_{-\infty}^{+\infty} \hat{f}(\nu)e^{i2\pi\nu x}d\nu} \tag{5.4}$$

Indichiamo con $\mathcal{F}$ l'operazione che fa passare da $f(x)$ a $\hat{f}(\nu)$

$$\mathcal{F}[f](\nu) = \hat{f}(\nu)$$

e con $\mathcal{F}^{-1}$ l'operazione inversa

$$\mathcal{F}^{-1}[\hat{f}](x) = f(x)$$

Si osserva che si passa da $\mathcal{F}$ a $\mathcal{F}^{-1}$ attraverso i seguenti scambi:

$$\begin{array}{ccc} f & \longmapsto & \hat{f} \\ x & \longmapsto & \nu \\ i & \longmapsto & -i \end{array} \tag{5.5}$$

Simmetrie

Si osservi che

$$\int_{-\infty}^{+\infty} f(x)e^{-i2\pi\nu x}\,dx = \int_{-\infty}^{+\infty} f(x)\cos(2\pi\nu x)\,dx - i\int_{-\infty}^{+\infty} f(x)\sin(2\pi\nu x)\,dx\,.$$

Perciò, se f è pari, $\int_{-\infty}^{+\infty} f(x)\sin(2\pi\nu x)\,dx = 0$ e $\int_{-\infty}^{+\infty} f(x)\cos(2\pi\nu x)\,dx = 2\int_0^{\infty} f(x)\cos(2\pi\nu x)\,dx$; perciò $\hat{f}(\nu)$ è pure funzione pari di ν e si ha

$$\hat{f}(\nu) = 2\int_0^{\infty} f(x)\cos(2\pi\nu x)\,dx$$

e, reciprocamente,

$$f(x) = 2\int_0^{\infty} \hat{f}(\nu)\cos(2\pi\nu x)d\nu\,.$$

Analogamente, se f è dispari, valgono le formule di reciprocità nei punti di continuità di f:

$$\hat{f}(\nu) = -2i\int_0^{\infty} f(x)\sin(2\pi\nu x)\,dx \iff f(x) = 2i\int_0^{\infty} \hat{f}(\nu)\sin(2\pi\nu x)d\nu\,.$$

Esempio

Funzioni impulsive.

$$f(x) = A \text{ se } |x| \le \frac{T}{2},\ f(x) = 0 \text{ se } |x| > \frac{T}{2}$$

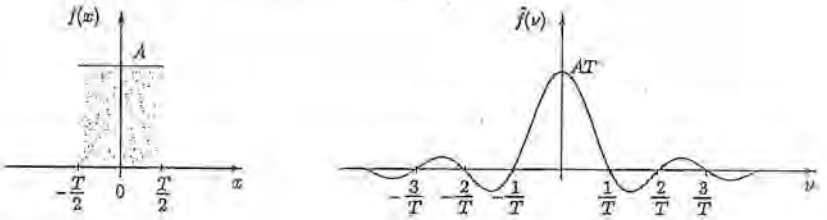


Figura A.13.

$$\hat{f}(\nu) = 2\int_0^{+\infty} f(x)\cos(2\pi\nu x)\,dx = 2\int_0^{\frac{T}{2}} A\cos(2\pi\nu x)\,dx = \frac{\sin(\pi\nu T)}{\pi\nu T}AT. \tag{5.6}$$

Impulso di Dirac. Il prodotto AT rappresenta l'area sottesa dagli impulsi dell'esempio 1.1. Mantenendo costante (= 1) questo prodotto e facendo tendere T a zero (e perciò A a $+\infty$) questi impulsi "tendono" all'impulso di Dirac δ; la trasformata di Fourier $\hat{f}(\nu)$ tende a 1 (ricorda il limite notevole $\lim_{T\to 0}\frac{\sin(\pi\nu T)}{\pi\nu T} = 1$). Così possiamo giustificare l'importante risultato

$$\boxed{\hat{\delta} = 1} \tag{5.7}$$

Questo risultato è stato ottenuto in modo puramente formale. Abbiamo già osservato che la distribuzione δ di Dirac non è una funzione; osserviamo ora che la sua trasformata non tende a zero all'infinito, come espresso dalla (5.2).

Funzioni esponenziali. Per esercizio lo studente può verificare le seguenti formule: sia $a > 0$ e consideriamo le funzioni

i) $f(x) = \begin{cases} 0 & \text{se } x \le 0 \\ e^{-ax} & \text{se } x > 0 \end{cases}$ ii) $f(x) = e^{-a|x|}$ iii) $f(x) = e^{-a|x|}\operatorname{sgn} x$.

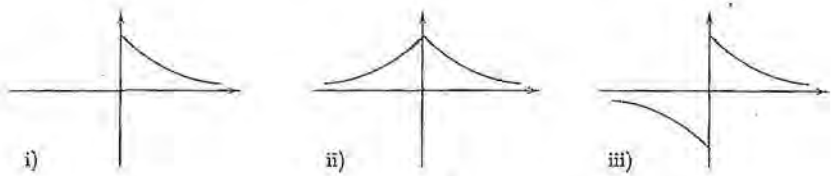


Figura A.14.

Si trova, per $\hat{f}(\nu)$, rispettivamente

$$\frac{1}{a + 2i\pi\nu} \qquad \frac{2a}{a^2 + 4\pi^2\nu^2} \qquad \frac{-4i\pi\nu}{a^2 + 4\pi^2\nu^2}.$$

6 PROPRIETÀ DELLA TRASFORMAZIONE DI FOURIER

Linearità

Se c_1 e c_2 sono costanti (complesse), si ha

$$(c_1 f + c_2 g)\hat{} = c_1\hat{f} + c_2\hat{g}$$

come immediatamente segue dalla proprietà di linearità dell'integrale (generalizzato).

Formula del ritardo

Sia $g(x) = f(ax - b)$, $a, b \in \mathbb{R}$.

$$\hat{g}(\nu) = \int_{-\infty}^{+\infty} e^{-i2\pi\nu x} f(ax - b)\,dx =$$

(posto $y = ax - b$ e tenendo conto che, se $a > 0$, i limiti di integrazione restano immutati mentre vengono scambiati se $a < 0$)

$$= \operatorname{sgn} a \int_{-\infty}^{+\infty} e^{-i2\pi\nu\frac{y+b}{a}} f(y)\frac{dy}{a} =$$

$$= \frac{e^{-i2\pi\frac{b}{a}\nu}}{|a|}\int_{-\infty}^{+\infty} e^{-i2\pi\frac{\nu}{a}y} f(y)\,dy = \frac{1}{|a|}e^{-i2\pi\frac{b}{a}\nu}\hat{f}\left(\frac{\nu}{a}\right)$$

In particolare, se $b = 0$, abbiamo

$$\boxed{\mathcal{F}[f(ax)](\nu) = \frac{1}{|a|}\hat{f}\left(\frac{\nu}{a}\right)} \tag{6.1}$$

Se $a = 1$ e $b = x_0$

$$\boxed{\mathcal{F}[f(x - x_0)](\nu) = e^{-2\pi i\nu x_0}\hat{f}(\nu)} \tag{6.2}$$

e anche (applicando le regole di scambio (5.5))

$$\boxed{\mathcal{F}^{-1}[\hat{f}(\nu - \nu_0)](x) = e^{2\pi i\nu_0 x} f(x)} \tag{6.3}$$

ESEMPIO

Impulso centrato in x_0

Sia $f(x)$ come in figura. Abbiamo allora

$$\hat{f}(\nu) = e^{-2\pi i\nu x_0}\frac{\sin(\pi\nu T)}{\pi\nu T}.$$

Se $T \to 0$, $f(x)$ "tende" alla distribuzione di Dirac concentrata in x_0, che si indica con $\delta(x - x_0)$; abbiamo così:

$$\boxed{\mathcal{F}[\delta(x - x_0)](\nu) = e^{-2\pi i\nu x_0}} \tag{6.4}$$

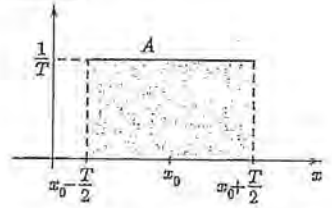


Figura A.15.

Trasformata della derivata

Sia f derivabile e calcoliamo

$$\mathcal{F}[f'](\nu) = \int_{-\infty}^{+\infty} f'(x)e^{-i2\pi\nu x}\,dx = \text{(con una integrazione per parti)}$$

$$= \left[f(x)e^{-2\pi i\nu x}\right]_{-\infty}^{+\infty} - \int_{-\infty}^{+\infty} f(x)\left(-2i\pi\nu e^{-2\pi i\nu x}\right)dx.$$

Se $\lim_{x\to\pm\infty} f(x) = 0$ (cosa che si potrebbe dimostrare per quasi tutte le funzioni derivabili e assolutamente integrabili su $\mathbb{R}$), abbiamo la formula

$$\boxed{\mathcal{F}[f'](\nu) = 2\pi i\nu\hat{f}(\nu)} \tag{6.5}$$

e, reciprocamente,

$$\mathcal{F}^{-1}[(\hat{f})'](x) = -2\pi i x f(x)\,.$$

Questa uguaglianza si può anche scrivere nella forma (applicando $\mathcal{F}$ a entrambi i membri e tenendo conto che $\mathcal{F}\mathcal{F}^{-1}[\hat{f}] = \hat{f}$)

$$\mathcal{F}[xf(x)](\nu) = \frac{i}{2\pi}\hat{f}'(\nu)\,. \tag{6.6}$$

Le precedenti formule possono essere iterate ottenendo

$$\boxed{\begin{aligned}\mathcal{F}[f^{(n)}(x)](\nu) &= (2\pi i\nu)^n\hat{f}(\nu)\\ \mathcal{F}[x^n f(x)](\nu) &= \left(\frac{i}{2\pi}\right)^n\hat{f}^{(n)}(\nu)\end{aligned}}$$

La (6.5) vale anche se f è continua ovunque e derivabile tranne che in un numero finito di punti angolosi, come nel prossimo esempio.

Esempio

Trasformata di un segnale triangolare. Sia $f(x)$ come in figura A.16

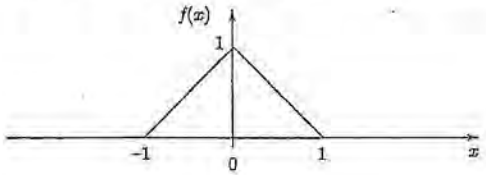


Figura A.16.

La sua derivata, $f'(x)$ è la funzione rappresentata in figura A.17

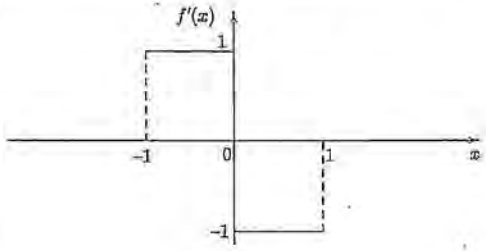


Figura A.17.

cioè è la somma di due impulsi, uno positivo (centrato in $-\frac{1}{2}$) e l'altro negativo (centrato in $\frac{1}{2}$).

Dalla (5.6) (con $T = 1$) otteniamo subito

$$\mathcal{F}[f'](\nu) = e^{-2\pi i\nu(-\frac{1}{2})}\frac{\sin\pi\nu}{\pi\nu} - e^{-2\pi i\nu(\frac{1}{2})}\sin\frac{\pi\nu}{\pi\nu} = \frac{\sin\pi\nu}{\pi\nu}(e^{i\pi\nu} - e^{-i\pi\nu}) = 2i\frac{\sin\pi\nu}{\pi\nu}\sin\pi\nu$$

(avendo applicato la formula di Eulero). Utilizzando infine la (6.5) otteniamo

$$\hat{f}(\nu) = \left[\frac{\sin(\pi\nu)}{\pi\nu}\right]^2\,. \tag{6.7}$$

Trasformata del prodotto di convoluzione

Vale un risultato analogo a quello visto per la trasformata di Laplace; se $f(x)$, $g(x)$ sono funzioni definite su $\mathbb{R}$ e

$$(f * g)(x) = \int_{-\infty}^{+\infty} f(y)\, g(x-y)\, dy$$

è il loro prodotto di convoluzione, allora risulta

$$\boxed{\mathcal{F}[f * g](\nu) = \hat{f}(\nu)\cdot\hat{g}(\nu)} \tag{6.8}$$

e, reciprocamente,

$$\mathcal{F}[f\cdot g](\nu) = (\hat{f} * \hat{g})(\nu)$$

cioè: *la trasformata di Fourier scambia tra di loro prodotto algebrico e prodotto di convoluzione.*

Esempio

Se $g(x) = \delta(x)$ (distribuzione di Dirac), si ha (ricordando che $\hat{\delta} = 1$)

$$\mathcal{F}[f * \delta] = \mathcal{F}[f]$$

e cioè:

$$(f * \delta)(x) = f(x)\,.$$

Se $g(x) = \delta(x - x_0)$ si ha (ricordando la (6.4) e la (6.2))

$$\mathcal{F}[f * \delta(x - x_0)](\nu) = e^{-2\pi i\nu x_0}\hat{f}(\nu) = \mathcal{F}[f(x - x_0)](\nu)$$

e cioè

$$f * \delta(x - x_0) = f(x - x_0)\,.$$

La convoluzione di una funzione $f(x)$ con l'impulso di Dirac $\delta(x-x_0)$ provoca una traslazione di x_0 unità lungo l'asse x.

 © 978-88-08-06485-1

Uguaglianza di Parseval

Vale, per la trasformata di Fourier, un risultato analogo ad uno già visto per le serie di Fourier. Se $f(x)$ rappresenta (l'ampiezza di) un segnale (funzione del tempo: qui la variabile x ha significato di coordinata temporale), l'integrale $\int_{-\infty}^{+\infty} f(x)^2\,dx$ (se convergente) si può interpretare come energia totale del segnale. Questa si può calcolare anche per mezzo della trasformata di Fourier di f; risulta infatti

$$\boxed{\int_{-\infty}^{+\infty} f(x)^2\,dx = \int_{-\infty}^{+\infty} |\hat{f}(\nu)|^2\,d\nu} \tag{6.9}$$

Questa formula (analoga alla (3.11) del teorema 7.12 del capitolo 7) si chiama *uguaglianza di Parseval*. Essa è un caso particolare della seguente (detta *teorema di Plancherel*), dove $f(x)$ e $g(x)$ sono due funzioni tali che gli integrali scritti siano convergenti:

$$\int_{-\infty}^{+\infty} f(x)\overline{g(x)}\,dx = \int_{-\infty}^{+\infty} \hat{f}(\nu)\overline{\hat{g}(\nu)}\,d\nu\,. \tag{6.10}$$

Esempio

Calcolare $\int_{-\infty}^{+\infty} \left(\frac{\sin t}{t}\right)^2 dt$. Ricordando la (6.7) abbiamo

$$\int_{-\infty}^{+\infty} \left(\frac{\sin t}{t}\right)^2 dt = \int_{-\pi/2}^{\pi/2} dx = \pi$$

Proprietà della trasformata di Fourier

Funzione	Trasformata	
$f(x)$	$\hat{f}(\nu)$	
$f(ax-b)$	$\frac{1}{\lvert a\rvert} e^{-i2\pi\frac{b}{a}\nu}\hat{f}\left(\frac{\nu}{a}\right)$	formula del ritardo
$f^{(n)}(x)\quad n=1,2,\ldots$	$(2\pi i\nu)^n \hat{f}(\nu)$	formule per le derivate
$x^n f(x)\quad n=1,2,\ldots$	$\left(\frac{i}{2\pi}\right)^n \hat{f}^{(n)}(\nu)$	
$(f*g)(x)$	$\hat{f}(\nu)\cdot\hat{g}(\nu)$	formula della convoluzione
$f(x)\cdot g(x)$	$(\hat{f}*\hat{g})(\nu)$	

7 UNA APPLICAZIONE: STUDIO DI UN CIRCUITO *RC*

Consideriamo un circuito *RC*: la tensione $s(t)$ che si stabilisce alle facce opposte del condensatore soddisfa l'equazione differenziale

$$\tau\dot{s}(t) + s(t) = e(t) \tag{7.1}$$

© 978-88-08-06485-1

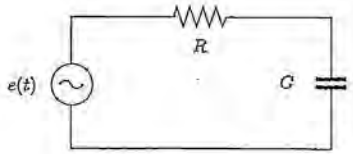


Figura A.18.

dove si è posto $\tau = RC$. Applicando all'equazione la trasformata di Fourier si ottiene

$$\tau(2\pi i\nu)\hat{s}(\nu) + \hat{s}(\nu) = \hat{e}(\nu)$$

e cioè

$$\hat{s}(\nu) = \frac{\hat{e}(\nu)}{1+2\pi i\nu\tau} = \hat{e}(\nu)\cdot\hat{h}(\nu)\,. \tag{7.2}$$

La funzione $\hat{h}(\nu) = 1/(1+2\pi i\nu\tau)$ è detta *funzione di trasferimento* del circuito. Essa permette di ottenere $s(t)$ dalla conoscenza di $e(t)$. Applicando infatti la (6.8) (teorema di convoluzione) otteniamo

$$s(t) = (h*e)(t)\,.$$

Nel nostro caso, essendo

$$\hat{h}(\nu) = \frac{1}{\tau}\,\frac{1}{\frac{1}{\tau}+2\pi i\nu}$$

otteniamo (ricordando l'esempio 5.3)

$$h(t) = \begin{cases} \frac{1}{\tau}\,e^{-\frac{t}{\tau}} & (t>0) \\ 0 & (t\le 0) \end{cases}$$

e perciò

$$s(t) = \int_0^{+\infty} e(t-t')\,\frac{1}{\tau}\,e^{-\frac{t'}{\tau}}dt' = \int_0^{+\infty} e(t')\,\frac{1}{\tau}\,e^{-\frac{t-t'}{\tau}}dt'\,.$$

Per esempio, se $e(t)$ è un impulso di durata T (fig. A.19) abbiamo

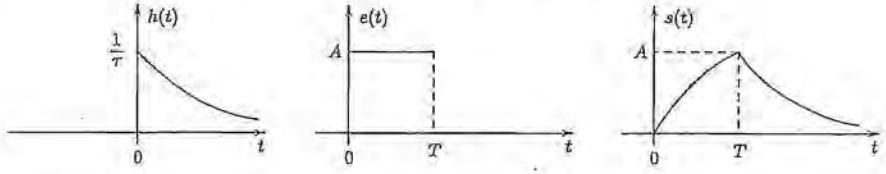


Figura A.19.

$$s(t) = \begin{cases} A\,(1-e^{-t/\tau}) & \text{se } t\le T \\ A\,e^{-t/\tau}(e^{\frac{T}{\tau}}-1) & \text{se } t\ge T\,. \end{cases}$$

Notiamo infine che se $e(t)$ è un impulso di Dirac: $e(t)=\delta(t)$ si ottiene

$$s(t) = (h*\delta)(t) = h(t)\,.$$

La funzione di trasferimento risulta pertanto la risposta del sistema all'impulso di Dirac.

 © 978-88-08-06485-1

Esercizi

Trasformata di Laplace

Trovare le $\mathcal{L}$-trasformate delle seguenti funzioni (assegnate per $x \geq 0$, nulle per $x < 0$):

i) $(x-2)^3$ ii) $e^{-x}(1+x^2)$ iii) $\sin(x-2)$ iv) $\sin 3x \cos x$ v) $\operatorname{Ch} \omega x$

$\mathcal{L}$-trasformare i seguenti segnali periodici

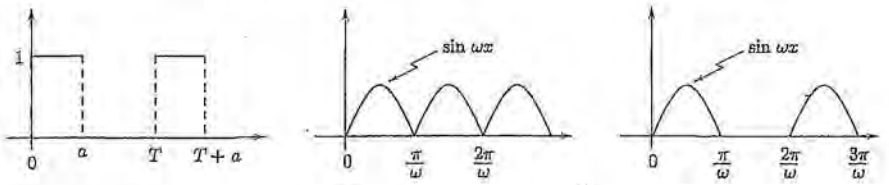


Figura A.20.

Trovare le funzioni la cui $\mathcal{L}$-trasformata è:

i) $\dfrac{1}{p^2+1}$ ii) $\dfrac{1}{(p+1)^2}$ iii) $\dfrac{p}{(p+1)^2(p-3)}$ iv) $\dfrac{p}{(p^2+1)^2}$

v) $\dfrac{1}{(p^2+\omega^2)^2}$ vi) $\dfrac{1}{(p^2+a^2)(p^2+b^2)}$ vii) $\dfrac{p}{(p-1)(p+2)}$

Utilizzare il teorema di convoluzione per trovare l'antitrasformata di:

i) $\dfrac{p}{(p^2+a^2)(p^2+b^2)}$ ii) $\dfrac{1}{p^n(p^2+a^2)}$

Risolvere i seguenti problemi di Cauchy:

i) $y'' + y = \sin \omega x \qquad y(0) = 0 \quad \text{e} \quad y'(0) = 1$

ii) $y'' + 2y' + y = e^{-2x} \qquad y(0) = 1 \quad \text{e} \quad y'(0) = 0$

Risolvere i seguenti sistemi:

i) $\begin{cases} y'(x) = 4y + 4z & y(0) = 1 \\ z'(x) = \ y + 4z & z(0) = 0 \end{cases}$ ii) $\begin{cases} y'(x) = -\ y + z + e^{-x} & y(0) = 0 \\ z'(x) = -9y - z & z(0) = 0 \end{cases}$

Trasformata di Fourier

Sia $f(x) = e^{-\pi x^2}$. Verificare che $f'(x) + 2\pi x f(x) = 0$. Applicare la trasformata di Fourier a questa equazione e ricavare $\hat{f}(\nu)$.
Calcolare poi $\int_0^{\infty} e^{-x^2} \cos(\omega x) dx$. In particolare, per $\omega = 0$, ritrovare il risultato del capitolo 5, paragrafo 2.

Calcolare la F-trasformata di $f_\sigma(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-x^2/2\sigma^2}$ $(\sigma > 0)$. Mostrare che $f_{\sigma_1} * f_{\sigma_2} = f_{\sqrt{\sigma_1^2+\sigma_2^2}}$. Ricavare che $\lim_{\sigma \to 0} f_\sigma = \delta$.

Applicare la trasformata di Fourier all'equazione differenziale del secondo ordine:

$$s''(t) + 2b\omega s'(t) + \omega^2 s(t) = \omega^2 e(t)$$

(b, ω costanti assegnate, $e(t)$ funzione nota). Si troverà

$$\hat{s}(\nu) = h(\nu)\hat{e}(\nu)$$

dove $h(\nu)$ è la funzione di trasferimento. Considerare i casi $b < 1$ e $b > 1$.
In particolare, trovare la risposta del sistema ad un impulso di Dirac.

Trovare la trasformata inversa di Fourier per

$$\hat{f}(\nu) = \frac{1}{(4\pi^2\nu^2+1)^2}$$

(ricordando che $\mathcal{F}(e^{-|x|}) = \frac{2}{1+4\pi^2\nu^2}$ e applicando il teorema di convoluzione).

POLITECNICO
Biblioteca Centrale
di Ingegneria
AC 2187